Final

E4.26
CS5.3

IMPERIAL COLLEGE LONDON

DEPARTMENT OF ELECTRICAL AND ELECTRONIC ENGINEERING
EXAMINATIONS 2010

MSc and EEE PART IV: MEng and ACGI

ESTIMATION AND FAULT DETECTION

Friday, 30 April 10:00 am

Time allowed: 3:00 hours

There are FIVE 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(s) :	R.B. Vinter
	Second Marker(s) :	D. Angeli

Information for candidates:

Some formulae relevant to the questions.

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

$$N(m,\sigma^2)(x) = \frac{1}{\sqrt{2\pi\sigma^2}} \exp\left(-\frac{(x-m)^2}{2\sigma^2}\right)$$

System equations:

$$\begin{aligned} x_t &= Fx_{t-1} + u^s + w_t \\ y_t &= Hx_t + u^o + v_t \,. \end{aligned}$$

Here, w_t and v_t are white noise sequences with covariances Q^s and Q^0 respectively.

The Kalman filter equations are

$$\begin{aligned} P_{t|t-1} &= FP_{t-1}F^T + Q^s \\ P_t &= P_{t|t-1} - P_{t|t-1}H^T(HP_{t|t-1}H^T + Q^o)^{-1}HP_{t|t-1}\,, \\ K_t &= P_{t|t-1}H^T(HP_{t|t-1}H^T + Q^o)^{-1}\,, \\ \hat{x}_t &= \hat{x}_{t|t-1} + K_t(y_t - \hat{y}_{t|t-1})\,, \end{aligned}$$

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

1. a) The position $y(t)$ of a floating object on a line at time t, moving under the action of fluid drag and a random wind force, is modelled as

$$\ddot{y}(t) = -\dot{y}(t) + w(t),$$

in which $\{w(t) | t \geq 0\}$ is unit intensity white noise, i.e. a stationary Gaussian process with covariance function $R_y(\tau) = \delta(\tau)$.

Derive the system matrix A and the input vector $\mathbf{b}$ of the stochastic state space equation

$$\dot{\mathbf{x}} = A\mathbf{x} + \mathbf{b}w(t),$$

for the state $\mathbf{x}(t) = (y(t), \dot{y}(t))^T$. [2]

For initial conditions $y(0) = \dot{y}(0) = 0$, calculate $cov\{\mathbf{x}(t)\}$ for arbitrary t. [8]

Show that, as $t \rightarrow \infty$,

$$var\{x_2(t)\} \rightarrow c \text{ and } var\{x_1(t)\} \rightarrow \infty .$$

for some (finite) constant c. [2]

Hint: to evaluate the exponential matrix e^{At}, solve $d/dt\,\dot{y}(t) = -\dot{y}(t)$ for $\dot{y}(t)$ and then $d/dt\,y(t) = \dot{y}(t)$ for $y(t)$, for arbitrary initial values of $\dot{y}(0)$ and $y(0)$.

b) Consider the Autoregressive Moving Average (ARMA) process

$$y_t + ay_{t-1} = e_t + e_{t-1},$$

in which a is a unknown number, $-1 < a < +1$, and $\{e_t\}$ is a white noise sequence such that $E[e_t^2] = 1$. Denote by $R_y(k)$ the covariance function. It is observed that

$$R_y(1) / R_y(0) = 0.25 .$$

What is the value of a? [8]

2. The position of an object on the line is described by the scalar random variable x. Two sensors provide measurements y_1 and y_2 of the position, which are modelled by the equations

$$y_1 = x + e_1 \quad \text{and} \quad y_2 = x + e_2\,,$$

where e_1 and e_2 are noise variables. It is assumed that x, e_1 and e_2 are independent random variables such that

$$x \sim N(0,\sigma^2),\ e_1 \sim N(0,\sigma_1^2) \ \text{ and } \ e_2 \sim N(0,\sigma_2^2)\,.$$

a) Determine formulae (expressed in terms of σ^2, σ_1^2 and σ_2^2) for the coefficients $a_1(\sigma^2,\sigma_1^2,\sigma_2^2)$ and $a_2(\sigma^2,\sigma_1^2,\sigma_2^2)$ in the least squares estimate of x given the two measurements:

$$\hat{x} \;=\; a_1(\sigma^2,\sigma_1^2,\sigma_2^2)\,y_1 \;+\; a_2(\sigma^2,\sigma_1^2,\sigma_2^2)\,y_2\,.$$

(You may quote the formulae for the solution to the standard linear, Gaussian least squares estimation problem.) [12]

b) For fixed σ^2 and σ_1^2 calculate the limits of the coefficients

$$a_i(\sigma^2,\sigma_1^2,\infty) \;=\; \lim_{\sigma_2^2\to\infty} a_i(\sigma^2,\sigma_1^2,\sigma_2^2) \quad i=1,2\,,$$

as $\sigma_2^2 \to \infty$. Comment on the nature of the limiting estimate

$$\hat{x}' \;=\; a_1(\sigma^2,\sigma_1^2,\infty)y_1 + a_2(\sigma^2,\sigma_1^2,\infty)y_2\,.$$

[4]

c) Now suppose that each sensor is a 'smart' sensor that processes the measurement it takes, and transmits to the base station, not its raw measurement, but its estimate of x and the scaled error covariance. Thus the base station receives $\hat{x}_i$ and γ_i, $i=1,2$, where

$$\hat{x}_i \;=\; E[x\,|\,y_i] \text{ and } \gamma_i \;=\; cov\{x\,|\,y_i\}\,/\sigma^2\,.$$

The base station fuses the local estimates $\hat{x}_1$ and $\hat{x}_2$ according to the formula

$$\hat{\hat{x}} \;=\; b_1(\gamma_1,\gamma_2)\hat{x}_1 \;+\; b_2(\gamma_1,\gamma_2)\hat{x}_2\,.$$

Determine the coefficients $b_1(\gamma_1,\gamma_2)$, $b_2(\gamma_1,\gamma_2)$ (as functions of the scaled variances γ_1 and γ_2) such that the 'fused' estimate $\hat{\hat{x}}$ coincides with the least squares estimate $\hat{x}$. [4]

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

$$\begin{aligned} x_t &= Fx_{t-1} + D_1 e_t \\ y_t &= Hx_t + D_2 e_t \end{aligned}$$

for $t = 1, 2, \ldots$, in which F, H, D_1 and D_2 are given $n \times n$, $m \times n$, $n \times r$ and $m \times r$ matrices respectively. $\{e_t\}$ is a sequence of r-vector random variables. Assume that $x_0, e_1, e_2, \ldots$ are independent and

$$x_0 \sim N(\hat{x_0}, P_0) \quad \text{and} \quad e_t \sim N(0, , I_{r \times r}) \quad \text{for each } t\,.$$

(This model permits the system and measurement noise to be correlated.) Write

$$\begin{aligned} &\hat{x}_t = E[x_t | y_{1:t}],\ P_t = cov\{x_t | y_{1:t}\}, \\ &x_{t|t-1} = E[x_t | y_{1:t-1}],\ y_{t|t-1} = E[y_t | y_{1:t-1}],\ P_{t|t-1} = cov\{x_t | y_{1:t-1}\}. \end{aligned}$$

a) Quoting the solution to the general linear least squares estimation problem, show that $\{\hat{x}_t, P_t\}$ can be recursively computed from the formulae:

$$\begin{aligned} P_{t|t-1} &= FP_{k-1}F^T + D_1 D_1^T \\ K_t &= \left[FP_{t-1}F^T H^T + D_1(HD_1 + D_2)^T\right]\left[HFP_{t-1}F^T H^T + (HD_1 + D_2)(HD_1 + D_2)^T\right]^{-1} \\ P_t &= P_{t|t-1} - K_t\left[HFP_{t-1}F^T + (HD_1 + D_2)D_1^T\right], \\ \hat{x}_t &= F\hat{x}_{t-1} + K_t(y_k - HF\hat{x}_{t-1}). \end{aligned}$$

[14]

b) At time t the position of an object x_t and that of a moving sensor platform z_t (along the line) at time t are governed by the equations

$$\begin{aligned} x_t &= ax_{t-1} + e_t^1 \\ z_t &= e_t^1 \end{aligned}$$

in which a is a constant and $\{e_t^1\}$ is a white noise sequence . Noisy measurements y_t are taken of x_t relative to the sensor platform position z_t:

$$y_t = x_t - z_t + e_t^2\,,$$

where $\{e_t^2\}$ is a white noise sequence, independent of $\{e_t^1\}$. Write $\mathbf{e}_t = [e_t^1, e_t^2]^T$. We assume that

$$\mathbf{e}_t \sim N(\mathbf{0}, I_{2\times 2})\,.$$

Find 2-vectors $\mathbf{d}_1$ and $\mathbf{d}_2$ such that $\{x_t\}$ and $\{y_t\}$ are governed by the equations

$$\begin{aligned} x_t &= ax_{t-1} + \mathbf{d}_1^T \mathbf{e}_t \\ y_t &= x_t + \mathbf{d}_2^T \mathbf{e}_t\,. \end{aligned}$$

Using the results of part (1), or otherwise, construct recursive equations for $\hat{x}_t = E[x_t | y_{1:t}]$ and $P_t = \text{var}[x_t | y_{1:t}]$. [6]

4. Consider the two step scalar signal and measurement processes governed by the equations

$$\begin{aligned} x_1 &= f(x_0) \qquad (4.1)\\ y_1 &= x_1 + e_1 \end{aligned}$$

in which $f(.)$ is a given (possibly nonlinear) function. It is assumed that the initial state x_0 and additive measurement noise term e_1 are independent random variables such that

$$x_0 \sim N(\hat{x}_0, p_0) \quad \text{and} \quad e_1 \sim N(0, r)$$

for some constants $\hat{x}_0$, $p_0 > 0$ and $r > 0$.

a) Describe the extended Kalman filter (EKF) for estimating x_1 and the error variance, given y_1, and explain the principles behind its construction. [8]

b) The statistical linear filter (SLF) is an alternative suboptimal filter to the EKF, obtained by applying the Kalman filter when the state equation (4.1) is approximated by the linear equation:

$$x_1 = u_0 + F(x_0 - \hat{x}_0) \qquad (4.2)$$

in which

$$\begin{aligned} u_0 &= E[f(x_0)]\\ F &= (E[f(x_0)x_0] - E[f(x_0)]\hat{x}_0)\, p_0^{-1}\\ (p_0 &= E[(x_0 - \hat{x}_0)^2]) \end{aligned}$$

Show that the right side of (4.2) is the linear least squares estimate of x_1 given x_0. [6]

c) Write out the equations for the EKF and the SLF filters when

$$f(x) = x^3 \text{ and } \hat{x}_0 = 0\,.$$

what filter would you expect to give the better estimate and why? [6]

Hint: You may use the fact that if x is a random variable such that $x \sim N(0, \sigma^2)$ then the fourth moment of x is $E[x^4] = 3\sigma^4$.

5. a) Consider the system and measurement processes

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

in which w_t and v_t are white noise sequences with covariances Q^s and Q^0 respectively.

Derive from the standard Kalman filter equations the following alternative update formulae for the error covariance Pt:

$$\begin{aligned} P_{t|t-1} &= FP_{t-1}F^T + Q^s \\ K_t &= P_{t|t-1}H^T(HP_{t|t-1}H^T + Q^o)^{-1}\,, \\ P_t &= (I - K_tH)P_{t|t-1}(I - K_tH)^T + K_tQ^oK_t^T\,. \end{aligned}$$

In what way are these formulae more useful than the standard update formulae?

[8]

b) A tracking system provides estimates $\hat{x}_A$ and $\hat{x}_B$ of the measured positions of two targets on the line, labelled A and B, and also the error variances $p_A = 1$ and $p_B = 1$.

Measurements x_1 and x_2 are received. Consider the two assignments

(H_0): x_1 (x_2) is the measured position of target A (target B),

(H_1): x_1 (x_2) is the measured position of target B (target A).

Writing $p_0(x_1,x_2) = p(x_1,x_2|(H_0))$ and $p_1(x_1,x_2) = p(x_1,x_2|(H_1))$, you may assume that

$$p_0(x_1,x_2) = N([x_A,x_B]^T, I_{2\times 2}),\; p_1(x_1,x_2) = N([x_B,x_A]^T, I_{2\times 2}).$$

Assume also that $x_A - x_B > 0$. Show that the inverse of the log-likelihood ratio $LLR(x_1,x_2) := \log_e\{p_1(x_1,x_2)\,/\,p_0(x_1,x_2)\}$ is

$$LLR^{-1}(x_1,x_2) = (\hat{x}_A - \hat{x}_B)(x_1 - x_2).$$

Treating (H_0) as the null hypothesis, show that the Neyman-Pearson Test for the track assignment (H_0), at the α significance level is [4]

$$x_1 - x_2 > c$$

where c is a constant such that

$$1 - erf\left(\frac{c - (\hat{x}_A - \hat{x}_B)}{\sqrt{2}}\right) = \alpha\,,$$

and the power of the test is [6]

$$erf\left(\frac{c + (\hat{x}_A - \hat{x}_B)}{\sqrt{2}}\right) = \alpha\,.$$

Here $erf(.)$ is the cumulative distribution function of the unit normal density [2]

$$erf(x) = \int_{-\infty}^{x} N(0,1)(y)dy\,.$$

E4.26

SOLUTIONS C5.7

1/5

Estimation and Fault Detection. 2010

1. (i) Let $x_1 = y$ and and $x_2 = \dot{y}$. Then

$$\begin{bmatrix} \dot{x}_1(t) \\ \dot{x}_2(t) \end{bmatrix} = \underbrace{\begin{bmatrix} 0 & 1 \\ 0 & -x_2 \end{bmatrix}}_{A} + \underbrace{\begin{bmatrix} 0 \\ 1 \end{bmatrix}}_{b} w(t)$$ [2]

For $w \equiv 0$, $x_2(t) = e^{-t} x_2(0)$. Then $x_1(t) = x_1(0) + \int_0^t e^{-s} ds\, x_2(0)$
$= x_1(0) + (1 - e^{-t}) x_2(0)$

So $e^{At} = \begin{bmatrix} 1 & 1 - e^{-t} \\ 0 & e^{-t} \end{bmatrix}$

If $\underline{x}(0) = 0$, $x(t) = 0 + \int_0^t e^{A(t-s)} b\, w(s)$ and

$$\operatorname{cov}\{\underline{x}(t)\} = \int_0^t \begin{bmatrix} 1 & 1 - e^{-(t-s)} \\ 0 & e^{-(t-s)} \end{bmatrix} \begin{bmatrix} 0 \\ 1 \end{bmatrix} [0 \ 1] \begin{bmatrix} 1 & 0 \\ 1 - e^{-(t-s)} & e^{-(t-s)} \end{bmatrix} ds$$

$$= \int_0^t \begin{bmatrix} 1 - e^{-s'} \\ e^{-s'} \end{bmatrix} [1 - e^{-s'} ; e^{-s'}]\, ds' = \int_0^t \begin{bmatrix} 1 - 2e^{-s} + e^{-2s} & e^{-s} - e^{-2s} \\ e^{-s} - e^{-2s} & e^{-2s} \end{bmatrix} ds$$

So $$R(t) = \begin{bmatrix} t - 2(1 - e^{-t}) + (1 - e^{-2t}) & (1 - e^{-t}) - (1 - e^{-2t}) \\ -e^{-2t} - e^{-t} & (1 - e^{-2t}) \end{bmatrix}$$ [8]

$$= \begin{bmatrix} t + 2e^{-t} - e^{-2t} - 1 & -(e^{-t} - e^{-2t}) \\ -(e^{-t} - e^{-2t}) & 1 - e^{-2t} \end{bmatrix}.$$

We see, as $t \to \infty$, $\operatorname{var}\{x_1(t)\} \to \infty$, and $\operatorname{var}\{x_2(t)\} \to$

[2]

(ii) $y_t + a y_{t-1} = e_t + e_{t-1}$

Multiply across by y_{t-1} and take $E\{..\} \Rightarrow R_y(1) + a R_y(0) = 0 + R_{ye}(0)$

Mult. by e_t, take $E\{..\} \Rightarrow R_{ye}(0) + 0 = 1 + 0$.

So $R_y(1) + a R_y(0) = 1$ — (A)

Also, $E(y_t + a y_{t-1})^2 = E(e_t + e_{t-1})^2$

$\Rightarrow R_y(0) + 2a R_y(1) + a^2 R_y(0) = 1 + 0 + 1 = 2$

$\Rightarrow (1 + a^2) R_y(0) + 2a R_y(1) = 2$ — (B)

Combining (A) and (B) gives $(1 + a^2) R_y(0) + 2a[1 - a R_y(0)] = 2$

$\Rightarrow (1 - a^2) R_y(0) = 2(1 - a) \Rightarrow R_y(0) = 2/(1 + a)$

So $R_y(1) = (1 + a - 2a)/1 + a = (1 - a)/(1 + a)$

So $R_y(1)/R_y(0) = \frac{1}{4} \Rightarrow \frac{1 - a}{2} = \frac{1}{4} \Rightarrow a = \frac{1}{2}$ [8]

2 (i) $y = [y_1\ y_2]^T = x\,[1\ 1]^T + [e_1, e_2]^T$. Since variables have zero mean

$$\hat{x} = E[x y^T]\, E[y y^T]^{-1} y \quad \text{where}$$

$E[xy^T] = \sigma^2 [1 \quad 1]$ and $E[yy^T] = \sigma^2 \begin{bmatrix} 1+\rho_1 & 1 \\ 1 & 1+\rho_2 \end{bmatrix}$

where $\rho_1 = \sigma_1^2/\sigma^2$, $\rho_2 = \sigma_2^2/\sigma^2$

So $$\hat{x} = \frac{1}{\rho_1+\rho_2+\rho_1\rho_2} [1 \quad 1] \begin{bmatrix} 1+\rho_2 & -1 \\ -1 & 1+\rho_2 \end{bmatrix} \begin{bmatrix} y_1 \\ y_2 \end{bmatrix}$$

$$= a_1 y_1 + a_2 y_2$$

where $a_1 = \underbrace{\frac{\rho_2}{\rho_1+\rho_2+\rho_1\rho_2}}$ and $a_2 = \underbrace{\frac{\rho_1}{\rho_1+\rho_2+\rho_1\rho_2}}$ $\left(\rho_1 = \sigma_1^2/\sigma^2,\ \rho_2 = \sigma_2^2/\sigma^2\right)$ [12]

(ii) $$b_1 = \lim_{\sigma_2^2\to\infty} a_1(\sigma^2, \sigma_1^2, \sigma_2^2) = \lim_{\sigma_2^2\to\infty} \frac{\sigma_2^2/\sigma^2}{\sigma_1^2/\sigma^2 + \sigma_2^2/\sigma^2\left(1+\sigma_1^2/\sigma^2\right)} = \underbrace{\frac{\sigma^2}{\sigma^2+\sigma_1^2}}$$

$$b_2 = \lim_{\sigma_2^2\to\infty} a_2(\sigma^2, \sigma_1^2, \sigma_2^2) = \lim_{\sigma_2^2\to\infty} \frac{\sigma_1^2/\sigma^2}{\sigma_1^2/\sigma^2 + \sigma_2^2/\sigma^2 + \frac{\sigma_1^2\sigma_2^2}{\sigma^4}} = \underbrace{0}$$

The limiting filter is

$$\hat{x}' = \frac{\sigma^2}{\sigma^2+\sigma_1^2}\, y_1 + 0$$

This is the least squares estimate of x given y_1 alone. This is appropriate since when $\sigma_2^2 \to \infty$, the second measurement provides no information about x and should be ignored. [4]

(iii) We know that $\hat{x}_i = (1+\rho_i)^{-1} y_i$ and $\gamma_i = \dfrac{\rho_i}{1+\rho_i}$

It follows

$$y_i = (1+\rho_i)\,\hat{x}_i \quad \text{and} \quad \rho_i = \frac{\gamma_i}{1-\gamma_i}$$

So $\hat{\hat{x}} = c_1 \hat{x}_1 + c_2 \hat{x}_2$ coincides with the least squares estimate if where

$$c_1 = \left.\frac{\rho_2(1+\rho_1)}{\rho_1+\rho_2+\rho_1\rho_2}\right|_{\rho_i = \frac{\gamma_i/\sigma^2}{1-\gamma_i^2/\sigma^2}} = \frac{\gamma_2}{\gamma_1+\gamma_2-\gamma_1\gamma_2}$$

and a similar formula may be derived for c_2

So we require

$$\underbrace{c_1 = \frac{\gamma_2}{\gamma_1+\gamma_2-\gamma_1\gamma_2}} \quad \text{and} \quad \underbrace{c_2 = \frac{\gamma_1}{\gamma_1+\gamma_2-\gamma_1\gamma_2}} \qquad [4]$$

3 (i) Fix t. Then writing $\hat{x}_{t|t-1} = E[x_t | y_{1:t-1}]$, $P_{t|t-1} = \mathrm{cov}[x_t | y_{1:t-1}]$ and $y_{t|t-1} = E[y_t | y_{1:t-1}]$ we have

$$\hat{x}_{t|t-1} = F\hat{x}_{t-1}, \quad y_{t|t-1} = H\hat{x}_{t|t-1}, \quad P_{t|t-1} = FP_{t-1}F^T + D_1D_1^T$$

We must calculate expectations $E[\cdot] = E[\cdot | y_{1:t-1}]$

$$E[(x_t - \hat{x}_{t|t-1})(y_t - y_{t|t-1})^T]$$
$$= E[(F(x_{t-1} - \hat{x}_{t-1}) + D_1 e_t)(HF(x_{t-1} - \hat{x}_{t-1}) + (HD_1 + D_2)e_t]$$
$$= FP_{t-1}F^TH^T + D_1(D_1^TH^T + D_2^T).$$

Also

$$E[(y_t - y_{t|t-1})(y_t - y_{t-1})]$$
$$= E[(HF(x_{t-1} - \hat{x}_{t-1}) + (HD_1 + D_2)e_t)(\ldots)^T]$$
$$= HFP_{t-1}F^TH^T + (HD_1 + D_2)(D_1^TH^T + D_2^T)$$

The standard "least squares" formulae now give

$$K_t = [FP_{t-1}F^TH^T + D_1(D_1^TH^T + D_2^T)][HFP_{t-1}F^TH^T + (HD_1 + D_2)(D_1^TH^T + D_2^T)]^{-1}$$
$$P_t = P_{t|t-1} - K_t[FP_{t-1}F^TH^T + D_1(D_1^TH^T + D_2^T)]^T$$

and

$$\hat{x}_t = \hat{x}_{t|t-1} + K_t[y_t - y_{t|t-1}].$$

[14]

(ii) Eliminating z_t from the state measurement eqns gives

$$\begin{cases} x_t = a x_{t-1} + e_t^1 \\ y_t = x_t - d e_t^1 + e_t^2 \end{cases}$$

This is an example of the "correlated noise" model in which

$$D_1 = d_1^T = [1 \quad 0] \quad \text{and} \quad D_2 = d_2^T = [-1, 1]$$

(also $H = 1$, $F = a$). In this case

$$D_1(D_1^TH^T + D_2^T) = 0 \quad \text{and} \quad (HD_1 + D_2)(HD_1 + D_2)^T = 1.$$

The estimator becomes

$$\hat{x}_t = a\hat{y}_{t-1} + a^2P_{t-1}[a^2P_t + 1]^{-1}(y_t - a\hat{x}_{t-1})$$

and

$$P_t = (a^2P_{t-1} + 1) - (a^2P_{t-1})^2 / (a^2P_{t-1} + 1)$$

[6]

4(i) The EKF is obtained by applying the Kalman filter to the approximate signal/measurement equations

$$\begin{cases} x_1 = f(\hat{x}_0) + \nabla f(\hat{x}_0)(x_0 - \hat{x}_0) \\ y_1 = x_1 + e_1 \end{cases}$$

Note that the right side of the state equation is the first order Taylor expansion of $f(\cdot)$ about the prior mean $\hat{x}_0$. This gives

$$\hat{x}_1 = f(\hat{x}_0) + K_1 [y_1 - f(\hat{x}_0)]$$

and $$P_1 = \nabla f P_0 \nabla f - \nabla f P_0 \nabla f [\nabla f P_0 \nabla f + r]^{-1} \nabla f P_0 \nabla f$$

for the estimate $\hat{x}_1$ and error covariance given y_1, in which

$$K_1 = \nabla f P_0 \nabla f [\nabla f P_0 \nabla f + r]^{-1}$$ [8]

(ii) $x_1 = f(x_0)$ and $x_0 \sim N(\hat{x}_0, P_0)$. So the linear least squares estimate of x_1 given x_0 is

$$E[x_1] + E[(x_1 - E[x_1])(x_0 - E[x_0])] \times E[(x_0 - E[x_0])(x_0 - E[x_0])]^{-1} (x_0 - \hat{x}_0)$$

$$= E[f(x_0)] + (E[f(x_0) x_0] - E[f(x_0)]\hat{x}_0) P_0^{-1} (x_0 - \hat{x}_0)$$

$$= \underline{u_0 + F(x_0 - \hat{x}_0)}.$$ [6]

(iii) When $f(x) = x^3$, $x_0 \sim N(0,1)$ and $e_1 \sim N(0,1)$

<u>EKF</u>

$f(\hat{x}_0) = \hat{x}^3 = 0$ $\nabla f(\hat{x}_0) = 3\hat{x}_0^2 = 0$. So $K_1 = P_1 = 0$

Estimate and error variance are $\hat{x}_1 = 0$ and $P_1 = 0$.
(does not depend on y_1)

<u>SLF</u>

$u_0 = E[x_0^3] = 0$, $F = E[x_0^4] P_0^{-1} = 3 P_0^{-1}$ (using hint)

Estimate and error variance are now

$$\hat{x}_1 = K_1 y_1$$

$$P_1 = 9/P_0 - \frac{(9/P_0)^2}{(9/P_0 + r)}$$

and $$K_1 = \frac{9}{P_0}\left(\frac{9}{P_0} + r\right)^{-1}$$

The SLF is superior, because it gives an estimate that is positively correlated with y_1 (which is clearly correct); the EKF does not depend on y at all. [6]

5/5

5 (i) $(I-K_tH)P_{t|t-1}(I-K_tH)^T + K_tQ_0K_t^T$

$= P_{t|t-1} - P_{t|t-1}H^T[..]^{-1}HP_{t|t-1} - P_{t|t-1}H^T[..]^{-1}HP_{t|t-1}$
$\quad + P_{t|t-1}H^T[..]^{-1}HP_{t|t-1}H^T[..]^{-1}HP_{t|t-1}$
$\quad + P_{t|t-1}H^T[..]^{-1}Q_0[..]^{-1}HP_{t|t-1}$

$= P_{t|t-1} - 2P_{t|t-1}H^T[..]^{-1}HP_{t|t-1} + P_{t|t-1}H^T[..]^{-1}\cancel{[HP_{t|t-1}H^T+Q_0]}\cancel{[..]}HP_{t|t-1}$

$= \underline{P_{t|t-1} - P_{t|t-1}H^T[HP_{t|t-1}H^T+Q_0]^{-1}HP_{t|t-1}}$ as required

The alternative error covariance update equation can be more suitable in computations, in presence of 'rounding' errors, because it automatically guarantees the updated P_t will be non-negative if P_{t-1} is non-negative [8]

(ii) $LLR^{-1}(x_1,x_2) = \dfrac{\log_e\left(\frac{1}{(2\pi)^{\frac{1}{2}}}\exp\{-\frac{1}{2}(|x_1-\hat{x}_A|^2+|x_2-\hat{x}_B|^2)\}\right)}{\log_e(\ \ldots\ (|x_1-\hat{x}_B|^2+|x_2-\hat{x}_A|^2)\})}$ $\left(=\dfrac{P_0(x_1,x_2)}{P_1(x_1,x_2)}\right)$

$= -\frac{1}{2}\left(|x_1-\hat{x}_A|^2+|x_2-\hat{x}_B|^2-|x_1-\hat{x}_B|^2-|x_2-\hat{x}_A|^2\right)$

$= -\frac{1}{2}\big(\cancel{x_1^2} - 2x_1\hat{x}_A + \cancel{\hat{x}_A^2} + \cancel{x_2^2} - 2x_2\hat{x}_B + \cancel{\hat{x}_B^2}$
$\quad \cancel{-x_1^2} + 2x_1\hat{x}_B \cancel{-\hat{x}_B^2} \cancel{-x_2^2} + 2x_2\hat{x}_A - \cancel{\hat{x}_A^2}\big)$

$= -\left(\hat{x}_A(-x_1+x_2) - \hat{x}_B(-x_1+x_2)\right) = \underline{(\hat{x}_A-\hat{x}_B)(x_1-x_2)}$ [4]

N-P test for validity of (H_0) is therefore (since $\hat{x}_A-\hat{x}_B>0$)
Accept H_0 if

$$\underline{(x_1-x_2) \geq c}$$

where c is such that
$P_0(x_1-x_2\geq c) = \alpha$ (probability of "false alarm" $=\alpha$)
But $E_0[x_1-x_2] = (\hat{x}_A-\hat{x}_B)$ and $\mathrm{cov}_1[x_1-x_2] = 1+1 = 2$
It follows

$P_0\left(\text{"}x_1-x_2\geq c\text{"} = \text{"}\dfrac{x_1-x_2-(\hat{x}_A-\hat{x}_B)}{\sqrt{2}} \geq \dfrac{c-(\hat{x}_A-\hat{x}_B)}{\sqrt{2}}\text{"}\right)$ $\sim N(0,1)$

$= \underline{1-\mathrm{erf}\left\{\dfrac{c-\hat{x}_A-\hat{x}_B)}{\sqrt{2}}\right\}} = \alpha$ [6]

same arguments as before

The "power" of the test is
$P_1(x_1-x_2\geq c) = \ldots = \underline{\mathrm{erf}\left\{\dfrac{c+(\hat{x}_A-\hat{x}_B)}{\sqrt{2}}\right\}}$ [2]