DEPARTMENT OF AERONAUTICAL ENGINEERING

A SURVEY OF FIVE RECURSIVE MINIMUM VARIANCE ESTIMATION PROCEDURES

BY

[Name]

[Date]
A survey of five recursive minimum variance estimation procedures.

by

Drs H.L. Jonkers

DELFT - THE NETHERLANDS
FEBRUARY 1970
Summary

Statistical estimation procedures are used for the accurate reconstruction of state trajectories of systems from inaccurate measurements of system input and output signals. A variety of estimation procedures has been designed for this purpose. As minimization of the variance of an estimate implies maximization of the estimation accuracy, all procedures in common use are designed to minimize the variance of the state vector estimate.

This report presents a survey of five recursive minimum variance estimation procedures. An attempt has been made to clarify theoretical similarities and practical differences between these procedures. Differences between these procedures will be shown to arise from differences in system model characteristics. Moreover it will be apparent from this survey that the choice of an estimation procedure with respect to a particular task depends largely upon the available knowledge of the systems statistics.

This study is mainly based on the work of S.L. Fagin [1].
Contents:

0. Introduction 2

1. Symbol definitions.
   1.o. General symbols and notations 3
   1.1. Alphabetic symbols 4

2. Survey of minimum variance estimation procedures.
   2.o. Tabulated survey 6

Appendix A. Relation between B(n), I(n) and P(n). 19

Appendix B. Estimation mechanism of the Kalman filter and the recursive linear regression method 21

References 26
Introduction.

The purpose of this brief report is merely to elucidate some special topics of minimum variance estimation.

In order to provide a concise survey of five recursive minimum variance estimation methods a table has been compiled of all the constituent relations. The methods under concern are:

1. The recursive least-squares-method (R.L.S.)
2. The recursive weighted least-squares method (R.W.L.S.)
3. The recursive weighted least-squares method with time-dependent weighting matrix (R.W(t).L.S.)
4. The recursive linear regression method (R.L.R.)
5. The discrete-time version of the Kalman filter.

The comparison of these estimation methods is mainly based on the work of Fagin [1]. No attempt has been made either to explain or to derive relations already explained in his article.

For more detailed information the reader is referred to the work of Sorensen [2], Nahi [3] and Deutsch [4].

In Chapter 1 a list of definitions will be provided to facilitate the interpretation of the tables.

Chapter 2 presents a tabulated survey of all relations pertaining to the five mentioned estimation procedures.

The relation between the integral observation matrix B(n), the weighting matrix Σ(n) and the estimation error covariance matrix P(n) will be derived in Appendix A.

The fundamental estimation mechanism of the Kalman filter will be discussed in Appendix B and it will be compared with the mechanism of some of the other estimation methods described in this report.
1. Symbol definitions.

1.0. General symbols and notations.

\[ \theta, \phi \] dummy symbols

\[ \approx \] equals by definition, denotes.

\[ \| \theta \| \] approximately equals

\[ \| \theta \| \] norm of \( \theta \)

\[ \text{Tr}(\theta) \] trace of a matrix

\[ \mathbf{E}[\theta] \] expected value of \( \theta \)

\[ \mathbf{E} \left[ \theta \theta^T \right] \] covariance matrix of \( \theta \) and \( \phi \)

\[ \hat{\theta} \] estimate of \( \theta \)
1.1. Alphabetic symbols.

\( B(t) \) \hspace{1cm} \text{time-dependent integral observation matrix}

\( ^{\hat{\Delta}} \hat{e} \) \hspace{1cm} \( \hat{x} - \hat{x} \), the estimation error vector

\( I \) \hspace{1cm} \text{identity matrix}

\( i,j,n \) \hspace{1cm} \text{integer-valued indices}

\( K(n+1) \) \hspace{1cm} \text{weighting or gain matrix}

\( M(n) \) \hspace{1cm} \text{time-dependent observation matrix at time } t_n.
\( M(n) \) is an \( r \times k \) matrix if \( \hat{x}(n) \) is a \( k \)-dimensional state vector and \( \hat{y}(n) \) is an \( r \)-dimensional vector-valued observation.

\( P(n+1) \) \hspace{1cm} \( \Delta E \left[ (\hat{e}(n+1)) (\hat{e}(n+1))^T \right] \), the error covariance matrix posterior to time \( t_{n+1} \)

\( P(n+1|n) \) \hspace{1cm} \text{the error covariance matrix predicted at time } t_n \text{ for } t_{n+1}

\( q(n) \) \hspace{1cm} \text{minimization criterion}

\( T(n+1) \) \hspace{1cm} \text{weighting matrix}

\( t \) \hspace{1cm} \text{time}

\( \hat{U}(t) \) \hspace{1cm} \text{vector-valued input signal = vector valued forcing function}
\( \mathcal{V}(n) \)  
vector-valued observation noise

\( v(n) \)  
scalar-valued observation noise

\( \Omega(n) \delta_{n,j} \)  
\( \tilde{\Delta} \mathbb{E} \left[ \hat{w}(n) \hat{w}^\text{T}(j) \right] \) = the covariance matrix of the plant noise

\( \tilde{W}(n) \)  
vector-valued white noise input signal (plant noise)

\( \hat{x}(n) \)  
state vector. The components of \( \hat{x}(n) \) are the quantities to be estimated by the applied estimation method. The choice of the appropriate estimation method is determined by the model of the system under consideration (Table 1, 2 and 3),

\( \hat{y}(n) \)  
\( \hat{\Delta} \begin{bmatrix} y_1(n) \\ \vdots \\ y_r(n) \end{bmatrix} \) = the vector valued observation at time \( t_n \).

\( y(n) \)  
scalar observation at time \( t_n \) for \( n = 1, 2, \ldots \)

\( \Gamma(n+1,n) \)  
input signal distribution matrix

\( \delta_{n,j} \)  
Kronecker delta

\( \Sigma^{-1} \)  
weighting matrix at time \( t_n \)

In some cases \( \Sigma(n) \hat{\Delta} \mathbb{E} \left[ \hat{v}(n) \hat{v}^\text{T}(n) \right] \) = the observation error covariance matrix

\( \sigma_y^2(n) \)  
\( \hat{\Delta} \mathbb{E} \left[ v^2(n) \right] \), the variance of the scalar valued observation error at time \( t_n \)

\( \Phi(n+1,n) \)  
state transition matrix
2. Survey of minimum variance estimation procedures.

2.0. Tabulated survey.

Each of the five estimation procedures under concern is characterized by appropriate versions of twelve basic relations. For each basic relation a tabulated survey has been compiled of its five pertaining versions. Practical differences between these estimation procedures and theoretical similarities will be apparent from these tables.

The basic relations characterizing the estimation procedures can be subdivided into five categories.

1. The system model, i.e. the relations describing the time-dependent behavior of the quantities to be estimated and the observation equations relating these quantities to those observed. There are three relations in this category:
a. The state equation
b. The integral observation equation relating n observations to the state vector.
c. The updating observation equation relating one observation to the state vector.
It will be evident that the integral and updating observation equations bear great similarity.

2. The minimization criterion, specified by two relations:
a. The so-called minimization criterion or risk function, i.e. the function to be minimized by the appropriate estimation procedure.
b. The weighting matrix definition.
3. The optimal non-recursive estimation procedure, i.e. the equation relating the optimal nonrecursive estimation of the state vector to a set of n observations.

4. The estimation error covariance matrix definition.

5. The recursive estimation procedure. This category includes five relations:
   a. The prediction at time \( t_n \) of the state vector at time \( t_{n+1} \).
   b. Prediction at time \( t_n \) of the error covariance matrix at time \( t_{n+1} \).
   c. Derivation of the gain matrix at time \( t_{n+1} \) from the predicted error covariance matrix.
   d. The recursive estimation relation providing the desired optimal state vector estimate.
   e. The error propagation equation relating the error covariance matrix at time \( t_{n+1} \) to the error covariance matrix at time \( t_n \) using the predicted error covariance matrix (see 5b), the gain matrix (see 5c) and the system noise covariance matrices.
<table>
<thead>
<tr>
<th>Model: State equation</th>
<th>Methods</th>
<th>Kalman filter</th>
</tr>
</thead>
<tbody>
<tr>
<td>$\dot{x} = 0 \rightarrow \dot{x}(n+1) = \dot{x}(n) = \cdots = \dot{x}$</td>
<td>R.L.S.</td>
<td>$x(n+1) = \xi(n+1,n) x(n) + \Gamma(n+1,n) u(n) + \Gamma(n+1,n) \bar{w}(n)$; $\dot{x}(n) = \dot{x}$</td>
</tr>
<tr>
<td>$\dot{x} = 0 \rightarrow \dot{x}(n+1) = \dot{x}(n) = \cdots = \dot{x}$</td>
<td>R.W.L.S.</td>
<td></td>
</tr>
<tr>
<td>$\dot{x} = 0$, or very small; $\dot{x}(n+1) = \dot{x}(n) = \cdots = \dot{x}$</td>
<td>R.W.(t), L.S.</td>
<td></td>
</tr>
<tr>
<td>$\dot{x} = 0 \rightarrow \dot{x}(n+1) = \dot{x}(n) = \cdots = \dot{x}$</td>
<td>R.L.R.</td>
<td></td>
</tr>
<tr>
<td>Methods</td>
<td>Model: The integral observation equation for n observations.</td>
<td>Table 2</td>
</tr>
<tr>
<td>-------------</td>
<td>-----------------------------------------------------------</td>
<td>---------</td>
</tr>
<tr>
<td>R.L.S.</td>
<td>$\hat{Y}(n) = B(n)\hat{X} + V(n)$; $\hat{Y}(n) \triangleq \begin{bmatrix} y(1) \ \vdots \ y(n) \end{bmatrix}$; $B(n) \triangleq \begin{bmatrix} M(1) \ \vdots \ M(n) \end{bmatrix}$; $V(n) \triangleq \begin{bmatrix} v(1) \ \vdots \ v(n) \end{bmatrix}$</td>
<td></td>
</tr>
<tr>
<td>R.W.L.S.</td>
<td>$\hat{Y}(n) = B(n)\hat{X} + V(n)$; $\hat{Y}(n) \triangleq \begin{bmatrix} y(1) \ \vdots \ y(n) \end{bmatrix}$; $B(n) \triangleq \begin{bmatrix} M(1) \ \vdots \ M(n) \end{bmatrix}$; $V(n) \triangleq \begin{bmatrix} v(1) \ \vdots \ v(n) \end{bmatrix}$</td>
<td></td>
</tr>
<tr>
<td>R.W(t).L.S.</td>
<td>$\hat{Y}(n) = B(n)\hat{X} + V(n)$; $\hat{Y}(n) \triangleq \begin{bmatrix} y(1) \ \vdots \ y(n) \end{bmatrix}$; $B(n) \triangleq \begin{bmatrix} M(1) \ \vdots \ M(n) \end{bmatrix}$; $V(n) \triangleq \begin{bmatrix} v(1) \ \vdots \ v(n) \end{bmatrix}$</td>
<td></td>
</tr>
<tr>
<td>R.L.R.</td>
<td>$\hat{Y}(n) = B(n)\hat{X} + V(n)$</td>
<td>$\hat{Y}(n) \triangleq \begin{bmatrix} \hat{Y}(1) \ \vdots \ \hat{Y}(n) \end{bmatrix}$; $\hat{Y}(n) \triangleq \begin{bmatrix} y_r(1) \ \vdots \ y_r(n) \end{bmatrix}$; $B(n) \triangleq \begin{bmatrix} M(1) \ \vdots \ M(n) \end{bmatrix}$</td>
</tr>
<tr>
<td></td>
<td>$V(n) \triangleq \begin{bmatrix} V(1) \ \vdots \ V(n) \end{bmatrix}$; $V(n) \triangleq \begin{bmatrix} v_r(1) \ \vdots \ v_r(n) \end{bmatrix}$</td>
<td></td>
</tr>
<tr>
<td>Kalman f.</td>
<td>No integral observation equation is defined because of the fact that the Kalman filter is used only as a recursive estimation procedure.</td>
<td></td>
</tr>
<tr>
<td>Methods</td>
<td>Model: The updating observation equation.</td>
<td></td>
</tr>
<tr>
<td>-----------</td>
<td>----------------------------------------------------------------------------------------</td>
<td></td>
</tr>
<tr>
<td>R.L.S.</td>
<td>$y(n+1) = M(n+1)\hat{x} + v(n+1); \quad M(n+1) \triangleq [b_1(n+1) \ldots b_k(n+1)]$</td>
<td></td>
</tr>
<tr>
<td>R.W.L.S.</td>
<td>$y(n+1) = M(n+1)\hat{x} + v(n+1); \quad M(n+1) \triangleq [b_1(n+1) \ldots b_k(n+1)]$</td>
<td></td>
</tr>
<tr>
<td>R.W(t).L.S.</td>
<td>$y(n+1) = M(n+1)\hat{x} + v(n+1); \quad M(n+1) \triangleq [b_1(n+1) \ldots b_k(n+1)]$</td>
<td></td>
</tr>
<tr>
<td>R.L.R.</td>
<td>$\hat{Y}(n+1) = M(n+1)\hat{x} + \hat{V}(n+1); \quad \hat{Y}(n+1) \triangleq \begin{bmatrix} y_1(n+1) \ \vdots \ y_r(n+1) \end{bmatrix}; \quad M(n+1) \triangleq \begin{bmatrix} b_{ij}(n+1) \end{bmatrix}_{i=1}^{r} \quad \begin{bmatrix} j=1 \ldots k \end{bmatrix}$</td>
<td></td>
</tr>
<tr>
<td></td>
<td>$\hat{V}(n+1) \triangleq \begin{bmatrix} v_1(n+1) \ \vdots \ v_r(n+1) \end{bmatrix}; \quad \Sigma(n+1) \triangleq \begin{bmatrix} \Sigma_{n+1,j} \end{bmatrix}_{n+1,j} \quad E \left[ \hat{V}(n+1)\hat{V}^T(j) \right]$</td>
<td></td>
</tr>
<tr>
<td>Kalman f.</td>
<td>$\hat{Y}(n+1) = M(n+1)\hat{x}(n+1) + \hat{V}(n+1); \quad \hat{Y}(n+1) \triangleq \begin{bmatrix} y_1(n+1) \ \vdots \ y_r(n+1) \end{bmatrix}; \quad M(n+1) \triangleq \begin{bmatrix} b_{ij}(n+1) \end{bmatrix}_{i=1}^{r} \quad \begin{bmatrix} j=1 \ldots k \end{bmatrix}$</td>
<td></td>
</tr>
<tr>
<td></td>
<td>$\hat{V}(n+1) \triangleq \begin{bmatrix} v_1(n+1) \ \vdots \ v_r(n+1) \end{bmatrix}; \quad \Sigma(n+1) \triangleq \begin{bmatrix} \Sigma_{n+1,j} \end{bmatrix}_{n+1,j} \quad E \left[ \hat{V}(n+1)\hat{V}^T(j) \right]$</td>
<td></td>
</tr>
<tr>
<td>Methods</td>
<td>Minimization Criterion: Risk function</td>
<td></td>
</tr>
<tr>
<td>--------------</td>
<td>--------------------------------------</td>
<td></td>
</tr>
<tr>
<td>R.I.S.</td>
<td>$q(n) = (\hat{x}(n) - \hat{y}(n))^T (\hat{x}(n) - \hat{y}(n))$</td>
<td></td>
</tr>
<tr>
<td>R.W.I.S.</td>
<td>$q(n) = (\hat{x}(n) - \hat{y}(n))^T \Sigma^{-1}(n) (\hat{x}(n) - \hat{y}(n))$</td>
<td></td>
</tr>
<tr>
<td>R.W.I. S.</td>
<td>$q(n) = (\hat{x}(n) - \hat{y}(n))^T \Sigma^{-1}(n) (\hat{x}(n) - \hat{y}(n))$</td>
<td></td>
</tr>
<tr>
<td>R.L.R.</td>
<td>$q(n) = (\hat{x}(n) - \hat{y}(n))^T \Sigma^{-1}(n) (\hat{x}(n) - \hat{y}(n))$</td>
<td></td>
</tr>
<tr>
<td>Kalman F.</td>
<td>$q(n) = \text{Tr}. [P(n)]$</td>
<td></td>
</tr>
<tr>
<td>Methods</td>
<td>Minimization criterion: Definition of weighting matrix</td>
<td></td>
</tr>
<tr>
<td>-----------</td>
<td>------------------------------------------------------</td>
<td></td>
</tr>
<tr>
<td>R.LS.</td>
<td>$\Sigma(n) = I ; \quad \Sigma^{-1}(n) = I$</td>
<td></td>
</tr>
<tr>
<td>R.W.L.S.</td>
<td>$\Sigma(n) = \begin{bmatrix} \frac{1}{\sigma_1} &amp; \cdots &amp; 0 \ \cdots &amp; \cdots &amp; \cdots \ 0 &amp; \cdots &amp; \frac{1}{\sigma_n} \end{bmatrix} \quad \Sigma^{-1}(n) = \begin{bmatrix} \frac{1}{\sigma_1} &amp; \cdots &amp; 0 \ \cdots &amp; \cdots &amp; \cdots \ 0 &amp; \cdots &amp; \frac{1}{\sigma_n} \end{bmatrix}$</td>
<td></td>
</tr>
<tr>
<td>R.W(t).L.S.</td>
<td>$\Sigma(n) = \begin{bmatrix} \frac{t_n - t_1}{\tau} &amp; \frac{c_1}{\sigma_1} &amp; \cdots &amp; 0 \ \cdots &amp; \cdots &amp; \cdots &amp; \cdots \ 0 &amp; \cdots &amp; \frac{t_n - t_n}{\tau} &amp; \frac{c_2}{\sigma_2} \end{bmatrix} \quad \Sigma^{-1}(n) = \begin{bmatrix} -\frac{t_n - t_1}{\tau} &amp; \frac{1}{\sigma_1} &amp; \cdots &amp; 0 \ \cdots &amp; \cdots &amp; \cdots &amp; \cdots \ 0 &amp; \cdots &amp; -\frac{t_n - t_n}{\tau} &amp; \frac{c_2}{\sigma_2} \end{bmatrix}$</td>
<td></td>
</tr>
<tr>
<td>R.L.R.</td>
<td>$\Sigma(n) = \begin{bmatrix} \Sigma(1) &amp; \cdots &amp; 0 \ \cdots &amp; \cdots &amp; \cdots \ 0 &amp; \cdots &amp; \Sigma(n) \end{bmatrix} \quad \Sigma(1) \delta_{ij} = E \left[ \hat{v}(i) \hat{v}^T(j) \right]$</td>
<td></td>
</tr>
<tr>
<td>Kalman f.</td>
<td>$\Sigma(n) = \begin{bmatrix} \Sigma(1) &amp; \cdots &amp; 0 \ \cdots &amp; \cdots &amp; \cdots \ 0 &amp; \cdots &amp; \Sigma(n) \end{bmatrix} \quad \Sigma(1) \delta_{ij} = E \left[ \hat{v}(i) \hat{v}^T(j) \right]$</td>
<td></td>
</tr>
<tr>
<td>Method</td>
<td>Optimal non-recursive estimation procedure</td>
<td></td>
</tr>
<tr>
<td>------------</td>
<td>------------------------------------------------------------------------------------------------------------</td>
<td></td>
</tr>
<tr>
<td>R.L.S.</td>
<td>$\hat{x}(n) = \left( B^T(n)B(n) \right)^{-1} B^T(n)\hat{Y}(n)$</td>
<td></td>
</tr>
<tr>
<td>R.W.L.S.</td>
<td>$\hat{x}(n) = \left( B^T(n)\Sigma^{-1}(n)B(n) \right)^{-1} B^T(n)\Sigma^{-1}(n)\hat{Y}(n)$</td>
<td></td>
</tr>
<tr>
<td>R.W(t).L.S.</td>
<td>$\hat{x}(n) = \left( B^T(n)\Sigma^{-1}(n)B(n) \right)^{-1} B^T(n)\Sigma^{-1}(n)\hat{Y}(n)$</td>
<td></td>
</tr>
<tr>
<td>R.L.R.</td>
<td>$\hat{x}(n) = \left( B^T(n)\Sigma^{-1}(n)B(n) \right)^{-1} B^T(n)\Sigma^{-1}(n)\hat{Y}(n)$</td>
<td></td>
</tr>
<tr>
<td>Kalman f.</td>
<td>This expression type has not been defined here because of the inclusion of the assumed system dynamics in the filter formulation (see tables 1, 8, 9 and 11). Moreover the Kalman filter is always used as an recursive estimator.</td>
<td></td>
</tr>
<tr>
<td>Method</td>
<td>Definition of error covariance matrix $P(n)$</td>
<td></td>
</tr>
<tr>
<td>-----------------</td>
<td>----------------------------------------------------------------</td>
<td></td>
</tr>
<tr>
<td>R.L.S.</td>
<td>$P(n) := (B^T(n)B(n))^{-1}$</td>
<td></td>
</tr>
<tr>
<td>R.W.L.S.</td>
<td>$P(n) := (B^T(n)\Sigma^{-1}(n)B(n))^{-1}$</td>
<td></td>
</tr>
<tr>
<td>R.W(t).L.S.</td>
<td>$P(n) := (B^T(n)\Sigma^{-1}(n)B(n))^{-1}$</td>
<td></td>
</tr>
<tr>
<td>R.L.R.</td>
<td>$P(n) := (B^T(n)\Sigma^{-1}(n)B(n))^{-1} = E \left[ (\hat{x} - \hat{x}(n))(\hat{x} - \hat{x}(n))^T \right]$</td>
<td></td>
</tr>
<tr>
<td>Kalman f.</td>
<td>$P(n) := E \left[ (\hat{x}(n) - \hat{x}(n))(\hat{x}(n) - \hat{x}(n))^T \right]$</td>
<td></td>
</tr>
<tr>
<td>Methods</td>
<td><strong>Table 8</strong> ( \text{Best prediction of } \hat{x}(n+1) \text{ at } t_n )</td>
<td><strong>Table 9</strong> ( \text{Best prediction of } P(n+1) \text{ at } t_n )</td>
</tr>
<tr>
<td>---------</td>
<td>-------------------------------------------------</td>
<td>-------------------------------------------------</td>
</tr>
<tr>
<td>R.L.S.</td>
<td>( \hat{x}(n+1</td>
<td>n) = \hat{x}(n) )</td>
</tr>
<tr>
<td>R.W.L.S.</td>
<td>( \hat{x}(n+1</td>
<td>n) = \hat{x}(n) )</td>
</tr>
<tr>
<td>R.W(t).L.S.</td>
<td>( \hat{x}(n+1</td>
<td>n) = \hat{x}(n) )</td>
</tr>
<tr>
<td>R.L.R.</td>
<td>( \hat{x}(n+1</td>
<td>n) = \hat{x}(n) )</td>
</tr>
<tr>
<td>Kalman f.</td>
<td>( \hat{x}(n+1</td>
<td>n) = \hat{E}(n+1,n)\hat{x}(n) + \Gamma(n+1,n)\hat{U}(n) )</td>
</tr>
<tr>
<td>Methods</td>
<td>Recursive estimation procedure: Gain matrix</td>
<td></td>
</tr>
<tr>
<td>--------------</td>
<td>---------------------------------------------------------------------------------</td>
<td></td>
</tr>
<tr>
<td>R.L.S.</td>
<td>$K(n+1) = P(n+1</td>
<td>n)M^T(n+1) \left(1 + M(n+1)P(n+1</td>
</tr>
<tr>
<td>R.W.L.S.</td>
<td>$K(n+1) = P(n+1</td>
<td>n)M^T(n+1)\left(\sigma^2(n+1) + M(n+1)P(n+1</td>
</tr>
<tr>
<td>R.W(t).L.S.</td>
<td>$K(n+1) = P(n+1</td>
<td>n)M^T(n+1)\left(\Sigma(n+1) + M(n+1)P(n+1</td>
</tr>
<tr>
<td>R.L.R.</td>
<td>$K(n+1) = P(n+1</td>
<td>n)M^T(n+1)\left(\Sigma(n+1) + M(n+1)P(n+1</td>
</tr>
<tr>
<td>Kalman f.</td>
<td>$K(n+1) = P(n+1</td>
<td>n)M^T(n+1)\left(\Sigma(n+1) + M(n+1)P(n+1</td>
</tr>
<tr>
<td>Methods</td>
<td>Recursive estimation procedure: estimation procedure.</td>
<td></td>
</tr>
<tr>
<td>--------------</td>
<td>--------------------------------------------------------</td>
<td></td>
</tr>
<tr>
<td>R.L.S.</td>
<td>( \hat{\mathbf{x}}(n+1) = \hat{\mathbf{x}}(n+1</td>
<td>n) + K(n+1) \left( y(n+1) - M(n+1)\hat{\mathbf{x}}(n+1</td>
</tr>
<tr>
<td>R.W.L.S.</td>
<td>( \hat{\mathbf{x}}(n+1) = \hat{\mathbf{x}}(n+1</td>
<td>n) + K(n+1) \left( y(n+1) - M(n+1)\hat{\mathbf{x}}(n+1</td>
</tr>
<tr>
<td>R.W(t).L.S.</td>
<td>( \hat{\mathbf{x}}(n+1) = \hat{\mathbf{x}}(n+1</td>
<td>n) + K(n+1) \left( y(n+1) - M(n+1)\hat{\mathbf{x}}(n+1</td>
</tr>
<tr>
<td>R.L.R.</td>
<td>( \hat{\mathbf{x}}(n+1) = \hat{\mathbf{x}}(n+1</td>
<td>n) + K(n+1) \left( \hat{y}(n+1) - M(n+1)\hat{\mathbf{x}}(n+1</td>
</tr>
<tr>
<td>Kalman f.</td>
<td>( \hat{\mathbf{x}}(n+1) = \hat{\mathbf{x}}(n+1</td>
<td>n) + K(n+1) \left( \hat{y}(n+1) - M(n+1)\hat{\mathbf{x}}(n+1</td>
</tr>
<tr>
<td>Methods</td>
<td>Recursive estimation procedure: Error propagation equation</td>
<td></td>
</tr>
<tr>
<td>------------------</td>
<td>-----------------------------------------------------------</td>
<td></td>
</tr>
<tr>
<td>R.L.S.</td>
<td>( P(n+1) = P(n+1</td>
<td>n) - K(n+1) M(n+1) P(n+1</td>
</tr>
<tr>
<td>R.W.L.S.</td>
<td>( P(n+1) = P(n+1</td>
<td>n) - K(n+1) M(n+1) P(n+1</td>
</tr>
<tr>
<td>R.W(t).L.S.</td>
<td>( P(n+1) = P(n+1</td>
<td>n) - K(n+1) M(n+1) P(n+1</td>
</tr>
<tr>
<td>R.L.R.</td>
<td>( P(n+1) = P(n+1</td>
<td>n) - K(n+1) M(n+1) P(n+1</td>
</tr>
<tr>
<td>Kalman f.</td>
<td>( P(n+1) = P(n+1</td>
<td>n) - K(n+1) M(n+1) P(n+1</td>
</tr>
</tbody>
</table>
Appendix A  Relation between $B(n)$, $\Sigma(n)$ and $P(n)$.

For the recursive weighted least-squares method or the recursive linear regression method for scalar valued observations $y(i)$, where $i = 1, 2, \ldots, n$, the minimum variance state vector estimate $\hat{X}(n)$ is:

(A-1) \[ \hat{X}(n) = \left( B^T(n) \Sigma^{-1}(n) B(n) \right)^{-1} B^T(n) \Sigma^{-1}(n) \hat{Y}(n) \]

Consider now the estimation error covariance matrix:

(A-2) \[ P(n) = E \left[ (\hat{X}(n) - \hat{X}(n))(\hat{X}(n) - \hat{X}(n))^T \right] \]

The minimum value of $\text{Tr} \{ P(n) \}$ is obtained if $\hat{X}(n)$ is the minimum variance state vector estimate:

(A-3) \[ \text{Min} \left[ \text{Tr} \{ P(n) \} \right] = \text{Tr} \left[ E \left[ \left( \hat{X}(n) - \left( B^T(n) \Sigma^{-1}(n) B(n) \right)^{-1} B^T(n) \Sigma^{-1}(n) \hat{Y}(n) \right)^T \right] \right] \]

The integral observation process was described in table 2 as:

(A-4) \[ \hat{Y}(n) = B(n) \hat{X}(n) + \hat{V}(n) \]
for the four classical estimators. Substitution of this relation in expression (A-3) and omission of the trace function yields the following result:

\[
(A-5) \quad P(n) = \left( B(n)^T \Sigma^{-1}(n) B(n) \right)^{-1} B(n)^T \Sigma^{-1}(n) \begin{bmatrix} \hat{\nu}(n) \hat{\nu}(n)^T \end{bmatrix} \cdot \Sigma^{-1}(n) B(n) \left( B(n)^T \Sigma^{-1}(n) B(n) \right)^{-1}.
\]

By definition:

\[
(A-6) \quad \Sigma(n) \triangleq \mathbb{E} \left[ \hat{\nu}(n) \hat{\nu}(n)^T \right]
\]

(see table 5). Here \( \Sigma \) is supposed to be an nxn matrix, not necessarily diagonal. Hence expression (A-5) reduces to:

\[
(A-7) \quad P(n) = \left( B(n)^T \Sigma^{-1}(n) B(n) \right)^{-1}
\]

This result holds if and only if \( \hat{X}(n) \) is the minimum variance state vector estimate.
Appendix B  Estimation mechanism of the Kalman filter and the recursive linear regression method.

The expressions for the recursive estimation procedures given in table 11 are quite similar. The estimation mechanism underlying these expressions is not similar for all minimum variance procedures. The main difference between the Kalman filter procedure and the other, classical procedures will be apparent from tables 8 and 9. The Kalman filter is designed to include system dynamics and system input noise statistics as well as observation noise statistics in the estimation process. The classical recursive weighted least-squares method and the recursive linear regression method only include observation error statistics in the estimation process.

To clarify the estimation mechanism of the Kalman filter, the filter equation (table 11) is reformulated:

\[(B-1) \hat{X}(n+1) = (I - K(n+1)M(n+1))\hat{X}(n+1|n) + K(n+1)\hat{Y}(n+1)\]

Left hand multiplication of this expression by \(M(n+1)\) yields:

\[(B-2) M(n+1)\hat{X}(n+1) = (M(n+1) - M(n+1)K(n+1)M(n+1))\hat{X}(n+1) + M(n+1)K(n+1)\hat{Y}(n+1)\]

or:

\[(B-3) \hat{Y}(n+1) = (I - M(n+1)K(n+1))\hat{Y}(n+1|n) + M(n+1)K(n+1)\hat{Y}(n+1)\]

\(')\) This transformation is introduced only for the sake of this discussion.
where the estimated observation:

\[(\text{B-4}) \quad \hat{Y}(n+1) = M(n+1)\hat{X}(n+1)\]

Now define:

\[(\text{B-5}) \quad T(n+1) = M(n+1)K(n+1)\]

Hence:

\[(\text{B-6}) \quad \hat{Y}(n+1) = (I - T(n+1))\hat{Y}(n+1|n) + T(n+1)\hat{Y}(n+1)\]

It will be evident from the expression for the Kalman filter gain matrix (Table 10) that:

\[(\text{B-7}) \quad T(n+1) = M(n+1)P(n+1|n)M^T(n+1).\]

\[\cdot \left[ \Sigma(n+1) + M(n+1)P(n+1|n)M^T(n+1) \right]^{-1}\]

According to table 9:

\[(\text{B-8}) \quad P(n+1|n) = \Phi(n+1,n)P(n)\Phi^T(n+1,n) + \]

\[\Gamma(n+1,n)\Omega(n)\Gamma^T(n+1,n)\]
Substitution of this expression in equation (B-7) yields:

\[(B-9) \quad T(n+1) = \left[ M(n+1) \Phi(n+1, n)P(n)\Phi^T(n+1, n)M^T(n+1) + 
+ M(n+1)\Gamma(n+1, n)\Omega(n)\Gamma^T(n+1, n)M^T(n+1) \right] \cdot \left[ \Sigma(n+1) + M(n+1)\Phi(n+1, n)P(n)\Phi^T(n+1, n)M^T(n+1) + 
+ M(n+1)\Gamma(n+1, n)\Omega(n)\Gamma^T(n+1, n)M^T(n+1) \right]^{-1} \]

It will now be shown that:

\[(B-10) \quad o \leq T(n+1) \leq I \]

If at time \( t_{n+1} \):

\[(B-11) \quad \Pi(n+1) = o \quad ') \]

then:

\[(B-12) \quad T(n+1) = I \]

') The inverse matrix part of the gain matrix does not exist if \( \Pi(n+1) = o \), while either \( \| P(n) \| \sim \| \Omega(n) \| \sim o \), or the rank of \( M^T(n+1)M(n+1) < r \) (the observation space)
\( \hat{Y}(n+1) \) is then determined by the second term of expression (B-6), i.e. the actual observation \( \bar{Y}(n+1) \). If:

\[
(B-13) \quad \Sigma(n+1) \gg M(n+1) \Gamma(n+1,n) \Omega(n) \Gamma^T(n+1,n)
\]

then:

\[
(B-14) \quad T(n+1) \approx 0
\]

\( \hat{Y}(n+1) \) is then determined by the first term, i.e. the predicted observation \( \hat{Y}(n+1|n) \), of expression (B-6). For any value of the matrix \( T(n+1) \) between the zero matrix 0 and the unit matrix I, \( \hat{Y}(n+1) \) equals a weighted sum of the predicted observation \( \hat{Y}(n+1|n) \) and the actual observation \( \bar{Y}(n+1) \).

The weighting matrix \( T(n+1) \) depends in fact on the proportion of the input signal noise covariance matrix \( \Omega(n) \) to the observation noise covariance matrix \( \Sigma(n+1) \). This proportion is effected by the state transition matrix \( \Phi(n+1|n) \) and the input signal distribution matrix \( \Gamma(n+1,n) \).

Remark

The state estimate \( \hat{X}(n+1) \) can be derived from \( \hat{Y}(n+1) \) by writing

\[
(B-15) \quad \hat{X}(n+1) = (M^T(n+1)M(n+1))^{-1} M^T(n+1) \hat{Y}(n+1)
\]

if \( M^T(n+1)M(n+1) \) is a nonsingular matrix, i.e. if \( \hat{X}(n+1) \) is observable.
In the recursive linear regression method and the recursive weighted least-squares method no input signal noise statistics is included. Hence:

\[(B-16) \quad \Omega(n) \equiv 0\]

(see expression B-9). If the recursive linear regression method yields a very small estimation error covariance matrix \(P(n)\):

\[(B-17) \quad \|P(n)\| \to 0\]

then

\[(B-18) \quad T(n+1) \to 0\]

Hence the estimate of \(\hat{X}(n+1)\) depends increasingly on the predicted observation \(\hat{Y}(n+1|n)\).

If:

\[(B-19) \quad \|P(n)\| \approx 0\]

new observations are omitted by these estimation procedures.
References:


In: "Advances in Control Systems" Vol. 3. Ed. C.T. Leondes


In: "Advances in Control Systems" Vol. 6 1968 Ed. C.T. Leondes
Academic Press. p. 95 — 158.