State estimation and the Kalman filter

PhD, David Di Ruscio
Telemark university college
Department of Technology
Systems and Control Engineering
N-3914 Porsgrunn, Norway
Fax: +47 35 57 52 50
Tel: +47 35 57 51 68
E-mail: David.Di.Ruscio@hit.no[1]

October 23, 2009

[1]The contents in this note are based on the lecture notes witch are used in a course in Advanced Control Theory.

Chapter 1

Introduction

This note is meant to cover parts of the syllabus in state estimation and Kalman filter theory in the System Identification and Optimal estimation course. Syllabus in optimal estimation is Sections 2.2,2.6.2 and 2.6.3. Section 2.1, 2.3, 2.4 and 2.5 is not syllabus in this course but syllabus in a course in Advanced control theory.

Chapter 2

State estimation and the Kalman filter

2.1 Continuous estimator and regulator duality

It can be shown that the solution to the Linear Quadratic optimal control problem is dual to the optimal minimum variance estimator problem, Kalman filter. This means that if we know the solution to the LQ optimal control problem, then we can directly write down the solution to the optimal estimator problem by using the duality principle. However, note that the LQ optimal control problem is a topic of a course in Advanced control theory.

The duality principle can be formulated in the following table

$$\begin{array}{ccc} \text{Regulator} & & \text{Estimator} \\ A & \rightarrow & A^T \\ B & \rightarrow & D^T \\ Q & \rightarrow & V \\ P & \rightarrow & W \\ G & \rightarrow & -K^T \\ A+BG & \rightarrow & (A^T - D^T K^T)^T \\ R & \rightarrow & X \\ -t & \rightarrow & t \\ \dot{R} & \rightarrow & -\dot{X} \end{array} \tag{2.1}$$

As we know from the solution of the LQ optimal control problem the Riccati equation is solved backward in time from the final time instant, i.e. recursively from the final value, $R(t_1) = S$. The solution to the dual minimum variance estimator problem is also containing a Riccati equation. The Riccati equation in the dual estimator problem is however solved forward in time with initial values given at the start time. This is the reason why we have specified $-t$ in the table for the LQ control problem and t in connection with the dual estimator problem.

2.2 Minimum variance estimation in linear continuous systems

Given a linear dynamic system described by

$$\dot{x} = Ax + Bu + v, \quad (2.2)$$

$$y = Dx + Eu + w, \quad (2.3)$$

where v is uncorrelated white process noise with zero mean and covariance matrix V and w is uncorrelated white measurements noise with zero mean and covariance matrix W, i.e. such that

$$V = \mathrm{E}(vv^T), \quad (2.4)$$

$$W = \mathrm{E}(ww^T). \quad (2.5)$$

We assume that A, B, D and E are known model matrices. Furthermore we assume that the covariance matrices V and W are known or specified and that the measurements vector y is measured and given. We also assume that the matrix pair A, D is observable. Since the state vector x is not measured it can be estimated in a so called state estimator or state observer.

The principle of duality in connection with the solution of the Linear Quadratic (LQ) optimal control problem can be used to find the solution to the optimal minimum variance estimation problem.

Note that we have from the duality principle that $\dot{R} \rightarrow \frac{dX}{d(-t)} = -\dot{X}$. using the duality principle we have that

$$\dot{X} = AX + XA^T - XD^TW^{-1}DX + V, \;\; X(t_0) \text{ given}, \quad (2.6)$$

which is a matrix Riccati equation which defines X. The Kalman filter gain matrix is then given by

$$K^T = W^{-1}DX. \quad (2.7)$$

Let us define the error between the actual state, x, and the estimated state, $\hat{x}$, as follows

$$\Delta x = x - \hat{x}. \quad (2.8)$$

It can be shown that the solution to the riccati equation, X, is the covariance matrix of the error between x and the estimate $\hat{x}$, i.e.

$$X = \mathrm{E}[(x - \hat{x})(x - \hat{x})^T] = \mathrm{E}[\Delta x \Delta x^T]. \quad (2.9)$$

The state estimator is then given by

$$\dot{\hat{x}} = A\hat{x} + Bu + K(y - \hat{y}), \quad (2.10)$$

$$\hat{y} = D\hat{x} + Eu. \quad (2.11)$$

$\hat{x}$ is the minimum variance estimate of the state vector x in the sense that X is minimized. Note also that $\hat{y}$ is the optimal prediction of the measurements vector y, given all old outputs y and given all old input vectors u as vell as the present input at the present time t.

The reason for that $\hat{y}$ is dependent of the input u at present time t is the direct feed through term matrix E. However E is in principle always zero for continuous systems, but a nonzero E may be the results of some model reduction procedures. Note also that a non zero E often is the case in discrete time systems due to sampling.

Equations (2.10) and (2.11) gives the following equation for the state estimate

$$\dot{\hat{x}} = (A - KD)\hat{x} + (B - KE)u + Ky, \tag{2.12}$$

where the initial state estimate $\hat{x}(t_0)$ is given.

Note that the eigenvalues of the matrix $A - KD$ defines the stability properties of the estimator. It make sense that K is so that $A - KD$ is stable, i.e., all eigenvalues in the left half of the complex plane. the reason for this is that $\hat{x}$ is given from a differential equation driven by known inputs u and known outputs y. Note also that when $A - kD$ is stable then the effect of wrong initial values $\hat{x}(t_0)$ will die out when $t \to \infty$.

Let us study the properties of the estimator by studying the excepted value of the error in the state estimate Δx. From the definition (2.8) we have that

$$\dot{\Delta x} = \dot{x} - \dot{\hat{x}}. \tag{2.13}$$

Using (2.2) and (2.10) gives

$$\dot{\Delta x} = Ax + Bu + v - [A\hat{x} + Bu + K(y - \hat{y})]. \tag{2.14}$$

using (2.3) and (2.11) gives

$$\dot{\Delta x} = Ax + Bu + v - [A\hat{x} + Bu + K(Dx + Eu + w - D\hat{x} - Eu)], \tag{2.15}$$

which gives

$$\dot{\Delta x} = (A - KD)\Delta x + v - Kw. \tag{2.16}$$

The excepted value of the estimated error, Δx, is then given by

$$\mathrm{E}\{\dot{\Delta x}\} = (A - KD)\mathrm{E}\{\Delta x\}. \tag{2.17}$$

The stability properties of the estimator can be analyzed by studying the estimation error when $t \to \infty$.

It can be shown that the minimum variance estimator is stable. This can be argued from the fact that the LQ optimal controller is stable (by properly choice of some weighting matrices) and that the optimal minimum variance estimator is dual to the LQ controller. Hence, a similar stability theorem exists for the optimal minimum variance estimator.

In the following a different argumentation for stability will be given. Assume that v and w is uncorrelated white noise stationary processes. Then the covariance matrices will be constant and positive definite, i.e., $V > 0$ and $W > 0$. Letting $t \to \infty$ then we have that X is a solution to the stationary algebraic matrix Riccati equation

$$AX + XA^T - XD^T W^{-1} DX + V = 0. \tag{2.18}$$

This can be written as a Lyapunov matrix equation, i.e.,

$$(A - KD)X + X(A - KD)^T = -(V + KWK^T). \tag{2.19}$$

From the discussion above it is clear that $X > 0$ and $V + KWK^T > 0$. From Lyapunovs stability theory we then know that $A - KD$ is a stable matrix, i.e. all eigenvalues of $A - KD$ lies in the left half of the complex plane.

It is clear that when $A-KD$ is a stable matrix then the excepted value is $\mathrm{E}\{\dot{\Delta x}\} = 0$. From (2.17) we then have that $0 = (A-KD)\mathrm{E}\{\Delta x\}$. This implies that $\mathrm{E}\{\Delta x\} = 0$.

Another alternative is to analyze the error from the solution of (2.17). We have

$$\lim_{t\to\infty} \mathrm{E}\{\Delta x\} = \lim_{t\to\infty} [e^{(A-KD)(t-t_0)}]\mathrm{E}\{\Delta x(t_0)\} = 0, \tag{2.20}$$

which is valid even if $\mathrm{E}\{\Delta x(t_0)\} \neq 0$.

2.3 Separasjonsteoremet: Kontinuerlig tid

Theorem 2.3.1 (Separasjonsteoremet: Kontinuerlig tid)
Gitt et lineært stokastisk system

$$\dot{x} = Ax + Bu + Cv, \tag{2.21}$$

$$y = Dx + w, \tag{2.22}$$

der v og w er ukorrelerte hvite prosesser wed null middelverdi og kovariansmatriser henholdsvis V og W.

Systemet skal reguleres slik at objektfunksjonalen

$$J = \frac{1}{2}\mathrm{E}\{x^T(t_1)Sx(t_1) + \int_{t_0}^{t_1} [x^T Q x + u^T P u] dt\}, \tag{2.23}$$

minimaliseres med hensyn på pådragsvektoren $u(t)$ i tidsintervallet $t_0 \leq t < t_1$.

Løsningen på dette stokastiske optimalreguleringsproblemet er gitt ved

$$u = G(t)\hat{x}. \tag{2.24}$$

G er den tilbakekoblingsmatrisen vi finner ved å løse det tilsvarende deterministiske LQ optimalreguleringsproblemet der hele tilstandsvektoren er kjent. Dvs. med $v = 0$ og $w = 0$ i (2.21) og (2.22) og samme LQ kriterium som i (2.23). Det er ikke behov for forventningsoperatoren $\mathrm{E}(\{\cdot\}$ i det deterministiske tilfellet. Dvs. at G er gitt av

$$G(t) = -P^{-1}B^T R \tag{2.25}$$

der R er den maksimale (positive) løsning av Riccati-ligningen

$$-\dot{R} = A^T R + RA - RBP^{-1}B^T R + Q, \;\; R(t_1) = S. \tag{2.26}$$

$\hat{x}$ er minimum varians estimatet av tilstandsvektoren x. $\hat{x}$ er gitt av Kalman-filteret for systemet som er gitt ved

$$\dot{\hat{x}} = A\hat{x} + Bu + K(y - D\hat{x}), \tag{2.27}$$

med gitt initialverdi $\hat{x}(t_0)$ og der Kalmanfilter forsterkningsmatrisen K er gitt ved

$$K(t) = XD^T W^{-1}, \tag{2.28}$$

og der X er den maksimale (positive) løsning av Riccati-ligningen

$$\dot{X} = AX + XA^T - XD^T W^{-1} DX + CVC^T, \;\; X(t_0) = \text{gitt}. \tag{2.29}$$

△

Ofte benytter vi en uendelig horisont $t_1 \rightarrow \infty$. Dette leder til den stasjonære Riccatiligningen ($\dot{R} = 0$) og den stasjonære Riccatligningen for X, dvs. med $\dot{X} = 0$ i (2.29). Vi har i dette tilfellet at G og K blir konstante matriser.

2.4 Kontinuerlig LQG regulator

Med en LQG (Linear Quadratic Gausian) regulator menes en regulator der vi benytter en LQ optimal tilbakekoblingsmatrise G og en tilbakekobling fra ett optimalt minimum variansestimat, $\hat{x}$, av tilstandsvektoren x. Denne strategien er vanlig der man ikke kan måle hele tilstandsvektoren. Prinsippet er gitt ved prosessmodellen

$$\dot{x} = Ax + Bu, \tag{2.30}$$
$$y = Dx, \tag{2.31}$$

tilstandsestimatoren

$$\dot{\hat{x}} = A\hat{x} + Bu + K(y - \hat{y}), \tag{2.32}$$
$$\hat{y} = D\hat{x}, \tag{2.33}$$

og regulatoren

$$u = G\hat{x}. \tag{2.34}$$

Vi vil nå gi en analyse av totalsystemet. Merk at denne analysen er gyldig for vilkårlige G og K.

Dette leder til det autonome systemet

$$\begin{bmatrix} \dot{x} \\ \dot{\hat{x}} \end{bmatrix} = \begin{bmatrix} A & BG \\ KD & A + BG - KD \end{bmatrix} \begin{bmatrix} x \\ \hat{x} \end{bmatrix}. \tag{2.35}$$

For å kunne studere stabilitetsegenskapene til totalsystemet er det hensiktsmessig å transformere systemet v.h.a.

$$\begin{bmatrix} x \\ x - \hat{x} \end{bmatrix} = \begin{bmatrix} x \\ \Delta x \end{bmatrix} = \begin{bmatrix} I & 0 \\ I & -I \end{bmatrix} \begin{bmatrix} x \\ \hat{x} \end{bmatrix}. \tag{2.36}$$

Dette gir det autonome systemet

$$\begin{bmatrix} \dot{x} \\ \dot{\Delta x} \end{bmatrix} = \overbrace{\begin{bmatrix} A + BG & -BG \\ 0 & A - KD \end{bmatrix}}^{\bar{A}_{tc}} \begin{bmatrix} x \\ \Delta x \end{bmatrix}. \tag{2.37}$$

fordi

$$\begin{bmatrix} I & 0 \\ I & -I \end{bmatrix}^{-1} = \begin{bmatrix} I & 0 \\ I & -I \end{bmatrix}. \tag{2.38}$$

Vi ser her at stabiliteten til totalsystemet er gitt ved de n egenverdiene til matrisen $A+BG$ og de n egenverdiene til estimator matrisen $A-KD$. Systemmatrisen $\bar{A}_{tc}$ har dermed $2n$ egenverdier. Det er en tommelfingerregel å tune estimatorforsterkningen K slik at egenverdiene til $A-KD$ ligger noe til venstre for egenverdiene til $A+BG$ i venstre del av det komplekse plan. Vi skal merke oss at dette resultatet bare gjelder dersom vi ikke har modellfeil. Dersom vi har modellfeil vil det i beste fall være en approksimasjon.

2.5 Discrete time LQG controller

2.5.1 Analysis of discrete time LQG controller

We will in this section discuss the discrete time LQG controller. We assume that the process is described by

$$x_{k+1} = Ax_k + B_p u_k, \tag{2.39}$$

$$y_k = Dx_k. \tag{2.40}$$

The controller is of the form

$$u_k = G\hat{x}_k. \tag{2.41}$$

where $\hat{x}_k$ is given by the state observer

$$\bar{y}_k = D\bar{x}_k \tag{2.42}$$

$$\hat{x}_k = \bar{x}_k + K(y_k - \bar{y}_k), \tag{2.43}$$

$$\bar{x}_{k+1} = A\hat{x}_k + Bu_k. \tag{2.44}$$

where $\bar{x}_0$ is given. Here $\bar{x}_k$ is defined as the a-priori estimate of x_k. Furthermore we define $\hat{x}_k$ as the a-posteriori estimate of x_k. We assume that the feedback matrix G is computed based on the model matrices A, B. The observer gain matrix K is computed based on the model matrices A, D.

We see that we have a perfect model is $B = B_p$. If $B \neq B_p$ then we have modeling errors. Let us in the following study the entire closed loop system. Putting (2.41) into (2.39) and (2.44) and we obtain

$$x_{k+1} = Ax_k + B_p G\hat{x}_k, \tag{2.45}$$

$$\bar{x}_{k+1} = (A + BG)\hat{x}_k. \tag{2.46}$$

We may now eliminate $\hat{x}_k$ from (2.45) and (2.46) by using (2.43).

$$x_{k+1} = (A + B_p GKD)x_k + B_p G(I - KD)\bar{x}_k, \tag{2.47}$$

$$\bar{x}_{k+1} = (A + BG)KDx_k + (A + BG)(I - KD)\bar{x}_k. \tag{2.48}$$

This means that we have an autonomous system

$$\begin{bmatrix} x_{k+1} \\ \bar{x}_{k+1} \end{bmatrix} = \overbrace{\begin{bmatrix} A + B_p GKD & B_p G(I - KD) \\ (A + BG)KD & (A + BG)(I - KD) \end{bmatrix}}^{A_{td}} \begin{bmatrix} x_k \\ \bar{x}_k \end{bmatrix}. \tag{2.49}$$

The entire system is stable if the $2n$ eigenvalues of the matrix A_{td} is located inside the unit circle in the complex plane. Let us use the transformation (2.36). This gives

$$\begin{bmatrix} x_{k+1} \\ x_{k+1} - \bar{x}_{k+1} \end{bmatrix} = \overbrace{\begin{bmatrix} A + B_p G & -B_p G(I - KD) \\ (B_p - B)G & A - AKD - (B_p - B)G(I - KD) \end{bmatrix}}^{\bar{A}_{td}} \begin{bmatrix} x_k \\ x_k - \bar{x}_k \end{bmatrix} \tag{2.50}$$

In case of a perfect model, i.e., $B = B_p$, we se that the eigenvalues of the total system is given by the n eigenvalues of the matrix $A + BG$ and the n egenvalues of the observer system matrix $A - AKD$.

This also means that in case of modeling errors we have to check the eigenvalues/poles of the system matrix for the entire system, i.e., $\bar{A}_{td}$ for different cases of model errors B_p.

Note also that a rule of thumb is that the eigenvalues of the observer matrix $A - AKD$ should be ten times faster than the eigenvalues of the controller feedback matrix $A + BG$.

2.6 The discrete Kalman filter

2.6.1 Innovation formulation of the Kalman filter

Given a process

$$x_{k+1} = Ax_k + Bu_k + v_k, \tag{2.51}$$

$$y_k = Dx_k + w_k, \tag{2.52}$$

where v_k is white process noise and w_k is white measurements noise with known covariance matrices.

First, let us present the apriori-aposteriori formulation of the discrete time optimal minimum variance Kalman filter as follows

$$\bar{y}_k = D\bar{x}_k \tag{2.53}$$

$$\hat{x}_k = \bar{x}_k + K(y_k - \bar{y}_k), \tag{2.54}$$

$$\bar{x}_{k+1} = A\hat{x}_k + Bu_k. \tag{2.55}$$

where $\bar{x}_0$ is a given initial value for the apriori or predicted state estimate. Here, $\bar{x}_k$ is defined as the apriori or predicted state estimate of the state vector x_k. Furthermore, $\hat{x}_k$ is defined as the aposteriori state estimate of x_k. The apriori-aposteriori Kalman filter is further discussed in Section 2.6.3.

Note that $\hat{x}_k$ can be eliminated from the estimator equation (2.55), i.e. an equivalent estimator for the predicted state $\bar{x}_k$ is given by

$$\bar{y}_k = D\bar{x}_k, \tag{2.56}$$

$$\bar{x}_{k+1} = A\bar{x}_k + Bu_k + \tilde{K}(y_k - \bar{y}_k). \tag{2.57}$$

$$= (A - \tilde{K}D)\bar{x}_k + Bu_k + \tilde{K}y_k, \tag{2.58}$$

where

$$\tilde{K} = AK. \tag{2.59}$$

It is the apriori estimate, $\bar{x}_k$ which is the essential state in the estimator. $\bar{x}_k$ is also referred to as the predicted state.

The dynamics of the estimator is in this case described by the eigenvalues of the matrix $A - \tilde{K}D = A - AKD$. the estimator given by (2.56)-(2.57) above gives the optimal one step ahead prediction $\bar{y}_k$ of the output y_k. This formulation is used if we only want to compute the prediction of the output y_k. As a rule of thumb we may say that $\tilde{K} = AK$ is the Kalman filter gain for the prediction of y_k and for computing the predicted state $\bar{x}_k$.

Note also that if we are using $y_k = \bar{y}_k + \varepsilon_k$ where the predicted output is given by $\bar{y}_k = D\bar{x}_k$ then we obtain the innovations formulation of the Kalman filter, i.e.,

$$\bar{x}_{k+1} = A\bar{x}_k + Bu_k + \tilde{K}\varepsilon_k, \tag{2.60}$$

$$y_k = D\bar{x}_k + \varepsilon_k. \tag{2.61}$$

This means that $\tilde{K} = AK$ is the kalman filter gain in the innovations formulation (2.60)-(2.61) and K is the Kalman filter gain in the apriori-aposteriori formulation (2.42)-(2.44) of the Kalman filter.

Note that the above equations easily is extended to be valid for a proper system in which $y_k = D\bar{x}_k + Eu_k + \varepsilon_k$.

2.6.2 Development of the Kalman filter on innovations form

Given a process

$$x_{k+1} = Ax_k + v_k, \tag{2.62}$$

$$y_k = Dx_k + w_k, \tag{2.63}$$

where v_k is white process noise and w_k is white measurements noise with covariance matrices given by

$$\mathrm{E}(\begin{bmatrix} v_k \\ w_k \end{bmatrix} \begin{bmatrix} v_k \\ w_k \end{bmatrix}^T) = \begin{bmatrix} V & R_{12} \\ R_{12}^T & W \end{bmatrix} \tag{2.64}$$

The Kalman filter on innovations form is then given by

$$\bar{x}_{k+1} = A\bar{x}_k + \tilde{K}\varepsilon_k, \quad (2.65)$$

$$y_k = D\bar{x}_k + \varepsilon_k. \quad (2.66)$$

Note that the Kalman filter gain $\tilde{K}$ in the innovations formulation is related to the Kalman filter gain K in the apriori-aposteriori formulation as $\tilde{K} = AK$.

When analyzing the Kalman filter the estimating error $\Delta x_k = x_k - \bar{x}_k$ is of great importance. The equations for the estimating errors are obtained from the above equations. i.e. from the process model and the Kalman filter above, i.e.,

$$\Delta x_{k+1} = A\Delta x_k + v_k - \tilde{K}\varepsilon_k, \quad (2.67)$$

$$\varepsilon_k = D\Delta x_k + w_k, \quad (2.68)$$

$$\Delta x_k = x_k - \bar{x}_k. \quad (2.69)$$

The equations for the estimating error are to be used in the following discussions.

Equation for computing $\tilde{K}$ in the predictor

The development which is given here is based on the fact that the innovations process ε_k is white noise when the optimal Kalman filter gain $\tilde{K}$ is used in the filter. Since ε_k is white it is independent and uncorrelated with the estimation error Δx_{k+1}. Hence, by demanding

$$\mathrm{E}(\Delta x_{k+1}\varepsilon_k^T) = 0, \quad (2.70)$$

then we can derive an expression for $\tilde{K}$. We have that

$$\begin{aligned}\Delta x_{k+1}\varepsilon_k^T &= (A\Delta x_k + v_k - \tilde{K}\varepsilon_k)\varepsilon_k^T \\ &= A\Delta x_k\varepsilon_k^T + v_k\varepsilon_k^T - \tilde{K}\varepsilon_k\varepsilon_k^T \\ &= A\Delta x_k(\Delta x_k^T D^T + w_k^T) + v_k(\Delta x_k^T D^T + w_k^T) - K\varepsilon_k\varepsilon_k^T. \quad (2.71)\end{aligned}$$

Using this in (2.70) gives

$$\mathrm{E}(A\Delta x_k\Delta x_k^T D^T + v_k w_k^T - \tilde{K}\varepsilon_k\varepsilon_k^T) = 0, \quad (2.72)$$

where we have used that $\mathrm{E}(\Delta x_k v_k^T) = 0$ and $\mathrm{E}(\Delta x_k w_k^T) = 0$. We have then obtained an equation

$$AXD^T + R_{12} - \tilde{K}\Delta = 0, \quad (2.73)$$

where

$$\Delta = \mathrm{E}(\varepsilon_k\varepsilon_k^T) = DXD^T + W. \quad (2.74)$$

This gives the following expression for the Kalman filter gain

$$\tilde{K} = (AXD^T + R_{12})(DXD^T + W)^{-1}. \quad (2.75)$$

This is the equation for th Kalman filter gain in the innovations formulation of the Kalman filter. We now have to find an expression for the covariance matrix of the estimation error, $X = \mathrm{E}(\Delta x_k\Delta x_k^T)$. It can be shown that X is given as the solution of a matrix Riccati equation.

Equation for computing $X = \mathbf{E}(\Delta x_k \Delta x_k^T)$

The derivation of the riccati equation for computing the covariance matrix X is based that we under stationary conditions have that

$$\mathrm{E}(\Delta x_{k+1}\Delta x_{k+1}^T) = \mathrm{E}(\Delta x_k \Delta x_k^T) = X. \tag{2.76}$$

From equations (2.67) and (2.68) we have that

$$\Delta x_{k+1} = A\Delta x_k + v_k - \tilde{K}(D\Delta x_k + w_k), \tag{2.77}$$

which gives

$$\Delta x_{k+1} = (A - \tilde{K}D)\Delta x_k + v_k - \tilde{K}w_k. \tag{2.78}$$

we have that the estimation error Δx_k is uncorrelated with the white noise processes v_k and w_k. We then have that

$$\begin{aligned}
\Delta x_{k+1}\Delta x_{k+1}^T &= [(A - \tilde{K}D)\Delta x_k + v_k - \tilde{K}w_k][(A - \tilde{K}D)\Delta x_k + v_k - \tilde{K}w_k]^T \\
&= (A - \tilde{K}D)\Delta x_k \Delta x_k^T (A - \tilde{K}D)^T + (v_k - \tilde{K}w_k)(v_k - \tilde{K}w_k)^T \\
&= (A - \tilde{K}D)\Delta x_k \Delta x_k^T (A - \tilde{K}D)^T + v_k v_k^T - v_k w_k^T \tilde{K}^T \\
&- \tilde{K}(v_k w_k^T)^T + \tilde{K}w_k w_k^T \tilde{K}^T.
\end{aligned} \tag{2.79}$$

Using the mean operator $\mathrm{E}(\cdot)$ on both sides of the equal sign gives

$$X = (A - \tilde{K}D)X(A - \tilde{K}D)^T + V - R_{12}\tilde{K}^T - \tilde{K}R_{12}^T + \tilde{K}W\tilde{K}^T, \tag{2.80}$$

which also can be written as follows

$$X = (A - \tilde{K}D)X(A - \tilde{K}D)^T + \begin{bmatrix} I & \tilde{K} \end{bmatrix} \begin{bmatrix} V & R_{12} \\ R_{12}^T & W \end{bmatrix} \begin{bmatrix} I & \tilde{K} \end{bmatrix}^T. \tag{2.81}$$

Note that (2.80) and (2.81) is a discrete matrix Lyapunov equation in X when $\tilde{K}$ is given. A Lyapunov equation is a linear equation. The Lyapunov equation can e.g. simply be solved by using the MATLAB control system toolbox function **dlyap**. By substituting the expression for the Kalman filter gain $\tilde{K}$ given by (2.75) into (2.81) gives the discrete Riccati equation for computing the covariance matrix X, i.e.,

$$\begin{aligned}
X &= AXA^T + V - \tilde{K}(AXD^T + R_{12})^T \\
&= AXA^T + V - (AXD^T + R_{12})(DXD^T + W)^{-1}(AXD^T + R_{12})^T.
\end{aligned} \tag{2.82}$$

The stationar Riccati equation can simply be solved for X by iterating (2.82) until convergence. Another elegant method is to iterate both (2.75) and (2.80) until convergence and computing both $\tilde{K}$ and X at the same time. this is illustrated and implemented in the MATLAB function **dlqe2.m**.

```
function [K,X,itnum]=dlqe2(A,C,D,V,W,R12);
% DLQE2
% [K,X]=dlqe2(A,C,D,V,W,R12);
```

```
% This function computes the Kalman gain K in the Kalman filter on
% innovations form, and the covariance matrix X of the estimation
% error, i.e. the error between the state and the predicted state.

X=C*V*C';                               % Initial covariance matrix.
K=(A*X*D'+R12)*pinv(D*X*D'+W);          % The corresponding Kalman gain.
it=100;                                 % Maximum number of iterations.
Tol=1e-8;                               % Tolerance for norm(X(i)-X(i-1)).

Xold=X*0;                               % Iterate for the solution X of
for i=1:it;                             % the discrete Riccati equation.
   K=(A*X*D'+R12)*pinv(D*X*D'+W);
   AKD=A-K*D;
   X=AKD*X*AKD'+V-R12*K'-K*R12'+K*W*K';
   if norm(X-Xold) <= Tol
      itnum=i;
      break
   end
   Xold=X;
end
K=(A*X*D'+R12)*pinv(D*X*D'+W);
```

2.6.3 derivation of the Kalman filter on apriori-aposteriori form

Given a process

$$x_{k+1} = Ax_k + v_k, \tag{2.83}$$

$$y_k = Dx_k + w_k, \tag{2.84}$$

where v_k is white process noise and w_k is white measurements noise with covariance matrices given by

$$\mathrm{E}(\begin{bmatrix} v_k \\ w_k \end{bmatrix} \begin{bmatrix} v_k \\ w_k \end{bmatrix}^T) = \begin{bmatrix} V & R_{12} \\ R_{12}^T & W \end{bmatrix}. \tag{2.85}$$

We here note that the process noise v_k may be correlated with the measurements noise w_k, i.e. $\mathrm{E}(v_k w_k^T) = R_{12}$.

The kalman filter on apriori-aposteriori form is basically used when we are out for the optimal state estimate of x_k. The filter is of the form

$$\bar{y}_k = D\bar{x}_k \tag{2.86}$$

$$\hat{x}_k = \bar{x}_k + K(y_k - \bar{y}_k), \tag{2.87}$$

$$\bar{x}_{k+1} = A\hat{x}_k + R_{12}\Delta^{-1}(y_k - \bar{y}_)), \tag{2.88}$$

where the initial predicted state $\bar{x}_0$ is given or specified. Here $\bar{x}_k$ is defined as the apriori state estimate of x_k. the estimate $\bar{x}_k$ is also often referred to as the predicted state. Furthermore we define $\hat{x}_k$ as the aposteriori state estimate of x_k. Apriori means known in advance, and aposteriori means the new information which is obtained by the updating in (2.87), i.e., by using the apriori information and the new information in the measurement y_k. The reason for that the state estimate is divided into two parts $\bar{x}_k$ and $\hat{x}_k$ is mainly because the system is discrete time, e.g. because of sampling.

The kalman filter gain K in the filter given by (2.86)-(2.88) above is given by

$$K_k = \bar{X}_k D^T (D\bar{X}_k D^T + W)^{-1}, \tag{2.89}$$
$$\hat{X}_k = (I - K_k D)\bar{X}_k (I - K_k D)^T + K_k W K_k^T, \tag{2.90}$$
$$\bar{X}_{k+1} = A\hat{X}_k A^T + V + Z_k, \tag{2.91}$$

where

$$Z_k = -R_{12}\Delta^{-1}R_{12}^T - AK_kR_{12}^T - R_{12}K_k^TA^T. \tag{2.92}$$

Note that (2.91) contain an extra term given by Z_k when the process and measurements noise is correlated, this term is not present when $R_{12} = 0$, which usually is the case.

In order to start the filter process we need an initial value for the covariance matrix $\bar{X}_0$, i.e. when we look at the filter at time $k = 0$. Note that the covariance matrices are defined as follows

$$\bar{X}_k = \mathrm{E}((x_k - \bar{x}_k)(x_k - \bar{x}_k)^T), \tag{2.93}$$
$$\hat{X}_k = \mathrm{E}((x_k - \hat{x}_k)(x_k - \hat{x}_k)^T). \tag{2.94}$$

Note that when the system is time invariant, i.e. when the system matrices A and D and the noise covariance matrices V, W og R_{12} are constant matrices, then the filter will be stationary and we will have that $\bar{X}_{k+1} = \bar{X}_k = \bar{X}$ and $K_k = K$ are constant matrices. Note also that (2.90) can be expressed as the following alternative

$$\hat{X}_k = \bar{X}_k - K_k D\bar{X}_k. \tag{2.95}$$

However, Equation (2.90) is to be preferred of numerical reasons due to the fact that all terms in (2.90) are symmetric and positive semidefinite. Hence, it is of higher probability that the final computed results is symmetric and positive semidefinite by using (2.90). The final computed covariance matrix $\hat{X}$ should be symmetric and positive semidefinite, i.e. symmetric and $\hat{X} \geq 0$

Equation for computing K_k in the filter

We take the updating given by (2.87) as the starting point and write

$$x_k - \hat{x}_k = x_k - \bar{x}_k - K\varepsilon_k. \tag{2.96}$$

Post multiplication with $\varepsilon_k^T = (y_k - \bar{y}_k)^T = (D(x_k - \bar{x}_k) + w_k)^T$ gives

$$(x_k - \hat{x}_k)((x_k - \bar{x}_k)^T D^T + w_k^T) = (x_k - \bar{x}_k)((x_k - \bar{x}_k)^T D^T + w_k^T) - K\varepsilon_k\varepsilon_k^T. \tag{2.97}$$

Using the mean operator E(·) on both sides of the equal sign in (2.97) gives

$$0 = \bar{X}D^T - K\mathrm{E}(\varepsilon_k \varepsilon_k^T), \tag{2.98}$$

because

$$\mathrm{E}((x_k - \hat{x}_k)(x_k - \bar{x}_k)^T) = 0, \tag{2.99}$$
$$\mathrm{E}((x_k - \hat{x}_k)w_k^T) = 0, \tag{2.100}$$
$$\mathrm{E}((x_k - \bar{x}_k)w_k^T) = 0, \tag{2.101}$$

when we are using the optimal Kalman filter gain K.

We then get from (2.98) that the optimal Kalman filter gain matrix in the filter is given by

$$K = \bar{X}D^T(D\bar{X}D^T + W)^{-1}. \tag{2.102}$$

Let us now compare (2.102) with the expression for $\tilde{K} = AK$ for the Kalman filter gain in the predictor given by Equation (2.75). As we see, the equations are consistent and the same when $R_{12} = 0$. However, (2.102) will be valid even when the process noise and the measurements noise are correlated, but we then have to take $\bar{X}$ given by (2.82).

Equation for computing $\hat{X}$

The updating equation (2.87) can be expressed as follows

$$\hat{x}_k = \bar{x}_k + K(y_k - \bar{y}_k) = (I - KD)\bar{x}_k + KDx_k + Kw_k. \tag{2.103}$$

We can then write the estimator error $x_k - \hat{x}_k$ as follows

$$\begin{aligned} x_k - \hat{x}_k &= x_k - ((I - KD)\bar{x}_k + KDx_k + Kw_k) \\ &= (I - KD)(x_k - \bar{x}_k) + Kw_k. \end{aligned} \tag{2.104}$$

This gives

$$\hat{X}_k = (I - KD)\bar{X}_k(I - KD)^T + KWK^T. \tag{2.105}$$

Equation for updating $\bar{X}_k$

We have earlier deduced the Riccati equation for computing $\bar{X}_k$ in connection with the Kalman filter on prediction and innovations form. Se Equations (2.80)-(2.82). By substituting the expression for $\hat{X}_k$ given by (2.90) into Equation (2.91) gives Equation (2.80). This proves Equation (2.91).

2.6.4 Summary

It is important to note that for discrete time systems, we have two formulations of the Kalman filter, one Kalman filter on innovations or prediction form, and one Kalman filter on apriori-aposteriori form for filtering or optimal state estimation. The Kalman filter gain in the innovations form is denoted $\tilde{K}$ and the Kalman filter gain in the filter is denoted K.

The relationship is given by $\tilde{K} = AK$ when the process noise v_k and the measurements noise w_k are uncorrelated, i.e. when $R_{12} = 0$. When the process noise and the measurements noise are correlated then the Kalman filter gain in the innovations form (the predictor) is given by

$$\tilde{K}_k = (A\bar{X}_k D^T + R_{12})(D\bar{X}_k D^T + W)^{-1},$$

and the gain in the filter used to compute the aposteriori state estimate is given by

$$K_k = \bar{X}_k D^T (D\bar{X}_k D^T + W)^{-1}.$$

As we see, the relationship is particularly simple and given by $\tilde{K}_k = AK_k$ when the noise are uncorrelated, i.e. when $R_{12} = 0$.

Chapter 3

The Kalman filter algorithm for discrete time systems

3.1 Contunuous time state space model

A continuous time nonlinear state space model can usually be written as

$$\dot{x} = f(x, u, v) \tag{3.1}$$

$$y = g(x, u) + w \tag{3.2}$$

where x is the state vector, u is the vector of known deterministic inputs, v is a process noise vector, w is a zero mean measurements noise vector, and y is a vector of measurements (observations).

This model is both driven by known deterministic inputs (u) and usually unknown process and measurements disturbances, (v and w).

3.2 Discrete time state space model

We will in this section formulate a discrete process model which can be used to design an Extended and possibly Augmented Kalman filter.

A discrete time model, which can be a discrete version of the continuous model, can usually be written as follows.

$$x_{t+1} = f_t(x_t, u_t, v_t) + dx_t \tag{3.3}$$

$$y_t = g_t(x_t, u_t) + w_t \tag{3.4}$$

where w_t is zero mean discrete measurements noise, dx_t is a zero mean process noise vector which also can represent unmodeled effects or uncertainity. The effect of adding the noise vector dx_t to the right hand side of the process noise is that it usually gives more tuning parameters in the process noise covariance matrix, which can result in a Kalman filter gain matrix with better properties of estimating the states.

We will next write this model on a form which is more convenient for nonlinear filtering (Extended Kalman filter, Jazwinski (1970)). The problem is the case when the process model function $f_t(\cdot)$ is a non-linear function of the pocess noise vector v_t. Assume that the statistical properties of v_t is known. In general, the statistical properties of the non linear function $f_t(v_t)$ is unknown. The idea is to augment a model for v_t with the process model such that the augmented model is linear in the process noise.

Assume the case when the process noise have known mean (or trend) $\bar{v}_t$ and that the noise can be modeled as

$$v_t = \bar{v}_t + dv_t \tag{3.5}$$

where dv_t is a zero mean white noise vector. The known mean process noise vector or trend $\bar{v}_t$ can be augmented into the vector of known deterministic inputs (u_t). The resulting model is then driven by both deterministic inputs (u_t and $\bar{v}_t$) and zero mean white process and measurements noise (dv_t and w_t). $f_t(\cdot)$ can in some cases be assumed to be a linear function of the white process noise vector (dv_t).

Assume next the better case when the process noise v_t can be modeled as a random walk (or drift), i.e.

$$v_{t+1} = v_t + dv_t \tag{3.6}$$

The vector v_t can be augmented into the state vector x_t. The resulting augmented model is linear in the process noise (dv_t).

The process model to be used in the filter is assumed to be of the following form, (i.e. linear in the process noise vector)

$$x_{t+1} = f_t(x_t, u_t) + \Omega_t v_t \tag{3.7}$$
$$y_t = g_t(x_t, u_t) + w_t \tag{3.8}$$

which is linear in terms of the unknown process and measurement white noise processes v_t and w_t, respectively. The input vector u_t is a collection of all (deterministic) known variables, including possibly measured process noise variables and manipulable process input variables. The system vector x_t can be an augmented vector of system states, possibly states in a process noise model and states in a parameter model, e.g. random walk (or drift) models.

Furthermore, the following statistical properties are assumed

$$\begin{array}{l} \mathrm{E}(v_t) = 0 \text{ and } \mathrm{E}(v_t v_j^T) = V\delta_{tj} \\ \mathrm{E}(w_t) = 0 \text{ and } \mathrm{E}(w_t w_j^T) = W\delta_{tj} \end{array} \quad \text{where} \quad \delta_{tj} = \begin{cases} 1 \text{ if } j = t \\ 0 \text{ if } j \neq t \end{cases} \tag{3.9}$$

The linearized discrete time state space model is defined as

$$dx_{t+1} = \Phi_t dx_t + \Delta_t du_t + \Omega_t dv_t \tag{3.10}$$
$$dy_t = D_t dx_t + E du_t + w_t \tag{3.11}$$

where dx_t, du_t, dv_t and dy_t are deviations around some vectors of variables.

3.3 The Kalman filter algorithm

The algorithm presented is a formulation of the Extended and possibly Augmented Kalman filter. The algorithm is formulated, step for step, such that it can be directly implemented in a computer.

Algorithm 3.3.1 (Extended Kalman filter algorithm)
Step 0. *Initial values.*
Specify the apriori state vector, $\bar{x}_t$, and the apriori state covariance matrix, $\bar{X}_t$. ($\bar{x}_t$ and $\bar{X}_t$ are usually given from the previous sample of this algorithm. Note that t is discrete time.)

Step 1. *Measurements model uppdate.*

$$\bar{y}_t = g_t(\bar{x}_t, u_t) \tag{3.12}$$

Step 2. *The Kalman filter gain matrix.*
Linearized measurements model matrix

$$D_t = \left. \frac{\partial g_t(x_t, u_t)}{\partial x_t} \right|_{\bar{x}_t, u_t} \tag{3.13}$$

Kalman filter gain matrix.

$$K_t = \bar{X}_t D_t^T (D_t \bar{X}_t D_t^T + W)^{-1} \tag{3.14}$$

Step 3. *Aposteriori state estimate.*

$$\hat{x}_t = \bar{x}_t + K_t(y_t - \bar{y}_t) \tag{3.15}$$

Step 4. *Apriori state uppdate.*

$$\bar{x}_{t+1} = f_t(\hat{x}_t, u_t) \tag{3.16}$$

Define the state transition and the disturbance input matrices.

$$\Phi_t = \left. \frac{\partial f_t(x_t, u_t) + \Omega_t v_t}{\partial x_t} \right|_{\hat{x}_t, u_t} \tag{3.17}$$

$$\Omega_t = \left. \frac{\partial f_t(x_t, u_t) + \Omega_t v_t}{\partial v} \right|_{\hat{x}_t, u_t} \tag{3.18}$$

Step 5. *State covariance matrices.*
Aposteriori state covariance matrix.

$$\hat{X}_t = (I - K_t D_t)\bar{X}_t (I - K_t D_t)^T + K_t W K_t^T \tag{3.19}$$

Apriori state covariance matrix uppdate.

$$\bar{X}_{t+1} = \Phi_t \hat{X}_t \Phi_t^T + \Omega_t V \Omega_t^T \tag{3.20}$$

△

Note that the matrix equation for the aposteriori state covariance matrix, Equation (3.19), is called the stabilized implementation, because it have better numerical properties than the other frequently used equations for $\hat{X}$, e.g.

$$\hat{X}_t = \bar{X}_t - \bar{X}_t D_t^T (D_t \bar{X}_t D_t^T + W)^{-1} D_t \bar{X}_t \tag{3.21}$$

$$\hat{X}_t = (I - K_t D_t) \bar{X}_t \tag{3.22}$$

The Algorithm 3.3.1 is all that is needed for the design of an Kalman filter application. Se also the next sections for pure details about implementation. However, for extreme accuracy of the computational results the (square root) algorithm by Bierman (1974) should be implemented

3.3.1 Example: parameter estimation

Assume the linear (measurement) equation

$$y_t = E_t u_t + w_t \tag{3.23}$$

where $y_t \in \Re^m$ and $u_t \in \Re^r$ are known. The error $w_t \in \Re^m$ is assumed to be a zero mean white noise process. $E_t \in \Re^{m \times r}$ is a matrix of unknown parameters. The problem adressed in this section is to estimate the (gain) matrix E_t.

We will first write the model into a more convenient form for parameter estimation. We have

$$\begin{bmatrix} y_1 \\ y_2 \\ \vdots \\ y_m \end{bmatrix}_t = \begin{bmatrix} e_1^T \\ e_2^T \\ \vdots \\ e_m^T \end{bmatrix}_t u_t = \begin{bmatrix} u^T e_1 \\ u^T e_2 \\ \vdots \\ u^T e_m \end{bmatrix}_t = \begin{bmatrix} u_t^T & 0 & \cdots & 0 \\ 0 & u_t^T & \ddots & \vdots \\ \vdots & \vdots & \ddots & 0 \\ 0 & 0 & \cdots & u_t^T \end{bmatrix} \begin{bmatrix} e_1 \\ e_2 \\ \vdots \\ e_m \end{bmatrix}_t \tag{3.24}$$

which can be written as

$$y_t = \varphi_t^T \theta_t \tag{3.25}$$

where $y_t \in \Re^m$ is a vector of observations, $\varphi_t^T \in \Re^{m \times r \cdot m}$ is a matrix of (regression) known variables and $\theta_t \in \Re^{r \cdot m}$ is a vector of unknown parameters.

Hence, the parameter vector θ_t is formed from the rows in the matrix E and the matrix φ_t^T is a matrix with the known (input) vector u_t^T on the "diagonal". Note that in the Multiple Input Single Output (MISO) case, we simply have $\varphi_t^T = u_t^T$ and $\theta_t = E^T$.

Assume that the parameter vector θ_t is slowly varying. A reasonable model is then a so called random walk (or drift), i.e.

$$\theta_{t+1} = \theta_t + v_t \tag{3.26}$$

where v_t is a zero mean white noise process.

Problem

Use the Kalman filter Alorithm 3.3.1 to write an algorithm for parameter estimation based on the models given by Equations (3.25) and (3.26). Express the parameter estimates in terms of the apriori parameter estimate vector, i.e. $\bar{\theta}_t$.

3.4 Implementation

The Kalman filter matrix equations that are computed at each sample (if required) is given by,

1. Stabilized Kalman measurement uppdate equations.

$$K = XD^T(DXD^T + W)^{-1} \tag{3.27}$$
$$\hat{X} = (I - KD)X(I - KD)^T + KWK^T \tag{3.28}$$

2. Time uppdate apriori covariance matrix equation.

$$X = \Phi\hat{X}\Phi^T + V \tag{3.29}$$

where for simplicity $X := \bar{X}$.

We will in what follows count the number of multiplications which is required for one sample of the actual implementation and then suggest efficient implementations of the algorithm where the number of multiplications is considerably reduced.

The stabilized Kalman measurement uppdate Equation (3.28) is implemented in the following steps. The resulting matrix dimension and the number of multiplications required is identified to the right of each equations.

Algorithm 3.4.1 ("Bulk" implementation)

$$\begin{array}{lll} WORK1 = I - KD & (n \times n) & n^2m \\ WORK2 = X\ WORK1^T & (n \times n) & n^3 \\ WORK3 = WORK1\ WORK2 & (n \times n) & n^3 \\ X \quad\quad = WORK3 + KWK^T & (n \times n) & 2n^2m \\ & Total & 2n^3 + 3n^2m \end{array} \tag{3.30}$$

△

The total number of multiplications for Equation (3.28) is then given by

$$2n^3 + 3n^2m \quad (= 400 \text{ for } n = 5 \text{ and } m = 2) \tag{3.31}$$

The term KWK^T can be implemented more effectively as follows

$$\begin{array}{lll} WORK1 = KW & (n \times m) & nm \\ WORK2 = WORK1\ K^T & (n \times n) & n^2m \end{array} \tag{3.32}$$

The total number of multiplications is in this case given by

$$2n^3 + 2n^2m + nm \quad (= 360 \text{ for } n = 5 \text{ and } m = 2) \tag{3.33}$$

Multiplications can be saved if the symmetry of the matrix terms $(I - KD)X(I - KD)^T$ and KWK^T are utilized. Only the lower or upper part of the latter terms needs to be computed.

Algorithm 3.4.2 (Computations of symmetrical parts only)

$$\begin{array}{llll} WORK1 & = I - KD & (n \times n) & n^2 m \\ WORK2 & = X\ WORK1^T & (n \times n) & n^3 \\ WORK3 & = WORK1\ WORK2 & (n \times n) & n\frac{n(n+1)}{2} \\ WORK1 & = K\ W & (n \times m) & nm \\ X & = WORK3 + WORK1\ K^T & (n \times n) & m\frac{n(n+1)}{2} \\ & & \mathit{Total} & \frac{3}{2}n^3 + \frac{3}{2}n^2 m + \frac{1}{2}n^2 + \frac{3}{2}nm \end{array} \tag{3.34}$$

△

The total number of multiplications is in this case given by

$$\frac{3}{2}n^3 + \frac{3}{2}n^2 m + \frac{1}{2}n^2 + \frac{3}{2}nm \quad (= 290 \text{ for } n = 5 \text{ and } m = 2) \tag{3.35}$$

In general, the most efficient implementation of Equation (3.28) with respect to the number of multiplications is probably as follows. However, both algorithms (3.4.1) and (3.4.2) are probably better conditioned with respect to positive definiteness of the computed covariance matrix.

Algorithm 3.4.3 (Biermans implementation)

$$\begin{array}{lllll} WORK1 & = XD^T & (n \times m) & n^2 m & \\ X & = X - K\ WORK1^T & (n \times n) & n^2 m & \\ WORK2 & = KW & (n \times m) & nm & \\ WORK1 & = XD^T - WORK2 & (n \times m) & n^2 m & \\ X & = X - WORK1\ K^T & (n \times n) & n^2 m & m\frac{n(n+1)}{2} \\ & & \mathit{Total} & (4n^2 + n)m & (\frac{5}{2}n^2 + \frac{3}{2}n)m \end{array} \tag{3.36}$$

△

Note that the matrix product XD^T used initially in Algorithm 3.4.3 is available from the computation of the gain matrix K. Therefore the total number of multiplications by Algorithm 3.4.3 can be reduced by $n^2 m$ for comparison with Algorithms 3.4.1 and 3.4.2. The total number of multiplications required to form the a posteriori state covariance matrix $\hat{X}$ is illustrated in the following table.

Table 1: Comparison of number of multiplications for $m = 2$

Algorithm	Total	N = 3	N = 5	Remarks
4.1	$2n^3 + 3n^2 m$	108	400	
4.2	$\frac{3}{2}n^3 + \frac{3}{2}n^2 m + \frac{1}{2}n^2 + \frac{3}{2}nm$	81	290	
4.3	$(3n^2 + n)m$	64	160	
4.3 Symmetrized	$(\frac{5}{2}n^2 + \frac{3}{2}n)m$	54	140	

(3.37)

The a priori state covariance uppdate matrix Equation (3.29) can be directly implemented with $2n^3$ multiplications or with $n^3 + n\frac{n(n+1)}{2} = \frac{3}{2}n^3 + \frac{1}{2}n^2$ if the symmetry of the resulting product $\Phi\hat{X}\Phi^T$ is utilized.

Note that the structure of the Φ matrix should be utilized if it is sparse. For the $N = 5$ and $M = 2$ example given in this note, only 36 multiplications are needed to form $\bar{X}$ compared to 250 (or 200 if symmetry is utilized) in the general case.

Skogn implementation: $72 + 400 + 250 = 722$.

Symetrical implementation: $67 + 290 + 200 = 557$.

Symetrical and structure: $67 + 290 + 36 = 393$.

4.3 symmetrized and structure: $67 + 140 + 36 = 243$.

References

Bierman, G. J. (1974). Sequential Square Root Filtering and Smoothing of Discrete Linear Systems. *Automatica*, Vol. 10, pp. 147-158.

Jazwinski, A. H. (1970). *Stochastic Processes and filtering Theory*. Academic Press, New York and London, 1970.