System Identification

(Course: SSY 230)

Exam 4th June 2019

Time: 14:00-18:00

Examiner: Jonas Sjöberg, phone 772 1855 (073-0346321)

Total number of credits is 50. Preliminary grade limits are 23 for grade 3, and 32 for grade 4, and 40 for grade 5.

Solutions should be clearly formulated so that it is easy to follow each step.

This is an "open book" exam. Help material:

- The course book
- Mathematical handbook such as BETA, handbook with physical formulas such as Physics Handbook
- Standard Calculator
- Lecture slides
- Written notes but no solved examples.

Solutions will be published on the course web after the exam.

Result of the exam is delivered in mail by the LADOK system.

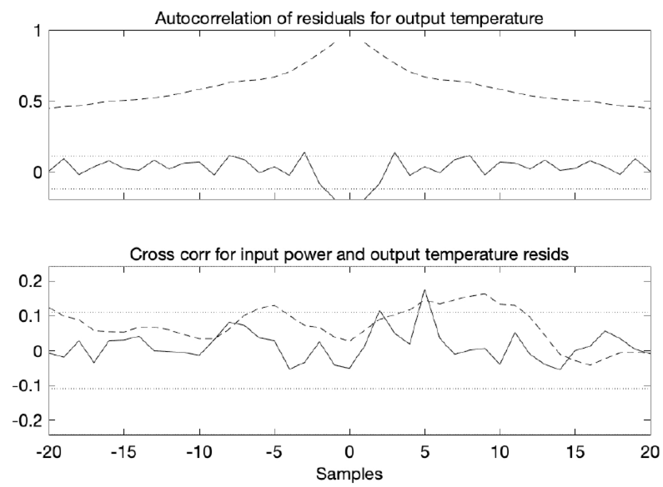
Evaluation of grading at the examiner's office, please email for an appointment. Latest 1st of Septembre.

Good Luck!

Department of Electrical Engineering Chalmers University

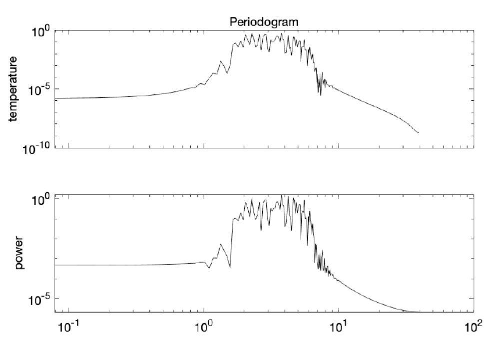
Problem 1

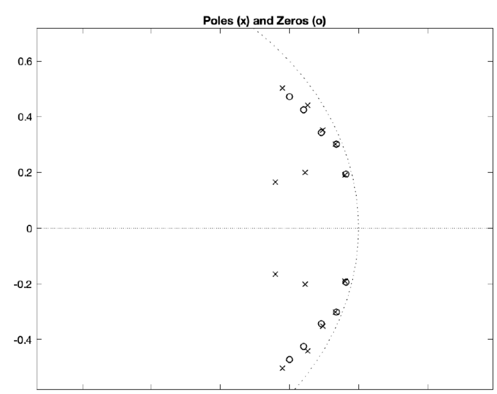
(a) In a system identification process an ARX model and an OE model have been identified. In the figure the auto-correlation of the residuals is shown for both models, and also the correlation betwen the input and the residuals.



Which one is for the ARX model and which is for the OE model? Correct motivation is required. (3p)

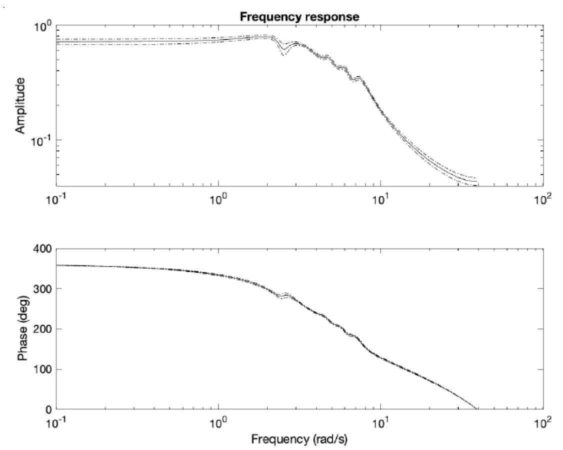
(b) A high order ARX model is identified with the data which spectra is shown to the left. The zeros and poles of the model are shown in the plot to the right. Why are they placed at a part of the unit circle?





(3p)

(c) In the figure the Bode plot including uncertainty is shown for the model in b). Although the main part of energy of the data is within a frequency band, the uncertainty of the model is not particularly large outside that band. Why?



(4p)

Problem 2
Consider the system

$$\mathcal{S}:\ y(t) = -a\,y(t-1) + b\,u(t-1) + v(t)$$

where $\{v(t)\}$ is a sequence of independent identically distributed stochastic variables, each with the probability density function $f(x) = \mu e^{-\mu x},\ \ x \geq 0$. Design the Maximum-Likelihood method that permits estimation of a and b. You do this by formulating the optimization problem, with criterion to be minimized and constraints. The solution of the minimization problem does not need to be given.

What is the difference in algorithmic complexity for an unknown μ and for a known value of μ?

(10p)

Problem 3

(a) Consider a moving average (MA) process $y(t) = b_1 u(t-1) + \cdots + b_m u(t-m) + v(t)$, and formulate the least-squares estimate of the process parameters $b_1, \ldots, b_m$ given N measurements $\{u(t), y(t)\}_{t=1}^{N}$. (4p)

(b) Consider the two possibilities, $v(t)$ white and colored, respectively. Can you guarantee a consistent estimate if you make some assumption on the cross-correlation of $v(t)$ and $u(t)$ for the two cases? (6p)

Problem 4

The P-matrix in the recursive least-squares (RLS) algorithm, except for the factor representing the noise variance, asymptotically represents the parameter covariance in the case of constant parameters and an uncorrelated noise sequence.

Modify/complement the RLS algorithm so that also an estimate of the noise variance is obtained.

Note that the noise variance estimate should be formulated so that the need of memory does not increase with time.

(10p)

Problem 5

Consider an MRAS where the plant is

$$G(s) = \frac{1}{s}$$

and the reference model

$$G_r(s) = \frac{\theta^0}{s}$$

The control is given by

$$u(t) = \theta\, r(t)$$

and the parameter adaption by

$$\theta(t) = -\gamma_1 r(t) e(t) - \gamma_2 \int^{t} r(\tau) e(\tau) d\tau$$

Determine the differential equation for $e(t)$ and discuss how γ_1 and γ_2 influence the convergence rate. Assume $r(t)$ to be constant.

Except γ_1 and γ_2, the differential equation can depend on $r(t)$, and θ_0. Dependence on $\theta(t)$ should be eliminated. To "obtain more equations" you take the derivative of the expressions of $e(t)$ and $\theta(t)$.

(10p)

Exam 4th June 2019: Solutions

1

(a) The ARX model can model the disturbance in difference from the OE model, and can hence have low auto-correlation of the residuals.

(b) The data is in a narrow frequency band and the fit of the model will be most important in that band. Hence, poles and zeros will be placed near the corresponding part of the unit circle.

(c) Apperently, the values of the parameters can be identified with a small uncertainty with the band limited data, and then the uncertainty becomes small everywhere.

2

We have:

$$v(t) = y(t) + a\,y(t-1) - b\,u(t-1) = y(t) - \theta\varphi(t)$$

where $\varphi^T(t) = [-y(t-1)\ u(t-1)]$ and $\theta = [a\ b]$. We have the probability function:

$$f(x) = \mu\, e^{-\mu\, x} \quad x \geq 0$$

So using Maximum Likelihood estimation with the data points, we get the definition of the estimate as

$$\arg\max_{\theta} \prod_{t=2}^{N} \mu\, e^{-\mu(y(t)-\theta\varphi(t))} = \mu^N e^{-\mu \sum_{t=2}^{N}(y(t)-\theta\varphi(t))}$$

Take the logarithm of this gives

$$\arg\max_{\theta} N \log \mu - \mu \sum_{t=2}^{N} (y(t) - \theta\, \varphi(t))$$

If μ is known we then want to maximize

$$\max_{\theta} -\mu \sum_{t=2}^{N} (y(t) - \theta, \varphi(t))$$

Subject to:

$$y(t) - \theta\,\varphi(t) \geq 0, \quad t = 2, \ldots, N$$

This can be simplified and expressed on matrix form

$$\max_{\theta} \theta \sum_{t=2}^{N} \varphi(t)$$

Subject to

$$\Phi\,\theta^T \leq Y$$

with

$$Y = \begin{bmatrix} y(2) \\ \vdots \\ y(N) \end{bmatrix}, \quad X = \begin{bmatrix} -y(1) & u(1) \\ \vdots & \vdots \\ -y(N-1) & u(N-1) \end{bmatrix}$$

If μ also needs to be estimated, the optimization problem becomes

$$\max_{\theta,\mu} \; N \log \mu - \mu \sum_{t=2}^{N} y(t) + \mu\,\theta \sum_{t=2}^{N} \varphi(t)$$

Subject to

$$\Phi\,\theta^T \leq Y, \quad \mu > 0$$

which is a much nastier problem.

3

(a) We have the system

$$y(t) = b_1 u(t-1) + ... + b_m u(t-m) + v(t)$$

and the model

$$\hat{y}(t) = \hat{b}_1 u(t-1) + ... + \hat{b}_m u(t-m) = \theta^T \varphi(t)$$

where $\theta = [b_1, \ldots, b_m]^T$ and $\varphi(t) = [u(t-1), \ldots, u(t-m)]^T$.

The least squares estimate is

$$\hat{\theta} = [\frac{1}{N}\sum_{t=1}^{N}\varphi(t)\varphi^T(t)]^{-1}\frac{1}{N}\sum_{t=1}^{N}\varphi(t)y(t)$$

(b) The LS estimate in (a) can be expressed as

$$\hat{\theta} = [\frac{1}{N}\sum_{t=1}^{N}\varphi(t)\varphi^T(t)]^{-1}\frac{1}{N}\sum_{t=1}^{N}\varphi(t)y(t) = [\frac{1}{N}\sum_{t=1}^{N}\varphi(t)\varphi^T(t)]^{-1}\frac{1}{N}\sum_{t=1}^{N}\varphi(t)(\varphi^T(t)\theta+v(t))$$
$$= \theta + [\frac{1}{N}\sum_{t=1}^{N}\varphi(t)\varphi^T(t)]^{-1}\frac{1}{N}\sum_{t=1}^{N}\varphi(t)v(t)$$

and hence

$$\hat{\theta} - \theta = [\frac{1}{N}\sum_{t=1}^{N}\varphi(t)\varphi^T(t)]^{-1}\frac{1}{N}\sum_{t=1}^{N}\varphi(t)v(t)$$

This expression goes to zero (assuming $u(t)$ is exciting the system) if

$$\frac{1}{N}\sum_{t=1}^{N}\varphi(t)v(t) \to 0 \;\text{ when } N \to \infty$$

This means that $v(t)$ must be uncorrelated with the components of $\varphi(t)$ for the estimate to be consistent, and this is independent on if $v(t)$ is white or colored.

- To have a consistent estimate in the case $v(t)$ is white, $u(t)$ cannot depend on future $v(t)$ but it can be correlated with past $v(t)$. That is, feedback is possible.
- In the case $v(t)$ is colored, then it must be uncorrelated with $u(t)$ since a colored $v(t)$ cannot be correlated with *only* future $u(t)$. That is, it must be an open-loop experiment.

4

Let p be the number of estimated parameters. We formulate the estimate of the noise variance up to time t and from that we separate the last contribution so that we can form it as an update of the estimate from time $t-1$, that is

$$\hat{\sigma}_t^2 = \frac{1}{t-p}\sum_{k=1}^{t}\varepsilon_k^2 = \frac{1}{t-p}(\sum_{k=1}^{t-1}\varepsilon_k^2 + \varepsilon_t^2) = \frac{t-1-p}{t-p}\frac{1}{t-1-p}\sum_{k=1}^{t-1}\varepsilon_k^2 + \frac{1}{t-p}\varepsilon_t^2 = \frac{t-1-p}{t-p}\hat{\sigma}_{t-1}^2 + \frac{1}{t-p}\varepsilon_t^2$$

5

We have adaption law,

$$\theta(t) = -\gamma_1 r(t)e(t) - \gamma_2 \int^t r(\tau)e(\tau)d\tau \tag{1}$$

plant

$$Y(s) = \frac{\theta}{s}R(s)$$

where θ is the unknown, possibly varying, parameter, and the reference model

$$Y_m(s) = \frac{\theta_0}{s}R(s)$$

The error is defined as the difference between the plant output and the output of the reference model

$$e = y - y_m \tag{2}$$

Since we are looking for the differencial equation of $e(t)$, we take the derivative of (2)

$$\frac{de}{dt} = (\theta(t) - \theta_0)r(t)$$

once more gives

$$\frac{d^2e}{dt^2} = \dot{\theta}(t)r(t) + (\theta(t) - \theta_0)\dot{r}(t) \tag{3}$$

Derivative of (1),

$$\dot{\theta}(t) = -\gamma_1\dot{r}(t)e(t) - \gamma_1 r(t)\dot{e}(t) - \gamma_2 r(t)e(t)$$

Insert this in (3)

$$\frac{d^2e}{dt^2} = (-\gamma_1\dot{r}(t)e(t) - \gamma_1 r(t)\dot{e}(t) - \gamma_2 r(t)e(t))r(t) + (\theta(t) - \theta_0)\dot{r}(t)$$

Re arrange this

$$\frac{d^2e}{dt^2} + \gamma_1 r^2(t)\dot{e}(t) + (\gamma_1 \dot{r}(t) r(t) + \gamma_2 r^2(t))e(t) = (\theta(t) - \theta_0)\dot{r}(t)$$

With a constant reference this becomes

$$\frac{d^2e}{dt^2} + \gamma_1 r^2(t)\dot{e}(t) + \gamma_2 r^2(t)e(t) = 0$$

So the error has second order dynamics whose properties are defined by γ_1, γ_2 and the constant value of $r(t)$.