Uppsala University
Dept. of Systems and Control

Final Exam in System Identification for F and STS

Date and time: Friday, March 16, 2007, at 9–14
Aiding material: The course book *System Identification* by Söderström and Stoica, the mathematical handbook *BETA*, and a calculator. *No copies of any material are allowed whatsoever.*
Examiner: Magnus Mossberg

Read the following information carefully.

- Give clear and structured solutions to the problems in Swedish or in English. Do not hand in anything else but one solution to each problem. Do not solve more than one problem on each sheet of paper and do not write on the back of the sheets. Write your name on every sheet of paper. Hand in the solutions in numerical order. In your solutions, you can refer to results and equations in the book by giving page- or equation numbers.
- If you find anything in a problem formulation unclear, see to that the assumptions you make when solving the problem are clearly stated in your solution.
- The problems are not necessarily given in an increased order of difficulty and subproblems are not necessarily dependent in the sense that their solutions build upon each other.
- The last problem is an alternative to the homework assignments. In case you hand in a solution to the last problem, you will be accounted for the best performance of the homework assignments and the last problem.

Good luck!

1. Consider the system

$$y(t) + a_0 y(t-1) = b_0 u(t-1) + w(t)$$

where the true and unknown parameters a_0 and b_0 are to be estimated. Here, $u(t)$ is zero mean white noise of variance σ^2 and

$$w(t) = e(t) + c_0 e(t-1)$$

where $e(t)$ is zero mean white noise of variance λ^2, uncorrelated with $u(t)$.

a) Show that the least squares estimate of a_0 and b_0 is not consistent. (3p)

b) Suggest a consistent estimator for a_0 and b_0. (2p)

2. A DC-level d in white noise $v(t)$, of zero mean and variance λ^2, is to be estimated from measurements of the signal

$$x(t) = d + v(t)$$

a) The estimator

$$\check{d} = \frac{1}{N}\sum_{t=1}^{N} x(t)$$

of d is suggested. What is the mean square error of $\check{d}$? Note that the mean square error is defined as the variance plus the squared bias. (3p)

b) Consider the estimator

$$\bar{d} = a\frac{1}{N}\sum_{t=1}^{N} x(t)$$

of d, where a is some constant. Find the a that minimizes the mean square error of the estimator $\bar{d}$. Is the estimator $\bar{d}$ realizable (that is, can it be written solely as a function of the data) with this choice of a? (3p)

c) Consider the criterion function

$$V_t(d) = \sum_{k=1}^{t} \lambda^{t-k}(x(k) - d)^2$$

where $\lambda < 1$. Give the estimator $\hat{d}(t)$ that minimizes $V_t(d)$, and find a recursive form

$$\hat{d}(t) = \hat{d}(t-1) + f(\lambda)\big(x(t) - \hat{d}(t-1)\big)$$

where $f(\lambda)$ is a function of λ. (Note that $f(\lambda)$ should be given explicitly.) (4p)

3. a) Consider the system

$$y(t) = b_1 u(t-1) + b_2 u(t-2) + e(t)$$

where $e(t)$ is white noise. Assume that the input signal is chosen as $u(t) \equiv u_0$, where u_0 is a constant. Is the system identifiable? (Motivate.) (1p)

b) Consider the system

$$y(t) + ay(t-1) = bu(t-1) + e(t)$$

where $e(t)$ is white noise. Assume that the system is controlled by a proportional controller

$$u(t) = -cy(t)$$

under the data collection (c is known). Show that the system is not identifiable and suggest a way to change the controller so that the system becomes identifiable. (2p)

4. Give brief answers to the following questions.

a) Assume that the transfer function for a system is estimated by spectral analysis and that the amplitude curve has two resonance peaks. If a parametric model for the transfer function is searched for, what model order is suitable to start with? (Motivate.) (1p)

b) Why is it easier to estimate the parameters in an ARX-structure than in an ARMAX-structure (assume that the number of parameters is the same in both cases)? (1p)

c) Why is it dangerous to choose a too high model order? (1p)

d) Within system identification, parametric- and non-parametric methods are distinguished. What is a common disadvantage with the results given by the non-parametric methods? (1p)

e) Assume that a system with a complicated disturbance dynamic is modelled with an ARX-model and an ARMAX-model, respectively. What can be expected for the A-polynomials of the two models and why? (1p)

f) Why is it important that a model passes a cross-validation test in a satisfactory way? (1p)

5. A system is given by

$$y(t) + a_0 y(t-1) = e(t)$$

where $|a_0| < 1$ and $e(t)$ is white noise with variance λ^2. It holds that

$$r_y(\tau) = \mathrm{E}\{y(t)y(t-\tau)\} = (-a_0)^\tau \frac{\lambda^2}{1-a_0^2}, \qquad \tau \geqslant 0$$

Consider the model

$$y(t) + a_1 y(t-1) + a_2 y(t-2) = e(t)$$

and let

$$\hat{\theta}_N = [\hat{a}_1 \quad \hat{a}_2]^T$$

denote the estimates of a_1 and a_2 given by the prediction error method.

a) Determine the variances of the estimates in $\hat{\theta}_N$ when the number of data N is large. (2p)

b) What is $\hat{\theta}_N$ when the number of data $N \to \infty$? Comment upon the result. (3p)

6. A system is described by

$$y(t) = G_0(q)u(t) + v(t)$$

where $v(t)$ is zero mean white noise with variance λ^2.

a) Determine an estimate of $G_0(e^{i\omega})$ based on the impulse-response estimate

$$\hat{g}(t) = \frac{y(t)}{\alpha}$$

where α is the amplitude of the pulse. Show that this estimate coincides with the empirical transfer function estimate

$$\hat{\hat{G}}_N(e^{i\omega}) = \frac{Y_N(\omega)}{U_N(\omega)}$$

where

$$Y_N(\omega) = \frac{1}{\sqrt{N}} \sum_{t=0}^{N-1} y(t)e^{-i\omega t} \quad \text{and} \quad U_N(\omega) = \frac{1}{\sqrt{N}} \sum_{t=0}^{N-1} u(t)e^{-i\omega t}$$

(3p)

b) Applying the input

$$u(t) = \begin{cases} \beta, & t \geqslant 0 \\ 0, & t < 0 \end{cases}$$

gives the output

$$y(t) = \beta \sum_{k=1}^{t} g_0(k) + v(t)$$

which motivates the impulse response estimate

$$\bar{g}(t) = \frac{y(t) - y(t-1)}{\beta}$$

Compute the error

$$\tilde{g}(t) = \bar{g}(t) - g_0(t)$$

and its variance. (3p)

7. Consider the ARMA-process

$$\begin{aligned} A(q^{-1})y(t) &= C(q^{-1})e(t) \\ A(q^{-1}) &= 1 + a_1 q^{-1} + \ldots + a_{na} q^{-na} \\ C(q^{-1}) &= 1 + c_1 q^{-1} + \ldots + c_{nc} q^{-nc} \end{aligned}$$

A method to estimate the AR-part is given as follows. Let

$$\hat{r}(\tau) = \frac{1}{N} \sum_{t=\tau}^{N} y(t)y(t-\tau)$$

Then solve for $\hat{a}_i$ from

$$\hat{r}(\tau) + a_1 \hat{r}(\tau - 1) + \ldots + a_{na} \hat{r}(\tau - na) = 0$$

where $\tau = nc + 1, nc + 2, \ldots, nc + na$. Show that this essentially is an application of the instrumental variable method using specific instruments. Which ones? (5p)

8. *This problem is an alternative to the homework assignments.*

a) Consider the criterion

$$C_p = \frac{\sum_{t=1}^{N} \varepsilon^2(t, \hat{\theta}_N)}{\hat{s}_N^2} - (N - 2p)$$

for selecting model structures. Here, p is the number of estimated parameters and $\hat{s}_N^2$ is an estimate of the innovations variance, normally taken as the normalized sum of prediction errors for the largest model structure considered. The criterion C_p is to be minimized with respect to p. Discuss the relationship between C_p and Akaike's information criterion (AIC). (4p)

b) It is known from a Bode plot that a second order system has a resonance peak at $\omega = 1$ rad/s. Where are the two poles located? (3p)

c) Consider the scalar analytic function $J(x)$ which has the quadratic approximation

$$Q(x) = J(x_k) + J^{'}(x_k)(x - x_k) + \frac{1}{2}J^{''}(x_k)(x - x_k)^2$$

around x_k, where $J^{'}(x)$ and $J^{''}(x)$ denote the first and second order derivatives of $J(x)$, respectively. Take x_{k+1} as the minimum point of the quadratic approximation and derive the Newton-Raphson algorithm. (3p)