IMPERIAL COLLEGE LONDON

E4.27
CS5.2
ISE4.41

DEPARTMENT OF ELECTRICAL AND ELECTRONIC ENGINEERING
EXAMINATIONS 2009

MSc and EEE/ISE PART IV: MEng and ACGI

IDENTIFICATION AND ADAPTIVE CONTROL Corrected Copy

Wednesday, 29 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) : S. Evangelou

Information for candidates:

The probability density $N(m, Q)$ of an n-vector, normal random variable with mean m and covariance matrix Q ($Q > 0$) is

$$N(m, Q)(x) = \frac{1}{(\sqrt{2\pi})^{\frac{n}{2}} (\det Q)^{\frac{1}{2}}} \exp -\frac{1}{2} \left((x - m)^T Q^{-1} (x - m) \right) .$$

In the case that $n = 1$, m is a scalar and $Q = \sigma^2$ ($\sigma^2 > 0$),

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

and, if X is a scalar random variable with probability density $N(m, \sigma^2)$,

$$\text{Prob}\{m - 2\sigma \leq X \leq m + 2\sigma\} \approx 0.95 .$$

1. Two coupled, stationary, chemical processes y_t and z_t are governed by the equations

$$\begin{aligned} y_t &= \alpha y_{t-1} + e_t \\ z_t &= \alpha z_{t-1} + \gamma y_{t-1} , \end{aligned}$$

in which e_t is a white noise process with variance $\sigma^2 > 0$. The constant α ($|\alpha| < 1$) is the reaction rate parameter for both processes. The constant γ is the coupling coefficient.

Let $x_t = (y_t,\ z_t)^T$. Develop a state space model for x_t, of the form

$$x_t = Fx_{t-1} + ge_t .$$

[2]

Show that $R = E[x_t x_t^T]$ satisfies the equation

$$R = FRF^T + bb^T\sigma^2$$

and derive formulae for the entries of R in terms of σ^2, α and γ. [4]

[6]

Now assume that the value of the reaction rate parameter is

$$\alpha = 0.5 .$$

Assume also that, by means of an identification experiment, it has been possible to establish the following relation between the variances $r_z = E[z_t^2]$ and $r_y = E[y_t^2]$:

$$\frac{r_z}{r_y} = 1.25 .$$

Determine the value of the coupling coefficient γ. [8]

2. An N-dimensional measurement vector is assumed to be modelled by the equation

$$y = x\theta + e$$

in which x is a known, non-zero, deterministic N-vector and e is a normal N-vector random variable zero mean and covariance matrix Q ($Q > 0$). θ is an unknown scalar parameter.

Consider the linear estimate $\hat{\theta}$ of θ given y:

$$\hat{\theta} = \left(x^T Q^{-1} x\right)^{-1} x^T Q^{-1} y \,.$$

Show that the estimate $\hat{\theta}$ is unbiased. [3]

Determine the variance of the estimate $\hat{\theta}$. [3]

Show that the estimate $\hat{\theta}$ minimizes the mean square error

$$E\left[|\hat{\hat{\theta}} - \theta|^2\right]$$

over all unbiased linear estimates $\hat{\hat{\theta}}$ of θ given y. [10]

Hint: Use the fact that an arbitrary linear, unbiased estimate $\hat{\hat{\theta}}$ can be expressed as

$$\hat{\hat{\theta}} = \left(x^T Q^{-1} x\right)^{-1} x^T Q^{-1} y \, + \, b^T y \,,$$

where b is an N-vector satisfying the condition:

$$b^T x = 0 \,.$$

Now assume that, for some integer $N \geq 2$, x is the N-vector $[1, \ldots, 1]^T$ and

$$Q^{-1} = \begin{bmatrix} 1 & 0.5 & 0 & 0 & . & 0 \\ 0.5 & 1 & 0.5 & 0 & . & 0 \\ . & . & . & . & . & \\ 0 & . & 0 & 0.5 & 1 & 0.5 \\ 0 & . & 0 & 0 & 0.5 & 1 \end{bmatrix} .$$

Determine a 0.95 confidence interval for θ, given the estimate $\hat{\theta}$. [4]

3. Let y_t be a stationary, ergodic, scalar process satisfying

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

in which e_t is a Gaussian, white noise process, with variance $\sigma^2 > 0$. The number a is an unknown, scalar parameter satisfying $|a| < 1$. Write

$$\hat{R}^N = \frac{1}{N}\sum_{t=0}^{N-1} y_t^2 \quad \text{(the sample covariance function of } y_t) .$$

What is the linear least squares estimate $\hat{a}_N$ of the parameter a, based on observations y_t, $t = 0, 1, \ldots, N$? [2]

Show that, as $N \to \infty$,

$$\hat{a}_N \to a .$$

[8]

Show further that

$$\hat{R}^N(\hat{a}_N - a) = -\sum_{i=1}^{N} y_{i-1}e_i .$$

[5]

Hence, or otherwise, show that

$$\mathrm{var}\{\hat{R}^N(\hat{a}_N - a)\} \le \frac{1}{N} \times \frac{\sigma^2}{(1-a^2)} ,$$

for all $N > 1$. [5]

Hint: To evaluate var$\{\hat{R}^N(\hat{a}_N - a)\}$, use the fact that

$$E[y_{t-1}e_t y_{t'-1}e_{t'}] = 0 \quad \textit{for } t' \neq t.$$

4 A control system relating the scalar control signal u_t and the scalar output signal y_t is modelled by the equations

$$y_t - ay_{t-1} = bu_t + e_t$$

in which e_t is a white noise sequence with variance σ^2. Here, a and b are scalar parameters. The true value of the parameter a is $a = 0$. b is non-zero.

In an identification experiment, the input signal is chosen to be samples of a process modelled as

$$u_t = v_t + 0.5v_{t-1}$$

in which v_t is a white noise sequence with unit variance, uncorrelated with e_t. It can be assumed that the joint process (y_t, u_t) is stationary and ergodic.

Let $(\hat{a}, \hat{b})$ be linear least squares estimate of (a, b), given $\{y_0, \ldots, y_N\}$ and $\{u_1, \ldots, u_N\}$, based on the assumption that both a and b are unknown parameters.

Let $\hat{\hat{b}}$ be the linear least squares estimate of b, based on the assumption that $a = 0$.

(a): Calculate $R_u(0) = E[u_t^2]$, $R_y(0) = E[y_t^2]$ and $R_{uy}(1) = E[u_t y_{t-1}]$. (In performing this calculation you should assume that $a = 0$ and b is an arbitrary non-zero number). [5]

(b): Obtain formulae for $\hat{b}$ and $\hat{\hat{b}}$, expressed in terms of sample convariances and cross-covariances of y_y and u_t, and the constant b. [7]

(c): Show that the conditional covariances $\hat{\gamma}$ and $\hat{\hat{\gamma}}$ of $\hat{b}$ and $\hat{\hat{b}}$ respectively, given $\{y_0, \ldots, y_N\}$ and $\{u_1, \ldots, u_N\}$, are, for N large, approximately

$$\hat{\gamma} = \frac{1}{N}(b^2 + \frac{4}{5})\sigma^2 \quad \text{and} \quad \hat{\hat{\gamma}} = \frac{1}{N} \times \frac{4}{5}\sigma^2 .$$

Comment on the relative magnitudes of the $\hat{\gamma}$ and $\hat{\hat{\gamma}}$. [7]

In (c) you should assume that sample covariances/cross-covariances can be replaced by covariances/cross-covariances. [1]

5. **(a):** Measurements y_1 and y_2 are taken at times $t = 1$ and $t = 2$ of a process governed by the ARMA model equations

$$y_t \;=\; e_t + c\,e_{t-1}\,,$$

in which $e_0 = 0$, and e_1 and e_2 are independent, zero mean, normal random variables, each with variance σ^2. c and σ^2 are unknown parameters.

Show that the 2-vector random variables $y = (y_1, y_2)^T$ and $e = (e_1, e_2)^T$ are related by

$$y \;=\; \begin{bmatrix} 1 & 0 \\ c & 1 \end{bmatrix} e\,.$$

Calculate the log likelihood function $L(c, \sigma^2; y)$ of c and σ^2:

$$L(c, \sigma^2; y) \;=\; \log_e\, p(y|c, \sigma^2)$$

in which $p(y|c, \sigma^2)$ denotes the probability density of y, given c and σ^2. [3]

Calculate the maximum likelihood estimates $\hat{c}$ and $\hat{\sigma^2}$ of c and σ^2; that is, the values of c and σ^2 maximizing $L(c, \sigma^2)$. [7]

Show that $\hat{\sigma^2}$ is an biased estimate of σ^2. [2]

Hint: Maximize the likelihood function first over c (for fixed σ^2) and then over σ^2.

(b): Describe the Generalized Least Squares Algorithm for estimating the parameters $a_1, \ldots, a_n$ and $d_1, \ldots, d_n$ in the model

$$\begin{cases} A(z)y_t &=& B(z)u_t + n_t \\ D(z)n_t &=& e_t \end{cases}$$

given measurements $y_1, \ldots, y_N$, $u_1, \ldots, u_N$ (and appropriate starting values). Here

$A(z) = 1 + a_1 z^{-1} + \ldots + a_n z^{-n}$, $B(z) = b_0 + b_1 z^{-1} + \ldots + b_n z^{-n}$, $D(z) = 1 + d_1 z^{-1} + \ldots + d_n z^{-n}$,

and $\{e_t\}$ is a white noise sequence with random variables, each with variance σ^2. [8]

Model Answers, Identification + Adaptive Control, 2009

1. Set $\underline{x}_t = \begin{pmatrix} y_t \\ z_t \end{pmatrix}$. Then $x_t^{(1)} = \alpha x_{t-1}^{(1)} + e_t$, $x_t^{(2)} = \gamma x_{t-1}^{(1)} + \alpha x_{t-1}^{(2)}$.

It follows

$$x_t = \begin{bmatrix} \alpha & 0 \\ \gamma & \alpha \end{bmatrix} x_{t-1} + \begin{bmatrix} 1 \\ 0 \end{bmatrix} e_t \qquad \left(E[e_t] = 0,\ \mathrm{var}\{e_t\} = \sigma^2 \right)$$

Write $x_t = F x_{t-1} + b e_t$, with

$$F = \begin{bmatrix} \alpha & 0 \\ \gamma & \alpha \end{bmatrix} \text{ and } b = \begin{bmatrix} 1 \\ 0 \end{bmatrix}$$

Write $R = E[x_t x_t^T]$. We have

$$E[x_t x_t^T] = E[F x_t x_t^T F^T] + b b^T \sigma^2$$

or

$$R = F R F^T + b b^T \sigma^2$$

Expand as

$$\begin{bmatrix} r_{11} & r_{12} \\ r_{12} & r_{22} \end{bmatrix} = \begin{bmatrix} \alpha & 0 \\ \gamma & \alpha \end{bmatrix} \begin{bmatrix} r_{11} & r_{12} \\ r_{12} & r_{22} \end{bmatrix} \begin{bmatrix} \alpha & \gamma \\ 0 & \alpha \end{bmatrix} + \sigma^2 \begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix}.$$

$$= \begin{bmatrix} \alpha & 0 \\ \gamma & \alpha \end{bmatrix} \begin{bmatrix} \alpha r_{11} & \gamma r_{11} + \alpha r_{12} \\ \alpha r_{12} & \gamma r_{12} + \alpha r_{22} \end{bmatrix} + \sigma^2 \begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix}$$

$$= \begin{bmatrix} \alpha^2 r_{11} + \sigma^2 & \alpha\gamma r_{11} + \alpha^2 r_{12} \\ \alpha\gamma r_{11} + \alpha^2 r_{12} & \gamma^2 r_{11} + \alpha\gamma r_{12} + \alpha\gamma r_{12} + \alpha^2 r_{22} \end{bmatrix}$$

Equating matrix entries gives

$$r_{11} = \alpha^2 r_{11} + \sigma^2, \quad r_{12} = \alpha\gamma r_{11} + \alpha^2 r_{12}$$

and $r_{22} = \gamma^2 r_{11} + 2\alpha\gamma r_{12} + \alpha^2 r_{22}$

Hence

$$\underline{r_{11} = \frac{\sigma^2}{1-\alpha^2}}, \quad r_{12} = \frac{\alpha\gamma}{(1-\alpha^2)} \times \frac{\sigma^2}{1-\alpha^2} = \frac{\alpha\gamma\sigma^2}{(1-\alpha^2)^2}$$

$$r_{22} = \frac{1}{(1-\alpha^2)} \left(\frac{\gamma^2\sigma^2}{1-\alpha^2} + \frac{2\alpha^2\gamma^2}{1-\alpha^2} \times \frac{\sigma^2}{1-\alpha^2} \right)$$

$$= \frac{1}{(1-\alpha^2)^3} \left(\gamma^2\sigma^2(1-\alpha^2) + 2\alpha^2\gamma^2\sigma^2 \right) = \underline{\frac{\sigma^2(1+\alpha^2)\gamma^2}{(1-\alpha^2)^3}}$$

If $\alpha = 1/2$

$$r_{11} = \tfrac{4}{3}\sigma^2, \quad r_{22} = \left(\frac{4^3}{3^3}\right) \times 1.25\,\gamma^2\sigma^2$$

and

$$r_{22}/r_{11} = \frac{4^2}{3^2} \times 1.25\,\gamma^2 = 1.25$$

So $\underline{\gamma = 3/4}$

2. Take $\hat{\theta} = (X^T Q^{-1} X)^{-1} X^T Q^{-1} y$, $\quad y = X\theta + e$, $\quad e \sim N(0, Q)$

Then $E[\hat{\theta}] = E[(X^T Q^{-1} X)^{-1} X^T Q^{-1}(X\theta + e)] = \theta + E[\underbrace{(X^T Q X)^{-1} X^T Q^{-1} e}_{=0}]$

We have shown $\hat{\theta}$ is unbiased.

$\hat{\theta} - \theta = (X^T Q^{-1} X)^{-1} X^T Q^{-1} e$ so

$$E|\hat{\theta} - \theta|^2 = E[(X^T Q^{-1} X)^{-2} X^T Q^{-1} e e^T Q^{-1} X] = (X^T Q^{-1} X)^{-2}(X^T Q^{-1} X)$$
$$= (X^T Q^{-1} X)^{-1}$$

An arbitrary linear, unbiased estimator can be written

$$\hat{\hat{\theta}} = (X^T Q^{-1} X)^{-1} X^T Q^{-1} y + b^T y$$

for some N-vector b. Because $\hat{\hat{\theta}}$ is unbiased

$$E[\hat{\hat{\theta}}] = \theta + (X^T Q^{-1} X)^{-1} X^T Q^{-1} \underbrace{E[e]}_{=0} + b^T X \theta + \underbrace{b^T E[e]}_{\to 0}$$

Since $E[\hat{\hat{\theta}}] = \theta$, for all θ, we must have

$$b^T X = 0 \quad \text{—— (1)}$$

Now examine mean square error of estimate

$$E[|\hat{\hat{\theta}} - \theta|^2] = E|(X^T Q^{-1} X)^{-1} X^T Q^{-1} e + \underbrace{b^T X \theta}_{\to 0 \text{ (by (1))}} + b^T e|^2$$
$$= E[[(X^T Q^{-1} X)^{-1} X^T Q^{-1} + b^T] e e^T [Q^{-1} X (X^T Q^{-1} X)^{-1} + b]]$$
$$= [\quad \cdots \quad] Q [\quad \cdots \quad]$$
$$= (X^T Q^{-1} X)^{-1} + 0 + 0 + b^T Q b$$

(we have used (1))

Since $b^T Q b \geq 0$, we have shown

$$E[|\hat{\hat{\theta}} - \theta|^2] \geq (X^T Q^{-1} X)^{-1} = E[|\hat{\theta} - \theta|^2].$$

This establishes that $\hat{\theta}$ is BLUE ("best linear, unbiased estimator.)

We know that $\hat{\theta}$ has mean θ and variance $(X^T Q^{-1} X)^{-1}$. Furthermore, it is an affine function of e, so it is normal: $\hat{\theta} \sim N(\theta, (X^T Q^{-1} X))$

It is now supposed that

$$X = (\underbrace{1, 1, \ldots, 1, 1}_{N \text{ terms}})^T \quad \text{and} \quad Q^{-1} = \begin{bmatrix} 1 & 0.5 & 0 & \\ 0.5 & 1 & 0.5 & \\ & 0.5 & 1 & \ddots \\ 0 & & 0.5 & 1 \end{bmatrix}$$

We see $X^T Q^{-1} X = [1, \ldots, 1] \begin{bmatrix} 1.5 \\ 2 \\ \vdots \\ 2 \\ 1.5 \end{bmatrix} = 2 \times 1.5 + (N-2) \times 2$
$$= N - 1 \quad (N > 2)$$

Hence $\hat{\theta} \sim N(\theta, \frac{1}{N-1})$. So $|\hat{\theta} - \theta| \leq 2/\sqrt{N-1}$ w.p. 0.95

We conclude

$$\hat{\theta} - \frac{2}{\sqrt{N-1}} \leq \theta \leq \hat{\theta} + \frac{2}{\sqrt{N-1}} \quad \text{w.p. } 0.95$$

7/5

3. $y_t = -a y_{t-1} + e_t$ — (1)

implies $\begin{pmatrix} y_1 \\ \vdots \\ y_N \end{pmatrix} = \begin{pmatrix} -y_0 \\ \vdots \\ -y_{N-1} \end{pmatrix} a + \begin{pmatrix} e_1 \\ \vdots \\ e_N \end{pmatrix}$; write as $\underline{y} = X\underline{a} + \underline{e}$

The linear l.s. estimate $\hat{a} = (X^T X)^{-1} X^T y = -\underbrace{\sum_{t=1}^{N} y_i y_0 \Big/ \sum_{t=1}^{N} y_{t-1}^2}$

But $N^{-1} \sum_i^N y_i y_0 \to R_y(1)$, $N^{-1} \sum_i^N y_i^2 \to R_y(0)$ as $N \to \infty$.

It follows that $\hat{a} \to -R_y(1) / R_y(0)$ $\neq 0$

But, from (1), $E\{y_t y_{t-1}\} = -a E[y_{t-1}^2] + \cancel{E[e_t y_{t-1}]}$,

whence $R_y(1) = -a R_y(0)$

We deduce that

$$\underset{\sim\sim\sim\sim\sim\sim\sim\sim\sim\sim}{\hat{a} \to a \quad \text{as } N \to \infty.}$$

We have

$$\left(\frac{1}{N} \sum_{t=1}^{N} y_{t-1}^2\right) \hat{a} = -\frac{1}{N} \sum_{t=1}^{N} y_t y_{t-1}$$

Multiply across 1 by y_{t-1} and summing gives

$$\frac{1}{N} \sum_{i=1}^{N} y_i y_{i-1} = -a \frac{1}{N} \sum_{i=1}^{N} y_{t-1}^2 + \frac{1}{N} \sum_{t=0}^{N} y_{t-1} e_t$$

It follows

$$\left(\frac{1}{N} \sum_{t=0}^{N-1} y_t^2\right)(\hat{a} - a) = \cancel{-\frac{1}{N} \sum y_t y_{t-1}} + \cancel{\frac{1}{N} \sum y_t y_{t-1}} - \frac{1}{N} \sum_{t=0}^{N} y_{t-1} e_t .$$

Hence

$$\hat{R}_y(0) \times (\hat{a} - a) = -\frac{1}{N} \sum_{t=0}^{N} y_{t-1} e_t$$

We see

$$E[\hat{R}_y(0) \times (\hat{a} - a)] = -E\left[\frac{1}{N} \ldots\right] = 0$$

Hence,

$$\operatorname{var}\{\hat{R}_y(0) \times (\hat{a} - a)\} = \frac{1}{N^2} E\left[\left(\sum_{t=1}^{N} y_{t-1} e_t\right) \times \left(\sum_{t=1}^{N} y_{t-1} e_t\right)\right]$$

$$= \frac{1}{N^2} E\left[\sum_{t=1}^{N} y_{t-1}^2\right] \sigma^2 + 0 + \ldots + 0 .$$

$$= \frac{\sigma^2}{N} R_y(0).$$

But, from (1) $R_y(0) = E[y_i^2] = E[(-a y_{t-1} + e_t)^2] = a^2 R_y(0) + \sigma^2$

Hence $R_y(0) = \sigma^2 / (1 - a^2)$

But then

$$\operatorname{var}\{\hat{R}_y(0) \times (\hat{a} - a)\} = \underset{\sim\sim\sim\sim\sim\sim}{\frac{1}{N} \times \frac{\sigma^4}{(1 - a^2)}}$$

4. (i) $R_u(0) = E[(v_t + \frac{1}{2}v_{t-1})(v_t + \frac{1}{2}v_{t-1})] = E[v_t^2] + 0 + 0 + \frac{1}{4}E[v_t^2] = \frac{5}{4}$

$R_y(0) = E[(b(v_t + \frac{1}{2}v_{t-1}) + e_t)^2] = b^2(1 + \frac{1}{4}) + \sigma^2 = \frac{5}{4}b^2 + \sigma^2$

$R_{uy}(L) = E[(v_t + \frac{1}{2}v_{t-1})(b(v_t + \frac{1}{2}v_{t-1}) + e_t)] = \frac{5}{4}b$

(ii) We write $\begin{bmatrix} y_1 \\ \vdots \\ y_N \end{bmatrix} = \begin{bmatrix} y_0 & u_1 \\ \vdots & \vdots \\ y_{N-1} & u_N \end{bmatrix} \begin{bmatrix} a \\ b \end{bmatrix} + \begin{bmatrix} e_1 \\ \vdots \\ e_n \end{bmatrix} \equiv y = X\theta + e$

Then $\begin{bmatrix} \hat{a} \\ \hat{b} \end{bmatrix} = (X^TX)^{-1}X^Ty = \begin{bmatrix} \frac{1}{N}\sum_{i=0}^{N-1} y_i^2 & \frac{1}{N}\sum_{i=1}^{N} u_i y_{i-1} \\ \frac{1}{N}\sum_{i=1}^{N} u_i y_{i-1} & \frac{1}{N}\sum_{i=1}^{N} u_i^2 \end{bmatrix}^{-1} \begin{bmatrix} \frac{1}{N}\sum_{i=1}^{N} y_i y_{i-1} \\ \frac{1}{N}\sum_{i=1}^{N} u_i y_i \end{bmatrix}$

$= \begin{bmatrix} \hat{R}_y(0) & \hat{R}_{uy}(1) \\ R_{uy}(L) & \hat{R}_u(0) \end{bmatrix}^{-1} \begin{bmatrix} R_y(1) \\ R_{uy}(L) \end{bmatrix}$. So $\hat{b} = [0\ 1] \begin{bmatrix} \hat{R}_y(0) & \hat{R}_{uy}(1) \\ \hat{R}_{uy}(1) & R_u(0) \end{bmatrix}^{-1} \begin{bmatrix} R_y(1) \\ R_{uy}(0) \end{bmatrix}$

If however we assume $a=0$, $\begin{bmatrix} y_1 \\ \vdots \\ y_N \end{bmatrix} = \begin{bmatrix} u_1 \\ \vdots \\ u_N \end{bmatrix} b + \begin{bmatrix} e_1 \\ \vdots \\ e_n \end{bmatrix} \equiv y = \bar{X}\bar{\theta} + e$

In this case, $\hat{\hat{b}} = (\bar{X}^T\bar{X})^{-1}\bar{X}^Ty = \left(\frac{1}{N}\sum_{i=1}^{N} u_i^2\right)^{-1}\left(\frac{1}{N}\sum_{i=1}^{N} u_i y_i\right)$

giving $\hat{\hat{b}} = \hat{R}_u(0)^{-1} R_{uy}(0)$.

(iii) The conditional variance of $\hat{b}$ (given $\{y_0, \dots, y_N\}$ and $\{u_1, \dots, u_N\}$):

$\hat{\gamma} = \sigma^2 [0\ 1](X^TX)^{-1}\begin{bmatrix} 0 \\ 1 \end{bmatrix} = \frac{\sigma^2}{N} \times [0\ 1]\begin{bmatrix} \frac{1}{N}\sum_i y_i^2 & \frac{1}{N}\sum_i u_i y_{i-1} \\ \frac{1}{N}\sum_i u_i y_{i-1} & \frac{1}{N}\sum_i u_i^2 \end{bmatrix}^{-1}\begin{bmatrix} 0 \\ 1 \end{bmatrix}$

$= \frac{\sigma^2}{N} \times \frac{1}{\left(\hat{R}_y(0)\hat{R}_u(0) - \frac{1}{N}\hat{R}_{uy}^2(0)\right)} \times \hat{R}_y(0)$

Also, the conditional variance of $\hat{\hat{b}}$ is

$\hat{\hat{\gamma}} = \sigma^2(\bar{X}^TX)^{-1} = \frac{\sigma^2}{N}\left(\frac{1}{N}\sum_i u_i^2\right)^{-1} = \frac{\sigma^2}{N}\hat{R}_u(0)^{-1}$

Using the results from (i), and assuming $\hat{R}_y(0) = R_y(0)$, etc, we have

$\hat{\gamma} = \frac{\sigma^2}{N} \frac{\left(\frac{5}{4}b^2 + \sigma^2\right)}{\frac{5}{4}\left(\frac{5}{4}b^2 + \sigma^2\right) - \left(\frac{5}{4}\right)^2 b^2} = \frac{1}{N}\left(b^2 + \frac{4}{5}\sigma^2\right)$

$\hat{\hat{\gamma}} = \frac{1}{N} \times \frac{4}{5}\sigma^2$

We observe that $\hat{\gamma} > \hat{\hat{\gamma}}$, reflecting the fact that the model (with a unknown) is over-parameterized and therefore gives rise to estimates of non-zero parameters with increased variance.

5/5

5. (a) $y_1 = e_1 + 0$ and $y_2 = e_2 + c e_1$, so $y = \begin{bmatrix} y_1 \\ y_2 \end{bmatrix} = \begin{bmatrix} 1 & 0 \\ c & 1 \end{bmatrix} \begin{bmatrix} e_1 \\ e_2 \end{bmatrix}$

The prob. density of y is:

$$p(y \mid c, \sigma^2) = \frac{1}{2\pi\, \sigma^2 \det Q} \exp\left\{ -\frac{\sigma^{-2}}{2} y^T Q^{-1} y \right\}$$

in which $Q = \begin{bmatrix} 1 & 0 \\ c & 1 \end{bmatrix} \begin{bmatrix} 1 & c \\ 0 & 1 \end{bmatrix} = \begin{bmatrix} 1 & c \\ c & 1+c^2 \end{bmatrix}$. Note, $\det Q = (1 + c^2 - c^2) = 1$

So $Q^{-1} = \begin{bmatrix} 1+c^2 & -c \\ -c & 1 \end{bmatrix}$.

$$L(c, \sigma^2) = \log_e p(y \mid c, \sigma^2) = -\log(2\pi) - \log(\sigma^2) - \frac{1}{2\sigma^2} (y_1, y_2) \begin{bmatrix} 1+c^2 & -c \\ -c & 1 \end{bmatrix} \begin{bmatrix} y_1 \\ y_2 \end{bmatrix}$$

$$= -\log(2\pi) - \log(\sigma^2) - \frac{1}{2\sigma^2} \left((1+c^2) y_1^2 - 2c\, y_1 y_2 + y_2^2 \right)$$

For fixed σ^2, the maximizing $c = \hat{c}$ satisfies

$$\partial/\partial c \left[\hat{c}^2 y_1^2 - 2\hat{c}\, y_1 y_2 \right] = 0 \implies \hat{c} = y_1 y_2 / y_1^2$$

Then

$$L(\hat{c}, \sigma^2) = -\log(2\pi) - \log(\sigma^2) - \frac{1}{2\sigma^2} \left(y_1^2 + y_2^2 - \frac{(y_1 y_2)^2}{y_1^2} \right)$$

$$= \text{const.} - \log(\sigma^2) - \frac{1}{2\sigma^2} y_1^2$$

The maximizing $\sigma^2 = \hat{\sigma}^2$ satisfies

$$\partial/\partial\sigma^2\, L(\hat{c}, \hat{\sigma}^2) = -\frac{1}{\hat{\sigma}^2} + \frac{1}{2\hat{\sigma}^4} y_1^2 = 0$$

Hence $\hat{\sigma}^2 = \frac{1}{2} y_1^2$.

We have shown that the maximum likelihood estimates are

$\hat{c} = y_1 / y_2$ and $\hat{\sigma}^2 = \frac{1}{2} y_1^2$.

We see $E[\hat{\sigma}^2] = \frac{1}{2} E[y_1^2] = \frac{1}{2} E[e_1^2] = \frac{1}{2}\sigma^2$. bias $= \frac{1}{2}\sigma^2$

(b) Generalized Least Squares Algorithm:

Choose $D^1(z)$ ($=1$, say). Compute $y_t' = D^0(z) y_t$, $u_t' = D^0(z) u_t$

Obtain LS estimate $A'(z), B'(z)$ of $A(z), B(z)$ for model

$$A(z)\, y_t' = B(z)\, u_t' + \text{"noise"}$$

given y_t', u_t'.

Calculate residuals $n_t' = A(z) y_t' - B(z) u_t'$.

Obtain LS estimate $D^2(z)$ of $D(z)$ for model

$$D(z)\, n_t' = 0 + \text{"noise"}$$

Repeat to obtain $(A^2(z), B^2(z))$, (D^3, A^3, B^3), ...