

Paper Number(s): **E4.26**  
**C2.2**

IMPERIAL COLLEGE OF SCIENCE, TECHNOLOGY AND MEDICINE  
UNIVERSITY OF LONDON

DEPARTMENT OF ELECTRICAL AND ELECTRONIC ENGINEERING  
EXAMINATIONS 2002

MSc and EEE PART IV: M.Eng. and ACGI

**ESTIMATION AND FAULT DETECTION**

(MSc only in 2002)

Wednesday, 1 May 10:00 am

There are SIX questions on this paper.

Answer FOUR questions.

Time allowed: 3:00 hours

**Examiners responsible:**

First Marker(s): Clark, J.M.C.

Second Marker(s): Allwright, J.C.

**Corrected Copy**

**Special instructions for invigilators:**

**None**

**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:

$$\begin{aligned}x_{k+1} &= Ax_k + Bu_k + Mv_k \\y_k &= Cx_k + Nw_k\end{aligned}$$

Here,  $v_k$  and  $w_k$  are standard white-noise sequences with identity covariance matrices.

The Kalman one-step-ahead predictor equations:

$$\begin{aligned}\hat{x}_{k+1|k} &= A\hat{x}_{k|k-1} + K(k)(y_k - C\hat{x}_{k|k-1}) \\K(k) &= AP_{k|k-1}(CP_{k|k-1}C^T + NN^T)^{-1} \\P_{k+1|k} &= AP_{k|k-1}A^T + MM^T - AP_{k|k-1}C^T(CP_{k|k-1}C^T + NN^T)^{-1}CP_{k|k-1}A^T\end{aligned}$$

The “completion of squares” identity for mean quadratic costs:

$$\begin{aligned}&E\left[\sum_{k=0}^{N-1} (x_k^T Q x_k) + x_N^T Q_N x_N\right] \\&= E\left[x_0^T S_0 x_0 + \sum_{k=0}^{N-1} (u_k + F_k x_k)^T (B^T S_{k+1} B + R)(u_k + F_k x_k)\right] + \sum_{k=0}^{N-1} \text{tr}(S_{k+1} M M^T) \\&\text{where for } k = 0, \dots, N-1,\end{aligned}$$

$$\begin{aligned}F_k &= (B^T S_{k+1} B + R)^{-1} B^T S_{k+1} A \\S_k &= A^T S_{k+1} A + Q - A^T S_{k+1} B (B^T S_{k+1} B + R)^{-1} B^T S_{k+1} A, \quad S_N = Q_N\end{aligned}$$

The algebraic Riccati equations:

$$S = A^T S A + Q - A^T S B (B^T S B + R)^{-1} B^T S A \quad (\text{control})$$

$$P = A^T P A + M M^T - A P C^T (C P A^T + N N^T)^{-1} C P A^T \quad (\text{filtering})$$

1. Suppose  $x_1(t)$  is the indefinite integral of a coloured noise process  $x_2(t)$  given by

$$\dot{x}_2 = -3x_2 + v,$$

where  $v(t)$  is continuous-time Gaussian white noise for which

$$E[v(t)v(s)] = \delta(t-s)$$

- (a) Show that the vector process  $x(t) = (x_1(t), x_2(t))^T$  satisfies, for  $t > s$ , the integral equation

$$x(t) = \begin{bmatrix} 1 & -\frac{1}{3}e^{-3(t-s)} \\ 0 & -e^{-3(t-s)} \end{bmatrix} x(s) + \int_s^t \begin{bmatrix} 1 & -\frac{1}{3}e^{-3(t-r)} \\ 0 & -e^{-3(t-r)} \end{bmatrix} \begin{pmatrix} 0 \\ 1 \end{pmatrix} v(r) dr$$

(Hint: first obtain the integral equation for  $x_2(t)$  alone). [8 marks]

- (b) Let  $x_k$  be the sampled process  $x(kh)$ . Show that  $x_k$  satisfies a difference equation of the form

$$x_{k+1} = \bar{A}x_k + \bar{v}_k$$

where  $\bar{v}_k$  is discrete-time vector white noise. Determine  $\bar{A}$  and the noise covariance  $Q = E[\bar{v}_k \bar{v}_k^T]$ .

Briefly explain why  $Q$  is non-singular. [7 marks]

- (c) Suppose that a Kalman filter has been constructed that generates the current conditional means  $\hat{x}_{k|k}$  of the sampled process from noisy measurements of its past and present values. Suppose that the value of  $x(t)$  at the intermediate time  $t = (k + \frac{1}{2})h$  is also of interest. Give an expression for its best estimate predicted at time  $kh$ . [5 marks]

- 2(a) Consider a sequence of observed random variables  $Y_1, Y_2, \dots, Y_n$  that are related to an unknown random variable  $X$  of mean  $m$  and variance  $p$  by

$$Y_k = X + N_k, \quad k = 1, \dots, n.$$

Here, each  $N_k$  is a zero-mean variable of variance  $q_k$  and the  $N_k$  and  $X$  are uncorrelated with each other. A feature of the linear least-squares estimate (LLSE) of a variable given a number of observed variables is that the error of estimation is uncorrelated with each of the observed variables. Use this characterization to establish that

$$\hat{X} = m + \frac{1}{p^{-1} + \sum_{k=1}^N q_k^{-1}} \left( \sum_{k=1}^N \frac{Y_k - m}{q_k} \right)$$

is the LLSE of  $X$  given  $Y_1, \dots, Y_n$  and determine the corresponding mean squared error  $\hat{p}$ . [12 marks]

- (b) An observer makes an initial estimate of the altitude of an aircraft as it passes directly overhead and then periodically estimates the altitude from measurements of its elevation as it recedes into the distance. The standard deviation of the initial estimate is 10 metres and that of the subsequent estimate computed from the  $k$ -th periodic measurement of the elevation is  $5k$  metres. If it is assumed that the aircraft remains at a constant altitude, what is the best accuracy, expressed as a standard deviation, that the observer can hope to achieve in his estimation of the aircraft altitude? [8 marks]

(Note:  $\sum_{k=1}^{\infty} \frac{1}{k^2}$  is approximately 1.64).

3. Suppose  $x_k$  and  $y_k$  are vector Gaussian processes satisfying the state space model

$$\begin{aligned}x_{k+1} &= Ax_k + Bu_k + Mv_k & x_0 \text{ normal} \\ y_k &= Cx_k + Nw_k .\end{aligned}$$

Here  $v_k$  and  $w_k$  are independent Gaussian white-noise sequences with zero means and identity covariances  $E[v_k v_k^T]$ ,  $E[w_k w_k^T]$ .  $u_k$  depends on  $y_k, y_{k-1}, \dots$ .

- (a) Establish the expressions for the one-step-ahead predictor  $\hat{x}_{k+1|k}$  and the Kalman gain  $K(k)$  that are given on page one. (You may use the fact that for vector normal random variables

$$E[X | Y] = EX + Cov(X, Y)Cov(Y)^{-1}(Y - EY). \quad [8 \text{ marks}]$$

- (b) Suppose  $z_k$  is a scalar controlled process described by

$$z_{k+1} = 0.2z_k + 0.3y_k + v_k$$

where  $y_k$  is an output measurement of the form

$$y_k = z_k + b + w_k .$$

Here,  $b$  is an unknown normal bias in the measurement and  $v_k$  and  $w_k$  are independent noise processes of unit variance.

An estimate of interest is the one-step-ahead predictor of  $z_k$ . Specify a state-space model from which an appropriate Kalman filter could be constructed.

[5 marks]

- (c) Suppose, for the model in (b), the value of the bias  $b$  is known. Determine the steady-state form of a first-order Kalman filter that generates  $\hat{z}_{k+1|k}$ . [7 marks]

4. Consider a stochastic linear system

$$x_{k+1} = Ax_k + Bu_k + Mv_k, \quad E[x_0] = 0, \quad \text{cov}[x_0] = P_0$$

and an averaged cost function of the special form

$$J_N^u = E\left[\sum_{k=0}^{N-1} (x_k^T Q x_k + u_k^T R u_k) + x_N^T S x_N\right]$$

where  $v_k$  is standard white noise and where  $S$  is assumed to be a positive definite solution of the control algebraic Riccati equation (ARE) on page one.

- (a) Using the “completion of squares” identity on page one, show that the control law that minimizes  $J_N^u$  over all control laws that are functions of the current state takes the time-invariant form

$$u_k = -F x_k$$

where

$$F = (B^T S B + R)^{-1} B^T S A. \quad [8 \text{ marks}]$$

- (b) The “rate” cost  $\bar{J}^u = \lim_{N \rightarrow \infty} \frac{1}{N} J_N^u$  is also minimized (over “stabilizing” control laws) by the control law in (a). Determine a formula for its optimal value. [4 marks]

- (c) Consider the scalar case where

$$x_{k+1} = x_k + u_k + m v_k, \quad x_0 = 0$$

and the “rate” cost is

$$\bar{J}^u = \lim_{N \rightarrow \infty} \frac{1}{N} \sum_{k=0}^{N-1} E[2x_k^2 + 4u_k^2].$$

Determine the optimal value of  $\bar{J}^u$  and the steady-state variances of  $x_k$  and  $u_k$  generated by the optimal law. [8 marks]

5. The behaviour of a scalar controlled process  $x_k$  is described by the equation

$$x_{k+1} = 0.5x_k + u_k + v_k, \quad x_0 = 0 \quad (1)$$

where  $v_k$  is independent Gaussian white noise of variance  $\sigma^2$ .

- (a)  $x_k$  is measured by a quantizer with an output  $y_k$  that takes values on a discrete set of levels  $\{kh, k \text{ an integer}\}$ . If the quantization parameter  $h$  is small compared with  $\sigma$ , the conditional probability density of  $x_k$  given  $y_k, y_{k-1}, \dots$ , can reasonably be approximated by the corresponding conditional probability density of the variable

$$\alpha \hat{x}_{k|k-1} + (1 - \alpha)y_k + b_k$$

where  $\hat{x}_{k|k-1}$  is the predicted mean of  $x_k$ ,  $\alpha = \frac{h^2}{12\sigma^2}$  and  $b_k$  is a uniform random variable on the interval  $[-\frac{h}{2}, \frac{h}{2}]$  independent of  $y_k$ . Show that the resulting conditional mean and variance of this variable are approximately the same as those of  $x_k$  given  $y_k, y_{k-1}, \dots$ , where now

$$y_k = x_k + w_k \quad (2)$$

and  $w_k$  is Gaussian white noise, independent of  $x_k$ , with zero mean and variance  $\frac{h^2}{12}$ . It may be assumed that the predicted covariance of  $x_k$  is approximately  $\sigma^2$ . [10 marks]

- (b) Suppose a cost  $E[\sum_0^{N-1} (q(x_k)^2 + r(u_k)^2)]$  is to be minimised. Explain what is meant by the “separation principle”. Demonstrate that it is valid under the assumption that observations  $y_k, y_{k-1}, \dots$  of the form (2) are available for the purposes of control at time  $k$ . If only the conditional means and covariances of  $x_k$  are available and this process may be conditionally non-Gaussian, as is the case in (a), describe how the separation principle has to be modified. [10 marks]

6. The presence or absence of a fault in a piece of equipment is assessed by a number of measurements  $y = (y_1, \dots, y_N)^T$  made at different locations. If no fault is present (" $F = 0$ "), the joint probability density of  $y$  is  $p_0(y)$ ; if a fault is present (" $F = 1$ "), the joint density is  $p_1(y)$ .
- (a) If the prior probabilities of a fault are  $P(F = 0) = \pi_0$ ,  $P(F = 1) = \pi_1$ , show that the Bayes test that minimizes the probability of error is:

$$\text{choose } F = 1 \quad \text{if } \frac{\pi_1 p_1(y)}{\pi_0 p_0(y)} \geq 1 ;$$

choose  $F = 0$  otherwise.

[10 marks]

- (b) Suppose that, if  $F = f$  ( $= 0$  or  $1$ ),

$y_1$  is normal with mean  $2f$  and variance 1

$y_2$  is normal with mean 0 and variance  $1 + f$

$y_3$  is uniformly distributed on  $[2f, 2f + 4]$

that, conditional on  $F$ , the measurements  $y_1, y_2, y_3$  are independent of each other and that  $\pi_0 = \pi_1 = 1/2$ .

Determine the outcome of the Bayes test in the two cases:

(i)  $y_1 = 1, y_2 = 0, y_3 = 3$

(ii)  $y_1 = 1, y_2 = 0, y_3 = 1$

and compute the probability of the decision being correct in the second case.

[10 marks]



## 1. Solution

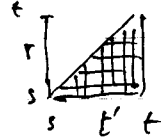
a) By the variation-of-constants formula

$$x_2(t) = e^{-3(t-s)} x_2(s) + \int_s^t e^{-3(t-r)} v(r) dr$$

But

$$\dot{x}_1(t) = x_1(s) + \int_s^t x_2(t') dt'$$

$$= x_1(s) + \int_s^t e^{-3(t'-s)} x_2(s) dt' + \int_s^t \int_s^{t'} e^{-3(t'-r)} v(r) dr dt'$$



$$+ \int_s^t \int_s^{t'} e^{-3(t'-r)} v(r) dr dt'$$

$$= x_1(s) + \frac{1}{3} e^{-3(t-s)} x_2(s)$$

$$+ \int_s^t \left( \int_r^t e^{-3(t'-r)} dt' \right) v(r) dr$$

$$= x_1(s) - \frac{1}{3} e^{-3(t-s)} x_2(s) + \int_s^t \left( -\frac{1}{3} e^{-3(t-r)} \right) v(r) dr$$

Combining the two gives the equation in (a).

b) Take  $s = kh$   $t = (k+1)h$ . From (a) we have

$$\text{that } x_{k+1} = \bar{A} x_k + \bar{v}_k \quad \text{where}$$

$$\bar{A} = \begin{bmatrix} 1 & -\frac{1}{3} e^{-3h} \\ 0 & -e^{-3h} \end{bmatrix}$$

$$\bar{v} = \int_s^t \begin{bmatrix} 1 & -\frac{1}{3} e^{-3(t-r)} \\ 0 & -e^{-3(t-r)} \end{bmatrix} \begin{bmatrix} 0 & 0 \\ 0 & 1 \end{bmatrix} \begin{bmatrix} 1 & 0 \\ -\frac{1}{3} e^{-3(t-r)} & -e^{-3(t-r)} \end{bmatrix} dr$$

$$= \int_s^t \begin{bmatrix} 0 & -\frac{1}{3} e^{-3(tr)} \\ 0 & -e^{-3(tr)} \end{bmatrix} \begin{bmatrix} 1 & 0 \\ -\frac{1}{3} e^{-3(tr)} & -e^{-3(tr)} \end{bmatrix} dr$$

$$= \int_0^h \begin{bmatrix} \frac{1}{9} e^{-6r} & \frac{1}{3} e^{-6r} \\ \frac{1}{3} e^{-6r} & e^{-6r} \end{bmatrix} dr = \frac{1}{6} (1 - e^{-6h}) \begin{bmatrix} \frac{1}{9} & \frac{1}{3} \\ \frac{1}{3} & 1 \end{bmatrix}$$

$$\hat{x}_{k+1/2|k} = E[x((k+1/2)h) | \text{past observations}] = \begin{bmatrix} 1 & -\frac{1}{3} e^{-3h/2} \\ 0 & -e^{-3h/2} \end{bmatrix} \hat{x}_{k|k}.$$

2002

2 Solution

a) The estimation error is

$$X - \hat{X} = X - m - \frac{\sum_{k=1}^n (Y_k - m)/q_k}{p^{-1} + \sum_{k=1}^n q_k^{-1}}$$

The correlation with  $Y_k$ , since  $X, N_k$  are uncorrelated is:

$$\begin{aligned} E[(X - \hat{X})(Y_k - m)] &= E\left[\left(X - m - \frac{\sum_{k=1}^n q_k^{-1}(X - m) + \sum_{k=1}^n N_k/q_k}{p^{-1} + \sum_{k=1}^n q_k^{-1}}\right)(X - m + N_k)\right] \\ &= p - \frac{\sum_{k=1}^n q_k^{-1} p}{p^{-1} + \sum_{k=1}^n q_k^{-1}} + \frac{q_k/q_k}{p^{-1} + \sum_{k=1}^n q_k^{-1}} \\ &= 0 \end{aligned}$$

This holds for all  $k$ ; so  $\hat{X}$  is the LLSE.

The mean squared error  $E[(X - \hat{X})^2]$

$$\begin{aligned} &= E\left[\left(\frac{p^{-1}(X - m) + \sum_{k=1}^n N_k/q_k}{p^{-1} + \sum_{k=1}^n q_k^{-1}}\right)^2\right] \\ &= \frac{p^{-2} E(X - m)^2 + \sum_{k=1}^n E N_k^2 / q_k^2}{(p^{-1} + \sum_{k=1}^n q_k^{-1})^2} = (p^{-1} + \sum_{k=1}^n q_k^{-1})^{-1} \end{aligned}$$

b) Take the initial variance of altitude  $X$  to be  $p = 100$

Variance of the measurement error  $N_k$  is  $q_k = 25k^2$ .

After  $n$  subsequent measurements, the LLSE  $\hat{X}$  of  $X$

has mean squared error (from part (a)) of

$$\frac{1}{\frac{1}{100} + \frac{1}{25} \sum_{k=1}^n \frac{1}{k^2}} = \frac{100}{1 + 4 \sum_{k=1}^n \frac{1}{k^2}}$$

As  $n$  increases, this decreases to

$$\frac{100}{1 + 4 \sum_{k=1}^{\infty} \frac{1}{k^2}} = \frac{100}{7.56}$$

So the smallest standard deviation is

$$\frac{10}{\sqrt{7.56}} \approx 3.6 \text{ metres}$$

2002

3. (solution)

- a) Prediction We may suppose  $x_k$  is conditionally Gaussian with mean  $\hat{x}_{k/k}$  and covariance  $P_{k/k}$ . Then it follows that

$$\hat{x}_{k+1/k} = A \hat{x}_{k/k} + B u_k$$

$$P_{k+1/k} = A P_{k/k} A^T + M M^T$$

as  $u_k$  is a function of  $(y_k, y_{k-1}, \dots)$ .

Updating Suppose we are conditioning on  $y_{k-1}, \dots$ ,

$$\text{Then } \text{Cov}(y_k / y_{k-1}, \dots) = C P_{k/k-1} C^T + N N^T$$

$$\text{Cov}(x_k, y_k / y_{k-1}, \dots) = P_{k/k-1} C^T$$

so

8

$$\hat{x}_{k/k} = \hat{x}_{k/k-1} + P_{k/k-1} C^T (C P_{k/k-1} C^T + N N^T)^{-1} \times (y_k - C \hat{x}_{k/k-1})$$

Combining this with the prediction equation gives

$$\hat{x}_{k+1/k} = A \hat{x}_{k/k-1} + \underbrace{A P_{k/k-1} C^T (C P_{k/k-1} C^T + N N^T)^{-1} (y_k - C \hat{x}_{k/k-1})}_{= K(k)} + B u_k$$

- b) Take  $x_k^1 = z_k$ ,  $x_k^2 = b$ . The state-space model that is suitable for filtering

is then. 
$$x_{k+1} = \begin{bmatrix} 0.2 & 0 \\ 0 & 1 \end{bmatrix} x_k + \begin{pmatrix} 0.3 \\ 0 \end{pmatrix} y_k + \begin{pmatrix} 1 \\ 0 \end{pmatrix} v_k$$

5

$$y_k = [1 \quad 1] x_k + w_k.$$

- c) If  $b$  is known we can replace  $y_k$  by  $\bar{y}_k = y_k - b = z_k + w_k$ .

$p = P_{k/k-1}$  (for  $z_k$ ) satisfies the ARE. where  $a = 0.2$ .

$$p = a^2 p + 1 - \frac{a^2 p}{p+1}$$

7

$$\text{or } 0.96 p(p+1) = p+1 - 0.04 p : 0.96 p^2 = 1$$

$$\Rightarrow p = 1.02 \Rightarrow K = \frac{(1.02) 0.2}{2.02} = 0.101 \text{ and } \hat{x}_{k+1/k} = 0.2 \hat{x}_{k/k-1} + 0.101 (y_k - b - \hat{x}_{k/k-1}).$$

2002

#### 4. Solution

(a) By the completion-of-squares identity

$$J_N^u = \text{tr}(SP_0) + \sum_{k=0}^{N-1} \text{tr}(S_{k+1} M M^T) + E \left[ \sum_{k=0}^{N-1} (u_k + F_k x_k)^T (B_{k+1}^T B_{k+1}^* + R) (u_k + F_k x_k) \right]$$

8 Since  $u_k$  may be a function of  $x_k$ ,  $u_k = -F_k x_k$  annihilates the last term and minimizes  $J_N^u$ .

The ARE implies that the solution to the difference Riccati eq.<sup>un</sup>

$$S_0 = S_1 = \dots = S_N = S \quad \text{and therefore that} \\ F_k = (B^T S B + R)^{-1} B^T S A.$$

(b) The optimal value of  $J_N^u$  in (a) is

$$J_N^0 = \text{tr}(SP_0) + N \text{tr}(S M M^T)$$

4

The optimal rate cost is then

$$\bar{J}^0 = \lim_{N \rightarrow \infty} \frac{1}{N} J_N^0 = \underline{\text{tr}(S M M^T)}$$

(c) The ARE becomes.

$$S = S + 2 - \frac{S^2}{S+4}$$

or

$$S^2 - 2S - 8 = 0$$

8

As  $S > 0$

$$S = 1 + \sqrt{9} = 4$$

and so

$$\underline{\bar{J}^0 = 4m^2}$$

The optimal control  $u_k = -F_k x_k = -\frac{S}{S+4} x_k = -\frac{1}{2} x_k$

Therefore

$$x_{k+1} = \frac{1}{2} x_k + m v_k$$

$$\text{Cor}(x_{k+1}) = \text{Cor}(x_k)$$

$$\text{and} \quad \text{Cor}(x_{k+1}) = \frac{1}{4} \text{Cor}(x_k) + m^2$$

gives

$$\text{Cor}(x_k) = \underline{\frac{4}{3} m^2}$$

$$\text{Cor}(u_k) = \text{Cor}\left(-\frac{1}{2} x_k\right) = \underline{\frac{m^2}{3}}$$

2002

### 5. solution

(a) The variable

$$\alpha \hat{x}_{k/k-1} + (1-\alpha) y_k + b_k$$

clearly has the conditional mean  $\alpha \hat{x}_{k/k-1} + (1-\alpha) y_k$

and conditional variance  $\frac{1}{h} \int_{-h/2}^{h/2} x^2 dx = \frac{h^2}{12}$

If  $y_k$  is given by (2) the 'updating' formula of filtering gives

$$\begin{aligned} \hat{x}_{k/k} &= \hat{x}_{k/k-1} + \frac{\text{Cov}(x_k, y_k | y_{k-1}, \dots)}{\text{Cov}(y_k | y_{k-1}, \dots)} (y_k - \hat{x}_{k/k-1}) \\ &= \hat{x}_{k/k-1} + \frac{P_{k/k-1}}{P_{k/k-1} + \frac{h^2}{12}} (y_k - \hat{x}_{k/k-1}) \end{aligned}$$

with  $P_{k/k-1} \approx \sigma^2$ ,

$$= \frac{\alpha \hat{x}_{k/k-1}}{1+\alpha} + \frac{y_k}{1+\alpha}$$

$$\approx \alpha \hat{x}_{k/k-1} + (1-\alpha) y_k$$

Similarly  $\text{Cov}(x_k | y_k, y_{k-1}, \dots) = \frac{h^2/12}{1+\alpha} \approx \frac{h^2}{12}$

b) The separation principle states that in the linear-quadratic-Gaussian set-up, the optimal control law is linear in the best estimate, and the coefficient  $F_k$  depends only on the 'control' parameters such as  $Q, R$ , (and  $A$  and  $B$ ). The best estimate & filter design depends only on statistical parameters such as  $\sigma^2, h^2$  and not on control parameters.

The key expression to be minimized in the completion-of-squares formula is

$$\begin{aligned} &\sum_{k=0}^{N-1} E[(u_k + F_k x_k)^T (B^T S_{k+1} B + R)(u_k + F_k x_k)] \\ &= \sum_{k=0}^{N-1} E[E[ \dots | y_k, y_{k-1}, \dots]] \\ &= \sum_k E[(u_k + F_k \hat{x}_{k/k})^T (B^T S_{k+1} B + R)(u_k + F_k \hat{x}_{k/k})] + \text{tr}(P_{k/k} (B^T S_{k+1} B + R) F_k^T F_k) \end{aligned}$$

which is minimized if  $u_k = -F_k \hat{x}_{k/k}$ .

In the non-Gaussian case, if the minimization is over

(linear (or affine) function of  $y_k, y_{k-1}, \dots$ , the principle still holds.

## 6 Solution

(a) Let  $\delta(y)$  denote the Bayes decision rule, with values 0 or 1.

Then the probability of error

$$\begin{aligned}
 &= P[\delta(y) \neq F] \\
 &= P(\delta(y) = 1 | F=0) \pi_0 + P(\delta(y) = 0 | F=1) \pi_1 \\
 &= \pi_0 \int_D p_0(y) dy + \pi_1 \int_{D^c} p_1(y) dy \quad \text{where } D = \{y: \delta(y) = 1\} \\
 &= \pi_0 \int_D p_0(y) dy + \pi_1 \left(1 - \int_D p_1(y) dy\right) \\
 &= \pi_0 \int_D \left(1 - \frac{\pi_1 p_1(y)}{\pi_0 p_0(y)}\right) p_0(y) dy + \pi_1
 \end{aligned}$$

This is minimized if  $D = \{y: 1 - \frac{\pi_1 p_1(y)}{\pi_0 p_0(y)} \leq 0\}$ , which gives the result.

$$(b) \frac{p_1(y_1)}{p_0(y_1)} = \exp\left(-\frac{(y_1-2)^2}{2} + \frac{y_1^2}{2}\right) = \exp(2y_1 - 2)$$

$$\frac{p_1(y_2)}{p_0(y_2)} = \sqrt{\frac{1}{2}} \exp\left(-\frac{y_2^2}{4} + \frac{y_2^2}{2}\right) = \sqrt{\frac{1}{2}} \exp \frac{y_2^2}{4}$$

$$\begin{aligned}
 \frac{p_1(y_3)}{p_0(y_3)} &= 0 & \text{if } 0 \leq y_3 \leq 2 \\
 &= 1 & \text{if } 2 < y_3 \leq 4 \\
 &= \infty & \text{if } 4 < y_3 \leq 6
 \end{aligned}$$

As the  $y_1, y_2, y_3$  are conditionally independent,  $\frac{p_1(y)}{p_0(y)} = \prod_{k=1}^3 \frac{p_1(y_k)}{p_0(y_k)}$ .

$$(i) \frac{p_1(y)}{p_0(y)} = 1 \cdot \sqrt{\frac{1}{2}} \cdot 1 \quad \text{So } \delta(y) = 0; F=0 \text{ is chosen}$$

$$(ii) \frac{p_1(y)}{p_0(y)} = 0 \quad \text{since } \frac{p_1(y_3)}{p_0(y_3)} = 0$$

So  $F=0$  is chosen

$$\text{Clearly } P(\delta(y) = F = 0 | y_3 = 1) = 1.$$