8976-6002-RU-000.

8976-6002-RU-000

STATISTICS OF ORBIT DETERMINATION – WEIGHTED LEAST SQUARES

T. A. Magness, J. B. McGuire

15 January 1962

SPACE TECHNOLOGY LABORATORIES, INC.
A SUBSIDIARY OF THOMPSON RAMO WOOLDRIDGE INC.
ONE SPACE PARK • REDONDO BEACH, CALIFORNIA

Rgl 30122

8976-6002-RU-000

STATISTICS OF ORBIT DETERMINATION — WEIGHTED LEAST SQUARES

T. A. Magness
J. B. McGuire

Prepared for: Jet Propulsion Laboratory
California Institute of Technology

Contract No. 950045

Approved: E. H. Tompkins
E. H. Tompkins
Associate Manager
Systems Analysis Department

SPACE TECHNOLOGY LABORATORIES, INC.
A Subsidiary of Thompson Ramo Wooldridge, Inc.
One Space Park
Redondo Beach, California

CONTENTS

1. INTRODUCTION

This report deals with the determination of free flight spacecraft trajectories from noisy tracking data. Although the observations may be correlated, the method of statistical estimation is restricted to weighted least squares. An earlier report [4] has already examined the relative merits of least squares and other more sophisticated estimation techniques. The emphasis in this report will be on special topics which arise in connection with systems analyses of space missions. Specifically, these topics are:

(i) A Priori Data. The handling of a priori information on parameters to be estimated from observational data is discussed.

(ii) Separation of Parameters. The separation of parameters into classes, for example orbital and non-orbital parameters, is studied. This includes the problem of simultaneously estimating two classes of parameters, and that of determining the degrading effects of uncertainties in parameters which are not solved for.

(iii) Midcourse Maneuvers

Tracking through a midcourse correction is treated both from the point of view of real time operation and preflight error analysis.

(iv) Up-dated Least Squares

A simplified least squares orbit determination technique suitable for onboard use is considered. The essence of this technique is to continually up-date the latest estimate of position and velocity, modifying this estimate as new data is taken.

2. BACKGROUND

Before proceeding to the specialized topics with which this report is primarily concerned, it is well to review briefly the basic statistical theory of orbit determination. This theory may be formulated as follows: A set of observations (radar tracking data, optical observations, etc.) denoted by an n-vector z is given. From the laws of mechanics and from geometrical considerations, the true value μ of this vector of observations is expressible as a known function of a set of p parameters $(p < n)$ denoted by the p-vector γ : $\mu = \mu(\gamma)$. In the simplest case, γ denotes six (or less) position and velocity components at a specified epoch. In general, however, γ may include other non-orbital parameters such as physical constants, biases in observations, tracking station coordinates, etc. The non-linear regression equation is then

$$z = \mu(\gamma) + w , \tag{1}$$

where w denotes an n-vector of noise. Given z, the functional form of μ, and the statistical properties of w, the problem is to estimate γ. In practice, (1) is linearized by expanding μ about an initial guess g_o. Letting $\Delta\gamma = \gamma - g_o$ and $\Delta z = z - \mu(g_o)$ (the components of Δz are called "residuals"), we have

$$\Delta z = \theta \Delta\gamma + w , \tag{2}$$

where θ is an nxp matrix of known regression coefficients, which are simply partial derivatives of μ with respect to γ. The non-linear problem (1) is then solved by iteratively solving linear equations such as (2). Since we are primarily concerned here with the solution to the linearized equation (2), we shall in the future occasionally drop the Δ's in (2) for the sake of notational convenience:

$$z = \theta\gamma + w . \tag{3}$$

It should be obvious when this convention is followed.

A "weighted least squares" (WLS) estimate of γ in (3) takes the form

$$\boxed{g_{WLS} = (\theta' W \theta)^{-1} \theta' W z} \tag{4}$$

where W is an $n \times n$ diagonal matrix of non-negative weights specified in advance. This unbiased estimate has the property that of all linear unbiased estimates g, g_{WLS} minimizes $F(g)$, the sum of squares of residuals weighted according to the matrix W,

$$F(g) = (z - \theta g)' W (z - \theta g) . \tag{5}$$

The p x p covariance matrix of g_{WLS} is

$$\boxed{G_{WLS} = (\theta' W \theta)^{-1} \theta' W R W \theta (\theta' W \theta)^{-1}} \tag{6}$$

where $R = Eww'$ is the noise covariance matrix ("E" denotes the mathematical expectation operator).

When the noise is uncorrelated, it is optimal to choose $W = R^{-1}$, a diagonal n x n matrix. For then $G_{WLS} = (\theta' W \theta)^{-1}$ and is a minimum among the covariance matrices of all linear unbiased estimates of γ. When the noise is correlated and W is restricted to being diagonal, the criterion for selecting W is less obvious. Golub [2] has considered the problem of optimizing W, subject to W being diagonal, and has solved this problem for the case in which the correlation matrix is exponential. For the general case in which there are several mutually uncorrelated data types, the authors [4] have proposed a "minimax" weighting which may be described as follows: let the observations be grouped into k consecutive data types which are mutually uncorrelated but are internally correlated. Let M_i be the diagonal matrix of reciprocal standard deviations for data type i, let θ_i be the matrix of regression coefficients corresponding to data type i, and let λ_i be the maximum eigenvalue (or an upper bound on the maximum eigenvalue) of the noise correlation matrix for data type i. Then the conventional LS estimate of γ based on data type i, only, is

$$g_i = (\theta_i' M_i^2 \theta)^{-1} \theta_i' M_i^2 z \tag{7}$$

and has covariance matrix G_i satisfying (see [4])

$$G_i \leq \lambda_i (\theta_i' M_i^2 \theta)^{-1} \tag{8}$$

The minimax estimate g_{MM} of γ is a linear combination of the g_i in which each g_i is weighted according to $(\theta_i' M_i^2 \theta_i)/\lambda_i$. The explicit formula for g_{MM} is as a WLS estimate (4) with weighting matrix W_{MM},

$$W_{MM} = \left[\begin{array}{c|c|c} M_1^2/\lambda_i & & \\ \hline & M_2^2/\lambda_2 & \\ \hline & & \ddots \quad M_k^2/\lambda_k \end{array}\right] \qquad (9)$$

(ideally, W_{MM} is chosen as the maximum diagonal $n \times n$ matrix satisfying $W \leq R^{-1}$.) The covariance matrix G_{MM} of g_{MM} then has the property that

$$\boxed{G_{MM} \leq (\theta' W_{MM} \theta)^{-1} \leq \lambda_i (\theta_i' M_i^2 \theta_i), \quad i = 1, \ldots, k.} \qquad (10)$$

In other words, the covariance matrix of the minimax estimate, G_{MM}, is never greater than the upper bound on the covariance matrix of the conventional LS estimate based on any single data type alone. This property is achieved by giving less weight to data types which are highly correlated and therefore may contain less information than their RMS values would indicate. It should be noted that if data types are weighted according to M_i^2, as in conventional LS analysis, this property cannot be assured. In fact, the authors have encountered actual cases in practical orbit determination in which the inclusion of additional highly correlated data, weighted according to its reciprocal variance, actually degraded the accuracy of the determination over not using that data at all. When the noise is uncorrelated, then of course the minimax solution coincides with the conventional LS solution which, in turn, is the minimum variance solution.

When the minimax least squares estimate is to be employed, it is sometimes conceptually convenient to replace the correlated noise w in (3) with "equivalent-or-worse" uncorrelated noise v for which the covariance matrix is the reciprocal of the minimax weighting matrix W_{MM}. The new regression equation is then

$$z = \theta\gamma + v \; . \qquad (11)$$

The optimal estimate of γ in (11) is now the minimum variance estimate (for which the formula is identical with the minimax estimate of γ in (3)):

$$g_{MV} = (\theta' W_{MM} \theta)^{-1} \theta' W_{MM} z \quad . \tag{12}$$

The new covariance matrix of g_{MV} is then simply

$$G_{MV} = (\theta' W_{MM} \theta)^{-1} \tag{13}$$

which is an upper bound on the covariance matrix of the minimax estimate. Thus the uncorrelated noise v may be said to be "equivalent-to-or-worse-than" the correlated noise w. Since g_{MV} and G_{MV}, above, are forms which are commonly computed in numerical orbit determination programs, we are thus led to a theory which is not only conceptually simple, but readily adaptable to analysis using existing computer programs. The concept of replacing correlated noise with equivalent-or-worse uncorrelated noise is not new and is often used to advantage in an heuristic way; the particular formulation proposed here, however, has the advantage of being both convenient and rigorous.

In concluding, we note that the actual success of a space mission may be judged on the basis of one's ability to control a set of mission parameters, denoted by a vector α of dimension of $q \leq p$, rather than on one's ability to control γ explicitly. For example, on a lunar mission α could denote the two impact parameter coordinates at the moon. It may nevertheless be more practical to perform the statistical estimation with respect to γ. When this is the case, and when variations in α are related to variations in γ by a known $q \times p$ matrix ϕ, $\Delta \alpha = \phi \Delta \gamma$, then the covariance matrix A of uncertainty in α is related to the covariance matrix G of uncertainty in γ by the well-known formula

$$A = \phi G \phi' \quad . \tag{14}$$

3. A PRIORI DATA

In the standard non-linear regression equation,

$$z = \mu(\gamma) + w , \tag{1}$$

one is given z, the functional form of μ, the statistical properties of w, and some criterion for specifying the weighting matrix W. We assume here that, in addition, the initial estimate g_o of γ has associated with it an a priori "information" matrix S. S is formed out of the reciprocal of the a priori covariance sub-matrix of those components of g_o for which a priori variances are available,* together with zeros in rows and columns corresponding to the remaining components of g_o. For example, if only the first two components of g_o are specified with a priori covariance matrix Λ_o, S takes the form

$$S = \left[\begin{array}{c|c} \Lambda_o^{-1} & 0 \\ \hline 0 & 0 \end{array}\right]$$

There is no restriction that S be non-singular. The problem then is to incorporate g_o into the WLS procedure for estimating γ. This may be done as follows:

Let $\Delta\gamma = \gamma - g_o$. Then the a priori estimate of $\Delta\gamma$ is zero, having information matrix S. The WLS estimate of $\Delta\gamma$, based on observations only, is, of course, $\Delta g_{WLS} = (\theta' W \theta)^{-1} \theta' W \Delta z$, where $\Delta z = z - \mu(g_o)$, and the covariance matrix of Δg_{WLS} is $G_{WLS} = (\theta' W \theta)^{-1} \theta' W R W \theta (\theta' W \theta)^{-1}$.

---

*If the actual covariance sub-matrix is not known, an upper bound on the sub-matrix may be used instead.

Optimal Combination

The optimal way to combine these two estimates of $\Delta\gamma$ is to weight each according to its information matrix. This leads to the estimate

$$\boxed{g^* = g_o + (G_{WLS}^{-1} + S)^{-1} G_{WLS}^{-1} \Delta g_{WLS}} \tag{15}$$

having covariance matrix

$$\boxed{G^* = (G_{WLS}^{-1} + S)^{-1}} \tag{16}$$

One may now iterate for the solution of the non-linear regression equation, each time weighting the a priori estimate g_o according to its information matrix S.

The disadvantage of the optimal procedure described above is that it requires explicit use of the noise covariance matrix R, which can make the computations excessively difficult when R is not diagonal (i.e., when the noise is correlated). Moreover, R may not be known explicitly. For these reasons we propose the following procedure when the observational noise is correlated:

Minimax Combination

Let W be the minimax weighting matrix W_{MM}, described in Section 2. Then we know that the covariance matrix of the WLS estimate of $\Delta\gamma$ based on observations, only, satisfies $G_{WLS} \leq (\theta' W \theta)^{-1}$. Now weight the a priori estimate according to S, but weight the WLS estimate based on observations, only, according to $\theta' W \theta$ to generate the combined estimate g:

$$g = g_o + (\theta' W \theta + S)^{-1} \theta' W \theta \Delta g_{WLS} \,. \tag{17}$$

Since $\theta' W \theta \Delta g_{WLS} = \theta' W \Delta z$, (17) reduces to a simple operation on the observational data:

$$\boxed{g = g_o + (\theta' W \theta + S)^{-1} \theta' W \Delta z.} \tag{18}$$

The covariance matrix of g is

$$\boxed{G = (\theta' W \theta + S)^{-1} (\theta' W R W \theta + S) (\theta' W \theta + S)^{-1}.} \tag{19}$$

Since $R \leq W_{MM}^{-1}$ (see Section 2), an upper bound on G is

$$\boxed{G \leq (\theta' W \theta + S)^{-1}} \tag{20}$$

which, in turn is bounded by $(\theta' W \theta)^{-1}$. In case S possesses an inverse, then $G \leq S^{-1}$, also, so that the minimax property of g_{WLS} is preserved in g. For this reason it is proper to call (18) a minimax utilization of a priori data. Of course, if the observations are actually uncorrelated with diagonal covariance matrix equal to W^{-1}, then all of the above formulas apply except that now equality holds in (20).

It should be noted that if additional iterations are called for in order to solve the non-linear regression equation, some of the simplicity of (18) is lost. The principle to follow then is in each iteration to weight g_o according to S and to weight the estimate based on observations only according to $\theta' W \theta$, evaluated at that iteration. However, the inequality (20), based on θ evaluated for the "nominal" trajectory, will usually be acceptably accurate as a final answer when the true trajectory is not far from nominal.

4. SEPARATION OF PARAMETERS

Sometimes the components of the vector parameter γ will fall quite naturally into two classes. For example, the components of γ may be classified as orbital and non-orbital, or the orbital parameters may be subdivided into position and velocity components. One may then be interested in actually determining only those components belonging to one class, but be concerned about the effect of neglecting or simultaneously estimating components in the other class. This section deals with this aspect of orbit determination.

4.1 Estimating Only One Class of Parameters

Let us write (3) in the form

$$z = \mu(\gamma_1, \gamma_2) + w \tag{21}$$

where γ_1 and γ_2 denote a separation of the total vector parameter γ into two sub-vector parameters, and suppose that we plan to estimate only

γ_1, either because we are unaware of any uncertainty in γ_2, or else because it is not practical to estimate both γ_1 and γ_2. (Often, the available observations will not contain enough "information" about the parameter γ_2 to allow its estimation.) Then the linear expansion of (21) about the initial guesses g_{10} and g_{20} of γ_1 and γ_2 takes the form

$$\Delta z = \theta_1 (\gamma_1 - g_{10}) + \theta_2 (\gamma_2 - g_{20}) + w \quad . \tag{22}$$

When g_{10} has a priori information matrix S_1, then by analogy with (18) the new estimate of γ_1 is

$$\boxed{g_{11} = g_{10} + (\theta_1' W \theta_1 + S)^{-1} \theta_1' W \Delta z} \tag{23}$$

where **W** is the weighting matrix. It is instructive to substitute from (22) for Δz, then add and subtract $(\theta_1' W \theta_1 + S_1)^{-1} \times S_1 (\gamma_1 - g_{10})$ to the right hand side of (23), to obtain

$$g_{11} = \gamma_1 + (\theta_1' W \theta_1 + S_1)^{-1} \Big[-S_1 (\gamma_1 - g_{10}) + \theta_1' W \theta_2 (\gamma_2 - g_{20}) + \theta_1' W w \Big] \quad . \tag{24}$$

The advantage of this formulation is that it shows explicitly that g_{11} is an unbiased estimate of γ_1 in which there are three sources of error: (i) the a priori estimate g_{10} of γ_1, (ii) the incorrect value g_{20} which was assumed for γ_2, and (iii) the random noise w on the observations.

We shall assume that the a priori estimates g_{10} and g_{20} are uncorrelated.* Then the covariance matrix of g_{11} in (24) is

$$\boxed{G_{11} = (\theta_1' W \theta_1 + S_1)^{-1} \Big[\theta_1' W R W \theta_1 + S_1 + \theta_1' W \theta_2 \Lambda_2 \theta_2' W \theta_1 \Big] \times (\theta_1' W \theta_1 + S_1)^{-1}} \tag{25}$$

*If g_{10} and g_{20} are cross-correlated, then the formula (25) for G_{11} is merely made more complicated by the presence of cross-correlated terms.

where Λ_2 is the a priori covariance matrix of uncertainty in g_{20}. When the noise w is uncorrelated with $R = W^{-1}$, then (25) reduces to

$$\boxed{G_{11} = (\theta_1' W \theta_1 + S_1)^{-1} + (\theta_1' W \theta_1 + S_1)^{-1} \theta_1' W \theta_2 \Lambda_2 \theta_2' W \theta_1 \times (\theta_1' W \theta_1 + S_1)^{-1} .} \tag{26}$$

If the noise is correlated but $R \leq W^{-1}$ (as in the case with minimax estimation), then the right side of (26) is an upper bound on G_{11}.

Formula (26) may be interpreted as follows: $(\theta_1' W \theta_1 + S_1)^{-1}$ is the covariance matrix of uncertainty in the estimate g_{11} resulting from uncertainty in the initial estimate of γ_1 and from noise on the tracking observations, only. This is the matrix which is commonly computed in tracking programs. The additional term on the right in (26) is a non-negative, symmetric matrix of the same order as $(\theta_1' W \theta_1 + S_1)^{-1}$ which shows the additional uncertainty in g_{11} resulting from having used the incorrect value of γ_2 in the WLS determination of γ_1. Thus, the effect of uncertainties in the γ_2 parameters on the orbit determination process may be examined by comparing the relative contributions of the two terms in equation (26). In high precision orbit determination work, the second term can play an important role in establishing the confidence which is to be assigned to the estimate g_{11}.

To carry this analysis one step further, let us assume now that the (vector) mission parameter α which one is interested in (see Section 2) is related to the parameters γ_1 and γ_2 by

$$\Delta \alpha = \phi_1 \Delta \gamma_1 + \phi_2 \Delta \gamma_2$$

Then the final uncertainty in α can be written as

$$\Delta \alpha = \phi_1 (\theta_1' W \theta_1 + S_1)^{-1} \left[\theta_1' W \theta_2 (\gamma_2 - g_{20}) - S_1 (\gamma_1 - g_{10}) + \theta_1' W w \right] + \phi_2 (\gamma_2 - g_{20}) \tag{27}$$

It is useful to regard the coefficient of $(\gamma_2 - g_{20})$ in (27) as the "total derivative" of α with respect to γ_2, since it shows how uncertainty in γ_2 affects α both explicitly and implicitly:

$$\phi_1 \, (\theta_1' W \theta_1 + S_1)^{-1} \, \theta_1' W \theta_2 + \phi_2 = \frac{d\alpha}{d\gamma_2} = \tilde{\phi}_2 \tag{28}$$

The final formula for the covariance matrix of uncertainty in α is thus

$$\boxed{\begin{aligned} A = {} & \phi_1 \, (\theta_1' W \theta_1 + S_1)^{-1} \, (\theta_1' W R W \theta_1 + S_1) \, (\theta_1' W \theta_1 + S_1)^{-1} \, \phi_1' \\ & + \tilde{\phi}_2 \Lambda_2 \, \tilde{\phi}_2' \end{aligned}} \tag{29}$$

As before, this equation simplifies in the obvious way to give an upper bound on A when $R \leq W^{-1}$.

4.2 Estimating Two Classes of Parameters Simultaneously

The problem of estimating two classes of parameters simultaneously is, in principle, completely covered by Sections 2 and 3. The purpose of this section is to put the results into a form which can be readily compared with the results of Section 4.1.

We shall write (22) as

$$z = \left[\theta_1 \mid \theta_2 \right] \left[\begin{array}{c} \gamma_1 \\ \hline \gamma_2 \end{array} \right] + w \tag{30}$$

dropping the Δ's for convenience. The WLS estimate of $\left[\begin{array}{c} \gamma_1 \\ \hline \gamma_2 \end{array} \right]$ is then

$$\left[\begin{array}{c} g_1 \\ \hline g_2 \end{array} \right] = \left[\begin{array}{c|c} \theta_1' W \theta_1 + S_{11} & \theta_1' W \theta_2 + S_{12} \\ \hline \theta_2' W \theta_1 + S_{21} & \theta_2' W \theta_2 + S_2 \end{array} \right]^{-1} \left[\begin{array}{c} \theta_1' W z \\ \hline \theta_2' W z \end{array} \right] \tag{31}$$

where $\left[\begin{array}{c|c} S_{11} & S_{12} \\ \hline S_{21} & S_{22} \end{array} \right]$ is the a priori information matrix on parameters.

The total covariance matrix of this estimate in the general case when the noise is correlated is given in Section 3 and need not be repeated. We are interested here in some simple relations which arise when the noise is uncorrelated, and will therefore make the simplifying assumption that $R = Eww' = W^{-1}$. (This is actually no great restriction, since if the noise were correlated it could be replaced with "equivalent-ot-worse" uncorrelated noise as shown in Section 2.) We shall now introduce the notation

$$C = \left[\begin{array}{c|c} \theta_1' W \theta_1 + S_{11} & \theta_1' W \theta_2 + S_{12} \\ \hline \theta_2' W \theta_1 + S_{21} & \theta_2' W \theta_2 + S_{22} \end{array}\right] = \left[\begin{array}{c|c} C_{11} & C_{12} \\ \hline C_{21} & C_{22} \end{array}\right], \quad C^{-1} = \left[\begin{array}{c|c} C^{11} & C^{12} \\ \hline C^{21} & C^{22} \end{array}\right], \tag{32}$$

$$\text{i.e., } C^{-1} = \left[\begin{array}{c|c} \begin{array}{c}\text{covariance of est. of } \gamma_1 \\ \text{parameters}\end{array} & \text{cross-covariance} \\ \hline \text{cross-covariance} & \begin{array}{c}\text{covariance of est. of } \gamma_2 \\ \text{parameters}\end{array} \end{array}\right]$$

Then C^{-1} is the total covariance matrix of $\left[\begin{array}{c} g_1 \\ \hline g_2 \end{array}\right]$ in (31), and C^{11}, C^{22} and C^{12} ($= (C^{21})'$), are the covariance and cross-covariance matrices of g_1 and g_2, individually.

It is useful to be able to express the above matrices in terms of the more easily computed matrices C_{11}^{-1}, C_{22}^{-1} and C_{12} ($= C_{21}'$). This can be done as follows: If we let $b_1 = \theta_1' W z$ and $b_2 = \theta_2' W z$, then (31) can be written as

$$g_1 = C^{11} b_1 + C^{12} b_2 \tag{33}$$

$$g_2 = C^{21} b_1 + C^{22} b_2 \tag{34}$$

or as the inverse transformation

$$b_1 = C_{11} g_1 + C_{12} g_2 \tag{35}$$

$$b_2 = C_{21} g_1 + C_{22} g_2 \tag{36}$$

Now eliminate g_1 from (35) and substitute into (36):

$$b_2 = C_{21} (C_{11}^{-1} b_1 - C_{11}^{-1} C_{12} g_2) + C_{22} g_2 .$$

This equation may now be solved for g_2,

$$g_2 = (C_{22} - C_{21} C_{11}^{-1} C_{12})^{-1} (b_2 - C_{21} C_{11}^{-1} b_1)$$

and the coefficients of b_1 and b_2 compared with those in (34), leading to

$$\boxed{C^{22} = (C_{22} - C_{21} C_{11}^{-1} C_{12})^{-1}} \tag{37}$$

and

$$\boxed{C^{21} = (C_{22} - C_{21} C_{11}^{-1} C_{12})^{-1} C_{21} C_{11}^{-1} = (C^{12})'} \tag{38}$$

Similarly, one can show that

$$\boxed{C^{11} = (C_{11} - C_{12} C_{22}^{-1} C_{21})^{-1}} \tag{39}$$

We have thus succeeded in expressing the components of the inverse of the matrix C in terms of the more easily computed sub-matrices C_{11}^{-1}, C_{22}^{-1} and C_{12} (= C_{21}) without explicitly inverting the entire matrix C. This is of interest if, for example, we are primarily interested in orbital parameters (denoted by γ_1), and therefore interested in the covariance matrix of orbital elements (C^{11}), but must decide between estimating or not estimating non-orbital parameters.

Actually, the case which we have treated here is slightly more general than the case treated in Section 4.1 in that we have not required the a priori estimates of γ_1 and γ_2 to be uncorrelated and we have not required S_{22} to possess an inverse. If we now add these restrictions, we can rewrite equation (26), using our new notation:

$$\boxed{G_{11} = C_{11}^{-1} + C_{11}^{-1} C_{12} S_{22}^{-1} C_{21} C_{11}^{-1}} \tag{26}$$

We can now distinguish three levels of sophistication in orbit determination:

(i) Estimate γ_1 only, ignoring uncertainties in γ_2. It is common practice to estimate only orbital parameters and give no attention to uncertainties in physical constants, station location, etc. The covariance matrix which is ordinarily assigned to the resulting estimate of γ_1 is $C_{11}^{-1} = (\theta_1' W \theta_1 + S_1)^{-1}$.

(ii) Estimate γ_1 only, as above, but include the effects of uncertainty in γ_2 in computing the covariance matrix of the estimate of γ_1. This leads to the correct formula for G_{11} as given by (26).

(iii) Estimate both γ_1 and γ_2. Then C^{11}, which is the "upper left hand sub-matrix" of C^{-1} in (32), is the covariance matrix for the estimate of γ_1.

The following relations will exist among the above covariance matrices:

$$\boxed{C_{11}^{-1} \leq C^{11} \leq G_{11} = C_{11}^{-1} + \text{(Non-neg. def. matrix)}.}$$

4.3 Discussion

In a typical problem, γ_1 in (21) may represent the six orbital parameters of position and velocity at a specified epoch and the non-orbital components γ_2 may represent such parameters as physical constants, tracking station coordinates, and biases in observations.

We note first that whenever $\theta_1' W \theta_2 = 0$, then (24) shows that there will be no interaction between g_{20} and the WLS estimate of γ_1, and that using an incorrect value of γ_2 in the estimation of γ_1 does not effect the answer to first order. This is confirmed in Section 4.2, which shows that when $\theta_1' W \theta_2 + S_{12} = C_{12} = 0$, then $C^{11} \doteq (C_{11})^{-1}$, $C^{22} = (C_{22})^{-1}$ and $C^{12} = 0$, so that there is no correlation between estimates of the two classes of parameters.

In order to evaluate the effects of either neglecting or solving for γ_2 in a general case, it is necessary to know the elements of the matrix θ_2. Evaluating these coefficients involves varying degrees of complexity, depending on the exact nature of γ_2:

(i) Biases. Partial derivatives of μ with respect to biases in observations are the simplest to evaluate, since in that case the partial derivative is either one or zero, depending on whether the bias is or is not associated with the particular observation in question.

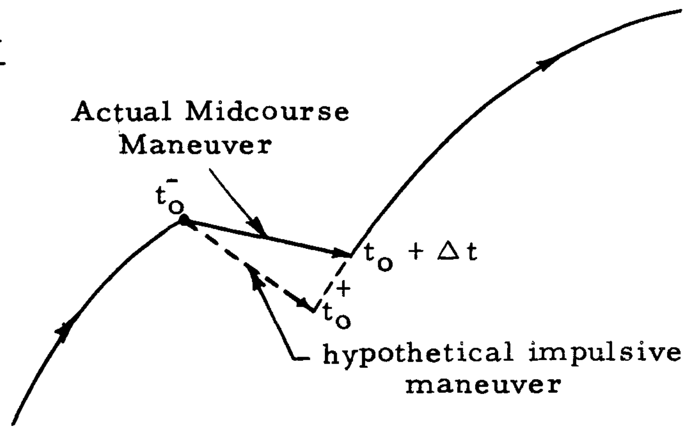
(ii) Station Coordinates. Partial derivatives with respect to tracking station coordinates are merely geometric transformations which do not involve the equations of motion of the spacecraft.

(iii) Physical Constants. Partial derivatives with respect to physical constants such as the mass of the earth or the astronomical unit are more difficult to evaluate, since they involve the actual equations of motion. These partial derivatives are usually evaluated through variational equations or use of analytic derivative formulas. The evaluation is normally carried out in such a way as to keep constant the various angles and angular rates associated with the solar system, since planetary angles and angular rates are generally known with sufficient accuracy that they may be considered as known, compared with other measurements. On the other hand, when the velocity of light c enters into an orbit determination problem, it usually enters merely as a scaling factor on range or range rate data, so that partial derivatives of μ with respect to c do not involve the equations of motion.

The analytic tracking accuracy prediction program (TAPP)*, which is under development at STL as a part of the Space Systems Analysis Study Contract, will generate partial derivatives of spacecraft observational data with respect to all of the above non-orbital parameter. Depending on how these partial derivatives are subsequently processed in the program, one will be able to simulate the fitting on non-orbital as well as orbital parameters, or else estimate statistically the degrading effect of using incorrect values for non-orbital parameters.

5. MIDCOURSE MANEUVERS

5.1 General Theory



The above diagram represents schematically a spacecraft trajectory in which there is a midcourse maneuver, i.e., a short powered flight, which

*This program makes use of a completely analytic (i.e., non-integrating) formulation and is designed for the pre-flight tracking and guidance analysis of space missions, rather than real-time operation.

begins at time t_o and lasts until $t_o + \Delta t$. We are concerned here with reconstructing this trajectory from tracking data taken during the free-flight portions of the trajectory (prior to t_o and after $t_o + \Delta t$) and from a priori information on the maneuver itself. Since the handling of any tracking data taken during powered flight is beyond the scope of this report, we shall assume that none is taken. For the same reason, we shall extrapolate the second portion of free flight back to time t_o and conceptually replace the actual maneuver with an equivalent "impulsive" correction in both position and velocity, which occurs instantaneously at t_o - see dashed curves. As a further simplification we shall assume that all tracking observations are uncorrelated. This is no great restriction, since if the tracking noise were correlated, it could be conceptually replaced with "equivalent-or-worse" uncorrelated noise (see Section 2).

As a matter of notation, γ_1 will denote true position and velocity at t_o - (just before the maneuver) and the subscript 1 will denote observations, etc., before the maneuver, $z_1 = \theta_1 \gamma_1 + w_1$, where $E w_1 w_1' = W_1^{-1}$, diagonal. Similarly, γ_2 will denote true position and velocity at t_o + (just after the equivalent impulsive maneuver) and the subscript 2 will denote (free-flight) observations, etc., after the maneuver, $z_2 = \theta_2 \gamma_2 + w_2$, where $E w_2 w_2' = W_2^{-1}$, diagonal, and $E w_1 w_2' \equiv 0$. Next let $\Lambda(a)$ be the 6 x 6 covariance matrix of execution errors, with "a" the commanded value of the maneuver. We shall discuss the origin of $\Lambda(a)$ later, merely noting here that it is a function of a.

As a rule, the statistical problem is to estimate γ_2, and this is the problem we shall consider. Assuming that execution errors are independent of tracking errors, we have two independent determinations of γ_2:

(i) From pre-midcourse tracking plus any a priori information, the estimate of γ_1 has covariance matrix $G_1 = (\theta_1' W_1 \theta_1 + S_1)^{-1}$. Since the maneuver has covariance matrix $\Lambda(a)$, the covariance matrix of uncertainty just after the hypothetical maneuver is $(\theta_1' W_1 \theta_1 + S_1)^{-1} + \Lambda(a)$. (Note: in practice, G_1 may actually be obtained by estimating orbital elements at a different epoch and then up-dating to epoch t_o. This in no way affects the results.)

(ii) From post-midcourse tracking, only, the estimate of γ_2 has covariance matrix $(\theta_2' W_2 \theta_2)^{-1}$.

Therefore the net covariance matrix of uncertainty of an estimate of γ_2 which is a linear combination of the above estimates (with each estimate being weighted inversely as its covariance matrix) is

$$\boxed{G_2(a) = (\theta_2' W_2 \theta_2 + ((\theta_1' W_1 \theta_1' + S_1)^{-1} + \Lambda(a))^{-1})^{-1}} \tag{40}$$

If the orbit determination is performed in real time or later, then a is a known vector, viz., the (hypothetical) commanded correction, and (40) can be evaluated numerically. If the analysis is carried out prior to the actual flight in order to perform an error analysis on the mission, then $G_2(a)$ must be averaged over the random variable a. When the pre-flight error analysis is done by Monte Carlo technique, this averaging can be done quite conveniently.* On the other hand, if the analysis is purely analytic, then it is more convenient to approximate this averaging by replacing $\Lambda(a)$ in (40) with its average value $\bar{\Lambda}$,

$$\bar{\Lambda} = \int df(a)\, \Lambda(a)$$

where $f(a)$ is the distribution function of a. The distribution $f(a)$ comes from a priori knowledge of how close to nominal the trajectory is likely to be, together with the guidance logic of which orbital parameters the midcourse maneuver is designed to correct. Although the philosophy of replacing $\Lambda(a)$ in (40) with its mean value is not strictly correct, this is not a critical point and it is highly questionable that it would be worth the effort to perform analytically the exact averaging. Therefore we may conclude that for a pre-flight analysis, the net covariance matrix of uncertainty in γ_2 is given by the approximate formula

$$\boxed{G_2 = (\theta_2' W_2 \theta_2 + ((\theta_1' W_1 \theta_1 + S_1)^{-1} + \bar{\Lambda})^{-1})^{-1}.} \tag{41}$$

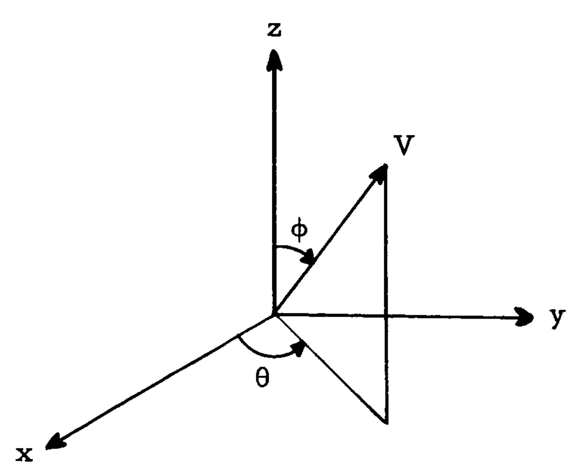
*The Monte Carlo technique is a mission analysis tool which encompasses injection, multi-midcourse, and terminal guidance analysis, as well as orbit determination. It will not be discussed here, since it is covered in other reports - e.g. [5].

It should be noted that the distribution of the estimate of γ_2, over the ensemble of all possible trajectories and all possible tracking data, is non-Gaussian and hence not completely characterized by the second moment matrix G_2. However, this does not mean that G_2 is not a meaningful matrix to generate. Furthermore, the degree of non-Gaussianness of this distribution will depend upon the extent to which $\bar{\Lambda}_2$ dominates G_2 in (41), and may be very slight.

5.2 The Matrices $\Lambda(a)$ and $\bar{\Lambda}$

The chief item which makes the analysis of orbit determination complicated when there is a midcourse maneuver is that the midcourse execution error is functionally dependent upon the maneuver which is commanded, which, for purposes of pre-flight analysis, is a random vector. This feature makes the real time (or post flight) analysis different from the pre-flight analysis, as we have just seen.

For the purpose of illustrating execution errors more concretely, we shall examine in detail a simplified model of a midcourse maneuver which, although not the most general, is typical of maneuvers occurring on many space missions. This model is characterized, first, by the fact that the actual correction may be considered impulsive (and therefore in velocity, only) and thus is identical with the hypothetical correction. The actual execution errors are most easily expressed in terms of a spherical coordinate system generated by the commanded correction vector V, shown below. The expressions for velocity errors in terms of basic error



sources are

$$\begin{aligned} \delta V &= \epsilon_1 + \epsilon_2 \left| V \right| \\ \delta V_\theta &= \epsilon_3 + \epsilon_4 \left| V \right| \\ \delta V_\phi &= \epsilon_5 + \epsilon_6 \left| V \right| \end{aligned} \tag{42}$$

where $\left| V \right|$ is the magnitude of V and where

ϵ_1 is a speed error due to engine shutdown,

ϵ_2 is a proportional speed error due to accelerometer,

ϵ_3 and ϵ_5 are lateral velocity errors due to control system,

ϵ_4 and ϵ_6 are "pointing" errors due to angular misalignment.

We assume that the ϵ_i are mutually uncorrelated with zero means and $E\,\epsilon_1^2 = k_1$, $E\,\epsilon_2^2 = k_2$, $E\,\epsilon_3^2 = E\,\epsilon_5^2 = k_3$, $E\,\epsilon_4^2 = E\,\epsilon_6^2 = k_4$. It is convenient to introduce a rotation U relating the above variations in spherical coordinates to variations in the rectangular reference coordinate system x, y, z:

$$\begin{bmatrix} \delta V_x \\ \delta V_y \\ \delta V_z \end{bmatrix} = U \begin{bmatrix} \delta V \\ \delta V_\theta \\ \delta V_\phi \end{bmatrix}$$

where

$$U = \begin{bmatrix} \sin\phi\cos\theta & -\sin\theta & \cos\phi\cos\theta \\ \sin\phi\sin\theta & \cos\theta & \cos\phi\sin\theta \\ \cos\phi & 0 & -\sin\phi \end{bmatrix}. \tag{43}$$

Thus the final form of the 3 x 3 execution velocity error covariance matrix, $\Sigma(V)$, is

$$\Sigma(V) = U \begin{bmatrix} k_1 & & \\ & k_3 & \\ & & k_3 \end{bmatrix} U' + V^2 U \begin{bmatrix} k_2 & & \\ & k_4 & \\ & & k_4 \end{bmatrix} U' . \tag{44}$$

The 6 x 6 matrix $\Lambda(a)$ which occurs in (40) is simply

$$\Lambda(V) = \left[\begin{array}{c|c} 0 & 0 \\ \hline 0 & \Sigma(V) \end{array}\right] \tag{45}$$

where the 6-vector a is in this example replaced by the 3-vector V.

When the 3-vector V is specified numerically, $\Sigma(V)$ and $\Lambda(V)$ can be evaluated numerically. On the other hand when V is specified by its probability distribution and it is required to compute $\overline{\Lambda}$, the calculation is not so simple. Lass and Solloway[3] and Gates[1] have proposed integration techniques for evaluating Λ which appear promising. For the present, the following special cases and approximations to $\overline{\Sigma}$ can be useful in computing $\overline{\Lambda}$:

(i) Degenerate Case. When the a priori distribution of V is actually one-dimensional along a direction characterized by θ_o and ϕ_o, then $U = U_o$, evaluated at θ_o and ϕ_o, and the mean value of $\Sigma(V)$ is

$$\overline{\Sigma} = U_o \begin{bmatrix} k_1 & & \\ & k_3 & \\ & & k_3 \end{bmatrix} U_o' + \overline{V^2}\, U_o \begin{bmatrix} k_2 & & \\ & k_4 & \\ & & k_4 \end{bmatrix} U_o' \tag{46}$$

where $\overline{V^2} = E\,|V|^2$. It frequently happens in practice that this is very nearly the case for the first midcourse correction after injection. Thus (46) may often be used as an approximate formula for execution errors during the first midcourse correction, with θ_o and ϕ_o denoting the direction of the maximum eigenvector of the a priori covariance matrix of V, and $\overline{V^2}$ as the trace of the a priori covariance matrix of V.

(ii) Symmetric Case. When the a priori distribution of V is spherically symmetric, then $|V|$, θ and ϕ are independent with θ uniformly distributed from 0 to 2π and ϕ distributed with frequency $(1/2) \sin \phi$ between 0 and π. Thus, averages with respect to V, θ and ϕ may be performed independently in (44), leading to

$$\boxed{\overline{\Sigma} = 1/3 \left[(k_1 + 2k_3) + \overline{V^2}\,(k_2 + 2k_4)\right] I_3} \tag{47}$$

In practice, the distribution of V for the second or third mid-course correction is likely to be very nearly symmetric. Thus (47) can often be used as an approximate formula for execution errors going with these later midcourse corrections.

(iii) Upper Bound on $\overline{\Sigma}$. Let $\lambda = \max(k_1, k_3)$ and $\mu = \max(k_2, k_4)$. Then it follows that

$$\boxed{\Sigma \leq (\lambda + \sqrt{2}\,\mu)\, I_3} \tag{48}$$

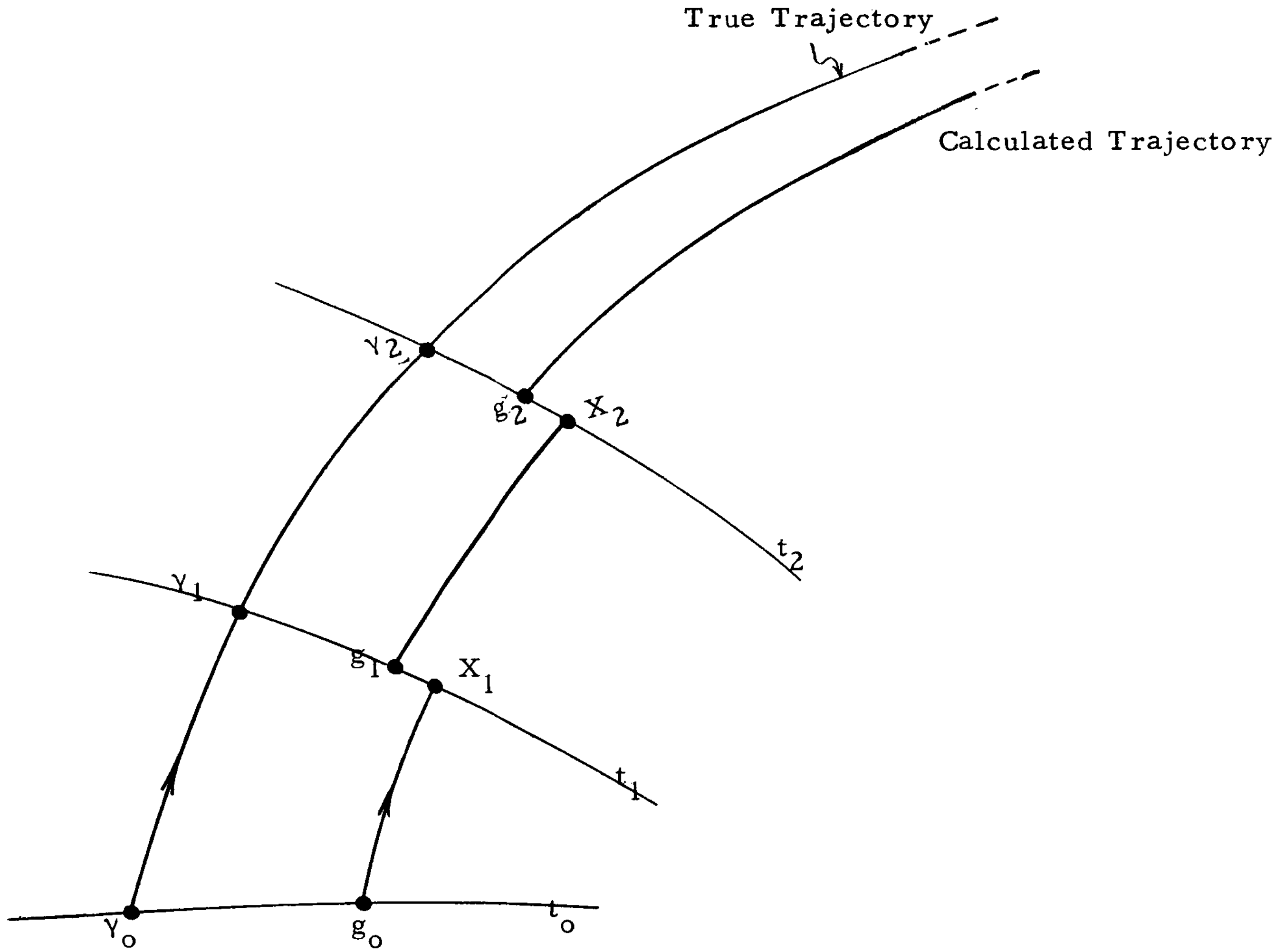
This easily computed upper bound on $\overline{\Sigma}$ can always be used as a conservative substitute for $\overline{\Sigma}$. If it happens that $k_1 = k_3$ and $k_2 = k_4$, then, of course, equality is attained in (48) and the formula is exact.

6. UPDATED LEAST SQUARES

This section describes an orbit determination technique which, because it does not require the storage of large quantities of observational data, is especially adapted to real time operation by an on-board computer. The essential features of this method have been proposed by Smith and Schmidt [6] who, because of the analogy between the estimation of orbits and the prediction of a time series by linear filtering, refer to this scheme as an "optimal filter" method. From our point of view, however, it is more natural to regard it as a least squares estimation procedure in which estimates of orbital parameters are continually updated and modified as new data arrives.

Consider the following estimation problem: an initial unbiased estimate g_o of spacecraft position and velocity at time t_o, together with an a priori 6 x 6 covariance matrix G_o of uncertainty in g_o, is provided. At each observation time t_k, $k = 1, 2, \ldots$, a q-vector of unbiased observations z_k is taken. (The dimension q may be different for different observation times.) We assume for the present that all observations are uncorrelated, and that the observations taken at time t_k are characterized by a known (diagonal) covariance matrix R_k, $k = 1, 2, \ldots$. Then at each observation time t_k, it is required to combine the old estimate of the orbit with the new data to form a new "best" estimate of position and velocity at t_k, and to determine the covariance matrix G_k of g_k. This concept is illustrated in the accompanying diagram, in which γ_o, γ_1, γ_2, ... denote the true position and velocity vector at t_o, t_1, t_2, ...

In the solution to this problem, it is sufficient to describe the calculations performed just after the k^{th} set of observations is taken. Let



Updated least squares orbit determination. γ_o, γ_1, ... denote the true position - velocity vectors at times t_o, t_1, ... g_k is the "best" estimate at time t_k, while X_k is g_{k-1} updated to time t_k.

X_k denote the result of integrating the equations of motion from t_{k-1} to t_k, using g_{k-1} as initial conditions. Then X_k serves as an initial estimate of γ_k for which the "a priori" covariance matrix, Λ_k, is G_{k-1} updated to time t_k:

$$\Lambda_k = \phi(k, k-1)\, G_{k-1}\, \phi'(k, k-1) \tag{49}$$

where $\phi(k, k-1)$ is a known 6 x 6 transition matrix satisfying

$$\Delta\gamma_k = \phi(k, k-1)\, \Delta\gamma_{k-1} \quad . \tag{50}$$

The observations z_k satisfy the non-linear regression equation

$$z_k = \mu_k(\gamma_k) + w_k \tag{51}$$

where μ_k is a known function of the orbital parameters γ_k, and w_k is noise for which $E w_k w_k' = R_k$. We may now proceed exactly as in Section 3. The linearized form of (51) is

$$\Delta z_k = \theta_k \Delta\gamma_k + w_k \tag{52}$$

where $\Delta\gamma_k = \gamma_k - X_k$, $\Delta z_k = z_k - \mu(X_k)$, and $\theta_k = \left(\frac{\partial\mu_k}{\partial\gamma_k}\right)$ is a q x p matrix of known coefficients. Setting $W_k = R_k^{-1}$, the new estimate of γ_k is

$$\boxed{g_k = X_k + G_k \theta_k' W_k \Delta z_k} \tag{53}$$

where G_k is the covariance matrix of g_k,

$$\boxed{G_k = (\theta_k' W_k \theta_k + \Lambda_k^{-1})^{-1} .} \tag{54}$$

We have described above the basic orbit determination technique. This technique can be generalized and/or modified to fit different situations. Some such modifications are described below.

(i) Matrix Identity. An equivalent formulation of (54) is as follows:

$$\boxed{G_k = \Lambda_k - \Lambda_k \theta_k' (W_k^{-1} + \theta_k \Lambda_k \theta_k')^{-1} \theta_k \Lambda_k .} \tag{55}$$

When $q < 6$, this formula has the computational advantage that the matrix which must be inverted is of order $q \times q$, rather than 6×6 as in (54). This new formula follows from (54) as a result of the matrix identity below:

Theorem. Let Λ and W be positive definite symmetric matrices of orders $p \times p$ and $q \times q$, respectively. Let θ be any $q \times p$ matrix. Then

$$(\Lambda^{-1} + \theta' W \theta)^{-1} = \Lambda - \Lambda \theta' (W^{-1} + \theta \Lambda \theta')^{-1} \theta \Lambda . \quad (56)$$

Proof: Consider first the special case $W = I_q$. By solving the matrix equation

$$(\Lambda^{-1} + \theta' \theta)(\Lambda - X) = I_p \quad (57)$$

for the unknown $p \times p$ matrix X, we obtain

$$\begin{aligned} X &= (\Lambda^{-1} + \theta' \theta)^{-1} \theta' \theta \Lambda \\ &= \Lambda (I_p + \theta' \theta \Lambda)^{-1} \theta' \theta \Lambda . \end{aligned} \quad (58)$$

Next we note that

$$(I_p + \theta' \theta \Lambda) \theta' (I_q + \theta \Lambda \theta')^{-1} = \theta' .$$

Substituting this expression for θ' in (58) leads to

$$\begin{aligned} X &= \Lambda (I_p + \theta' \theta \Lambda)^{-1} (I_p + \theta' \theta \Lambda) \theta' (I_q + \theta \Lambda \theta')^{-1} \theta \Lambda \\ &= \theta' (I_q + \theta \Lambda \theta')^{-1} \theta \Lambda \end{aligned} \quad (59)$$

which proves the theorem in this special case. The general case may now be proved by substituting $\theta = W^{-1/2} \tilde{\theta}$ into (56), which reduces the general case to the special case just examined, Q. E. D.

(ii) Correlated Observations. Suppose that the q observations taken at time t_k are correlated with non-diagonal covariance matrix R_k, but that observations taken at different times are uncorrelated. Then it is optimal to use the non-diagonal weighting matrix $W_k = R_k^{-1}$ in (53). Equation (54) for the covariance matrix of the estimate is still valid, using $W_k = R_k^{-1}$. If it is not convenient to use a non-diagonal weighting matrix, or if R_k is not known explicitly, then a "minimax" diagonal weighting matrix may be used. When this is done, the right hand side of (54) becomes an upper bound on the covariance matrix of the estimate.

If observations are correlated in time as well as instantaneously, then diagonal weighting matrices may still be used, but they should be scaled down according to the minimax principle enunciated in Section 2 to insure that highly correlated data types are not overly weighted.

(iii) Midcourse Maneuvers. The (real time) handling of midcourse maneuvers can be incorporated quite easily into the updated least squares routine. For example, if a maneuver occurs just prior to time t_k, then the commanded correction should be added to the estimate X_k, and the covariance matrix of execution errors should be added to Λ_k, to form a new a priori estimate and covariance matrix at t_k.

REFERENCES

[1] C. R. Gates, "An Approximation to the Second Moment of Maneuver Errors," Jet Propulsion Laboratory Technical Memo No. 312-155, December 26, 1961.

[2] G. H. Golub, "Comparison of the Variance of Minimum Variance and Weighted Least Squares Regression Coefficients," Space Technology Laboratories Document 9850.2-56, NN-171, 6 November 1961.

[3] H. Lass and C. B. Solloway, "On an Integral Arising in the Theory of the Midcourse Maneuver," Jet Propulsion Laboratories Technical Memo No. 312-150, December 5, 1961.

[4] T. A. Magness and J. B. McGuire, "Statistics of Orbit Determination - Correlated Observations," Space Technology Laboratories Document No. 8976-6001-RU-000, 15 December 1961.

[5] T. A. Magness, P. A. Penzo, P. Steiner, W. H. Pace, "Trajectory and Guidance Considerations for Two Lunar Return Missions Employing Radio Command Midcourse Guidance," Space Technology Laboratories Document 8976-0007-RU-000, 29 September 1961.

[6] G. L. Smith and S. F. Schmidt, "The Application of Statistical Filter Theory to Optimal Trajectory Determination On-Board a Circumlunar Vehicle," paper presented at the American Astronautical Society Meeting, San Francisco, 3 August 1961.