File Copy

AD 766283

BRL R 1658

BRL

REPORT NO. 1658

LEAST SQUARES ADJUSTMENT WITH FINITE RESIDUALS FOR NON-LINEAR CONSTRAINTS AND PARTIALLY CORRELATED DATA

by

Dr. Aivars Celmiņš

July 1973

Approved for public release; distribution unlimited.

USA BALLISTIC RESEARCH LABORATORIES
ABERDEEN PROVING GROUND, MARYLAND

Destroy this report when it is no longer needed.
Do not return it to the originator.

Secondary distribution of this report by originating or sponsoring activity is prohibited.

Additional copies of this report may be obtained from the National Technical Information Service, U.S. Department of Commerce, Springfield, Virginia 22151.

The findings in this report are not to be construed as an official Department of the Army position, unless so designated by other authorized documents.

B A L L I S T I C R E S E A R C H L A B O R A T O R I E S

REPORT NO. 1658

July 1973

LEAST SQUARES ADJUSTMENT WITH FINITE RESIDUALS FOR NON-LINEAR CONSTRAINTS AND PARTIALLY CORRELATED DATA.

DR. AIVARS CELMIŅŠ

Applied Mathematics Laboratory

Approved for public release; distribution unlimited.

RDT&E Project No. 1T161102A14B

A B E R D E E N P R O V I N G G R O U N D, M A R Y L A N D

B A L L I S T I C R E S E A R C H L A B O R A T O R I E S

REPORT NO. 1658

AKCelmins/cas
Aberdeen Proving Ground
July 1973

LEAST SQUARES ADJUSTMENT WITH FINITE RESIDUALS FOR NON-LINEAR CONSTRAINTS AND PARTIALLY CORRELATED DATA.

ABSTRACT

The subject of this paper is the adjustment ("curve fitting") of data by the least squares method. Based on general formulas, which are derived in this paper, a new algorithm for the least squares method is established. It can be applied to cases where more than one observable contain observational errors and where the postulated relation between the observables is non-linear. Also taken into account are accuracies of the observations and correlations between the components of each observation vector. New formulas are derived for the estimation of the variance-covariance matrix of the fitted parameters. It is shown that the conventionally used estimation formula is theoretically wrong except for very limited special cases. Numerical tests of the algorithm demonstrate its accuracy and exceptional convergence characteristics. They also show that the conventional estimate of the variance-covariance matrix of the parameters is a very bad approximation to the theoretically correct estimate derived in this paper.

TABLE OF CONTENTS

		Page
	ABSTRACT	3
	LIST OF FIGURES	7
	LIST OF TABLES	9
	LIST OF SYMBOLS	11
1.	INTRODUCTION	13
2.	COMPUTATION OF RESIDUALS AND PARAMETERS	14
3.	ESTIMATED VARIANCES AND COVARIANCES OF THE PARAMETERS	23
4.	EXAMPLES	28
	REFERENCES	53
	APPENDIX A - COMPUTER PROGRAM COLSGN	55
	DISTRIBUTION LIST	61

LIST OF FIGURES

Page

Figure 1. Linear Fit to Pearson's Data. 35

Figure 2. Cubic Fit to Pearson's Data 36

Figure 3. Quintic Fit to Pearson's Data 37

Figure 4. Fitting of Pseudo-Cassinian Curves to Correlated Data 38

Figure 5. Fitting of Pseudo-Cassinian Curves to Independent Data. 39

LIST OF TABLES

Page

Table I. Pearson's Data and York's Weights. 40

Table II. Linear Fit to Pearson's Data with Unit Weights 41

Table III. Linear Fit to Pearson's Data with York's Weights . . .42

Table IV. Cubic Fit to Pearson's Data with Unit Weights. 43

Table V. Cubic Fit to Pearson's Data with York's Weights. . . .44

Table VI. Quintic Fit to Pearson's Data with Unit Weights. . . .45

Table VII. Quintic Fit to Pearson's Data with York's Weights. . .46

Table VIII. Input for Acoustic Amplification Problem 47

Table IX. Parameters of Acoustic Amplification Problem 49

Table X. Data for Fitting Generalized Cassinian Curve 50

Table XI. Results of Fitting Pseudo-Cassinian Curve to 51
Correlated Observations 51

Table XII. Results of Fitting Pseudo-Cassinian Curve to
Observations with Unit Weights. 52

LIST OF SYMBOLS

$A = \mathrm{grad}_{\xi} F$ - n-dimensional vector of partial derivatives of F with respect to the components of the observable ξ.

$A_x = \dfrac{\partial^2 F}{\partial \xi^i \partial \xi^k}$ - (nxn)-matrix of second order derivatives of F.

$A_t = \dfrac{\partial^2 F}{\partial \xi^i \partial \theta^k}$ - (nxp)-matrix of second order derivatives of F.

$B = \mathrm{grad}_{\theta} F$ - p-dimensional vector of partial derivatives of F with respect to the components of the parameter vector θ.

$B_t = \dfrac{\partial^2 F}{\partial \theta^i \partial \theta^k}$ - (pxp)-matrix of second order derivatives of F.

$B_x = A_t^T$

c_j - n-dimensional least squares correction of the observation X_j.

C_j - approximation to c_j, n-dimensional vector.

e_{lin} - linearization error in Section 2.

$f(\Theta)$ - an additive part of the constraint functional.

$F(\xi, \theta)$ - constraint functional.

$g_j = 1/(A_j^T R_j A_j)$ - weight of observation X_j.

G - (nxn)-matrix in Section 3.

H - (pxn)-matrix in Section 3.

I - unit matrix.

k_j - correlate (Lagrange multiplier) corresponding to observation X_j.

m_o, $\hat{m}_o$ - estimated standard errors of weight one.

M_j - (pxp)-matrix in Section 3.

n - dimension of the space of observables.

N - (pxp)-normal equation matrix.

p - dimension of the parameter vector.

r - number of points observed in the space of observables.

R_j - (nxn)-variance-covariance matrix of the components of X_j.

t - p-dimensional least squares value of the parameter vector.

T - approximation to t; p-dimensional vector.

V - estimated variance-covariance matrix of the components of t; (pxp)-matrix.

V_{conv} - conventional approximation to V; (pxp)-matrix.

W - weighted sum of correlate squares, corresponding to the weighted sum of adjustment squares for not correlated data.

$\tilde{W}$ - modified W.

x_j - n-dimensional vector of the least squares value of X_j.

X_j - n-dimensional vector of observations.

Γ_1 - (nxn)-matrix in Section 3.

Γ_2 - (nxp)-matrix in Section 3.

ε_j = $c_j - C_j$ - correction of C_j; n-dimensional vector.

θ - p-dimensional parameter vector.

Θ - (pxp)-matrix in Section 3.

ξ - n-dimensional coordinate vector in the space of observables.

τ = T - t - correction of T; p-dimensional vector.

1. INTRODUCTION

Mathematical models of causal relations between measurable variables are often expressed in form of functional equations between the variables. Typically the functionals contain a number of parameters, the values of which are determined such that the corresponding relation between the variables agrees closely with experiments. (In case of only two variables this is often called "curve fitting".) A most common technique for the determination of the parameters is the method of least squares. In this paper we shall discuss an algorithm for a fairly general type of the least squares method, applicable to partially correlated observations*, non-linear functionals and finite residuals. Particularly for cases with finite residuals the theory of the least squares method has not been sufficiently developed. On the other hand, even linear curve fitting with errors in both variables cannot be correctly treated under the usual assumption of negligible residuals.

Numerical techniques for problems similar to those considered here have been suggested by several authors. Non-linear problems with correlated data have been treated by Brown (1955)[1]**, Tienstra (1956)[2], and Grossmann (1961)[3]. These authors linearize the constraint functionals by using a truncated Taylor expansion and treat the problem as one with linear constraints. As will be shown later, such an approach generally will not give a correct least squares solution, unless an iteration is included in the algorithm. Only Tienstra mentions the possibility of an iteration, but he does not derive the necessary formulas.

Not correlated data with non-linear constraints are treated in several textbooks, e.g., Demming (1944)[4], Arley & Buch (1950)[5] and Wolberg (1967)[6]. None of these authors consider the necessary iteration processes. Usually one iteration step or iteration of the parameters only is suggested.

---

*The partial correlation considered is described in Section 2.
**References are listed on page 53.

For not correlated data and special forms of the constraint functionals different algorithms have been suggested in a number of publications. Some of the recent papers are by Wentworth (1965)[7], Barieau & Dalton (1966)[8], York (1966)[9], Williamson (1968)[10], O'Neil et al. (1969)[11], Southwell (1969)[12], and Powell & Macdonald (1972)[13]. Of these, Powell & Macdonald's algorithm has convergence properties superior to others. We shall, therefore, compare our numerical results with those reported by Powell & Macdonald. The most advanced theoretical basis for a computing algorithm for general non-linear constraints is given in the paper by Barieau & Dalton[8]. However, their theory is restricted to the case where only one observable contains observational errors. Moreover, the description of the iteration process is in Reference 8 not sufficiently clear.

The estimates of parameter variances are seldom computed correctly. Of the authors mentioned above, Barieau & Dalton, York, Williamson, and Southwell derive correct error expressions. However, these are applicable to the corresponding special cases only.

In Section 2 of this paper we shall demonstrate the theoretical deficiencies of some widely used algorithms and derive necessary formulas for a simple iterative algorithm. The algorithm, which is described in detail in Appendix A is applicable to problems with general constraints, and it has convergence characteristics which are equal or superior to those of other algorithms published. The estimation of parameter variances is treated in Section 3. Section 4 contains examples for the application of the algorithm. Some results are compared in Section 4 with those of other authors.

2. COMPUTATION OF RESIDUALS AND PARAMETERS

Let the mathematical model be given by a functionalrelationship,

$$F(\xi,\theta) = 0, \qquad (1)$$

where ξ is the n-dimensional position vector in the space of the observable variables and θ is the p-dimensional unknown parameter vector. Let r

points X_j $(j = 1,2,\ldots,r)$ be observed in the sample space and $r>p$. Let the observational errors of each point X_j be characterized by a corresponding variance-covariance matrix R_j with the dimension (nxn). That is, we assume no correlation between the different observations X_j, and known variances and covariances of the components of each X_j. Our aim is to compute a "best estimate" t of the parameter vector θ, compatible with the given data X_j and R_j. The "best estimate" we define as that value of θ which is obtained by a least squares analysis of the data.[6]

The least squares analysis is based upon an adjustment of the observed data X_j. Let x_j be the adjusted values and

$$x_j = X_j + c_j \tag{2}$$

The corrections (adjustments, negative residuals) c_j and the best estimate t of the parameter vector are thereby defined as solutions of the following minimization problem with r constraints:[1,2,3]

$$\begin{cases} W = \sum_{j=1}^{r} c_j^T R_j^{-1} c_j = \text{min.}, & (3) \\ F(X_j + c_j, t) = 0, \quad (j = 1,2,\ldots,r). & (4) \end{cases}$$

A commonly used technique for the solution of such a minimization problem makes use of Lagrange multipliers k_j, called also "correlates" in the context of data adjustment[1-5]. According to that technique one minimizes instead of W the functional

$$\tilde{W} = (1/2) \sum_{j=1}^{r} c_j^T R_j^{-1} c_j - \sum_{j=1}^{r} k_j F(X_j + c_j, t), \tag{5}$$

subject to the constraints (4). The correlates k_j are determined simultaneously with the other unknowns, c_j and t.

For further discussion we need some assumptions about the functional F. Since we intend to linearize F by truncating its Taylor expansion, we assume that F can be expanded in a Taylor series within a finite neighborhood of each point (X_j, t). For obvious reasons these neighborhood regions must be sufficiently large to include the points (x_j,t) as

well. In all practical cases these assumptions are satisfied because the normal purpose of the adjustment is to obtain a analytic relation between the ξ-components. F is therefore usually chosen such that it is analytic at least within the above mentioned regions.

Let $A = \text{grad}_{\xi} F$ be the n-dimensional vector of the partial derivatives of F with respect to the components of ξ and $B = \text{grad}_{\theta} F$ be the p-dimensional vector of the partial derivatives of F with respect to the components of θ. We assume that neither A nor B vanish within the above mentioned neighborhoods of (X_j, t). This condition excludes some singular problems (B=0) and singular points (A=0) from our considerations.

We obtain necessary conditions for a minimum of $\tilde{W}$ by setting its partial derivatives with respect to the components of c_j and t equal to zero. Using the symbols A and B these conditions can be expressed by

$$R_j \partial\tilde{W}/\partial c_j = c_j - k_j R_j A(X_j + c_j, t) = 0, \quad (j = 1,2,\ldots r) \quad (6)$$

and

$$\partial\tilde{W}/\partial t = \sum_{j=1}^{r} k_j B(X_j + c_j, t) = 0. \quad (7)$$

Equation (6) is a system of $n \cdot r$ equations, while Equation (7) consists of p equations. Together with the r Equations (4) we have thus $(n + 1)r + p$ equations for an equal number of unknowns, namely, $n \cdot r$ components of the c_j, r correlates k_j and p components of t. Usually the equation system (4), (6) and (7) is non-linear.

In principle it can be solved by any numerical method designed for non-linear systems of equations. In practice such an approach is seldom considered because the number of equations is often very large (due to a large r). Instead, the equations are linearized and the solution obtained by an iteration process. The structure of the linearized equations makes such an approach particularly attractive. As will be shown later, the linearized equations can be manipulated in such a manner that most of the unknowns appear in explicit expressions. The only simultaneous equations remaining are those in a system of p linear equations.

The proper linearization process is the main subject of this section. We shall derive the linearized equations and show that their solutions converge indeed to the solutions of the original problem. In contrast, the usual linearized formulas[1-6] yield wrong solutions, unless F is linear in ξ and B is independent of ξ.

We start our considerations by outlining the usual linearization process and pointing out some of its errors.

Let T be an approximation to t and,

$$\tau = t - T \tag{8}$$

Expanding F in a Taylor series at the point (X_j, T) we obtain,

$$F(X_j + c_j, T + \tau) = F_{jo} + A_{jo}^T c_j + B_{jo}^T \tau + \ldots, \tag{9}$$

where $F_{jo} = F(X_j, T)$, $A_{jo} = A(X_j, T)$, and $B_{jo} = B(X_j, T)$.

The supremum of the error introduced by truncating the series after the linear terms is equal to one half of a maximum value of the second order terms. Let the matrices of the second order derivatives of F be denoted as follows,

$$A_x = \left\{\frac{\partial^2 F}{\partial \xi^i \partial \xi^k}\right\}, \quad A_t = B_x^T = \left\{\frac{\partial^2 F}{\partial \xi^i \partial \theta^k}\right\},$$
$$B_t = \left\{\frac{\partial^2 F}{\partial \theta^i \partial \theta^k}\right\}, \tag{10}$$

where ξ^i, ξ^k, θ^i, and θ^k, are the components of ξ and θ, respectively. The error of the linearization is negligible, if

$$|A_{jo}^T c_j + B_{jo}^T \tau| >> |c_j^T A_x c_j + 2c_j^T A_t \tau + \tau^T B_t \tau|_{max}, \tag{11}$$

whereby the second order derivatives on the right hand side of Equation (11) are evaluated at such a point between (X_j, T) and $(X_j + c_j, T + \tau)$, which yields the maximum value for the expression.

The condition (11) is obviously satisfied if $|c_j|$ and $|\tau|$ are sufficiently small. Let us assume, that the parameters T and the adjustments c_j are computed by an iterative process. If that process converges, then $T \to t$, that is $\tau \to 0$, while the c_j approach their respective least squares values. In the limit the linearization error e_{lin} is then bounded by:

$$|e_{lin}| \leq 1/2 |c_j^T A_x c_j|_{max}. \tag{12}$$

The right hand side of Equation (12) is zero only if F is a linear function of ξ. Hence only in that case we can be confident that the linearization of the constraints F = 0 does not affect the Equation (4).

This result is insofar important as very often the constraint functionals F are linear in the parameters θ but not in the observables ξ.

We shall now consider cases where F is not a linear function of ξ and, moreover, make no assumptions about the size of the corrections c_j. Particularly, we will not assume that,

$$||A_{jo}|| \gg ||A_x c_j||_{max} \tag{13}$$

is satisfied, which condition would justify the linearization (9) in case $A_x \neq 0$. The traditional assumption at this point is, that the $||c_j||$ are sufficiently small so that (13) is satisfied. We note in passing that (13) might not be satisfied even if the $||c_j||$ are "small" because that condition depends on the formulation of the constraint functional F. Thus, the same c_j's can satisfy (13) for one functional F and fail to satisfy (13) if F is formulated differently.

An obvious way to reduce linearization errors is to expand F at a point which is closer to the solution than (X_j, T). This approach has been suggested, for instance, by Tienstra[2] without deriving the corresponding formulas. Powell & Macdonald [13] use the same idea for an algorithm which is different from ours.

Let C_j be an approximation to c_j and,

$$x_j = X_j + C_j + \varepsilon_j. \tag{14}$$

Linearizing the equations (4) at $(X_j + C_j, T)$, we make the conventional assumption, that the second order therms are small compared to the linear terms. Let $F_j = F(X_j + C_j, T)$, $A_j = A(X_j + C_j, T)$ and $B_j = B(X_j + C_j, T)$. The linearization conditions we may express then by the following inequalities, which we assume to be satisfied for $j = 1,2,\ldots,r$:

$$||A_j|| >> ||A_x \varepsilon_j + A_t \tau||_{max} \tag{15}$$

and

$$||B_j|| >> ||\varepsilon_j^T B_x^T + \tau^T B_t||_{max}, \tag{16}$$

whereby the maximum is to be taken between the points $(X_j + C_j, T)$ and $(X_j + C_j + \varepsilon_j, T + \tau)$. Note that, unless A_j and B_j vanish, the inequalities (15) and (16) can be satisfied for sufficiently small ε_j, and τ. (According to our assumptions about F, the A_j and B_j do not vanish in a neighborhood of $\varepsilon_j = 0$ and $\tau = 0$.)

With (15) and (16) satisfied we reduce Equations (4), (6), and (7) to the following system of linear equations for ε_j, k_j, and τ:

$$F_j + A_j^T \varepsilon_j + B_j^T \tau = 0, \quad (j = 1,2,\ldots,r), \tag{17}$$

$$C_j + \varepsilon_j - k_j R_j A_j = 0, \quad (j = 1,2,\ldots,r), \tag{18}$$

and

$$\sum_{j=1}^{r} k_j B_j = 0 \tag{19}$$

We shall now simplify this system of equations. Let,

$$g_j = \frac{1}{A_j^T R_j A_j} \tag{20}$$

be the "weight" of the observation j[1,2]. Multiplying Equation (18) from the left by A_j^T and subtracting the result from Equation (17) we obtain an explicit formula for the correlates k_j:

$$k_j = g_j (A_j^T C_j - F_j - B_j^T \tau), \qquad (j = 1, 2, \ldots, r) \tag{21}$$

Substitution of Equation (21) into Equation (19) yields,

$$\sum_{j=1}^{r} g_j B_j B_j^T \tau = \sum_{j=1}^{r} g_j (A_j^T C_j - F_j) B_j. \tag{22}$$

Substitution of Equation (21) into Equation (18) furnishes an explicit formula for the corrections of the observations:

$$C_j + \varepsilon_j = g_j (A_j^T C_j - F_j - B_j \tau) R_j A_j \qquad (j = 1, 2, \ldots, r). \tag{23}$$

Equation (22) corresponds to the "normal equations" in linear least squares problems. It furnishes the correction τ of T, if C_j and T are given. After the computation of τ, the corresponding corrections $C_j + \varepsilon_j$ of the observations can be computed by Equation (23) and, after replacing T by $T + \tau$ anc C_j by $C_j + \varepsilon_j$, the process repeated. The correlates k_j, if needed, can be computed by Equation (21) at each cycle. Particulars about an iteration algorithm based on Equations (21) through (23) are given in Appendix A and numerical examples shown in Section 4. Here we note only that we have found advantageous to iterate Equation (22) several times (by replacing T by $T + \tau$ in all arguments) before using Equation (23). Also, Equation (23) should be iterated (by replacing C_j by $C_j + \varepsilon_j$) before returning to Equation (22).

If the iteration converges then by definition $C_j \to c_j$ and $T \to t$. The set of Equations (21), (22), and (23) is equivalent to the set of Equations (17), (18), and (19). Hence, at the limit c_j and t satisfy the equations

$$\left.\begin{aligned} F_j &= 0, \\ c_j &= k_j R_j A_j, \\ \sum_{j=1}^{r} k_j B_j &= 0. \end{aligned}\right\} \tag{24}$$

Therefore, the k_j, x_j and t are indeed solutions of the original Equations (4), (6) and (7). The correlates converge thereby to the values

$$k_j = g_i \, A_j^T \, c_j. \tag{25}$$

When the iteration starts, then the $C_j = 0$ and the right hand side of Equation (22) is equal to $-\sum_j g_j F_j B_j$, that is, equal to the commonly used right hand side of normal equations. However, as the iteration proceeds, the F_j's approach zero and the right hand side of the normal equations approaches $\sum g_j A_j^T C_j B_j$, which is the expression in Equation (7). This is the main difference between the commonly used algorithms and ours. Of the cited references only Barieau & Dalton [8] use a right hand side corresponding to Equation (7).

The conventional normal equations for correlated observations are[1,2,3]

$$\sum_j g_{jo} B_{jo} B_{jo}^T \tau = \sum_j k_{jo} B_{jo} \tag{26}$$

with

$$k_j = - g_{jo} F_{jo} \tag{27}$$

The Equations (26) and (27) are iterated until $\tau = 0$. The adjustments are then computed with

$$c_j = k_j \, R_j \, A(X_j, t). \tag{28}$$

The difference between the results obtained by this process and by our algorithm can be readily seen if we compare the corresponding formulas for the correlates, namely Equations (21) and (27). These formulas are identical at convergence only if F is a linear function of ξ. The gradients A_j are then independent of ξ and the c_j, too, are computed by identical formulas (cf. Equation (24) and (28)). However, instead of $\sum_j k_j B(X_j + c_j, t) = 0$, the iteration of Equations (26) and (27) furnishes $\sum_j k_j B(X_j, t) = 0$. Hence, unless F is linear in ξ <u>and</u> B is independent of ξ, the results will not satisfy Equations (7).

We note in passing that these conditions are met in the common case of "curve fitting with errors in one coordinate", whereby the constraints must be solved explicitly for that coordinate. The constraints have then the form,

$$F(\xi,\theta) = \xi - f(\theta) = 0 \tag{29}$$

The functionals f contain generally different constants (the error free coordinates) for different observations. Such a functional F is indeed linear in ξ and the $B_j = \text{grad}_\theta\ f_j(\theta)$ are independent of ξ.

We have thus shown that while usual normal equations do not yield the least squares solution in the general case, only trival changes of these equations are necessary to insure correct solutions for finite corrections. Moreover the employment of the equation system (20) through (23) does not require any special form of the constraint functional F. It is also interesting to note that our equations do not contain higher than first order derivatives of F.

The use of higher order derivatives is advocated by Powell & Macdonald[13]. It is claimed that such derivatives (but only, if they are computed numerically!) will increase the rate of convergence. According to the derivation of our equations, higher order derivatives can be retained either in Equation (18), or in Equation (19), or both, without making the problem non-linear in ε and τ. Inclusion of the term $A_x\varepsilon + A_t\ \tau$ in Equation (18) however, contradicts the assumption (15), which was made for the linearization (17) of F. Similarly the terms $B_x\varepsilon + B_t\tau$ in Equation (19) can be neglected because of the assumption (16). Hence the inclusion of these terms, without using a second degree approximation for F, would introduce an inconsistency in the linearization process. Since it also leads to much more complicated equations we have not attempted to code and test the performance of such an approach. As far as the rate of convergence is concerned, our algorithm was found to perform equally good or better than that of Powell & Macdonald as shown by the examples in Section 4.

3. ESTIMATED VARIANCES AND COVARIANCES OF THE PARAMETERS

An essential assumption which is usually made in error analysis is, that all variances are sufficiently small to permit a linearization of relations involving the variances. We shall carry out an error analysis for the least squares parameters under this assumption. In our case such an assumption contradicts in a sense the assumption made in the previous Section, namely that the estimated corrections c_j are finite. We note, however, that a similar contradiction is inherent in any least squares error analysis. Assuming, for instance, normal distribution of observation errors we admit the possibility of very large errors and yet estimate their effects using a linear model. In our case we did not impose any conditions on the sizes of the corrections c_j in order to obtain correct formulas for the least squares solutions. Once the correct solutions are obtained an error analysis may be very well based on a "small deviation" theory. As usual, the results will not be applicable to large uncertainities but they will correctly express the sensitivity of the results to observational errors.

According to Equations (24) and (25) we have the following relations between the least squares values c_j and t:

$$\sum_{j=1}^{r} g_j B_j A_j^T c_j = 0, \tag{30}$$

$$F_j = 0, \quad j = 1,\ldots,r \tag{31}$$

$$c_j = g_j R_j A_j A_j^T c_j, \quad j = 1,\ldots r. \tag{32}$$

We are interested in variations dX_j of the observations X_j, corresponding variations dt of the parameter t, and variations dc_j of the residuals c_j, such that the Equations (30) through (32) hold. We obtain the desired relations between dX_j, dt and dc_j by differentiating these equations and setting the differentials equal to zero. To simplify the notation we shall omit the index j from our formulas.

Differentiation of Equations (31) yields thus,

$$A^T dX + A^T dc + B^T dt = 0. \tag{33}$$

Differentiation of Equation (30) yields after some manipulation and use of Equation (25)

$$\begin{aligned} &\Sigma\{(-gBc^T A_t + kB_x)\ (dX + dc) + gBA^T dc\} + \\ &\Sigma\,(-gBc^T A_t + kB_t)\ dt = 0 \end{aligned} \tag{34}$$

In order to estimate the variance of t we have to express dt in terms of the observation variations dX, that is, we have to eliminate dc from the equations. The product $A^T dc$ can be eliminated from Equations (34) with the aid of Equation (33). The remaining terms containing dc can be eliminated by making use of the differential of Equations (32), which is

$$\begin{aligned} dc = &-kR(gAA^T R - I)\ A_x(dX + dc) + \\ &+ gRAA^T dc - kR(gAA^T R - I)A_t dt. \end{aligned} \tag{35}$$

After elimination of $A^T dc$ from Equation (35) with the aid of Equation (33) and rearanging of terms we obtain

$$\begin{aligned} &\{I + kR(gAA^T R - I)A_x\}\ dc = \\ &= R\{-gAA^T - k\,(gAA^T R - I)A_x\}dX + \\ &+ R\{-gAB^T - k\,(gAA^T R - I)\ A_t\}dt. \end{aligned} \tag{36}$$

Equation (36) can be solved for the dc if,

$$\det\ \{\ I + kR(gAA^T R - I)A_x\} \neq 0 \tag{37}$$

The condition (37), replaces for our analysis the usual assumptions,

$$I + kR(gAA^T R - I)A_x \approx I \tag{38}$$

or

$$A_x = 0 \tag{39}$$

It is essentially a condition for the sizes of the residuals c, because $k = gA^T c$, according to Equation (25). Let,

$$G = I + kR\ (gAA^T R - I)A_x, \tag{40}$$

$$\Gamma_1 = G^{-1}R\left\{-gAA^T - k(gAA^TR - I)A_x\right\} \tag{41}$$

and

$$\Gamma_2 = G^{-1}R\left\{-gAB^T - k(gAA^TR - I)A_t\right\} \tag{42}$$

Equation (36) yields then under the assumption (37),

$$dc = \Gamma_1\, dx + \Gamma_2\, dt. \tag{43}$$

Elimination of dc between Equation (33), (34), and (43) gives finally the desired relation between dt and dX:

$$\begin{aligned} &\Sigma\left\{gBB^T + gBc^TA_t - kB_t + (gBc^TA_x - k\,B_x)\Gamma_2\right\} dt = \\ &= \Sigma\left\{-gBA^T - (gBc^TA_x - k\,B_x)\,(I + \Gamma_1)\right\} dX. \end{aligned} \tag{44}$$

This equation is of the form

$$\Theta dt = \sum_{j=1}^{r} H_j\, dX_j. \tag{45}$$

with obvious meanings of Θ and H_j. We apply the general error propagation formulas[1,2,3] to Equation (45) in two steps. First, we compute the contribution M_j of the observation X_j to the variance estimate of t. M_j is according to our assumptions about the observations equal to[1,2]:

$$M_j = m_o^2\, H_j R_j H_j^T \quad , \tag{46}$$

where m_o^2 is a proportionality factor to be discussed later. The estimated variance-covariance matrix of the components of t is then equal to

$$V = m_o^2\ \Theta^{-1}\ (\sum_{j=1}^{r} M_j)\ (\Theta^{-1})^T. \tag{47}$$

It can be easily verified, that in case we neglect all second order derivatives in Equation (44), V becomes equal to the conventionally used variance-covariance matrix,

$$V_{conv} = m_o^2\ (\sum_{j=1}^{r} g_j B_j B_j^T)^{-1} \tag{48}$$

which is proportional to the inverse of the normal equation matrix in Equation (22).

We note, however, that even if the constraint functional $F(\xi, \theta)$ is linear in both arguments, ξ and θ, the matrix $A_t = B_x^T$ does not vanish. Hence the conventional variance-covariance matrix $V_{conv} \neq V$ even in linear multi-dimensional adjustment problems.

We pointed out in Section 2 that conventional least squares algorithms provide correct results (parameters and residuals) if the constraint functional is of the form,

$$\xi - f(\theta) = 0 \tag{49}$$

with one-dimensional ξ. For the variance-covariance matrix V we obtain in that case the formulas,

$$\Theta = \Sigma\,(gBB^T - kB_t), \tag{50}$$

$$M_j = m_o^2\, g_j\, B_j\, B_j^T \tag{51}$$

and

$$V = m_o^2\, [\,\Sigma\,(gBB^T - kB_t)]^{-1}\, (\Sigma\, gBB^T)\, \{[\,\Sigma\,(gBB^T - kB_t)]^{-1}\}^T \tag{52}$$

This matrix is equal to the conventional V_{conv} of Equation (48) only if $B_t = 0$, that is, if the functional $f(\theta)$ in Equation (49) is linear in θ.

In summary then the conventional least squares algorithms can be safely used only if the following conditions are all satisfied.

(a) In each observation X_j only one component is adjusted.

(b) The constraint functional is of the form (49).

(c) The functional $f(\theta)$ is linear in the parameter θ.

In all other cases Equation (47) should be used instead of Equation (48). Examples in Section 4 will show that in general the importance of the higher order derivatives cannot be assessed in advance. Either Equation (47) or Equation (48) can furnish larger diagonal elements of the matrix. Moreover, corresponding off-diagonal elements of V and V_{conv}, respectively, can even have opposite signs. These facts makes V_{conv} virtually useless as an approximation to the estimated variance-covariance matrix.

The complexity of the formulas for V is of little practical consequence, because these formulas are evaluated only once for each adjustment problem. The establishing of the necessary second derivative expressions can be a more serious problem. That problem, however, might be attacked by using a symbol manipulation code for the computation of complicated derivatives.

The proportionality factor m_o^2 ("variance of weight one") should be set equal to one, if the variances R_j of the data are well known. Otherwise, mostly the value,

$$\hat{m}_o^2 = \frac{1}{r-p} W = \frac{1}{r-p} \sum_{j=1}^{r} c_j^T R_j^{-1} c_j. \tag{53}$$

is used[1,3,4,6] as an approximate to m_o^2. Macdonald (1969)[14] has shown that a better estimate of m_o^2 is obtained if W in Equation (53) is corrected by a term, which in his case ($n = 2$, $R_j = I$) is the square of an algebraic sum of the weighted residual norms $||c_j||$. The sign of each term is thereby assumed positive or negative, depending on the position of X_j above or below the curve $F = 0$. In our case a corresponding correction can be obtained by observing that W is the weighted sum of the correlate squares, namely,

$$\begin{aligned} W &= \Sigma c^T R^{-1} c = \Sigma k A^T R R^{-1} R A k = \\ &= \Sigma k^2 A^T R A = \Sigma \frac{k^2}{g} \end{aligned} \tag{54}$$

Following Macdonald we compute the algebraic average of the terms $k_j/\sqrt{g_j}$ by,

$$\bar{k} = \frac{1}{r} \sum_{j=1}^{r} \frac{k_j}{\sqrt{g_j}} \tag{55}$$

and define m_o^2 by

$$\begin{aligned} m_o^2 &= \frac{1}{r-p} \sum_{j=1}^{r} \left(\frac{k_j}{\sqrt{g_j}} - \bar{k} \right)^2 = \\ &= \frac{1}{r-p} (W - r\bar{k}^2). \end{aligned} \tag{56}$$

The correlates k_j are related to the residuals by,

$$k_j = g_j c_j^T A_j = g_j c_j^T \operatorname{grad}_\xi F_j. \tag{57}$$

The sign of k_j depends, therefore, on the location of the observation X_j with respect to the hypersurface F=0: for all X_j on one side of the hypersurface the products $c_j \operatorname{grad}_\xi F_j$ have equal sign, if we assume that the adjustments are such, that the straight line from X_j to x_j does not penetrate the hypersurface. (If it does, our solution is not the best least squares solution. Such singular cases we do not consider here.) Hence the sign convention in Equation (56) is the same as the one used by Macdonald for curve fitting.

4. EXAMPLES

In this Section we shall present some examples of computations by the method described in previous Sections. First, we shall compare the performance of our method with that of Powell & Macdonald[13], second, we shall demonstrate the importance of correct variance estimates in an example of recent work at BRL and, third, we shall give an example of approximation by multivalued curves. For all computations presented here, the computer code was used which is described in Appendix A.

In the first group of examples polynomials are fitted in a two-dimensional space to a data set of 10 points. Let the coordinates in the space of observables be called x and y, respectively. The polynomial constraints are then of the form,

$$F(x,y;\theta) = y - \sum_{i=1}^{p} \theta_i \, x^{i-1} = 0$$

The data (Table I) are taken from Reference 13. The coordinates of the data points were originally given by Pearson (1901)[14] and the weights are due to York (1966)[9]. For our calculations we computed the variance-covariance matrices R from York's weights by the formulas,

$$r_{11} = 1/w_1 \,,$$

$$r_{22} = 1/w_2 \,,$$

$$r_{12} = r_{21} = 0,$$

where r_{ik} are the elements of R. We have also fitted Pearson's data assuming unit weights. In those cases all the R's were assumed to be unit matrices.

Whenever it was possible, we have compared our results with those of Powell & Macdonald[13]. Those authors have shown that their algorithm is better for fitting Pearson's data than a number of other methods. Therefore, a comparison with Powell & Macdonald's results is a sensible test for our method. In the sequel we shall refer to Powell & Macdonald's method by calling it the PM algorithm.

Table II displays the results of a linear fit to Pearson's data with unit weights. The crucial end condition for the iteration in this and all other examples was that the absolute changes of the parameters between cycles become less than 10^{-7} times a "crude standard error" of the corresponding parameter. (For particulars see Appendix A.) The "crude standard error" was thereby obtained by taking the inverse of the normal equation matrix times m_o^2 as an approximation to the variance-covariance matrix. These crude error estimates were computed at the end of every cycle for the sole purpose to check the end conditions. The final standard error estimates were obtained after completed iteration using the formulas derived in Section 3. The differences between both estimates show that second order derivatives have an effect on error estimates even in this rather trivial linear case.

Another result which is typical for all our computations is the fact that W becomes stationary within computing accuracy (about 15 decimal digits) before the end of the iterations. This indicates, that near its minimum W is relatively insensitive to small changes of the parameters.

The results of the computation are displayed graphically in Figure 1a, where the observed points, corrected observations, the error ellipse for each observation, and the fitted line with its one standard error confidence limits are plotted. The error ellipses are circles in the present case with unit weights. The factor m_o, computed with Equation (56), is included in the graphical as well as numerical results in this and other examples.

Table III and Figure 1b show the results of a linear fit to Pearson's data with York's weights. This case can be compared with similar calculations by the PM algorithm,[13] the results of which are included in Table III. The comparison indicates a slightly better rate of convergence for our method. The respective final parameter values are equal within five places for both methods. There are, however, significant differences in the error estimates computed by our method and by Powell & Macdonald. Our estimates, which are computed using the formulas of Section 3, agree with those computed by Williamson (1968)[10] and Southwell (1969)[12]. The estimates of the PM algorithm are obtained in the same manner as our "crude error estimates", that is, by inverting the normal equation matrix. The Powell & Macdonald error estimates differ from our crude estimates because the normal equation matrix of the PM method is different from ours. Elements of that matrix consist in the PM method of certain second order derivatives of W which are computed by numerical differentiation. No theoretical basis for the computation of error estimates in this manner is given in Reference 13. Moreover, in the PM algorithm certain derivatives are neglected in order to simplify the equations. We conclude, therefore, that the error estimates obtained by Powell & Macdonald are smaller because significant terms in the error equations have been neglected.

Tables IV and V and Figures 2c and 2b give the results of fitting a cubic to Pearson's data. In case of unit weights the results can again be compared with those of the PM algorithm. The respective parameters computed by either method agree within the five places given in Reference 13. The rates of convergence are comparable. The error estimates are again different. Our formulas furnish estimates which are bigger than the conventionally computed crude estimates, while the estimates given by Powell & Macdonald are smaller than our crude ones.

Tables VI and VII, and Figures 3A and 3B give the results of quintic fits to Pearson's data. The comparison with the PM algorithm in the case of unit weights (Table VI) reveals a weakness of that algorithm in addition to the inaccurate error estimates. This time the

parameter values computed by Powell & Macdonald are different from ours. The relative magnitudes of the differences are as large as $2 \cdot 10^{-2}$. Differences of such magnitude are suspicious because the iteration end conditions in both algorithms required that the relative changes of the parameters between cycles be less than 10^{-6} (less than 10^{-7} for the PM algorithm). In order to find out which is the "better" parameter set we have recalculated the W value for the parameters of the PM algorithm. (Powell & Macdonald give that value with five places only.) The result, shown in Table VI, indicates that our parameter set produces a smaller W value. The reason why the PM algorithm fails in this case to converge to a minimum of W is likely to be found in the computation of derivatives of W by numerical differentiation. Since near its minimum W is rather insensitive to changes of parameter values, the numerically computed second derivatives have low accuracies. One would expect that this behavior of W has a detrimental effect on the results of the PM algorithm, particularly when the number of parameters is not very small. Our results, on the other hand, are not influenced by the above mentioned behavior of W, because the numerical value of W is not used in our calculations.

Our next example illustrates the importance of accurate error estimation procedures. In this example the constraint functional has the form,

$$F = \alpha - \alpha_o \frac{f}{f_o}\left[\left(\frac{p_o}{p}\right)^n \frac{f}{f_o}\right]^a \cdot \exp\left\{1-\left[\left(\frac{p_o}{p}\right)^n \frac{f}{f_o}\right]^a\right\} \cdot \left(\frac{p}{p_o}\right)^m \cdot \exp\left\{1-\left(\frac{p}{p_o}\right)^m\right\}$$

where,

α, f and p are observables,

α_o, f_o, p_o and a are parameters,

and

$n = m = 0.44$.

This approximation problem arises in the evaluation of acoustic amplification measurements. (About the physical background of such measurements see Ibiricu (1966)[16].) The data set to be fitted is given in Table VIII and the results are shown in Table IX. The crude error estimates are in this case of the same order as our accurate estimates. However, comparing corresponding elements in the full variance-covaraince matrices we notice that one of the off-diagonal elements has a wrong sign in the inverted normal equation matrix. Hence the use of that matrix as an approximation to the variance-covariance matrix can lead to serious errors, even if the diagonal elements (i.e., parameter error estimates) are of the right order. If different methods for variance estimates are compared numerically, then obviously the full variance-covariance matrices should be compared and not only the diagonal elements. Fitting the same functional to other data sets we obtained in some cases parameter error estimates which were up to three times larger than the corresponding crude estimates, while in other cases the crude estimates were larger. It is thus obvious that the conventional crude error estimates are practically useless, because one never knows whether they are too large or to small. Moreover, the signs of the crude off-diagonal covariance estimates can be wrong.

Our next example demonstrates the usefullness of the general formulation of the constraint functional in the case of multi-valued functions. In order to make the presentation of the results simple we have chosen a two-dimensional space of observables (x,y - plane) and fitted a generalized Cassinian curve. This example shows also how correlations between observables naturally enter the calculations.

We assumed that the distance r and the angles of sight φ were measured independently and that the corresponding error estimates were e_r and e_φ, respectively. The cartesian coordinates of an observation (r,φ) are then given by,

$$x = r \cos\varphi$$
$$y = r \sin\varphi$$

while the variance-covariance matrix for x and y is,

$$R = \begin{pmatrix} e_r^2 \cos^2\varphi + r^2 e_\varphi^2 \sin^2\varphi & (e_r^2 - r^2 e_\varphi^2) \sin\varphi \cos\varphi \\ (e_r^2 - r^2 e_\varphi^2) \sin\varphi \cos\varphi & e_r^2 \sin^2\varphi + r^2 e_\varphi^2 \cos^2\varphi \end{pmatrix}$$

The cartesian coordinates of the (assumed) observations and the error estimates e_r and e_φ are given in Table X.

The constraint functional was formulated as follows,

$$F = [(x - x_1)^2 + (y - y_1)^2] \cdot [(x - x_2)^2 + (y - y_2)^2 \cdot b] - a = 0$$

This functional depends on six parameters, namely x_1, y_1, x_2, y_2, b and a. For b = 1 and a > 0 we have a regular Cassinian curve. A reformulation of F = 0 in polar coordinates is not necessary because our method takes the correlations between x and y into account.

The curves F = 0 are multiple valued for both coordinates, x and y. Since in our method the constraint functional need not be solved for one observable, no further manipulation of the constraint functional is necessary, except for the computation of derivatives of F.

The results of the adjustment are shown in Table XI and in Figures 4a and 4b. Comparing the estimated variance-covariance matrix with the inverted normal equation matrix we notice in this case again differences of signs of some off-diagonal elements. A consequence of these and other differences between corresponding elements can be seen by comparing the confidence curves in Figure 4a with those of Figure 4b.

In order to test the significance of the correlations between x and y we have adjusted the same data assuming unit weights of observations. The results are given in Table XII and Figures 5a and 5b. As expected, the final parameter values are different from those of Table XI.

The purpose of the last example was to illustrate the treatment of correlated data and to show the influence of the correlations on the results. Since the theory presented in this paper is restricted to constant variance-covariance matrices R_j, we have carried out all calculations under that assumption. Constant R_j are typical for situations where the variances and covariances of the data are obtained by an analysis of the measurement process and/or by linear coordinate transformations.

In the case described in the last example, however, such a treatment is strictly speaking, not correct. The R_j were, namely obtained from the original variance-covariance matrices (diagonal and constant in the example) by a non-linear coordinate transformation. One consequence of such a transformation is, that the elements of R_j are not constants but depend on the coordinates x and y, and are subject to change whenever x and y are corrected. A more serious consequence is that the basic equations of Sections 2 and 3 must be modified if non-linear coordinate transformations are involved. We intend to treat such transformations in a forthcomming paper.

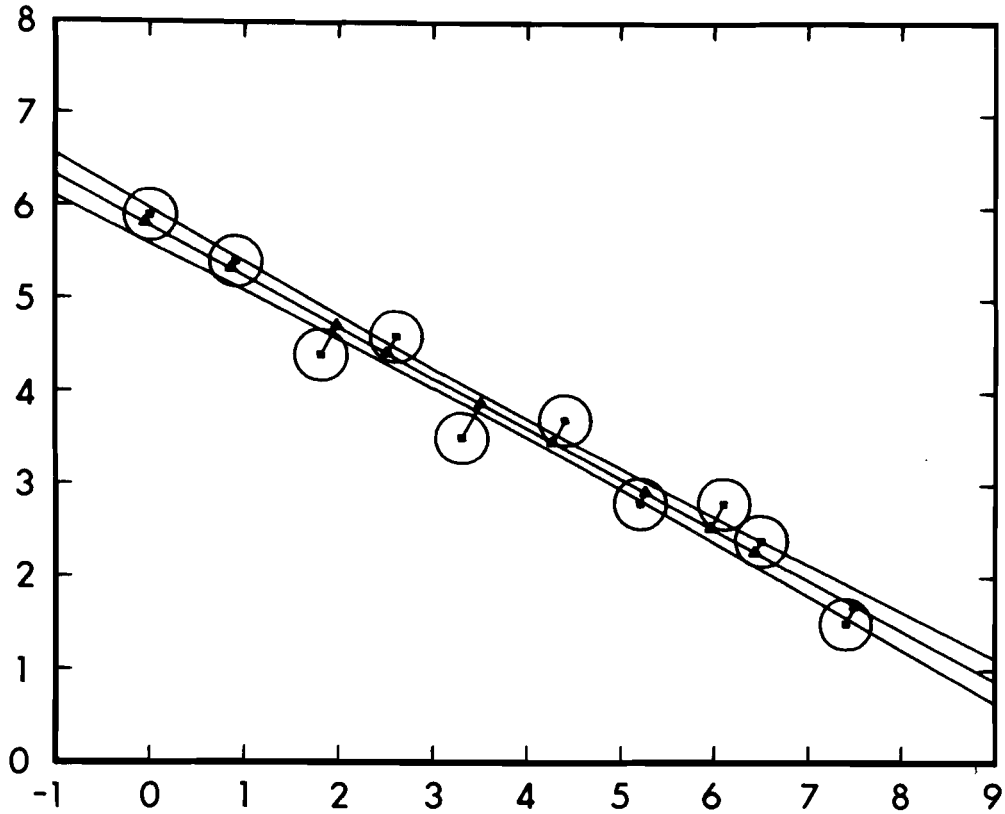

FIGURE 1a.

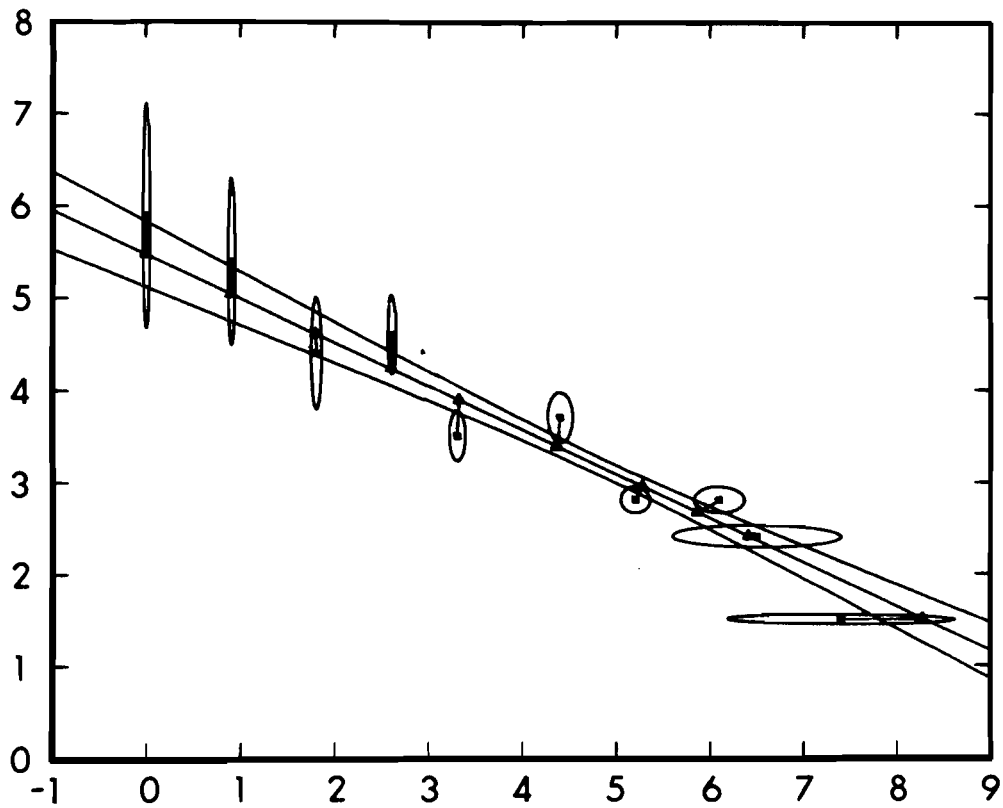

FIGURE 1b.

LINEAR FIT TO PEARSON'S DATA

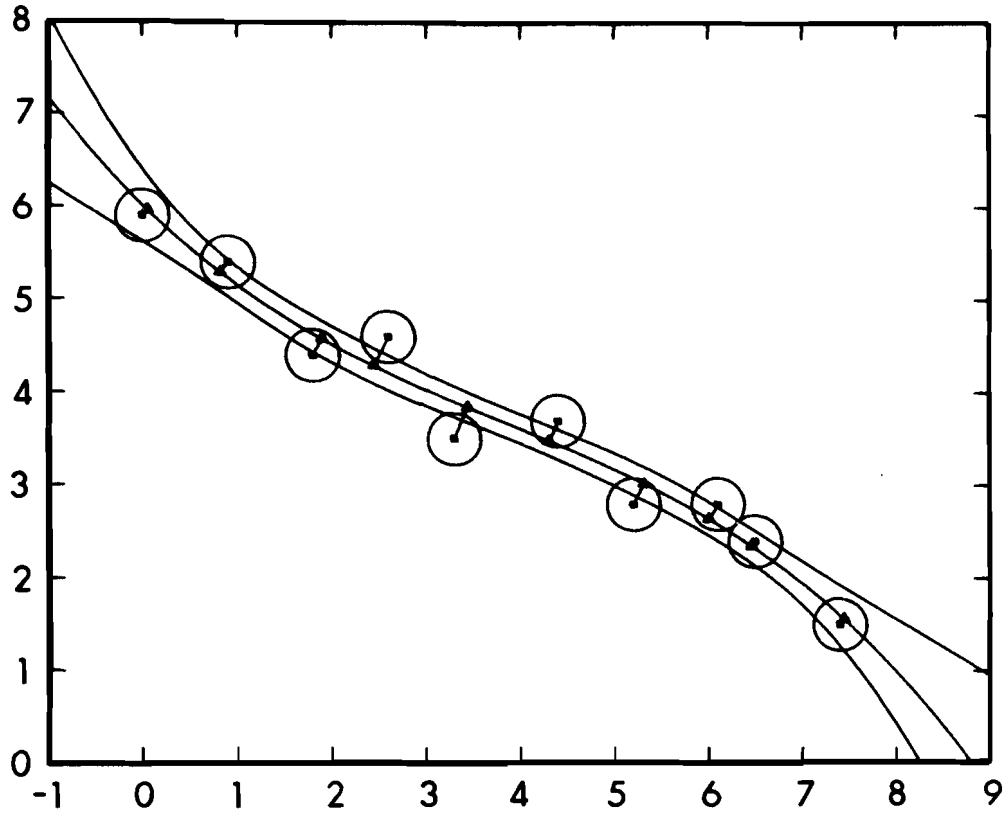

FIGURE 2a.

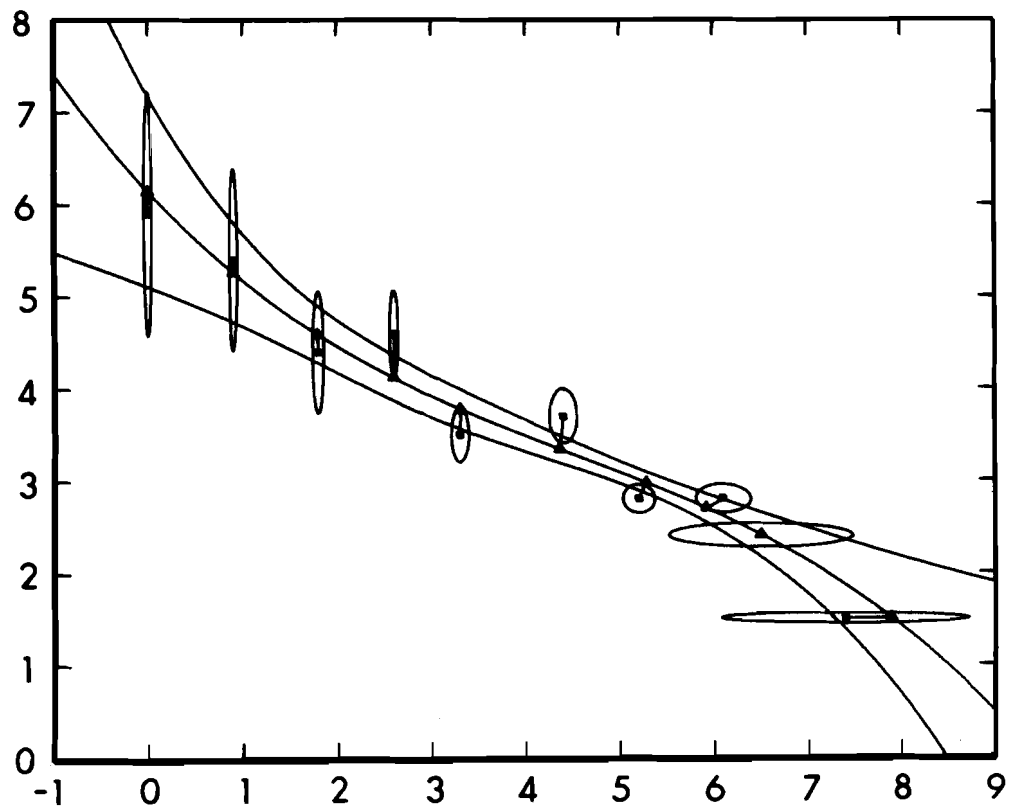

FIGURE 2b.

CUBIC FIT TO PEARSON'S DATA

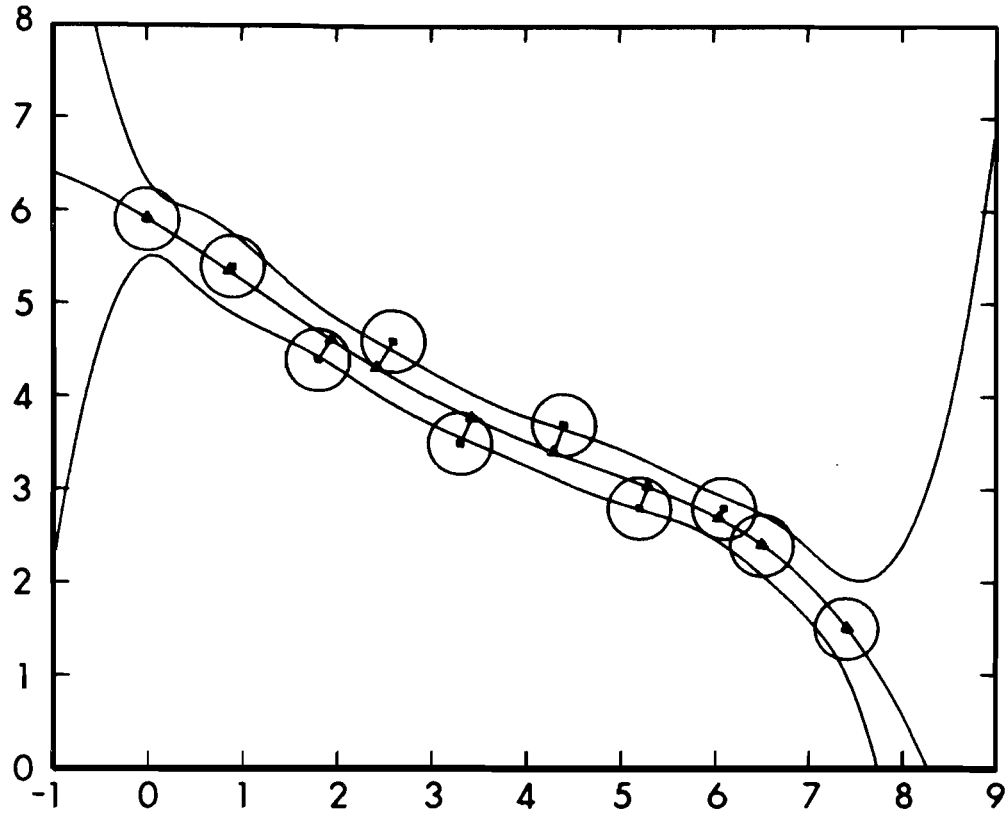

FIGURE 3a.

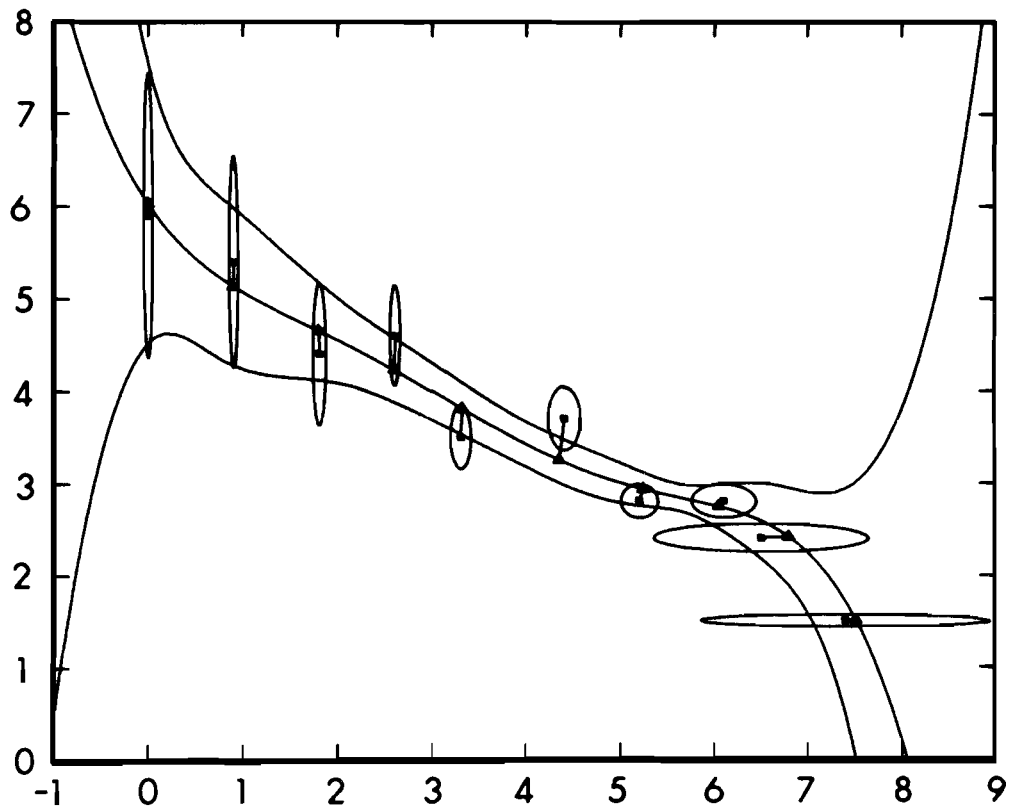

FIGURE 3b.

QUINTIC FIT TO PEARSON'S DATA

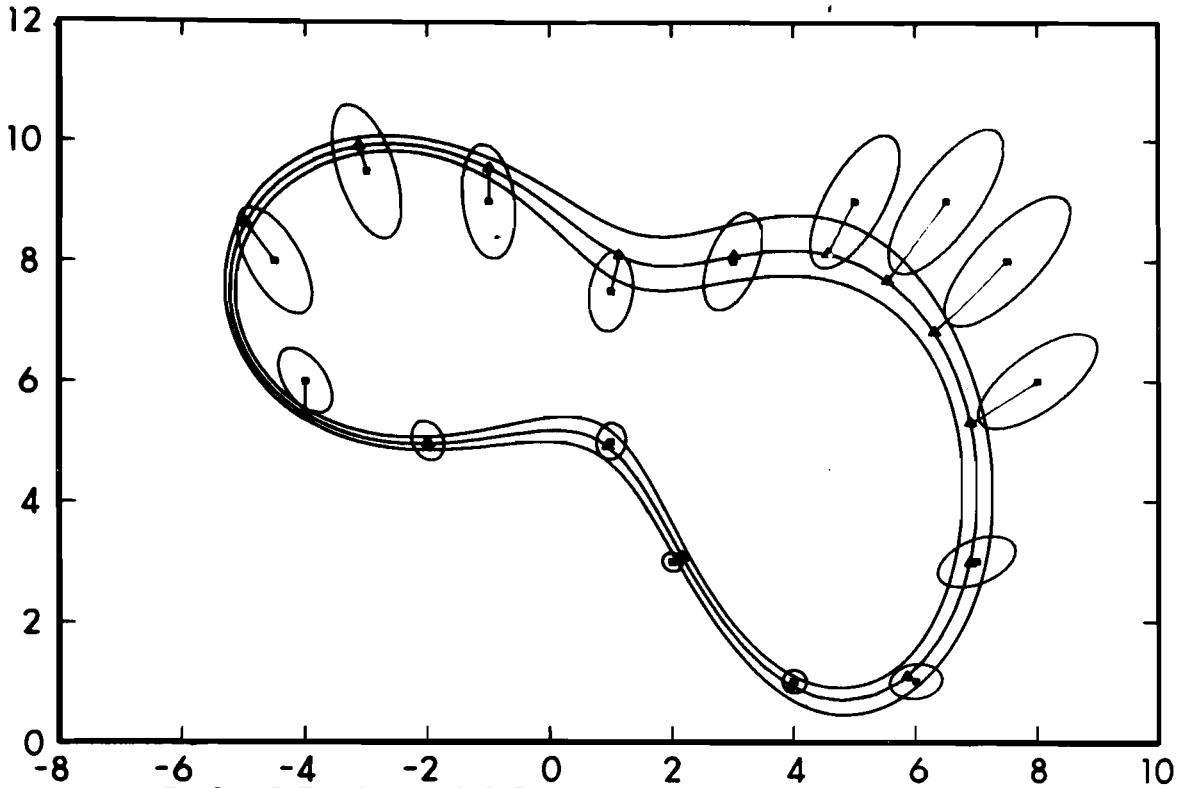

FIGURE 4a. CORRECT ERROR ESTIMATES.

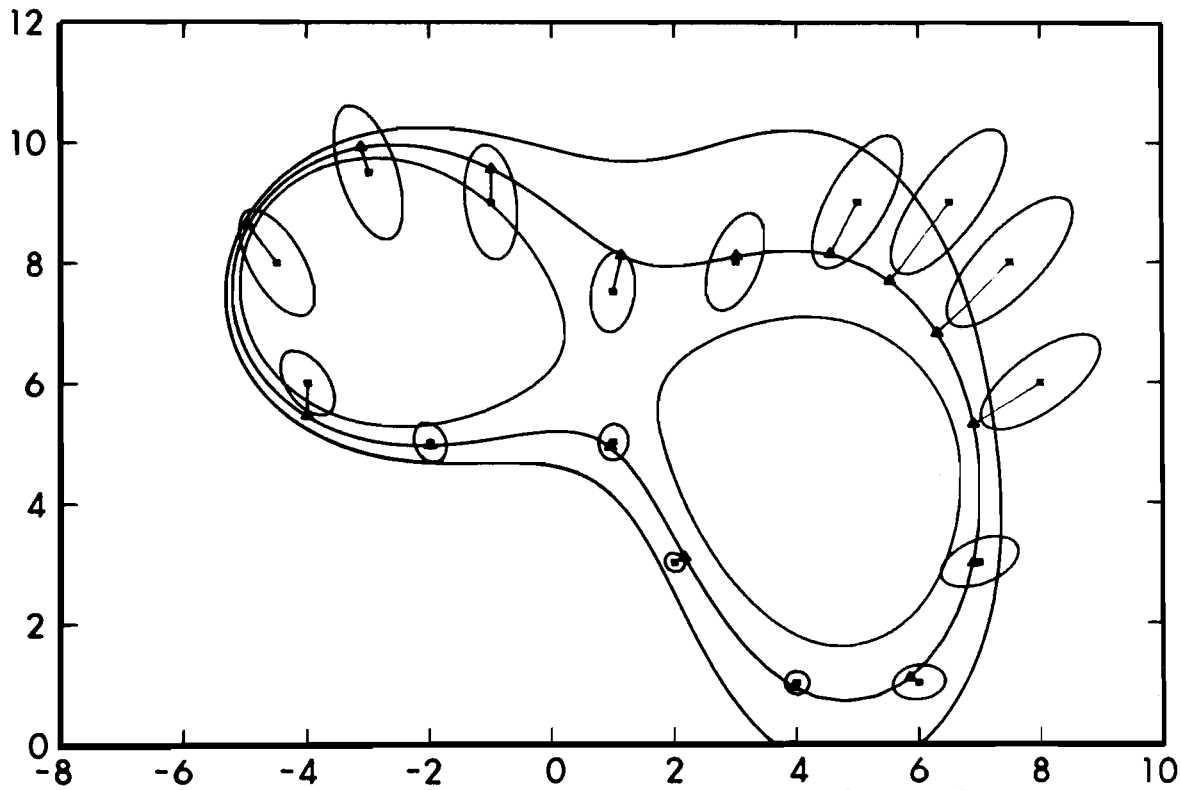

FIGURE 4b. CONVENTIONAL ERROR ESTIMATES.

FITTING OF PSEUDO-CASSINIAN CURVES

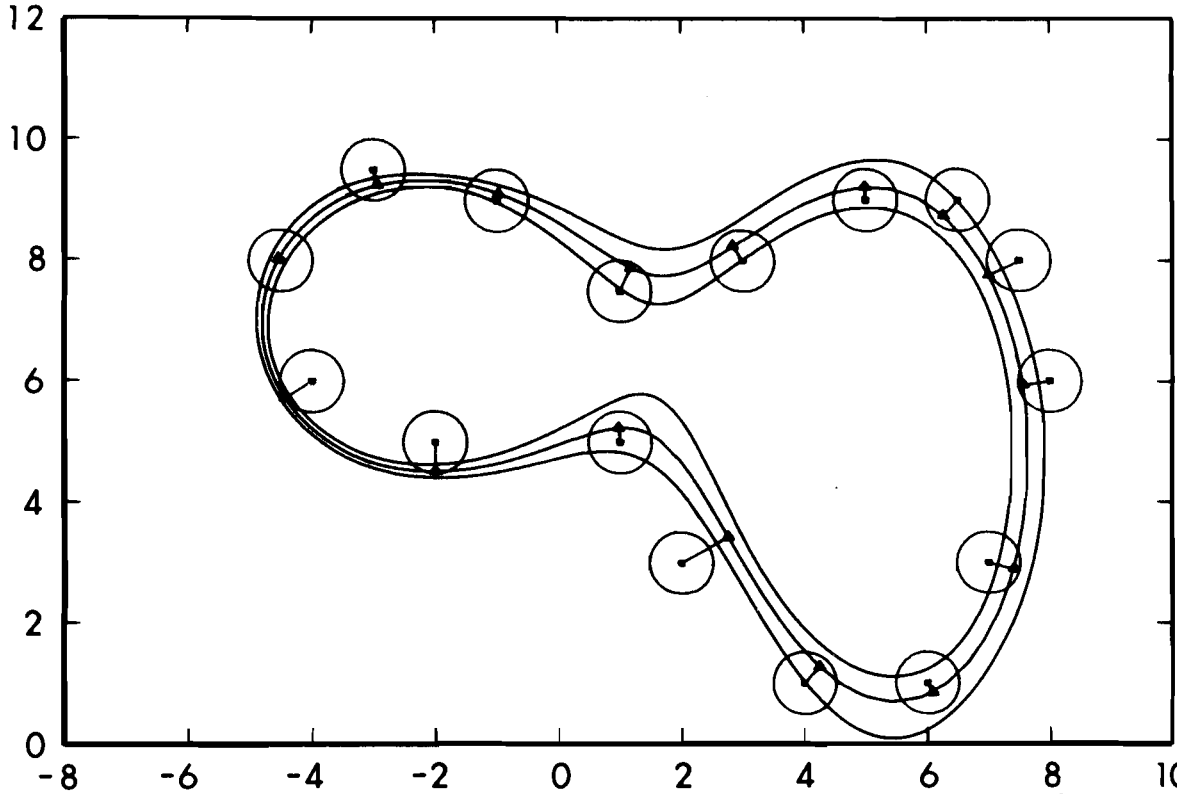

FIGURE 5a. CORRECT ERROR ESTIMATES.

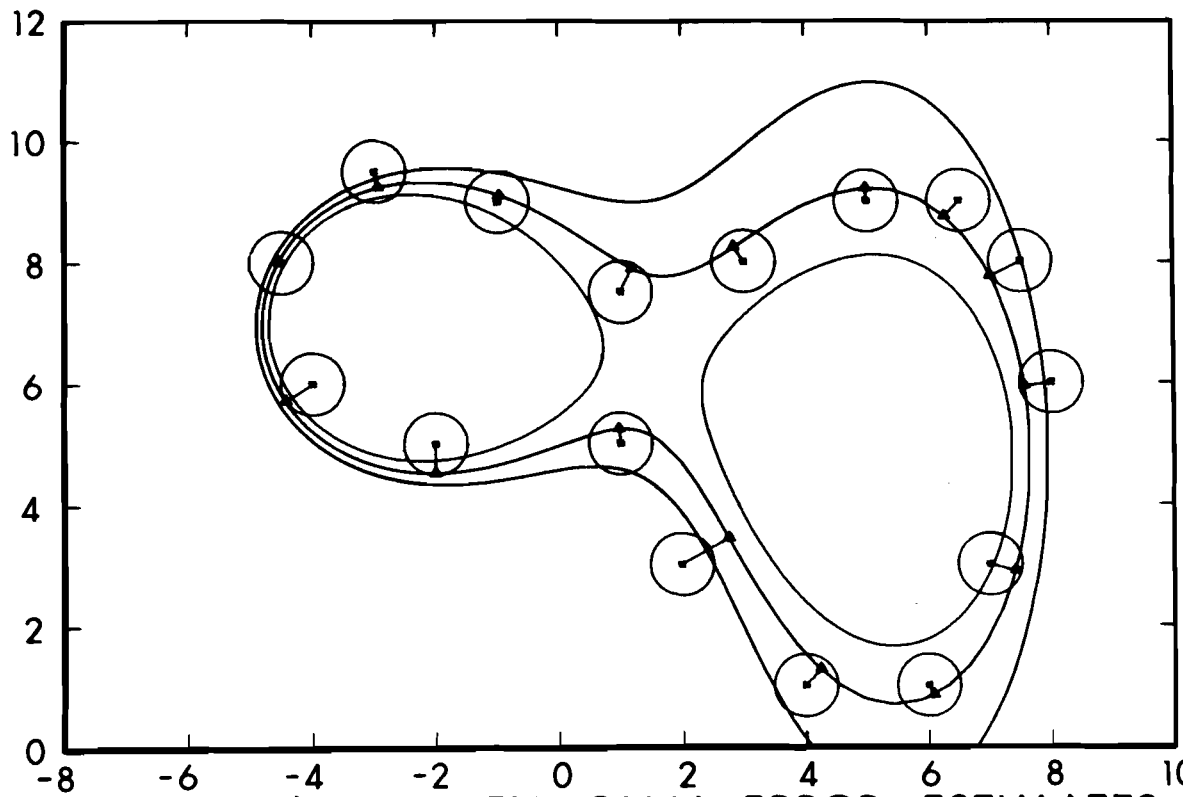

FIGURE 5b. CONVENTIONAL ERROR ESTIMATES.

FITTING OF PSEUDO-CASSINIAN CURVES

TABLE I

PEARSON'S DATA AND YORK'S WEIGHTS

Nr	$X_1(=x)$	$X_2(=y)$	W_1	W_2
1	0.0	5.9	1000.0	1.0
2	0.9	5.4	1000.0	1.8
3	1.8	4.4	500.0	4.0
4	2.6	4.6	800.0	8.0
5	3.3	3.5	200.0	20.0
6	4.4	3.7	80.0	20.0
7	5.2	2.8	60.0	70.0
8	6.1	2.8	20.0	70.0
9	6.5	2.4	1.8	100.0
10	7.4	1.5	1.0	500.0

TABLE II

LINEAR FIT TO PEARSON'S DATA WITH UNIT WEIGHTS

Cycle Nr.	T_1	T_2	W
0	0.	0.	154.
1	5.761 185 19	-0.539 577 275	0.620 119 637
2	5.783 849 45	-0.545 510 327	0.618 572 871
3	5.784 042 12	-0.545 560 765	0.618 572 759
4	5.784 043 76	-0.545 561 193	0.618 572 759
5	5.784 043 77	-0.545 561 197	0.618 572 759
Standard Error	0.1917	0.04277	
Crude St. Error	0.1899	0.04223	

$W = 0.618\ 572\ 759\ 437$

Last Change of $W = 4 \cdot 10^{-16}$

$\overline{k}^2 = 8 \cdot 10^{-35}$

$m_o = 0.278\ 067\ 6$

Variance-Covariance Matrix

$$\begin{pmatrix} 3.673 \cdot 10^{-2} & -6.989 \cdot 10^{-3} \\ -6.989 \cdot 10^{-3} & 1.830 \cdot 10^{-3} \end{pmatrix}$$

TABLE III

LINEAR FIT TO PEARSON'S DATA WITH YORK'S WEIGHTS

Cycle Nr.	Present Method		Powell & Mac Donald[13]	
	T_1	T_2	T_1	T_2
0	0	0	5.396 1	-0.463 45
1	5.396 016 28	-0.463 441 754	5.398 2	-0.463 88
2	5,479 941 92	-0.480 540 364	5.477 5	-0.479 98
3	5.479 909 72	-0.480 533 321	5.479 9	-0.480 53
4	5.479 910 23	-0.480 533 409	5.479 9	-0.480 53
5	5.479 910 22	-0.480 533 407	-	-
Standard Error	0.3549	0.07004	0.252	0.0496
Crude St. Error	0.3585	0.07048		

W = 11.866 353 1941

Last change of W $3 \cdot 10^{15}$

$\bar{k}^2 = 4.573 \cdot 10^{-3}$

$m_o = 1.215\ 556$

Variance - Covariance Matrix

$$\begin{pmatrix} 1.259 \cdot 10^{-1} & -2.392 \cdot 10^{-2} \\ -2.392 \cdot 10^{-2} & 4.905 \cdot 10^{-3} \end{pmatrix}$$

W = 11.866 353

$m_o = 1.218$

TABLE IV

CUBIC FIT TO PEARSON'S DATA WITH UNIT WEIGHTS

Cycle Nr.	T_1	T_2	T_3	T_4
0	0.	0.	0.	0.
1	5.998 796 76	-1.004 987 63	0.157 037 148	$-1.371\ 381\ 65 \cdot 10^{-2}$
2	6.015 242 55	-1.000 068 79	0.152 558 564	-1.324 759 43
3	6.015 277 65	-0.999 855 467	0.152 477 162	-1.324 093 64
4	6.015 265 21	-0.999 837 200	0.152 472 109	-1.324 056 72
5	6.015 263 84	-0.999 835 470	0.152 471 633	-1.324 053 08
6	6.015 263 74	-0.999 835 356	0.152 471 604	-1.324 052 87
7	6.015 263 73	-0.999 835 347	0.152 471 602	-1.324 052 86
Standard Error	0.3868	0.4400	0.1341	$1.153 \cdot 10^{-2}$
Crude St. Err.	0.3663	0.4098	0.1276	$1.121 \cdot 10^{-2}$
P & M[13] St. Err.	0.265	0.298	0.0918	$0.80 \cdot 10^{-2}$

W = 0.485 152 486 927

Last Change of W = $8 \cdot 10^{-17}$

$\overline{K}^2 = 1.404 \cdot 10^{-7}$

$m_o = 0.284\ 356\ 3$

Variance-Covariance Matrix Times 10^3
(Upper Half Shown Only)

$$\begin{pmatrix} 1496. & -140.9 & 35.59 & -2.637 \\ & 193.6 & -56.87 & 4.586 \\ & & 17.99 & -1.521 \\ & & & 0.1329 \end{pmatrix}$$

TABLE V

CUBIC FIT TO PEARSON'S DATA WITH YORK'S WEIGHTS

Cycle Nr.	T_1	T_2	T_3	T_4
0	0.	0.	0.	0.
1	6.028 421 70	-0.916 872 84	0.085 556 903	$-0.432\ 499\ 89 \cdot 10^{-2}$
2	6.258 917 93	-1.256 138 96	0.203 086 978	-1.566 544 00
3	6.130 137 20	-1.095 986 50	0.154 063 801	-1.133 672 56
4	6.142 062 26	-1.108 148 28	0.157 130 861	-1.155 755 74
5	6.142 213 53	-1.108 214 45	0.157 114 212	-1.155 325 34
6	6.142 353 27	-1.108 387 74	0.157 166 014	-1.155 767 22
7	6.142 316 13	-1.108 335 36	0.157 148 606	-1.155 604 69
8	6.142 334 88	-1.108 360 68	0.157 156 748	-1.155 678 74
9	6.142 326 99	-1.108 349 92	0.157 153 261	-1.155 646 83
10	6.142 329 74	-1.108 353 56	0.157 154 414	-1.155 657 20
11	6.142 329 38	-1.108 353 17	0.157 154 310	-1.155 656 39
12	6.142 329 42	-1.108 353 23	0.157 154 333	-1.155 656 62
13	6.142 329 40	-1.108 353 20	0.157 154 320	-1.155 656 51
Standard Error	1.028	0.7692	0.1794	$1.324 \cdot 10^{-2}$
Crude Error	1.034	0.8214	0.2102	$1.702 \cdot 10^{-2}$

W = 10.486 904 057 7

Last Change of W = 10^{-14}

$\bar{k}^2 = 2.352 \cdot 10^{-3}$

$m_o = 1.320\ 567$

Variance-Covariance Matrix Times 10^3
(Upper Half Shown Only)

$$\begin{pmatrix} 1058. & -730.8 & 149.6 & -9.334 \\ & 591.7 & -133.4 & 8.984 \\ & & 32.19 & -2.305 \\ & & & 0.1753 \end{pmatrix}$$

TABLE VI

QUINTIC FIT TO PEARSON'S DATA WITH UNIT WEIGHTS

i	Present Method			Powell & MacDonald[13]	
	Parameters t_i	Last ΔT_i	Standard Errors	Parameters t_i	Standard Errors
1	5.914 825 96	$-6 \cdot 10^{-9}$	0.411 9	5.915	0.288
2	-0.603 166 896	$5 \cdot 10^{-8}$	1.748 0	-0.6034	1.20
3	$-8.032\ 030\ 78 \cdot 10^{-2}$	$-5 \cdot 10^{-8}$	$168.9 \cdot 10^{-2}$	$-8.003 \cdot 10^{-2}$	$117 \cdot 10^{-2}$
4	$2.632\ 202\ 02 \cdot 10^{-2}$	$2 \cdot 10^{-8}$	$60.13 \cdot 10^{-2}$	$2.622 \cdot 10^{-2}$	$41.7 \cdot 10^{-2}$
5	$-8.277\ 185\ 40 \cdot 10^{-4}$	$-3 \cdot 10^{-9}$	$896.8 \cdot 10^{-4}$	$-8.119 \cdot 10^{-4}$	$623 \cdot 10^{-4}$
6	$-1.675\ 050\ 59 \cdot 10^{-4}$	$1 \cdot 10^{-10}$	$47.46 \cdot 10^{-4}$	$-1.683 \cdot 10^{-4}$	$33 \cdot 10^{-4}$
	Number of Cycles: 10 (Starting with $T_i \equiv 0$) W = 0.450 325 667 217 Last Change of W = $3 \cdot 10^{-16}$ $\bar{k}^2 = 1.136 \cdot 10^{-8}$ m_o = 0.335 531 50			Number of Cycles: 10 (Starting with Approxim.) W = 0.450 329 283 54 (Computed using (A-1), (A-2), and (A-4)) m_o = 0.33 55	

Variance - Covariance Matrix Times 10^3
(Upper Half Shown Only)

$$\begin{pmatrix} 169.7 & -428.5 & 316.1 & -96.51 & 12.95 & -0.6330 \\ & 3056. & -2869. & 977.3 & -139.6 & 7.100 \\ & & 2854. & 1005. & 146.8 & -7.583 \\ & & & 361.6 & -53.62 & 2.803 \\ & & & & 8.043 & -0.4241 \\ & & & & & 2.253 \cdot 10^{-2} \end{pmatrix}$$

TABLE VII

QUINTIC FIT TO PEARSON'S DATA WITH YORK'S WEIGHTS

i	Parameters t_i	Last ΔT_i	Standard Errors	Crude St. Errors
1	6.029 451 86	$2 \cdot 10^{-8}$	1.508	1.503
2	-1.530 034 23	$-2 \cdot 10^{-7}$	3.539	3.419
3	0.817 877 33	$2 \cdot 10^{-7}$	2.805	2.647
4	-0.294 920 02	$-7 \cdot 10^{-8}$	0.9164	0.8548
5	$4.698\ 541\ 20 \cdot 10^{-2}$	$1 \cdot 10^{-8}$	$13.16 \cdot 10^{-2}$	$12.30 \cdot 10^{-2}$
6	$-2.666\ 420\ 13 \cdot 10^{-3}$	$-6 \cdot 10^{-10}$	$6.876 \cdot 10^{-3}$	$6.528 \cdot 10^{-3}$

NUMBER OF CYCLES: 13 (Starting with $T_i \equiv 0$)

W = 9.505 013 741 86

Last Change of W = $9 \cdot 10^{-14}$

$\bar{k}^2 = 1.931 \cdot 10^{-3}$

m_o = 1.539 944

Variance-Covariance Matrix Times 10^3
(Upper Half Shown Only)

$$\begin{pmatrix} 2274. & -3861. & 2268. & -602.3 & 74.09 & -3.430 \\ & 12520. & -9536. & 2934. & -397.5 & 19.71 \\ & & 7869. & -2535. & 354.2 & -17.96 \\ & & & 839.8 & -119.7 & 6.159 \\ & & & & 17.31 & -0.9006 \\ & & & & & 4.728 \cdot 10^{-2} \end{pmatrix}$$

TABLE VIII

INPUT FOR ACOUSTIC AMPLIFICATION PROBLEM

α	e_α	f[Hz]	e_f	p[Pa]	e_p
14	2.1	500	20	$1.482 \cdot 10^6$	$0.08274 \cdot 10^6$
29	3.5	1000	30	$1.482 \cdot 10^6$	$0.08274 \cdot 10^6$
57	3.4	2000	50	$1.482 \cdot 10^6$	$0.08274 \cdot 10^6$
82	5.3	3000	60	$1.482 \cdot 10^6$	$0.08274 \cdot 10^6$
103	7.2	4000	80	$1.482 \cdot 10^6$	$0.08274 \cdot 10^6$
122	9.1	5000	100	$1.482 \cdot 10^6$	$0.08274 \cdot 10^6$
133	10.6	6000	120	$1.482 \cdot 10^6$	$0.08274 \cdot 10^6$
135	11.4	7000	140	$1.482 \cdot 10^6$	$0.08274 \cdot 10^6$
131	11.8	8000	160	$1.482 \cdot 10^6$	$0.08274 \cdot 10^6$
125	11.9	9000	180	$1.482 \cdot 10^6$	$0.08274 \cdot 10^6$
117	11.7	10000	200	$1.482 \cdot 10^6$	$0.08274 \cdot 10^6$
8.5	1.3	500	20	$2.861 \cdot 10^6$	$0.1241 \cdot 10^6$
18	2.1	1000	30	$2.861 \cdot 10^6$	$0.1241 \cdot 10^6$
37	2.2	2000	50	$2.861 \cdot 10^6$	$0.1241 \cdot 10^6$
55	3.6	3000	60	$2.861 \cdot 10^6$	$0.1241 \cdot 10^6$
73	5.1	4000	80	$2.861 \cdot 10^6$	$0.1241 \cdot 10^6$
88	6.6	5000	100	$2.861 \cdot 10^6$	$0.1241 \cdot 10^6$
99	7.9	6000	120	$2.861 \cdot 10^6$	$0.1241 \cdot 10^6$
104	8.8	7000	140	$2.861 \cdot 10^6$	$0.1241 \cdot 10^6$
105	9.5	8000	160	$2.861 \cdot 10^6$	$0.1241 \cdot 10^6$

TABLE VIII

(con't)

α	e_α	f[Hz]	e_f	p[Pa]	e_p
103	9.8	9000	180	$2.861 \cdot 10^6$	$0.1241 \cdot 10^6$
101	10.1	10000	200	$2.861 \cdot 10^6$	$0.1241 \cdot 10^6$
13	1.0	2000	50	$5.619 \cdot 10^6$	$0.2206 \cdot 10^6$
24	1.6	3000	60	$5.619 \cdot 10^6$	$0.2206 \cdot 10^6$
28	2.0	4000	80	$5.619 \cdot 10^6$	$0.2206 \cdot 10^6$

TABLE IX

PARAMETERS OF ACOUSTIC AMPLIFICATION PROBLEM

	PARAMETERS		Last Changes	Standard Errors	Crude Standard Errors
	Initial Values	Final Values			
α_o	40.	39.854 580 8	$9 \cdot 10^{-8}$	3.165	3.703
f_o	725.	724.680 299	$2 \cdot 10^{-6}$	63.13	76.64
p_o	$1.93 \cdot 10^5$	$1.903\ 796\ 24 \cdot 10^5$	$6 \cdot 10^{-5}$	$0.1200 \cdot 10^5$	$0.1206 \cdot 10^5$
a	0.63	0.634 854 890	1.10^{-9}	0.03392	0.04131

Number of Cycles: 8

W=16.496 783 247 4

Last Change of W = $2 \cdot 10^{-14}$

$\overline{k}^2 = 0.00768$

$m_o = 0.881\ 111\ 5$

Variance-Covariance Matrix
(Upper Half Only)

$$\begin{pmatrix} 10.02 & 133.0 & -1.249 \cdot 10^4 & 0.09830 \\ & 3986. & 3.340 \cdot 10^5 & 1.553 \\ & & 1.440 \cdot 10^8 & -51.08 \\ & & & 1.151 \cdot 10^{-3} \end{pmatrix}$$

Crude Variance-Covariance Matrix
(Upper Half Only)

$$\begin{pmatrix} 13.71 & 219.1 & -0.8234 \cdot 10^4 & 0.1433 \\ & 5873. & 4.117 \cdot 10^5 & 2.606 \\ & & 1.455 \cdot 10^8 & 35.22 \\ & & & 1.706 \cdot 10^{-3} \end{pmatrix}$$

TABLE X

DATA FOR FITTING GENERALIZED CASSINIAN CURVE

X	Y	X	Y
1.0	5.0	5.0	9.0
2.0	3.0	3.0	8.0
4.0	1.0	1.0	7.5
6.0	1.0	-1.0	9.0
7.0	3.0	-3.0	9.5
8.0	6.0	-4.5	8.0
7.5	8.0	-4.0	6.0
6.5	9.0	-2.0	5.0

Standard Errors of Observations

$$e_r = 0.02\ (x^2 + y^2)$$

$$e_\varphi = 0.08$$

TABLE XI

RESULTS OF FITTING PSEUDO - CASSINIAN CURVE TO CORRELATED OBSERVATIONS

	PARAMETERS		Last Changes	Standard Errors	Crude Standard Errors
	Initial Values	Final Values			
x_1	-2.	-3.246 408 5	$-4\cdot10^{-8}$	0.4386	0.4472
y_1	7.	7.606 215 9	$8\cdot10^{-9}$	0.1616	0.3261
x_2	5.	5.097 509 9	$2\cdot10^{-9}$	0.1929	0.2307
y_2	4.5	3.855 190 1	$7\cdot10^{-9}$	0.2832	0.3083
a	200.	437.692 47	$6\cdot10^{-6}$	48.76	99.06
b	0.25	0.376 844 61	$1\cdot10^{-9}$	0.1324	0.09642

Number of Cycles: 21

W = 3.469 719 340 38

Last Change of W = $7\cdot10^{-15}$

$\bar{k}^2 = 1.84\cdot10^{-3}$

m_o - 0.586 531 8

Variance-Covariance Matrix (Upper Half Only)

$$\begin{pmatrix} 0.1924 & -4.213\cdot10^{-2} & -5.487\cdot10^{-3} & -5.812\cdot10^{-2} & -0.1379 & -4.736\cdot10^{-2} \\ & 2.612\cdot10^{-2} & 2.153\cdot10^{-3} & 3.820\cdot10^{-2} & 0.6076 & 1.420\cdot10^{-2} \\ & & 3.720\cdot10^{-2} & -3.611\cdot10^{-3} & 8.487 & 6.952\cdot10^{-3} \\ & & & 8.024\cdot10^{-2} & -0.9958 & 1.613\cdot10^{-2} \\ & & & & 2.377\cdot10^{3} & 2.130 \\ & & & & & 1.753\cdot10^{-2} \end{pmatrix}$$

Crude Varinace-Covariance Matrix (Upper Half Only)

$$\begin{pmatrix} 0.2000 & -4.478\cdot10^{-2} & -6.525\cdot10^{-3} & -5.720\cdot10^{-2} & -30.94 & -0.389\cdot10^{-2} \\ & 10.63\cdot10^{-2} & 17.12\cdot10^{-3} & -2.015\cdot10^{-2} & 20.29 & 1.700\cdot10^{-2} \\ & & 5.324\cdot10^{-2} & -15.57\cdot10^{-3} & 14.88 & 10.89\cdot10^{-3} \\ & & & 9.503\cdot10^{-2} & 0.6343 & -1.535\cdot10^{-2} \\ & & & & 9.814\cdot10^{3} & 5.536 \\ & & & & & 0.929\cdot10^{-2} \end{pmatrix}$$

TABLE XII

RESULTS OF FITTING PSEUDO-CASSINIAN CURVE
TO OBSERVATIONS WITH UNIT WEIGHTS

	PARAMETERS		Last Changes	Standard Errors	Crude Standard Errors
	Initial Values	Final Values			
x_1	-2.	-2.887 709 0	$-4\cdot10^{-9}$	0.8572	0.3152
y_1	7.	6.983 339 1	$1\cdot10^{-8}$	0.1360	0.2468
x_2	5.	5.765 751 0	8.10^{-9}	0.2297	0.3351
y_2	4.5	4.505 450 5	$-5\cdot10^{-9}$	0.4386	0.3637
a	200.	414.933 17	$-1\cdot10^{-6}$	45.89	66.01
b	0.25	0.252 214 55	$-2\cdot10^{-9}$	0.1792	0.0580

Number of Cycles: 10

W = 2.674 613 584 39

Last Change of W = $4\cdot10^{-14}$

$\bar{k}^2 = 5.75\cdot10^{-14}$

m_o = 0.516 275 9

Variance-Covariance Matrix (Upper Half Only)

$$\begin{pmatrix} 0.7348 & -3.641\cdot10^{-2} & -7.694\cdot10^{-2} & 8.775\cdot10^{-2} & 1.195 & 1.462\cdot10^{-1} \\ & 1.850\cdot10^{-2} & -2.735\cdot10^{-3} & 4.034\cdot10^{-2} & -8.008\cdot10^{-1} & 5.392\cdot10^{-3} \\ & & 5.272\cdot10^{-2} & -3.376\cdot10^{-2} & 7.652 & 1.963\cdot10^{-2} \\ & & & 1.924\cdot10^{-1} & -3.930 & -2.075\cdot10^{-2} \\ & & & & 2.106\cdot10^{3} & 1.250 \\ & & & & & 3.210\cdot10^{-2} \end{pmatrix}$$

Crude Variance-Covariance Matrix (Upper Half Only)

$$\begin{pmatrix} 0.0993 & -0.261\cdot10^{-2} & 1.384\cdot10^{-2} & -1.671\cdot10^{-2} & -13.82 & -0.046\cdot10^{-1} \\ & 6.090\cdot10^{-2} & -0.428\cdot10^{-3} & -1.961\cdot10^{-2} & 20.22\cdot10^{-1} & 1.869\cdot10^{-3} \\ & & 5.528\cdot10^{-2} & -0.468\cdot10^{-2} & 7.525 & 0.453\cdot10^{-2} \\ & & & 1.323\cdot10^{-1} & 0.0012 & 0.553\cdot10^{-2} \\ & & & & 4.353\cdot10^{3} & 2.281 \\ & & & & & 0.336\cdot10^{-2} \end{pmatrix}$$

REFERENCES

1. Duane Brown, "A Matrix Treatment of the General Problem of Least Squares Considering Correlated Observations", BRL Report 937, 1955.

2. J. M. Tienstra, "Theory of the Adjustment of Normally Distributed Observations", N.V. Uitgeverij "Argus", Amsterdam, 1956.

3. Walter Grossmann, "Grundzüge der Ausgleichsrechnung", 2nd edition, Springer Verlag, Berlin, 1961.

4. W. Edwards Demming, "Statistical Adjustment of Data", John Wiley, New York, 1944.

5. Niels Arley & K. Rander Buch, "Introduction to the Theory of Probability and Statistics", John Wiley, New York, 1950.

6. John R. Wolberg, "Prediction Analysis". D. Van Nostrand, Princeton, 1967.

7. W. E. Wentworth, "Rigorous Least Squares Adjustment", Journal of Chemical Education, Vol. 42 (1965), pp. 96-103 and 162-167.

8. Robert E. Barieau & B.J. Dalton "Nonlinear Regression and the Principle of Least Squares", Bureau of Mines Rep. of Investigations Nr 6900, 1966.

9. Derek York, "Least Squares Fitting of a Straight Line", Canadian Journal of Physics, Vol. 44 (1966), pp 1079-1086.

10. J. H. Williamson, "Least-Squares Fitting of a Straight Line", Canadian Journal of Physics, Vol. 46 (1968), pp 1845-1847.

11. M. O'Neil, I.G. Sinclair & Francis J. Smith, "Polynomial Curve Fitting When Abscissas and Ordinates are Both Subject to Error", The Computer Journal, Vol. 12 (1969), pp. 52-56.

12. W. H. Southwell, "Fitting Experimental Data", Journal of Computational Physics, Vol. 4 (1969), pp. 465-474.

13. D. R. Powell & J. R. Macdonald, "A Rapidly Convergent Iterative Method for the Solution of the Generalized Nonlinear Least Squares Problem", The Computer Journal Vol. 15 (1972), pp 148-155.

14. J. Ross Macdonald, "Review of Some Experimental and Analytical Equations of State", Rev. of Modern Physics, Vol. 41 (1969), pp 316-349.

15. K. Pearson, "On Lines and Planes Closest Fit to Systems of Points in Space", Phil. Mag., Vol 2 (1901), pp. 559-572.

16. M.M. Ibiricu, "Stable and Unstable Regimes in Solid Propellant Combustion", XVII International Astronautical Congress, Madrid 1966, Vol. 3, pp. 183-189.

APPENDIX A

Computer Program COLSGN

The numerical solution of the least squares equations (22) and (23) of Section 2 can be implemented on a computer in different ways. In this Appendix we shall describe the implementation of the solution by the BRL Applied Mathematics Laboratory's code COLSGN (COrrelated Least Squares with General constraints), and give the rationale for the particular implementation chosen.

The computation in COLSGN are initiated by evaluating the following equations:

$$g_j = \frac{1}{A_j^T R_j A_j} , \qquad (j = 1,2,\ldots,r), \tag{A-1}$$

$$k_j = g_j(A_j^T C_j - F_j) , \quad (j = 1,2,\ldots,r), \tag{A-2}$$

$$\sum_{j=1}^{r} g_j B_j B_j^T \tau = \sum_{j=1}^{r} k_j B_j . \tag{A-3}$$

These equations correspond to Equations (20), (21) and (22), respectively, whereby the Equations (21) for the correlates are slightly modified. Generally the $C_j \equiv 0$ at the beginning, but it was found adventageous for some problems to provide an option to start the calculations with estimated non-zero values of the C_j.

The Equations (A-3) are solved for τ, the parameter vector T replaced by $T + \tau$ in all arguments and a new τ determined using Equations (A-1) (A-2) and (A-3). In other words, Equations (A-1), (A-2) and (A-3) are used for iteration of the parameters, keeping the residuals fixed.

After that iteration has come to a halt (iteration end conditions will be discussed later), the residuals are computed by a modified Equation (23), namely,

$$C_j + \varepsilon_j = k_j \; R_j \; A_j \quad (j = 1,2,\ldots,r) \tag{A-4}$$

The C_j are then replaced by $C_j + \varepsilon_j$ in all arguments and new $C_j + \varepsilon_j$ computed using the Equations (A-1), (A-2) and (A-4). Thus, during this phase of computations, the residuals are iterated, keeping thereby the parameters fixed. The iteration is continued until an end criterion is satisfied. This completes then one iteration cycle. The next cycle is started by a new computation of the parameters using the new values of C_j and Equations (A-1), (A-2) and (A-3).

Let the three iterations which are involved in the computing process be called "parameter iteration", "residual iteration" and "cycle iteration", respectively. The end conditions for these iterations are purely numerical in the COLSGN code and, therefore, arbitrary to a certain degree. They are established mainly by numerical experiments.

One obvious criterion for the cycle convergence is that the weighted sum of the residual squares, namely,

$$W = \sum_{j=1}^{r} C_j^T \; R_j^{-1} C_j \tag{A-5}$$

becomes stationary. Such a condition alone, however, is not suitable for numerical purposes, because in the vicinity of a minimum W is little influenced by small changes of the variables (C_j and T). As a second criterion, therefore, a requirement was introduced that the changes of the components of the parameter vector T between cycles become smaller than a threshold. That threshold was set arbitrarily at 10^{-7} of a crude estimate of the corresponding standard error. Particularly, COLSGN stops the cycle iteration, if all of the following conditions are satisfied:

$$|\Delta W| < 10^{-12} \cdot m_o \; W \tag{A-6}$$

$$|\Delta T_j| < 10^{-7} \cdot m_o \sqrt{Q_{jj}} \; , \; j = 1,\ldots,p \tag{A-7}$$

where ΔW and ΔT_j are the changes of W and T_j, respectively, by the last cycle, m_o^2 is the variance of weight one (see Equation (56), Section 3) and Q_{jj} is a diagonal element of the matrix inverse to the normal equation

matrix in Equation (A-3). As pointed out in Sections 3 and 4, $m_o\sqrt{Q_{jj}}$ are poor estimates for the standard errors of parameters. Nevertheless they are used by COLSGN in its end conditions because the Q_{jj} are readily available during the iterations and because an error even by an order of magnitude would not have serious influence on the results.

The conditions (A-7) are safer to apply than for instance, the requirements that the relative values $|\Delta T_j/T_j|$ be less than a threshold, because some of the parameters might approach zero at convergence. (Such convergence criteria are used by Powell & Macdonald[13]). In all test examples computed, the conditions (A-7) were more stringent that (A-6), the latter being satisfied several cycles before (A-7).

The iteration and criterion for the parameter iteration should merely insure that the nonlinear effects of the constraint functional are reduced by the iteration. This is the case, if the correction τ of T is "sufficiently" small. During the first cycles, on the other hand, the end condition for the parameter iteration should not be too stringent, because the parameter values computed during theses cycles are only preliminary approximations. They will be corrected anyway in each subsequent cycle after the computation of new residuals. Finally, the "smallness" of τ should be in some relation to the present accuracy of W.

With these considerations in mind, the end condition for the parameter iteration was formulated as follows:

$$\tau^T \cdot \sum_{j=1}^{r} k_j B_j < 10^{-a} \cdot m_o \cdot W_{parameters} \qquad \text{(A-8)}$$

with

$$W_{parameters} = \sum_{j=1}^{r} \frac{k_j^2}{g_j} \qquad \text{(A-9)}$$

and

$$a < \min\,(12,\ \text{Cycle Nr} + 2). \qquad \text{(A-10)}$$

The factor 10^{-a} in (A-8) takes care of an increase in the accuracy requirement as the cycle number increases. $W_{parameter}$ is according to Section 3 an approximation to W. (W cannot be computed by Equation (A-5) during the parameter iteration, because the C_j do not change during that iteration.) The left hand side of Equation (A-8) is roughly proportional to the change of $W_{parameter}$ due to the change τ of the parameter T. Hence, (A-8) requires that an estimate of the relative change of W due to τ should be less than a threshold.

The above mentioned relation between τ and $W_{parameter}$ can be derived as follows. Neglecting the changes of A_j (and, therefore, those of g_j) due to τ we obtain from (A-2) for the changes Δk_j of k_j,

$$\Delta k_j \approx -g_j \tau^T B_j. \tag{A-11}$$

Substituting (A-11) into (A-9) we obtain,

$$\Sigma \frac{(k_j + \Delta k_j)^2}{g_j} \approx \Sigma \frac{k_j^2}{g_j} - 2\tau^T \quad \Sigma k_j B_j + \Sigma g_j (\tau^T B_j)^2,$$

where the first term on the right hand side is twice the left hand side of the condition (A-8), with negative sign. Incidentally the expression $\tau^T \Sigma k_j B_j$ is, because of Equation (A-3), equal to $\tau^T N \tau$, where N is the positive definite normal equation matrix. Hence that expression can be considered as the square of a norm of τ. Therefore, the condition (A-8) can also be interpreted as a condition for a norm of τ.

As the cycle number increases the condition (A-8) becomes more stringent. On the other hand, the corrections become smaller and, therefore, the linearized constraint functional an ever better approximation. Usually after a few initial cycles the condition (A-8) is satisfied by the first parameter iteration step.

The initial guess T of the parameters can be quite different from the least squares solution t. Therefore, COLSGN carries out in the first cycle at least three parameter iterations, regardless whether

(A-8) is satisfied or not. Experiments show that more than three parameter iterations are necessary in any cycle only if the current approximation T is very bad. In those cases the system (A-1), (A-2) and (A-3) may even fail to converge at all. Convergence can then often be achieved, if the weights g_j are kept constant during the iterations. Therefore, the weights of the third parameter iteration are used for any futher iteration within each cycle. (Only (A-2) and (A-3) are used for the 4th, 5th etc. parameter iterations in each cycle.)

For the residual iteration similar considerations hold as for the parameter iteration. Because for this iteration, W can be computed by Equation (A-5), also its change due to the ε_j can be computed directly. The corresponding formula is,

$$\Delta W_\varepsilon = \sum_{j=1}^{r} \varepsilon_j^T \; R_j^{-1} (2C_j + \varepsilon_j). \tag{A-13}$$

The iteration end condition for the residual iteration is accordingly

$$|\Delta W_\varepsilon| < 10^{-a} \, m_o \cdot W \, , \tag{A-14}$$

where a is defined by (A-10). Usually after a few initial cycles this condition, too, is satisfied by the first iteration step.

Estimates of the variances and covariances of the parameters are computed after completed iteration by a straight forward evaluation of the formulas of Section 3.

DISTRIBUTION LIST

No. of Copies	Organization
12	Commander Defense Documentation Center ATTN: TIPCR Cameron Station Alexandria, Virginia 22314
1	Director Defense Advanced Research Projects Agency 1400 Wilson Boulevard Arlington, Virginia 22209
1	Director Weapons Systems Evaluation Group Washington, DC 20305
1	Commander U.S. Army Materiel Command ATTN: AMCDL, Dr. R. Dillaway 5001 Eisenhower Avenue Alexandria, Virginia 22304
1	Commander U.S. Army Materiel Command ATTN: AMCDL, N. Klein 5001 Eisenhower Avenue Alexandria, Virginia 22304
1	Commander U.S. Army Materiel Command ATTN: AMCDMA, MG J. R. Guthrie 5001 Eisenhower Avenue Alexandria, Virginia 22304
1	Commander U.S. Army Materiel Command ATTN: AMCRD, MG S. C. Meyer 5001 Eisenhower Avenue Alexandria, Virginia 22304
1	Commander U.S. Army Materiel Command ATTN: AMCRD Dr. J.V.R. Kaufman 5001 Eisenhower Avenue Alexandria, Virginia 22304
1	Commander U.S. Army Materiel Command ATTN: AMCRD-TE 5001 Eisenhower Avenue Alexandria, Virginia 22304
1	Commander U.S. Army Materiel Command ATTN: AMCRD-TP 5001 Eisenhower Avenue Alexandria, Virginia 22304
1	Commander U.S. Army Materiel Command ATTN: AMCRD-T, Dr. J. El-Bisi 5001 Eisenhower Avenue Alexandria, Virginia 22304
2	Commander U.S. Army Materiel Command ATTN: AMCRD-TC, Mr. Crellin Mr. Miller 5001 Eisenhower Avenue Alexandria, Virginia 22304
1	Commander U.S. Army Materiel Command ATTN: AMCRD-RP 5001 Eisenhower Avenue Alexandria, Virginia 22304
1	Commander U.S. Army Materiel Command ATTN: AMCRD-W 5001 Eisenhower Avenue Alexandria, Virginia 22304

DISTRIBUTION LIST

No. of Copies	Organization
1	Commander U.S. Army Materiel Command ATTN: AMCRD-WI, T. Cosgrove 5001 Eisenhower Avenue Alexandria, Virginia 22304
1	Commander U.S. Army Materiel Command ATTN: AMCPA, D. J. Shearin 5001 Eisenhower Avenue Alexandria, Virginia 22304
1	Commander U.S. Army Aviation Systems Command ATTN: AMSAV-E 12th and Spruce Streets St. Louis, Missouri 63166
1	Director U.S. Army Air Mobility Research and Development Laboratory Ames Research Center Moffett Field, California 94035
1	Commander U.S. Army Electronics Command ATTN: AMSEL-RD Fort Monmouth, New Jersey 07703
4	Commander U.S. Army Missile Command ATTN: AMSMI-R AMSMI-RBL AMSMI-XB (2 cys) Redstone Arsenal, Alabama 35809
1	Commander U.S. Army Tank Automotive Command ATTN: AMSTA-RHFL Warren, Michigan 48090
3	Commander U.S. Army Mobility Equipment Research & Development Center ATTN: Tech Docu Cen, Bldg. 315 AMSME-RZT Dwight Shaw, Bldg. 354 Fort Belvoir, Virginia 22060
3	Commander U.S. Army Armament Command ATTN: MAJ Bob B. Lukens CPT George B. Martin, Jr. Mr. J. Selman Rock Island, Illinois 61202
1	Commander U.S. Army Picatinny Arsenal ATTN: SMUPA-MI, Otto Wolf Dover, New Jersey 07801
1	Commander Deseret Test Center ATTN: E. G. Peterson Dugway Proving Ground Dugway, Utah 84022
1	Commander U.S. Army Rock Island Arsenal Rock Island, Illinois 61202
2	Commander U.S. Army Watervliet Arsenal ATTN: SWEWV-RD SWEWV-RDR, Mr. Benet Watervliet, New York 12189
1	Director U.S. Army Advanced Materiel Concepts Agency 2461 Eisenhower Avenue Alexandria, Virginia 22314

DISTRIBUTION LIST

No. of Copies	Organization
4	Commander U.S. Army Harry Diamond Laboratories ATTN: AMXDO-TD/002 AMXDO-EM, Mr. J. M. Clodfelter AMXDO-RC Arthur Hausner Norman A. Eisenberg Washington, DC 20438
1	Commander U.S. Army Materials and Mechanics Research Center ATTN: AMXMR-ATL Watertown, Massachusetts 02172
3	Commander U.S. Army Natick Laboratories ATTN: AMXRE, Dr. D. Sieling Mr. Gerald Hertweck E. W. Ross Natick, Massachusetts 01762
2	Commander USA Training & Doctrine Command ATTN: ATCD-CS-M-AR DCS-CD Fort Belvoir, Virginia 22060
1	Commander U.S. Army Combined Arms Development Activity Fort Leavenworth, Kansas 66027
1	Commander U.S. Army Combat Developments Experimentation Command Fort Ord, California 93941
3	HQDA (DARD-DDC, LTC Shelton; DARD-MSN-C; LTC R. B. Henry) Washington, DC 20310
2	HQDA (DARD-ART, MAJ Manderville; DARD-ARP-P, MAJ Burns) Washington, DC 20310
1	Director Engineering Handbook Officer Box EM, Duke Station Durham, North Carolina 27706
1	Chief, US Army Strategy and Tactica Analysis Group 8120 Woodmont Avenue Bethesda, Maryland 20014
1	Mathematics Research Center ATTN: C. de Boor Madison, Wisconsin 53706
1	Chief of Naval Operations ATTN: OP-03EG Department of the Navy Washington, DC 20350
5	Commander U.S. Naval Air Systems Command ATTN: AIR-604 (3 cys) AIR-06221 AIR-5321 Washington, DC 20360
3	Commander U.S. Naval Ordnance Systems Command ATTN: ORD-0632 ORD-035 ORD-5524 Washington, DC 20360
3	Commander U.S. Naval Weapons Center ATTN: Code 753 Code 12 Code 4034 China Lake, California 93555

DISTRIBUTION LIST

No. of Copies	Organization
1	Commander U.S. Naval Ordnance Laboratory Silver Spring, Maryland 20910
1	Commander U.S. Naval Weapons Laboratory Dahlgren, Virginia 22448
2	Commandant U.S. Marine Corps Washington, DC 20380
1	Director Development Center (MCDEC) ATTN: CK, Gnd Cbt Div Quantico, Virginia 22134
1	HQ USAF (AFCSAI) Washington, DC 20330
1	AFATL (DLR) Eglin AFB Florida 32542
1	AFATL (DLRD) Eglin AFB Florida 32542
1	AFATL (DLRV) Eglin AFB Florida 32542
1	AFATL (DLMJ, W. Dittrich) Eglin AFB Florida 32542

No. of Copies	Organization
1	ADTC (ADBRL-2) Eglin AFB Florida 32542
1	ASD (ASJ) Wright-Patterson AFB Ohio 45433
2	ARINC Research Corporation ATTN: Mr. D. E. Rowe Mr. L. B. Greene 2551 Riva Road Annapolis, Maryland 21401
1	United Aircraft Research Labs ATTN: G. Buckel East Hartford, Connecticut 06108
1	Worcester Polytechnic Institute Department of Mathematics ATTN: Paul Davis Worcester, Massachusetts 01609

Aberdeen Proving Ground

Ch, Tech Lib
Marine Corps Ln Ofc
Dir, MTD
ATTN: STEAP-MT-TI
Cmdr, USATECOM
ATTN: AMSTE-BC, Mr. Morrow
Cmdr, USALWL
ATTN: CRDLWL-6B
Cmdr, USAEA
ATTN: Dr. W. Sacco, SMUEA-BL-S
Dir, USAMSAA

Unclassified

Security Classification

DOCUMENT CONTROL DATA - R & D

(Security classification of title, body of abstract and indexing annotation must be entered when the overall report is classified)

1. ORIGINATING ACTIVITY *(Corporate author)*
USA Ballisitic Research Laboratories
Aberdeen Proving Ground, Maryland 21005

2a. REPORT SECURITY CLASSIFICATION
Unclassified

2b. GROUP

3. REPORT TITLE
Least Squares Adjustment with Finite Residuals for Non-Linear and Partially Correlated Data.

4. DESCRIPTIVE NOTES *(Type of report and inclusive dates)*

5. AUTHOR(S) *(First name, middle initial, last name)*
Aivars K. Celmins

6. REPORT DATE
July 1973

7a. TOTAL NO. OF PAGES
64

7b. NO. OF REFS
16

8a. CONTRACT OR GRANT NO.

b. PROJECT NO. 1T161102A14B

c.

d.

9a. ORIGINATOR'S REPORT NUMBER(S)
BRL REPORT NO. 1658

9b. OTHER REPORT NO(S) *(Any other numbers that may be assigned this report)*

10. DISTRIBUTION STATEMENT
Approved for public release; distribution unlimited.

11. SUPPLEMENTARY NOTES

12. SPONSORING MILITARY ACTIVITY
US Army Materiel Command
5001 Eisenhower Avenue
Alexandria, Virginia 22304

13. ABSTRACT

The subject of this paper is the adjustment ("curve fitting") of data by the least squares method. Based on general formulas, which are derived in this paper, a new algorithm for the least squares method is established. It can be applied to cases where more than one observable contain observational errors and where the postulated relation between the observables is non-linear. Also taken into account are accuracies of the observations and correlations between the components of each observation vector. New formulas are derived for the estimation of the variance-covariance matrix of the fitted parameters. It is shown that the conventionally used estimation formula is theoretically wrong except for very limited special cases. Numerical tests of the algorithm demonstrate its accuracy and exceptional convergence characteristics. They also show that the conventional estimate of the variance-covariance matrix of the parameters is a very bad approximation to the theoretically correct estimate derived in this paper.

DD FORM 1 NOV 65 1473 REPLACES DD FORM 1473, 1 JAN OBSOLETE FOR ARMY USE.

Unclassified

Security Classification

14. KEY WORDS	LINK A ROLE	LINK A WT	LINK B ROLE	LINK B WT	LINK C ROLE	LINK C WT
Non-linear least squares						
Correlated observations						
Parameter variances and covariances						
Standard errors of parameters						
Curve fitting						