E4.26
C2.2

IMPERIAL COLLEGE LONDON

DEPARTMENT OF ELECTRICAL AND ELECTRONIC ENGINEERING
EXAMINATIONS 2003

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

ESTIMATION AND FAULT DETECTION

Tuesday, 13 May 10:00 am

Time allowed: 3:00 hours

There are SIX questions on this paper.

Answer FOUR questions.

Corrected Copy
Q4

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

Examiners responsible	First Marker(s) :	J.M.C. Clark
	Second Marker(s) :	J.C. Allwright

Special instruction for invigilators: **None**

Information for candidates

Some formulae relevant to the questions

The normal $N(m, \sigma^2)$ density: $p(y) = \frac{1}{\sqrt{2\pi\sigma^2}}\exp(-\frac{(y-m)^2}{2\sigma^2})$

System equations:

$$x_{k+1} = Ax_k + Bu_k + Mv_k$$
$$y_k = Cx_k + Nw_k$$

Here, v_k and w_k are standard white-noise sequences.

The Kalman filtering equations:

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

The average quadratic cost identity:

$$E\left[\sum_{k=0}^{N-1}(x_k^TQx_k + u_k^TRu_k) + x_N^TQx_N\right]$$
$$= E\left[x_0^TS_0x_0 + \sum_{k=0}^{N-1}(u_k + F_kx_k)^T(B^TS_{k+1}B + R)(u_k + F_kx_k)\right] + \sum_{k=0}^{N-1} tr(S_{k+1}MM^T).$$

where for $k = 0, ..., N-1$,

$$F_k = (B^TS_{k+1}B + R)^{-1}B^TS_{k+1}A$$
$$S_k = A^TS_{k+1}A + Q - A^TS_{k+1}B(B^TS_{k+1}B + R)^{-1}B^TS_{k+1}A, \; S_N = Q_N$$

The algebraic Riccati equations:

$$S = A^TSA + Q - A^TSB(B^TSB + R)^{-1}B^TSA, \qquad (control)$$
$$P = APA^T + MM^T - APC^T(CPC^T + NN^T)^{-1}CPA^T, \qquad (filtering)$$

1. Suppose that $x_1(t)$ is a doubly integrated white-noise process

$$\frac{d^2x_1}{dt^2} = v,$$

where $v(t)$ is Gaussian white noise with covariance function $E[v(t)v(s)] = \delta(t-s)$.

(a) Show that, if x_2 is taken to be dx_1/dt, the vector process $x(t) = (x_1(t), x_2(t))^T$ satisfies the integral equation

$$x(t) = \begin{bmatrix} 1 & t-s \\ 0 & 1 \end{bmatrix} x(s) + \int_s^t \begin{bmatrix} 1 & t-r \\ 0 & 1 \end{bmatrix} \begin{bmatrix} 0 \\ 1 \end{bmatrix} v(r)dr, \quad t \geq s.$$

(Hint: for the noiseless case, find by direct integration the fundamental matrix relating initial states to current states). [6]

(b) Suppose $x(0)$ is known. Give expressions for the mean and covariance of $x(t)$. [7]

(c) At a particular time t a single noise measurement is made of $x_1(t)$:

$$y_1 = x_1(t) + w,$$

where w is an independent normal random variable of zero mean and variance Q. Show that the conditional mean $E[x(t)|y_1]$ takes the form

$$E[x(t)|y_1] = x(0) + K_t(y_1 - x_1(0))$$

and give an expression for the gain vector K_t. Briefly explain why the first component of $K_t \to 1$ as $t \to \infty$. [7]

2. The state of a linear system

$$x_{k+1} = Ax_k \quad k = 0, 1, 2, \ldots$$

is measured by a series of observations

$$y_k = Cx_k + w_k, \quad k = 1, 2, \ldots$$

corrupted by a Gaussian white-noise process w_k of variance Q. The unknown initial state x_0 is normal with mean $\hat{x}_{0|0}$ and covariance $P_{0|0}$.

(a) Use the following version of Bayes' identity for conditional probability densities

$$\begin{aligned}\log p(x_0|y_k, y_{k-1}, \ldots) \; + \; \log p(y_k|y_{k-1}, \ldots) \\ = \log p(y_k|x_0, y_{k-1}, \ldots) \; + \; \log p(x_0|y_{k-1}, \ldots)\end{aligned}$$

to establish the recursive "information" filter

$$\begin{aligned} P_{0|k}^{-1}\hat{x}_{0|k} &= P_{0|k-1}^{-1}\hat{x}_{0|k-1} + (A^k)^T C^T Q^{-1} y_k \,, \\ P_{0|k}^{-1} &= P_{0|k-1}^{-1} + (A^k)^T C^T Q^{-1} C A^k \,, \end{aligned}$$

where $\hat{x}_{0|k}$ and $P_{0|k}$ are the conditional mean and covariance of x_0 given $y_k, y_{k-1}, \cdots$ [8]

(b) In the case where x_k is scalar, $Q = C = 1$ and $A = 0.99$, determine the limiting value of the conditional standard deviation of x_0 given $y_k, y_{k-1}, \ldots$, as $k \to \infty$ [6]

(c) In the case where x_k is scalar, $Q = C = 1$ and $A = 1.01$, show that $x_{0|k}$ is a consistent estimator of x_0 in the sense that the mean square error vanishes as $k \to \infty$. [6]

(Hint: the identity

$$1 + a + \ldots + a^{N-1} = \frac{a^N - 1}{a - 1}, \quad a \neq 1 \,,$$

is relevant to parts (b) and (c).)

3. A set of $2N+1$ measurements are made of a signal changing at a constant rate. These are modelled as

$$Y_k = A + Bk + w_k, \quad k \in \{-N, -N+1 \ldots, -1, 0, 1, \ldots, N\}$$

where A and B are unknown random variables representing a mean level and slope, and the w_k are uncorrelated noise terms of zero mean and variance Q, that are independent of A and B.

(a) Show that the estimates for A and B given by

$$\hat{A}_N = \frac{1}{2N+1} \sum_{k=-N}^{N} Y_k$$

and

$$\hat{B}_N = \frac{\sum_{k=1}^{N} k(Y_k - Y_{-k})}{2\sum_{k=1}^{N} k^2}$$

are unbiassed in the sense that

$$E[\hat{A}_N | A, B] = A, \quad E[\hat{B}_N | A, B] = B.$$

[6]

(b) Determine the conditional variance of $\hat{A}_N$ and $\hat{B}_N$ given A and B and show that $\hat{A}_N$ and $\hat{B}_N$ are uncorrelated. [7]

(c) An alternative conditionally unbiassed estimate of the slope B is

$$\bar{B}_N = \frac{Y_N - Y_{-N}}{2N}.$$

Show that this is asymptotically "inefficient" as an estimate of B when compared with $\hat{B}_N$, in the sense that

$$R_N = \frac{E[(\hat{B}_N - B)^2 | A, B]}{E[(\bar{B}_N - B)^2 | A, B]} \to 0 \quad \text{as } N \to \infty.$$

[7]

(Hint: You may find the inequality

$$\sum_{k=1}^{N} k^2 > \frac{N^3}{3}$$

useful in your argument.)

4. Suppose the scalar controls u_k of a linear system

$$x_{k+1} = Ax_k + bu_k$$

are to be chosen to ensure that the first component x_k^1 of the state tracks a scalar reference process z_k modelled by

$$z_{k+1} = az_k + v_k, \quad |a| < 1$$

where v_k is standard white noise of variance 1. The process z_k is measured via a noisy observation process

$$y_k = z_k + w_k$$

in which w_k is also standard whote noise of variance 1. The controls u_k, which may depend on x_k and $y_0, \ldots, y_k$, are chosen to minimise the cost function

$$E[\sum_{k=0}^{N-1}(x_k^1 - z_k)^2 + Ru_k^2].$$

(a) Show that the optimal control law also minimizes the cost function

$$E[\sum_{k=0}^{N-1}(x_k^1 - \hat{z}_{k|k})^2 + Ru_k^2]$$

where $\hat{z}_{k|k} = E[z_k|y_k, y_{k-1}, \cdots]$ [10]

(b) The control problem can be reformulated as a complete information LQG problem with the cost given in (a), in which $(x_k), \hat{z}_{k|k}$ is taken as an extended state. The subsystem of state equations for $\hat{z}_{k|k}$ are given by the Kalman filter.

Assuming that the prior distribution of z_0 is such that the coefficients of the Kalman filter are time-invariant, use the equations at the front of this paper to derive the coefficients of the filtering equation for $\hat{z}_{k|k}$ in the case where $a = 0.99$. This equation is a component of the system of stochastic difference equations for the extended state. Determine the variance of its noise term. [10]

changed to $(x_k^T, \hat{z}_{k|k}^T)^T$ before the start of the exam.

5. Consider the application of the first-order extended Kalman filter and the statistical linearization filter to the scalar process x_k:

$$
\begin{aligned}
x_{k+1} &= f(x_k) + v_k, \quad x_k \in R \\
y_k &= x_k + w_k .
\end{aligned}
$$

Here, y_k is an observation process and v_k and w_k are white-noise processes, of respective variances Q_s and Q_o, that are independent of x_k.

(a) The first-order extended Kalman filter is the ordinary Kalman filter for x_k applied to a modified model in which $f(x_k)$ is replaced by its first-order Taylor-series expansion. Determine an expresson for the one-step predictor $\hat{x}_{k+1|k}$ given by this filter in terms of the current state estimate $\hat{x}_{k|k}$. [4]

(b) The statistical linearization filter is based on the approximate model

$$
x_{k+1} = E_{k|k}[f(x_k)] + R_k P_{k|k}^{-1}(x_k - \hat{x}_{k|k}) + v_k
$$

where

$$
R_k = E_{k|k}[f(x_k)x_k] - E_{k|k}[f(x_k)]\hat{x}_{k|k}
$$

and where $E_{k|k}[.]$ denotes expectation with respect to the $N(\hat{x}_{k|k}, P_{k|k})$ normal density, $\hat{x}_{k|k}$ and $P_{k|k}$ being estimates of the conditional mean and variance of x_k given $y_k, y_{k-1}, \ldots$. Determine a predictor equation for $\hat{x}_{k+1|k}$ and the corresponding equation for $P_{k+1|k}$. [6]

(c) In the case where $f(x) = x - x^3$, express the predictor equation derived in (b) in terms of $\hat{x}_{k|k}$ and $P_{k|k}$.

(Hint: for normal $N(0,1)$ random variables X, $EX^3 = 0$). [5]

(d) What is your opinion on which filter is likely to give more accurate estimates for the example in (c)? [5]

6. For the controlled process

$$x_{k+1} = Ax_k + Bu_k + Mw_k, \quad x_0 = x,$$

where w_k is zero-mean Gaussian white noise with $Ew_j w_k^T = I\delta_{jk}$, the control feedback law for u_k is chosen to minimise $J_{0,N}^u(x)$, where

$$J_{0,N}^u(x) = E\left[\sum_{j=0}^{N-1}(x_j^T Q x_j + u_j^T R u_j) + x_N^T Q_N x_N \mid x_0 = x\right].$$

(a) Give a sufficient condition under which the algebraic form of the control Riccati equation (at the front of this paper) has a unique positive semi-definite solution. [3]

(b) Suppose this condition holds and that $Q_N = S$ is the unique solution. Using the average-quadratic-cost identity (also at the front of the paper) show that the optimal cost at $k = 0$ is

$$\min_u J_{0,N}^u = x^T S x + N\text{trace}(M^T S M)$$

[6]

(c) Taking Q_N to be S, as in (b), determine the value of the time-averaged cost:

$$\lim_{N\to\infty} \frac{\min_u J_{0,N}^u(x)}{N}.$$

For what class of control laws is this the minimum value of the time-average cost? [5]

(d) Suppose, for the set-up of (a), the only information available at time k is the set of noisy measurements $y_0, y_1, \ldots, y_{k-1}$, where

$$y_k = Cx_k + v_k$$

and v_k is white noise independent of w_k. Explain what is meant by the "separation" principle, and describe in qualitative terms the form of the control law in this particular case. [6]