IMPERIAL COLLEGE LONDON

E4.26
CS5.3

DEPARTMENT OF ELECTRICAL AND ELECTRONIC ENGINEERING
EXAMINATIONS 2009

MSc and EEE Part IV: MEng. and ACGI

Corrected Copy

ESTIMATION AND FAULT DETECTION

Monday, 27th April 2:30 pm

Time allowed: 3:00 hours

There are SIX questions on this paper.

Answer FOUR questions.

All questions carry equal marks

Any special instructions for invigilators and information for candidates are on page 1.

Examiners responsible First Marker R. B. Vinter

Second Marker D. Angeli

Information for candidates:

Some formulae relevant to the questions.

The normal $N(m, \sigma^2)$ density:

$$p(y) = \frac{1}{\sqrt{2\pi\sigma^2}} \exp\left(-\frac{(y-m)^2}{2\sigma^2}\right)$$

System equations:

$$\begin{aligned} x_k &= Fx_{k-1} + u^s + w_k \\ y_k &= Hx_k + u^o + v_k \,. \end{aligned}$$

Here, w_k and v_k are white noise sequences with covariances Q^s and Q^0 respectively.

The Kalman filter equations are

$$\begin{aligned} P_{k|k-1} &= FP_{k-1}F^T + Q^s \\ P_k &= P_{k|k-1} - P_{k|k-1}H^T(HP_{k|k-1}H^T + Q^o)^{-1}HP_{k|k-1} \,, \\ K_k &= P_{k|k-1}H^T(HP_{k|k-1}H^T + Q^o)^{-1} \,, \\ \hat{x}_k &= \hat{x}_{k|k-1} + K_k(y_k - \hat{y}_{k|k-1}) \,, \end{aligned}$$

in which $\hat{x}_{k|k-1} = F\hat{x}_{k-1} + u^s$ and $\hat{y}_{k|k-1} = H\hat{x}_{k|k-1} + u^o$

1. The relative concentrations x_1 and x_2 of two chemical reagents in a reactor are described by the differential equations

$$\begin{aligned}\dot{x}_1(t) &= -\alpha(x_1(t) - x_2(t)) + w(t) \\ \dot{x}_2(t) &= -\alpha(x_2(t) - x_1(t)) - w(t)\,,\end{aligned}$$

in which $\alpha > 0$ is a constant and $w(t)$ is scalar, 'continuous time' white noise with intensity σ^2. Show that, for all t,

$$x_1(t) + x_2(t) = c \qquad (1)$$

for some constant c. Assume that $c = 1$. (This means that the equations are consistent with an interpretation of x_1 and x_2 as 'relative concentrations'.) [4]

For same fixed sample period h, measurements y_k are taken of $x_1(kh)$, k=...,-1,0,+1,.... The measurements are modelled as

$$y_k = x_1(kh) + e_k \qquad (2)$$

where e_k is some white noise process independent of $w(t)$, with intensity σ_e^2. Develop a stochastic difference equation for y_k which includes as inputs a discrete time white noise process v_k related to $w(t)$ and e_k. The equation coefficients will depend on α, σ^2 and σ_e^2.

Can y_k be regarded as a 'biased' ARMA model, i.e. as a process satisfying the equations

$$y_k = -a_1 y_{k-1} + b_1 e'_k + b_2 e'_{k-1} + \xi\,,$$

in which e'_k is a white noise process and ξ is a constant? You should briefly explain your answer, without doing any calculations. [2]

Hint: Write $z(t) = x_1(t) - x_2(t)$. Then

Step 1: Derive a scalar stochastic differential equation for $z(t)$. [3]

Step 2: Derive a stochastic difference equation for the sampled values $z(kh)$. [8]

Step 3: Derive stochastic difference equations for $x_1(kh)$ and y_k, using (1) and (2). [3]

2. An object is located at a random position $\mathbf{x}$ lying on a straight line passing through two distinct points $\mathbf{x}_0$ and $\mathbf{x}_1$ in the plane. A sensor provides a measurement $\mathbf{y}$ of $\mathbf{x}$. $\mathbf{y}$ is modelled as a random variable satisfying

$$y = \mathbf{x} + \mathbf{n},$$

where $\mathbf{n}$ is a random variable, independent of $\mathbf{x}$, with $E[\mathbf{n}] = 0$ and $\text{cov}\{\mathbf{n}\} = \sigma_n^2 I_{2\times 2}$.

Assume that the mean of $\mathbf{x}$ is the midpoint of the line joining $\mathbf{x}_0$ and $\mathbf{x}_1$ and the variance of its displacement along the line is σ^2.

Show that, if $\mathbf{x}$ is modelled as

$$\mathbf{x} = \mathbf{x}_0 + (\frac{1}{2} + \alpha||\mathbf{x}_1 - \mathbf{x}_0||^{-1})\,(\mathbf{x}_1 - \mathbf{x}_0), \tag{3}$$

in which α is a scalar random variable with

$$E[\alpha] = 0 \quad \text{and} \quad \text{var}\{\alpha\} = \sigma^2,$$

then $\mathbf{x}$ has the specified mean and variance properties. [3]

($||\mathbf{z}|| = (z_1^2 + \ldots + z_n^2)^{\frac{1}{2}}$ is the Euclidean length of the n-vector $\mathbf{z}$.)

Determine the linear least squares estimate $\hat{\alpha}$ of α given $\mathbf{y}$. [4]

Show that the mean square estimation error is

$$E|\alpha - \hat{\alpha}|^2 = \frac{\sigma_n^2}{\sigma^2 + \sigma_n^2} \times \sigma^2 .$$

[7]

Construct an estimate $\hat{\mathbf{x}}$ of $\mathbf{x}$ given $\mathbf{y}$ from (3) and determine the mean square error $E||\mathbf{x} - \hat{\mathbf{x}}||^2$. [4]

Briefly explain why the estimator $\hat{\mathbf{x}}$ of $\mathbf{x}$ given $\mathbf{y}$ minimizes the least squares estimation error over all linear estimators that take values on the line through $\mathbf{x}_0$ and $\mathbf{x}_1$. [2]

You may use standard formulae of linear least squares estimation, and also the matrix identity, valid for any $\beta \geq 0$ and any n-vector $\mathbf{z}$,

$$\mathbf{z}^T \left[\beta^2 I_{n\times n} + \mathbf{z}\mathbf{z}^T\right]^{-1} \mathbf{z} = \left(\frac{1}{\beta^2 + ||\mathbf{z}||^2}\right) ||\mathbf{z}||^2 .$$

3a: Signal and measurement processes x_k and y_k respectively are described by

$$\begin{aligned} x_t &= Fx_{t-1} + w_t \\ y_t &= Hx_t + v_t \end{aligned}$$

for $t = 1, 2, \ldots$. Here, $\{w_t\}$ and $\{v_t\}$ are independent white noise sequences. Denote their covariances by Q^s and Q^o respectively. It is assumed that $\{w_t\}$, $\{v_t\}$ and x_0 are independent.

Under what conditions on the modelling parameters do the predicted error covariances, error covariance matrices and the Kalman gain matrices $P_{k|k-1}$, P_k and K_k converge: [2]

$$P_{k|k-1} \to S, \quad P_k \to P, \quad \text{and} \quad K_k \to K?$$

Derive equations for the limiting matrices S, P and K. [8]

3b: An object moves along the line. Measurements y_k are taken of its position at sample times kT, $k = 0, 1, 2 \ldots$ (T is the sample period.) Assume that the position x_k, velocity v_k and measurement y_k at time kh are modelled by the equations:

$$x_k = x_{k-1} + Tv_{k-1}, \quad v_k = v_{k-1} + w_k$$

and

$$y_k = x_k + e_k$$

in which w_k and e_k are independent white noise processes with intensity σ_s^2 and σ_0^2 respectively.

The widely used alpha-beta filter recursively computes estimates $\hat{x}_k$ and $\hat{v}_k$ of x_k and v_k, given $y_{1:k}$, by means of the following equations:

$$\begin{aligned} \hat{x}_k &= \hat{x}_{k-1} + T\hat{v}_{k-1} + \alpha\left[y_k - (\hat{x}_{k-1} + T\hat{v}_{k-1})\right] \\ \hat{v}_k &= \hat{v}_{k-1} + \beta\left[y_k - (\hat{x}_{k-1} + T\hat{v}_{k-1})\right] \end{aligned}$$

in which α and β are design parameters.

Show that standard conditions are satisfied under which the Kalman filter parameters converge. [4]

Show that alpha-beta filter can be interpreted as the asymptotic Kalman filter, if α and β are chosen to be:

$$\alpha = \frac{s_{11}}{s_{11} + \sigma_0^2} \quad \text{and} \quad \beta = \frac{s_{12}}{s_{11} + \sigma_0^2},$$

where $\{s_{ij}\}$ are the entries of the matrix $S = \lim_{k \to} P_{k|k-1}$. (For this part of the question, it is not necessary to derive the formulae for s_{11} and s_{12}.) [6]

4. Consider the signal and measurement processes described by the equations

$$\begin{aligned} x_t &= F x_{t-1} + w_t \\ y_t &= H x_t + v_t \end{aligned}$$

for $t = 1, 2, \ldots$, in which $\{w_t\}$ and $\{v_t\}$ are independent Gaussian white noise sequences, with covariances Q^s and Q^o respectively, independent of $x_0 \sim N(\hat{x}_0, P_0)$. Write, for times $t \geq 1$, $s \geq 0$,

$$\begin{aligned} \hat{x}_{t|s} &= E[x_t|y_{1:s}], \quad P_{t|s} = \operatorname{cov}\{x_t|y_{1:s}\} \\ \hat{y}_{t|s} &= E[y_t|y_{1:s}] \end{aligned}$$

and write, briefly, $\hat{x}_t = \hat{x}_{t|t}$, $P_t = P_{t|t}$.

Derive the following equations relating the one-step-backwards smoothed estimate $\hat{x}_{t|t+1}$ of x_t and its error covariance $P_{t|t+1}$ to the un-smoothed estimate $\hat{x}_t$ and its error covariance P_t:

$$\begin{aligned} \hat{x}_{t|t+1} &= \hat{x}_t + P_t F^T H^T \left(H(FP_tF^T + Q^s)H^T + Q^o\right)^{-1} (y_{t+1} - HF\hat{x}_t) \\ P_{t|t+1} &= P_t - P_t F^T H^T \left(H(FP_tF^T + Q^s)H^T + Q^o\right)^{-1} HFP_t. \end{aligned}$$

You should take the following steps in your derivation:

Step 1: Calculate

$$E[x_t|y_{1:t}],\ E[y_{t+1}|y_{1:t}],\ \operatorname{cov}\{x_t, y_{t+1}|y_{1:t}\} \text{ and } \operatorname{cov}\{y_{t+1}|y_{1:t}\}.$$

[8]

Step 2: Apply the standard formulae for the solution to the 'static' linear least squares estimation problem and for the error covariance. [4]

Now suppose that the processes $\{x_t\}$ and $\{y_t\}$ are scalar (write $f = F$, $h = H$, $\sigma_s^2 = Q^s$, $\sigma_m^2 = Q^o$, $p_t = P_t$, etc.) Suppose also that the smoothed estimate will only be used if it gives a sufficiently large reduction in error variance, i.e. the percentage reduction in the error covariance that results from using the smoothed estimate $\hat{x}_{t|t+1}$ instead of the un-smoothed estimate $\hat{x}_t$ is at least $\alpha \times 100\%$, i.e.

$$\frac{P_t - P_{t|t+1}}{P_t} \geq \alpha\,.$$

Show P_t must satisfy:

$$P_t \geq \frac{h^2\sigma_s^2 + \sigma_m^2}{h^2 f^2} \times \frac{\alpha}{1-\alpha}$$

[8]

5. (a): Consider the vector signal and scalar measurement processes described by

$$\begin{aligned} x_t &= Fx_{t-1} + w_t \\ y_t &= \psi(x_t) + v_t \end{aligned}$$

for $t = 1, 2, \ldots$ Here, $\{w_t\}$ and $\{v_t\}$ are independent, Gaussian, white noise sequences, with covariances Q^s and σ_m^2 respectively, independent of $x_0 \sim N(\hat{x}_0, P_0)$. F is a given matrix and $\psi(x)$ is a given function.

State the equations of the Extended Kalman Filter (EKF) for the recursive computation of an estimate $\hat{x}_t$ of x_t given $y_{1:t}$ and an approximation to the error covariance. Explain how it is related to the Kalman filter, and the approximations made in its construction. [14]

(b): The position x_t^1 of a target and the position x_t^2 of a moving sensor, in one dimension, are described by the scalar equations

$$x_{t+1}^1 = x_t^1 + w_{t+1}^1 \quad \text{and} \quad x_{t+1}^2 = w_{t+1}^2$$

where $\{w_t^1\}$ and $\{w_t^2\}$ are Gaussian white noise processes with variances σ_1^2 and σ_2^2. The sensor provides noisy measurements of the relative position of the target to the sensor:

$$y_{t+1} = \psi(x_{t+1}) + v_t$$

where v_t is a Gaussian, white noise process with variance σ_m^2 and $\psi(x)$ is the nonlinear function

$$\psi(x) = |x_1 - x_2|^3 \quad (x = (x_1, x_2)^T) .$$

It is assumed that $\{w_t^1\}$, $\{w_t^2\}$, $\{v_t\}$, x_0^1 and x_0^2 are independent.

Derive the Extended Kalman Filter equations giving an estimate $\hat{x}_t = (\hat{x}_t^1, \hat{x}_t^2)^T$ of the joint position $x_t = (x_t^1, x_t^2)^T$ of the target and sensor, and an approximation of the error covariance. [6]

6. The output from a measuring device is modelled as a non-zero mean, stationary, scalar stochastic process y_t governed by the equations

$$y_t = 0.5y_{t-1} + d + e_t \ .$$

Here, e_t is a Gaussian, white noise process with $e_t \sim N(0, \sigma^2 = 4/3)$ and d is a constant.

Determine the mean m_y and variance σ_y^2 of the process y_t. (m_y will depend on d.) [3]

[5]

Now consider two hypotheses

(H_0) : a fault has not occurred, in which case $d = 0$,
(H_1) : a fault has occurred, in which case $d = 3$.

A single measurement y_t is taken (at some time t). Construct a Neyman-Pearson decision function

$$\delta(y_t) = \begin{cases} 1 & \text{accept } (H_1) \\ 0 & \text{accept } (H_0) \end{cases}$$

which will detect the occurrence of a fault at the 0.05 significance level, i.e. the decision function is such that the probability of the occurrence of a false alarm is 0.05. [8]

Determine the power of the test, i.e. the probability that a fault will be detected, if it has occurred. [4]

You may use the data below, listing some relevant values of the function $F(z)$:

$$F(z) = \frac{1}{\sqrt{2\pi}} \int_z^{+\infty} e^{-x^2/2} dx.$$

(Note that, for a positive number $\bar{z}$, $F(-\bar{z}) = 1 - F(\bar{z})$.)

z	0	0.1	0.2	0.3	0.35	0.45	0.5	1.65
F(z)	0.5	0.47	0.42	0.38	0.36	0.33	0.31	0.05

Estimation and Fault Detection Exam, 2009

1. $\frac{d}{dt}(x_1+x_2) = -\alpha(x_1-x_2)+v_t - \alpha(x_2-x_1) - v_t = 0$

This implies $x_1(t)+x_2(t) =$ 'constant' for all t. We set

$$x_1(t)+x_2(t) = 1$$

Define $z = x_1 - x_2$. Then

$$\frac{d}{dt} z = -\alpha(x_1-x_2) + v_t + \alpha(x_2-x_1) + v_t = -2\alpha z + 2v$$

By the variation of constants formula

$$z(kh) = a\, z([k-1]h) + v_k ,$$

where

$$a = e^{-2\alpha h} \quad \text{and} \quad v_k = \int_{(k-1)h}^{kh} e^{-2\alpha(kh-s)} \times 2v(s)\, ds$$

By the properties of the stoch. integral, $\{v_k\}$ is a sequence of independent, zero-mean, Gaussian r.v.'s with var.:

$$\text{var}\{v_k\} = \int_{(k-1)h}^{kh} e^{-2\alpha(kh-s)} \times 2 \times 2 \times e^{-2\alpha(kh-s)} ds \times \sigma^2$$

$$= 4 \times \int_0^h e^{-4\alpha s'} ds' \quad (\text{in terms of the "dummy variable" } s' = kh - s)$$

$$= \frac{4}{4\alpha}\left(1 - e^{-4\alpha h}\right)\sigma^2$$

We deduce from $z = x_1 - x_2$ and $x_1 + x_2 = 1$ that

$x_1 = z + x_2 = z - x_1 + 1$, whence

$$x_1 = \tfrac{1}{2} z + \tfrac{1}{2}$$

It follows

$$y_k = x_1(kh) + e_k = \tfrac{1}{2} z(kh) + \tfrac{1}{2} + e_k \quad \text{— (1)}$$

$$y_{k-1} = x_1((k-1)h) + e_{k-1} = \tfrac{1}{2} z((k-1)h) + \tfrac{1}{2} + e_{k-1} \quad \text{— (2)}$$

Subtracting $a \times (2)$ from (1) gives

$$\underline{y_k - a y_{k-1}} = \tfrac{1}{2}(1-a) + \tfrac{1}{2}v_k + e_k - a e_{k-1} \quad \left(\text{in which } a = e^{-2\alpha h},\ \text{var}\{v_k\} = \underbrace{\tfrac{1}{\alpha}\left(1 - e^{-4\alpha h}\right)}_{\sigma_1^2}\right)$$

The noise process in this model is

$$w_k = \tfrac{1}{2} v_k + e_k - a e_{k-1}$$

This has covariance function $R_w(0) = \frac{1}{4}(1+a^2)\sigma_1^2 + \sigma^2$, $R_w(1) = a\sigma_1^2$ and $R_w(l) = 0$ for $l \geq 2$. We can conclude that w_k has the same 2^{nd} order statistics as some first order Moving Average model w'_k. It follows y_k has the same 2^{nd} order statistics as the "biased" ARMA model

$$y'_k = a\, y'_{k-1} + \tfrac{1}{2}(1-a) + w'_k$$

2. The model for x can be written

$$x \; (= x(\alpha)) = \tfrac{1}{2}(x_0 + x_1) + \alpha b \quad — (1)$$

where b is the unit-length vector $b = (x_1 - x_0) \times \frac{1}{\|x_1 - x_0\|}$. We see $E[x] = \frac{1}{2}(x_0 + x_1)$ (the correct mean). Also, displacement along the line is $d(\bar{\bar{\alpha}}) - d(\bar{\alpha}) = \| x(\bar{\bar{\alpha}}) - x(\bar{\alpha}) \| = (\bar{\bar{\alpha}} - \bar{\alpha}) \|b\| = (\bar{\bar{\alpha}} - \bar{\alpha}) \times 1$

It follows $\mathrm{var}\{\alpha\} = 1^2 \times \mathrm{var}\{d(\alpha)\} = \sigma^2$ (correct variance)

$y = \frac{1}{2}(x_0 + x_1) + \alpha b + n$, and $E\{\alpha\} = 0$, $\mathrm{var}\{\alpha\} = \sigma^2$

We see that $E\{y\} = \frac{1}{2}(x_1 + x_2)$

$\mathrm{cov}\{\alpha, y\} = \sigma^2 b^T$, $\mathrm{cov}\{\alpha\} = \sigma^2$, $\mathrm{cov}\{y\} = \sigma^2 b b^T + \sigma_n^2 I$

The LLSE $\hat{\alpha}$ of α given y is therefore

$$\underbrace{\hat{\alpha} = \sigma^2 b^T (\sigma^2 b b^T + \sigma_n^2 I)^{-1}}(y - \underbrace{\tfrac{1}{2}(x_1 + x_2)})$$

The error variance is

$$E|\alpha - \hat{\alpha}|^2 = \sigma^2 - \sigma^2 (\sigma b)^T [\sigma_n^2 I + (\sigma b)(\sigma b)^T]^{-1} (\sigma b)$$

$$= \sigma^2 - \sigma^2 \times \frac{\|\sigma b\|^2}{\sigma_n^2 + \|\sigma b\|^2} \quad \text{(from the matrix identity)}$$

$$= \sigma^2 - \sigma^2 \times \left(\frac{\sigma^2}{\sigma_n^2 + \sigma^2} \right) \quad (\text{since } \|b\| = 1)$$

$$= \underline{(\sigma_n^2 + \sigma^2)^{-1} \times \sigma_n^2}$$

From (1) the corresponding estimate $\hat{x}$ of x given y is

$$\hat{x} = \tfrac{1}{2}(x_0 + x_1) + \sigma^2 b^T (\sigma^2 b b^T + \sigma_n^2 I)^{-1} b \, (y - \tfrac{1}{2}(x_1 + x_2))$$

$$= \tfrac{1}{2}(x_0 + x_1) + \underbrace{\frac{\sigma^2}{\sigma_n^2 + \sigma^2} (y - \tfrac{1}{2}(x_1 + x_2))}$$

$$E\|x - \hat{x}\|^2 = E\|(\alpha - \hat{\alpha}) b\|^2 = E|\alpha - \hat{\alpha}|^2 = \underbrace{\frac{\sigma_n^2}{\sigma_n^2 + \sigma^2}}$$

An arbitrary linear function of y that lies on the line must have the form

$$d(y) = \tfrac{1}{2}(x_0 + x_1) + \alpha(y)\, b$$

where $\alpha(y)$ is a linear function of α and

$E\|x - d(y)\|^2 = E\|(\alpha - \alpha(y))\, b\|^2 = E|\alpha - \alpha(y)|^2$. So the mean square errors of the estimates $d(y)$ of x and $\alpha(y)$ of y are the same. Since $\hat{\alpha}$ is the LLSE of α, $\hat{x}$ is the (constrained) LLSE of x.

3. From the Kalman Filter equations

$$P_k = P_{k|k-1} - P_{k|k-1} H^T [H P_{k|k-1} H^T + Q^s]^{-1} H P_{k|k-1},$$

$$K_k = P_{k|k-1} H^T [\text{—}]^{-1}, \quad P_{k+1|k} = F P_k F^T + Q^s$$

$P_k \to P$, $P_{k|k-1} \to S$ and $K_k \to K$ as $k \to \infty$ if (F, H) is <u>observable</u>, i.e. $\begin{bmatrix} H \\ HF \\ \vdots \\ HF^{n-1} \end{bmatrix}$ <u>has full column rank</u> (n = state dimension)

Eqns for the asymptotic values of $P_{k|k-1}$, P_k and K_k are obtained by replacing P_k by P, etc., in the above eqns:

$$\left.\begin{aligned} S &= FPF^T + Q^s \\ &= FSF^T + FSH^T [HSH^T + Q^o]^{-1} HSF + Q^s \\ P &= S - SH^T [HSH^T + Q^o]^{-1} HS \\ K &= SH [HSH^T + Q^o]^{-1}. \\ \text{and} \quad \hat{x}_k &= F\hat{x}_{k-1} + K [y_k - HF\hat{x}_{k-1}] \end{aligned}\right\}$$

Write $x_k^1 = x_k$, $x_k^2 = v_k$. Then

$$\begin{bmatrix} x_k^1 \\ x_k^2 \end{bmatrix} = \underbrace{\begin{bmatrix} 1 & T \\ 0 & 1 \end{bmatrix}}_{F} \begin{bmatrix} x_{k-1}^1 \\ \dot{x}_{k-1} \end{bmatrix} + \begin{bmatrix} 0 \\ 1 \end{bmatrix} w_k \quad (\operatorname{var}\{w_k\} = \sigma_s^2)$$

$$y_k = \underbrace{[1 \ \ 0]}_{h^T} \begin{bmatrix} x_k^1 \\ x_k^2 \end{bmatrix} + v_k \quad (\operatorname{var}\{v_k\} = \sigma_o^2)$$

The observability condition is

$\operatorname{rank} \begin{bmatrix} h^T \\ h^T F \end{bmatrix} = 2$. But $\begin{bmatrix} h^T \\ h^T F \end{bmatrix} = \begin{bmatrix} 1 & 0 \\ 1 & T \end{bmatrix}$, a non-singular matrix. So the observability condition is satisfied, and the asymptotic Kalman filter exists.

The asymptotic Kalman gain is

$$K = S h [h S h^T + \sigma_o^2]^{-1}$$

$$= \begin{bmatrix} s_{11} & s_{12} \\ s_{12} & s_{22} \end{bmatrix} [1 \ 0] \Big/ \Big(\overbrace{[1 \ 0] \begin{bmatrix} s_{11} & s_{12} \\ s_{12} & s_{22} \end{bmatrix} \begin{bmatrix} 1 \\ 0 \end{bmatrix} + \sigma_o^2}^{s_{11} + \sigma_o^2} \Big)$$

$$= \left[\frac{s_{11}}{s_{11} + \sigma_o^2}, \ \frac{s_{12}}{s_{11} + \sigma_o^2} \right]^T$$

It follows the parameters in the alpha-beta filter are

$$\underline{\alpha = (s_{11} + \sigma_o^2)^{-1} s_{11}} \quad \text{and} \quad \underline{\beta = (s_{11} + \sigma_o^2)^{-1} s_{12}}$$

4. From $x_t = Fx_{t-1} + w_t$ and $y_t = Hx_t + v_t$ we deduce

$E[x_t | y_{1:t}] = \hat{x}_t$, by definition

$E[y_{t+1} | y_{1:t}] = H E[x_{t+1} | y_{1:t}] + E[v_{t+1} | y_{1:t}] = H\hat{x}_{t+1|t} + 0$

$= HF\hat{x}_t$

$\mathrm{cov}\{x_t, y_{t+1} | y_{1:t}\} = E[(x_t - \hat{x}_t)(Hx_{t+1} + v_{t+1} - HF\hat{x}_t)^T | y_{1:t}]$

$= E[(x_t - \hat{x}_t)(H(F[x_t - \hat{x}_t] + w_{t+1}) + v_{t+1})^T | y_{1:t}]$

$= P_t F^T H^T + 0 + 0$

$\mathrm{cov}\{y_{t+1} | y_{1:t}\}$

$= E[H(F[x_t - \hat{x}_t] + Hw_{t+1} + v_{t+1})(\ldots)^T | y_{1:t}]$

$= H(FP_t F^T + Q^s)H^T + Q^o.$

The standard Linear Least Squares formulae

("$\hat{x} = \mathrm{cov}\{x,y\}\mathrm{cov}\{y\}^{-1}(y - m_y) + m_x$

and $\mathrm{cov}[x - \hat{x}] = \mathrm{cov}\{x\} - \mathrm{cov}\{x,y\}\mathrm{cov}\{y\}^{-1}\mathrm{cov}\{y,x\}$")

in which we interpret $E[\ldots] = E[\ldots | y_{1:t}]$, $y = y_{t+1}$ give:

$$\hat{x}_{t|t+1} = \hat{x}_t + P_t F^T H^T \left(H(FP_t F^T + Q^s)H^T + Q^o\right)^{-1}(y_{t+1} - HF\hat{x}_t)$$

and

$$P_{t|t+1} = P_t - P_t F^T H^T \left(H(FP_t F^T + Q^s)H^T + Q^m\right)^{-1} HFP_t$$

In the scalar case, the relative reduction in error variance is

$$\frac{h^2 f^2 P_t^2}{h^2 f^2 P_t + h^2 \sigma_s^2 + \sigma_m^2} \times \frac{1}{P_t}$$

This must be at least α, requiring

$$h^2 f^2 P_t \geq (h^2 f^2 P_t + h^2 \sigma_s^2 + \sigma_m^2) \times \alpha$$

or

$$P_t \geq \frac{(h^2 \sigma_s^2 + \sigma_m^2)}{h^2 f^2} \times \frac{\alpha}{1-\alpha}$$

5. Signal + measurement eqns: $x_{t+1} = F x_t + w_{t+1}$, $y_{t+1} = \psi(x_{t+1}) + v_{t+1}$.
At time t, the predicted state $\hat{x}_{t+1|t}$ $(= E[x_{t+1} | y_{1:t}]) = F\hat{x}_t$.
The EKF generates an approximation $\hat{x}_{t+1}$ to the estimated state $E[x_{t+1} | y_{1:t+1}]$ and the error covariance P_{t+1} by linearizing the $\psi(\cdot)$ function about the predicted state, i.e. we replace the meas. eqn. by

$$y_{t+1} = \nabla_x \psi\,(x_{t+1} - \hat{x}_{t+1|t}) + \psi(\hat{x}_{t+1|t}) + v_{t+1}$$

This equation can be written

$$y_{t+1} = h^T x_{t+1} + u_{t+1} + v_{t+1}, \quad \begin{cases} h^T = \nabla_x \psi(F\hat{x}_t) \\ u_{t+1} = -h^T \hat{x}_{t+1|t} + \psi(\hat{x}_{t+1|t}) \end{cases}$$

The predicted output is then

$$y_{t+1|t} = h^T \hat{x}_{t+1|t} + u_{t+1} = 0 + \psi(\hat{x}_{t+1|t}) = \psi(F\hat{x}_t)$$

The update formulae for P_t and K_t (the Kalman gain) are now given by the usual formulae

$$\hat{x}_{t+1} = F\hat{x}_t + K_t\,(y_{t+1} - \psi(F\hat{x}_t)), \quad P_{t+1} = P_{t+1|t} - P_{t+1|t}\, h\, (h^T P_{t+1|t} h + \sigma_w^2)^{-1} h^T P_{t+1|t}$$

$$P_{t+1|t} = F P_t F^T + Q^s, \quad K_t = P_{t+1|t}\, h\, (h^T P_{t+1|t} h + \sigma_w^2)^{-1},$$

(where, as before, $h^T = \nabla_x \psi(F\hat{x}_t)$). (1)

(b) In the 'moving platform' problem, the signal and measurement equations are:

$$x_{t+1} = F x_t + w_{t+1}, \quad y_{t+1} = \psi(x_{t+1}) + v_{t+1}$$

with $F = \begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix}$, $w_t \sim N(0, \begin{bmatrix} \sigma_1^2 & 0 \\ 0 & \sigma_2^2 \end{bmatrix})$, $v_t \sim N(0, \sigma_w^2)$ and

$$\psi(x^1, x^2) = |x^1 - x^2|^3$$

Here $\nabla\psi(x^1, x^2)(\hat{x}_{t+1|t}) = (\partial/\partial x_1\, \psi, \partial/\partial x_2\, \psi) = 3\,|\hat{x}^1_{t+1|t} - \hat{x}^2_{t+1|t}|^2\, [1, -1]$

But $\hat{x}_{t+1|t} = F\hat{x}_t \Rightarrow \hat{x}^1_{t+1|t} = \hat{x}^1_t$ and $\hat{x}^2_{t+1|t} = 0$

It follows that the EKF takes the form (*), with

$\underline{F = \begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix}}$, $\underline{Q^s = \begin{bmatrix} \sigma_1^2 & 0 \\ 0 & \sigma_2^2 \end{bmatrix}}$, $\underline{\sigma_w^2 \text{ as specified}}$

and

$$\underline{h^T = 3\,|\hat{x}^1_t|^2\, [1, -1]}$$

6. Take expectations across the equation $y_t = -\frac{1}{2} y_{t-1} + d + e_t$:

$$E\{y_t\} = -\frac{1}{2} E[y_{t-1}] + d + 0 = -\frac{1}{2} E[y_t] + d$$

(by stationarity). So $m_y = E[y_t] = \frac{2}{3} \times d$

$$\underline{m_y = \frac{2}{3} d}$$

Also, writing $y_t' = y_t - m_y$, y_t' has zero mean and $\text{cov}\{y_t\} = E[(y_t')^2]$. So, since $y_t' = -\frac{1}{2} y_{t-1}' + e_t$

$$E[(y_t')^2] = E[(-\tfrac{1}{2} y_{t-1}' + e_t)^2]$$

$$= \frac{1}{4} E[y_{t-1}'] + 0 + E[e_t^2] = \frac{1}{4} E[(y_t')^2] + \frac{3}{4}.$$

Hence $3/4\ \sigma_y^2 = 3/4$, whence $\underline{\sigma_y^2 = 1}$

For A an event write $P_0(A)$ and $p_0(y)$ for $\underbrace{\text{prob}(A) \text{ and }}_{\text{prob. density of } y_t}$ under (H_0)

$P_1(A)$ and $p_1(y)$ " " " under (H_1)

Since $m_y = 0$ (under H_0) and $m_y = \frac{2}{3} \times d = \frac{2}{3} \times 3 = 2$ (under H_1)

$p_0(y_t) = \frac{1}{2\pi} e^{-\frac{1}{2} y_t^2}$ and $p_1(y_t) = \frac{1}{2\pi} e^{-\frac{1}{2}(y_t - 2)^2}$

The log likelihood ratio is

$$LLR(y_t) = \log_e \left(\frac{p_1(y_t)}{p_0(y_t)} \right) = -\frac{1}{2} [\cancel{y^2} - 4y_t + 4 - \cancel{y^2}]$$

$$= 2y_t - 2$$

To achieve the required test significance, choose η such that $P_0(\text{"}LLR(y_t) \geq \eta\text{"}) = \alpha = 0.05$. Then

$$\text{"}LLR(y_t) \geq \eta\text{"} \equiv 2y_t - 2 \geq \eta \equiv \text{"}y_t \geq \frac{\eta}{2} + 1\text{"}$$

But $y_t \sim N(0, 1)$ under (H_0). So

$$P_0[y_t \geq \frac{\eta}{2} + 1] = F(\frac{\eta}{2} + 1) = 0.05$$

where $F(c) = \int_c^\infty N(0, 1)(y')\, dy'$.

Also, since $y_t \sim N(2, 1)$ under (H_1),

$$P_1\{y_t \geq \frac{\eta_1}{2} + 1\} = P_1[y_t - 2 \geq \frac{\eta_1}{2} + 1 - 2] = F[\frac{\eta}{2} + 1 - 2]$$

From data, $F(\frac{\eta}{2} + 1) = 0.05 \Rightarrow \frac{\eta}{2} + 1 = 1.65 \Rightarrow \underline{\eta = 1.3}$

Then

"power of test" $= F(\frac{\eta}{2} + 1 - 2) = F(-0.35) = 1 - F(0.35)$

$= 1 - 0.36 = \underline{0.64}$