RO-S-68-2

A TUTORIAL DERIVATION OF RECRUSIVE WEIGHTED LEAST SQUARES STATE-VECTOR ESTIMATION THEORY (KALMAN THEORY)

(Special Report)

by

James S. Pappas

AUGUST 1968

U. S. ARMY TEST AND EVALUATION COMMAND
ANALYSIS AND COMPUTATION DIRECTORATE
DEPUTY FOR NATIONAL RANGE OPERATIONS
WHITE SANDS MISSILE RANGE, NEW MEXICO

DISTRIBUTION OF THIS DOCUMENT IS UNLIMITED

A TUTORIAL DERIVATION OF RECRUSIVE WEIGHTED LEAST SQUARES STATE-VECTOR ESTIMATION THEORY (KALMAN THEORY)

(Special Report)

by

James S. Pappas

Approved by
GUENTHER HINTZE
Director

August 1968

U. S. ARMY TEST AND EVALUATION COMMAND

ANALYSIS AND COMPUTATION DIRECTORATE
DEPUTY FOR NATIONAL RANGE OPERATIONS
WHITE SANDS MISSILE RANGE, NEW MEXICO

CONTENTS

ABSTRACT

This is one of a series of tutorial reports deriving the discrete recursive Kalman estimation equations. The series proceeds from unweighted to weighted least squares parameter estimation in a vector space setting to the stage-wise updating, to time varying parameter or state variable estimation. The derivations have been stage-wise tutorial in an attempt to make the theory accessible to the "non-specialist" in optimization theory.

INTRODUCTION

Modern state-vector recursive estimation theory (in the sense of Kalman) has been derived from many viewpoints: conditional probability, orthogonal projections, least squares, minimum variance, maximum likelihood, etc. The individual user of the theory may understand the derivation of the equations obtained by any one of the above theories or possibly by a number of approaches depending upon the individuals background.

This report is the third of a tutorial series deriving the estimation equations in a vector space weighted lease squares setting. The essential areas for understanding the theory via this approach are:

1. Unweighted Least Squares Parameter-Vector Estimation and the Variance-of-the-Estimate Matrix.

2. Discrete Matrix Recursive Methods apply to (1) for Real Time (on line) Computer Processing.

3. Weighted Least Squares Parameter Estimation and Variance-of-the-Estimate Matrix for Correlated Noise.

4. Discrete Matrix Recursive Methods applied to (3) for Real Time Computer Processing.

5. Recursive Weighted Least Squares State-Vector Estimation Theory (Kalman Theory).

Items 1 and 2 are completed and published in Reference 4. Item 3 is completed and published in Reference 6. Item 4 is not completed yet. Item 5 is the contents of the current report.

This report is restricted to full-rank matrices so that the reader may understand the basic principles of the theory. Non-full rank matrices processed recursively require a knowledge and background in Psuedo Inverses, and many more sophisticated linear algebra concepts. The minimization criterion is derived by taking the partial derivative of a scalar with respect to a pxk weighting matirx-the classical gradient approach.

The derivations can be achieved completely algebraicly in terms of orthogonal projections relating to spaces and sub-spaces. Many sophisticated paper exist in the literature using this approach. The algebraic approach may not be as readily accessible to the engineer trained in classical mathematics.

NOTATION

The notations used in the report is an effort to blend the notation of Friedman for inner-products and dyadic products with the current journal-literature on vector-spaces, psuedo-inverses, state-vectors, etc.

$\underset{p\times k}{X}$ Capital letters designate matrices of size p rows and k columns.

$|k\rangle$: when p = 1, the matrix is called a column vector, and we use Friedmans symbol to distinguish this matrix.

$\langle p|x$: when k = 1, the matrix is a row vector of dimension p.

$\langle p|x \quad y|p\rangle$ "inner-product" or scalar product of two vectors.

$y|p\rangle\langle p|x$: "outer-product" or dyadic product of two vectors.

$\underset{p\times k}{X} = [x|p\rangle_1, \ldots, x|p\rangle_k]$. Matrix X partitioned into a row k-tuple of column vectors from a p-space.

$\underset{p\times k}{X} = \begin{bmatrix} 1\langle k|x \\ \vdots \\ p\langle k|x \end{bmatrix}$ Matrix X partitioned into a p-column tuple of row vectors from a k-space

x small x is a scalar

x^i scalar from a column vector

x_j scalar from a row vector

Scalar here is a "real field" element.

Consider the system of two vector equations

$$|x(k+1)\rangle = \Phi(k+1, k)|x(k)\rangle + |f(k)\rangle + |u(k)\rangle$$

and

$$|z(k)\rangle = \underset{m\times p}{H}(k)|x(k)\rangle + |v(k)\rangle$$

where:

$\underset{p}{x(k)}$, $\underset{p}{x(k+1)}$ are p-dimensional column vectors describing the states at stage k and stage k+1.

$\Phi(k+1, k)$ is a pxp state transition matrix.

$\underset{p}{f(k)}$ is a p-dimensional deterministic forcing vector for which we can write a vector function.

$\underset{p}{u(k)}$ is a p-dimensional uncertainty or noise vector, it is the composite of the random noises and the variables we fail to model.

$\underset{m}{z(k)}$ is the m-dimensional observation vector, m is less than or equal to p.

$\underset{mxp}{H(k)}$ is the known matrix describing how the state vector is functionally related to the observation vector (if the instruments were noise free).

$\underset{m}{v(k)}$ is an m-dimensional additive instrument noise vector.

I. ONE PARAMETER FAMILY OF DISCRETE TRAJECTORIES

This section considers the one parameter family of i trajectories generated by i different values of the initial condition parameter $x_i(1)$. Each of the i trajectories consists of a sequence of k different points or stages.

We consider

(1) The homogeneous case

(2) The deterministically forced case.

Homogeneous System

Consider the system of scalar equations

$$x(k+1) = \Phi(k+1, k)x(k) + f(k) + u(k) \tag{1}$$

$$z(k) = h(k)x(k) + v(k) \tag{2}$$

where the following constraints apply

$$f(k) = u(k) = v(k) = 0 \tag{3}$$

$$h(k) = h_o \text{ a constant} \tag{4}$$

$$\phi(k+1,k) = \phi_o \text{ a constant} \tag{5}$$

then equation (1) and (2) become

$$x(k+1) = \phi_0\, x(k) \tag{6}$$

$$z(k) = h_0\, x(k) \tag{7}$$

If further, we consider the measurement to be exact

$$z(k) = x(k) \tag{8}$$

or

$$h_0 = 1 \tag{9}$$

then we need only consider

$$x(k+1) = \phi_0 x(k) \tag{10}$$

Case i. The most elementary case occurs when

$$\phi_0 = 1, \tag{11}$$

then

$$x(k+1) = x(k) \tag{12}$$

and for values of k

$$x(2) = x(1) \tag{13}$$

$$x(3) = x(2) = x(1) \tag{14}$$

$$x(k) = x(1) \tag{15}$$

hence we can plot

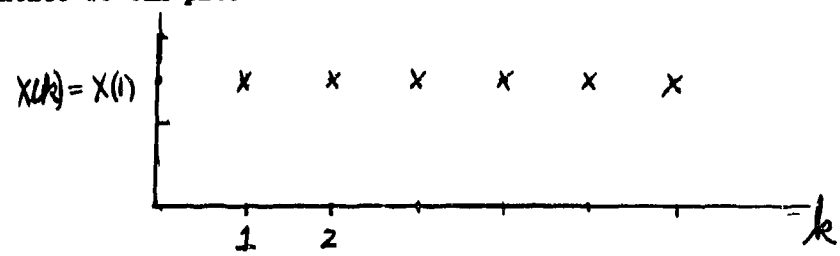

Fig. (1)

If we vary the initial condition we obtain a sequence of trajectories

$$x_i(k) = x_i(1) \tag{16}$$

for

$$x_1(1),\ x_2(1) \cdots x_i(1),\ x_{i+1}(1) \cdots$$

as shown in Fig. (2)

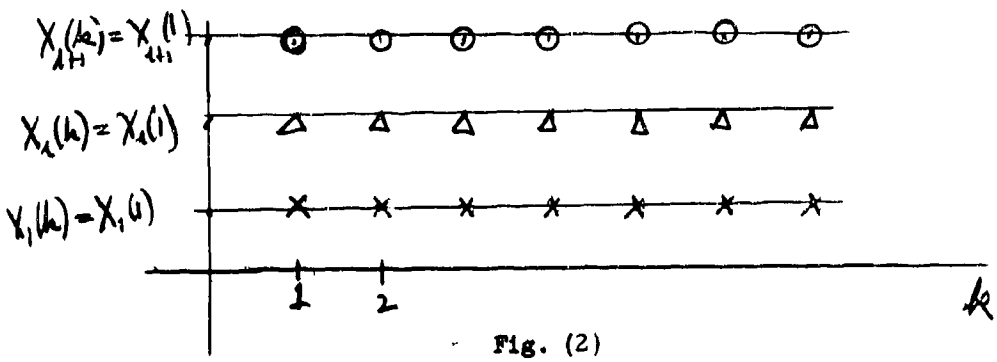

Fig. (2)

For any constant ϕ_0, we obtain

$$
\begin{aligned}
x(2) &= \phi_0\, x(1) \\
x(3) &= \phi_0\, x(2) = \phi_0(\phi_0\, x(1)) = \phi_0^2\, x(1) \\
&\vdots \\
x(k) &= \phi_0^{k-1}\, x(1).
\end{aligned}
\tag{17}
$$

If ϕ_0 is less than one, the trajectory of Equation ( ), for a given initial condition decreases, for example

$$
\begin{aligned}
\phi_0 &= 1/2 \\
x(2) &= \tfrac{1}{2}\, x(1) \\
x(3) &= (\tfrac{1}{2})^2\, x(1) = \tfrac{1}{4}\, x(1) \\
x(4) &= \tfrac{1}{8}\, x(1) \\
&\vdots \\
x(k) &= (\tfrac{1}{2})^{k-1}\, x(1)
\end{aligned}
\tag{18}
$$

The sequence of k vectors are

$$
\begin{pmatrix} 1 \\ x(1) \end{pmatrix} \quad \begin{pmatrix} 2 \\ \frac{1}{2}\, x(1) \end{pmatrix} \quad \begin{pmatrix} 3 \\ \frac{1}{4}\, x(1) \end{pmatrix} \quad \cdots \quad \begin{pmatrix} k \\ (\frac{1}{2})^{k-1} x(1) \end{pmatrix} \tag{19}
$$

which plot as shown in Fig. (3)

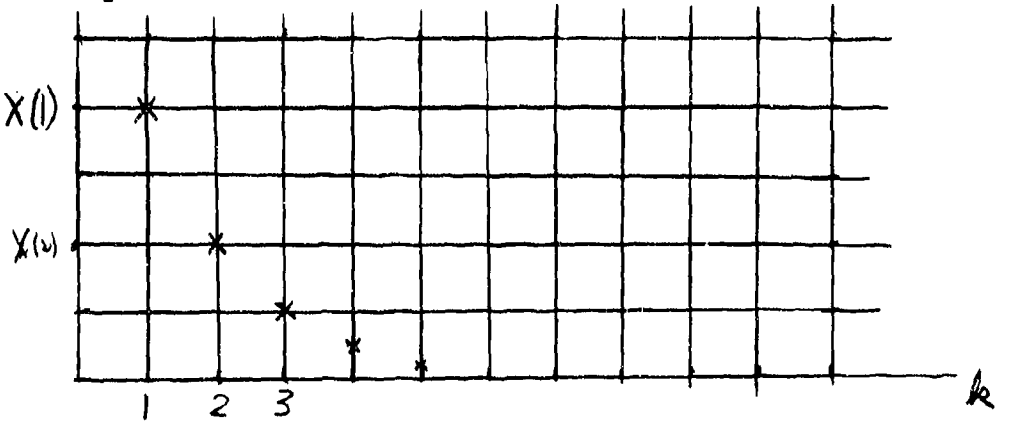

Fig. (3) Contracting or Decaying Sequence $\phi_0 = 1/2$

The first noisy observation is

$$z_{jn}(1) = h(1)\, x_j(1) + v_n(1)$$

Using (3) and (6) in (4)

$$\tilde{z}_{jn}(1, 0) = h(1)[x_j(1) - \hat{x}_j(1, 0)] + v_n(1). \tag{7}$$

Using (7) in (5)

$$\hat{x}_{jn}(1, 1) = \hat{x}_j(1, 0) + w(1)h(1)[x_j(1) - \hat{x}_j(1, 0)] + w(1)v_n(1) \tag{8}$$

$$= [1 - w(1)h(1)]\, \hat{x}_j(1, 0) + w(1)h(1)x_j(1) + w(1)\, v_n(1).$$

We now define <u>two new errors</u> as differences of the previous variables.

$$x_j(1) - \hat{x}_j(1, 0) = \tilde{x}_j(1, 0) \tag{9}$$

and

$$x_j(1) - \hat{x}_{jn}(1, 1) = \tilde{x}_{jn}(1, 1). \tag{10}$$

The variable $x_j(1)$ is the actual (but unknown) process state variable that exists on trajectory j. The variable $\hat{x}_{jn}(1,1)$ is the estimate of $x_j(1)$, having used the first noisy observation.

Hence $\tilde{x}_{jn}(1, 1)$ is the error in our estimate of the state at stage 1 based on one measurement.

From equation (3) we see that if our initial guess $\hat{x}_j(1, 0)$ is wrong, then $\hat{z}_j(1, 0)$ will also be wrong and after our observation z(1) comes in and our computed error $\tilde{z}_j(1, 0)$ of equation (4) is large, then by equation (5) we will have a correction term which is the product $(w(1)\tilde{z}_j(1, 1))$ of a large term times the weight value w(1).

Continuing with the derivation of the weights, we obtain the error term by equation (8) in equation (11) as

$$\tilde{x}_{jn}(1, 1) = x_j(1) - \hat{x}_j(1, 1) = x_j(1) - w(1)h(1)x_j(1) \tag{11}$$

$$- [(1-w(1)h(1)]\, \hat{x}_j(1, 0) - w(1)v_n(1).$$

or

$$\tilde{x}_{jn}(1, 1) = [1-w(1)h(1)][x_j(1) - \hat{x}_j(1, 0)] - w(1)v_n(1). \tag{12}$$

Using equation (9) in equation (12)

$$\tilde{x}_{jn}(1,1) = [1 - w(1)h(1)]\,\tilde{x}_j(1, 0) - w(1)\, v_n(1). \tag{13}$$

If we square the error of equation (13)

$$\tilde{x}^2_{jn}(1, 1) = [1-w(1)h(1)]^2\tilde{x}^2_j(1, 0) - 2w(1)v_n(1)[1-w(1)h(1)]\hat{x}_j(1, 0) \tag{14}$$
$$+ w^2(1)v_n{}^2(1).$$

If we take the partial derivative of equation (14) with respect to w(1) and equate to zero we obtain

$$\frac{\partial \tilde{x}^2_{jn}(1, 1)}{\partial w(1)} = 2[1 - w(1)h(1)]\,(-h(1))\,\tilde{x}^2_j(1, 0) \tag{15}$$
$$- 2v_n(1)[1 - w(1)h(1)]\hat{x}_j(1, 0)$$
$$+ 2w(1)\, v^2_n(1) = 0$$

Taking the expected value over all experiments in which j and n are varying we obtain

$$E\left\{\frac{\partial \tilde{x}^2_{jn}(1,1)}{\partial w(1)}\right\} = -[1 - w(1)h(1)]\; E\left\{\tilde{x}^2_j(1,0)\right\} h(1) \tag{16}$$
$$+ w(1)\; E\left\{v^2_n(1)\right\} = 0$$

Define the variances

$$E\left\{\tilde{x}^2_j(1, 0)\right\} = \sigma_{\tilde{x}\tilde{x}}(1, 0) = p(1, 0) \tag{17}$$

$$E\left\{v^2_n(1)\right\} = \sigma_{vv}(1) \tag{18}$$

We follow a number of current Kalman papers and papers about Kalman Estimation and designate the variance in state as p(1, 0).

The cross term of equation (15) under the expectation operator is zero when the following conditions hold,

$$E\left\{[1 - w(1)h(1)]\; v_n(2)\; \hat{x}_j(1, 0)\right\} \tag{19}$$
$$= [1 - w(1)h(1)]\; E\left\{v_n(1)\; \hat{x}_j(1, 0)\right\}$$

We now assume that our initial guess of $\hat{x}(1,0)$ is independent of $v_n(1)$ as n ranges over all allowable values, that is

$$E\left\{v_n(1)\; \hat{x}_j(1, 0)\right\} = 0. \tag{20}$$

Under these assumptions equation (16) and (18) become

$$0 = -[1 - w(1)h(1)]h(1)\ p(1, 0) + w(1)\sigma_{vv}(1) \tag{21}$$

or

$$-h(1)p(1, 0) + w(1)h^2(1)\ p(1, 0) + w(1)\sigma_{vv}(1) = 0 \tag{22}$$

or

$$w(1)[h^2(1)\ p(1, 0) + \sigma_{vv}(1)] = h(1)\ p(1, 0) \tag{23}$$

$$\boxed{w(1) = h(1)\ p(1, 0)[h^2(1)\ p(1, 0) + \sigma_{vv}(1)]^{-1}} \tag{24}$$

The value of p(1, 0) which by equation (17) is

$$p(1, 0) = E\left\{[x_j(1) - \hat{x}_j(1, 0)]^2\right\} = \text{guess} \tag{25}$$

can be obtained by guess or by a "learning process", or by a-priori knowledge about the system.

The variance of the estimate of state at stage 1 based on the observation at stage one can also be obtained by equation (14) as

$$E\left\{\hat{x}^2_{jn}(1, 1)\right\} = [1 - w(1)h(1)]^2\ E\left\{\hat{x}^2_j(1, 0)\right\} + w^2(1)\ E\left\{v^2_n(1)\right\} \tag{26}$$

when the cross terms and zero, that is

$$E\left\{v_n(1)\ \hat{x}_j(1, 0)\right\}\ 2w(1)[1 - w(1)h(1)] = 0. \tag{27}$$

The variance terms of equation ( ) will be denoted as

$$\sigma_{\tilde{x}\tilde{x}}(1, 1) = E\left\{\tilde{x}^2_{jn}(1, 1)\right\} \equiv p\ (1, 1) \tag{28}$$

$$E\left\{v^2_n(1)\right\} = \sigma_{vv}(1) \tag{29}$$

The p(1, 1) designation is in keeping with the now near classical notation.

Expanding the terms of equation (26)

$$\begin{aligned} p(1, 1) &= [1 - w(1)h(1)]^2\ p(1, 0) + w^2(1)\sigma_{vv}(1) \\ &= [1 - 2w(1)h(1) + w^2(1)h^2(1)]\ p(1, 0) + w^2(1)\ \sigma_{vv}(1) \\ &= p(1, 0) - 2w(1)h(1)\ p(1, 0) \\ &\quad + w^2(1)[h^2(1)\ p(1, 0) + \sigma_{vv}(1)]. \end{aligned} \tag{30}$$

Consider the last term in equation (30) and use equation (24) for $w^2(1)$, then

$$w^2(1)[h^2(1)\,p(1,0) + \sigma_{vv}(1)] \tag{31}$$
$$= h^2(1)\,p^2(1,0)[h^2(1)\,p(1,0) + \sigma_{vv}(1)]^{-2}\,[h^2(1)\,p(1,0) + \sigma_{vv}(1)]$$
$$= h^2(1)\,p^2(1,0)[h^2(1)\,p(1,0) + \sigma_{vv}(1)]^{-1}$$
$$= h(1)\,p(1,0)\left\{h(1)\,p(1,0)\,[h^2(1)\,p(1,0) + \sigma_{vv}(1)]^{-1}\right.$$

Replacing the bracket term of equation (31) by w(1) in equation (24) we obtain

$$w^2(1)[h^2(1)\,p(1,0) + \sigma_{vv}(1)] = h(1)\,p(1,0)\,w(1) \tag{32}$$

Using (32) in equation (30) we obtain

$$p(1,1) = p(1,0) - 2w(1)\,h(1)\,p(1,0) + h(1)\,p(1,0)\,w(1) \tag{33}$$
$$= p(1,0) - w(1)\,h(1)\,p(1,0).$$

$$\boxed{p(1,1) = p(1,0)[1 - w(1)h(1)]}$$

also using equation (24) in equation (33) we can write p(1, 1) as

$$\boxed{p(1,1) = p(1,0) - p(1,0)h(1)[h(1)p(1,0)h(1) + \sigma_{vv}(1)]^{-1}p(1,0)h(1)} \tag{34}$$

Equation (34) is the scalar variance (one dimensional uncertainty ellipsoid) of the estimate of the state using all of the current observations.

We can now predict what the next state variable value should be by propogating forward with the known dynamics as

$$\hat{x}(2,1) = \Phi(2,1)\,\hat{x}(1,1) + f(1). \tag{35}$$

The prediction of the next observation is

$$\hat{z}(2,1) = h(2)\,\hat{x}(2,1). \tag{36}$$

The elliposoid of uncertainty of the predicted state is

$$p(2,1) = E\left\{\tilde{x}^2(2,1)\right\} \tag{37}$$

where

$$\tilde{x}(2,1) = x(2) - \hat{x}(2,1) \tag{38}$$

Using equation (1) and equation (35) in equation (38)

$$\hat{\tilde{x}}(2,1) = \phi(2, 1)\, \tilde{x}(1, 1) + u(1) \tag{39}$$

$$\tilde{x}^2(2, 1) = \phi(2, 1)\, \tilde{x}^2(1, 1)\phi(2, 1) + 2\phi(2, 1)\, \tilde{x}(1, 1)\, u(1) + u^2(1) \tag{40}$$

The expected value over all experiments of equation (40) is

$$E\left\{\tilde{x}^2(2, 1)\right\} = p(2, 1) = \phi(2, 1)\, p(1, 1)\phi(2, 1) + \sigma_{uu}(1) \tag{41}$$

where

$$\sigma_{uu}(1) = E_n\left\{u_n^2(1)\right\} \tag{42}$$

Cross term statistical independence assumptions are

$$E\left\{\tilde{x}(1, 1)\, u(1)\right\} = 0. \tag{43}$$

etc.

In conclusion, at stage one, we do:

Given

$$\sigma_{vv}(1)$$

guess or estimate

$$p(1, 0) \text{ and } \hat{x}(1, 0)$$

compute

$$\hat{z}(1, 0) = h(1)\, \hat{x}(1, 0) \tag{44}$$

Measure and compute <u>error</u>

$$\tilde{z}(1, 0) = z(1) - \hat{z}(1, 0) \tag{45}$$

Compute weight w(1) equation ( )

$$w(1) = h(1)\, p(1, 0)[h(1)p(1, 0)h(1) + \sigma_{vv}(1)]^{-1}. \tag{46}$$

Update or correct state-estimate

$$\hat{x}(1, 1) = \hat{x}(1, 0) + w(1)\, \tilde{z}(1, 0) \tag{47}$$

Compute ellipsoid of uncertainty of state estimation (48)

$$p(1, 1) = p(1,0)[1 - w(1)h(1)]$$

Predict or propogate one stage into future with known dynamics the following three variables:

Predict next stage.

$$\hat{x}(2, 1) = \phi(2, 1)\, \hat{x}(1, 1) + f(1) \tag{49}$$

Predict next observation.

$$\hat{z}(2, 1) = h(2)\, \hat{x}(2, 1) \tag{50}$$

Predict ellipsoid of uncertainty in state

$$p(2, 1) = \phi(2, 1)\, p(1, 1)\, \phi(2, 1) + \sigma_{uu}(1). \tag{51}$$

Wait for stage k = 2 and second observation to come in.

Stage k = 2.

The second observation is taken at stage 2 and is

$$z_{jn}(2) = h(2)\, x_j(2) + v_n(2) \tag{52}$$

where, as before, $z_{jn}(2)$ is known, but $x_j(2)$ and $v_n(2)$ are unknown.

We compute the error signal

$$\tilde{z}_{jn}(2, 1) = z_{jn}(2) - \hat{z}_{jn}(2, 1)\ , \tag{53}$$

The sub-scripts j, n can now be dropped for simplicity of notation, where it is henceforth understood that $\hat{x}(2, 1)$ for example means the best estimate of the true jth trajectory based on the sequence of instrument noises $\{v_n\}$.

Using equation (50) and equation (52) in equation (53)

$$\tilde{z}(2, 1) = h(2)[x(2) - \hat{x}(2, 1)] + v(2) \tag{54}$$

we now have computed the error in the observation, hence we can correct our estimate of what the state variable should be based on the use of the second observation, that is

$$\hat{x}(2, 2) = \hat{x}(2, 1) + w(2)\, \tilde{z}(2, 1). \tag{55}$$

As before, we now seek the "feed-back" or correction weight w(2) at stage 2. Before deriving the w(2) expression, use (54) in (55)

$$\hat{x}(2, 2) = \hat{x}(2, 1) + w(2)h(2)[x(2) - \hat{x}(2, 1)] + w(2)\, v(2) \tag{56}$$

$$\hat{x}(2, 2) = [1 - w(2)h(2)]\, \hat{x}(2, 1) + w(2)\, h(2)\, x(2) + w(2)\, v(2). \tag{57}$$

From the error in state estimation at stage 2 using the second observation as before

$$\tilde{x}(2,2) = x(2) - \hat{x}(2, 2) = [1-w(2)\, h(2)]\tilde{x}(2,1) + w(2)\, v(2) \tag{58}$$

Multiply equation (58) by itself to obtain the square of the error term as

$$\begin{aligned}\tilde{x}^2(2, 2) &= [1 - w(2)\, h(2)]^2\, \tilde{x}^2(2, 1) \\ &\quad + 2w(2)\, v(2)[1 - w(2)\, h(2)]\, \tilde{x}(2, 1) \\ &\quad + w^2(2)\, v^2(2).\end{aligned} \tag{59}$$

Observe that the error and the square of the error in equation (59) is a function of w(2).

In order to select a w(2) which will minimize the square of the error we take the partial derivative of equation (59) with respect to w(2) and equate this "gradient" term to zero, hence

$$\begin{aligned}\frac{\partial \tilde{x}^2(2, 2)}{\partial w(2)} &= -2h(2)[1 - w(2)\, h(2)]\, \tilde{x}^2(2, 1) \\ &\quad + 2v(2)[1 - w(2)h(2)]\, \tilde{x}(2, 1) \\ &\quad + 2w(2)\, v(2)\, \big(-h(2)\big)\, \tilde{x}(2, 1) \\ &\quad + 2w(2)\, v^2(2) = 0\end{aligned} \tag{60}$$

Since we want w(2) to be the same regardless of how many times we repeat the experiment; or, from a "single-test" stand-point, w(2) should be selected to hold regardless of the unknown sequence driving the two variables at this stage $u_j(2)$ and $v_n(2)$. Consequently we take the expected value over all admissible vectors $\{u\}_j$ and $\{v\}_n$ and obtain

$$E\left\{\frac{\partial \tilde{x}^2(2,2)}{\partial w(2)}\right\} = [1 - w(2)\, h(2)]\, (-h(2))\, p(2, 1) + w(2)\sigma_{vv}(2) = 0. \tag{61}$$

The assumption of statistical independence of the variables $\tilde{x}(2, 1)$ and v(2) is again assumed

$$E\,\{\tilde{x}(2, 1)\, v(2)\} = 0 \tag{62}$$

also the assumption about $\{v\}_n$

$$E_n\left\{\begin{pmatrix} v_n(1) \\ v_n(2) \end{pmatrix} [v_n(1),\, v_n(2)]\right\} \tag{63}$$

$$= E_n\left\{\begin{bmatrix} u_n(1)\, v_n(1) & v_n(1)\, v_n(2) \\ v_n(2)\, v_n(1) & v_n(2)\, v_n(2) \end{bmatrix}\right\} = \begin{bmatrix} \sigma_{vv}(1) & 0 \\ 0 & \sigma_{vv}(2) \end{bmatrix}$$

The assumptions of equation (63) can be relaxed but then one needs a deeper knowledge of multi-linear algebras, matrix-packaging and partitioning, matrix psuedo-inverse, etc. Hence such "highly-colored" or correlated noise cases will not be discussed in this paper. The majority of published papers on Kalman theory make the implied Gaussian assumptions implied by equation (63) for arbitrary large k. Computer storage problems for correlated noise for large k are also a problem.

Solving equation (60) for w(2)

$$w(2)[h(2)\ p(2,\ 1)\ h(2) + \sigma_{vv}(2)] = h(2)\ p(2,\ 1) \tag{64}$$

or

$$\boxed{w(2) = h(2)\ p(2,1)\ [h(2)\ p(2,\ 1)\ h(2) + \sigma_{vv}(2)]^{-1}} \tag{65}$$

The second weight can be computed since $h(2)$, $\sigma_{vv}(2)$ are assumed known and $p(2, 1)$ was computed at stage 1.

We can now derive the expression for $p(2, 2)$, using equation (65) and taking the expected value over all experiments

$$p(2,\ 2) = E\ \{\tilde{x}^2(2,\ 2)\} = [1 - w(2)h(2)]^2\ p(2,\ 1) + w^2(2)\sigma_{vv}(2), \tag{66}$$

with assumptions

$$E\ \{v(2)\ \tilde{x}(2,\ 1)\} = 0. \tag{67}$$

Expanding equation (66)

$$p(2,\ 2) = [1 - 2w(2)\ h(2) + w^2(2)\ h^2(2)]\ p(2,\ 1) + w^2(2)\sigma_{vv}(2) \tag{68}$$

$$p(2,\ 2) = p(2,\ 1) - 2w(2)\ h(2)\ p(2,\ 1) \tag{69}$$

$$+ w^2(2)\ [h(2)\ p(2,\ 1)\ h(2) + \sigma_{vv}(2)]$$

Consider the last term of equation (69) using equation (65) for w(2)

$$w^2(2)\ [h(2)\ p(2,\ 1)\ h(2) + \sigma_{vv}(2)] \tag{70}$$

$$= h^2(2)\ p^2(2,\ 1)[h^2(2)\ p(2,\ 1) + \sigma_{vv}(2)]^{-2}[h^2(2)p(2,1) + \sigma_{vv}(2)]$$

$$= h^2(2)\ p^2(2,\ 1)\ [h^2(2)\ p(2,\ 1) + \sigma_{vv}(2)]^{-1}$$

and by equation (65) for w(2) we obtain

$$w^2(2)[h(2)\, p(2,1)\, h(2) + \sigma_{vv}(2)] = h(2)\, p(2,1)\, w(2) \tag{71}$$

Using equation (71) in equation (69)

$$p(2,2) = p(2,1) - 2w(2)\, h(2)\, p(2,1) + h(2)\, p(2,1)\, w(2) \tag{72}$$

or

$$p(2,2) = p(2,1) - p(2,1)\, h(2)\, w(2) \tag{73}$$

$$p(2,2) = p(2,1)\,[1 - w(2)\, h(2)] \tag{74}$$

Equation (73) can also be written as by equation (65) in equation (73)

$$p(2,2) = p(2,1) - p(2,1)\,[h^2(2)p(2,1) + \sigma_{vv}(2)]^{-1}\, h(2)p(2,1) \tag{75}$$

We can now predict at stage k = 3 the state

$$\hat{x}(3,2) = \phi(3,2)\, \hat{x}(2,2) + f(2) \tag{76}$$

and the observation

$$\hat{z}(3,2) = h(3)\, \hat{x}(3,2) \tag{77}$$

and the state-variance term

$$p(3,2) = E\{\tilde{x}^2(3,2)\}. \tag{78}$$

By equation (1) and equation (76)

$$x(3) = \phi(3,2)\, x(2) + f(2) + u(2) \tag{79}$$

$$\hat{x}(3,2) = \phi(3,2)\, \hat{x}(2,2) + f(2) \tag{80}$$

and subtracting

$$x(3) - \hat{x}(3,2) = \tilde{x}(3,2) = \phi(3,2)\, \tilde{x}(2,2) + u(2) \tag{81}$$

Squaring equation (81)

$$\tilde{x}^2(3,2) = \phi(3,2)\, \tilde{x}^2(2,2)\phi(3,2) + 2\phi(3,2)\, \tilde{x}(2,2)\, u(2) + u^2(2). \tag{82}$$

Taking the expected value

$$E\{\hat{x}^2(3, 2)\} = \phi(3, 2)\, E\{\hat{x}^2(2, 2)\}\, \phi(3, 2) + E\{u^2(2)\} \tag{83}$$

or

$$p(3, 2) = \phi(3, 2)\, p(2,2)\, \phi(3, 2) + \sigma_{uu}(2) \tag{84}$$

Summarizing the steps at stage k = 2, then we:

measure and compute error

$$\tilde{z}(2, 1) = z(2) - \hat{z}(2, 1) \tag{85}$$

compute weight w(2) by equation (65)

$$w(2) = h(2)\, p(2, 1)[h(2)\, p(2, 1)\, h(2) + \sigma_{vv}(2)]^{-1} \tag{86}$$

compute corrected estimate of state

$$\hat{x}(2, 2) = \hat{x}(2, 1) + w(2)\, \tilde{z}(2, 1) \tag{87}$$

compute variance of state equation (74)

$$p(2, 2) = p(2, 1) - w(2)\, h(2)\, p(2, 1) \tag{88}$$

Predict (update) state via dynamics equation (80)

$$\hat{x}(3, 2) = \phi(3, 2)\, \hat{x}(2, 2) + f(2) \tag{89}$$

Predict observation at next stage

$$\hat{z}(3, 2) = h(3)\, \hat{x}(3, 2) \tag{90}$$

Predict next stage state-variance

$$p(3, 2) = \phi(3, 2)\, p(2, 2)\, \phi(3, 2) + \sigma_{uu}(2) \tag{91}$$

wait for next stage or third observation to arrive.

Stage k.

The derivations of the equations will not be repeated for stage k, the relations will be based on the mathematical process of reasoning by analogy. The treatment of the multi-variable or vector case will derive the relations at stage k, but will not develop the stage-wise logic at k = 1 and k = 2.

We have available from previous stage predictions

$\hat{x}(k, k-1)$
$\hat{z}(k, k-1)$
$p(k, k-1)$

and stored $\sigma_{vv}(k)$, $\sigma_{uu}(k)$

Measure and compute error

$$\tilde{z}(k, k-1) = z(k) - \hat{z}(k, k-1) \tag{92}$$

Compute w(k)

$$w(k) = h(k)\, p(k, k-1)[h^2(k)\ \ p(k, k-1) + \sigma_{vv}(k)]^{-1} \tag{93}$$

Compute corrected state estimate

$$\hat{x}(k, k) = \hat{x}(k, k-1) + w(k)\, \tilde{z}(k, k-1) \tag{94}$$

Compute variance

$$p(k, k) = p(k, k-1)\, [1 - w(k)\, h(k)] \tag{95}$$

Predict next state

$$\hat{x}(k + 1, k) = \phi(k + 1, k)\, \hat{x}(k, k) + f(k) \tag{96}$$

Predict next observation

$$\hat{z}(k + 1, k) = h(k+1)\, \hat{x}(k+1, k) \tag{97}$$

Predict variance of state

$$p(k+1, k) = \phi(k+1, k)\, p(k, k)\, \phi(k+1, k) + \sigma_{uu}(k). \tag{98}$$

We define the noise variances in the notation of the many Kalman oriented papers, that is

$$\sigma_{uu}(k) = q(k) \tag{99}$$

$$\sigma_{vv}(k) = r(k) \tag{100}$$

The three familiar equations can be written as

$$\begin{aligned}\hat{x}(k+1, k+1) &= \phi(k+1, k)\, \hat{x}(k, k) \\ &+ p(k+1, k)\, h(k+1)[h(k+1)\, p(k+1, k)\, h(k+1) + r(k)]^{-1} \\ &\quad [z(k+1) - h(k+1)\, \phi(k+1, k)\, \hat{x}(k, k)]\end{aligned} \tag{101}$$

$$p(k+1, k) = \phi(k+1, k)\, p(k, k)\, \phi(k+1, k) + q(k) \tag{102}$$

$$p(k+1, k+1) = p(k+1, k) - p(k+1, k)\, h(k+1) \tag{103}$$
$$[h(k+1)\, p(k+1, k)\, h(k+1) + r(k)]^{-1} h(k+1)\, p(k+1, k)$$

We shall now define in words the meanings at stage k of the variables and rewrite the equations using the distinguishing j and n.

$$x_j(k+1) = \phi(k+1, k)\, x_j(k) + f(k) + u_j(k) \tag{104}$$

$$z_{jn}(k) = h(k)\, x_j(k) + v_n(k) \tag{105}$$

$$\hat{x}_{jn}(k+1, k+1) = \phi(k+1, k)\, \hat{x}_{jn}(k, k)$$
$$+ p(k+1,k)h(k)[h^2(k+1)\, p(k+1) + r(k)]^{-1}$$
$$[z_{jn}(k+1) - h(k+1)\phi(k+1,k)\, \hat{x}_{jn}(k, k)] \tag{106}$$

$$p(k+1, k) = \phi(k+1, k)\, p(k, k)\phi(k+1, k) \tag{107}$$

$$p(k+1, k+1) = p(k+1, k) - p(k+1, k)\, h(k+1) \tag{108}$$
$$[h^2(k+1)\, p(k+1, k) + r(k)]^{-1} h(k+1)\, p(k+1, k)$$

$x(k) = x_j(k)$ is the true (unknown) value of the process state at stage k as a result of the unknowns $u_j(1), \ldots u_j(k)$ forcing the system.

$z_j(k) = z_{jn}(k)$ is measurement of the true noise process $x_j(k)$ with additive unknown measurement noise $v_n(k)$.

$\hat{x}_{jn}(k, k) = \hat{x}(k, k)$ is the best estimate of the state at stage k of the jth trajectory based on past observations up to stage k, that is recursively we have used noisy

$$z_{jn}(1), z_{jn}(2) \ldots z_{jn}(k)$$

made noisy by $v_n(1), \ldots u_n(k)$.

$\hat{x}_{jn}(k+1, k) = \hat{x}(k+1, k)$ is the best estimate of the state of the jth trajectory at stage k+1, based on observations only up to k. Also interpreted as the prediction of the state at next stage k+1, based on current stage k and past measurements.

III. VECTOR ESTIMATION EQUATIONS

This section derives the Kalman Estimation Equations for the multivariable case using matrix analysis methods. The derivation techniques are the same as used in the previous scalar case. The essential difference lies in the minimization methods. The variance of the estimate in state for the scalar case is a scalar valued function of a scalar argument w(k). For the vector case, the trace of the variance matrix of the estimate of state is likewise a scalar-valued function, but a function of a matrix of p rows and m columns W(k). The minimization of a scalar-valued function with respect to this matrix.

One can arrive at the equations via strictly algebraic concepts of orthogonal projection matrices etc., in which one does not have to enter into discussions of partial derivaties, continuity of continous variables and gradients. Since the majority of expected readers are assumed to be more familiar with the least-squares criterion via gradients, this report will stick strictly with this method.

The general linearized vector equations are

$$x(k+1)_{(p)} = \underset{p}{\Phi}(k+1,\ k)\ x(k)_{(p)} + \underset{pxg}{B(k)}\, f(k)_{(g)} + \underset{pxq}{N(k)}\, u(k)_{(q)} \qquad (1)$$

$$z(k)_{(m)} = \underset{mxp}{H(k)}\ x(k)_{(p)} + v(k)_{(m)} \qquad (2)$$

The deterministic k-varying vector $f(k)_{(g)}$ is of dimension g less than or equal to p, and gets distributed or cross-coupled into all p state variables $x(k+1)_{(p)}$ via the functional relations of $\underset{pxg}{B(k)}$.

The same statements apply to the noise input vector $u(k)_{(q)}$.

The reader should keep in mind the families of trajectories accurately described by the j and n indices, that is

$$x(k+1)_{(p)_j} = \underset{p}{\Phi}(k+1,\ k)\ x(k)_{(p)_j} + B(k)f(k)_{(g)} + N(k)u(k)_{(q)_j} \qquad (3)$$

$$z_{jn}(k)_{(m)} = H(k)\ x(k)_{(p)_j} + v(k)_{(m)_n} \qquad (4)$$

As before, the accurate descriptions designated by j and n will be dropped for simplicity of representation.

The equations are developed as a "recursive process" or an "on-line" processor; that is, as the observations "role in" the mechanized computer-estimator utilizes the data, and discards it or stores it on tape or what have you. All past data is sequentially accumulated in the "memory of the

III. VECTOR ESTIMATION EQUATIONS

This section derives the Kalman Estimation Equations for the multivariable case using matrix analysis methods. The derivation techniques are the same as used in the previous scalar case. The essential difference lies in the minimization methods. The variance of the estimate in state for the scalar case is a scalar valued function of a scalar argument w(k). For the vector case, the trace of the variance matrix of the estimate of state is likewise a scalar-valued function, but a function of a matrix of p rows and m columns W(k). The minimization of a scalar-valued function with respect to this matrix.

One can arrive at the equations via strictly algebraic concepts of orthogonal projection matrices etc., in which one does not have to enter into discussions of partial derivaties, continuity of continous variables and gradients. Since the majority of expected readers are assummed to be more familiar with the least-squares criterion via gradients, this report will stick strictly with this method.

The general linearized vector equations are

$$x(k+1)_{(p)} = \underset{p}{\Phi}(k+1,\ k)\ x(k)_{(p)} + \underset{pxg}{B(k)}\ f(k)_{(g)} + \underset{pxq}{N(k)}\ u(k)_{(q)} \qquad (1)$$

$$z(k)_{(m)} = \underset{mxp}{H(k)}\ x(k)_{(p)} + v(k)_{(m)} \qquad (2)$$

The deterministic k-varying vector $f(k)_{(g)}$ is of dimension g less than or equal to p, and gets distributed or cross-coupled into all p state variables $x(k+1)_{(p)}$ via the functional relations of $\underset{pxg}{B(k)}$.

The same statements apply to the noise input vector $u(k)_{(q)}$.

The reader should keep in mind the families of trajectories accurately described by the j and n indices, that is

$$x(k+1)_{(p)_j} = \underset{p}{\Phi}(k+1,\ k)\ x(k)_{(p)_j} + B(k)f(k) + N(k)u(k)_j \qquad (3)$$

$$z_{jn}(k) = H(k)\ x(k)_j + v(k)_n \qquad (4)$$

As before, the accurate descriptions designated by j and n will be dropped for simplicity of representation.

The equations are developed as a "recursive process" or an "on-line" processor; that is, as the observations "role in" the mechanized computer-estimator utilizes the data, and discards it or stores it on tape or what have you. All past data is sequentially accumulated in the "memory of the

math-ware" via up dated estimates and variance matrices etc.

Stage k.

Suppose we are at stage k and have computed during stage k-1, the following

$$\hat{x}(k-1, k-1)$$

$$P(k-1, k-1)$$

and predicted via dynamics

$$\hat{x}(k, k-1) = \Phi(k, k-1)\, \hat{x}(k-1, k-1) + B(k-1)\, f(k-1) \tag{5}$$

$$\hat{z}(k, k-1) = H(k)\, \hat{x}(k, k-1) \tag{6}$$

$$\hat{z}(k, k-1) = H(k)\, \phi(k, k-1)\, \hat{x}(k-1, k-1) + H(k)\, B(k-1)\, f(k-1) \tag{7}$$

$$P(k, k-1) = \Phi(k, k-1)\, P(k-1, k-1)\, \Phi^{T}(k, k-1) + N(k-1)\, Q(k-1)\, N^{T}(k-1). \tag{8}$$

We now receive the kth observation

$$z(k) = H(k)\, x(k) + v(k) \tag{9}$$

where $z(k)$ is known but $x(k)$ and $v(k)$ are unknown. We can compute an observation error vector by equation (7) and (9) as

$$\tilde{z}(k, k-1) = H(k)\, [x(k) - \hat{x}(k, k-1)] + v(k) \tag{10}$$

we now can correct the estimate in the state vector based on the observable and computable estimate in the observation vector as

$$\hat{x}(k, k)_{(p)} = \hat{x}(k, k-1)_{(p)} + W(k)\, \tilde{z}(k, k-1)_{(m)} \tag{11}$$

where the weighting matrix W(k) at stage k has p rows and m columns.

We next seek a procedure for selecting at each stage a pxm weighting matrix W(k).

Using equation (10) in equation (11)

$$\begin{aligned} \hat{x}(k, k) &= \hat{x}(k, k-1) + W(k)\, \{H(k)[x(k) - \hat{x}(k, k-1)] + v(k)\} \\ &= [\underset{pxp}{I} - \underset{pxm}{W(k)}\, \underset{mxp}{H(k)}]\, \hat{x}(k, k-1) + W(k)H(k)x(k) + W(k)v(k) \end{aligned} \tag{12}$$

If we now define the "unknown" error vectors

$$\langle x(k) \rangle - \langle \hat{x}(k, k) \rangle = \langle \tilde{x}(k, k) \rangle \tag{13}$$

and

$$\langle x(k) \rangle - \langle \hat{x}(k, k-1) \rangle = \langle \tilde{x}(k, k-1) \rangle \tag{14}$$

then theoretically equation (12) in (13) yields

$$\langle \tilde{x}(k, k) \rangle = [I - W(k)H(k)]\langle x(k) \rangle - [I - W(k)H(k)]\langle \hat{x}(k, k-1) \rangle - W(k)\langle v(k) \rangle \tag{15}$$

$$\langle \tilde{x}(k, k) \rangle = [I - W(k)H(k)]\langle \tilde{x}(k, k-1) \rangle - W(k)\langle v(k) \rangle \tag{16}$$

Transposing (16) we obtain

$$\langle \tilde{x}(k, k) = \langle \tilde{x}(k, k-1) [I - H^T(k) W^T(k)] - \langle v(k) W^T(k) \tag{17}$$

The dyadic product of equation (16) and (17) yields

$$\begin{aligned} \langle \tilde{x}(k, k) \rangle \langle \tilde{x}(k, k) = \{[I - W(k)H(k)] \langle \tilde{x}(k, k-1) \rangle - W(k)\langle v(k) \rangle\} \\ \{\langle \tilde{x}(k, k-1)[I - H^T(k) W^T(k)] - \langle v(k) W^T(k)\} \\ = [I - W(k)H(k)]\langle \tilde{x}(k, k-1) \rangle \langle \tilde{x}(k, k-1)[I - H^T(k)W^T(k)] \\ - [I - W(k)H(k)]\langle \tilde{x}(k, k-1) \rangle \langle v(k)W^T(k) \\ - W(k)\langle v(k) \rangle \langle \tilde{x}(k, k-1)[I - H^T(k)W^T(k)] \\ + W(k)\langle v(k) \rangle \langle v(k)W^T(k). \end{aligned} \tag{18}$$

The square of the magnitude of the error vector $\langle \tilde{x}(k, k) \rangle$ is given as the inner-product of equation (16) and equation (17) or, as the trace of the outer-product of equation (18) as

$$\begin{aligned} \langle \tilde{x}(k, k) \tilde{x}(k, k) \rangle = \langle \tilde{x}(k, k-1) \tilde{x}(k, k-1) \rangle - 2\langle \tilde{x}(k, k-1)W(k)H\tilde{x}(k,k-1) \rangle \\ + \langle \tilde{x}(k,k-1)H^T(k)W^TWH(k) \tilde{x}(k, k-1) \rangle \\ - 2\langle \tilde{x}(k, k-1)W(k) v(k) \rangle \\ + 2\langle \tilde{x}(k, k-1)H^T(k) W^T(k) W(k) v(k) \rangle \\ + \langle v(k) W^T(k) W(k) v(k) \rangle \end{aligned} \tag{19}$$

Equation (19) is a scalar valued function of a matrix argument W(k) of size pxm. We shall take the partial derivative of the scalar with respect to the matrix W(k),

$$\frac{\partial}{\partial W} \left\langle \tilde{x}(k, k)\, \tilde{x}(k, k) \right\rangle \quad \text{term by term.}$$

By equation (19), there are six additive terms, each term will be handled via the "gradient" methods in Appendix B.

The first term is not a function of

W(k), hence

$$\frac{\partial}{\partial W(k)} \left\langle \tilde{x}(k, k-1)\, \tilde{x}(k, k-1) \right\rangle = [0] \tag{20}$$

The second term is by equation (b-61)

$$\frac{\partial}{\partial W(k)} \left\{ -2 \left\langle \tilde{x}(k, k-1)\, W(k)\, H(k)\, \tilde{x}(k, k-1) \right\rangle \right\} = \underset{mxp}{-2H(k)} \tilde{x}(k,k-1) \times \tilde{x}(k,k-1) \tag{21}$$

The third term is

$$\frac{\partial}{\partial W} \left\{ \left\langle \tilde{x}(k, k-1)\, H^T(k)\, W^T(k)\, W(k)\, H(k)\, \tilde{x}(k, k-1) \right\rangle \right\} \tag{22}$$
$$= 2H(k)\, \tilde{x}(k, k-1) \times \tilde{x}(k, k-1)\, H^T(k)\, W^T(k)$$

The latter derivation is based on equation (b-80) and setting

$$\langle c = \langle \tilde{x}(k, k-1)\, H^T(k) \tag{23}$$
$$b(m\rangle = H(k)\, \tilde{x}(k, k-1)\rangle$$

The fourth term is by equation (b-56) (24)

$$\frac{\partial}{\partial W} \left\{ -2 \left\langle \tilde{x}(k, k-1)\, W(k)\, v(k) \right\rangle \right\} = -2v(k) \times \tilde{x}(k, k-1) \tag{25}$$

The fifth term by equation b-80 is

$$\frac{\partial}{\partial W} \left\{ 2 \left\langle \tilde{x}(k, k-1)\, H^T(k)\, W^T(k)\, W(k)\, v(k) \right\rangle \right\} \tag{26}$$
$$= \{H(k)\, \tilde{x}(k, k-1) \times v(k) + v(k) \times \tilde{x}(k, k-1) H^T\}\, W^T$$

based on setting

$$\langle c = \langle \hat{x}(k) \rangle^T, \; v(k) \rangle = b \rangle \tag{27}$$

in equation (b-80)

The sixth term is

$$\frac{\partial}{\partial W} \{ \langle v(k)\, W^T(k)\, W(k)\, v(k) \rangle \} = 2v(k) \times v(k)\, W^T(k) \tag{28}$$

Utilizing the above six expressions in equation ($_{19}$) after the partial derivative has been taken

$$\frac{\partial}{\partial W} \left\{ \langle \tilde{x}(k, k)\, \tilde{x}(k, k) \rangle \right\} \tag{29}$$

$$= -2H(k)\, \tilde{x}(k, k-1) \times \tilde{x}(k, k-1)$$

$$+ 2H(k)\, \tilde{x}(k, k-1) \times \tilde{x}(k, k-1) H^T(k)\, W^T(k)$$

$$- 2v(k) \times \tilde{x}(k, k-1)$$

$$+ \{H(k)\, \tilde{x}(k, k-1) \times v(k) + v(k) \times \tilde{x}(k, k-1) H^T(k)\}\, W^T(k)$$

$$+ 2v(k) \times v(k)\, W^T(k) = \underset{mxp}{[0]}$$

The expected value over all experiments and allowable values of j and n yields

$$E\left\{ \frac{\partial}{\partial W} \langle \tilde{x}(k, k)\, \tilde{x}(k\, k) \rangle \right\} = -2H(k)\, E\{\tilde{x}(k, k-1) \times \tilde{x}(k, k-1)\} \tag{30}$$

$$+2H(k)\, E\{\tilde{x}(k, k-1) \times \tilde{x}(k, k-1)\}\, H^T(k)\, W^T(k)$$

$$-2E\{v(k) \times \tilde{x}(k, k-1)\}$$

$$+ [H(k)\, E\{\tilde{x}(k, k-1) \times v(k) + v(k) \times \tilde{x}(k, k-1)\, H^T(k)]\, W^T(k)$$

$$+ 2E\{v(k) \times v(k)\}\, W^T(k) = [0]$$

If we use the notation

$$P(k, k) = \Sigma_{\tilde{x}\tilde{x}}(k, k) = E\{\tilde{x}(k, k)\,\tilde{x}(k, k)\} \tag{31}$$

$$P(k, k-1) = \Sigma_{\tilde{x}\tilde{x}}(k, k-1) = E\{\tilde{x}(k, k-1)\,\tilde{x}(k, k-1)\} \tag{32}$$

$$Q(k) = E\{u(k)\,u(k)\} \tag{33}$$

$$R(k) = E\{v(k)\,v(k)\} \tag{34}$$

and assume that the conventional statistical independence assumptions hold,

$$E\{\tilde{x}(k, k-1)\,u(k)\} = [0] \tag{35}$$

$$E\{v(k)\,v(k-1)\} = [0] \tag{36}$$

$$E\{v(k)\,u(k)\} = [0], \text{ etc.} \tag{37}$$

By equation (31) through (37) in equation (30)

$$\frac{\partial(\mathrm{tr}\, P(k,k))}{\partial W(k)} = -2H(k)\, P(k, k-1) + 2H(k)\, P(k, k-1)\, H^T(k)\, W^T(k) + 2\Sigma_{vv}(k)\, W^T(k) = [0] \tag{38}$$

$$[H(k)\, P(k, k-1)\, H^T(k) + \Sigma_{vv}(k)]\, W^T(k) = H(k)\, P(k, k-1) \tag{39}$$

Transposing

$$W(k)[H(k)\, P(k, k-1)\, H^T(k) + \Sigma_{uu}(k)] = P(k, k-1)\, H^T(k) \tag{40}$$

Inverting

$$W(k) = P(k, k-1)\, H^T(k)\, [H(k)\, P(k, k-1)\, H(k) + \Sigma_{vv}(k)]^{-1} \tag{41}$$

or

$$\boxed{W(k) = P(k, k-1)\, H^T(k)[H(k)\, P(k, k-1) H^T(k) + R(k)]^{-1}} \tag{42}$$

$$W^T(k) = [H(k)\, P(k, k-1)\, H^T(k) + R(k)]^{-1}\, H(k)\, P(k, k-1) \tag{43}$$

The pxp matrix variance of the estimate of state (the p-space ellipsoid of uncertainty) can be obtained by taking the expected value over all experiments of the dyadic product of equation (18),

$$P(k, k) = E\{(\tilde{x}(k, k) - x(k, k))$$ (44)

$$= [I - W(k) H(k)] P(k, k-1)[I - H^T(k) W^T(k)]$$

$$+ W(k) R(k) W^T(k)$$

Multiplying out the terms of equation (44) we obtain

$$P(k, k) = P(k, k-1) - P(k, k-1) H^T(k) W^T(k)$$ (45)

$$- W(k) H(k) P(k, k-1) + W(k) H(k) P(k, k-1) H^T(k) W^T(k)$$

$$+ W(k) R(k) W^T(k)$$

$$= P(k, k-1) - P(k, k-1) H^T(k) W^T(k)$$

$$- W(k) H(k) P(k, k-1)$$

$$+ W(k)[H(k) P(k, k-1) H^T(k) + R(k)] W^T(k)$$

Consider the last term of the above equation and equation (42) for W(k) with the transpose (43), then the last term becomes

$$W(k)[H(k) P(k, k-1) H^T(k) + R(k)] W^T(k)$$ (46)

$$= P(k, k-1) H^T(k)[H(k) P(k, k-1) H^T(k) + R(k)]^{-1}[H(k) P(k,k-1) H^T(k) + R(k)]$$

$$x[H(k) P(k,k-1) H^T(k) + R(k)]^{-1} H(k) P(k, k-1)$$

$$= P(k,k-1) H^T(k)[H(k) P(k,k-1) + R(k)]^{-1} H(k) P(k, k-1)$$

$$= W(k) H(k) P(k, k-1).$$

Using the above expression for the last term in equation (45) we obtain

$$P(k, k) = P(k,k-1) - P(k, k-1) H^T(k) W^T(k)$$ (47)

$$- W(k) H(k) P(k, k-1) + W(k) H(k) P(k, k-1)$$

or

$$P(k,k) = P(k, k-1) - P(k, k-1) H^T(k) W^T(k)$$ (48)

$$P(k,k) = P(k, k-1)[I - H^T(k) W^T(k)]$$ (49)
$$\text{pxp} \quad \text{pxm} \quad \text{mxp}$$

The matrix P(k,k-1) was predicted and computed during stage k-1.

We can now predict the next stage (k+1) state vector via the deterministic process-dynamics

$$\hat{x}(k+1, k)\rangle = \phi(k+1, k)\,\hat{x}(k, k)\rangle + B(k)\,f(k)\rangle . \tag{50}$$

The next stage prediction of the observation vector is

$$\hat{z}(k+1, k)\rangle = H(k+1)\,\hat{x}(k+1, k)\rangle . \tag{51}$$

The error in the state vector at stage k+1 based on the prediction of equation (50) is

$$\tilde{x}(k+1, k)\rangle = x(k+1)\rangle - \hat{x}(k+1, k)\rangle \tag{52}$$

where the unknown state vector is

$$x(k+1)\rangle = \Phi(k+1, k)x(k)\rangle + B(k)\,f(k)\rangle + N(k)\,u(k)\rangle \tag{53}$$

and the unknown error is by equation (50) and equation (53) in equation (52)

$$\begin{aligned}\tilde{x}(k+1, k)\rangle &= \phi(k+1, k)x(k)\rangle + B(k)\,f(k)\rangle \\ &\quad + N(k)\,u(k)\rangle - \phi(k+1, k)\hat{x}(k, k)\rangle - B(k)\,f(k)\rangle \\ &= \phi(k+1, k)[x(k)\rangle - \hat{x}(k, k)\rangle] + N(k)\,u(k)\rangle\end{aligned} \tag{54}$$

Using equation (14) in equation (54)

$$\tilde{x}(k+1, k)\rangle = \phi(k+1, k)\,\tilde{x}(k, k)\rangle + N(k)\,u(k)\rangle \tag{55}$$

The transpose of equation (55) is

$$\langle\tilde{x}(k+1, k) = \langle\tilde{x}(k, k)\,\phi^{T}(k+1, k) + \langle u(k)\,N^{T}(k) \tag{56}$$

The dyadic product of equation (55) and equation (56) is

$$\begin{aligned}&\tilde{x}(k+1, k)\rangle\langle\tilde{x}(k+1, k) \\ &= \phi(k+1, k)\tilde{x}(k, k)\rangle\langle\tilde{x}(k, k)\Phi^{T}(k+1, k) \\ &+ \phi(k+1, k)\,\tilde{x}(k, k)\rangle\langle u(k)\,N^{T}(k) \\ &+ N(k)\,u(k)\rangle\langle\tilde{x}(k, k)\Phi^{T}(k+1, k) \\ &+ N(k)\,u(k)\rangle\langle u(k)\,N^{T}(k)\end{aligned} \tag{57}$$

The expectation over all experiments of equation (57) is

$$E\{\tilde{x}(k+1, k)\tilde{x}(k+1, k)\} = P(k+1, k) \tag{58}$$
$$= \phi(k+1, k)\, E\{\tilde{x}(k, k)\tilde{x}(k, k)\}\, \phi^T(k+1, k)$$
$$+ N(k)\, E\{u(k)u(k)\}\, N^T(k).$$

The statistical independence assumption was invoked:

$$E\{\tilde{x}(k, k)u(k)\} = [0].$$

Define the process noise variance matrix

$$E\{u(k)u(k)\} = Q(k) \tag{59}$$

and equation (58) becomes

$$P(k+1, k) = \phi(k+1, k)\, P(k, k)\, \phi^T(k+1, k) + N(k)\, Q(k)\, N^T(k). \tag{60}$$

We may now summarize the equations and the computations to be performed at the kth stage as the kth stage summary.

We have available from stage k-1

$$\hat{x}(k, k-1)$$
$$\hat{z}(k, k-1) = H(k)\, \hat{x}(k, k-1)$$
$$P(k, k-1).$$

Receive $z(k)$

Compute error in observation

$$\tilde{z}(k, k-1) = z(k) - \hat{z}(k, k-1) \tag{61}$$

Compute weight matrix W(k) by equation (42)

$$W(k) = P(k, k-1)\, H^T(k)[H(k)\, P(k, k-1)\, H^T(k) + R(k)]^{-1} \tag{62}$$

Correct state estimate by equation (11)

$$\hat{x}(k, k) = \hat{x}(k, k-1) + W(k)\, \tilde{z}(k, k-1) \tag{63}$$

Compute new state variance by equation (49)

$$P(k, k) = P(k, k-1)[I - H^T(k)\, W^T(k)] \tag{64}$$

Predict to stage k+1, by equation (50)

$$\hat{x}(k+1, k)\rangle = \Phi(k+1, k)\, \hat{x}(k, k)\rangle + B(k)\, f(k)\rangle \tag{65}$$

Predict observation by equation (51)

$$\hat{z}(k+1, k)\rangle = H(k+1)\, \hat{x}(k+1, k)\rangle \tag{66}$$

Predict state variance by equation (60)

$$P(k+1, k) = \Phi(k+1, k)\, P(k, k)\, \Phi^T(k+1, k) + N(k)\, Q(k)\, N^T(k) \tag{67}$$

wait for stage k+1 and new measurement vector.

The equations can be substituted and juggled around to obtain alternate expressions, for example using equation (42) and (61) in (63) we obtain

$$\hat{x}(k, k)\rangle = \hat{x}(k, k-1)\rangle \tag{68}$$
$$+ P(k, k-1)\, H^T(k)[H(k)\, P(k, k-1)\, H^T(k) + R(k)]^{-1}$$
$$x\{z(k)\rangle - H(k)\, \Phi(k, k-1)\, \hat{x}(k-1, k-1)\rangle - H(k)B(k-1)f(k-1)\rangle\}$$

Many similar variations of the above systems of equations occur in the literature.

Stage k = 1.

The vector starting system of equations can be derived from equation for k = 1,

$$\hat{x}(1, 0)\rangle = \text{intelligent guess} \tag{69}$$

$$\hat{z}(1,0)\rangle = H(1)\, \hat{x}(1,0)\rangle \tag{70}$$

and

$$P(1, 0) = \text{intelligent guess based on experience about the process.} \tag{71}$$

Receive $z(1)\rangle$

Compute error

$$\tilde{z}(1, 0)\rangle = z(1)\rangle - \hat{z}(1, 0)\rangle \tag{72}$$

Compute first weight

$$W(1) = P(1, 0)\,[H^T(1)\,P(1, 0)\,H^T(1) + R(1)]^{-1} \tag{73}$$

Correct state

$$\hat{x}(1, 1)\rangle = \hat{x}(1, 0)\rangle + W(1)\,\tilde{z}(1, 0)\rangle \tag{74}$$

Compute state variance matrix by equation 49

$$P(1, 1) = P(1, 0)\,[I - H^T(1)\,W^T(1)] \tag{75}$$

Predict stage 2 by equation (50)

$$\hat{x}(2, 1)\rangle = \Phi(2, 1)\,\hat{x}(1, 1)\rangle + B(1)\,f(1)\rangle \tag{76}$$

Predict stage 2 observation

$$\hat{z}(2, 1)\rangle = H(2)\,\hat{x}(2, 1)\rangle \tag{77}$$

Predict state variance matrix by equation 67

$$P(2, 1) = \Phi(2, 1)\,P(1, 1)\,\Phi^t(2, 1) + N(1)Q(1)N^T(1) \tag{78}$$

ETC.

IV. EQUATION SUMMARY FOR COMPUTER APPLICATION

This section summarizes the equations of the previous sections and points out how to compute mechanize the estimation equations to recursively estimate the state vector as the observations "roll into the computer". Precomputation of the sequence of weighting matrices for large dynamical systems is a necessity.

The dynamical process is described by the state vector equation

$$\overrightarrow{X(k+1)} = \Phi(k+1, k)\overrightarrow{X(k)} + B\overrightarrow{f(k)} + B\overrightarrow{f(k)} + N(k)\ \overrightarrow{U(k)} \tag{1}$$

and a system of noisy instruments whose outputs are functionally related to the states by the observation equation

$$\overrightarrow{Z(k)} = H(k)\overrightarrow{X(k)} + \overrightarrow{V(k)} \ . \tag{2}$$

The system of estimation equations are:

The state vector prediction equation

$$\overrightarrow{X(k+1, k)} = \Phi(k+1, k)\overrightarrow{\hat{X}(k,k)} + B(k)\overrightarrow{f(k)} . \tag{3}$$

The observation prediction

$$\overrightarrow{\hat{Z}(k+1, k)} = H(k+1)\overrightarrow{\hat{X}(k+1,k)} \ . \tag{4}$$

The Observation error

$$\overrightarrow{\gamma Z(k+1, k)} = \overrightarrow{Z(k+1)} - \overrightarrow{\hat{Z}(k+1,k)} \tag{5}$$

and the correction to the predicted states at k+1 after the $(k+1)^{th}$ observation is available

$$\overrightarrow{\hat{X}(k+1, k+1)} = \overrightarrow{\hat{X}(k+1, k)} + W(k+1)\ \overrightarrow{\gamma Z(k+1,k)} \ . \tag{6}$$

The sequence of weighting matrices W(k) can be precomputed and stored in memory. The weights are:

$$W(k+1) = P(k+1, k)H(k+1)\left[H(k+1)P(k+1, k)H^T(k+1) + R(k+1)\right]^{-1} \quad (7)$$

where

$$P(k+1, k) = \Phi(k+1,k)P(k,k)\Phi^T(k+1, k) + N(k+1)Q(k+1)N^T(k+1) \ . \quad (8)$$

The block diagram is shown in Figure (1) as the conventional feedback (discreet) system.

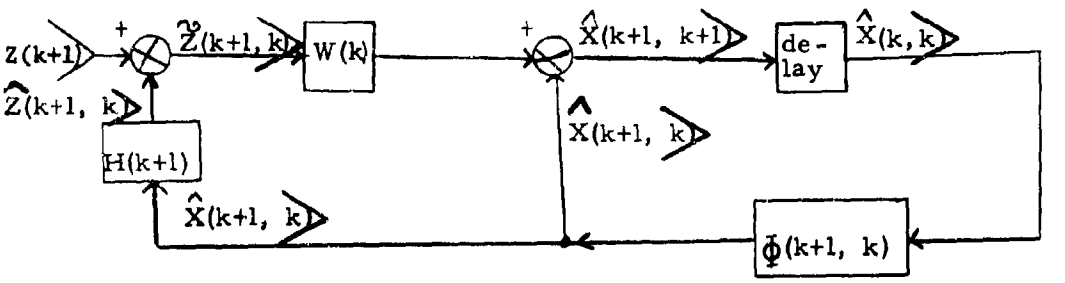

FIGURE (1) - DISCREET FEEDBACK BLOCK

By re-arranging the positions of the feedback blocks one can obtain a flow-block which looks more familiar to a digital programmer as shown in Figure (2).

Tests or applications in which one can plan or design the experiment and the times k, k+1, etc. at which instrument-data will be used to estimate appear to admit of pre-computing the weights. If the estimation times are not pre-designed one must compute the weights on-line.

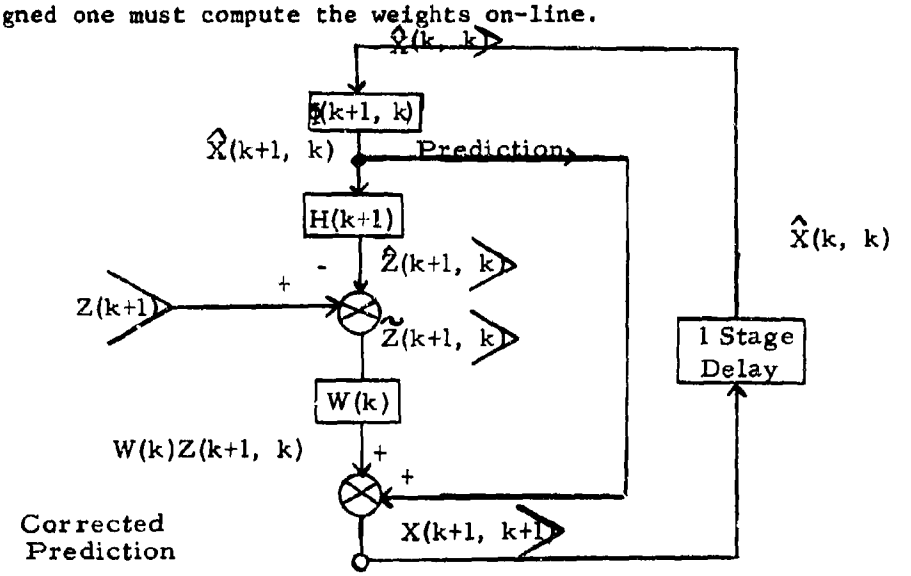

Figure (3) - Flow Block of Estimator

APPENDIX A - MATRIX TRACE PROPERTIES

The trace of a matrix, the trace of the product of two matrices, and the trace of a matrix-sum are useful notions to aid the development of the topics of Appendix B.

Consider a matrix $\underset{pxm}{A}$ of p rows and m columns where $m < p$ and a matrix $\underset{mxp}{B}$, then the product

$$\underset{pxp}{Q_1} = \underset{pxm}{A}\ \underset{mxp}{B} \tag{1}$$

is a pxp matrix.

The matrices A and B can be partitioned into their row and column spaces as shown

$$A = \left[\overset{1}{\underset{1}{\text{a}}}, \ \ldots\ , \overset{}{\underset{m}{\text{a}}} \right] = \begin{bmatrix} \overset{1}{\underset{m}{\text{a}}} \\ \vdots \\ \overset{p}{\underset{m}{\text{a}}} \end{bmatrix} \tag{2}$$

$$B = \begin{bmatrix} \overset{1}{\underset{p}{\text{b}}} \\ \vdots \\ \overset{m}{\underset{p}{\text{b}}} \end{bmatrix} = \left[\overset{}{\underset{1}{\text{b}}}, \ \ldots\ , \overset{}{\underset{p}{\text{b}}} \right] \tag{3}$$

The product Q can be written as a matrix of inner-products

$$Q_1 = AB = \begin{bmatrix} \overset{1}{\underset{m}{\text{a}}} \\ \vdots \\ \overset{p}{\underset{m}{\text{a}}} \end{bmatrix} \left[\overset{}{\underset{1}{\text{b}}}, \ \ldots\ , \overset{}{\underset{p}{\text{b}}} \right] \tag{4}$$

$$= \begin{bmatrix} \overset{1}{\underset{m}{\text{a}}}\ \overset{}{\underset{1}{\text{b}}} & , \ \ldots\ , & \overset{1}{\underset{m}{\text{a}}}\ \overset{}{\underset{p}{\text{b}}} \\ \vdots & & \\ \overset{p}{\underset{m}{\text{a}}}\ \overset{}{\underset{1}{\text{b}}} & , \ \ldots\ , & \overset{p}{\underset{m}{\text{a}}}\ \overset{}{\underset{p}{\text{b}}} \end{bmatrix} \tag{5}$$

or as a sum of dyads (outer products)

$$Q_1 = AB = \begin{bmatrix} a(p)_1 , \ldots , a(p)_m \end{bmatrix} \begin{bmatrix} {}^{1}(p)b \\ \vdots \\ {}^{m}(p)b \end{bmatrix} \tag{6}$$

$$Q_1 = a(p)_1 {}^{1}(p)b + \ldots + a(p)_m {}^{m}(p)b \tag{7}$$

Equation (7) expresses Q_1 as a sum of m rank-one matrices.

If we commute the product we obtain a square mxm matrix

$$\underset{m\times m}{Q_2} = \underset{m\times p}{B}\ \underset{p\times m}{A}$$

and as before Q_2 can be written as a matrix of inner-products

$$\underset{m\times m}{Q_2} = \begin{bmatrix} {}^{1}(p)b \\ \vdots \\ {}^{m}(p)b \end{bmatrix} \begin{bmatrix} a(p)_1 , \ldots , a(p)_m \end{bmatrix} \tag{8}$$

$$= \begin{bmatrix} {}^{1}(p)b\ a(p)_1 , \ldots , {}^{1}(p)b\ a(p)_m \\ {}^{m}(p)b\ a(p)_1 , \ldots , {}^{m}(p)b\ a(p)_m \end{bmatrix} \tag{9}$$

or as a sum of dyadic products

$$Q_2 = \begin{bmatrix} b(m)_1 , \ldots , b(m)_p \end{bmatrix} \begin{bmatrix} {}^{1}(m)a \\ \vdots \\ {}^{p}(m)a \end{bmatrix} \tag{10}$$

$$= b(m)_1 {}^{1}(m)a + \ldots + b(m)_p {}^{p}(m)a \tag{11}$$

Clearly matrix multiplication is not commutative, that is

$$\underset{p\times p}{AB} \neq \underset{m\times m}{BA} \tag{12}$$

in fact the matrices are not even of the same size.

However the trace of both products are equal, that is

$$\mathrm{tr}(\underset{p\times p}{AB}) = \mathrm{tr}(\underset{m\times m}{BA}) \tag{13}$$

The following will clarify the above relation.

If we have a column vector $x(p\rangle$ and a row vector $\langle p)y$ of the same dimension p then the dyadic product is the square, rank-one, matrix D of p rows and p columns

$$\underset{p\times p}{D} = x(p\rangle\langle p)y = \begin{pmatrix} x^1y_1 & \cdots & x^1y_p \\ x^py_1 & \cdots & x^py_p \end{pmatrix} \tag{14}$$

If we commute the product of Equation (14) we obtain

$$\underset{1\times 1}{d} = \langle p)\, yx\, (p\rangle = y_1x^1 + y_2x^2 + \cdots + y_px^p \tag{15}$$

a scalar.

When the elements y_i and x^i are real field elements the products commute, hence

$$y_ix^i = x^iy_i \tag{16}$$

and Equation (15) [the inner product] can be written as the sum of the main diagonal terms of $x\rangle\langle y$, which the conventional definition of the trace (tr) of a matrix, hence

$$\mathrm{tr}\left[x\rangle\langle y\right] = \langle y\, x\rangle. \tag{17}$$

The dyadic product is not as mysterious as many novices might imagine; in fact, if we write Equation (14) as

$$\begin{aligned} \underset{p\times p}{D} = x\rangle\langle y &= x\rangle\begin{pmatrix} y_1, & y_2 & \cdots & y_p \end{pmatrix} \\ &= \begin{bmatrix} x\rangle y_1, & x\rangle y_2 & \cdots & x\rangle y_p \end{bmatrix} \end{aligned} \tag{18}$$

we see that the matrix D when partitioned into its column space is a row of p parallel column vectors - all p of the vectors lie on a line, hence $x\rangle\langle y$ is said to have rank one - that is, there is only one linearly independent vector in the row "package" of column vectors.

APPENDIX B

GRADIENTS OF SCALARS WITH RESPECT TO MATRICES

This appendix develops the gradient of a scalar-valued function with respect to a vector variable and also with respect to a matrix variable.

Case 1. $q = \langle p|a\,x|p\rangle$. Consider the scalar q which is the inner-product

$$q = \langle p|a\,x|p\rangle \tag{b-1}$$

where $\langle a$ is a fixed p dimensional row vector and $|x\rangle$ is a variable column vector, or q is said to be a scalar-valued variable which is a function of the vector variable $|x\rangle$.

In equation (b-1) q may be considered to have vector factors $\langle a$ and $|x\rangle$.

If we have a dyad

$$Q = |x\rangle\langle a \tag{b-2}$$

then it was shown in appendix A that

$$\text{tr } Q = q \tag{b-3}$$

or

$$\text{tr}\left[|x\rangle\langle a\right] = \langle a\,x\rangle = q \tag{b-4}$$

The differential of equation (b-2) is

$$d\,Q = |dx\rangle\langle a \tag{b-5}$$

and the trace of (b-5) is

$$\text{tr } dQ = \text{tr}\left[|dx\rangle\langle a\right] = \langle a\,dx\rangle = dq \tag{b-6}$$

We may now ask to express the differential matrix dQ in terms of vector factors $|dx\rangle$ and a gradient vector, that is

$$dQ = |dx\rangle\left\langle\frac{\partial q}{\partial x}\right| \tag{b-7}$$

such that the trace of equation (b-7) is

$$dq = \text{tr } dQ = \text{tr}\left[|dx\rangle\left\langle\frac{\partial q}{\partial p}\right|\right] = \left\langle\frac{\partial q}{\partial x}\,dx\right\rangle \tag{b-8}$$

If we take the trace of AB by Equation (5) as the sum of diagonals we obtain

$$tr(AB) = \langle m|a\,b|m\rangle_1 + \cdots\cdots + \langle m|a\,b|m\rangle_p \tag{12}$$

If we take the trace of dyadic sum decomposition of AB given by Equation (7) we obtain

$$tr(AB) = tr\left[\, a(p\rangle_1 \langle p)b + \cdots\cdots + a(p\rangle_m \langle p)b \,\right] \tag{13}$$

The trace of a sum of matrices is the sum of the traces, hence by Equation (17)

$$tr(AB) = tr\, a(p\rangle_1 \langle p)b + tr\, a(p\rangle_2 \langle p)b + \cdots\cdots + tr\, a(p\rangle_m \langle p)b \tag{14}$$

$$tr(AB) = \langle p|b\, a(p\rangle_1 + \langle p|b\, a(p\rangle_2 + \cdots\cdots + \langle p|b\, a(p\rangle_m \tag{15}$$

Equation (12) is a sum of p inner-products of m-dimensional vectors and Equation (15) is a sum of m inner-products of p-dimensional vector.

The sum of the main diagonal terms of Equation (9) is

$$tr(BA) = \langle p|b\, a(p\rangle_1 + \cdots\cdots + \langle p|b\, a(p\rangle_m \tag{16}$$

which by Equation (15) and Equation (16)

$$tr(\underset{pp}{AB}) = tr(\underset{m\times m}{BA}) \tag{17}$$

By equation (b-7) and (b-5) we can state

$$\left\langle \frac{\partial q}{\partial x} \right\rangle = \langle a. \tag{b-9}$$

We arrive at the result of equation (b-9) directly from (1)

$$dq = \langle a \, dx \rangle = \left\langle \frac{\partial q}{\partial x} \, dx \right\rangle \tag{b-10}$$

hence

$$\langle a = \left\langle \frac{\partial q}{\partial x} \right. . \tag{b-11}$$

Also one can consider the gradient as an operator $\left\langle \frac{\partial}{\partial p} \right.$

$$q \left\langle \frac{\partial}{\partial q} \right. = \langle a \rangle \left\langle \frac{\partial}{\partial q} \right. = \langle a \left[\rangle \left\langle \frac{\partial}{\partial q} \right. \right] \tag{b-12}$$

The dyadic-type operator

$$\rangle \left\langle \frac{\partial}{\partial x} \right. = \begin{pmatrix} x^1 \\ x^2 \\ \cdot \\ \cdot \\ \cdot \\ x^p \end{pmatrix} \left(\frac{\partial}{\partial x_1}, \ldots \frac{\partial}{\partial x_p} \right) \tag{b-13}$$

$$= \begin{bmatrix} \frac{\partial x^1}{\partial x_1} & \frac{\partial x^1}{\partial x_2}, \ldots & \frac{\partial x^1}{\partial x_p} \\ & & \\ \frac{\partial x^p}{\partial x_1} & & \frac{\partial x^p}{\partial x_p} \end{bmatrix} \tag{b-14}$$

when the coordinates are independent of each other, then

$$\rangle \left\langle \frac{\partial}{\partial x} \right. = I . \tag{b-15}$$

Hence

$$q \left\langle \frac{\partial}{\partial x} \right. = \left\langle \frac{\partial q}{\partial x} \right. = \left\langle a \right. \tag{b-16}$$

In conclusion:

if $q = \langle a\ x \rangle$ (b-17)

then

$$\left\langle \frac{\partial q}{\partial x} \right. = \left\langle a \right. \tag{b-18}$$

Case 2. $q = \langle x\ x \rangle$.

When q is quadratic we can write q as the trace of the dyad

$$Q = |x\rangle\langle x| \tag{b-19}$$

for

$$\text{tr } Q = \text{tr } (|x\rangle\langle x|) = \langle x\ x \rangle = q. \tag{b-20}$$

The differential of the dyad

$$dQ = |dx\rangle\langle x| + |x\rangle\langle dx| \tag{b-21}$$

$$dq = \text{tr } dQ = \langle x\ dx \rangle + \langle dx\ x \rangle = 2 \langle x\ dx \rangle = \left\langle \frac{\partial q}{\partial x}\ dx \right\rangle \tag{b-22}$$

hence

$$\left\langle \frac{\partial q}{\partial x} \right. = 2 \left\langle x \right. \tag{b-23}$$

Case 3. $q = \langle x\ B\ x \rangle$ (b-24)

For this case we have two different matrices

$$Q_1 = |Bx\rangle\langle x| \tag{b-25}$$

and

$$Q_2 = |x\rangle\langle Bx| \tag{b-26}$$

which under the trace operation map down to the same scalar

$$q = \text{tr } Q_1 = \text{tr } Q_2 = \langle x\ B\ x \rangle \tag{b-27}$$

The differential of $Q_2 = Q$ is

$$dQ = dx \langle xB + x \rangle \langle dxB \tag{b-28}$$

The trace of (b-28) is

$$\text{tr } dQ = \langle x\ B\ dx \rangle + \langle dx\ Bx \rangle \tag{b-29}$$

The differential of (b-24) is

$$dq = \langle dx\ Bx \rangle + \langle xB\ dx \rangle = \text{tr } Q \tag{b-30}$$

$$dq = \langle xB^T\ dx \rangle + \langle xB\ dx \rangle \tag{b-31}$$

$$dq = \langle x \left[B + B^T \right] dx \rangle \tag{b-32}$$

we have

$$dq = \left\langle \frac{\partial q}{\partial x}\ dx \right\rangle \tag{b-33}$$

and by (b-32) and (b-33)

$$\left\langle \frac{\partial q}{\partial x} \right\rangle = \langle x \left[B + B^T \right] \tag{b-34}$$

and for symmetric B

$$B = B^T \tag{b-35}$$

then

$$\left\langle \frac{\partial q}{\partial x} \right\rangle = 2 \langle xB \tag{b-36}$$

Case 4. $q = \langle p)_{pxm} a\ X\ b(m \rangle$ (b-37)

The scalar q is a function of the matrix X of p-rows and m columns.

The scalar q can be written as the trace of the matrix

$$\underset{mxm}{Q} = b(m \rangle \langle p)_{pxm} a\ X \tag{b-38}$$

The differential of Q is

$$dQ = |b\rangle\langle a|\, dX \tag{b-39}$$

By equation (b-37), differentiating

$$dq = \langle a\, dX\, b\rangle = tr\, dQ. \tag{b-40}$$

We seek a gradient matrix $\frac{\partial q}{\partial X}$ of m rows and p columns as one of the factors of dQ that is

$$\underset{m\times m}{dQ} = \underset{m\times p}{\frac{\partial q}{\partial X}}\ \underset{p\times m}{dX} \tag{b-41}$$

such that

$$trdQ = dq = \langle a\, dX\, b\rangle \tag{b-42}$$

Clearly by equation (b-39) and (b-41) if

$$\underset{m\times p}{\frac{\partial q}{\partial X}} = |b(m)\rangle\langle(p)|a \tag{b-43}$$

then (b-42) is satisfied.

An alternate, more direct, approach is given below. Partition X into a row of column vectors (all "contravariant" vectors), then

$$q = \langle(p)|a\left[x(p)_1, \ldots x(p)_m\right]|b(m)\rangle \tag{b-44}$$

$$= \left[\langle a|x(p)_1\rangle, \langle a|x(p)_2\rangle, \ldots \langle a|x(p)_m\rangle\right]|b(m)\rangle$$

$$= \left[\langle a|x(p)_1\rangle, \langle a|x(p)_2\rangle, \ldots \langle a|x(p)_m\rangle\right]\begin{bmatrix} b^1 \\ b^2 \\ \cdot \\ \cdot \\ \cdot \\ b^m \end{bmatrix}$$

$$q = \langle a\, x \rangle_1 b^1 + \langle a\, x \rangle_2 b^2 + \ldots + \langle a\, x \rangle_m b^m$$
$$= q_1(x)_1 + \ldots\ldots\ldots + q_m(x)_m \qquad \text{(b-45)}$$

where each q_i is a function of a single column vector $\langle x \rangle_1$.

The scalar differential of q is

$$dq = \left\langle \frac{\partial q}{\partial x}\, dx \right\rangle_1 + \left\langle \frac{\partial q}{\partial x}\, dx \right\rangle_2 + \ldots + \left\langle \frac{\partial q}{\partial x}\, dx \right\rangle_m \qquad \text{(b-46)}$$

$$dq = \left[\left\langle \frac{\partial q}{\partial x} \right\rangle_1 \left\langle \frac{\partial q}{\partial x} \right\rangle_2 \ldots \left\langle \frac{\partial q}{\partial x} \right\rangle_m \right] \begin{bmatrix} \langle dx \rangle_1 \\ \langle dx \rangle_2 \\ \vdots \\ \langle dx \rangle_m \end{bmatrix} \qquad \text{(b-47)}$$

Equation (b-47) can be written as

$$dq = \mathrm{tr} \left\{ \begin{bmatrix} \langle p \rangle_1 \frac{\partial q}{\partial x} \\ \vdots \\ \langle p \rangle_m \frac{\partial q}{\partial x} \end{bmatrix} \left[\langle dx \rangle_1 \ldots \langle dx \rangle_m \right] \right\} \qquad \text{(b-48)}$$

$$= \mathrm{tr} \left\{ \begin{bmatrix} \left\langle \frac{\partial q}{\partial x}\, dx \right\rangle_1 & \ldots & \langle p\, dx \rangle_m \\ \vdots & & \\ \left\langle \frac{\partial q}{\partial x}\, dx \right\rangle_1 & & \left\langle \frac{\partial q}{\partial x}\, dx \right\rangle_m \end{bmatrix} \right\} \qquad \text{(b-49)}$$

The differential of X is a row of column vectors

$$\underset{pxm}{d\,X} = \left[\, dx\overset{1}{(p)}, \ldots . \; dx\overset{m}{(p)} \,\right] \tag{b-50}$$

and the gradient matrix is a column of row gradient-vectors.

$$\frac{\partial q}{\partial X} = \begin{pmatrix} \overset{1}{(p)}\frac{\partial q}{\partial x} \\ \vdots \\ \overset{m}{(p)}\frac{\partial q}{\partial x} \end{pmatrix} \tag{b-51}$$

From the foregoing we write

$$dQ = \underset{mxp}{\frac{\partial q}{\partial X}}\; dX \tag{b-52}$$

and

$$dq = tr\, dQ = tr\left[\frac{\partial q}{\partial X}\; dX\right] \tag{b-52}$$

By equation (b-45), (b-46) and (b-16)

$$\begin{aligned} \overset{1}{(p)}\frac{\partial q}{\partial x} &= \overset{1'}{\left(\frac{\partial q_1}{\partial x}\right)} = b^1 a \\ &\vdots \\ \overset{m}{(p)}\frac{\partial q}{\partial x} &= \overset{m'}{\left(\frac{\partial q_m}{\partial x}\right)} = b^m a\,. \end{aligned} \tag{b-53}$$

Packaging the row vector gradients of (b-53) into the column of (b-51) we obtain

$$\frac{\partial q}{\partial X} = \begin{pmatrix} b^1 \overset{1}{(p)} a \\ b^2 \overset{2}{(p)} a \\ \vdots \\ b^m \overset{m}{(p)} a \end{pmatrix} = \begin{pmatrix} b^1 \\ b^2 \\ \vdots \\ b^m \end{pmatrix} \overset{p}{(p)} a \tag{b-54}$$

or

$$\frac{\partial q}{\partial X}_{mxp} = b(m\rangle\langle p)a \tag{b-55}$$

hence in conclusion

$$\boxed{\begin{array}{l}\text{if } q = \langle p)a \underset{pxm}{X} \quad b(m\rangle \\ \\ \text{then } \dfrac{\partial q}{\partial X}_{mxp} = b(m\rangle\langle p)a.\end{array}} \tag{b-56}$$

Case 5. $$q = \langle p)a \underset{pxm}{X} \underset{mxp}{B} a(p\rangle \tag{b-57}$$

For this case we set

$$\underset{mxp}{B}\, a(p\rangle = b(m\rangle \tag{b-58}$$

as in equation (b-56), then

$$q = \langle a \underset{pxm}{X} b(m\rangle \tag{b-59}$$

and we obtain the case 4, hence

$$\frac{\partial q}{\partial X}_{mxp} = \underset{mxp}{B}\, a(p\rangle\langle p)a \tag{b-60}$$

or

$$\boxed{\begin{array}{l}\text{if } q = \langle p)a \underset{pxm}{X} \underset{mxp}{B} a(p\rangle \\ \\ \text{then } \dfrac{\partial q}{\partial X}_{mxp} = \underset{mxp}{B}\, a(p\rangle\langle p)a\end{array}} \tag{b-61}$$

Case 6. $q = \underset{pxm}{\langle p)a}\ \underset{mxp}{X\ X^T b(p\rangle}$. (b-62)

This case is the matrix analog of the quadratic vector case of equation (b-24).

We can partition X into its column space and X^T into its row space and obtain

$$q = \langle p)a \left[x(p\rangle_1, \ldots x(p\rangle_m \right] \begin{bmatrix} {}^1\langle p)x \\ \vdots \\ {}^m\langle p)x \end{bmatrix} b(p\rangle \tag{b-63}$$

and

$$q = \langle n \left[x(p\rangle_1 {}^1\langle p)x + \ldots + x(p\rangle_m {}^m\langle p)x \right] b(p\rangle \tag{b-64}$$

Distributing the two end vectors over the dyadic-sum decomposition of XX^T we obtain

$$q = \langle a\, x\rangle_1 {}^1\langle x\, b\rangle + \langle a\, x\rangle_2 {}^2\langle x\, b\rangle + \ldots + \langle a\, x\rangle_m {}^m\langle x\, b\rangle \tag{b-65}$$

Because of inner-product commutativity

$$\langle x\, b\rangle = \langle b\, x\rangle \tag{b-66}$$

hence

$$q = \langle a\, x\rangle_1 \langle b\, x\rangle_1 + \ldots + \langle a\, x\rangle_m \langle b\, x\rangle_m \tag{b-67}$$

$$= p_1(x\rangle_1)\, q_1(x\rangle_1) + \ldots + p_m(x\rangle_m)\, q_m(x\rangle_m)$$

hence the scalar q is a sum of products of scalars $p_i\ q_i$.

We have as before

$$dq = \left[\frac{\partial q}{\partial x}, \frac{\partial q}{\partial x}, \ldots \frac{\partial q}{\partial x} \right] \begin{bmatrix} {}^1dx\rangle \\ \vdots \\ {}^m dx\rangle \end{bmatrix} = \text{tr } dQ \tag{b-68}$$

where dQ is as in equation (b-41).

$$\left\langle \frac{\partial q}{\partial x} \right\rangle = \left\langle \frac{\partial (p_i q_i)}{\partial x} \right\rangle = q_i \left\langle \frac{\partial p_i}{\partial x} \right\rangle + p_i \frac{\partial q_i}{\partial x} \tag{b-69}$$

and

$$p_i = \langle a \, x \rangle_i \tag{b-70}$$

$$\left\langle \frac{\partial p_i}{\partial x} \right\rangle_i = \langle a \tag{b-71}$$

$$q_i = \langle p | b \, x \rangle_i \tag{b-72}$$

$$\left\langle p | \frac{\partial q_i}{\partial x} \right\rangle = \langle p | b \tag{b-73}$$

Using (b-70), (b-71), (b-72), (b-73) in (b-69)

$$\left\langle \frac{\partial q}{\partial x} \right\rangle = q_i \langle a + p_i \langle b \tag{b-74}$$

$$= \langle x \, b \rangle_i \langle a + \langle x \, a \rangle_i \langle b$$

$$\left\langle \frac{\partial q}{\partial x} \right\rangle = \left\langle x \left[b(p) \langle p | a + a(p) \langle p | b \right] \right\rangle_i \tag{b-75}$$

Packaging (b-75) into the gradient matrix of equation (b-51)

$$\underset{m \times p}{\frac{\partial q}{\partial x}} = \begin{bmatrix} \langle x^1 \left[b \langle a + a \langle b \right] \\ \vdots \\ \langle x^m \left[b \langle a + a \langle b \right] \end{bmatrix} \tag{b-76}$$

or

$$\frac{\partial q}{\partial X} = \begin{bmatrix} \langle 1 \rangle x \\ \vdots \\ \langle m \rangle x \end{bmatrix} \left[\rangle b \langle a + a \rangle \langle b \right] \tag{b-77}$$

$$\frac{\partial q}{\partial X} = X^T \left[\rangle b \langle a + a \rangle \langle b \right] \tag{b-78}$$

In conclusion

$$\text{if } q = \langle p \rangle a \underset{p \times m}{X} \underset{m \times p}{X^T} b \langle p \rangle$$

$$\text{then } \underset{m \times p}{\frac{\partial q}{\partial X}} = \underset{m \times p}{X^T} \left[\rangle b \langle a + a \rangle \langle b \right] \tag{b-79}$$

In a similar fashion it can be shown that

$$\text{if } q = \langle m \rangle c \underset{m \times p}{X^T} \underset{p \times m}{X} b \langle m \rangle \tag{b-80}$$

$$\text{then } \underset{m \times p}{\frac{\partial q}{\partial X}} = \left[c \langle m \rangle \langle m \rangle b + b \langle m \rangle \langle c \right] \underset{m \times p}{X^T}$$

Consider the pxp matrix L which has factors as shown

$$\underset{pxp}{L} = \underset{pxk}{B} \; \underset{kxp}{X} \tag{b-81}$$

where X is a variable matrix.

If we factor B into its column space and X into its row space

$$L = [b(p)_1, \ldots b(p)_k] \begin{bmatrix} 1 \\ \langle p \rangle x \\ \vdots \\ k \\ \langle p \rangle x \end{bmatrix} \tag{b-82}$$

$$= b(p)_1 \langle p \rangle^1 x + \ldots + b(p)_k \langle p \rangle^k x \tag{b-83}$$

The trace of L is

$$\text{tr } L = \ell = \langle x \, b \rangle_1^1 + \ldots + \langle x \, b \rangle_k^k \tag{b-84}$$

The differential of (b-81) is

$$dL = B \, d \, X. \tag{b-85}$$

The factors of dL can also be expressed as

$$dL = \underset{pxk}{\frac{\partial (\text{tr} L)}{\partial X}} \; \underset{kxp}{dX} \tag{b-86}$$

where the pxk gradient matrix is

$$\underset{pxk}{\frac{\partial \ell}{\partial X}} = \left[\frac{\partial \ell}{\partial x(p)_1}, \frac{\partial \ell}{\partial x(p)_2}, \ldots \frac{\partial \ell}{\partial x(p)_k} \right] \tag{b-87}$$

The differential of Equation (b-84) is

$$d(\mathrm{tr}\ L) = d\ell = d\ell_1 + \ldots + d\ell_k \tag{b-88}$$

$$= \left\langle dx \frac{\partial \ell}{\partial x} \right\rangle_1^1 + \ldots + \left\langle dx \frac{\partial \ell}{\partial x} \right\rangle_K^k \tag{b-89}$$

where

$$\left\langle \frac{\partial \ell}{\partial x} \right\rangle_1 = \left\langle \frac{\partial \ell_1}{\partial x} \right\rangle_1 = \left\langle \frac{\partial}{\partial x} \right\rangle_1 \left\langle x\ b \right\rangle_1^1 = \langle b \rangle_1 \tag{b-90}$$

$$\vdots$$

$$\left\langle \frac{\partial \ell}{\partial x} \right\rangle_k = \left\langle \frac{\partial \ell_k}{\partial x} \right\rangle_k = \left\langle \frac{\partial}{\partial x} \right\rangle_k \left\langle x\ b \right\rangle_k^k - \langle b \rangle_k$$

and

$$\frac{\partial \ell}{\partial X} = [\langle b(p) \rangle_1, \ldots \langle b(p) \rangle_k] = \underset{pxk}{B} \tag{b-91}$$

In summary,

If

$$\underset{pxp}{L} = \underset{(pxk)}{B}\underset{(kxp)}{X} \tag{b-92}$$

then

$$\underset{pxk}{\frac{\partial(\mathrm{tr}L)}{\partial X}} = \underset{pxk}{B} \tag{b-93}$$

APPENDIX C

MINIMIZATION

Consider the linear surface

$$\ell = \langle b\ x \rangle \tag{1}$$

and the quadratic surface

$$q = \langle xQx \rangle \tag{2}$$

and the difference

$$q - \ell = \phi. \tag{3}$$

If ℓ is a constant, $\ell = \ell_o$, then we seek a vector $x\rangle$ that lies on the linear surface and on the quadratic surface such that difference in the linear surface and the quadratic surface is a minimum.

Differentiating

$$d\phi = dq - d\ell \tag{4}$$

and

$$d\phi = \left\langle \frac{\partial \phi}{\partial x}\, dx \right\rangle \tag{5}$$

$$= \left\langle \frac{\partial q}{\partial x}\, dx \right\rangle - \left\langle \frac{\partial \ell}{\partial x}\, dx \right\rangle \tag{6}$$

$$= \left\langle \left[\frac{\partial q}{\partial x} - \left\langle \frac{\partial \ell}{\partial x} \right] dx \right\rangle \tag{7}$$

or

$$\left\langle \frac{\partial \phi}{\partial x} = \left\langle \frac{\partial q}{\partial x} - \left\langle \frac{\partial \ell}{\partial x} \right. \right. \right. \tag{8}$$

If we equate the gradient vector to zero

$$\left\langle \frac{\partial q}{\partial x} = \left\langle \frac{\partial \ell}{\partial x} \right. \right. . \tag{9}$$

By equation (B-36) and equation (B-9)

$$2\langle xQ = \langle b \tag{10}$$

and solving for $\langle x$

$$\langle x = \langle \frac{bQ^{-1}}{2} . \tag{11}$$

Multiplying equation (11) by $b\rangle$ and using equation (C-1)

$$\langle x\, b\rangle = \langle \frac{b\, Q^{-1}\, b}{2}\rangle = \ell o \tag{12}$$

or

$$\frac{1}{2} = \frac{\ell o}{\langle bQ^{-1}b\rangle} \tag{13}$$

Using (13) in (11)

$$\langle x = \ell o \frac{\langle b\, Q^{-1}}{\langle bQ^{-1}b\rangle} . \tag{14}$$

If

$$\ell o = 1$$

then

$$\langle x = \frac{\langle bQ^{-1}}{\langle bQ^{-1}b\rangle} \tag{15}$$

REFERENCES

1. Dwyer, Paul S., "Some Applications of Matrix Derivatives on Multi-Variate Analysis", American Statistical Association Journal, June 1967.

2. Kalman, R. E., "New Methods in Weiner Filtering Theory", Proceedings of the First Symposium in Engineering Applications of Random Function Theory and Probability. John Wiley and Sons, 1963.

3. Lee, C. K. Robert, "Optimal Estimation, Identification and Control", The M.I.T. Press, Cambridge, Mass., 1964.

4. Pappas, James S., "Application of the Kalman Filter To Sequential Optimal Parameter Estimation Via Householder's Matrix Inversion Method", R-O-S-67-1, Special Report, USATECOM, Analysis and Computation Directorate, Deputy for NRO, White Sands Missile Range, New Mexico, June 1967.

5. Pappas, J. S., Diaz, A., "A Math Model for Computing Noise Variance Matrices for a System of Radar Trackers", R-O-S-67-2, Special Report USATECOM, Analysis and Computation Directorate, Deputy for NRO, White Sands Missile Range, New Mexico, July 1967.

6. Pappas, J. S., "A Vector Space Derivation Using Dyads-of-Weighted Least Squares For Correlated Noise, RO-S-68-1, Special Report, USATECOM, Analysis and Computation Directorate, Deputy for NRO, White Sands Missile Range, New Mexico, June 1967.

7. Scheffe, Henry, "The Analysis of Variance", John Wiley, 1959.

8. Scott, Maceo T., Rede, Edmundo, and Wygant, Marthe, "Discrete Recursive Estimation: An Optimum Automated Data Processing Technique for Multi-Radar Tracking Systems", Technical Report Nr. 5, Analysis and Computation Directorate, White Sands Missile Range, New Mexico, April 1967.

9. Friedman, B., "Principles and Techniques of Applied Mathematics", John Wiley, 1956.

DISTRIBUTION

Special Report Nr. RO-S-68-2, "A Tutorial Deviation of Recursive Weighted Least Squares State-Vector Estimation Theory (Kalman Theory)", UNCLASSIFIED, National Range Operations, White Sands Missile Range, New Mexico 88002, August 1968.

White Sands Missile Range New Mexico (STEWS)	Number of Copies
National Range Operations	
Analysis and Computation Directorate (Record Copy)	1
(Ref. Copy)	1
Math Services Division, ATTN: Mr. Pappas	69
Technical Library	10
Defense Documentation Center Cameron Station Alexandria, Virginia 22314	20

UNCLASSIFIED

Security Classification

DOCUMENT CONTROL DATA - R&D

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

1. ORIGINATING ACTIVITY *(Corporate author)*
ANALYSIS AND COMPUTATION DIRECTORATE
DEPUTY FOR NATIONAL RANGE OPERATIONS
WHITE SANDS MISSILE RANGE, NEW MEXICO 88002

2a. REPORT SECURITY CLASSIFICATION
UNCLASSIFIED

2b. GROUP

3. REPORT TITLE
A TUTORIAL DERIVATION OF RECURSIVE WEIGHTED LEASE SQUARES STATE-VECTOR ESTIMATION THEORY (KALMAN THEORY)

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

5. AUTHOR(S) *(Last name, first name, initial)*
PAPPAS, JAMES S.

6. REPORT DATE	7a. TOTAL NO. OF PAGES	7b. NO. OF REFS
AUGUST 1968	67	9

8a. CONTRACT OR GRANT NO.	9a. ORIGINATOR'S REPORT NUMBER(S)
b. PROJECT NO.	
c.	9b. OTHER REPORT NO(S) *(Any other numbers that may be assigned this report)*
d.	

10. AVAILABILITY/LIMITATION NOTICES
DISTRIBUTION OF THIS DOCUMENT IS UNLIMITED

11. SUPPLEMENTARY NOTES	12. SPONSORING MILITARY ACTIVITY

13. ABSTRACT

This is one of a series of tutorial reports deriving the discrete recursive Kalman estimation equations. The series proceeds from unweighted to weighted least squares parameter estimation in a vector space setting to the stage-wise updating, to time varying parameter or state variable estimation. The derivations have been stage-wise tutorial in an attempt to make the theory accessible to the "non-specialist" in optimization theory.

DD FORM 1 JAN 64 1473

UNCLASSIFIED

Security Classification

14. KEY WORDS	LINK A		LINK B		LINK C	
	ROLE	WT	ROLE	WT	ROLE	WT
1. Matrices						
2. Weighted least squares						
3. Correlated noise						
4. Kalman Filtering						
5. Optimal Estimator						
6. State vector						

INSTRUCTIONS

1. **ORIGINATING ACTIVITY:** Enter the name and address of the contractor, subcontractor, grantee, Department of Defense activity or other organization *(corporate author)* issuing the report.

2a. **REPORT SECURITY CLASSIFICATION:** Enter the overall security classification of the report. Indicate whether "Restricted Data" is included. Marking is to be in accordance with appropriate security regulations.

2b. **GROUP:** Automatic downgrading is specified in DoD Directive 5200.10 and Armed Forces Industrial Manual. Enter the group number. Also, when applicable, show that optional markings have been used for Group 3 and Group 4 as authorized.

3. **REPORT TITLE:** Enter the complete report title in all capital letters. Titles in all cases should be unclassified. If a meaningful title cannot be selected without classification, show title classification in all capitals in parenthesis immediately following the title.

4. **DESCRIPTIVE NOTES:** If appropriate, enter the type of report, e.g., interim, progress, summary, annual, or final. Give the inclusive dates when a specific reporting period is covered.

5. **AUTHOR(S):** Enter the name(s) of author(s) as shown on or in the report. Enter last name, first name, middle initial. If military, show rank and branch of service. The name of the principal author is an absolute minimum requirement.

6. **REPORT DATE:** Enter the date of the report as day, month, year; or month, year. If more than one date appears on the report, use date of publication.

7a. **TOTAL NUMBER OF PAGES:** The total page count should follow normal pagination procedures, i.e., enter the number of pages containing information.

7b. **NUMBER OF REFERENCES:** Enter the total number of references cited in the report.

8a. **CONTRACT OR GRANT NUMBER:** If appropriate, enter the applicable number of the contract or grant under which the report was written.

8b, 8c, & 8d. **PROJECT NUMBER:** Enter the appropriate military department identification, such as project number, subproject number, system numbers, task number, etc.

9a. **ORIGINATOR'S REPORT NUMBER(S):** Enter the official report number by which the document will be identified and controlled by the originating activity. This number must be unique to this report.

9b. **OTHER REPORT NUMBER(S):** If the report has been assigned any other report numbers *(either by the originator or by the sponsor)*, also enter this number(s).

10. **AVAILABILITY/LIMITATION NOTICES:** Enter any limitations on further dissemination of the report, other than those imposed by security classification, using standard statements such as:

(1) "Qualified requesters may obtain copies of this report from DDC."

(2) "Foreign announcement and dissemination of this report by DDC is not authorized."

(3) "U. S. Government agencies may obtain copies of this report directly from DDC. Other qualified DDC users shall request through _____."

(4) "U. S. military agencies may obtain copies of this report directly from DDC. Other qualified users shall request through _____."

(5) "All distribution of this report is controlled. Qualified DDC users shall request through _____."

If the report has been furnished to the Office of Technical Services, Department of Commerce, for sale to the public, indicate this fact and enter the price, if known.

11. **SUPPLEMENTARY NOTES:** Use for additional explanatory notes.

12. **SPONSORING MILITARY ACTIVITY:** Enter the name of the departmental project office or laboratory sponsoring *(paying for)* the research and development. Include address.

13. **ABSTRACT:** Enter an abstract giving a brief and factual summary of the document indicative of the report, even though it may also appear elsewhere in the body of the technical report. If additional space is required, a continuation sheet shall be attached.

It is highly desirable that the abstract of classified reports be unclassified. Each paragraph of the abstract shall end with an indication of the military security classification of the information in the paragraph, represented as *(TS)*, *(S)*, *(C)*, *or (U)*.

There is no limitation on the length of the abstract. However, the suggested length is from 150 to 225 words.

14. **KEY WORDS:** Key words are technically meaningful terms or short phrases that characterize a report and may be used as index entries for cataloging the report. Key words must be selected so that no security classification is required. Idenfiers, such as equipment model designation, trade name, military project code name, geographic location, may be used as key words but will be followed by an indication of technical context. The assignment of links, rules, and weights is optional.