FEDERAL SYSTEMS CENTER
Houston Operations

INTRODUCTION TO TRAJECTORY ESTIMATION FOR RTCC PROGRAMMERS

N70-17416 (ACCESSION NUMBER)	(THRU)
139 (PAGES)	1 (CODE)
CR-102079 (NASA CR OR TMX OR AD NUMBER)	30 (CATEGORY)

FACILITY FORM 602

IBM

H69-0009-R

NASA CR 102079

INTRODUCTION TO TRAJECTORY ESTIMATION FOR RTCC PROGRAMMERS

by
Robert G. Rich
Department of Mathematical Analysis

Approved by

H. L. Norman

Herbert L. Norman
Manager, Department of Mathematical Analysis

Submitted to

National Aeronautics and Space Administration
Manned Spacecraft Center
Houston, Texas 77058

Contract No. NAS 9-996

Federal Systems Division
International Business Machines Corporation
1322 Space Park Drive
Houston, Texas 77058

PREFACE

This paper results from a voluntary evening course in trajectory estimation at the IBM Real Time Computer Complex, Manned Spaceflight Center (RTCC, MSC). It is written for programmers and navigators assigned to implement the navigation system, but who may arrive without previous knowledge of the subject. These people need to understand the applied system as soon as possible without necessarily becoming experts in all the individual disciplines. The attempt, therefore, is to include all necessary background material and provide compact, simple instruction on how to formulate the trajectory estimation problem for solution by a digital computer. This brief treatment certainly is not a substitute for formal study of trajectory estimation from texts in estimation theory and astrodynamics.

A sufficient background for understanding the presented material is a B. S. in mathematics, science, or engineering, including courses in differential equations, matrix algebra, and vector analysis. Some introduction to celestial mechanics and probability theory is helpful but not necessary.

The approach is first to review some useful facts about matrices and vectors and formulate partial derivatives, first-order Taylor series, Newton's method of successive approximations, and quadratic forms all in matrix notation. Then the estimation equations are derived from fundamentals without relying on any previous background in probability. The derivation is simplified by assuming that the dynamic model of the spacecraft trajectory is perfect. Later on, since model errors are inevitable, methods are suggested for empirically tuning the system to improve its performance.

Attention is focused on the derivation of the estimation equations; and many associated problems of a complete, implemented system are not included. For example, the manual does not explain numerical methods for integrating the equations of motion or calculating the state transition matrix. Other problems such as editing observations, calculating refraction and local vertical, and programming for displays are not mentioned.

Most of the theory is contained in the first fifteen sections. Beyond that is a collection of applications and ideas that may be interesting (or even useful).

I feel that I have only partially accomplished my purpose in writing this manual. Hopefully, a future revision would have increased scope, clarity, and simplicity. There are bound to be mistakes, and I would be grateful to anyone who sends in corrections.

I would like to acknowledge the contributions to this document made by Herbert L. Norman. He reviewed the entire text and suggested countless corrections, deletions, improvements, and additions. Although we were concerned mainly with the Apollo processor, he also contributed items of interest from his association with the Vanguard, Mercury, and Gemini programs.

TABLE OF CONTENTS

TABLE OF CONTENTS (Continued)

INTRODUCTION TO TRAJECTORY ESTIMATION FOR RTCC PROGRAMMERS

1. INTRODUCTION

The navigational problem considered by this paper is to determine where the spacecraft is and where it is going. If a navigator had exact knowledge of initial conditions and acting forces and a perfect solution to the equations of motion, trajectory estimation would not be needed. Unfortunately this is not the case. Measuring techniques used to determine initial conditions suffer from hardware and environmental limitations. External forces due to gravity, drag, thrusting, and venting are not known precisely. And integration techniques are such that predictions tend to diverge from the truth after a time, due to truncation and round-off errors and errors in the known forces. In view of these limitations a navigator must have some statistical means of resolving measurements into a best estimate of initial conditions, and he must do this at regular intervals to re-estimate current conditions. This is just a fancy way of describing any navigator's traditional task of using measurements to determine a fix and velocity vector.

Our problem, then, is to formulate a mathematical method of processing radar and optical measurements to estimate the position and velocity of a spacecraft. The spacecraft may be in either free flight (power off) or a powered maneuver, as long as the equations of motion are known. For example, if the spacecraft is in free flight and tracked in an earth-centered inertial frame, the equation of motion is

$$1.1 \qquad \ddot{\bar{r}} = \frac{-\mu \bar{r}}{|r|^3} + g(\bar{r}, \dot{\bar{r}}, t)$$

where $\bar{r}$ and $\dot{\bar{r}}$ are the position and velocity of the spacecraft, t is time, μ is the gravitational constant, and g is a function describing perturbations from the Keplerian motion. For the purpose of this paper we are not concerned with the formula for $\ddot{\bar{r}}$ (1. 1) or its derivation. We only need to know that $\ddot{\bar{r}}$ is a function of $\bar{r}$, $\dot{\bar{r}}$, and t, where $\bar{r}$ and $\dot{\bar{r}}$ are the trajectory parameters to be estimated. For a powered maneuver we only need to know what additional trajectory parameters are used in the formula for $\ddot{\bar{r}}$ to describe the thrusting forces and changing mass. The estimated trajectory parameters become the initial conditions for integrating the equation of motion to predict new (a priori) values of the parameters at a future time.

The spacecraft may be observed from earth or from another spacecraft; or the spacecraft itself may measure quantities related to other bodies. The measurements are range, range-rate, and various angles, all of which can be formulated from a knowledge of the geometry and dynamics. The actual measurements and times are transmitted to the memory of a digital computer where they are available to the processor. The program solves a system of equations (called a filter) expressing the best estimates of the parameters as functions of the measurements. The computations for this are executed at the command of a controller. The filter is said to be sequential (or stepwise, or recursive) because it is used repeatedly while navigating.

The next several sections contain some fundamentals which should be understood before proceeding with the derivation of the filter. The advanced student at his own option may omit those sections with which he is already familiar.

2. MATRICES

A matrix is a rectangular array of elements with certain mathematical properties. Most of the properties which are important to us are listed below. [1]

If A is a matrix and a_{ij} is the element in the i^{th} row and j^{th} column, then

2.1 $A = (a_{ij}) \qquad (i = 1, \ldots, m), \; (j = 1, \ldots, n)$

<u>Addition</u>

2.2 $A + B = (a_{ij} + b_{ij})$

<u>Subtraction</u>

2.3 $A - B = (a_{ij} - b_{ij})$

<u>Multiplication</u> (α a scalar)

2.4 $\alpha A = A\alpha = (\alpha a_{ij})$

Let

$$\begin{matrix} A = (a_{ij}) \\ B = (b_{jk}) \end{matrix} \qquad \left\{ \begin{matrix} i = 1, \ldots, m \\ j = 1, \ldots, n \\ k = 1, \ldots, p \end{matrix} \right.$$

Then

2.5 $$AB = \left(\sum_{j=1}^{n} a_{ij} \, b_{jk} \right) = (c_{ik}) = C$$

Also

2.6 $(AB)C = A(BC)$ (associative)

2.7 $AB \neq BA$ (not commutative unless A and B are both diagonal matrices)

Identity

2.8 $I = (\delta_{ij})$ $\delta_{ij} = \begin{cases} 1 & (i = j) \\ 0 & (i \neq j) \end{cases}$

2.9 $AI = IA = A$

Zero

$A = \emptyset \rightarrow$ every $a_{ij} = 0$ and

2.10 $AB = BA = \emptyset$ $(B \neq \emptyset)$

Transpose

The transpose of A is written A^T.

2.11 $A = (a_{ij}) \longleftrightarrow (a_{ji}) = A^T$

Symmetric

2.12 $A = A^T \longleftrightarrow a_{ij} = a_{ji}$

Skew - symmetric

2.13 $A = -A^T \longleftrightarrow a_{ij} = -a_{ji} \rightarrow a_{ii} = 0$

Inverse

2.14 $B = A^{-1} \longleftrightarrow AB = BA = I$

also

2.15 $(AC)^{-1} = C^{-1}A^{-1}$

and

2.16 $(A^T)^{-1} = (A^{-1})^T = A^{-T}$

Partitioning (an example)

Let

$$\left.\begin{array}{l} A = (a_{ij}), \ B = (b_{ij}) \quad (i, j = 1, \ldots, n) \\ A_{11} = (a_{ij}), \ B_{11} = (b_{ij}) \quad (i, j = 1, \ldots, m) \end{array}\right\} \quad (m < n)$$

$$A = \begin{bmatrix} A_{11} & A_{12} \\ A_{21} & A_{22} \end{bmatrix}, \qquad B = \begin{bmatrix} B_{11} & B_{12} \\ B_{21} & B_{22} \end{bmatrix}$$

Then

$$A+B = \begin{bmatrix} A_{11} + B_{11} & A_{12} + B_{12} \\ A_{21} + B_{21} & A_{22} + B_{22} \end{bmatrix} \tag{2.17}$$

$$AB = \begin{bmatrix} A_{11}B_{11} + A_{12}B_{21} & A_{11}B_{12} + A_{12}B_{22} \\ A_{21}B_{11} + A_{22}B_{21} & A_{21}B_{12} + A_{22}B_{22} \end{bmatrix} \tag{2.18}$$

There are several ways of formulating the inverse of a partitioned symmetric matrix. The way suggested here can be proved easily. [1]

Let

$$A = A^T \rightarrow A_{12}^T = A_{21}$$

and

$$A^{-1} = B \rightarrow B_{12}^T = B_{21}$$

Then

2.19 $$\begin{bmatrix} A_{11} & A_{12} \\ A_{12}^T & A_{22} \end{bmatrix}^{-1} = \begin{bmatrix} B_{11} & B_{12} \\ B_{12}^T & B_{22} \end{bmatrix}$$

where

2.20 $$B_{11} = \left[A_{11} - A_{12}A_{22}^{-1}A_{12}^T\right]^{-1}$$

2.21 $$B_{22} = \left[A_{22} - A_{12}^T A_{11}^{-1}A_{12}\right]^{-1}$$

2.22 $$B_{12} = -B_{11}A_{12}A_{22}^{-1}$$

2.23 $$B_{12}^T = -B_{22}A_{12}^T A_{11}^{-1}$$

3. VECTORS

Let $\hat{i}$, $\hat{j}$, $\hat{k}$ be unit basis vectors in an orthogonal inertial frame. Then a position vector may be expressed

3.1 $\bar{r} = \hat{i}x + \hat{j}y + \hat{k}z$

or equivalently

3.2 $$\bar{r} = [\hat{i}\ \hat{j}\ \hat{k}]\begin{bmatrix} x \\ y \\ z \end{bmatrix}$$

Since we are using matrix notation throughout, it is convenient to omit the inertial basis vectors and express the vector as the ordered column of its components. Then

3.3 $$\bar{r} = \begin{bmatrix} x \\ y \\ z \end{bmatrix}$$, a column vector

and

3.4 $\bar{r}^T = [x, y, z]$, a row vector.

Addition

3.5 $$\bar{r}_1 + \bar{r}_2 = \begin{bmatrix} x_1 \\ y_1 \\ z_1 \end{bmatrix} + \begin{bmatrix} x_2 \\ y_2 \\ z_2 \end{bmatrix} = \begin{bmatrix} x_1 + x_2 \\ y_1 + y_2 \\ z_1 + z_2 \end{bmatrix}$$

Dot product

3.6 $$\bar{r}_1 \cdot \bar{r}_2 = \bar{r}_1^T\bar{r}_2 = [x_1, y_1, z_1]\begin{bmatrix} x_2 \\ y_2 \\ z_2 \end{bmatrix} = x_1x_2 + y_1y_2 + z_1z_2$$

(a scalar)

Cross product

Let

3.7 $\bar{r}^T = [x, y, z]$; $\bar{v}^T = [\dot{x}, \dot{y}, \dot{z}]$

Then by the definition of vector analysis

3.8 $$\bar{r} \times \bar{v} = \begin{vmatrix} \hat{i} & \hat{j} & \hat{k} \\ x & y & z \\ \dot{x} & \dot{y} & \dot{z} \end{vmatrix} \longleftrightarrow \begin{bmatrix} y\dot{z} - z\dot{y} \\ z\dot{x} - x\dot{z} \\ x\dot{y} - y\dot{x} \end{bmatrix}$$

With every 3-dimensional vector, $\bar{r}$, there is associated a skew-symmetric matrix, $\overset{\approx}{r}$, as follows:

3.9 $$\bar{r} = \begin{bmatrix} x \\ y \\ z \end{bmatrix} \longleftrightarrow \begin{bmatrix} 0 & -z & y \\ z & 0 & -x \\ -y & x & 0 \end{bmatrix} = \overset{\approx}{r}$$ [2]

Now the cross-product can be expressed

3.10 $$\bar{r} \times \bar{v} \longleftrightarrow \overset{\approx}{r}\bar{v} = \begin{bmatrix} 0 & -z & y \\ z & 0 & -x \\ -y & x & 0 \end{bmatrix} \begin{bmatrix} \dot{x} \\ \dot{y} \\ \dot{z} \end{bmatrix} = \begin{bmatrix} y\dot{z} - z\dot{y} \\ z\dot{x} - x\dot{z} \\ x\dot{y} - y\dot{x} \end{bmatrix}$$ [2]

and the result of 3.10 is the same as 3.8.

The following equivalencies can all be proved easily using 3.6 and 3.10. [2]

3.11 $\overset{\approx}{r}\bar{v} = -\overset{\approx}{v}\bar{r}$

3.12 $\bar{r} \times (\bar{r} \times \bar{v}) \longleftrightarrow \overset{\approx}{r}\overset{\approx}{r}\bar{v}$

3.13 $\widetilde{\widetilde{r \times v}} = \overset{\approx}{r}\overset{\approx}{v} - \overset{\approx}{v}\overset{\approx}{r}$

3.14 $(\bar{r} \times \bar{v}) \times \bar{r} \longleftrightarrow (\overset{\approx}{r}\overset{\approx}{v} - \overset{\approx}{v}\overset{\approx}{r})\bar{r} = \overset{\approx}{r}\overset{\approx}{v}\bar{r}$

3.15 $\bar{w} \cdot \bar{r} \times \bar{v} \longleftrightarrow \bar{w}^T \overset{\approx}{r}\bar{v}$

In this manner any combination of dot and cross products is equivalent to a product of skew-symmetric matrices and vectors.

Rotating frames

Although we can omit the basis vectors of the inertial frame, it may be necessary to express the basis vectors of a rotating frame. Let

3.16 $\bar{\rho}_R^T = [\rho_1 \ \rho_2 \ \rho_3]$ be a vector expressed relative to a rotating frame,

and

$\bar{\rho}_I^T$ be the same vector expressed in the inertial frame.

Also let $[\hat{e}_1 \ \hat{e}_2 \ \hat{e}_3]$ be the unit basis vectors in the rotating frame. Each $\hat{e}_i$ can be expressed in the inertial frame, e.g.,

3.17 $\hat{e}_1^T = [e_{1x} \ e_{1y} \ e_{1z}]$

and

3.18 $$[\hat{e}_1 \ \hat{e}_2 \ \hat{e}_3] = \begin{bmatrix} e_{1x} & e_{2x} & e_{3x} \\ e_{1y} & e_{2y} & e_{3y} \\ e_{1z} & e_{2z} & e_{3z} \end{bmatrix} \equiv T$$

Then

3.19 $\bar{\rho}_I = [\hat{e}_1 \ \hat{e}_2 \ \hat{e}_3] [\rho_1 \ \rho_2 \ \rho_3]^T$

3.20 $\bar{\rho}_I = T\bar{\rho}_R$

State vectors

Up to here we have been discussing vectors which can be plotted in Cartesian 3-space, but abstractly a vector can have many more elements than three. The trajectory parameters to be estimated, for example, are expressed as an ordered column of functionally independent basis elements, called the state vector, S.

3.21 $$S = \begin{bmatrix} x \\ y \\ z \\ \dot{x} \\ \dot{y} \\ \dot{z} \end{bmatrix} = \begin{bmatrix} \bar{r} \\ \bar{v} \end{bmatrix} \qquad \text{(by partitioning)}$$

(Note: For notational convenience later on the bar over S and certain other vectors is omitted.)

The basis elements of S are chosen so that they are functionally independent, e.g.,

$$\frac{\partial y}{\partial x} = 0 \; , \quad \frac{\partial x}{\partial \dot{z}} = 0 \; , \quad \text{etc.}$$

This choice results in more convenient formulations. If there are more trajectory parameters to be estimated, then the corresponding basis elements are adjoined to S, and the dimension of S is increased accordingly.

Observation vectors

Each element of the computed observation vector is a function of the basis elements of S, i.e., the trajectory parameters. Consider the vector modeling three measurements at time t_i:

3.22 $$\beta_i^T = [\beta_{i1} \quad \beta_{i2} \quad \beta_{i3}]$$

Here each β_{ij} is a scalar function of the trajectory parameters, i.e.,

3.23 $$\beta_i = \beta(S_i) \qquad\qquad (S_i \equiv S_{t_i})$$

Magnitude of a vector (example)

3.24 $$r = |\bar{r}| = \sqrt{\bar{r}^T \bar{r}} = \sqrt{x^2 + y^2 + z^2}$$

where

$$\bar{r}^T = [x \;\; y \;\; z]$$

Unit vector (examples)

3.25 $$\hat{r} = \frac{\bar{r}}{|\bar{r}|} = \frac{\bar{r}}{r}$$

If

$$\bar{e} = \overset{\approx}{r} v$$

then

3.26 $$\hat{e} = \frac{\overset{\approx}{r} \bar{v}}{|\overset{\approx}{r} \bar{v}|}$$

Dyadic (example)

3.27 $$\bar{r}\bar{r}^T = \begin{bmatrix} x \\ y \\ z \end{bmatrix} [x \;\; y \;\; z] = \begin{bmatrix} x^2 & xy & xz \\ xy & y^2 & yz \\ xz & yz & z^2 \end{bmatrix}$$

Note that the determinant of 3.27 is zero, and the matrix has no inverse.

Quadratic form

Let A be an n x n symmetric matrix of constants and $\bar{u}$ be an n x 1 vector of variable elements. Then the scalar function of elements of $\bar{u}$,

3.28 $$\varphi = \bar{u}^T A \bar{u} ,$$

is a quadratic form. If $\varphi > 0$ for all $\bar{u}$, then both φ and A are said to be positive definite. If $\varphi \geq 0$ for all $\bar{u}$, then both φ and A are said to be positive semi-definite.

4. PROBLEMS

4.1 $$A = \begin{bmatrix} 3 & -3 & 1 \\ 0 & 1 & 2 \\ 2 & -4 & 4 \end{bmatrix} ; \quad B = \begin{bmatrix} 2 & 0 & 6 \\ 4 & -1 & 5 \\ 1 & 1 & 3 \end{bmatrix} ; \quad C = \begin{bmatrix} 2 & -3 & 4 \\ 1 & 1 & -2 \\ 0 & 2 & 3 \end{bmatrix}$$

$$\alpha = 2$$

Compute:

(a) $A + B$
(b) $A - B$
(c) αA
(d) $B\alpha$
(e) AB
(f) BA
(g) $A(BC)$
(h) $(AB)C$
(i) AI
(j) IA
(k) A^T

4.2 (a) Give an example of a symmetric matrix.
(b) Give an example of a skew-symmetric matrix.

Let

4.3 $$A_{11} = \begin{bmatrix} 3 & -3 \\ 0 & 1 \end{bmatrix}, \quad A_{12} = \begin{bmatrix} 1 \\ 2 \end{bmatrix}, \quad A_{21} = [2, \ -4], \quad A_{22} = [4]$$

$$B_{11} = \begin{bmatrix} 2 & 0 \\ 4 & -1 \end{bmatrix}, \quad B_{12} = \begin{bmatrix} 6 \\ 5 \end{bmatrix}, \quad B_{21} = [1, \ 1], \quad B_{22} = [3]$$

Then

$$A = \begin{bmatrix} A_{11} & A_{12} \\ A_{21} & A_{22} \end{bmatrix}, \quad B = \begin{bmatrix} B_{11} & B_{12} \\ B_{21} & B_{22} \end{bmatrix}$$

Compute AB using this partition and compare the result with 4.1e.

4.4 $$\bar{r} = \begin{bmatrix} 3 \\ 1 \\ 2 \end{bmatrix}, \quad \bar{v} = \begin{bmatrix} 1 \\ 5 \\ -3 \end{bmatrix}, \quad \Omega = \begin{bmatrix} 4 & 1 & -3 \\ 1 & 2 & 0 \\ -3 & 0 & 3 \end{bmatrix}$$

Compute (without using $\hat{i}$, $\hat{j}$, $\hat{k}$):

(a) $\bar{r}^T \bar{v}$

(b) $\bar{r}\bar{r}^T$

(c) $|\bar{r}| = r$

(d) $\hat{r}$

(e) $\bar{r}^T \Omega$

(f) $\Omega \bar{r}$

(g) $\bar{r}^T \Omega \bar{r}$

(h) $\tilde{\tilde{r}} \bar{v}$

(i) $(\bar{r} \times \bar{v}) \times \bar{r}$

(j) $|S|$, where $S^T = [\bar{r}^T, \bar{v}^T]$

4.5 If

$$\bar{r}^T = [x \quad y \quad z]$$

and

$$\Omega^{-1} = \begin{bmatrix} a^2 & 0 & 0 \\ 0 & b^2 & 0 \\ 0 & 0 & c^2 \end{bmatrix}$$

Give an interpretation of the equation

$$\bar{r}^T \Omega \bar{r} = 1.$$

5. PARTIAL DERIVATIVES

This section shows some convenient methods of differentiating scalars and vectors which are expressed as functions of vectors and matrices. The rules are simple and often will be demonstrated by an example rather than defined.

Let

5.1 $\varphi \equiv$ a scalar

5.2 $S^T \equiv \left[\bar{r}^T, \bar{v}^T\right]$

5.3 $\bar{r}^T \equiv [x, y, z]$

5.4 $\bar{v}^T \equiv [\dot{x}, \dot{y}, \dot{z}]$

Derivative of a scalar

Clearly

$$\varphi = \varphi^T$$

By definition the derivative of φ with respect to several variables is a row vector, e.g.,

5.5 $$\frac{\partial \varphi}{\partial \bar{r}} = \left[\frac{\partial \varphi}{\partial x}, \frac{\partial \varphi}{\partial y}, \frac{\partial \varphi}{\partial z}\right]$$

Then $\left(\frac{\partial \varphi}{\partial \bar{r}}\right)^T$ is a column, and by definition

5.6 $$\left(\frac{\partial \varphi}{\partial \bar{r}}\right)^T \equiv \frac{\partial \varphi}{\partial \bar{r}}^T$$

Derivative of a vector

The partial derivative of a vector with respect to several variables is a matrix. Let

$$\bar{r}_0^T = \left[x_0, y_0, z_0\right]$$

and

$$\bar{r} = \bar{r}(\bar{r}_0)$$

Then

5.7 $$\frac{\partial \bar{r}}{\partial \bar{r}_0} = \begin{bmatrix} \frac{\partial x}{\partial x_0} & \frac{\partial x}{\partial y_0} & \frac{\partial x}{\partial z_0} \\ \frac{\partial y}{\partial x_0} & \frac{\partial y}{\partial y_0} & \frac{\partial y}{\partial z_0} \\ \frac{\partial z}{\partial x_0} & \frac{\partial z}{\partial y_0} & \frac{\partial z}{\partial z_0} \end{bmatrix}$$

Another example:

$$\bar{w} = \begin{bmatrix} u \\ v \end{bmatrix}$$

where

$$u = u(x, y, z)$$

$$v = v(x, y, z)$$

Then

5.8 $$\frac{\partial \bar{w}}{\partial \bar{r}} = \begin{bmatrix} \frac{\partial u}{\partial x} & \frac{\partial u}{\partial y} & \frac{\partial u}{\partial z} \\ \frac{\partial v}{\partial x} & \frac{\partial v}{\partial y} & \frac{\partial v}{\partial z} \end{bmatrix}$$

Derivative of a matrix

The partial derivative of a matrix with respect to one variable is a matrix.

5.9 $$A = \begin{bmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{bmatrix} \longrightarrow \frac{\partial A}{\partial x} = \begin{bmatrix} \frac{\partial a_{11}}{\partial x} & \frac{\partial a_{12}}{\partial x} \\ \frac{\partial a_{21}}{\partial x} & \frac{\partial a_{22}}{\partial x} \end{bmatrix}$$

The derivative of a matrix with respect to several variables can be expressed in tensor notation. We are able to avoid this type of derivative and use matrix notation throughout.

Derivative of the dot product

The dot product is a scalar; so the derivative is a row vector. (5.5) Let $\bar{u}$, $\bar{v}$, $\bar{w}$ be three vectors. Recall

$$\bar{u}^T\bar{v} = \bar{v}^T\bar{u} .$$

Then

$$5.10 \qquad \frac{\partial(\bar{u}^T\bar{v})}{\partial\bar{w}} = \bar{u}^T\frac{\partial\bar{v}}{\partial\bar{w}} + \bar{v}^T\frac{\partial\bar{u}}{\partial\bar{w}}$$

It follows that

$$5.11 \qquad \frac{\partial(\bar{u}^T\bar{u})}{\partial\bar{w}} = 2\bar{u}^T\frac{\partial\bar{u}}{\partial\bar{w}}$$

Another example: (5.2, 5.3, 5.4)

$$\frac{\partial(\bar{r}^T\bar{v})}{\partial S} = \bar{r}^T\frac{\partial\bar{v}}{\partial S} + \bar{v}^T\frac{\partial\bar{r}}{\partial S}$$

$$= \bar{r}^T[0, I] + \bar{v}^T[I, 0]$$

$$5.12 \qquad = \left[\bar{v}^T, \bar{r}^T\right]$$

Here we took advantage of partitioning, i.e.,

$$\frac{\partial\bar{r}}{\partial S} = \left[\frac{\partial\bar{r}}{\partial\bar{r}}, \frac{\partial\bar{r}}{\partial\bar{v}}\right] = [I, 0] \text{ , etc.}$$

Derivative of a quadratic form

The quadratic form is a scalar; so the derivative is a row vector.

Let Ω be a symmetric matrix of constants, and

$$\varphi \equiv \overline{v}^T \Omega \overline{u} = \overline{u}^T \Omega \overline{v}$$

Then

5.13 $$\frac{\partial \varphi}{\partial S} = \overline{v}^T \Omega \frac{\partial \overline{u}}{\partial S} + \overline{u}^T \Omega \frac{\partial \overline{v}}{\partial S}$$

Let Ω be (6 x 6) and $\varphi \equiv S^T \Omega S$, then

$$\frac{\partial \varphi}{\partial S} = 2S^T \Omega \frac{\partial S}{\partial S} = 2S^T \Omega I = 2S^T \Omega$$

Or let Ω be (3 x 3) and $\varphi \equiv \overline{r}^T \Omega \overline{r}$, then

$$\frac{\partial \varphi}{\partial S} = 2\overline{r}^T \Omega \frac{\partial \overline{r}}{\partial S} = 2\overline{r}^T \Omega [I, \emptyset]$$

Derivative of the product of a scalar and vector

Note that the product is commutative.

$$\varphi \overline{u} = \overline{u} \varphi$$

Then

5.14 $$\frac{\partial (\varphi \overline{u})}{\partial S} = \varphi \frac{\partial \overline{u}}{\partial S} + \overline{u} \frac{\partial \varphi}{\partial S}$$

Another example:

Find $\frac{\partial \hat{\rho}}{\partial S}$

$$\frac{\partial \hat{\rho}}{\partial S} = \frac{\partial}{\partial S} \left(\frac{\overline{\rho}}{\sqrt{\overline{\rho}^T \overline{\rho}}} \right) = \frac{1}{\sqrt{\overline{\rho}^T \overline{\rho}}} \frac{\partial \overline{\rho}}{\partial S} + \overline{\rho} \frac{\partial}{\partial S} \left(\frac{1}{\sqrt{\overline{\rho}^T \overline{\rho}}} \right)$$

$$= \frac{1}{\sqrt{\overline{\rho}^T \overline{\rho}}} \frac{\partial \overline{\rho}}{\partial S} + \overline{\rho} \left(-\frac{1}{2} \right) \frac{1}{(\overline{\rho}^T \overline{\rho})^{3/2}} (2) \overline{\rho}^T \frac{\partial \overline{\rho}}{\partial S}$$

5.15
$$\frac{\partial\hat{\rho}}{\partial S} = \left[I - \hat{\rho}\hat{\rho}^T\right]\frac{1}{|\rho|}\frac{\partial\bar{\rho}}{\partial S}$$

And taking the transpose of 5.15

5.16
$$\frac{\partial\hat{\rho}^T}{\partial S} = \frac{1}{|\rho|}\frac{\partial\bar{\rho}^T}{\partial S}\left[I - \hat{\rho}\hat{\rho}^T\right]$$

Derivative of the cross product

Recall $\bar{r}\ \mathrm{x}\ \bar{v} \longleftrightarrow \overset{\approx}{r}\bar{v} = -\overset{\approx}{v}\bar{r}$ (3.11)

Then

5.17
$$\frac{\partial\left(\overset{\approx}{r}\bar{v}\right)}{\partial S} = \overset{\approx}{r}\frac{\partial\bar{v}}{\partial S} - \overset{\approx}{v}\frac{\partial\bar{r}}{\partial S} = \overset{\approx}{r}[0, I] - \overset{\approx}{v}[I, 0] = [-\overset{\approx}{v}, \overset{\approx}{r}]$$

Derivative of $\bar{w}\ \mathrm{x}\ (\bar{r}\ \mathrm{x}\ \bar{v}) \longleftrightarrow \overset{\approx}{w}\overset{\approx}{r}\bar{v}$

$\overset{\approx}{w}\overset{\approx}{r}\bar{v} = -\overset{\approx}{w}\overset{\approx}{v}\bar{r} = -(\overset{\approx}{r}\overset{\approx}{v} - \overset{\approx}{v}\overset{\approx}{r})\bar{w}$ (3.11 - 14)

Then

5.18
$$\frac{\partial(\overset{\approx}{w}\overset{\approx}{r}\bar{v})}{\partial S} = \overset{\approx}{w}\overset{\approx}{r}\frac{\partial\bar{v}}{\partial S} - \overset{\approx}{w}\overset{\approx}{v}\frac{\partial\bar{r}}{\partial S} - \overset{\approx}{r}\overset{\approx}{v}\frac{\partial\bar{w}}{\partial S} + \overset{\approx}{v}\overset{\approx}{r}\frac{\partial\bar{w}}{\partial S}$$

Let $\bar{w} = \bar{r}$, then

$$\frac{\partial(\overset{\approx}{w}\overset{\approx}{r}\bar{v})}{\partial S} = \overset{\approx}{r}\overset{\approx}{r}\frac{\partial\bar{v}}{\partial S} - \overset{\approx}{r}\overset{\approx}{v}\frac{\partial\bar{r}}{\partial S} + [\overset{\approx}{v}\overset{\approx}{r} - \overset{\approx}{r}\overset{\approx}{v}]\frac{\partial\bar{r}}{\partial S}$$

$$= \overset{\approx}{r}\overset{\approx}{r}[0, I] - \overset{\approx}{r}\overset{\approx}{v}[I, 0] + [\overset{\approx}{v}\overset{\approx}{r} - \overset{\approx}{r}\overset{\approx}{v}][I, 0]$$

5.19
$$\frac{\partial(\overset{\approx}{r}\overset{\approx}{r}\bar{v})}{\partial S} = [\overset{\approx}{v}\overset{\approx}{r} - 2\overset{\approx}{r}\overset{\approx}{v}, \overset{\approx}{r}\overset{\approx}{r}]$$

Derivative of $\bar{w}^T \overset{\approx}{r} \bar{v}$

$$\bar{w}^T \overset{\approx}{r} \bar{v} = -\bar{w}^T \overset{\approx}{v} \bar{r} = -\bar{v}^T \overset{\approx}{r} \bar{w}$$

5.20 $$\frac{\partial\left(\bar{w}^T \overset{\approx}{r} \bar{v}\right)}{\partial S} = \bar{w}^T \overset{\approx}{r} \frac{\partial \bar{v}}{\partial S} - \bar{w}^T \overset{\approx}{v} \frac{\partial \bar{r}}{\partial S} - \bar{v}^T \overset{\approx}{r} \frac{\partial \bar{w}}{\partial S}$$

The gradient of a scalar

From vector analysis

5.21 $$\nabla \varphi = \hat{i} \frac{\partial \varphi}{\partial x} + \hat{j} \frac{\partial \varphi}{\partial y} + \hat{k} \frac{\partial \varphi}{\partial z} \longleftrightarrow \begin{bmatrix} \frac{\partial \varphi}{\partial x} \\ \frac{\partial \varphi}{\partial y} \\ \frac{\partial \varphi}{\partial z} \end{bmatrix} = \frac{\partial \varphi}{\partial \bar{r}}^T$$

The gradient of a vector

From vector analysis

$$\nabla \bar{v} = \hat{i} \frac{\partial \bar{v}}{\partial x} + \hat{j} \frac{\partial \bar{v}}{\partial y} + \hat{k} \frac{\partial \bar{v}}{\partial z}$$

$$= \hat{i}\left(\hat{i}\frac{\partial v_x}{\partial x} + \hat{j}\frac{\partial v_y}{\partial x} + \hat{k}\frac{\partial v_z}{\partial x}\right) + \hat{j}\left(\hat{i}\frac{\partial v_x}{\partial y} + \hat{j}\frac{\partial v_y}{\partial y} \; \hat{k}\frac{\partial v_z}{\partial y}\right) + \hat{k}\left(\hat{i}\frac{\partial v_x}{\partial z} + \hat{j}\frac{\partial v_y}{\partial z} + \hat{k}\frac{\partial v_z}{\partial z}\right)$$

5.22 $$\nabla \bar{v} = [\hat{i} \;\; \hat{j} \;\; \hat{k}] \begin{bmatrix} \frac{\partial v_x}{\partial x} & \frac{\partial v_y}{\partial x} & \frac{\partial v_z}{\partial x} \\ \frac{\partial v_x}{\partial y} & \frac{\partial v_y}{\partial y} & \frac{\partial v_z}{\partial y} \\ \frac{\partial v_x}{\partial z} & \frac{\partial v_y}{\partial z} & \frac{\partial v_z}{\partial z} \end{bmatrix} \begin{bmatrix} \hat{i} \\ \hat{j} \\ \hat{k} \end{bmatrix} \longleftrightarrow \frac{\partial \bar{v}^T}{\partial \bar{r}}$$

The gradient operator of matrix calculus

Some authors use this notation. We explain its use, but we shall not use it further. Let $\bar{x}^T = [x_1, x_2, \ldots, x_n]$. Then

5.23 $$\nabla_{\bar{x}} \equiv \begin{bmatrix} \frac{\partial}{\partial x_1} \\ \frac{\partial}{\partial x_2} \\ \vdots \\ \frac{\partial}{\partial x_n} \end{bmatrix}$$ [3, 4]

and

$$\nabla_{\bar{x}} \bar{v}^T \equiv \begin{bmatrix} \frac{\partial \bar{v}^T}{\partial x_1} \\ \frac{\partial \bar{v}^T}{\partial x_2} \\ \vdots \\ \frac{\partial \bar{v}^T}{\partial x_n} \end{bmatrix}$$

Suppose we take the gradient of $\bar{v}^T$ with respect to $\bar{r}$, then

$$\nabla = \begin{bmatrix} \frac{\partial}{\partial x} \\ \frac{\partial}{\partial y} \\ \frac{\partial}{\partial z} \end{bmatrix}$$

and

$$5.24 \quad \nabla \bar{v}^T = \begin{bmatrix} \frac{\partial}{\partial x} \\ \frac{\partial}{\partial y} \\ \frac{\partial}{\partial z} \end{bmatrix} [v_x \quad v_y \quad v_z] = \frac{\partial \bar{v}^T}{\partial \bar{r}}$$

Note that $\nabla \bar{v}^T$ of 5.24 is equivalent to the $\nabla \bar{v}$ of 5.22. By either of the above definitions the gradient is equivalent to a partial derivative, and from now on we shall use the partial derivative notation.

Chain rule (example)

If $\bar{r} = \bar{r}(\bar{u})$ and $\bar{u} = \bar{u}(\bar{w})$, then

$$5.25 \quad \frac{\partial \bar{r}}{\partial \bar{u}} \frac{\partial \bar{u}}{\partial \bar{w}} = \frac{\partial \bar{r}}{\partial \bar{w}}$$

Let

$$\bar{r} = \begin{bmatrix} x \\ y \end{bmatrix}, \quad \bar{u} = \begin{bmatrix} u_1 \\ u_2 \end{bmatrix}, \quad \bar{w} = \begin{bmatrix} w_1 \\ w_2 \end{bmatrix}$$

$$\begin{bmatrix} \frac{\partial x}{\partial u_1} & \frac{\partial x}{\partial u_2} \\ \frac{\partial y}{\partial u_1} & \frac{\partial y}{\partial u_2} \end{bmatrix} \begin{bmatrix} \frac{\partial u_1}{\partial w_1} & \frac{\partial u_1}{\partial w_2} \\ \frac{\partial u_2}{\partial w_1} & \frac{\partial u_2}{\partial w_2} \end{bmatrix} = \begin{bmatrix} \frac{\partial x}{\partial w_1} & \frac{\partial x}{\partial w_2} \\ \frac{\partial y}{\partial w_1} & \frac{\partial y}{\partial w_2} \end{bmatrix}$$

since, e.g.,

$$\frac{\partial x}{\partial w_1} = \frac{\partial x}{\partial u_1} \frac{\partial u_1}{\partial w_1} + \frac{\partial x}{\partial u_2} \frac{\partial u_2}{\partial w_1}$$

Then from 5.25 it follows that

5.26 $\frac{\partial \bar{r}}{\partial \bar{u}} \frac{\partial \bar{u}}{\partial \bar{r}} = I$

Summary of rules for differentiation

(a) $\frac{\partial \varphi}{\partial S}$ is a row vector.

(b) $\frac{\partial \bar{\bar{w}}}{\partial S}$ is a matrix.

(c) $\frac{\partial \left(\bar{u}^T \bar{v} \right)}{\partial \bar{w}} = \bar{u}^T \frac{\partial \bar{v}}{\partial \bar{\bar{w}}} + \bar{v}^T \frac{\partial \bar{u}}{\partial \bar{\bar{w}}}$

(d) $\frac{\partial (\varphi \bar{u})}{\partial \bar{w}} = \varphi \frac{\partial \bar{u}}{\partial \bar{\bar{w}}} + \bar{u} \frac{\partial \varphi}{\partial \bar{\bar{w}}}$

(e) $\frac{\partial \left(\bar{v}^T \Omega \bar{u} \right)}{\partial \bar{\bar{w}}} = \bar{v}^T \Omega \frac{\partial \bar{u}}{\partial \bar{\bar{w}}} + \bar{u}^T \Omega^T \frac{\partial \bar{v}}{\partial \bar{\bar{w}}}$

(and $\Omega = \Omega^T$ if symmetric)

(f) For expressions such as $\bar{r}^T \overset{\approx}{u} \overset{\approx}{v} \bar{w}$ etc., find equal expressions so that each element of the expression is permuted to a vector on the right. Take the sum of these with each right-hand element differentiated.

(g) $\frac{d}{dt} \left(\frac{\partial \bar{u}}{\partial S} \right) = \frac{\partial \dot{\bar{u}}}{\partial S}$

(h) If A and B are matrices, then

$\frac{d}{dt} (AB) = \dot{A}B + A\dot{B}$

5.27 Problems

Define

$$S = \begin{bmatrix} \bar{r} \\ \bar{v} \end{bmatrix}, \quad \bar{r} = \begin{bmatrix} x \\ y \\ z \end{bmatrix}, \quad \bar{v} = \begin{bmatrix} \dot{x} \\ \dot{y} \\ \dot{z} \end{bmatrix}$$

$$\bar{\rho} = \bar{R} - \bar{r}$$

$$\dot{\bar{\rho}} = \bar{V} - \bar{v}$$

$\bar{R}$ and $\bar{V}$ are functionally independent of S.

$\begin{bmatrix} \bar{r} \\ \bar{v} \end{bmatrix}$ is the vector of trajectory parameters of the CSM.

$\begin{bmatrix} \bar{R} \\ \bar{V} \end{bmatrix}$ is the vector of trajectory parameters of the LM.

Then

$\rho = \sqrt{\bar{\rho}^T \bar{\rho}}$ is a range measurement from the CSM → LM, and $\frac{d\rho}{dt}$ is a range-rate measurement, where $\frac{d\rho}{dt} = \hat{\rho}^T \dot{\bar{\rho}}$.

(a) Find $\frac{\partial \rho}{\partial S}$ ans. $\left\{ = \hat{\rho}^T \frac{\partial \bar{\rho}}{\partial S} = \left[-\hat{\rho}^T, 0 \right] \right\}$

(b) Show that $\frac{d\rho}{dt} = \hat{\rho}^T \dot{\bar{\rho}}$

(c) Find $\frac{\partial \left(\hat{\rho}^T \dot{\bar{\rho}} \right)}{\partial S} = \hat{\rho}^T \frac{\partial \dot{\bar{\rho}}}{\partial S} + \dot{\bar{\rho}}^T \frac{\partial \hat{\rho}}{\partial S}$

(d) Find $\frac{\partial\left(\hat{\rho}^T\dot{\bar{\rho}}\right)}{\partial S} = \frac{\partial}{\partial S}\left(\frac{\bar{\rho}^T\dot{\bar{\rho}}}{\sqrt{\bar{\rho}^T\bar{\rho}}}\right) = \frac{1}{\sqrt{\bar{\rho}^T\bar{\rho}}}\frac{\partial}{\partial S}\left(\bar{\rho}^T\dot{\bar{\rho}}\right) + \bar{\rho}^T\dot{\bar{\rho}}\frac{\partial}{\partial S}\left(\frac{1}{\sqrt{\bar{\rho}^T\bar{\rho}}}\right)$

(e) Find $\frac{\partial}{\partial S}\left(\frac{d\rho}{dt}\right) = \frac{d}{dt}\left[\frac{\partial\rho}{\partial S}\right]$

Hint: Use the results of part (a) to solve for $\frac{\partial}{\partial S}\left(\frac{d\rho}{dt}\right)$.

Parts (c), (d), and (e) all have the same answer (of course) as follows:

$$\left[\frac{\dot{\bar{\rho}}^T}{\rho}\left(\hat{\rho}\hat{\rho}^T - I\right),\ -\hat{\rho}^T\right]$$

In parts (f), (g), and (h) let

$$\hat{e}_1 = \hat{r}\ ,\quad \hat{e}_3 = \frac{\bar{r} \times \bar{v}}{|\bar{r} \times \bar{v}|}\ ,\quad \hat{e}_2 = \frac{(\bar{r} \times \bar{v}) \times \bar{r}}{|(\bar{r} \times \bar{v}) \times \bar{r}|}$$

(f) Show that $\frac{\partial\hat{e}_1}{\partial S} = \frac{1}{r}\left[I - \hat{r}\hat{r}^T\right][I,\ 0]$

(g) $$\frac{\partial\hat{e}_3}{\partial S} = \frac{1}{e_3}\left[I - \hat{e}_3\hat{e}_3^T\right]\left[-\overset{\approx}{v},\ \overset{\approx}{r}\right]$$

(h) $$\frac{\partial\hat{e}_2}{\partial S} = \frac{1}{e_2}\left[I - \hat{e}_2\hat{e}_2^T\right]\left[2\overset{\approx}{r}\overset{\approx}{v} - \overset{\approx}{v}\overset{\approx}{r},\ -\overset{\approx}{r}\overset{\approx}{r}\right]$$

In parts (i) through (n) let

$$2\varphi = (\tilde{S} - S)^T\Gamma^{-1}(\tilde{S} - S)$$

where $\tilde{S}$ is a constant vector and Γ^{-1} is a symmetric matrix of constants (6 x 6) and partitioned as

$$\Gamma^{-1} = \begin{bmatrix} G_{11} & G_{12} \\ G_{12}^{T} & G_{22} \end{bmatrix} \qquad \text{where each } G_{ij} \text{ is } (3 \times 3).$$

Show that

(i) $$\frac{\partial \varphi}{\partial \bar{r}} = -(\tilde{S} - S)^{\Gamma} \begin{bmatrix} G_{11} \\ G_{12}^{T} \end{bmatrix}$$

(j) $$\frac{\partial \varphi}{\partial \bar{v}} = -(\tilde{S} - S)^{T} \begin{bmatrix} G_{12} \\ G_{22} \end{bmatrix}$$

(k) $$\frac{\partial \varphi}{\partial S} = -(\tilde{S} - S)^{T} \Gamma^{-1}$$

(l) $$\frac{\partial^2 \varphi}{\partial \bar{r}^2} = G_{11}$$

(m) $$\frac{\partial^2 \varphi}{\partial \bar{v}^2} = G_{22}$$

(n) $$\frac{\partial^2 \varphi}{\partial S^2} = \Gamma^{-1}$$

Note that $\hat{r} \times \hat{v} \longleftrightarrow \hat{\overset{\approx}{r}}\,\hat{v} = -\hat{\overset{\approx}{v}}\,\hat{r}$

Show that

$$\text{(o)}\quad \frac{\partial(\hat{r} \times \hat{v})}{\partial S} = \left[-\frac{\hat{\overset{\approx}{v}}}{r}\left(I - \hat{r}\hat{r}^T\right),\ \frac{\hat{\overset{\approx}{r}}}{v}\left(I - \hat{v}\hat{v}^T\right) \right]$$

Hint: Use the proved formula

$$\frac{\partial \hat{u}}{\partial S} = \frac{1}{u}\left[I - \hat{u}\hat{u}^T\right]\frac{\partial \bar{u}}{\partial S}$$

What are the dimensions of the matrix answer?

$$\text{(p)}\quad \frac{\partial\left(\hat{r}^T\hat{v}\right)}{\partial S} = \left[\frac{\hat{v}^T}{r}\left(I - \hat{r}\hat{r}^T\right),\ \frac{\hat{r}^T}{v}\left(I - \hat{v}\hat{v}^T\right) \right]$$

What are the dimensions of the matrix answer?

6. TAYLOR SERIES

Complete discussions of the Taylor series can be found in almost any text on advanced calculus. The only purpose here is to show how to express the first-order Taylor series in matrix form. [8]

One dependent and two independent variables

In Section 3 the superscript carat was used to denote a unit vector. In this section it is used to denote a close estimate of a scalar or vector, as follows. Let y be a scalar function of two scalar variables

$$y = y(x_1, x_2)$$

and let $\hat{x}_1$ and $\hat{x}_2$ be close approximations of x_1 and x_2 so that a linear approximation, $\hat{y}$, of y is valid:

$$\hat{y} = y\left(\hat{x}_1, \hat{x}_2\right)$$

Then a first-order expansion of $\hat{y}$ about y is

6.1 $$\hat{y} = y + \frac{\partial y}{\partial x_1}\left(\hat{x}_1 - x_1\right) + \frac{\partial y}{\partial x_2}\left(\hat{x}_2 - x_2\right)$$

Define

$$x = \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} \quad , \hat{x} = \begin{bmatrix} \hat{x}_1 \\ \hat{x}_2 \end{bmatrix}$$

Then 6.1 can be written

6.2 $$\hat{y} = y + \frac{\partial y}{\partial x}\left(\hat{x} - x\right)$$

Three dependent and three independent variables

Let

$$y_i = y_i(x_1, x_2, x_3) \qquad (i = 1, \ldots, 3)$$

$$\hat{y}_i = y_i\left(\hat{x}_1, \hat{x}_2, \hat{x}_3\right)$$

$\hat{y}_i - y_i$ is small enough to allow linear approximations. Then

6.3 $$\hat{y}_i = y_i + \frac{\partial y_i}{\partial x_1}\left(\hat{x}_1 - x_1\right) + \frac{\partial y_i}{\partial x_2}\left(\hat{x}_2 - x_2\right) + \frac{\partial y_i}{\partial x_3}\left(\hat{x}_3 - x_3\right)$$

Define

$$Y = \begin{bmatrix} y_1 \\ y_2 \\ y_3 \end{bmatrix}, \quad X = \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix}$$

Then 6.3 can be written

6.4 $$\hat{Y} = Y + \frac{\partial Y}{\partial X}(\hat{X} - X)$$

n dependent and n independent variables

$$y_i = y_i(x_1, x_2, \ldots, x_n) \qquad (i = 1, \ldots, n)$$

$$\hat{y}_i = y_i\left(\hat{x}_1, \hat{x}_2, \ldots, \hat{x}_n\right)$$

$\hat{y}_i - y_i$ is small. Then

6.5 $$\hat{y}_i = y_i + \frac{\partial y_i}{\partial x_1}\left(\hat{x}_1 - x_1\right) + \ldots + \frac{\partial y_i}{\partial x_n}\left(\hat{x}_n - x_n\right)$$

Define

$$Y = \begin{bmatrix} y_1 \\ y_2 \\ \vdots \\ y_n \end{bmatrix}, \quad X = \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix}$$

Then 6.5 can be written the same as 6.4

$$6.6 \qquad \hat{Y} = Y + \frac{\partial Y}{\partial X} (\hat{X} - X)$$

This extension to n variables is apparent without a formal proof. The saving in notation is obvious when 6.6 is compared to 6.5. Another necessary assumption is that the functions are continuous in the region of the expansion and that $\frac{\partial Y}{\partial X}$ exists.

7. NEWTON'S METHOD

Newton's method of successive linear approximations can be used to get the solution to n non-linear equations in n unknowns. There is a lot of theory written about this method, particularly in connection with convergence properties [13]. Although this is a very worthwhile subject to study, for our purpose it is sufficient just to demonstrate the method and comment on the convergence criteria.

One equation and one unknown

Let y be some non-linear function of x.

7.1 $y = y(x)$

and there exists some value, $\hat{x}$, of x such that

7.2 $\hat{y} = y(\hat{x}) = 0$

Then find $\hat{x}$.

Let $\tilde{x}$ be a close approximation of $\hat{x}$ such that $\hat{x} - \tilde{x}$ is small and linear approximations are valid. Express $\hat{y}$ as a first-order Taylor series expansion (6.1).

7.3 $\hat{y} = \tilde{y} + \frac{dy}{d\tilde{x}} (\hat{x} - \tilde{x}) = 0$ where $\tilde{y} = y(\tilde{x})$ and $\frac{dy}{d\tilde{x}} = \tilde{y}' = y'(\tilde{x})$

Then

7.4 $\hat{x} = \tilde{x} - \left(\frac{dy}{d\tilde{x}}\right)^{-1} \tilde{y}$

Equation 7.4 can be re-written for iteration, where subscript, n, indicates the n^{th} iteration.

7.5 $x_{n+1} = x_n - \frac{dx}{dy_n} y_n$

If convergence criteria are satisfied after n iterations, we consider that

$$\hat{x} = x_n$$

The manner of convergence by this succession of linear approximations is illustrated in Figure 7.1.

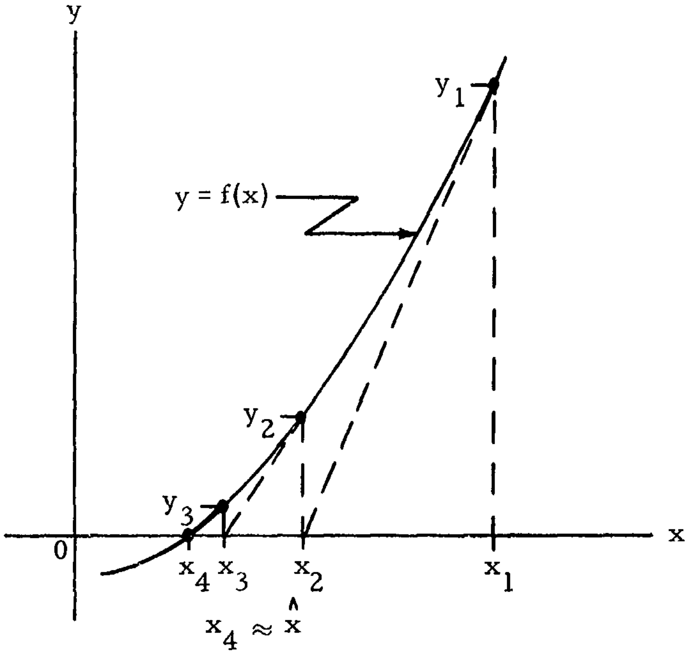


Figure 7.1

n equations and n unknowns

To illustrate this we shall solve a problem which will confront us later on.

Consider the following system of n non-linear equations in n unknowns.

7.6 $\Phi(\hat{S}) = 0$

Find $\hat{S}$.

By our notation $\Phi(\hat{S}) = \hat{\Phi}$, and S is an n-element state vector.

7.7 $S^T = [x_1, \ldots, x_n]$

7.8 $\Phi = \Phi(S)$

or equivalently

7.9 $$\begin{bmatrix} \Phi_1 \\ \vdots \\ \Phi_n \end{bmatrix} = \begin{bmatrix} \Phi_1(x_1, \ldots, x_n) \\ \vdots \\ \Phi_n(x_1, \ldots, x_n) \end{bmatrix}$$

Let $\tilde{S}$ be a close first approximation of $\hat{S}$ such that $\hat{S} - \tilde{S}$ is small and linear approximations are valid. Then an approximation of $\hat{\Phi}$ is

7.10 $$\hat{\Phi} = \tilde{\Phi} + \frac{\partial \Phi}{\partial \tilde{S}} (\hat{S} - \tilde{S}) = 0 \qquad (6.6)$$

where $\tilde{\Phi} = \Phi(\tilde{S})$

Solving for $\hat{S}$

7.11 $$\hat{S} = \tilde{S} - \left(\frac{\partial \Phi}{\partial \tilde{S}}\right)^{-1} \tilde{\Phi}$$

Assume that $\frac{\partial \Phi}{\partial \tilde{S}}$ is non-singular.

Since $\hat{S}$ is a closer approximation to the solution than $\tilde{S}$ in 7.11, we can rewrite 7.11 for iteration.

7.12 $$S_{n+1} = S_n - \left(\frac{\partial \Phi}{\partial S_n}\right)^{-1} \Phi_n$$

If convergence occurs after n iterations, consider that

$$\hat{S} = S_n$$

A theory exists [13] which shows that Newton's method will converge under certain conditions, but it is difficult and time consuming to determine if these conditions are met. For our purpose it is sufficient to assume that the conditions are satisfied, and the method will converge. Computer programming will stop the process in occasional cases of non-convergence.

8. PROBLEMS

8.1 If a, b, c, d are scalars, show that

$$\begin{bmatrix} a & b \\ c & d \end{bmatrix}^{-1} = \frac{1}{ad - bc} \begin{bmatrix} d & -b \\ -c & a \end{bmatrix}$$

8.2 Consider the system of equations

$$y = x^2$$

$$y = x$$

Using initial conditions as given below, perform the first iteration toward a solution by Newton's method, i.e., find S_1.

Hint:

$$S_n = \begin{bmatrix} x_n \\ y_n \end{bmatrix}$$

$$\Phi_n = \begin{bmatrix} x_n^2 - y_n \\ x_n - y_n \end{bmatrix}$$

$$S_1 = S_0 - \left(\frac{\partial \Phi}{\partial S_0}\right)^{-1} \Phi_0 \qquad (7.12)$$

$$\left.\begin{aligned} &(a)\quad S_0^T = \left[\frac{1}{4}, 1\right] \\ &(b)\quad S_0^T = \left[\frac{3}{4}, 1\right] \\ &(c)\quad S_0^T = \left[\frac{1}{2}, 1\right] \end{aligned}\right\}$$

In each case find S_1. What do you conclude from the results?

9. FURTHER PROPERTIES OF SYMMETRIC MATRICES

Some properties of positive definite and semi-definite matrices are discussed. The proofs of the statements are not difficult and they are available in standard tests.

Let

$\overline{x}$ be a vector (n x 1)

Γ a symmetric matrix (n x n)

$\varphi = \overline{x}^T \Gamma \overline{x}$

Then if $\varphi > 0$ for all $\overline{x} \neq 0$, Γ is said to be positive definite, written

9.1 $\Gamma > \emptyset$

If $\varphi \geq 0$ for all $\overline{x} \neq 0$, Γ is said to be positive semi-definite, written

9.2 $\Gamma \geq \emptyset$

Then

9.3 $\Gamma > \emptyset \longrightarrow \Gamma \geq \emptyset$

Also it is true that

9.4 $\Gamma > \emptyset \longrightarrow |\Gamma| > 0$

and

9.5 $\Gamma > \emptyset \longrightarrow \Gamma^{-1}$ exists.

Let λ_i be an eigenvalue of Γ.

9.6 $\Gamma > \emptyset \longrightarrow \lambda_i > 0$

9.7 $\Gamma \geq \emptyset \longrightarrow \lambda_i \geq 0$

Also

9.7A $\Gamma > \emptyset \longleftrightarrow \Gamma^{-1} > \emptyset$

Let

A be (n x p) of rank r, and

$\Omega > \emptyset$ (n x n) and symmetric.

Then the following are true:

9.8 $A^T A \geq \emptyset \qquad (n < p)$

9.9 $A^T A \geq \emptyset \qquad (r < p \leq n)$

9.10 $A^T A > \emptyset \qquad (r = p \leq n)$

9.11 $A^T \Omega A \geq \emptyset \qquad (n < p)$

9.12 $A^T \Omega A \geq \emptyset \qquad (r < p \leq n)$

9.13 $A^T \Omega A > \emptyset \qquad (r = p \leq n)$

All of these (9.8 - 13) are symmetric. Equation 3.27 is an example of 9.8. In derivations which follow it is necessary to compute forms such as $\left(A^T \Omega A\right)^{-1}$ and also to be assured that $A^T \Omega A > \emptyset$. Line 9.13 shows that the necessary and sufficient condition is $(r = p \leq n)$.

9.14 Problem

Let

$$A = \begin{bmatrix} 1 & 2 \\ 2 & 1 \\ 1 & 2 \end{bmatrix} \quad , \; B = \begin{bmatrix} 1 & 2 \\ 2 & 4 \end{bmatrix}$$

$$\Omega = \begin{bmatrix} 1 & -1 & 0 \\ -1 & 2 & -1 \\ 0 & -1 & 2 \end{bmatrix} \quad , \Gamma = \begin{bmatrix} 1 & -1 \\ -1 & 2 \end{bmatrix}$$

(a) Compute $|\Omega|$, $|\Gamma|$, $|B|$.

(b) From the answers to (a) comment on the existence of Ω^{-1}, Γ^{-1}, B^{-1}.

(c) Compute $\left[A\Gamma A^T\right]^{-1}$ if it exists. Classify it according to 9.11, 9.12, or 9.13.

(d) Compute $\left[B\Gamma B^T\right]^{-1}$ if it exists. Classify as in (c).

(e) Compute $\left[A^T\Omega A\right]^{-1}$ if it exists. Classify as in (c).

10. MINIMIZATION OF A QUADRATIC FORM AND SOLUTION BY NEWTON'S METHOD

This method will be used later on in deriving the Bayes filter.

Let

10.1 $S^T = [x_1, x_2, \dots, x_p]$

10.2 $\alpha^T = [\alpha_1, \alpha_2, \dots, \alpha_n]$

10.3 $\alpha = \alpha(S)$, i.e.,

10.4 $\alpha_i = \alpha_i(x_1, x_2, \dots, x_p)$

10.5 $R > 0$ and a symmetric matrix of constants (n x n).

Then from 10.5 it follows that

10.6 $R^{-1} > 0$ and symmetric.

Consider the quadratic form,

10.7 $2\varphi = \alpha^T R^{-1} \alpha$,

where

10.8 $\varphi = \varphi(S)$.

Find $\hat{S}$, the value of S such that the scalar φ is a minimum,

10.9 $\varphi_{min} = \varphi(\hat{S})$

We use the classical method.

Let S_1 be the solution of $\frac{\partial \varphi}{\partial S} = 0$. Then $\varphi(S_1)$ is an extremum. If in addition $\frac{\partial^2 \varphi}{\partial S^2} > 0$, then $\varphi(S_1)$ is a minimum and $S_1 = \hat{S}$.

10.10 $\Phi \equiv \frac{\partial \varphi}{\partial S}^T = \frac{\partial \alpha}{\partial S}^T R^{-1} \alpha$

10.11 $\Phi = \Phi(S)$

and

10.12 $\hat{\Phi} = \Phi(\hat{S}) = 0$

The solution to 10.12 will render φ an extremum. Disregard second order partials in taking the second derivative:

10.13 $\frac{\partial \Phi}{\partial S} = \frac{\partial \alpha}{\partial S}^T R^{-1} \frac{\partial \alpha}{\partial S}$

Assume that $\frac{\partial \alpha}{\partial S}$ is $(n \times p)$ of rank r and $(r = p \leq n)$. Then by 9.13 $\frac{\partial \Phi}{\partial S} > 0$, assuring that the extremum is a minimum and $\left(\frac{\partial \Phi}{\partial S}\right)^{-1}$ exists.

Assume that 10.12 is a system of non-linear equations and $\tilde{S}$ is a close first approximation to the solution, $\hat{S}$. Then by 7.11

10.14 $\hat{S} = \tilde{S} - \left(\frac{\partial \Phi}{\partial \tilde{S}}\right)^{-1} \tilde{\Phi}$

10.15 $\hat{S} = \tilde{S} - \left[\frac{\partial \alpha}{\partial \tilde{S}}^T R^{-1} \frac{\partial \alpha}{\partial \tilde{S}}\right]^{-1} \frac{\partial \alpha}{\partial \tilde{S}}^T R^{-1} \tilde{\alpha}$

where

$\tilde{\alpha} = \tilde{\alpha}(S)$,

or iteratively as in 7.12

10.16 $$S_{n+1} = S_n - \left[\frac{\partial \alpha}{\partial S_n}^T R^{-1} \frac{\partial \alpha}{\partial S_n} \right]^{-1} \frac{\partial \alpha}{\partial S_n}^T R^{-1} \alpha_n$$

11. THE STATE TRANSITION MATRIX

Let $S_i^T = [x_i, y_i, z_i, \dot{x}_i, \dot{y}_i, \dot{z}_i]$ be the true value of the state vector at time t_i.

Let $\tilde{S}_i^T = \left[\tilde{x}_i, \tilde{y}_i, \tilde{z}_i, \tilde{\dot{x}}_i, \tilde{\dot{y}}_i, \tilde{\dot{z}}_i\right]$ be a close estimate of S_i^T such that $(\tilde{S} - S)_i$ is small and linear approximations are valid. It is also true that the state vector at time t_j is a function of the state vector at time t_i, written

11.1 $$S_j = S_j(S_i)$$

Then using a first-order Taylor series expansion as in 6.6

11.2 $$\tilde{S}_j = S_j + \frac{\partial S_j}{\partial S_i}(\tilde{S}_i - S_i)$$

or

11.3 $$(\tilde{S} - S)_j = \frac{\partial S_j}{\partial S_i}(\tilde{S} - S)_i$$

The derivative, $\frac{\partial S_j}{\partial S_i}$, is the transformation matrix which relates a small deviation in the state vector at time t_j to a small deviation in the state vector at time t_i. This is called the state transition matrix. In expanded notation, the state transition matrix relating the deviation vector at time t to time t_0 is written

11.4 $$\frac{\partial S}{\partial S_0} = \begin{bmatrix} \frac{\partial x}{\partial x_0} & \frac{\partial x}{\partial y_0} & \frac{\partial x}{\partial z_0} & \frac{\partial x}{\partial \dot{x}_0} & \frac{\partial x}{\partial \dot{y}_0} & \frac{\partial x}{\partial \dot{z}_0} \\ \frac{\partial y}{\partial x_0} & \frac{\partial y}{\partial y_0} & \frac{\partial y}{\partial z_0} & \frac{\partial y}{\partial \dot{x}_0} & \frac{\partial y}{\partial \dot{y}_0} & \frac{\partial y}{\partial \dot{z}_0} \\ \frac{\partial z}{\partial x_0} & \frac{\partial z}{\partial y_0} & \frac{\partial z}{\partial z_0} & \frac{\partial z}{\partial \dot{x}_0} & \frac{\partial z}{\partial \dot{y}_0} & \frac{\partial z}{\partial \dot{z}_0} \\ \frac{\partial \dot{x}}{\partial x_0} & \frac{\partial \dot{x}}{\partial y_0} & \frac{\partial \dot{x}}{\partial z_0} & \frac{\partial \dot{x}}{\partial \dot{x}_0} & \frac{\partial \dot{x}}{\partial \dot{y}_0} & \frac{\partial \dot{x}}{\partial \dot{z}_0} \\ \frac{\partial \dot{y}}{\partial x_0} & \frac{\partial \dot{y}}{\partial y_0} & \frac{\partial \dot{y}}{\partial z_0} & \frac{\partial \dot{y}}{\partial \dot{x}_0} & \frac{\partial \dot{y}}{\partial \dot{y}_0} & \frac{\partial \dot{y}}{\partial \dot{z}_0} \\ \frac{\partial \dot{z}}{\partial x_0} & \frac{\partial \dot{z}}{\partial y_0} & \frac{\partial \dot{z}}{\partial z_0} & \frac{\partial \dot{z}}{\partial \dot{x}_0} & \frac{\partial \dot{z}}{\partial \dot{y}_0} & \frac{\partial \dot{z}}{\partial \dot{z}_0} \end{bmatrix}$$

This idea is readily extended to state vectors of any dimension.

11.5 Problem (A)

Given:

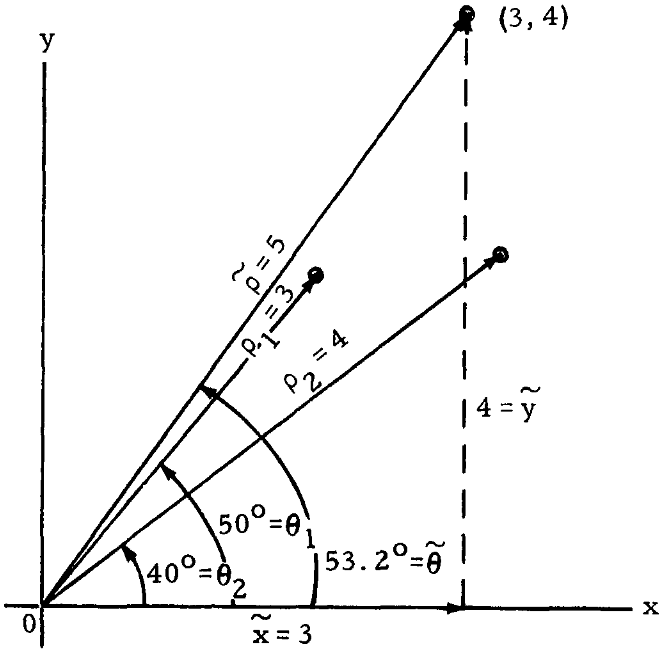
(a) An x-y cartesian frame

(b) Radar station at (0, 0)

(c) State vector = S = $\begin{bmatrix} x \\ y \end{bmatrix}$

(d) A priori estimate of the location of an object is

$$\tilde{S} = \begin{bmatrix} \tilde{x} \\ \tilde{y} \end{bmatrix} = \begin{bmatrix} 3 \\ 4 \end{bmatrix}$$



(e) $\alpha = \begin{bmatrix} \theta \\ \rho \end{bmatrix}$ is the true angle and range of the object, i.e., $\alpha = \alpha(S)$

(f) $\alpha_1 = \begin{bmatrix} 50^\circ \\ 3 \end{bmatrix}$ a radar observation

(g) $\alpha_2 = \begin{bmatrix} 40^\circ \\ 4 \end{bmatrix}$ another radar measurement

(h) $\tilde{\alpha} = \alpha(\tilde{S})$

Find:

(a) A better estimate of S.

Solution:

We shall do this by the method of least squares, i.e , we shall find the value of S which minimizes the sum of the squares of the residuals. Residuals are $(\theta_i - \theta)$ and $(\rho_i - \rho)$. Do one iteration only of Newton's method with $\tilde{S}$ as the first estimate. The sum of squares of residuals is written as a quadratic form:

$$2\varphi = \left[(\alpha_1 - \alpha)^T, (\alpha_2 - \alpha)^T\right] \begin{bmatrix} \alpha_1 - \alpha \\ \alpha_2 - \alpha \end{bmatrix}$$

$$= (\alpha_1 - \alpha)^T(\alpha_1 - \alpha) + (\alpha_2 - \alpha)^T(\alpha_2 - \alpha)$$

$$\Phi = \frac{\partial\varphi}{\partial S}^T = -\frac{\partial\alpha}{\partial S}^T [(\alpha_1 - \alpha) + (\alpha_2 - \alpha)]$$

$$\frac{\partial\Phi}{\partial S} = \frac{2\partial\alpha}{\partial S}^T \frac{\partial\alpha}{\partial S}$$

Then

$$\hat{S} = \tilde{S} - \left(\frac{\partial \Phi}{\partial \tilde{S}}\right)^{-1} \tilde{\Phi} \qquad (10.14)$$

$$\hat{S} = \tilde{S} + \left[2\frac{\partial \alpha}{\partial \tilde{S}}^{T} \frac{\partial \alpha}{\partial \tilde{S}}\right]^{-1} \frac{\partial \alpha}{\partial \tilde{S}}^{T} [\delta\alpha_1 + \delta\alpha_2]$$

Assume that $\left(\frac{\partial \alpha}{\partial \tilde{S}}\right)^{-1}$ exists, then

$$\hat{S} = \tilde{S} + \left(\frac{\partial \alpha}{\partial \tilde{S}}\right)^{-1} \left(\frac{\delta\alpha_1 + \delta\alpha_2}{2}\right)$$

* Now go ahead and compute the first iteration, i.e., compute $\hat{S} = \hat{S}(\tilde{S})$ where

$$\tilde{S} = \begin{bmatrix} 3 \\ 4 \end{bmatrix}$$

Problem (B)

Do the second iteration.

12. STATISTICAL THEORY

This section is prepared for those who need to understand trajectory estimation but lack a foundation in statistical theory. Such a scant treatment as this is only a shortcut to understanding the main subject and certainly not a substitute for formal study. For the previously uninitiated, statistical theory provides a new realm for mathematical imagination, where ideas may be beautiful and apparently simple, yet elusive. The student, however, should not be deluded by this apparent simplicity into dismissing the subject lightly as trivial. Tenacious pondering of the new notions must lead to feelings of frustration and inadequacy, followed by awareness and respect, and eventually appreciation and even astonishment — if he gets the right answer!

First consider a simple problem. Suppose we have three urns, each containing an infinite number of balls of different colors, assorted as follows:

I		II		III	
.1	blue	.2	red	.1	violet
.2	red	.2	yellow	.3	pink
.3	yellow	.6	blue	.5	red
.4	green			.1	black

Let the first letter of the color denote the color, i.e., B ⟷ blue, etc.

In each of the following selections one ball will be chosen at random from urn I, urn II, and urn III in that order.

P is the probability of making a selection.

Then

$$P(R, B, P) = (.2)(.6)(.3) = .036$$

$$P(R, R, R) = (.2)(.2)(.5) = .02$$

$$P(G, Y, V) = (.4)(.2)(.1) = .008$$

To generalize this idea consider a set of p urns, U_i, each containing an infinite number of named elements. One random sample, α_i, is taken from each U_i. And n_i is the decimal part of U_i which is named α_i. Then we have

$$\{U_i\} \qquad (i = 1, \ldots, p)$$

$$\alpha_i \varepsilon U_i$$

12.1 $$P(\alpha_1, \ldots, \alpha_p) = n_1 \cdots n_p$$

Suppose each n_i is a function of a set of parameters, S, and we took the sample $\{\alpha_i\}$ in order to find the most probable value of S.

$$n_i = n_i(S)$$

Then we would try to find the solution, $\hat{S}$, which would maximize $P(\alpha_1, \ldots, \alpha_p)$. This is the elementary principal which we use in processing radar measurements to get a better estimate of the state vector of a spacecraft.

So now we are just beginning to consider the problem of using radar measurements to get a better estimate of trajectory parameters. Let each measurement be modeled as a scalar function of the state vector. Later this will be extended to include vector functions, where several scalar measurements can be the elements of a measurement vector. Each measurement can be thought of as a random sample from an urn, one measurement only from each urn. In the example above we listed the assortment of colored balls in each urn. Analogous to this we need a way of listing the assortment of radar measurement values in each "urn". The assumption here is that the normal density function as shown below is a valid representation of the "assortment". A discussion of the normal density function for one random variable follows.

Let "urn" U be the set of elements (scalar measurements) represented by all values along the α- axis in Figure 12.1. Partition U according to a partition of the α- axis into short intervals such as δ. Let one value on δ, say α, be the label attached to every value on δ. Then α is the value assigned to every measurement represented by a point on δ. Let β be the mean value of all the elements of U. Finally, let the contents of U be distributed according to the normal density function, f(α) (12.2), where the cross-hatched area represents the decimal part of U labeled α.

12.2 $$f(\alpha) = \frac{1}{\sqrt{2\pi}\,\sigma_\alpha} \exp \left\{ -\frac{1}{2} \left(\frac{\alpha - \beta}{\sigma_\alpha} \right)^2 \right\}$$

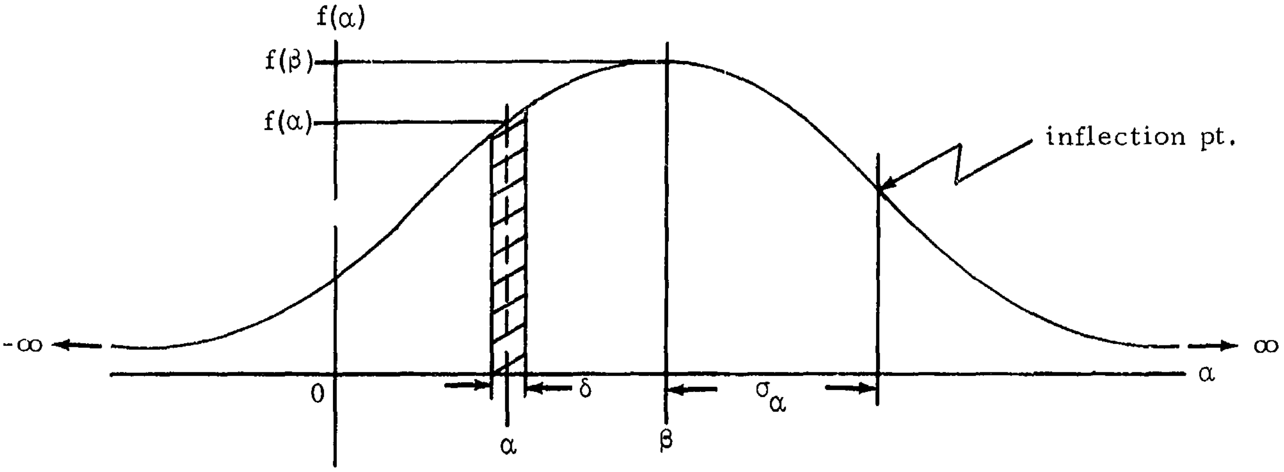


Figure 12. 1

Before discussing this curve further let us define the statistical expectation operator, E. If α is distributed in accordance with f(α), and g(α) is continuous almost everywhere on $-\infty < \alpha < \infty$, then

12.3 the mean value of $g(\alpha) \equiv E[g(\alpha)] = \int_{-\infty}^{\infty} g(\alpha)f\alpha)d\alpha$

Now return to Figure 12. l. The curve is symmetric. Using either the gamma function or a table of definite integrals it can be shown easily that

(a) $\int_{-\infty}^{\infty} f(\alpha)d\alpha = 1$.

(' $E(\alpha) = \int_{-\infty}^{\infty} \alpha f(\alpha)d\alpha = \beta$, where β is the mean value of α.

(c) $E(\beta) = \beta$

(d) $E(\alpha - \beta) = 0$

(e) $E[(\alpha - \beta)^2] = \sigma_\alpha^2$

where σ_α^2 is called the variance, and $\sigma_\alpha = \sqrt{\sigma_\alpha^2}$ is called the standard deviation.

(f) Approximately 2/3 of all the measurements in U have values on $\beta - \sigma_\alpha < \alpha < \beta + \sigma_\alpha$.

It is assumed that $\delta << \sigma_\alpha$. Note that δ arises from the limit of accuracy in reading the measuring instrument. For example, if we measured distance with a scale readable to the nearest tenth of a foot, we would have measurements 5.3, 5.4, 5.5, etc., but not 5.37. If the true measurement were 5.37 it would have the label 5.4. Thus in Figure 12.1 any measurement falling on δ should be labeled α. The cross-hatched area is the probability of choosing α, i.e.,

12.4 $$P(\alpha) = \int_{\alpha-\frac{1}{2}\delta}^{\alpha+\frac{1}{2}\delta} f(\alpha)d\alpha \approx f(\alpha)\delta$$

Note that the curve is completely determined by β and σ_α. The standard deviation σ_α, determines the shape (fat or thin), and the mean value, β, determines the position along the α-axis.

Suppose now that we have p independent measurements, $\{\alpha_i\}$ $(i = 1, \ldots, p)$, such that each measurement can be considered to be a sample from a separate "urn,"

$\alpha_i \epsilon U_i$

Then

12.5 $$f(\alpha_i) = \frac{1}{\sqrt{2\pi}\,\sigma_{\alpha_i}} \exp \left\{ -\frac{1}{2} \left(\frac{\alpha_i - \beta_i}{\sigma_{\alpha_i}} \right)^2 \right\} \qquad (12.2)$$

and

12.6 $$P(\alpha_i) \approx f(\alpha_i)\delta_i \qquad (12.4)$$

The joint probability is determined as in 12.1:

12.7 $$P(\alpha_1, \ldots, \alpha_p) \approx f(\alpha_1)\delta_1 \cdots f(\alpha_p)\delta_p = f(\alpha_1) \cdots f(\alpha_p)\delta_1 \cdots \delta_p$$

$$= \frac{1}{(2\pi)^{p/2}\sigma_{\alpha_1} \cdots \sigma_{\alpha_p}} \exp \left\{ -\frac{1}{2} \left[\left(\frac{\alpha_1 - \beta_1}{\sigma_{\alpha_1}} \right)^2 + \cdots + \left(\frac{\alpha_p - \beta_p}{\sigma_{\alpha_p}} \right)^2 \right] \right\} \delta_1 \cdots \delta_p$$

Define

12.8 $$f(\alpha_1, \ldots, \alpha_p) = f(\alpha_1) \cdots f(\alpha_p)$$

Since α_i and α_j are functionally independent $(i \neq j)$,

12.9 $$\int_{-\infty}^{\infty} \cdots \int f(\alpha_1, \ldots, \alpha_p) d\alpha_1 \cdots d\alpha_p = \int_{-\infty}^{\infty} f(\alpha_1) d\alpha_1 \cdots \int_{-\infty}^{\infty} f(\alpha_p) d\alpha_p = 1$$

Then 12.8 is the multivariate normal density function and

12.10 $$E[g(\alpha_1, \ldots, \alpha_p)] = \int_{-\infty}^{\infty} \cdots \int g(\alpha_1, \ldots, \alpha_p) f(\alpha_1, \ldots, \alpha_p) d\alpha_1 \cdots d\alpha_p$$

Again due to functional independence

12.11 $$E(\alpha_i) = \beta_i$$

12.12 $E[(\alpha_i - \beta_i)(\alpha_j - \beta_j)] = \begin{cases} \sigma_{\alpha_i}^2 & (i = j) \\ 0 & (i \neq j) \end{cases}$

Now we shall re-write equations 12.7 - 12.12 in matrix form. Define:

12.13 $\alpha^T = [\alpha_1, \ldots, \alpha_p]$

12.14 $R \equiv \begin{bmatrix} \sigma_{\alpha_1}^2 & & 0 \\ & \ddots & \\ 0 & & \sigma_{\alpha_p}^2 \end{bmatrix}$

Then

12.15 $P(\alpha) \approx \frac{1}{(2\pi)^{p/2}|R|^{1/2}} \exp\left\{-\frac{1}{2}(\alpha - \beta)^T R^{-1}(\alpha - \beta)\right\} \delta_1 \cdots \delta_p$ (12.7)

12.16 $f(\alpha) = \frac{1}{(2\pi)^{p/2}|R|^{1/2}} \exp\left\{-\frac{1}{2}(\alpha - \beta)^T R^{-1}(\alpha - \beta)\right\}$ (12.8)

12.17 $\int_{-\infty}^{\infty} \cdots \int f(\alpha) d\alpha_1 \cdots d\alpha_p = 1$ (12.9)

12.18 $E[g(\alpha)] = \int_{-\infty}^{\infty} \cdots \int g(\alpha) f(\alpha) d\alpha_1 \cdots d\alpha_p$ (12.10)

12.19 $E(\alpha) = \beta$ (12.11)

12.20 $E[(\alpha - \beta)(\alpha - \beta)^T] = R$ (12.12)

The covariance matrix, R, is still diagonal and errors in the measurements, α_i and α_j $(i \neq j)$, are said to be uncorrelated.

Consider a non-singular linear transformation, T, such that

12.21 $\alpha' = T\alpha$

Then

12.22 $R' = E[(\alpha' - \beta')(\alpha' - \beta')^T] = TE[(\alpha - \beta)(\alpha - \beta)^T]T^T = TRT^T$

The matrix, R', is non-diagonal (except for particular choices of T), and errors in the pseudomeasurements α_i' and α_j' are said to be correlated. We shall show that equations 12.15 - 12.20 can be expressed in the new coordinate system simply by inserting primes over the variables.

12.23 $$\begin{cases} \alpha \longleftrightarrow \alpha' \\ \beta \longleftrightarrow \beta' \\ R \longleftrightarrow R' \\ g(\alpha) \longleftrightarrow g'(\alpha') \\ \delta_i \longleftrightarrow \delta_i' \end{cases}$$

Define

12.24 $2\varphi = (\alpha - \beta)^T R^{-1}(\alpha - \beta)$

This quadratic form is invariant under the transformation, as follows:

12.25 $$\begin{aligned} 2\varphi &= (\alpha - \beta)^T R^{-1}(\alpha - \beta) \\ &= (\alpha - \beta)^T T^T T^{-T} R^{-1} T^{-1} T(\alpha - \beta) \\ &= (\alpha' - \beta')^T R'^{-1}(\alpha' - \beta') && (12.21) \\ &= 2\varphi' \end{aligned}$$

The normal density function transforms as

12.26 $$\begin{aligned} f(\alpha) &= \frac{1}{(2\pi)^{p/2}|R|^{1/2}} e^{-\varphi} \\ &= \frac{|T|}{(2\pi)^{p/2}|R'|^{1/2}} e^{-\varphi} && (12.22, 12.25) \\ &= |T| f(\alpha') \end{aligned}$$

The differential hyper-volume of the definite integral transforms as

12.27 $$d\alpha_1 \cdots d\alpha_p = \left|\frac{\partial \alpha}{\partial \alpha'}\right| d\alpha_1' \cdots d\alpha_p' = \frac{d\alpha_1' \cdots d\alpha_p'}{|T|}$$ [8]

Combining 12.26 and 12.27 gives

12.28 $$f(\alpha) d\alpha_1 \cdots d\alpha_p = f(\alpha') d\alpha_1' \cdots d\alpha_p'$$

Using 12.28 it can be shown that equations 12.15 - 12.20 are expressed in the new coordinate system simply by mapping the variables as in 12.23. Then $f(\alpha')$ is the multivariate normal density function for variables with correlated errors and $P(\alpha')$ is the probability of selecting the random vector, α'.

From here on measurement errors are considered uncorrelated; so the measurement covariance matrix is diagonal. One exception is correlated doppler measurement errors to be discussed later.

We have shown that the normal density function for p measurements with correlated errors is

12.29 $$f(\alpha') = \frac{1}{(2\pi)^{p/2} |R'|^{1/2}} \exp \left\{ -\frac{1}{2} (\alpha' - \beta')^T R'^{-1} (\alpha' - \beta') \right\}$$

Now we wish to express the normal density function for n trajectory parameters with correlated errors. Let

$$S = \begin{bmatrix} x_1 \\ \vdots \\ x_n \end{bmatrix}$$ the mean (true) value of an n-parameter state vector

$$\tilde{S} = \begin{bmatrix} \tilde{x}_1 \\ \vdots \\ \tilde{x}_n \end{bmatrix}$$ an estimate of S

12.30 $$\tilde{\Gamma} = E[(\tilde{S} - S)(\tilde{S} - S)^T]$$ the state covariance matrix

If n = 1, the normal density function is

12.31 $$f(\tilde{x}) = \frac{1}{\sqrt{2\pi}\,\sigma_{\tilde{x}}} \exp\left\{-\frac{1}{2}\left(\frac{\tilde{x} - x}{\sigma_{\tilde{x}}}\right)^2\right\} \qquad (12.2)$$

Starting with 12.31 and repeating the procedure which led to 12.29, the multivariate normal density function for the state vector is

12.32 $$f(\tilde{S}) = \frac{1}{(2\pi)^{n/2}|\tilde{\Gamma}|^{1/2}} \exp\left\{-\frac{1}{2}(\tilde{S} - S)^T\tilde{\Gamma}^{-1}(\tilde{S} - S)\right\}$$

Note that the state covariance matrix will seldom, if ever, be diagonal. It can be diagonalized, but this is time consuming for large order matrices and is not done. One thing more: For the purpose of deriving 12.32 we should consider that the transformation, T, (12.21) was orthogonal ($TT^T = I$). Then the elements of S will be functionally independent. This results in simpler mathematical formulations. To emphasize this remember that the elements of α were assumed to be functionally independent, but the elements of $\alpha' = T\alpha$ are not functionally independent unless $T^{-1} = T^T$. Notation for the elements of 12.30 is

12.33 $$\tilde{\Gamma} = \begin{bmatrix} \sigma^2_{\tilde{x}_1} & \sigma_{\tilde{x}_1\tilde{x}_2} & \cdots & \sigma_{\tilde{x}_1\tilde{x}_p} \\ \sigma_{\tilde{x}_1\tilde{x}_2} & \sigma^2_{\tilde{x}_2} & & \vdots \\ \vdots & & \ddots & \\ \sigma_{\tilde{x}_1\tilde{x}_p} & \cdots & & \sigma^2_{\tilde{x}_p} \end{bmatrix}$$

The variances, $\sigma^2_{\tilde{x}_i}$, may be expressed

12.34 $$\sigma^2_{\tilde{x}_i} \equiv \sigma_{\tilde{x}_i\tilde{x}_i}$$

The elements $\sigma_{\tilde{x}_i\tilde{x}_j}$ $(i \neq j)$ are called covariances. Rewrite 12.32

12.35 $$f(\tilde{S}) = \frac{1}{(2\pi)^{p/2}|\tilde{\Gamma}|^{1/2}} e^{-\varphi}$$

where

12.36 $$2\varphi = (\tilde{S} - S)^T\tilde{\Gamma}^{-1}(\tilde{S} - S)$$

Abstractly, this (12.36) is the equation of a hyper-ellipsoid with p principal axes. If $\tilde{\Gamma}$ is diagonal, then the principal axes are aligned with the coordinate axes, and the errors in the trajectory parameters are uncorrelated.

Now we are finally at the point where we can process a set of radar measurements to get a better estimate of the state vector. Let

12.37 $$2\varphi = (\alpha - \beta)^T R^{-1}(\alpha - \beta)$$

and rewrite 12.15

12.38 $$P(\alpha) \approx \frac{1}{(2\pi)^{p/2}|R|^{1/2}} e^{-\varphi}\delta_1 \cdots \delta_p$$

The a priori estimate of the state vector is $\tilde{S}$. The measurement vector is α. We need to find the value of S which will make $P(\alpha)$ a maximum. All terms in $P(\alpha)$ are constants except $\beta = \beta(S)$. Obviously, $P(\alpha)$ is a maximum when 2φ is a minimum. So to get a better estimate of S, we find the value $\hat{S}$ which minimizes 2φ. Review our thinking a moment. We can never know the true value of the state vector; so our best assumption is that the true value equals the mean value, S. Our current estimate of S is $\tilde{S}$. Our better estimate will be $\hat{S}$. Now find $\hat{S}$. (See Section 10.)

12.39 $$2\varphi = (\alpha - \beta)^T R^{-1}(\alpha - \beta)$$

12.40 $$\Phi = \frac{\partial\varphi}{\partial S}^T = -\frac{\partial\beta}{\partial S}^T R^{-1}(\alpha - \beta)$$

12.41 $\frac{\partial\Phi}{\partial S} = {\frac{\partial\beta}{\partial S}}^T R^{-1} \frac{\partial\beta}{\partial S}$ (disregarding 2nd order partials)

12.42 $\hat{\Phi} = \Phi(\hat{S}) = 0$

12.43 $\tilde{\Phi} = \Phi(\tilde{S})$

12.44 $\hat{\Phi} = \tilde{\Phi} + \frac{\partial\Phi}{\partial\tilde{S}}(\hat{S} - \tilde{S}) = 0$

12.45 $\hat{S} = \tilde{S} - \left(\frac{\partial\Phi}{\partial\tilde{S}}\right)^{-1}\tilde{\Phi}$

12.46 $\hat{S} = \tilde{S} + \left[{\frac{\partial\beta}{\partial\tilde{S}}}^T R^{-1}\frac{\partial\beta}{\partial\tilde{S}}\right]^{-1}{\frac{\partial\beta}{\partial\tilde{S}}}^T R^{-1}(\alpha - \tilde{\beta})$

or iteratively

12.47 $S_{n+1} = S_n + \left[{\frac{\partial\beta}{\partial S_n}}^T R^{-1}\frac{\partial\beta}{\partial S_n}\right]^{-1}{\frac{\partial\beta}{\partial S_n}}^T R^{-1}(\alpha - \beta_n)$

Since R is diagonal we can write 12.47 as

12.48 $S_{n+1} = S_n + \left[\sum_{i=1}^{p}{\frac{\partial\beta_i}{\partial S_n}}^T R_i^{-1}\frac{\partial\beta_i}{\partial S_n}\right]^{-1}\sum_{i=1}^{p}{\frac{\partial\beta_i}{\partial S_n}}^T R_i^{-1}(\alpha_i - \beta_{in})$

where each α_i is a subvector of α and

12.49 $R_i = E\left[(\alpha_i - \beta_i)(\alpha_i - \beta_i)^T\right]$

This (12.48) is a convenient formulation to program, since the procedure is to measure a specified set of quantities at each time t_i. For example,

$$\alpha_i = \begin{bmatrix} \alpha_{i1} \\ \alpha_{i2} \\ \alpha_{i3} \\ \alpha_{i4} \end{bmatrix} = \begin{bmatrix} \text{range} \\ \text{azimuth} \\ \text{elevation} \\ \text{range-rate} \end{bmatrix}_{t_i}$$

as measured from a radar station at time t_i.

If we converge after n iterations, then consider $S_n = \hat{S}$. Now to find $\hat{\Gamma}$, express $\hat{S}$ as a function of S, using a first-order Taylor series as in 12.45

$$12.50 \quad \hat{S} = S - \left(\frac{\partial \Phi}{\partial S}\right)^{-1} \Phi$$

$$12.51 \quad \hat{\Gamma} = E\left[(\hat{S} - S)(\hat{S} - S)^T\right] = E\left\{\left(\frac{\partial \Phi}{\partial S}\right)^{-1} \Phi\Phi^T \left(\frac{\partial \Phi}{\partial S}\right)^{-1}\right\}$$

$$\hat{\Gamma} = \left(\frac{\partial \Phi}{\partial S}\right)^{-1} \frac{\partial \beta}{\partial S}^T R^{-1} E\left[(\alpha - \beta)(\alpha - \beta)^T\right] R^{-1} \frac{\partial \beta}{\partial S} \left(\frac{\partial \Phi}{\partial S}\right)^{-1}$$

which can be reduced by 12.20 and 12.41 to

$$12.52 \quad \hat{\Gamma} = \left(\frac{\partial \Phi}{\partial S}\right)^{-1} = \left[\frac{\partial \beta}{\partial S}^T R^{-1} \frac{\partial \beta}{\partial S}\right]^{-1}$$

Then 12.52 is the new estimate of the state covariance matrix, computed as a function of $\hat{S}$.

13. SEQUENTIAL ESTIMATION — THE BAYES FILTER

So far we showed how to process a set of radar measurements to get a better estimate of the state vector, and we found the state covariance matrix associated with this estimate. This can be extended to fit the real situation where batches of measurements are processed sequentially to estimate a state vector changing with time. First review the propagation of small deviations of the state vector as in section 11.

Define

$(\hat{S} - S)_i$ the error in the best estimate at t_i

$(\tilde{S} - S)_j$ the error in the apriori estimate at t_j

$$\hat{\Gamma}_i = E\left[(\hat{S} - S)(\hat{S} - S)^T\right]_i$$

$$\tilde{\Gamma}_j = E\left[(\tilde{S} - S)(\tilde{S} - S)^T\right]_j$$

Then $(t_i < t_j)$

13.1 $$(\tilde{S} - S)_j = \frac{\partial S_j}{\partial S_i}(\hat{S} - S)_i$$ and

13.2 $$\tilde{\Gamma}_j = \frac{\partial S_j}{\partial S_i}\hat{\Gamma}_i\frac{\partial S_j}{\partial S_i}^T \qquad (11.3)$$

That is, the best estimates at t_i are propagated to t_j, where they are a priori estimates. Let α^* be a measurement vector.

Substituting S for $\tilde{S}$, write 12.46:

13.3 $$(\hat{S} - S) = \left[\frac{\partial \beta^*}{\partial S}^T R^{*-1}\frac{\partial \beta^*}{\partial S}\right]^{-1}\frac{\partial \beta^*}{\partial S}^T R^{*-1}(\alpha - \beta)^*$$

This substitution is valid, because $\tilde{S}$ is any good first guess; and hopefully the true value, S, would be a good first guess (if not, we are in trouble).

Equation 13.3 is a linear approximation of the error in the state vector estimate at time, t, after processing measurement batch, α^*. Also

$$13.4 \qquad \hat{\Gamma} = \left[\frac{\partial\beta^{*T}}{\partial S} R^{*-1} \frac{\partial\beta^*}{\partial S}\right]^{-1} \qquad (12.52)$$

Partition α^* into two non-empty subvectors

$$13.5 \qquad \alpha^* = \begin{bmatrix}\alpha_1 \\ \alpha\end{bmatrix} \quad ; \quad \begin{matrix} E\left[(\alpha - \beta)(\alpha - \beta)^T\right]_1 = R_1 \; ; \\ E\left[(\alpha - \beta)(\alpha - \beta)^T\right] = R \end{matrix}$$

and

$$R^* = \begin{bmatrix} R_1 & 0 \\ 0 & R \end{bmatrix}$$

Choose a time $t_j < t$, which is an appropriate time to process α_1. Then (since R is diagonal):

$$13.6 \qquad (\hat{S} - S)_j = \left[\frac{\partial\beta_1^T}{\partial S_j} R_1^{-1} \frac{\partial\beta_1}{\partial S_j}\right]^{-1} \frac{\partial\beta_1^T}{\partial S_j} R_1^{-1} (\alpha - \beta)_1 \quad \text{and} \qquad (13.3)$$

$$13.7 \qquad \hat{\Gamma}_j = \left[\frac{\partial\beta_1^T}{\partial S_j} R_1^{-1} \frac{\partial\beta_1}{\partial S_j}\right]^{-1} \qquad (13.4)$$

Also 13.3 can be written

$$13.8 \qquad (\hat{S} - S) = \left[\frac{\partial\beta_1^T}{\partial S} R_1^{-1} \frac{\partial\beta_1}{\partial S} + \frac{\partial\beta^T}{\partial S} R \frac{\partial\beta}{\partial S}\right]^{-1} \left[\frac{\partial\beta_1^T}{\partial S} R_1^{-1} (\alpha - \beta)_1 + \frac{\partial\beta^T}{\partial S} R^{-1} (\alpha - \beta)\right] \qquad (13.5)$$

where

$$13.9 \quad \frac{\partial \beta_1}{\partial S}^T R_1^{-1} \frac{\partial \beta_1}{\partial S} = \frac{\partial S_j}{\partial S}^T \frac{\partial \beta_1}{\partial S_j}^T R_1^{-1} \frac{\partial \beta_1}{\partial S_j} \frac{\partial S_j}{\partial S} = \frac{\partial S_j}{\partial S}^T \hat{\Gamma}_j^{-1} \frac{\partial S_j}{\partial S} = \tilde{\Gamma}^{-1} \qquad (13.2,\ 13.7)$$

$$13.10 \quad \frac{\partial \beta_1}{\partial S}^T R_1^{-1} (\alpha - \beta)_1 = \frac{\partial S_j}{\partial S}^T \frac{\partial \beta_1}{\partial S_j}^T R_1^{-1} (\alpha - \beta)_1 = \frac{\partial S_j}{\partial S}^T \hat{\Gamma}_j^{-1} (\hat{S} - S)_j \qquad (13.6) .$$

$$= \frac{\partial S_j}{\partial S}^T \hat{\Gamma}_1^{-1} \frac{\partial S_j}{\partial S} \frac{\partial S}{\partial S_j} (\hat{S} - S)_j$$

$$= \tilde{\Gamma}^{-1} (\tilde{S} - S) \qquad (13.1,\ 13.2)$$

Substitute 13.9, 13.10 into 13.8:

$$13.11 \quad (\hat{S} - S) = \left[\tilde{\Gamma}^{-1} + \frac{\partial \beta}{\partial S}^T R^{-1} \frac{\partial \beta}{\partial S} \right]^{-1} \left[\tilde{\Gamma}^{-1} (\tilde{S} - S) + \frac{\partial \beta}{\partial S}^T R^{-1} (\alpha - \beta) \right] ,$$

where the a priori $\tilde{\Gamma}$ and $\tilde{S}$ come from processing past measurements and α is the next measurement vector to be processed. Note that 13.3 and 13.11 are equal (if first order approximations are valid), although 13.3 was obtained by processing α^* at t, and 13.11 is from processing α_1 at t_j and α at t. This can be extended by induction to show that the final $(\hat{S} - S)$ (after processing all of α^*) is independent of the batching partition and times of processing. This idea is emphasized by an algebraic proof in the final section. Since R is diagonal, 13.11 can be written

$$13.12 \quad (\hat{S} - S) = \left[\tilde{\Gamma}^{-1} + \sum_i \frac{\partial \beta_i}{\partial S}^T R_i^{-1} \frac{\partial \beta_i}{\partial S} \right]^{-1} \left[\tilde{\Gamma}^{-1} (\tilde{S} - S) + \sum_i \frac{\partial \beta_i}{\partial S}^T R_i^{-1} (\alpha - \beta)_i \right] \qquad (12.48)$$

where each α_i is a subvector of α such that all elements of α_i were measured at t_i. This is the form of the Bayes sequential filter used by the RTCC, MSC for Apollo trajectory determination.

Also

$$13.13 \quad \hat{\Gamma} = \left[\tilde{\Gamma}^{-1} + \sum_i \frac{\partial \beta_i}{\partial S}^T R_i^{-1} \frac{\partial \beta_i}{\partial S} \right]^{-1} \qquad (13.4,\ 13.12)$$

which is the error matrix associated with the estimate in 13.12.

14. FORMULATION OF MEASUREMENTS

Consider a vector, $[x \quad y \quad z]^T$, expressed in a right-hand, rectangular frame. If this frame is rotated positively through angle θ about the x-axis, then the same vector is expressed as $[x' \quad y' \quad z']^T$ in the rotated frame.

14.1 $$\begin{bmatrix} x' \\ y' \\ z' \end{bmatrix} = \begin{bmatrix} 1 & 0 & 0 \\ 0 & \cos\theta & \sin\theta \\ 0 & -\sin\theta & \cos\theta \end{bmatrix} \begin{bmatrix} x \\ y \\ z \end{bmatrix}$$

If the positive rotation through angle θ is about the y-axis, then

14.2 $$\begin{bmatrix} x' \\ y' \\ z' \end{bmatrix} = \begin{bmatrix} \cos\theta & 0 & -\sin\theta \\ 0 & 1 & 0 \\ \sin\theta & 0 & \cos\theta \end{bmatrix} \begin{bmatrix} x \\ y \\ z \end{bmatrix}$$

If the rotation is about the z-axis, then

14.3 $$\begin{bmatrix} x' \\ y' \\ z' \end{bmatrix} = \begin{bmatrix} \cos\theta & \sin\theta & 0 \\ -\sin\theta & \cos\theta & 0 \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} x \\ y \\ z \end{bmatrix}$$

Now we can formulate some representative radar measurements. Some of these measurements are now in the Apollo Trajectory Estimation (ATE) program, while others are just good possibilities for future programs. Also, some of the fine points of the formulations are omitted.

Azimuth and elevation measurements are expressed in a topocentric, (x', y', z'), frame centered at the radar station. The x'-axis points east; the y'-axis, north; the z'-axis, to the zenith.

14.4 $\bar{r} = \begin{bmatrix} x \\ y \\ z \end{bmatrix}$ geocentric (inertial) position of the spacecraft. [10]

14.5 $\bar{r}_s = \begin{bmatrix} x_s \\ y_s \\ z_s \end{bmatrix}$ geocentric position of the radar station [10]

14.6 $\bar{r}' = \begin{bmatrix} x' \\ y' \\ z' \end{bmatrix}$ topocentric position of the spacecraft [10]

14.7 φ latitude of radar station

14.8 θ right ascension of radar station meridian

14.9 ρ range of spacecraft from radar station

The position of the spacecraft in the topocentric frame is

14.10 $\bar{r}' = T(\bar{r} - \bar{r}_s)$ [10]

where

14.11 $$T = \begin{bmatrix} 1 & 0 & 0 \\ 0 & \cos(90-\varphi) & \sin(90-\varphi) \\ 0 & -\sin(90-\varphi) & \cos(90-\varphi) \end{bmatrix} \begin{bmatrix} \cos(\theta+90) & \sin(\theta+90) & 0 \\ -\sin(\theta+90) & \cos(\theta+90) & 0 \\ 0 & 0 & 1 \end{bmatrix}$$

14.12 $$T = \begin{bmatrix} -\sin\theta & \cos\theta & 0 \\ -\sin\varphi\cos\theta & -\sin\varphi\sin\theta & \cos\varphi \\ \cos\varphi\cos\theta & \cos\varphi\sin\theta & \sin\varphi \end{bmatrix}$$ [10]

The azimuth measurement, A. is

14.13 $A = \tan^{-1}\left(\frac{x'}{y'}\right)$ [10]

The elevation measurement, E, is

14.14 $$E = \tan^{-1}\left(\frac{z'}{\sqrt{\rho^2 - z'^2}}\right) = \tan^{-1}\left(\frac{z'}{\sqrt{x'^2 + y'^2}}\right)$$ [10]

The range measurement, ρ, is

14.15 $$\rho = \left[(x - x_s)^2 + (y - y_s)^2 + (z - z_s)^2\right]^{1/2}$$ [10]

Now consider some measurements taken from the LM on the surface of the moon to the CSM in orbit.

Define

14.16 $$\bar{r}_G = \begin{bmatrix} x_G \\ y_G \\ z_G \end{bmatrix} = \begin{bmatrix} r \cos \varphi \cos \lambda \\ r \cos \varphi \sin \lambda \\ r \sin \varphi \end{bmatrix}$$ selenographic position of LM on the moon [11]

14.17 $$\bar{r}_L = \begin{bmatrix} x_L \\ y_L \\ z_L \end{bmatrix}$$ selenocentric (inertial) position of the LM on the moon [11]

14.18 L libration matrix, such that [11]

14.19 $$\bar{r}_L = L^T \bar{r}_G \qquad (L^T L = I)$$ [11]

14.20 $$\bar{r}_C = \begin{bmatrix} x_C \\ y_C \\ z_C \end{bmatrix}$$ selenocentric position of CSM in orbit [11]

14.21 $\bar{r}_{CL} = \bar{\rho} = \bar{r}_C - \bar{r}_L = \begin{bmatrix} x_{CL} \\ y_{CL} \\ z_{CL} \end{bmatrix}$ [11]

Then the following three measurements are from the LM to the CSM.

The pseudomeasurement D in [11] is

14.22 $D = \sin^{-1}\left(\frac{z_{CL}}{|\rho|}\right)$ [11]

The pseudomeasurement HA in [11] is

14.23 $HA = \tan^{-1}\left(\frac{y_{CL}}{x_{CL}}\right)$ [11]

The range measurement is

14.24 $|\rho| = \sqrt{\bar{\rho}^T \bar{\rho}}$

Now consider some measurements between the CSM and LM when they are both in orbit.

Define

14.25 $S^T = [x \ y \ z \ \dot{x} \ \dot{y} \ \dot{z}]$ inertial state vector of the CSM

14.26 $L^T = [x_L \ y_L \ z_L \ \dot{x}_L \ \dot{y}_L \ \dot{z}_L]$ inertial state vector of the LM (not the $\bar{r}_L$ and L of 14.17 and 14.18)

Then

14.27 $S = \begin{bmatrix} \bar{r} \\ \dot{\bar{r}} \end{bmatrix}$, $L = \begin{bmatrix} \bar{r}_L \\ \dot{\bar{r}}_L \end{bmatrix}$

14.28 $\bar{\rho} = \bar{r}_L - \bar{r}$

14.29 $\dot{\bar{\rho}} = \dot{\bar{r}}_L - \dot{\bar{r}}$

The range measurement is

14.30 $\rho = \sqrt{\bar{\rho}^T \bar{\rho}}$

The range-rate measurement is

14.31 $\frac{d\rho}{dt} = \frac{d}{dt}\left(\sqrt{\bar{\rho}^T \bar{\rho}}\right)$

The space-craft coordinate system centered in the CSM is as follows:

14.32 $\hat{e}_1 = \frac{\bar{r}}{|\bar{r}|}$

14.33 $\hat{e}_3 = \frac{\bar{r} \times \bar{v}}{|\bar{r} \times \bar{v}|}$

14.34 $\hat{e}_2 = \frac{(\bar{r} \times \bar{v}) \times \bar{r}}{|(\bar{r} \times \bar{v}) \times \bar{r}|}$

Then direction cosines from the CSM to the LM are

14.35 $\bar{\beta} = \begin{bmatrix} \beta_1 \\ \beta_2 \\ \beta_3 \end{bmatrix} = \begin{bmatrix} \hat{\rho}^T \hat{e}_1 \\ \hat{\rho}^T \hat{e}_2 \\ \hat{\rho}^T \hat{e}_3 \end{bmatrix}$

Now we shall formulate the doppler measurement as used in the ATE program. Another formulation will be given later when discussing the Kalman filter powered flight processor. See Figure 14.1. [10]

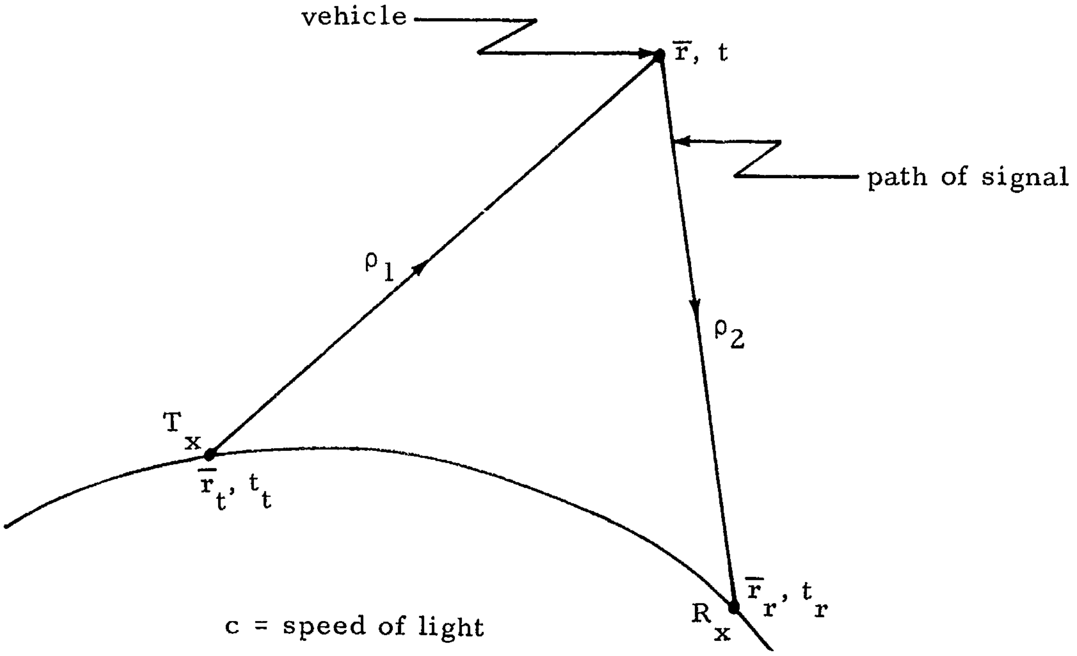


Figure 14.1

14.36 $t = t_r - \frac{\rho_2}{c}$

14.37 $\rho_2 = |\bar{r}(t) - \bar{r}_r(t_r)|$

Initialize with $t = t_r$ and then iterate using 14.36 and 14.37 to find ρ_2.

14.38 $t_t = t - \frac{\rho_1}{c}$

14.39 $\rho_1 = |\bar{r}(t) - \bar{r}_t(t_t)|$

Initialize with $t_t = t$ and then iterate using 14.38 and 14.39 to find ρ_1.

Define

c speed of light

τ counting interval

t_r is the doppler time tag (at the end of the counting interval) and is the time the signal is received at the ceiving station

f doppler frequency

$\omega_3 = 10^6$ Hertz $= 10^6$ cps, a bias constant

b a bias which can be estimated

ν the transmitting frequency

ω_4 a constant for signal adjustment

14.40 $\bar{\rho}_4 = \bar{r}\left(t_r - \frac{\rho_4}{c}\right) - \bar{r}_r(t_r)$

14.41 $\bar{\rho}_3 = \bar{r}\left(t_r - \frac{\rho_4}{c}\right) - \bar{r}_t\left(t_r - \frac{\rho_3 + \rho_4}{c}\right)$

14.42 $\bar{\rho}_2 = \bar{r}\left(t_r - \tau - \frac{\rho_2}{c}\right) - \bar{r}_r(t_r - \tau)$

14.43 $\bar{\rho}_1 = \bar{r}\left(t_r - \tau - \frac{\rho_2}{c}\right) - \bar{r}_t\left(t_r - \tau - \frac{\rho_1 + \rho_2}{c}\right)$

Then the computed measurement is

14.44 $f = (\omega_3 + b) + \frac{\omega_4 \nu}{c\tau}\left[(\rho_3 + \rho_4) - (\rho_1 + \rho_2)\right]$ [10]

Note that

$$\frac{\rho_3 - \rho_1}{\tau} \approx \dot{\rho}_3 \quad , \quad \frac{\rho_4 - \rho_2}{\tau} \approx \dot{\rho}_4$$

and

14.45 $$f = (\omega_3 + b) + \frac{\omega_4 \nu}{c} [\dot{\rho}_3 + \dot{\rho}_4]$$ [10]

It should be understood that a doppler measurement is not a discrete observation at a discrete time, but rather a counting process over a time interval τ. For mathematical convenience, however, we create an average frequency change over the counting interval, affix an average time, and treat this pseudomeasurement as a discrete observation. The pseudomeasurement corresponding to 14.44 is

14.46 $$f = \frac{K(t_r) - K(t_r - \tau)}{\tau}$$

where $K(t_r)$ is the doppler count at t_r. The average time t_* associated with f is the vehicle time for an imaginary signal received at the counting interval mid-time, $\left(t_r - \frac{\tau}{2}\right)$. Then

14.47 $$t_* = t_r - \frac{\tau}{2} - \frac{\rho_4^*}{c}$$

where

14.48 $$\bar{\rho}_4^* = \bar{r}\left(t_r - \frac{\tau}{2} - \frac{\rho_4^*}{c}\right) - \bar{r}_r\left(t_r - \frac{\tau}{2}\right)$$

15. PARTIAL DERIVATIVES OF MEASUREMENTS

Some of the following derivatives are used in the Apollo trajectory processor; others are just typical examples.

15.1 $$A = \tan^{-1}\left(\frac{x'}{y'}\right) \qquad (14.13)$$

Find $\frac{\partial A}{\partial S}$

$$\frac{\partial A}{\partial S} = \frac{\partial A}{\partial S'}\frac{\partial S'}{\partial S}$$

15.2 $$\frac{\partial A}{\partial S} = \frac{1}{1 + \left(\frac{x'}{y'}\right)^2}\left[\frac{1}{y'}\frac{\partial x'}{\partial S'} + x' \frac{\partial}{\partial S'}\left(\frac{1}{y'}\right)\right]\frac{\partial S'}{\partial S}$$

$$= \frac{y'^2}{\rho^2 - z'^2}\left\{\left[\frac{1}{y'}, 0, 0, 0, 0, 0\right] + \left[0, -\frac{x'}{y'^2}, 0, 0, 0, 0\right]\right\}\frac{\partial S'}{\partial S}$$

$$= \frac{-1}{\rho^2 - z'^2}[-y', x', 0, 0, 0, 0]\frac{\partial S'}{\partial S}$$

15.3 $$\frac{\partial S'}{\partial S} = \begin{bmatrix} \frac{\partial \bar{r}'}{\partial \bar{r}} & \frac{\partial \bar{r}'}{\partial \dot{\bar{r}}} \\ \frac{\partial \dot{\bar{r}}'}{\partial \bar{r}} & \frac{\partial \dot{\bar{r}}'}{\partial \dot{\bar{r}}} \end{bmatrix}$$

$$= \begin{bmatrix} -\sin\theta & \cos\theta & 0 & 0 & 0 & 0 \\ -\sin\varphi\cos\theta & -\sin\varphi\sin\theta & \cos\varphi & 0 & 0 & 0 \\ \cos\varphi\cos\theta & \cos\varphi\sin\theta & \sin\varphi & 0 & 0 & 0 \\ 0 & 0 & 0 & -\sin\varphi & \cos\theta & 0 \\ 0 & 0 & 0 & -\sin\varphi\cos\theta & -\sin\varphi\sin\theta & \cos\varphi \\ 0 & 0 & 0 & \cos\varphi\cos\theta & \cos\varphi\sin\theta & \sin\varphi \end{bmatrix}$$

15.4 $$\frac{\partial A}{\partial S} = \frac{-1}{\rho^2 - z'^2}\,[y' \sin\theta - x' \sin\varphi \cos\theta,\ -x' \sin\varphi \sin\theta - y' \cos\theta,\ x' \cos\varphi,\ 0,\ 0,\ 0]$$ [10]

15.5 $$E = \tan^{-1}\left(\frac{z'}{\sqrt{\rho^2 - z'^2}}\right)$$ (14.14)

Find $\frac{\partial E}{\partial S}$

15.6 $$\frac{\partial E}{\partial S} = \frac{1}{1 + \frac{z'^2}{\rho^2 - z'^2}}\left[\frac{1}{\sqrt{\rho^2 - z'^2}}\left(\frac{\partial z'}{\partial S'}\right) + z'\frac{\partial}{\partial S'}\left(\frac{1}{\sqrt{\rho^2 - z'^2}}\right)\right]\frac{\partial S'}{\partial S}$$

$$= \frac{\rho^2 - z'^2}{\rho^2}\left\{\frac{1}{\sqrt{\rho^2 - z'^2}}[0, 0, 1, 0, 0, 0] - \frac{z'}{2\left(\rho^2 - z'^2\right)^{3/2}}[2x', 2y', 0, 0, 0, 0]\right\}\frac{\partial S'}{\partial S}$$

$$= \frac{1}{\rho^2\sqrt{\rho^2 - z'^2}}\left\{\left(\rho^2 - z'^2\right)[0, 0, 1, 0, 0, 0] - z'[x', y', 0, 0, 0, 0]\right\}\frac{\partial S'}{\partial S}$$

$$= \frac{1}{\rho^2\sqrt{\rho^2 - z'^2}}\left\{\left[-x'z',\ -y'z',\ \rho^2 - z'^2,\ 0, 0, 0\right]\right\}\frac{\partial S'}{\partial S}$$

$$= \frac{1}{\rho^2\sqrt{\rho^2 - z'^2}}\Big[x'z' \sin\theta + y'z' \sin\varphi\cos\theta + \left(\rho^2 - z'^2\right)\cos\varphi\cos\theta,$$

$$-x'z' \cos\theta + y'z' \sin\varphi\sin\theta + \left(\rho^2 - z'^2\right)\cos\varphi\sin\theta,$$

$$-y'z' \cos\varphi + \left(\rho^2 - z'^2\right)\sin\varphi,\ 0, 0, 0\Big]$$

15.7 $$\frac{\partial E}{\partial S} = \frac{1}{\rho^2\sqrt{\rho^2 - z'^2}}\Big[-z'(x - x_s) + \rho^2 \cos\varphi\cos\theta,\ -z'(y - y_s) + \rho^2 \cos\varphi\sin\theta,$$

$$-z'(z - z_s) + \rho^2 \sin\varphi,\ 0, 0, 0\Big]$$ [10]

where

15.8 $$\begin{bmatrix} x - x_s \\ y - y_s \\ z - z_s \end{bmatrix} = \begin{bmatrix} -\sin\theta & -\sin\varphi\cos\theta & \cos\varphi\cos\theta \\ \cos\theta & -\sin\varphi\sin\theta & \cos\varphi\sin\theta \\ 0 & \cos\varphi & \sin\varphi \end{bmatrix} \begin{bmatrix} x' \\ y' \\ z' \end{bmatrix}$$ (14.10, 14.12)

Find $\frac{\partial \rho}{\partial S}$

15.9 $$\rho = \left[\left(\bar{r} - \bar{r}_s\right)^T\left(\bar{r} - \bar{r}_s\right)\right]^{1/2} = \sqrt{\bar{\rho}^T\bar{\rho}}$$ (14.15)

15.10 $$\frac{\partial \rho}{\partial S} = \frac{1}{2\rho} 2\bar{\rho}^T \frac{\partial \bar{\rho}}{\partial S} = \hat{\rho}^T[I, \emptyset] = \left[\hat{\rho}^T, \emptyset\right]$$

$$= \frac{1}{\rho}\left[x - x_s, y - y_s, z - z_s, 0, 0, 0\right]$$ [10]

15.11 $$D = \sin^{-1} \frac{z_{CL}}{\rho}$$ (14.22)

Find $\frac{\partial D}{\partial Q}$ where $Q = \begin{bmatrix} \varphi \\ \lambda \\ r \end{bmatrix}$

First find $\frac{\partial D}{\partial \bar{r}_L}$; then

15.12 $$\frac{\partial D}{\partial Q} = \frac{\partial D}{\partial \bar{r}_L} \frac{\partial \bar{r}_L}{\partial \bar{r}_G} \frac{\partial \bar{r}_G}{\partial Q}, \quad \text{where } \frac{\partial \bar{r}_L}{\partial \bar{r}_G} = L^T$$ (14.19)

15.13 $$\frac{\partial \bar{r}_G}{\partial Q} = \begin{bmatrix} -r \sin \varphi \cos \lambda & -r \cos \varphi \sin \lambda & \cos \varphi \cos \lambda \\ -r \sin \varphi \sin \lambda & r \cos \varphi \cos \lambda & \cos \varphi \sin \lambda \\ r \cos \varphi & 0 & \sin \varphi \end{bmatrix} \qquad (14.16)$$

15.14 $$\frac{\partial D}{\partial \bar{r}_L} = \frac{\rho}{\sqrt{\rho^2 - z_{CL}^2}} \left[z_{CL} \frac{\partial}{\partial \bar{r}_L}\left(\frac{1}{\rho}\right) + \frac{1}{\rho} \frac{\partial z_{CL}}{\partial \bar{r}_L} \right]$$

$$= \frac{\rho}{\sqrt{\rho^2 - z_{CL}^2}} \left[z_{CL}(-1)\left(\frac{1}{\rho^2}\right)\hat{\rho}^T \frac{\partial \bar{\rho}}{\partial \bar{r}_L} + \frac{1}{\rho}[0,\ 0,\ -1] \right]$$

$$= \frac{\rho}{\sqrt{\rho^2 - z_{CL}^2}} \left[\frac{z_{CL}}{\rho^2} \hat{\rho}^T + \frac{z_{CL}}{\rho^3} \left(0,\ 0,\ \frac{-\rho^2}{z_{CL}}\right) \right]$$

$$= \frac{z_{CL}}{\rho^2 \sqrt{\rho^2 - z_{CL}^2}} \left\{ \left[\bar{\rho}^T\right] - \left[0,\ 0,\ \frac{\rho^2}{z_{CL}}\right] \right\}$$

$$= \frac{z_{CL}}{\rho^2 \sqrt{\rho^2 - z_{CL}^2}} \left[x_{CL},\ y_{CL},\ \frac{z_{CL}^2 - \rho^2}{z_{CL}} \right] \qquad [11]$$

15.15 $$HA = \tan^{-1}\left(\frac{y_{CL}}{x_{CL}}\right) \qquad (14.23)$$

Find $\frac{\partial(HA)}{\partial Q}$

$$\frac{\partial(HA)}{\partial \bar{r}_L} = \frac{x_{CL}^2}{x_{CL}^2 + y_{CL}^2}\left[y_{CL} \frac{\partial}{\partial \bar{r}_L}\left(\frac{1}{x_{CL}}\right) + \frac{1}{x_{CL}} \frac{\partial y_{CL}}{\partial \bar{r}_L} \right]$$

$$= \frac{x_{CL}^2}{\rho^2 - z_{CL}^2}\left[y_{CL}(-1)\frac{1}{x_{CL}^2}[-1, 0, 0] + \frac{1}{x_{CL}}[0, -1, 0] \right]$$

$$= \frac{1}{\rho^2 - z_{CL}^2}\left[y_{CL}, -x_{CL}, 0 \right] \qquad [11]$$

Then

15.16 $$\frac{\partial(HA)}{\partial Q} = \frac{\partial(HA)}{\partial \bar{r}_L} \frac{\partial \bar{r}_L}{\partial \bar{r}_G} \frac{\partial \bar{r}_G}{\partial Q} \qquad (15.12)$$

Find Derivatives of Relative Measurements

15.17 $$S = \begin{bmatrix} \bar{r} \\ \dot{\bar{r}} \end{bmatrix} \quad ; \quad L = \begin{bmatrix} \bar{r}_L \\ \dot{\bar{r}}_L \end{bmatrix} \qquad (14.27)$$

15.18 $$\begin{bmatrix} \bar{\rho} \\ \dot{\bar{\rho}} \end{bmatrix} = \begin{bmatrix} \bar{r}_L - \bar{r} \\ \dot{\bar{r}}_L - \dot{\bar{r}} \end{bmatrix} \qquad \begin{matrix} (14.28) \\ (14.29) \end{matrix}$$

15.19 $$\frac{\partial \rho}{\partial S} = \frac{\partial}{\partial S}\left(\sqrt{\bar{\rho}^T \bar{\rho}}\right) = \frac{1}{2} \frac{1}{\sqrt{\bar{\rho}^T \bar{\rho}}} 2\bar{\rho}^T \frac{\partial \bar{\rho}}{\partial S} = \hat{\rho}^T \frac{\partial \bar{\rho}}{\partial S} \qquad (14.30)$$

$$= \hat{\rho}^T[-I, 0] = \left[-\hat{\rho}^T, 0\right]$$

15.20 $$\frac{\partial \rho}{\partial L} = \hat{\rho}^T \frac{\partial \bar{\rho}}{\partial L} = \hat{\rho}^T [I, 0] = \left[\hat{\rho}^T, 0\right]$$

15.21 $$\frac{d\rho}{dt} = \frac{d}{dt}\left(\sqrt{\bar{\rho}^T \bar{\rho}}\right) = \frac{1}{2} \frac{1}{\sqrt{\bar{\rho}^T \bar{\rho}}} 2 \bar{\rho}^T \dot{\bar{\rho}} = \hat{\rho}^T \dot{\bar{\rho}}$$ (14.31)

15.22 $$\frac{\partial}{\partial S}\left(\hat{\rho}^T \dot{\bar{\rho}}\right) = \hat{\rho}^T \frac{\partial \dot{\bar{\rho}}}{\partial S} + \dot{\bar{\rho}}^T \frac{\partial \hat{\rho}}{\partial S}$$

$$= \hat{\rho}^T [0, -I] + \dot{\bar{\rho}}^T \frac{1}{\rho}\left[I - \hat{\rho}\hat{\rho}^T\right] [-I, 0]$$

$$= \left[\frac{\dot{\bar{\rho}}^T}{\rho}\left(\hat{\rho}\hat{\rho}^T - I\right), -\hat{\rho}^T\right]$$

15.23 $$\frac{\partial}{\partial L}\left(\hat{\rho}^T \dot{\bar{\rho}}\right) = \hat{\rho} \frac{\partial \dot{\bar{\rho}}}{\partial L} + \dot{\bar{\rho}}^T \frac{\partial \hat{\rho}}{\partial L}$$

$$= \hat{\rho}^T [0, I] + \dot{\bar{\rho}}^T \frac{1}{\rho}\left[I - \hat{\rho}\hat{\rho}^T\right] [I, 0]$$

$$= \left[\frac{\dot{\bar{\rho}}^T}{\rho}\left(I - \hat{\rho}\hat{\rho}^T\right), \hat{\rho}^T\right]$$

15.24 $$\bar{\beta} = \begin{bmatrix} \beta_1 \\ \beta_2 \\ \beta_3 \end{bmatrix} = \begin{bmatrix} \hat{\rho}^T \hat{e}_1 \\ \hat{\rho}^T \hat{e}_2 \\ \hat{\rho}^T \hat{e}_3 \end{bmatrix}$$ (14.35)

Find $\frac{\partial \bar{\beta}}{\partial S}$ and $\frac{\partial \bar{\beta}}{\partial L}$

15.25 $$\frac{\partial \hat{e}_1}{\partial L} = \frac{\partial \hat{e}_2}{\partial L} = \frac{\partial \hat{e}_3}{\partial L} = 0$$ (14.32 - 34)

15.26 $$\frac{\partial \hat{e}_1}{\partial S} = \frac{\partial \hat{r}}{\partial S} = \frac{1}{r}\left[I - \hat{r}\hat{r}^T\right][I,\ 0]$$

$$\bar{e}_3 = \bar{r} \times \bar{v} = \overset{\approx}{r}\,\bar{v} = -\overset{\approx}{v}\,\bar{r}$$

$$\frac{\partial \bar{e}_3}{\partial S} = \overset{\approx}{r}\frac{\partial \bar{v}}{\partial S} - \overset{\approx}{v}\frac{\partial \bar{r}}{\partial S} = [-\overset{\approx}{v},\ \overset{\approx}{r}]$$

15.27 $$\frac{\partial \hat{e}_3}{\partial S} = \frac{1}{e_3}\left[I - \hat{e}_3\hat{e}_3^T\right][-\overset{\approx}{v},\ \overset{\approx}{r}]$$

$$\bar{e}_2 = (\bar{r} \times \bar{v}) \times \bar{r} = \overset{\approx}{r}\overset{\approx}{v}\bar{r} = -\overset{\approx}{r}\overset{\approx}{r}\bar{v} = (\overset{\approx}{r}\overset{\approx}{v} - \overset{\approx}{v}\overset{\approx}{r})\bar{r}$$

$$\frac{\partial \bar{e}_2}{\partial S} = \overset{\approx}{r}\overset{\approx}{v}[I,\ 0] - \overset{\approx}{r}\overset{\approx}{r}[0,\ I] + [\overset{\approx}{r}\overset{\approx}{v} - \overset{\approx}{v}\overset{\approx}{r}][I,\ 0]$$

$$= [2\overset{\approx}{r}\overset{\approx}{v} - \overset{\approx}{v}\overset{\approx}{r},\ -\overset{\approx}{r}\overset{\approx}{r}]$$

15.28 $$\frac{\partial \hat{e}_2}{\partial S} = \frac{1}{e_2}\left[I - \hat{e}_2\hat{e}_2^T\right][2\overset{\approx}{r}\overset{\approx}{v} - \overset{\approx}{v}\overset{\approx}{r},\ -\overset{\approx}{r}\overset{\approx}{r}]$$

15.29 $$\frac{\partial \hat{\rho}}{\partial S} = \frac{1}{\rho}\left[I - \hat{\rho}\hat{\rho}^T\right][-I,\ 0]$$

15.30 $$\frac{\partial \hat{\rho}}{\partial L} = \frac{1}{\rho}\left[I - \hat{\rho}\hat{\rho}^T\right][I,\ 0]$$

Then

15.31 $$\frac{\partial\beta_1}{\partial S} = \frac{\partial\left(\hat{\rho}^T\hat{e}_1\right)}{\partial S} = \hat{\rho}^T\frac{\partial\hat{e}_1}{\partial S} + \hat{e}_1^T\frac{\partial\hat{\rho}}{\partial S}$$

$$= \frac{\hat{\rho}^T}{r}\left[I - \hat{r}\hat{r}^T\right][I, \emptyset] + \frac{\hat{e}_1^T}{\rho}\left[I - \hat{\rho}\hat{\rho}^T\right][-I, \emptyset]$$

$$= \left[\frac{\hat{\rho}^T}{r}\left(I - \hat{r}\hat{r}^T\right) - \frac{\hat{e}_1^T}{\rho}\left(I - \hat{\rho}\hat{\rho}^T\right), \emptyset\right]$$ (15.26, 15.29)

15.32 $$\frac{\partial\beta_1}{\partial L} = \frac{\partial\left(\hat{\rho}^T\hat{e}_1\right)}{\partial L} = \hat{\rho}^T\frac{\partial\hat{e}_1}{\partial L} + \hat{e}_1^T\frac{\partial\hat{\rho}}{\partial L}$$

$$= \left[\frac{\hat{e}_1^T}{\rho}\left(I - \hat{\rho}\hat{\rho}^T\right), \emptyset\right]$$ (15.25, 15.30)

In the same manner:

15.33 $$\frac{\partial\beta_2}{\partial S} = \frac{\partial\left(\hat{\rho}^T\hat{e}_2\right)}{\partial S} = \hat{\rho}^T\frac{\partial\hat{e}_2}{\partial S} + \hat{e}_2^T\frac{\partial\hat{\rho}}{\partial S}$$

15.34 $$\frac{\partial\beta_2}{\partial L} = \frac{\partial\left(\hat{\rho}^T\hat{e}_2\right)}{\partial L} = \hat{e}_2^T\frac{\partial\hat{\rho}}{\partial L}$$

15.35 $$\frac{\partial\beta_3}{\partial S} = \frac{\partial\left(\hat{\rho}^T\hat{e}_3\right)}{\partial S} = \hat{\rho}^T\frac{\partial\hat{e}_3}{\partial S} + \hat{e}_3^T\frac{\partial\hat{\rho}}{\partial S}$$

15.36 $$\frac{\partial\beta_3}{\partial L} = \frac{\partial\left(\hat{\rho}^T\hat{e}_3\right)}{\partial L} = \hat{e}_3^T\frac{\partial\hat{\rho}}{\partial L}$$

15.37 $$f = (\omega_3 + b) + \frac{\omega_4\nu}{c}\left(\dot{\rho}_3 + \dot{\rho}_4\right)$$ (14.45)

15.38 $\frac{\partial f}{\partial b} = 1$ (Used only when b is adjoined to the state vector, in order to be estimated along with the trajectory parameters)

15.39 $$\frac{\partial f}{\partial S} = \frac{\omega_4 \nu}{c}\left[\frac{\partial \dot{\rho}_3}{\partial S} \quad \frac{\partial \dot{\rho}_4}{\partial S}\right]$$

As shown before:

$$\dot{\rho} = \frac{d}{dt}\sqrt{\bar{\rho}^T \bar{\rho}} = \hat{\rho}^T \dot{\bar{\rho}}$$

15.40 $$\frac{\partial}{\partial S}\hat{\rho}^T \dot{\bar{\rho}} = \hat{\rho}^T \frac{\partial \dot{\bar{\rho}}}{\partial S} + \dot{\bar{\rho}}^T \frac{\partial \hat{\rho}}{\partial S}$$

$$= \hat{\rho}^T [0 \ I] + \dot{\bar{\rho}} \frac{1}{\rho}\left[I - \hat{\rho}\hat{\rho}^T\right][I \ 0]$$

15.41 $$\frac{\partial \dot{\rho}}{\partial S} = \left[\frac{\dot{\bar{\rho}}^T}{\rho}\left(I - \hat{\rho}\hat{\rho}^T\right), \ \hat{\rho}^T\right] = \left[\frac{\dot{\bar{\rho}}^T}{\rho} - \frac{\dot{\rho}\bar{\rho}^T}{\rho^2}, \ \hat{\rho}^T\right]$$

Then

15.42 $$\frac{\partial f}{\partial S} = \frac{\omega_4 \nu}{c}\left[\frac{\dot{\bar{\rho}}_3^T}{\rho_3} - \frac{\dot{\rho}_3 \bar{\rho}_3^T}{\rho_3^2} + \frac{\dot{\bar{\rho}}_4^T}{\rho_4} - \frac{\dot{\rho}_4 \bar{\rho}_4^T}{\rho_4^2}, \ \hat{\rho}_3^T + \hat{\rho}_4^T\right]$$ (15.39, 15.41) [10]

To evaluate 15.42 $\bar{\rho}$ and $\dot{\bar{\rho}}$ are computed at time, t_* (14.48), by the same iterative method used in obtaining 14.40 and 14.41.

15.43 $$\bar{\rho}_4^* = \bar{r}\left(t_r - \frac{\tau}{2} - \frac{\rho_4^*}{c}\right) - \bar{r}_r\left(t_r - \frac{\tau}{2}\right)$$ (14.46) [10]

15.44 $$\dot{\bar{\rho}}_4^* = \dot{\bar{r}}\left(t_r - \frac{\tau}{2} - \frac{\rho_4^*}{c}\right) - \dot{\bar{r}}_r\left(t_r - \frac{\tau}{2}\right)$$ [10]

15.45 $$\overline{\rho}_3^* = \overline{r}\left(t_r - \frac{\tau}{2} - \frac{\rho_4^*}{c}\right) - \overline{r}_t\left(t_r - \frac{\tau}{2} - \frac{\rho_3^* + \rho_4^*}{c}\right)$$ [10]

15.46 $$\dot{\overline{\rho}}_3^* = \dot{\overline{r}}\left(t_r - \frac{\tau}{2} - \frac{\rho_4^*}{c}\right) - \dot{\overline{r}}_t\left(t_r - \frac{\tau}{2} - \frac{\rho_3^* + \rho_4^*}{c}\right)$$ [10]

where

$$t_r - \frac{\tau}{2} - \frac{\rho_4^*}{c} = t_* \qquad (14.47)$$

16. ESTIMATING THE TRAJECTORIES OF TWO SPACECRAFT SIMULTANEOUSLY, USING BOTH GROUND AND ONBOARD OBSERVATIONS

Early planners intended to estimate the Apollo trajectory by processing onboard observations along with those received from the sparsely-located and costly earth tracking stations. This was for two reasons: (1) It is possible for a spacecraft to complete several earth orbits out of sight of the tracking network. (2) The geometry at lunar distances precludes the successful use of earth-based measurements other than doppler, which by itself may not reliably determine a lunar trajectory. Sometime later, in order to minimize dependency on telemetry and to simplify computer programs, the decision was made to estimate the trajectories of the CSM and LM separately, using only earth-based radar data. Systems of the future, however, will probably rely more on onboard observations, and then such measurements between neighboring spacecraft may be used to adjust both trajectories simultaneously. This would be an accurate way to determine their relative state vector when far from the inertial origin. The mathematics for this is discussed for possible future use.

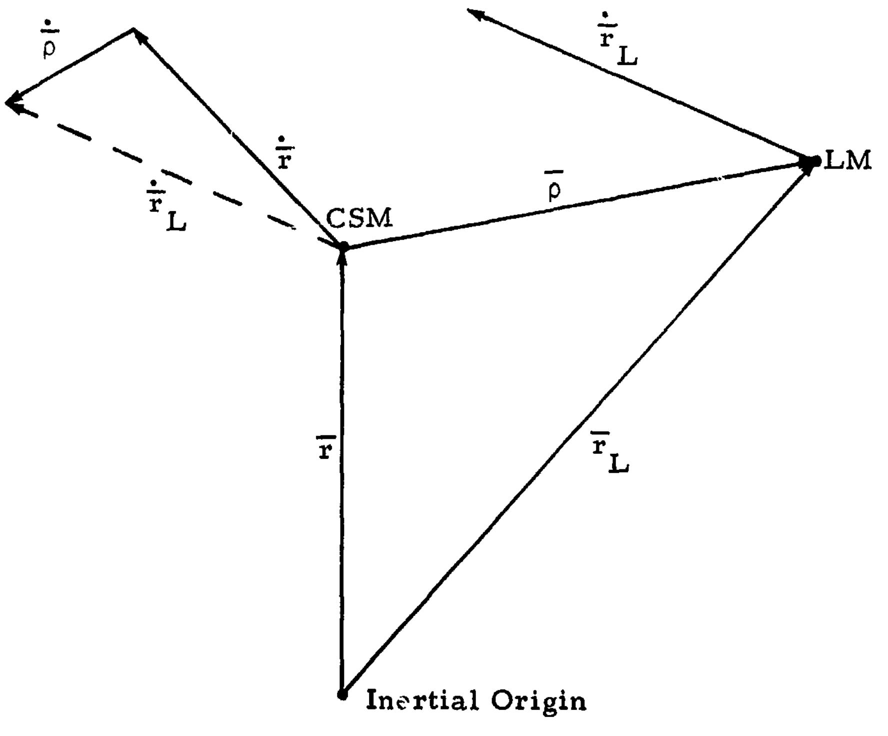


Figure 16.1

Define (See Figure 16.1):

16.1 $S = \begin{bmatrix} \bar{r} \\ \dot{\bar{r}} \end{bmatrix}$ CSM state vector

16.2 $L = \begin{bmatrix} \bar{r}_L \\ \dot{\bar{r}}_L \end{bmatrix}$ LM state vector

16.3 $B = \begin{bmatrix} \bar{\rho} \\ \dot{\bar{\rho}} \end{bmatrix}$ relative state vector

First consider that an estimate, $\hat{B}$, of the relative state vector is desired during a lunar rendezvous, and onboard observations between the spacecraft are available. In the Apollo program $\hat{S}$ and $\hat{L}$ are estimated separately by the following two equations.

16.4 $$(\hat{S} - S) = \left[\tilde{\Gamma}^{-1} + \frac{\partial \beta}{\partial S}^T R^{-1} \frac{\partial \beta}{\partial S} \right]^{-1} \left[\tilde{\Gamma}^{-1} (\tilde{S} - S) + \frac{\partial \beta}{\partial S}^T R^{-1} (\alpha - \beta) \right] \quad (13.11)$$

16.5 $$(\hat{L} - L) = \left[\tilde{\Gamma}_L^{-1} + \frac{\partial \beta}{\partial L}^T R^{-1} \frac{\partial \beta}{\partial L} \right]^{-1} \left[\tilde{\Gamma}_L^{-1} (\tilde{L} - L) + \frac{\partial \beta}{\partial L}^T R^{-1} (\alpha - \beta) \right] \quad (13.11)$$

If the CSM ephemeris is assumed well-known and the LM ephemeris uncertain (which is a real possibility), relative measurements between the two spacecraft could be used in 16.5 to find $\hat{L}$. Then

16.6 $\hat{B} = \hat{L} - \hat{S}$

During rendezvous, equation 16.6 requires the subtraction of very large, nearly-equal quantities, but this is handled accurately enough by the IBM 360 in double precision. This simple procedure gives adequate results in this case. We can conceive, however, that in the future situations may arise where the more general approach would be useful. That is, every measurement would be used to adjust the entire twelve-element state vector.

First choose the twelve functionally independent basis elements of the state vector; then all other elements in the space will be functions of these. A possible choice is $\left[S^T, L^T\right]$; then B would be a function of S and L. But in order to avoid the subtraction of 16.6 and estimate B directly, choose $H^T \equiv \left[S^T, B^T\right]$ as the state vector to be estimated. Now all elements of H are functionally independent and the elements of L are functions of S and B, i.e.,

$$16.7 \quad L = S + B$$

$$16.8 \quad H = \begin{bmatrix} S \\ B \end{bmatrix}$$

$$16.9 \quad \tilde{\Gamma}_H \equiv \begin{bmatrix} \tilde{\Gamma} & \tilde{\omega} \\ \tilde{\omega}^T & \tilde{\Gamma}_B \end{bmatrix} \equiv E\left[(\tilde{H} - H)(\tilde{H} - H)^T\right]$$

From 13.11 and 13.13 the equations for estimating $\hat{H}$ are

$$16.10 \quad (\hat{H} - H) = \left[\tilde{\Gamma}_H^{-1} + \frac{\partial\beta}{\partial H}^T R^{-1} \frac{\partial\beta}{\partial H}\right]^{-1} \left[\tilde{\Gamma}_H^{-1}(\tilde{H} - H) + \frac{\partial\beta}{\partial H}^T R^{-1}(\alpha - \beta)\right]$$

and

$$16.11 \quad \hat{\Gamma}_H = \left[\tilde{\Gamma}_H^{-1} + \frac{\partial\beta}{\partial H}^T R^{-1} \frac{\partial\beta}{\partial H}\right]^{-1}$$

The partitioned forms of 16.10 and 16.11 are useful as references in later sections.

Define:

$$16.12 \quad \begin{cases} A \equiv \dfrac{\partial \beta}{\partial S}^{T} R^{-1} \dfrac{\partial \beta}{\partial S} \\[2ex] C \equiv \dfrac{\partial \beta}{\partial S}^{T} R^{-1} \dfrac{\partial \beta}{\partial B} \\[2ex] M \equiv \dfrac{\partial \beta}{\partial B}^{T} R^{-1} \dfrac{\partial \beta}{\partial B} \\[2ex] N \equiv \dfrac{\partial \beta}{\partial S}^{T} R^{-1} (\alpha - \beta) \\[2ex] D \equiv \dfrac{\partial \beta}{\partial B}^{T} R^{-1} (\alpha - \beta) \end{cases}$$

Then, using 16.10, 16.11, 16.12,

$$16.13 \quad \begin{bmatrix} \hat{S} - S \\ \hat{B} - B \end{bmatrix} = \left\{ \begin{bmatrix} \tilde{\Gamma} & \tilde{\omega} \\ \tilde{\omega}^{T} & \tilde{\Gamma}_{B} \end{bmatrix}^{-1} + \begin{bmatrix} A & C \\ C^{T} & M \end{bmatrix} \right\}^{-1} \left\{ \begin{bmatrix} \tilde{\Gamma} & \tilde{\omega} \\ \tilde{\omega}^{T} & \tilde{\Gamma}_{B} \end{bmatrix}^{-1} \begin{bmatrix} \tilde{S} - S \\ \tilde{B} - B \end{bmatrix} + \begin{bmatrix} N \\ D \end{bmatrix} \right\}$$

and

$$16.14 \quad \begin{bmatrix} \hat{\Gamma} & \hat{\omega} \\ \hat{\omega}^{T} & \hat{\Gamma}_{B} \end{bmatrix} = \left\{ \begin{bmatrix} \tilde{\Gamma} & \tilde{\omega} \\ \tilde{\omega}^{T} & \tilde{\Gamma}_{B} \end{bmatrix}^{-1} + \begin{bmatrix} A & C \\ C^{T} & M \end{bmatrix} \right\}^{-1}$$

Note that the partitioned matrices can be inverted by 2.19.

The partitioned state transition matrix for propagating the covariance (16.14) from time, t_o, to time, t, is

$$16.15 \quad \frac{\partial H}{\partial H_o} = \begin{bmatrix} \dfrac{\partial S}{\partial S_o} & \dfrac{\partial S}{\partial B_o} \\[2ex] \dfrac{\partial B}{\partial S_o} & \dfrac{\partial B}{\partial B_o} \end{bmatrix}$$

It would be convenient to express 16.15 in terms of $\frac{\partial S}{\partial S_o}$ and $\frac{\partial L}{\partial L_o}$, since these derivatives can be computed by methods discussed in Section 11. As a worthwhile exercise, we shall derive the required expression in two different ways. First, suppose we had chosen $\left[S^T, L^T\right]$ as the basis elements; then a small deviation in the state vector at time, t, would be related to a small deviation at time, t_o, as

$$16.16 \quad \begin{bmatrix} \delta\tilde{S} \\ \\ \delta\tilde{L} \end{bmatrix} = \begin{bmatrix} \frac{\partial S}{\partial S_o} & \frac{\partial S}{\partial L_o} \\ \\ \frac{\partial L}{\partial S_o} & \frac{\partial L}{\partial L_o} \end{bmatrix} \begin{bmatrix} \delta\hat{S}_o \\ \\ \delta\hat{L}_o \end{bmatrix} = \begin{bmatrix} \frac{\partial S}{\partial S_o} & 0 \\ \\ 0 & \frac{\partial L}{\partial L_o} \end{bmatrix} \begin{bmatrix} \delta\hat{S}_o \\ \\ \delta\hat{L}_o \end{bmatrix}$$

where

$$16.17 \quad \begin{bmatrix} \delta\tilde{S} \\ \delta\tilde{L} \end{bmatrix} \equiv \begin{bmatrix} \tilde{S} - S \\ \tilde{L} - L \end{bmatrix}_t$$

and

$$16.18 \quad \begin{bmatrix} \delta\hat{S}_o \\ \delta\hat{L}_o \end{bmatrix} \equiv \begin{bmatrix} \hat{S} - S \\ \hat{L} - L \end{bmatrix}_{t_o}$$

Notice that $\frac{\partial S}{\partial L_o} = \frac{\partial L}{\partial S_o} = 0$, since S at time, t, is functionally independent of L at time, t_o, and vice versa. This is apparent from examining the equations of motion, remembering:

$$16.19 \quad S = \begin{bmatrix} \bar{r} \\ \dot{\bar{r}} \end{bmatrix}, \; L = \begin{bmatrix} \bar{r}_L \\ \dot{\bar{r}}_L \end{bmatrix}, \; B = \begin{bmatrix} \bar{\rho} \\ \dot{\bar{\rho}} \end{bmatrix} = \begin{bmatrix} \bar{r}_L - \bar{r} \\ \dot{\bar{r}}_L - \dot{\bar{r}} \end{bmatrix}$$

where

$$16.20 \quad \ddot{\bar{r}} = -\frac{\mu\bar{r}}{r^3} + g(\bar{r}, \dot{\bar{r}}, q, t) \qquad (1.1)$$

16.21 $\ddot{\bar{r}}_L = -\frac{\mu \bar{r}_L}{r_L^3} + g_L\left(\bar{r}_L, \dot{\bar{r}}_L, m, t\right)$

16.22 $\ddot{\bar{p}} - \ddot{\bar{r}}_L - \ddot{\bar{r}}$

Equation 16.16 can be mapped into a propagation of $\left[\delta\tilde{S}^T \quad \delta\tilde{B}^T\right]$ as follows:

16.23 $$\begin{bmatrix} I & \emptyset \\ -I & I \end{bmatrix}\begin{bmatrix} \delta\tilde{S} \\ \delta\tilde{L} \end{bmatrix} = \begin{bmatrix} I & \emptyset \\ -I & I \end{bmatrix}\begin{bmatrix} \frac{\partial S}{\partial S_o} & \emptyset \\ \emptyset & \frac{\partial L}{\partial L_o} \end{bmatrix}\begin{bmatrix} I & \emptyset \\ I & I \end{bmatrix}\begin{bmatrix} I & \emptyset \\ -I & I \end{bmatrix}\begin{bmatrix} \delta\hat{S}_o \\ \delta\hat{L}_o \end{bmatrix}$$

16.24 $$\begin{bmatrix} \delta\tilde{S} \\ \delta\tilde{B} \end{bmatrix} = \begin{bmatrix} \frac{\partial S}{\partial S_o} & \emptyset \\ \frac{\partial L}{\partial L_o} - \frac{\partial S}{\partial S_o} & \frac{\partial L}{\partial L_o} \end{bmatrix}\begin{bmatrix} \delta\hat{S}_o \\ \delta\hat{B}_o \end{bmatrix}$$

The matrix in 16.24 is the desired expression of $\frac{\partial H}{\partial H_o}$ (16.18). Now this same expression will be derived by a more direct method. Assume that $\left[S^T \quad B^T\right]$ is the set of functionally independent basis elements and propagation is as

16.25 $$\begin{bmatrix} \delta\tilde{S} \\ \delta\tilde{B} \end{bmatrix} = \begin{bmatrix} \frac{\partial S}{\partial S_o} & \frac{\partial S}{\partial B_o} \\ \frac{\partial B}{\partial S_o} & \frac{\partial B}{\partial B_o} \end{bmatrix}\begin{bmatrix} \delta\hat{S}_o \\ \delta\hat{B}_o \end{bmatrix}$$

From 16.20 S is a function of S_o.

16.26 $S = S(S_o)$

Then

16.27 $\frac{\partial S}{\partial S_o} = \frac{\partial S}{\partial S_o}$

16.28 $\frac{\partial S}{\partial B_o} = \emptyset$

From 16.22 B is a function of S_o and B_o,

16.29 $B = B(S_o, B_o) = L - S$ (where $L_o = S_o + B_o$)

$= L[L_o(S_o, B_o)] - S(S_o)$

Then, since $\frac{\partial L}{\partial L_o} = \frac{\partial L}{\partial B_o} = I$,

16.30 $\frac{\partial B}{\partial S_o} = \frac{\partial L}{\partial L_o}\frac{\partial L_o}{\partial S_o} - \frac{\partial S}{\partial S_o} = \frac{\partial L}{\partial L_o} - \frac{\partial S}{\partial S_o}$

and

16.31 $\frac{\partial B}{\partial B_o} = \frac{\partial L}{\partial L_o}\frac{\partial L_o}{\partial B_o} = \frac{\partial L}{\partial L_o}$

Substituting 16.27, 28, 30, 31 into 16.25 gives 16.24 again.

This completes the discussion of the general method for using onboard observations to estimate relative trajectories. The formulation could be modified in many ways to fit the requirements of specific situations. The process leading to 16.6 is an example of such a modification.

17. MODIFICATION OF THE STATE COVARIANCE MATRIX

Up till now we have accounted for observational errors, assuming, however, that the forces acting on a spacecraft are modeled perfectly as functions of precisely-known physical parameters. Actually our knowledge of these factors is limited, and for simplicity of computations we do not always even use the best model available. Questions arise, therefore, as to how we can account for any adverse effects on the estimation. It is not intuitively obvious that anything bad should occur, but on the contrary it seems that the estimates should always continue to improve as more measurement batches are processed. Historically, in the initial testing of the Bayes estimation programs, the covariance matrix, indeed, did get smaller and smaller, indicating a more accurate estimate of the state vector; the sequence of estimates, on the other hand, initially converged rapidly toward the true value, approached a minimum error after about two orbits, and then slowly began to diverge. The estimation process is equivalent to the method of generalized weighted least squares, where the a priori state vector represents a pseudomeasurement weighted by the a priori inverse state covariance matrix. This weighting matrix grows with each sequential step; so estimates become increasingly dominated by the a priori state, until the effect of new measurements is negligible. This situation implies that the estimates are always improving, which would be true if the dynamic model were perfect. The neglected errors of the real model, however, cause the propagated estimate of the state vector to depart farther from the truth. Hopefully, this would be corrected by processing the next batch, but the dilemma is met when the effect of the next batch becomes negligible. Then the shrinking determinant of the state covariance matrix ceases to truly represent the growing state estimate error, which is induced by propagation and uncorrected by estimation. A major problem in implementing this program is how to consider nodel errors in a way to achieve optimum estimates with errors correctly represented by the covariance matrix. All the tried methods have involved modification of the state covariance matrix. The simplest way is to consider that the origin of model errors is unknown; then multiply the matrix by a scalar > 1 when the determinant appears too small. A frequently-used manual control for this is in the real time system. Another approach is to guess the most likely sources of error, such as atmospheric drag, fuel venting and gravitational constant, and derive a term to be added to the state covariance matrix in propagation. This way, used in the Gemini program, was justified as an application of proper corrections to respective components. It took a lot of computing time, however, and seemed no more effective than the first method. A variation of the latter, which considers the model parameters in propagation of covariance, is in the Apollo program (Section 19). It has also presented many problems and has not yet proved completely satisfactory. Another approach (Section 20), as yet untried, is exponential downweighting of past data with respect to time.

This reduces the observation arc length to one that can be accurately reproduced by the model, and it also has the advantage of producing estimates independent of measurement batching and times of processing. [9]

The remainder of this section will present a general modification of a state covariance matrix with the intent of determining what can be done, what does it mean geometrically, and what are some reasonable criteria for evaluating any scheme for altering the state covariance matrix.

Define

17.1	S	true value of state vector, (p x 1)
17.2	$\tilde{S}$	estimate of S
17.3	$\delta\tilde{S} = \tilde{S} - S$	state error vector
17.4	$\tilde{\Gamma} = E(\delta\tilde{S}\delta\tilde{S}^T)$	state covariance matrix
17.5	T	a non-singular transformation with complex elements, (p x p)
17.6	T^*	conjugate transpose of T

Then the most general modification possible of $\delta\tilde{S}$ can be represented by $T\delta\tilde{S}$; the most general modification of $\tilde{\Gamma}$, by

17.7 $\tilde{\Gamma}_* = E(T\delta\tilde{S}\delta\tilde{S}^T T^*)$

The problem is to choose the matrix, T, to modify $\tilde{\Gamma}$ in a manner justified numerically as an advantage to the processor. For now, however, we shall be concerned with developing criteria to show whether a particular choice of T is reasonable, rather than with making the choice. Actual choices will be made at the end of this section and tested against the criteria. To start with we assume that T is diagonal. After all we are trying to preserve the past history, $\tilde{\Gamma}$, as nearly as possible, merely giving it an empirical "nudge" to correct some dilemma in the processor. To do this we should choose the simplest transformation possible. If T were non-diagonal, the change in $\widetilde{\Gamma}$ would probably be complicated, drastic, and difficult to justify. With this assumption the jth diagonal element of T is the complex number

17.8 $$T_{jj} = \zeta_j + i\widetilde{\gamma}_j$$

where ζ_j is a scalar constant to be chosen and $\widetilde{\gamma}_j$ is a zero-mean random variable uncorrelated with state noise, such that

17.9 $$E\left(\widetilde{\gamma}_j\widetilde{\gamma}_k\right) \equiv \begin{cases} \eta_j^2 & (j = k) \\ 0 & (j \neq k) \end{cases}$$

and η_j^2 is a scalar constant to be chosen. Defining $\widetilde{\gamma}_j$ as a random variable in this way ensures that the modified matrix, $\widetilde{\Gamma}_*$ (17.7), will have real elements, whereas $\widetilde{\gamma}$ as a constant would result in complex elements.

If $S^T = [x_1, \ldots, x_p]$ and

17.10 $$\widetilde{\Gamma} = [\sigma_{ij}] \qquad (i, j = 1, \ldots, p)$$

Then by 17.7 the elements of $\widetilde{\Gamma}$ correspond to the elements of $\widetilde{\Gamma}_*$ as

17.11 $$\sigma_{ij} \longrightarrow \begin{cases} \zeta_i\zeta_j\sigma_{ij} & i \neq j \\ \left(\zeta_i^2 + \eta_i^2\right)\sigma_{ii} & i = j \end{cases}$$

17.12 (Note: $\sigma_i^2 \equiv \sigma_{ii}$)

Note that the matrix $\tilde{\Gamma}_*$ is still positive definite, since

17.13 $$|\tilde{\Gamma}_*| \geq \zeta_1^2 \cdots \zeta_p^2 \, |\tilde{\Gamma}| > 0$$

The geometrical meaning of this transformation will be illustrated by considering a three-dimensional state error vector,

17.14 $$\delta\tilde{S}^T = [\delta\tilde{x}, \ \delta\tilde{y}, \ \delta\tilde{z}]$$

and a matrix, T, which modifies only the $\delta\tilde{z}$ component,

17.15 $$T = \begin{bmatrix} 1 & & 0 \\ & 1 & \\ 0 & & \zeta + i\tilde{\gamma} \end{bmatrix}$$

(Note: $\zeta + i\tilde{\gamma} \equiv \zeta_3 + i\tilde{\gamma}_3$) (17.8)

Then the modified covariance matrix is

17.16 $$\tilde{\Gamma}_* = \begin{bmatrix} \sigma_x^2 & \sigma_{xy} & \zeta\sigma_{xz} \\ \sigma_{xy} & \sigma_y^2 & \zeta\sigma_{yz} \\ \zeta\sigma_{xz} & \zeta\sigma_{yz} & (\zeta^2 + \eta^2)\sigma_z^2 \end{bmatrix} \qquad (17.11)$$

Let ρ_{ij} be a correlation coefficient, i.e.,

17.17 $$\sigma_{ij} = \rho_{ij}\sigma_i\sigma_j$$

Then

17.18 $$\tilde{\Gamma}_* = \begin{bmatrix} \sigma_x & & 0 \\ & \sigma_y & \\ 0 & & \sigma_z \end{bmatrix} \begin{bmatrix} 1 & \rho_{xy} & \zeta\rho_{xz} \\ \rho_{xy} & 1 & \zeta\rho_{yz} \\ \zeta\rho_{xz} & \zeta\rho_{yz} & \zeta^2 + \eta^2 \end{bmatrix} \begin{bmatrix} \sigma_x & & 0 \\ & \sigma_y & \\ 0 & & \sigma_z \end{bmatrix}$$

Also

$$17.19 \quad \tilde{\Gamma}_* = \begin{bmatrix} \sigma_x & & 0 \\ & \sigma_y & \\ 0 & & \sqrt{\zeta^2 + \eta^2}\,\sigma_z \end{bmatrix} \begin{bmatrix} 1 & \rho_{xy} & \frac{\zeta}{\sqrt{\zeta^2 + \eta^2}}\rho_{xz} \\ \rho_{xy} & 1 & \frac{\zeta}{\sqrt{\zeta^2 + \eta^2}}\rho_{yz} \\ \frac{\zeta}{\sqrt{\zeta^2 + \eta^2}}\rho_{xz} & \frac{\zeta}{\sqrt{\zeta^2 + \eta^2}}\rho_{yz} & 1 \end{bmatrix} \begin{bmatrix} \sigma_x & & 0 \\ & \sigma_y & \\ 0 & & \sqrt{\zeta^2 + \eta^2}\,\sigma_z \end{bmatrix}$$

The quadratic form associated with 17.19 is

$$17.20 \quad 2\varphi = \delta\tilde{S}^T \tilde{\Gamma}_*^{-1} \delta\tilde{S}$$

$$17.21 \quad 2\varphi = \begin{bmatrix} \delta\tilde{x}, & \delta\tilde{y}, & \frac{\delta\tilde{z}}{\sqrt{\zeta^2 + \eta^2}} \end{bmatrix} \begin{bmatrix} \frac{1}{\sigma x} & & 0 \\ & \frac{1}{\sigma y} & \\ 0 & & \frac{1}{\sigma z} \end{bmatrix} \begin{bmatrix} 1 & \rho_{xy} & \frac{\zeta}{\sqrt{\zeta^2 + \eta^2}}\rho_{xz} \\ \rho_{xy} & 1 & \frac{\zeta}{\sqrt{\zeta^2 + \eta^2}}\rho_{yz} \\ \frac{\zeta}{\sqrt{\zeta^2 + \eta^2}}\rho_{xz} & \frac{\zeta}{\sqrt{\zeta^2 + \eta^2}}\rho_{yz} & 1 \end{bmatrix}^{-1} \begin{bmatrix} \frac{1}{\sigma x} & & 0 \\ & \frac{1}{\sigma y} & \\ 0 & & \frac{1}{\sigma z} \end{bmatrix} \begin{bmatrix} \delta\tilde{x} \\ \delta\tilde{y} \\ \frac{\delta\tilde{z}}{\sqrt{\zeta^2 + \eta^2}} \end{bmatrix}$$

Inspection of equations 17.16 and 17.21 shows the constraints imposed in the choice of ζ and η^2 and also the geometric significance.

Some examples follow:

a. First notice that if $\zeta = 1$ and $\eta^2 = 0$, then $T = I$, $\tilde{\Gamma}_* = \tilde{\Gamma}$, and the quadratic form is unchanged. If various values of ζ and η^2 exist to cure the same problem in the filter, then the choice should be the values closest to these fundamental values.

b. A constraint is that $|\zeta| \leq \sqrt{\zeta^2 + \eta^2}$. This means that a multiplier, $\zeta^2 + \eta^2 < 1$, cannot be used on σ_z^2, if $\zeta = 1$ for the off-diagonal terms. For example, initially in the Apollo program an attempt was made to modify an augmented state covariance matrix, $\begin{bmatrix} \tilde{\Gamma} & \tilde{\omega} \\ \tilde{\omega}^T & \tilde{\Gamma}_q \end{bmatrix}$, by multiplying $\tilde{\Gamma}_q$ by a positive scalar < 1; the resulting matrix was not always positive definite.

$$\begin{bmatrix} \tilde{\Gamma} & \tilde{\omega} \\ \tilde{\omega}^T & \tilde{\Gamma}_q \end{bmatrix} \not\rightarrow \begin{bmatrix} \tilde{\Gamma} & \tilde{\omega} \\ \tilde{\omega}^T & (1 + \eta^2)\tilde{\Gamma}_q \end{bmatrix} \qquad (0 < 1 + \eta^2 \not< 1)$$

c. If $\zeta = 0$ and $\eta^2 = 1$, then the covariance matrix remains unchanged except that elements multiplied by ζ are zeroed. An example of this is the valid procedure (under our rules) used in the Apollo program to modify the matrix as

$$\begin{bmatrix} \tilde{\Gamma} & \tilde{\omega} \\ \tilde{\omega}^T & \tilde{\Gamma}_q \end{bmatrix} \rightarrow \begin{bmatrix} \tilde{\Gamma} & 0 \\ 0 & \tilde{\Gamma}_q \end{bmatrix}$$

d. If $\eta^2 = 0$, then the $\delta\tilde{z}$ component is re-scaled (multiplied by $\frac{1}{\zeta}$), and correlation coefficients are not altered (17.21). Or equivalently, a row and column of the covariance matrix is multiplied by ζ (17.16). This was used instead of b, above, to decrease the value of $\tilde{\Gamma}_q$ as

$$\begin{bmatrix} \tilde{\Gamma} & \tilde{\omega} \\ \tilde{\omega}^T & \tilde{\Gamma}_q \end{bmatrix} \rightarrow \begin{bmatrix} \tilde{\Gamma} & \zeta\tilde{\omega} \\ \zeta\tilde{\omega}^T & \zeta^2\tilde{\Gamma}_q \end{bmatrix} \qquad (0 < \zeta < 1)$$

Another example, used in the Apollo program, and also in exponential downweighting, is the multiplication of the entire matrix by a scalar as

$$\tilde{\Gamma} \rightarrow \zeta^2\tilde{\Gamma} \qquad (1 < \zeta^2)$$

e. If $\sqrt{\zeta^2 + \eta^2} = 1$, then the $\delta\tilde{z}$ correlation coefficients are multiplied by ζ, and no components are re-scaled (17.21). For example (not used in the Apollo program)

$$\begin{bmatrix} \tilde{\Gamma} & \tilde{\omega} \\ \tilde{\omega}^T & \tilde{\Gamma}_q \end{bmatrix} \longrightarrow \begin{bmatrix} \tilde{\Gamma} & \zeta\tilde{\omega} \\ \zeta\tilde{\omega}^T & \tilde{\Gamma}_q \end{bmatrix} \qquad (0 \le \zeta < 1)$$

f. Now refer to equation 17.16. If $\zeta = 1$ and η is chosen so that $\eta^2 \sigma_z^2 = k$, a constant, then

$$17.22 \quad \tilde{\Gamma}_* = \begin{bmatrix} \sigma_x^2 & \sigma_{xy} & \sigma_{xz} \\ \sigma_{xy} & \sigma_y^2 & \sigma_{yz} \\ \sigma_{xz} & \sigma_{yz} & \sigma_z^2 + k \end{bmatrix}$$

All we have done is add a constant to a main diagonal element. This is the method used in the LM powered flight processor [12]. This method was arrived at by considering the model errors in the derivation of the filter. It is interesting to note that we can arrive at the same method empirically, using the rules of this section.

From this we can compute the effects on re-scaling and correlation resulting from various other choices of ζ and η^2. We start out knowing that theoretically the best ζ and η^2 are as in (a) above, and any deviations from this should be justified numerically.

18. ESTIMATION OF MEASUREMENT MODEL BIASES

The mathematical model of a measurement may be a function of a bias element such as the scalar, b, in 14.44. Although b is essentially a constant, its value may drift slightly and is not known precisely; so the best estimate of it is used in computation. Because of this the filter is designed to allow the trajectory controller to alter the process at any sequential step so as to include estimation of the bias elements. Mathematically this is done by adjoining the bias elements to the state vector, then estimating this augmented state vector, and finally contracting the augmented state vector and covariance matrix back to their original dimensions. After this the filter proceeds in the usual manner (unless interrupted again), and the new values of the bias elements are used in modeling measurements. In the following discussion we show how to alter the filter to estimate bias elements and then return it to the original form.

Define

18.1 $b \equiv \begin{bmatrix} b_1 \\ \vdots \\ b_q \end{bmatrix}$ mean value of the bias vector, the elements of which are bias constants.

Conforming to previous notation:

18.2 $\tilde{b}$ a priori estimate of b

18.3 $\delta\tilde{b} = \tilde{b} - b$

18.4 $\hat{b}$ improved estimate of b

18.5 $\delta\hat{b} = \hat{b} - b$

18.6 $t_i < t_j < t_k$ anchor times, where t_i is the time of estimating b, and t_j and t_k are the next two anchor times.

From here on the derivation is just like equations 16.8 - 16.14 with B replaced by b. The augmented state vector to be estimated at time, t_i, is

18.7 $H^T = [S^T, b^T]$ (16.8)

18.8 $$\tilde{\Gamma}_H = E\left(\delta\tilde{H}\delta\tilde{H}^T\right) = \begin{bmatrix} \tilde{\Gamma} & \tilde{\omega} \\ \tilde{\omega}^T & \tilde{\Gamma}_b \end{bmatrix} \qquad (16.9)$$

Define

18.9 $$\begin{bmatrix} \tilde{\Gamma} & \tilde{\omega} \\ \tilde{\omega}^T & \tilde{\Gamma}_b \end{bmatrix}^{-1} \equiv \begin{bmatrix} \tilde{G} & \tilde{V} \\ \tilde{V}^T & \tilde{\Omega} \end{bmatrix}$$

The augmented state vector, $\tilde{H}$, is formed at time, t_i, and since we have no prior knowledge of $\tilde{\omega}$, assume it to be zero.

18.10 $$\tilde{\omega} = 0$$

Then 18.9 becomes

18.11 $$\begin{bmatrix} \tilde{\Gamma} & 0 \\ 0 & \tilde{\Gamma}_b \end{bmatrix}^{-1} = \begin{bmatrix} \tilde{G} & 0 \\ 0 & \tilde{\Omega} \end{bmatrix}$$

Combining 16.13 and 18.11, the filter for estimating the augmented state vector is

18.12 $$\begin{bmatrix} \hat{S} - S \\ \hat{b} - b \end{bmatrix} = \left\{ \begin{bmatrix} \tilde{G} & 0 \\ 0 & \tilde{\Omega} \end{bmatrix} + \begin{bmatrix} A & C \\ C^T & M \end{bmatrix} \right\}^{-1} \left\{ \begin{bmatrix} \tilde{G} & 0 \\ 0 & \tilde{\Omega} \end{bmatrix} \begin{bmatrix} \tilde{S} - S \\ \tilde{b} - b \end{bmatrix} + \begin{bmatrix} N \\ D \end{bmatrix} \right\}$$

where

18.13 $$\begin{bmatrix} \tilde{G} + A & C \\ C^T & \tilde{\Omega} + M \end{bmatrix} = \begin{bmatrix} \hat{G} & \hat{V} \\ \hat{V}^T & \hat{\Omega} \end{bmatrix}$$

The solution to 18.12, $\begin{bmatrix}\hat{S}^T & \hat{b}^T\end{bmatrix}$, is the improved estimate of the augmented state vector at time, t_i, and 18.13 is the associated covariance matrix. These (18.12, 18.13) are propagated from time, t_i, to time, t_j, where the next batch of measurements will be processed to estimate S only.

That is,

18.14 $$\begin{bmatrix}\hat{S} \\ \hat{b}\end{bmatrix}_i \longrightarrow \begin{bmatrix}\tilde{S} \\ \tilde{b}\end{bmatrix}_j$$

and

18.15 $$\begin{bmatrix}\hat{G} & \hat{V} \\ \hat{V}^T & \hat{\Omega}\end{bmatrix}_i \longrightarrow \begin{bmatrix}\tilde{G} & \tilde{V} \\ \tilde{V}^T & \tilde{\Omega}\end{bmatrix}_j$$

The quadratic form to be minimized with respect to S at time, t_j, is

18.16 $$2\varphi = \left[(\tilde{S} - S)^T (\tilde{b} - b)^T\right]\begin{bmatrix}\tilde{G} & \tilde{V} \\ \tilde{V}^T & \tilde{\Omega}\end{bmatrix}\begin{bmatrix}\tilde{S} - S \\ \tilde{b} - b\end{bmatrix} + (\alpha - \beta)^T R^{-1}(\alpha - \beta)$$

Using the methods of Section 12,

18.17 $$\Phi = \frac{\partial\varphi}{\partial S}^T = [-I \quad 0]\begin{bmatrix}\tilde{G} & \tilde{V} \\ \tilde{V}^T & \tilde{\Omega}\end{bmatrix}\begin{bmatrix}\tilde{S} - S \\ \tilde{b} - b\end{bmatrix} - \frac{\partial\beta}{\partial S}^T R^{-1}(\alpha - \beta)$$

18.18 $$\frac{\partial\Phi}{\partial S} = \tilde{G} + \frac{\partial\beta}{\partial S}^T R^{-1}\frac{\partial\beta}{\partial S}$$

Using Newton's method (Section 12):

18.19 $$\hat{S}_\epsilon - S = -\left(\frac{\partial\Phi}{\partial S}\right)^{-1}\Phi$$

18.20 $$\hat{S}_\varepsilon - S = \left[\tilde{G} + \frac{\partial\beta}{\partial S}^T R^{-1} \frac{\partial\beta}{\partial S}\right]^{-1} \left[\tilde{G}(\tilde{S} - S) + \tilde{V}(\tilde{b} - b) + \frac{\partial\beta}{\partial S}^T R^{-1}(\alpha - \beta)\right]$$

where $\hat{S}_\varepsilon$ is the vector to which we would converge if we knew the value of $\tilde{b} - b$. Since we do not know this, we define

18.21 $$\hat{S} \equiv \hat{S}_\varepsilon - \hat{Q}^{-1}\tilde{V}\delta\tilde{b}$$

and

18.22 $$\hat{Q} \equiv \tilde{G} + \frac{\partial\beta}{\partial S}^T R^{-1} \frac{\partial\beta}{\partial S}$$

Then equation 18.20 is expressed as

18.23 $$(\hat{S} - S) = \hat{Q}^{-1}\left[\tilde{G}(\tilde{S} - S) + \frac{\partial\beta}{\partial S}^T R^{-1}(\alpha - \beta)\right] \qquad (18.20\text{-}22)$$

where convergence is to the vector, $\hat{S}$.

Now we show that $\hat{Q} = \hat{G}$, as follows:

By the method of 12.51 and 12.52,

18.24 $$\hat{\Gamma} = E\left(\delta\hat{S}\delta\hat{S}^T\right) = \hat{Q}^{-1}\left[\tilde{G}\tilde{\Gamma}\tilde{G} + \frac{\partial\beta}{\partial S}^T R^{-1} \frac{\partial\beta}{\partial S}\right]\hat{Q}^{-1} \qquad (18.23)$$

18.25 $$\hat{\omega} = E\left(\delta\hat{S}\delta\tilde{b}^T\right) = \hat{Q}^{-1}\tilde{G}\tilde{\omega} \qquad (18.23)$$

18.26 $$\hat{\Gamma}_b = \tilde{\Gamma}_b$$

18.27 $$\begin{bmatrix} \hat{G} & \hat{V} \\ \hat{V}^T & \hat{\Omega} \end{bmatrix} = \begin{bmatrix} \hat{\Gamma} & \hat{\omega} \\ \hat{\omega}^T & \tilde{\Gamma}_b \end{bmatrix}^{-1} \qquad (18.9, 18.24\text{-}6)$$

18.28 $\hat{G}^{-1} = \hat{\Gamma} - \hat{\omega}\tilde{\Gamma}_b^{-1}\hat{\omega}^T$ (2.20, 18.27)

and similarly

18.29 $\tilde{G}^{-1} = \tilde{\Gamma} - \tilde{\omega}\tilde{\Gamma}_b^{-1}\tilde{\omega}^T$ (2.20, 18.9)

from which

18.30 $\tilde{\omega}\tilde{\Gamma}_b^{-1}\tilde{\omega}^T = \tilde{\Gamma} - \tilde{G}^{-1}$

Substituting 18.24 and 18.25 into 18.28 gives

18.31 $\hat{G}^{-1} = \hat{Q}^{-1}\left[\tilde{G}\tilde{\Gamma}\tilde{G} + \frac{\partial\beta}{\partial S}^T R^{-1}\frac{\partial\beta}{\partial S}\right]\hat{Q}^{-1} \;\Big|\; - \hat{Q}^{-1}\tilde{G}\tilde{\omega}\tilde{\Gamma}_b^{-1}\tilde{\omega}^T\tilde{G}\hat{Q}^{-1}$

18.32 $\Big|\; - \hat{Q}^{-1}\tilde{G}\left(\tilde{\Gamma} - \tilde{G}^{-1}\right)\tilde{G}\hat{Q}^{-1}$ (18.30)

18.33 $\Big|\; - \hat{Q}^{-1}\left[\tilde{G}\tilde{\Gamma}\tilde{G} - \tilde{G}\right]\hat{Q}^{-1}$ (18.32)

18.34 $\hat{G}^{-1} = \hat{Q}^{-1}\left[\tilde{G} + \frac{\partial\beta}{\partial S}^T R^{-1}\frac{\partial\beta}{\partial S}\right]\hat{Q}^{-1}$ (18.22, 18.33)

18.35 $\hat{G}^{-1} = \hat{Q}^{-1}$

which was to be proved, and 18.23 can be written

18.36 $(\hat{S} - S) = \hat{G}^{-1}\left[\tilde{G}(\tilde{S} - S) + \frac{\partial\beta}{\partial S}^T R^{-1}(\alpha - \beta)\right]$ (18.35)

This equation (18.36) was derived using a rather fundamental approach, starting with 18.14 and 18.15 and forming the quadratic form, 18.16. A quicker derivation which provides less insight is as follows: As in 18.14 and 18.15 we start with the a priori quantities

$$\begin{bmatrix} \tilde{S} \\ \tilde{b} \end{bmatrix}_j \quad \text{and} \quad \begin{bmatrix} \tilde{G} & \tilde{V} \\ \tilde{V}^T & \tilde{\Omega} \end{bmatrix}_j$$

Then as in 18.12, the filter for estimating the augmented state vector at time, t_j, is

18.37 $$\begin{bmatrix} \hat{S} - S \\ \hat{b} - b \end{bmatrix} = \left\{ \begin{bmatrix} \tilde{G} & \tilde{V} \\ \tilde{V}^T & \tilde{\Omega} \end{bmatrix} + \begin{bmatrix} A & C \\ C^T & M \end{bmatrix} \right\}^{-1} \left\{ \begin{bmatrix} \tilde{G} & \tilde{V} \\ \tilde{V}^T & \tilde{\Omega} \end{bmatrix} \begin{bmatrix} \tilde{S} - S \\ \tilde{b} - b \end{bmatrix} + \begin{bmatrix} N \\ D \end{bmatrix} \right\}$$

where

18.38 $$\begin{bmatrix} \hat{G} & \hat{V} \\ \hat{V}^T & \hat{\Omega} \end{bmatrix} = \begin{bmatrix} \tilde{G} + A & \tilde{V} + C \\ \tilde{V}^T + C^T & \tilde{\Omega} + M \end{bmatrix} = \begin{bmatrix} \hat{\Gamma} & \hat{\omega} \\ \hat{\omega}^T & \hat{\Gamma}_b \end{bmatrix}^{-1}$$

Modifying this filter to estimate S only is equivalent to setting

18.39 $\hat{b} = \tilde{b} = b$

Substituting 18.39 into 18.37 gives

18.40 $$\begin{bmatrix} \hat{S} - S \\ 0 \end{bmatrix} = \begin{bmatrix} \hat{\Gamma} & \hat{\omega} \\ \hat{\omega}^T & \hat{\Gamma}_b \end{bmatrix} \left\{ \begin{bmatrix} \tilde{G} & \tilde{V} \\ \tilde{V}^T & \tilde{\Omega} \end{bmatrix} \begin{bmatrix} \tilde{S} - S \\ 0 \end{bmatrix} + \begin{bmatrix} N \\ D \end{bmatrix} \right\}$$

from which

18.41 $\delta\hat{S} = \hat{\Gamma}(\tilde{G}\delta\tilde{S} + N) + \hat{\omega}(\tilde{V}^T \delta\tilde{S} + D)$

18.42 $0 = \hat{\omega}^T (\tilde{G}\delta\tilde{S} + N) + \hat{\Gamma}_b(\tilde{V}^T \delta\tilde{S} + D)$

Solving 18.42 for $\tilde{V}^T \delta\tilde{S} + D$ and substituting into 18.41 gives

18.43 $\delta\hat{S} = (\hat{\Gamma} - \hat{\omega}\hat{\Gamma}_b^{-1}\hat{\omega}^T)(\tilde{G}\delta\tilde{S} + N)$

and

18.44 $(\hat{S} - S) = \hat{G}^{-1}[\tilde{G}(\tilde{S} - S) + N]$ (18.43)

which is the same as 18.36. This latter derivation is worth remembering for those cases where the state vector is frequently augmented and contracted. Then 18.37 can be the basic filter, which is altered by the input, 18.39.

Equation 18.36 provides the estimate at time, t_j. The state vector and covariance matrix have been reduced to the original dimensions, and the bias vector has no effect on subsequent estimates. Notice that the solution of 18.36 requires propagation only of the partition, $\hat{G}_i$, of 18.15 as

$$\hat{G}_i \longrightarrow \tilde{G}_j \, .$$

The augmented state transition matrix is

18.45 $$\begin{bmatrix} \frac{\partial S_j}{\partial S_i} & \frac{\partial S_j}{\partial b} \\ \frac{\partial b}{\partial S_i} & \frac{\partial b}{\partial b} \end{bmatrix} = \begin{bmatrix} \frac{\partial S_j}{\partial S_i} & 0 \\ 0 & I \end{bmatrix}$$

and the inverse is

18.46 $$\begin{bmatrix} \frac{\partial S_i}{\partial S_j} & 0 \\ 0 & I \end{bmatrix}$$

The transition of 18.15 is

18.47
$$\begin{bmatrix} \tilde{G} & \tilde{V} \\ \tilde{V}^T & \tilde{\Omega} \end{bmatrix}_j = \begin{bmatrix} \frac{\partial S_i^T}{\partial S_j} & 0 \\ 0 & I \end{bmatrix} \begin{bmatrix} \hat{G} & \hat{V} \\ \hat{V}^T & \hat{\Omega} \end{bmatrix}_j \begin{bmatrix} \frac{\partial S_i}{\partial S_j} & 0 \\ 0 & I \end{bmatrix}$$

from which

18.48
$$\tilde{G}_j = \frac{\partial S_i^T}{\partial S_j} \hat{G}_i \frac{\partial S_i}{\partial S_j}$$

or equivalently

18.49
$$\tilde{G}_j^{-1} = \frac{\partial S_j}{\partial S_i} \hat{G}_i^{-1} \frac{\partial S_j^T}{\partial S_i}$$

From here on the sequential estimation procedure is defined by 18.36 and 18.49, regardless of whatever label is assigned to the matrix, G. So replace the letter, G, by Γ^{-1} in 18.36 and 18.49, and we have returned to the original filter and notation.

The results of all the above details can be summarized in the following very simple procedure. After estimating the augmented state vector at time, t_i, we have

18.50
$$\begin{bmatrix} \hat{\Gamma} & \hat{\omega} \\ \hat{\omega}^T & \hat{\Gamma}_b \end{bmatrix}_i = \begin{bmatrix} \hat{G} & \hat{V} \\ \hat{V}^T & \hat{\Omega} \end{bmatrix}^{-1} \qquad (18.9)$$

Then the filter to estimate S at time, t_j, is

18.51
$$(\hat{S} - S)_j = \left[\tilde{\Gamma}_j^{-1} + \frac{\partial \beta^T}{\partial S_j} R^{-1} \frac{\partial \beta}{\partial S_j}\right]^{-1} \left[\tilde{\Gamma}_j^{-1}(\tilde{S} - S)_j + \frac{\partial \beta^T}{\partial S_j} R^{-1}(\alpha - \beta)\right]$$

where

18.52 $\tilde{\Gamma}_j = \frac{\partial S_j}{\partial S_i} \hat{G}_i^{-1} \frac{\partial S_j}{\partial S_i}^T$ (18.36, 18.49)

and subsequent estimates and notation are as in Section 13.

In the ATE program this procedure has been modified. Equation 18.51 is used, but 18.52 is replaced by

18.53 $\tilde{\Gamma}_j = \frac{\partial S_j}{\partial S_i} \hat{\Gamma}_i \frac{\partial S_j}{\partial S_i}^T$

This is equivalent to setting the partition $\hat{\omega}_i = \emptyset$ in 18.50, as in example (c), Section 17. This modification does not give any computational advantage but it is permissible by the rules of Section 17 as long as subsequent estimates are not degraded significantly.

19. CONSIDERING DYNAMIC MODEL PARAMETERS IN PROPAGATION OF COVARIANCE

In Section 17 we mentioned that the state covariance matrix is modified in the ATE program by considering the variances of dynamic model parameters in propagation, even though the parameters are not estimated. In this section we show that these considerations, by themselves, leave the filter unchanged. The ATE method is achieved, however, by including an empirical modification to the augmented state covariance matrix.

Definitions

19.1 $j = i + 1 \qquad (i = 0, 1, \ldots)$

19.2 t_i time of processing the ith sequential batch of measurements

19.3 t_j the time of processing the next sequential batch after time, t_i

19.4 S_i state vector estimated at time, t_i

19.5 $q = \begin{bmatrix} \mu \\ c_d \end{bmatrix}$ vector of dynamic model parameters

19.6 $\begin{bmatrix} S \\ q \end{bmatrix}_i$ augmented state vector at time, t_i

19.7 $\begin{bmatrix} \tilde{\Gamma} & \tilde{\omega} \\ \tilde{\omega}^T & \tilde{\Gamma}_q \end{bmatrix}_i = \begin{bmatrix} \tilde{G} & \tilde{V} \\ \tilde{V}^T & \tilde{\Omega} \end{bmatrix}_i^{-1}$ a priori augmented state covariance matrix

19.8 $(\alpha - \beta)_i$ vector of residuals of the i^{th} batch

19.9 $A_i \equiv \dfrac{\partial \beta}{\partial S_i}^T R_i^{-1} \dfrac{\partial \beta}{\partial S_i}$

19.10 $N_i \equiv \dfrac{\partial \beta}{\partial S_i}^T R_i^{-1} (\alpha - \beta)_i$

Note that generally, the notation is as in Section 18, except that b is replaced by q.

Initial Assumptions

The basic filter is modified using these assumptions:

(1) $\tilde{\omega}_o = \emptyset$ (19.7)

(2) S, only, is estimated

(3) The updated augmented state covariance matrix is propagated as

$$\begin{bmatrix} \hat{\Gamma} & \hat{\omega} \\ \hat{\omega}^T & \tilde{\Gamma}_q \end{bmatrix}_i \longrightarrow \begin{bmatrix} \tilde{\Gamma} & \tilde{\omega} \\ \tilde{\omega}^T & \tilde{\Gamma}_q \end{bmatrix}_j$$

Note that $\tilde{\Gamma}_{q_o}$ is input to the program, and then $\tilde{\Gamma}_{q_o} = \tilde{\Gamma}_{q_i}$ for every i. This is because S, only, is estimated.

The Modified Filter

The a priori quantities at time, t_i, are

$$19.11 \quad \begin{bmatrix} \tilde{S} \\ \tilde{q} \end{bmatrix}_i \quad \text{and} \quad \begin{bmatrix} \tilde{\Gamma} & \tilde{\omega} \\ \tilde{\omega}^T & \tilde{\Gamma}_q \end{bmatrix}_i = \begin{bmatrix} \tilde{G} & \tilde{V} \\ \tilde{V}^T & \tilde{\Omega} \end{bmatrix}_i^{-1} \qquad (19.7)$$

with only S to be estimated. We showed in Section 18 that the filter for this is

$$19.12 \quad (\hat{S} - S)_i = [\tilde{G}_i + A_i]^{-1}[\tilde{G}_i(\tilde{S} - S)_i + N_i] \qquad (18.36)$$

where

$$19.13 \quad \hat{G}_i = \tilde{G}_i + A_i \qquad (18.22, 18.35)$$

and the necessary propagation of covariance is

$$19.14 \quad \tilde{G}_j^{-1} = \frac{\partial S_j}{\partial S_i} \hat{G}_i^{-1} \frac{\partial S_j}{\partial S_i}^T \qquad (18.49,\ 19.27,\ 19.31)$$

So the modified filter which results from the three initial assumptions is defined by 19.12, 19.13, and 19.14.

What to Prove

The basic filter derived in Section 13 is

$$19.15 \quad (\hat{S} - S)_i = \left[\tilde{\Gamma}_i^{-1} + A_i\right]^{-1} \left[\tilde{\Gamma}_i^{-1} (\tilde{S} - S)_i + N_i\right]$$

where

$$19.16 \quad \hat{\Gamma}_i^{-1} = \tilde{\Gamma}_i^{-1} + A_i$$

and propagation of covariance is

$$19.17 \quad \tilde{\Gamma}_j = \frac{\partial S_j}{\partial S_i} \hat{\Gamma}_i \frac{\partial S_j}{\partial S_j}^T$$

(Note: Do not confuse $\tilde{\Gamma}_i$ in 19.15 with $\tilde{\Gamma}_i$ of 19.11. They are not the same, except at time, t_o.)

In order to prove that the basic and modified filters (19.12 and 19.15) are identical, it is necessary and sufficient to show that

$$19.18 \quad \tilde{G}_k = \tilde{\Gamma}_k^{-1} \qquad (k = 0, 1, \ldots)$$

for all k.

Proof

Proof is by induction on k.

k = 0:

19.19 $\tilde{\omega}_o = 0$ (assumption 1)

19.20 $\tilde{G}_o = \left[\tilde{\Gamma}_o - \tilde{\omega}_o \tilde{\Gamma}_q^{-1} \tilde{\omega}_o^T\right]^{-1} = \tilde{\Gamma}_o^{-1}$ (19.19)

k = i:

19.21 $\tilde{G}_i = \tilde{\Gamma}_i^{-1}$ (induction hypothesis)

19.22 $\tilde{G}_j = \left(\frac{\partial S_j}{\partial S_i} \hat{G}_i^{-1} \frac{\partial S_j^T}{\partial S_i}\right)^{-1}$ (19.14)

19.23 $= \left[\frac{\partial S_j}{\partial S_i} \left(\tilde{G}_i + A_i\right)^{-1} \frac{\partial S_j^T}{\partial S_i}\right]^{-1}$ (19.13)

19.24 $= \left[\frac{\partial S_j}{\partial S_i} \left(\tilde{\Gamma}_i^{-1} + A_i\right)^{-1} \frac{\partial S_j^T}{\partial S_i}\right]^{-1}$ (19.21)

19.25 $= \left[\frac{\partial S_j}{\partial S_i} \hat{\Gamma}_i \frac{\partial S_j^T}{\partial S_i}\right]^{-1}$ (19.16)

19.26 $= \tilde{\Gamma}_j^{-1}$ (19.17)

Therefore, the modified filter is identical to the basic filter.

An Empirical Modification

In the ATE program the modified filter is used together with the empirical modification of Section 17, example (c). By assumption 3, covariance is propagated as

$$19.27 \quad \begin{bmatrix}\tilde{\Gamma} & \tilde{\omega} \\ \tilde{\omega}^T & \tilde{\Gamma}_q\end{bmatrix}_j = \begin{bmatrix}\frac{\partial S_j}{\partial S_i} & \frac{\partial S_j}{\partial q} \\ 0 & I\end{bmatrix} \begin{bmatrix}\hat{\Gamma} & \hat{\omega} \\ \hat{\omega}^T & \tilde{\Gamma}_q\end{bmatrix}_i \begin{bmatrix}\frac{\partial S_j^T}{\partial S_i} & 0 \\ \frac{\partial S_j^T}{\partial q} & I\end{bmatrix} = \begin{bmatrix}\tilde{G} & \tilde{V} \\ \tilde{V}^T & \tilde{\Omega}\end{bmatrix}_j^{-1} \qquad [10]$$

$$19.28 \quad = \begin{bmatrix}\frac{\partial S_j}{\partial S_i}\hat{\Gamma}_i\frac{\partial S_j^T}{\partial S_i} + \frac{\partial S_j}{\partial S_i}\hat{\omega}_i\frac{\partial S_j^T}{\partial q} + \frac{\partial S_j}{\partial q}\hat{\omega}_i^T\frac{\partial S_j^T}{\partial S_i} + \frac{\partial S_j}{\partial q}\tilde{\Gamma}_q\frac{\partial S_j^T}{\partial q} & \frac{\partial S_j}{\partial S_i}\hat{\omega}_i + \frac{\partial S_j}{\partial q}\tilde{\Gamma}_q \\ \hat{\omega}_i^T\frac{\partial S_j^T}{\partial S_i} + \tilde{\Gamma}_q\frac{\partial S_j^T}{\partial q} & \tilde{\Gamma}_q\end{bmatrix}$$

The partition, $\tilde{G}_j$, for use in the filter (19.12) is computed as

$$19.29 \quad G_j^{-1} = \tilde{\Gamma}_j - \tilde{\omega}_j\tilde{\Gamma}_q^{-1}\tilde{\omega}_j^T$$

At this point is the empirical modification. The partition $\tilde{\omega}_j$ is set equal to zero, so that

$$19.30 \quad \tilde{G}_j^{-1} = \frac{\partial S_j}{\partial S_i}\hat{\Gamma}_i\frac{\partial S_j^T}{\partial S_i} + \frac{\partial S_j}{\partial S_i}\hat{\omega}_i\frac{\partial S_j^T}{\partial q} + \frac{\partial S_j}{\partial q}\hat{\omega}_i^T\frac{\partial S_j^T}{\partial S_i} + \frac{\partial S_j}{\partial q}\tilde{\Gamma}_q\frac{\partial S_j^T}{\partial q} \qquad (19.28)$$

Note that, without setting $\tilde{\omega}_j = 0$, substituting the partitions of 19.28 into 19.29 gives

$$19.31 \quad \tilde{G}_j^{-1} = \frac{\partial S_j}{\partial S_i}\hat{G}_i^{-1}\frac{\partial S_j^T}{\partial S_i}$$

which is equation 19.14.

Comments

Certainly the use of 19.30 (rather than 19.31) complicates the filter. The question is, however, does it help cure the problems discussed in Section 17? Of course, some experimentation would be required with any empirical method in order to obtain satisfactory results. For example, in the Gemini program the matrix, $\tilde{G}$, was modified as

$$19.32 \quad \tilde{G}_j^{-1} = \frac{\partial S_j}{\partial S_i} \hat{\Gamma}_i \frac{\partial S_j^T}{\partial S_i} + \frac{\partial S_j}{\partial q} \tilde{\Gamma}_q \frac{\partial S_j^T}{\partial q} \qquad [3]$$

This program (19.32) was tuned to give excellent results by adjusting the elements of $\tilde{\Gamma}_q$. Based on this success, then, it was reasonable to hope that 19.30 could be used in the Apollo program as a more versatile version of 19.32. Due, perhaps, to the greater model errors in the Apollo, primarily arising from an inadequate model of SIVB venting, the method so far has not been completely successful. Further adjustment of the values of $\tilde{\Gamma}_q$ may improve the effect. Exponential downweighting of data, explained in the next section, is another method which should be considered, particularly when measurements are processed in batches. Variations of 19.32 work well when measurements are processed singly, as in the Kalman filter (Section 21).

20. EXPONENTIAL DOWNWEIGHTING OF PAST DATA

In Section 13 we derived the sequential, weighted, least-squares filter, and in Section 17 we discussed the assumption that the equations of motion model the trajectory "perfectly". For the purpose of our derivation this assumption is equivalent to saying that first-order error propagation is valid. Since the model is not perfect, however, there is some trajectory arc length beyond which the assumption does not hold. This problem can be avoided by letting the observation weight decrease exponentially with time at an appropriate rate; so then, in effect, the filter is always applied to a segment of past trajectory short enough to conform to the assumption. This method is simple to implement and adjust, and estimates do not depend on observation timing or times of processing.

The method works as follows: If R_i is the covariance matrix of a measurement vector at time t_i; t_0, the anchor time for convergence; e, the base of Napierian logorithms; and $\lambda \geq 0$, a chosen scalar constant; then the modified covariance matrix is mapped from time t_0 to t_1 as

$$\frac{\partial S_1}{\partial S_0} e^{\lambda (t_1 - t_0)} \left[e^{\lambda (t_0 - t_i)} R_i \right] \frac{\partial S_1^T}{\partial S_0} = \frac{\partial S_1}{\partial S_0} e^{\lambda (t_1 - t_i)} R_i \frac{\partial S_1^T}{\partial S_0}$$

where S_i is the vector of functionally independent trajectory parameters at time t_i. Thus the multiplier, $e^{\lambda \Delta t}$, is always used when mapping covariance over the interval Δt. We prove that with first-order approximations valid (as required by our Bayes trajectory processor) the mathematical consistency is retained. That is, if we partition a finite set of measurements into non-empty subsets for sequential processing, the final estimate is independent of the partition, the sequential order, and times of processing. The following is the first step of a proof by induction. In the last section we present an algebraic proof.

Let α be a p-element measurement vector. Then from 13.3 a better estimate at time t_1 is

$$20.1 \quad (\hat{S} - S)_1 = \left[\frac{\partial \beta^T}{\partial S_1} R^{-1} \frac{\partial \beta}{\partial S_1} \right]^{-1} \frac{\partial \beta^T}{\partial S_1} R^{-1} (\alpha - \beta)$$

Since the observations are uncorrelated in time and R is a diagonal matrix, equation 20.1 can be written

$$20.2 \qquad (\hat{S} - S)_1 = \left[\sum_{i=1}^{p} \frac{\partial \beta_i^T}{\partial S_1} R_i^{-1} \frac{\partial \beta_i}{\partial S_1} \right]^{-1} \left[\sum_{i=1}^{p} \frac{\partial \beta_i^T}{\partial S_1} R_i^{-1} (\alpha - \beta)_i \right]$$

This (20.2) is modified using exponential downweighting as

$$20.3 \qquad (\hat{S} - S)_1 = \left[\sum_{i=1}^{p} \frac{\partial \beta_i^T}{\partial S_1} e^{\lambda(t_i - t_1)} R_i^{-1} \frac{\partial \beta_i}{\partial S_1} \right]^{-1} \left[\sum_{i=1}^{p} \frac{\partial \beta_i^T}{\partial S_1} e^{\lambda(t_i - t_1)} R_i^{-1} (\alpha - \beta)_i \right]$$

and equivalently

$$20.4 \qquad (\hat{S} - S)_1 = \left[\sum_{i=1}^{k} \frac{\partial \beta_i^T}{\partial S_1} e^{\lambda(t_i - t_1)} R_i^{-1} \frac{\partial \beta_i}{\partial S_1} + \sum_{k+1}^{p} \frac{\partial \beta_i^T}{\partial S_1} e^{\lambda(t_i - t_1)} R_i^{-1} \frac{\partial \beta_i}{\partial S_1} \right]^{-1}$$

$$\left[\sum_{i=1}^{k} \frac{\partial \beta_i^T}{\partial S_1} e^{\lambda(t_i - t_1)} R_i^{-1} (\alpha - \beta)_i + \sum_{i=k+1}^{p} \frac{\partial \beta_i^T}{\partial S_1} e^{\lambda(t_i - t_1)} R_i^{-1} (\alpha - \beta)_i \right]$$

where $(0 < k < p)$

If the first k measurements were processed at time $t_0 < t_1$, we would have

$$20.5 \qquad (\hat{S} - S)_0 = \left[\sum_{i=1}^{k} \frac{\partial \beta_i^T}{\partial S_0} e^{\lambda(t_i - t_0)} R_i^{-1} \frac{\partial \beta_i}{\partial S_0} \right]^{-1} \left[\sum_{i=1}^{k} \frac{\partial \beta_i^T}{\partial S_0} e^{\lambda(t_i - t_0)} R_i^{-1} (\alpha - \beta)_i \right]$$

Now consider the expression from 20.4:

20.6 $$\sum_{i=1}^{k} \frac{\partial \beta_i^T}{\partial S_1} e^{\lambda(t_i - t_1)} R_i^{-1} \frac{\partial \beta_i}{\partial S_1}$$

$$= \frac{\partial S_0^T}{\partial S_1} e^{\lambda(t_0 - t_1)} \left[\sum_{i=1}^{k} \frac{\partial \beta_i^T}{\partial S_0} e^{\lambda(t_i - t_0)} R_i^{-1} \frac{\partial \beta_i}{\partial S_0} \right] \frac{\partial S_0}{\partial S_1}$$

$$= \frac{\partial S_0^T}{\partial S_1} e^{\lambda(t_0 - t_1)} \hat{\Gamma}_0^{-1} \frac{\partial S_0}{\partial S_1} = \tilde{\Gamma}_1^{-1} \qquad (20.5)$$

And also from 20.4:

20.7 $$\sum_{i=1}^{k} \frac{\partial \beta_i^T}{\partial S_1} e^{\lambda(t_i - t_1)} R_i^{-1} (\alpha - \beta)_i$$

$$= \frac{\partial S_0^T}{\partial S_1} e^{\lambda(t_0 - t_1)} \left[\sum_{i=1}^{k} \frac{\partial \beta_i^T}{\partial S_0} e^{\lambda(t_i - t_0)} R_i^{-1} (\alpha - \beta)_i \right]$$

$$= \frac{\partial S_0^T}{\partial S_1} e^{\lambda(t_0 - t_1)} \hat{\Gamma}_0^{-1} (\hat{S} - S)_0 \qquad (20.5)$$

$$= \frac{\partial S_0^T}{\partial S_1} e^{\lambda(t_0 - t_1)} \hat{\Gamma}_0^{-1} \frac{\partial S_0}{\partial S_1} \frac{\partial S_1}{\partial S_0} (\hat{S} - S)_0$$

$$= \tilde{\Gamma}_1^{-1} (\tilde{S} - S)_1$$

Substituting 20.6 and 20.7 into 20.4 gives

20.8
$$(\hat{S} - S)_1 = \left[\tilde{\Gamma}_1^{-1} + \sum_{i=k+1}^{p} \frac{\partial\beta_i^T}{\partial S_1} e^{\lambda(t_i - t_1)} R_i^{-1} \frac{\partial\beta_i}{\partial S_1}\right]^{-1} \left[\tilde{\Gamma}_1^{-1}(\tilde{S} - S)_1 + \sum_{i=k+1}^{p} \frac{\partial\beta_i^T}{\partial S_1} e^{\lambda(t_i - t_1)} R_i^{-1} (\alpha - \beta)_i\right]$$

This (20.8) is the sequential estimation formula, where the first k measurements were processed at time t_0 and the rest were processed at time t_1. It is equivalent to 20.3, where all p measurements were processed at time t_1. This is easily extended by induction to show that 20.3 is the final estimate at t_1 after all of α is processed, regardless of the batching partition and times of processing.

In implementing this method λ should be adjustable during tracking. The value should be large enough so that the segment of trajectory considered conforms to the model, yet small enough so that past data is not needlessly wasted. The value of λ should increase with the uncertainties in the model. For example, an earth orbit with drag and venting uncertainties would require a larger value of λ than a precisely-modeled earth-moon trajectory. Appropriate values of λ for different mission phases and vehicle configurations can be determined empirically with data from prior missions. Also λ can be made adjustable during the tracking by a manual entry in the program. Preliminary experimentation with this method showed that, when the model did not conform to the true orbit, the estimate was improved by inserting some small $\lambda > 0$. Of course downweighting vanished when $\lambda = 0$. Also the sequential estimate was the same as the estimate obtained by processing all observations in one step.

21. THE KALMAN FILTER

The theoretical derivation of the Kalman filter considers errors in the dynamic model. If we assume that the model is perfect, then the Kalman filter becomes just another algorithm for the sequential, weighted, least-squares filter already derived. We shall show this relationship and then mention some advantages of each of the two methods. [5]

The Kalman filter is derived directly from the Bayes filter (13.11) as follows: First write the Bayes filter.

$$21.1 \quad (\hat{S} - S) = \left[\tilde{\Gamma}^{-1} + \sum_i \frac{\partial \beta_i}{\partial S}^T R_i^{-1} \frac{\partial \beta_i}{\partial S}\right]^{-1} \left[\tilde{\Gamma}^{-1}(\tilde{S} - S) + \sum_i \frac{\partial \beta_i}{\partial S}^T R_i^{-1} (\alpha - \beta)_i\right] \qquad (13.11)$$

where α_i is the vector of measurements taken at time t_i, and

$$21.2 \quad \hat{\Gamma} = \left[\tilde{\Gamma}^{-1} + \sum_i \frac{\partial \beta_i}{\partial S}^T R_i^{-1} \frac{\partial \beta_i}{\partial S}\right]^{-1} \qquad (13.12)$$

Choose to process each measurement vector singly as it is received; so there is only one measurement vector in each batch. Accept the first iteration of 21.1, rather than iterating until convergence criteria are satisfied. Then 21.1 and 21.2 can be written

$$21.3 \quad (\hat{S} - \tilde{S}) = \left[\tilde{\Gamma}^{-1} + \frac{\partial \beta}{\partial S}^T R^{-1} \frac{\partial \beta}{\partial S}\right]^{-1} \left[\frac{\partial \beta}{\partial S}^T R^{-1} (\alpha - \tilde{\beta})\right] \qquad (21.1)$$

$$21.4 \quad \hat{\Gamma} = \left[\tilde{\Gamma}^{-1} + \frac{\partial \beta}{\partial S}^T R^{-1} \frac{\partial \beta}{\partial S}\right]^{-1} \qquad (21.2)$$

where α now denotes one of the α_i with the subscript dropped and $\tilde{\beta} = \beta(\tilde{S})$ is the measurement vector computed as a function of the a priori state. The Kalman filter is another algorithm for computing 21.3 and 21.4 as follows.

Consider the following three equations:

$$\left.\begin{aligned}
&21.5 \quad K \equiv \tilde{\Gamma}\frac{\partial\beta}{\partial S}^{T}\left[\frac{\partial\beta}{\partial S}\tilde{\Gamma}\frac{\partial\beta}{\partial S}^{T} + R\right]^{-1} \\
&21.6 \quad \hat{S} = \tilde{S} + K(\alpha - \tilde{\beta}) \\
&21.7 \quad \hat{\Gamma} = \left[I - K\frac{\partial\beta}{\partial S}\right]\tilde{\Gamma}
\end{aligned}\right\} \text{Kalman filter}$$

To show that 21.3 and 21.6 are equivalent we prove the following identity:

$$\left[\tilde{\Gamma}^{-1} + \frac{\partial\beta}{\partial S}^{T} R^{-1}\frac{\partial\beta}{\partial S}\right]^{-1}\frac{\partial\beta}{\partial S}^{T} R^{-1} \overset{?}{=} \tilde{\Gamma}\frac{\partial\beta}{\partial S}^{T}\left[\frac{\partial\beta}{\partial S}\tilde{\Gamma}\frac{\partial\beta}{\partial S}^{T} + R\right]^{-1} = K$$

Multiplying on the left by $\left[\tilde{\Gamma}^{-1} + \frac{\partial\beta}{\partial S}^{T} R^{-1}\frac{\partial\beta}{\partial S}\right]$ and on the right by $\left[\frac{\partial\beta}{\partial S}\tilde{\Gamma}\frac{\partial\beta}{\partial S}^{T} + R\right]$ gives

$$\frac{\partial\beta}{\partial S}^{T} R^{-1}\left[\frac{\partial\beta}{\partial S}\tilde{\Gamma}\frac{\partial\beta}{\partial S}^{T} + R\right] = \left[\tilde{\Gamma}^{-1} + \frac{\partial\beta}{\partial S}^{T} R^{-1}\frac{\partial\beta}{\partial S}\right]\tilde{\Gamma}\frac{\partial\beta}{\partial S}^{T}$$

$$\frac{\partial\beta}{\partial S}^{T} R^{-1}\frac{\partial\beta}{\partial S}\tilde{\Gamma}\frac{\partial\beta}{\partial S}^{T} + \frac{\partial\beta}{\partial S}^{T} = \frac{\partial\beta}{\partial S}^{T} R^{-1}\frac{\partial\beta}{\partial S}\tilde{\Gamma}\frac{\partial\beta}{\partial S}^{T} + \frac{\partial\beta}{\partial S}^{T}$$

Define $M \equiv \frac{\partial\beta}{\partial S}^{T}$ of dimension m x n, $(m > n)$. This is commonly the case. For example, in the Kalman powered flight filter for the LM lunar ascent and descent the measurement vector has four elements and the state vector, twenty-one. [12]

To show that 21.4 and 21.7 are equivalent, prove the following identity:

$$\left[\tilde{\Gamma}^{-1} + MR^{-1}M^{T}\right]^{-1} \overset{?}{=} \left\{I - \tilde{\Gamma}M\left[M^{T}\tilde{\Gamma}M + R\right]^{-1}M^{T}\right\}\tilde{\Gamma}$$

Multiplying on the left by $\left[\tilde{\Gamma}^{-1} + MR^{-1}M^T\right]$ and on the right by $\tilde{\Gamma}^{-1}$ gives

$$\tilde{\Gamma}^{-1} = \tilde{\Gamma}^{-1} + MR^{-1}M^T - M\left[M^T\tilde{\Gamma}M + R\right]^{-1}M^T - MR^{-1}M^T\tilde{\Gamma}M\left[M^T\tilde{\Gamma}M + R\right]^{-1}M^T$$

$$0 = M\left[R^{-1} - \left(M^T\tilde{\Gamma}M + R\right)^{-1} - R^{-1}M^T\tilde{\Gamma}M\left(M^T\tilde{\Gamma}M + R\right)^{-1}\right]M^T$$

Now multiply on the left by M^{-1} and on the right by M^{-T}. (M has a left inverse, and M^T has a right inverse.)

$$0 = R^{-1} - \left(M^T\tilde{\Gamma}M + R\right)^{-1} - R^{-1}M^T\tilde{\Gamma}M\left(M^T\tilde{\Gamma}M + R\right)^{-1}$$

Multiply on the right by $M^T\tilde{\Gamma}M + R$.

$$0 = R^{-1}\left(M^T\tilde{\Gamma}M + R\right) - I - R^{-1}M^T\tilde{\Gamma}M$$

$$0 = R^{-1}\left(M^T\tilde{\Gamma}M + R - M^T\tilde{\Gamma}M\right) - I$$

$$0 = 0$$

Comparing equations 21.1 and 21.2 with 21.5, 21.6, 21.7 we can summarize some of the major differences in the weighted, least-squares (Bayes) and Kalman filters.

The Bayes filter iterates until convergence, but the Kalman accepts the first iteration. The iteration of the Bayes filter solves a system of non-linear equations by producing a sequence of linear approximations converging to the final solution. So if the Bayes iterates more than once, it normally produces a better answer than the Kalman. We say "normally" because if the first guess is not close enough, it is possible to have non-convergence or convergence to the wrong answer. [13]

The Bayes filter can collect measurements and process them in batches at arbitrary times, whereas the Kalman must process each measurement separately at the time of the measurement. If the Kalman observations are close together so that the propagation time interval is very small, it may be difficult to modify the covariance matrix as a function of time. This is because the modification is too small to appear in the computer. The Bayes filter avoids this problem by choosing anchor times sufficiently far apart. The problem with the Kalman filter can be resolved, however, by modifying the covariance matrix at predetermined time intervals, rather than at the observation time.

The Bayes filter is particularly well adapted to estimating free-flight trajectories of long duration, where the observations actually are received in batches. Then each batch can be processed as it is received to update the state vector. The Kalman filter is particularly desirable when the observations are coming in continually and the trajectory characteristics are such that point-by-point processing of data is required, e.g., the LM powered flight processor. [12]

The Bayes filter requires inversion of matrices with order of the state vector; the Kalman, with order of the measurement vector. So the Kalman is very useful in avoiding inversion of large order matrices. For example, in the Kalman filter, LM, powered flight processor [12] the state vector has 21 elements; the measurement vector, 4 elements.

See Battin [6] for a discussion of trajectory estimation using the Kalman filter.

22. CORRELATED DOPPLER MEASUREMENTS

Up till now the sequential filters have been derived assuming that the measurement errors are uncorrelated in time. Depending on the particular problem, it becomes considerably more difficult to develop sequential filters for time correlated measurements and this subject alone provides a sizeable area for study [7]. We need not be concerned with this theory now, however, because all our measurements are assumed to be uncorrelated except for the very simple case of doppler (range-rate) observations discussed below.

From equation 14.44 the doppler frequency at time t_j is computed as

22.1 $$f_j = (\omega_3 + b) + \frac{\omega_4 \nu}{c(t_i - t_k)} \left[(\rho_1 + \rho_2)_i - (\rho_1 + \rho_2)_k\right]$$

where

22.2 $$t_j = \frac{t_i + t_k}{2} \qquad (t_i - t_k > 0)$$

Define

22.3 K_i actual measurement of cycle count at time, t_i

22.4 δK_i zero-mean, random error in K_i, δK_i and δK_k uncorrelated $(i \neq k)$

Then

22.5 $$K_i - K_k = f_j(t_i - t_k) + \delta K_i - \delta K_k$$

and

22.6 $$\frac{K_i - K_k}{t_i - t_k} = f_j + \frac{\delta K_i - \delta K_k}{t_i - t_k}$$

The actual measurements here are K_i and K_k, and the pseudomeasurement is

$$\frac{K_i - K_k}{t_i - t_k}$$

From now on, for simplification, consider the pseudomeasurement to be $K_i - K_k$. This does not affect our discussion of correlated measurement errors.

Let k = i - 1 in 22.5 and consider the following sequence of pseudomeasurements at times, $t_1 < t_2 < \ldots < t_n$,

22.7 $\{K_i - K_{i-1} = f_j(t_i - t_{i-1}) + \delta K_i - \delta K_{i-1}\}$ $\quad (i, j = 1, \ldots, n)$

where

22.8 $$E(\delta K_i \delta K_j) = \begin{cases} \sigma^2 & (i = j) \\ 0 & (i \neq j) \end{cases}$$

It follows that the covariance matrix associated with 22.7 is

22.9 $$\underset{(n \times n)}{R} = E\left\{\begin{bmatrix} \delta K_1 - \delta K_0 \\ \vdots \\ \delta K_n - \delta K_{n-1} \end{bmatrix} [(\delta K_1 - \delta K_0) \cdots (\delta K_n - \delta K_{n-1})]\right\}$$

$$= \sigma^2 \begin{bmatrix} 2 & -1 & . & \emptyset \\ -1 & 2 & \ddots & -1 \\ & \ddots & \ddots & \\ \emptyset & & -1 & 2 \end{bmatrix} \quad (n \times n)$$

Since this (22.9) is not a diagonal matrix, the pseudomeasurements cannot be processed sequentially by the methods of Section 13. We solve this problem as follows.

Define

22.10 $\Delta K_i \equiv (\omega_3 + b)(t_i - t_0) + \frac{\omega_4 \nu}{c} (\rho_1 + \rho_2)_i$

22.11 $J_{i-1} \equiv \Delta K_{i-1} + \delta K_{i-1}$

Substituting 22.10 into 22.7 gives

22.12 $\{K_i - K_{i-1} = \Delta K_i - \Delta K_{i-1} + \delta K_i - \delta K_{i-1}\}$ $\quad (i = 1, \ldots, n)$

Substituting 22.11 into 22.12 gives the sequence of pseudomeasurements modeled as

22.13 $\{K_i - K_{i-1} = \Delta K_i - J_{i-1} + \delta K_i\}$ $(i = 1, \ldots, n)$

Adjoin J_{i-1} to the state vector as an element to be estimated with the processing of $K_i - K_{i-1}$; then the errors in this sequence (22.13) are uncorrelated, and the covariance matrix is

$$R = \sigma^2 I \qquad (n \times n)$$

By combining 22.11 and 22.12 again the improved estimate, $\hat{J}_{i-1}$, is propagated to become the a priori estimate, $\tilde{J}_i$, as

22.17 $\tilde{J}_i = (K_i - K_{i-1}) + \hat{J}_{i-1}$

and from 22.11

22.18 $\tilde{J}_0 = \Delta K_0$

This way of processing the pseudomeasurements was presented to show how it can be done, but it is really clumsy compared to the following equivalent method which uses the actual measurements [12].

Combining 22.10 and 22.5 we can model the sequence of actual measurements as

22.19 $\{K_i = \Delta K_i - \Delta K_k + K_k + \delta K_i - \delta K_k\}$ $(k < i = 1, \ldots, n)$

If we choose $k = 0$, the covariance matrix associated with 22.19 is

22.20 $$\underset{(n \times n)}{R} = \sigma^2 \begin{bmatrix} 2 & -1 & \cdots & -1 \\ -1 & 2 & \ddots & \vdots \\ \vdots & \ddots & \ddots & -1 \\ -1 & \cdots & -1 & 2 \end{bmatrix} \qquad (n \times n)$$

If $k = i - 1$, then the covariance matrix associated with 22.19 is the same as 22.9. Neither matrix, however, is diagonal.

Define

22.21 $I_k \equiv \Delta K_k - K_k + \delta K_k$

Substituting this into 22.19, each measurement is modeled as

22.22 $K_i = \Delta K_i - I_k + \delta K_i \qquad (k < i = 1, \ldots, n)$

Adjoin I_k to the state vector as an element to be estimated; then the errors in measurements 22.22 are uncorrelated and the covariance matrix is $R = \sigma^2 I$.

Substituting 22.21 in to 22.22 with $k = i$ gives

22.23 $I_i = I_k$

Therefore, we can write

$$I_k = I_0$$

which is a constant to be re-estimated at each sequential step. From 22.21 the a priori value for the first step is

22.24 $\tilde{I}_0 = \Delta K_0 - K_0$

and measurements are modeled as

22.25 $K_i = \Delta K_i - I_0 + \delta K_i \qquad (i = 1, \ldots, n) \qquad (22.22)$

Another way of arriving at 22.25 is as follows.

Replace the first member only of sequence 22.12 by 22.22, as

22.26
$$\left\{ \begin{aligned} K_1 \quad &= \Delta K_1 - I_0 + \delta K_1 \\ K_2 - K_1 &= \Delta K_2 - \Delta K_1 + \delta K_2 - \delta K_1 \\ &\vdots \\ K_n - K_{n-1} &= \Delta K_n - \Delta K_{n-1} + \delta K_n - \delta K_{n-1} \end{aligned} \right.$$

Consider I_0 as a trajectory parameter to be estimated; then the covariance matrix associated with 22.26 is

$$22.27 \quad R = \sigma^2 \begin{bmatrix} 1 & -1 & & & 0 \\ -1 & 2 & \ddots & & \\ & \ddots & \ddots & \ddots & \\ & & \ddots & \ddots & -1 \\ 0 & & & -1 & 2 \end{bmatrix}$$

This is just like 22.9 except for the first diagonal element.

Define

$$22.28 \quad \delta\beta_i \equiv \delta K_i - \delta K_{i-1} \qquad (i = 1, \ldots, n), \ (\delta K_0 = 0)$$

The quadratic form associated with 22.28 is

$$22.29 \quad 2\varphi = [\delta\beta_1 \cdots \delta\beta_n] \frac{1}{\sigma^2} \begin{bmatrix} 1 & -1 & & & 0 \\ -1 & 2 & \ddots & & \\ & \ddots & \ddots & \ddots & \\ & & \ddots & \ddots & -1 \\ 0 & & & -1 & 2 \end{bmatrix}^{-1} \begin{bmatrix} \delta\beta_1 \\ \vdots \\ \delta\beta_n \end{bmatrix} \qquad (22.27)$$

which cannot be processed sequentially, since the matrix is not diagonal. Consider the following:

$$22.30 \quad \begin{bmatrix} 1 & -1 & & & 0 \\ -1 & 2 & \ddots & & \\ & \ddots & \ddots & \ddots & \\ & & \ddots & \ddots & -1 \\ 0 & & & -1 & 2 \end{bmatrix}^{-1} = \left\{ \begin{bmatrix} 1 & & & & 0 \\ -1 & 2 & & & \\ & \ddots & \ddots & & \\ & & \ddots & \ddots & \\ 0 & & & -1 & 2 \end{bmatrix} \begin{bmatrix} 1 & -1 & & & 0 \\ & 2 & \ddots & & \\ & & \ddots & \ddots & \\ & & & \ddots & -1 \\ 0 & & & & 2 \end{bmatrix} \right\}^{-1}$$

$$= \begin{bmatrix} 1 & 1 & \cdots & 1 \\ & 1 & & 1 \\ & & \ddots & \vdots \\ 0 & & & 1 \end{bmatrix} \begin{bmatrix} 1 & & & 0 \\ 1 & 1 & & \\ \vdots & \vdots & \ddots & \\ 1 & 1 & \cdots & 1 \end{bmatrix}$$

Define

22.31 $$\begin{bmatrix}\delta y_1\\ \cdot\\ \cdot\\ \delta y_n\end{bmatrix} \equiv \begin{bmatrix}1 & & & 0\\ 1 & 1 & & \\ \vdots & \vdots & \ddots & \\ 1 & 1 & \cdots & 1\end{bmatrix}\begin{bmatrix}\delta\beta_1\\ \cdot\\ \cdot\\ \cdot\\ \delta\beta_n\end{bmatrix}$$

Then 22.29 can be written

22.32 $$2\varphi = [\delta y_1 \cdots \delta y_n]\,\frac{1}{\sigma^2}\begin{bmatrix}1 & & 0\\ & 1 & \\ 0 & & \ddots\ 1\end{bmatrix}\begin{bmatrix}\delta y_1\\ \cdot\\ \cdot\\ \delta y_n\end{bmatrix}$$

and the $\{\delta y_i\}$ can be processed sequentially by the methods of Section 13. But applying the transformation of 22.31 to 22.26, we see that

$$y_i = K_i = \Delta K_i - I_0 + \delta K_i$$

as in 22.25. We have arrived at this result in different ways to show the possibility of using ingenuity to develop sequential estimators for correlated measurement errors. See Blum [7] for a more complete discussion.

23. ALGEBRAIC PROOF OF SEQUENTIAL PROPERTIES

A method was presented in Section 20 for downweighting data exponentially as a function of time in the Bayes filter. An explanation of the method and an algebraic proof that the sequential properties are retained were given in a previous paper [9]. This section is essentially a copy of the paper [9] with some minor improvements.

Our purpose is to define and prove a procedure for downweighting past data within the Bayes processor. We do this by first reviewing a derivation of the Bayes equation without downweighting. Here we introduce some new definitions to simplify writing the equation. Then we present an algebraic proof of mathematical consistency. Finally we extend this proof to include the case where data is downweighted by the prescribed formula. Following the proof is a brief discussion of some practical aspects of implementation.

Definitions

23.1	S_i	True value of the state vector (vector of functionally independent trajectory parameters) at time t_i
23.2	$\tilde{S}_i$	A priori estimate of S_i
23.3	$\hat{S}_i$	Improved estimate of S_i
23.4	$(\tilde{S} - S)_i$	Small deviation of a priori estimate of state vector from the true value at time t_i
23.5	$(\hat{S} - S)_i$	Small deviation of improved estimate of state vector from the true value at time t_i
23.6	E	Statistical expectation operator
23.7	$\tilde{\Gamma}_i \equiv E\left[(\tilde{S} - S)(\tilde{S} - S)^T\right]_i$	A priori state covariance matrix at time t_i
23.8	$\hat{\Gamma}_i \equiv E\left[(\hat{S}-S)(\hat{S}-S)^T\right]_i$	Improved estimate of state covariance matrix at time t_i
23.9	α_i	An observation vector at time t_i

23.10	$\beta_i \equiv \beta\ (S_i)$	An observation vector whose elements are determined as functions of S_i. Assume dimension $\beta_i \leq$ dimension S_i. This causes no loss in generality, because if dimension $\beta_i >$ dimension S_i, then β_i can be partitioned into subvectors conforming to the assumption.
23.11	$(\alpha - \beta)_i$	An observation residual at time t_i
23.12	$R_i \equiv E\left[(\alpha-\beta)(\alpha-\beta)^T\right]_i$	Observation covariance matrix at time t_i
23.13	$\frac{\partial S_k}{\partial S_j} \equiv \frac{\partial S_{t_k}}{\partial S_{t_j}}$,	i.e., subscript k on a partial derivative implies subscript t_k
23.14	$a_{ij} \equiv \frac{\partial S_j}{\partial \beta_i} (\alpha - \beta)_i$	
23.15	$W_{ij} \equiv \frac{\partial \beta_i^T}{\partial S_j} R_i^{-1} \frac{\partial \beta_i}{\partial S_j}$	
23.16	$M_j \equiv \{a_{ij}\}$	
23.17	2φ	Quadratic form
23.18	$\Phi \equiv \left(\frac{\partial \varphi}{\partial S_i}\right)^T$	
23.19	λ	A chosen scalar $(\lambda \geq 0)$
23.20	e	Base of Napierian logarithms
23.21	T	As a superscript, indicates the transpose of a matrix or vector

Derivation

There is a 1-1 correspondence between the elements of M_k and M_j as

23.22 $$a_{ik} = \frac{\partial S_k}{\partial S_j} a_{ij} \qquad (23.14,\ 23.16)$$

and also

$$23.23 \qquad W_{ik} = \frac{\partial S_j}{\partial S_k}^T \quad W_{ij} \quad \frac{\partial S_j}{\partial S_k}$$

Consider that we have a finite set of observations. Assuming a normal distribution of estimate errors about the true values and using the method of maximum liklihood, the quadratic form to be minimized with respect to S_j, the vector to be estimated, is

$$23.24 \qquad 2\varphi = \sum_i (\alpha - \beta)_i^T R_i^{-1} (\alpha - \beta)_i = \sum_i a_{ij}^T \, W_{ij} a_{ij}$$

Note that definition 23.10 implies that

$$\frac{\partial \beta_i}{\partial S_j} \frac{\partial S_j}{\partial \beta_i} = I$$

If an estimate of the state vector exists, it is included in the set of observations. For example

$(\alpha - \beta)_j^T \, R_j^{-1} \, (\alpha - \beta)_j = (\tilde{S} - S)_j^T \, \tilde{\Gamma}_j^{-1} \, (\tilde{S} - S)_j$, and we can keep the equation in the simple form (23.24).

Neglecting terms higher than first order,

$$23.25 \quad \Phi \equiv \left(\frac{\partial \varphi}{\partial S_j} \right)^T = -\sum_i \frac{\partial \beta_i}{\partial S_j}^T \, R_i^{-1} \, (\alpha - \beta)_i$$

$$= -\sum_i W_{ij} a_{ij} \qquad . \qquad (23.24)$$

$$23.26 \quad \frac{\partial \Phi}{\partial S_j} = \sum_i \frac{\partial \beta_i}{\partial S_j}^T \, R_i^{-1} \, \frac{\partial \beta_i}{\partial S_j} = \sum_i W_{ij} \qquad (i = 1, \ldots, n)$$

Assume that our best estimate of Φ, $\hat{\Phi} = 0$. Also assume the matrix (23.26) is positive definite, assuring that the solution to $\hat{\Phi} = 0$ will minimize (23.24) and also that $\left(\frac{\partial \Phi}{\partial S_j}\right)^{-1}$ exists. Since the solution to $\hat{\Phi} = 0$ minimizes the quadratic form and it is desirable to express the partial derivatives with respect to the true state, we expand $\hat{\Phi}$ in a Taylor series about Φ rather than the usual expansion of Φ about $\hat{\Phi}$.

23.27 $\hat{\Phi} = \Phi + \frac{\partial \Phi}{\partial S_j} (\hat{S} - S)_j = 0$ and the Bayes estimation equation is

23.28 $$(\hat{S} - S)_j = -\left(\frac{\partial \Phi}{\partial S_j}\right)^{-1} \Phi = \left[\sum_i W_{ij}\right]^{-1} \left[\sum_i W_{ij} a_{ij}\right] \qquad (23.27)$$

and assuming observation errors are serially uncorrelated

23.29 $$\hat{\Gamma}_j = E\left[(\hat{S} - S)(\hat{S} - S)^T\right]_j = \left[\sum_i W_{ij}\right]^{-1} \qquad (23.8,\ 23.28)$$

Now we show that if we partition a finite set of observations into non-empty subsets for sequential processing by 23.28, the final $\hat{S}_A$ at t_A is independent of the partition, the sequential order, and times of processing.

Consider an algebraic system $(M_j, *)$ where

23.30 $$M_j = \{a_{ij}\} \qquad (23.16)$$

Also consider $\tilde{S}_j$ and $\hat{S}_j$ as observations so that

23.31 $$\{(\tilde{S} - S)_j, (\hat{S} - S)_j\} \subset M_j$$

Let * be a binary operation such that

23.32 $$a_{1j} * a_{2j} = \left[\sum_{i=1}^{2} W_{ij}\right]^{-1} \left[\sum_{i=1}^{2} W_{ij} a_{ij}\right] = (\hat{S} - S)_j \qquad (23.28)$$

Note that $\left[\sum_{i=1}^{2} W_{ij}\right]^{-1}$ exists either as a true inverse or a pseudoinverse. See Deutsch [4].

Define this operation (23.32) to be the processing on M_j of the observations taken at times t_1 and t_2.

23.33 Clearly $*$ is commutative.

Show that $*$ is associative, i.e., that

$$(a_{1j} * a_{2j}) * a_{3j} = a_{1j} * (a_{2j} * a_{3j})$$

23.34 $$(a_{1j} * a_{2j}) * a_{3j} = (\hat{S} - S)_j * a_{3j} \qquad (23.32)$$

23.35 $$= \left[\hat{\Gamma}_j^{-1} + W_{3j}\right]^{-1} \left[\hat{\Gamma}_j^{-1} (\hat{S} - S)_j + W_{3j} a_{3j}\right] \qquad (23.28)$$

23.36 $$= \left[\sum_{i=1}^{3} W_{ij}\right]^{-1} \left[\sum_{i=1}^{3} W_{ij} a_{ij}\right] = (\hat{\hat{S}} - S)_j$$

Note that $(\hat{\hat{S}} - S)_j$ in 23.36 has the double carat superscript to distinguish it from the $(\hat{S} - S)_j$ of 23.34. Also,

23.37 $$a_{1j} * (a_{2j} * a_{3j}) = (a_{2j} * a_{3j}) * a_{1j} \qquad (23.33)$$

Evaluating the right side of 23.37 is the same as evaluating 23.34 after permutation of the "i" subscripts, and the result is again 23.36.

Show the isomorphism,

23.38 $$(M_j, *) \cong (M_k, *)$$

The 1-1 correspondence, $a_{ij} \longleftrightarrow a_{ik}$, is clear from the mapping. (23.22)

Also

23.39 $$\frac{\partial S_k}{\partial S_j} a_{1j} * \frac{\partial S_k}{\partial S_j} a_{2j} = a_{1k} * a_{2k} = (\hat{S} - S)_k$$

$$(?) = \frac{\partial S_k}{\partial S_j} (\hat{S} - S)_j = \frac{\partial S_k}{\partial S_j} (a_{1j} * a_{2j})$$

23.40 $$\frac{\partial S_k}{\partial S_j} (\hat{S} - S)_j = \frac{\partial S_k}{\partial S_j} \left[\sum_{i=1}^{2} W_{ij} \right]^{-1} \frac{\partial S_k^T}{\partial S_j} \frac{\partial S_j^T}{\partial S_k} \left[\sum_{i=1}^{2} W_{ij} a_{ij} \right]$$

23.41 $$= \left[\sum_{i=1}^{2} W_{ik} \right]^{-1} \left[\sum_{i=1}^{2} W_{ik} a_{ik} \right] = (\hat{S} - S)_k$$

It follows that if we partition a finite set of observations into non-empty subsets for sequential processing:

a. Because of the isomorphism the image of the process is always on M_A, and the final $\hat{S}_A$ is the same as if all the processing were on M_A.

b. Since * is associative, $\hat{S}_A$ is independent of the partition.

c. Since * is commutative, $\hat{S}_A$ is independent of the sequential order.

Next we extend the proof to include the method for exponential downweighting of data. (Section 20.) Re-define

23.42 $$W_{ij} \equiv e^{\lambda(t_i - t_j)} \frac{\partial \beta_i^T}{\partial S_j} R_i^{-1} \frac{\partial \beta_i}{\partial S_j}$$

and map

23.43 $$W_{ik} = \frac{\partial S_j^T}{\partial S_k} e^{\lambda(t_j - t_k)} R_{ij} \frac{\partial S_j}{\partial S_k} \qquad (0 \le \lambda)$$

Show that

23.44 $(M_j\ ,\ *) \cong (M_k\ ,\ *)$ still holds.

23.45 $$\frac{\partial S_k}{\partial S_j}\ (\hat{S} - S)_j$$

$$= \frac{\partial S_k}{\partial S_j}\left[\sum_{i=1}^{2} W_{ij}\right]^{-1} \frac{\partial S_k}{\partial S_j}^{T} e^{\lambda(t_k - t_j)}\ e^{\lambda(t_j - t_k)}\ \frac{\partial S_j}{\partial S_k}^{T}\left[\sum_{i=1}^{2} W_{ij}\, a_{ij}\right]$$

23.46 $$= \left[\sum_{i=1}^{2} W_{ik}\right]^{-1}\left[\sum_{i=1}^{2} W_{ik}\, a_{ik}\right] = (\hat{S} - S)_k$$

The rest of the definitions, proof, and results still hold.

Implementation

Assume that we have a set of m observations taken at times $\{t_1, t_2, \ldots, t_m\}$. Also at t_0 we have a priori estimates $\tilde{S}_0$ and $\tilde{\Gamma}_0$ of the state vector and its covariance. We wish to obtain $\hat{S}_m$ and $\hat{\Gamma}_m$ at t_m, the time of the last observation. This is a natural situation as we proceed along a trajectory. As we have shown, the time of processing is arbitrary as long as the result is mapped to t_m. We choose to estimate $\hat{S}_0$ and $\hat{\Gamma}_0$ at t_0 and map these to $\hat{S}_m$ and $\hat{\Gamma}_m$ at t_m. Rewrite the following equations:

23.47 $$(\hat{S} - S)_j = \left[\sum_i W_{ij}\right]^{-1}\left[\sum_i W_{ij}\, a_{ij}\right] \qquad (23.28)$$

23.48 $$W_{ij} = e^{\lambda(t_i - t_j)} \frac{\partial \beta_i^T}{\partial S_j} R_i^{-1} \frac{\partial \beta_i}{\partial S_j} \qquad (23.42)$$

23.49 $$a_{ij} = \frac{\partial S_j}{\partial \beta_i}\ (\alpha - \beta)_i \qquad (23.14)$$

Substituting 23.48 and 23.49 into 23.47,

$$23.50 \quad (\hat{S} - S)_j = \left[\sum_{i=0}^{m} \frac{\partial \beta_i^{T}}{\partial S_j} e^{\lambda (t_i - t_j)} R_i^{-1} \frac{\partial \beta_i}{\partial S_j} \right]^{-1} \left[\sum_{i=0}^{m} \frac{\partial \beta_i^{T}}{\partial S_j} e^{\lambda (t_i - t_j)} R_i^{-1} (\alpha - \beta)_i \right]$$

Letting $t_j = t_0$ and expressing the a priori values

$$23.51 \quad (\hat{S} - S)_0 = \left[\tilde{\Gamma}_0^{-1} + \sum_{i=1}^{m} \frac{\partial \beta_i^{T}}{\partial S_0} e^{\lambda (t_i - t_0)} R_i^{-1} \frac{\partial \beta_i}{\partial S_0} \right]^{-1} \left[\tilde{\Gamma}_0^{-1} (\tilde{S} - S)_0 + \sum_{i=1}^{m} \frac{\partial \beta_i^{T}}{\partial S_0} e^{\lambda (t_i - t_0)} R_i^{-1} (\alpha - \beta)_i \right]$$

From 23.51 and 23.29

$$23.52 \quad \hat{\Gamma}_0 = \left[\tilde{\Gamma}_0^{-1} + \sum_{i=1}^{m} \frac{\partial \beta_i^{T}}{\partial S_0} e^{\lambda (t_i - t_0)} R_i^{-1} \frac{\partial \beta_i}{\partial S_0} \right]^{-1}$$

Equation 23.51 is a Taylor series expansion valid in terms of any vector S_0 in the region of convergence about $\hat{S}_0$. To find $\hat{S}_0$ we set $S_0 = S_{0n}$, which is the current best estimate of $\hat{S}_0$, and then iterate until convergence.

$$23.53 \quad S_{0n+1} = S_{0n} + \left[\tilde{\Gamma}_0^{-1} + \sum_{i=1}^{m} \frac{\partial \beta_i^{T}}{\partial S_{0n}} e^{\lambda (t_i - t_0)} R_i^{-1} \frac{\partial \beta_i}{\partial S_{0n}} \right]^{-1} \left[\tilde{\Gamma}_0^{-1} (\tilde{S} - S_n)_0 + \sum_{i=1}^{m} \frac{\partial \beta_i^{T}}{\partial S_{0n}} e^{\lambda t_i - t_0)} R_i^{-1} (\alpha - \beta_n)_i \right]$$

If the convergence criteria are met after n iterations, then consider that

23.54 $\hat{S}_0 = S_n$

Then $\hat{S}_0$ is the initial conditions for integrating the equations of motion from t_0 to t_m to obtain $\hat{S}_m$. By inspection of 23.51, 23.52, and 23.54, after the n iterations consider

23.55 $$\hat{\Gamma}_0 = \left[\tilde{\Gamma}_0^{-1} + \sum_{i=1}^{m} \frac{\partial \beta_i^T}{\partial S_{0n}} e^{\lambda (t_i - t_0)} R_i^{-1} \frac{\partial \beta_i}{\partial S_{0n}} \right]^{-1}$$

Then $\hat{\Gamma}_0$ is mapped to $\hat{\Gamma}_m$ as

23.56 $$\hat{\Gamma}_m = \frac{\partial S_m}{\partial S_0} e^{\lambda (t_m - t_o)} \hat{\Gamma}_0 \frac{\partial S_m}{\partial S_0}^T$$

Inspection of the above shows that the well-known Bayes is the same as before, the only alteration being the method of downweighting data.

REFERENCES

1. Hohn, F. E., Elementary Matrix Algebra. New York: The Macmillan Company, 1958.

2. Ericksen, W. L. and Colson, H., Class Notes on Analytical Dynamics. Dayton, Ohio: AFIT, 1961.

3. Ditto, F. H., "Non-Linear Trajectory Estimation in Real Time for Project Gemini." Real Time Systems Seminar, Houston, Texas: IBM Corporation, 1966.

4. Deutsch, R., Estimation Theory. Englewood Cliffs, N. J.: Prentice-Hall, Inc., 1965.

5. Goodyear, W. H., Class Notes on Trajectory Estimation. Houston, Texas: IBM Corporation, 1964.

6. Battin, R. H., Astronautical Guidance. New York: McGraw-Hill Book Company, 1964.

7. Blum, M., "Best Linear Unbiased Estimation by Recursive Methods", J. Soc. Indust. Appl. Math., 14, No. 1 (Jan., 1966), 167-180.

8. Kaplan, W., Advanced Calculus. Reading, Mass.: Addison-Wesley Publishing Company, Inc., 1952.

9. Rich, R. G., "A Method for Downweighting Data with Respect to Time in a Bayes Trajectory Processor", RTCC Math. Dev. and Support, 12-022 (Aug., 1968). Houston, Texas: IBM Corporation.

10. Schiesser, E. R., deVezin, H. G., Savely, R. T., and Oles, M. J., Basic Equations and Logic for the Real-Time Ground Navigation Program for the Apollo Lunar Landing Mission, MSC Internal Note No. 68-FM-100 (Apr. 15, 1968). Houston, Texas: Manned Spacecraft Center.

11. Flanagan, P. F. and Austin, G. A., RTCC Requirements for Mission G: Landing Site Determination Using Rendezvous Radar and Optical Observations, MSC Internal Note No. 69-FM-92 (May 29, 1969). Houston, Texas: Manned Spacecraft Center.

12. Lear, W.M., deVezin, H.G. Jr., Wylie, A.D., and Schiesser, E.R., RTCC Requirements for Mission G: MSFN Tracking Data Processor for Powered Flight Lunar Ascent/Descent Navigation, MSC Internal Note No. 69-FM-36 (Feb. 7, 1969). Houston, Texas: Manned Spacecraft Center.

13. Henrici, P., Discrete Variable Methods in Ordinary Differential Equations. New York: John Wiley and Sons, Inc., 1962.