System Identification, 1RT875

08 March 2011

Find below the questions for the exam on system identification. The exam is **open book**, and students may use textbook, solutions manual and any (paper-version) of a compendium with mathematical formulas, as well as the slides used in the lectures and exercise sessions. It is **not** allowed to use anything to communicate (as e.g. an electronic devise - please do switch off your mobile). Neither laptop nor (advanced) electronic calculator might be used.

- **Time:** 08 March 2011, 8am-1pm.
- **Location:** (F4Sy) Lokal: Gimogatan 4, Sal 2
- **Answers:** (i) put you number on each page, (ii) answer each full question on separate pages. Give an actual answer to the questions as stated, and mark the key steps in your reasoning.

Ex.1 (6/30 pt):

Consider the system

$$Y_t = a_0 + W_t,$$

where $\{W_t\}_t$ are i.i.d. stochastic assumed to be white Gaussian noise with zero mean and variance σ_t^2 (Note that the variance is variable for different t). Let $\{y_t\}$ be n samples of the stochastic process $\{Y_t\}$.

1. Consider a model

   $$y_t = a + \epsilon_t,$$

   and let a_{LS} be the Least Squares estimate. Derive an analytical expression for its bias and variance.

2. What would the BLUE a_{BLUE} for this problem be? derive an expression for its variance.

3. Consider 3 cases:

   - (a) $\sigma_t^2 = 1$
   - (b) $\sigma_t^2 = t$ (Note that $\sum_{t=1}^{\infty} \frac{1}{t} \to \infty$)
   - (c) $\sigma_t^2 = t^2$

Compare the a_{LS} and a_{BLUE} on those 3 cases. When are either of them consistent?

Ex.2 (6/30 pt):

Consider the following system for given $m > 1$

$$Y_t = D_t + c_1 D_{t-1} + \cdots + c_m D_{t-m}$$

where $\{D_t\}_t$ are i.i.d. white, zero man stochastic variables with variance σ^2. Derive the auto-covariance function of Y_t.

1. What is the variance of Y_t?

2. What is the auto-covariance for $\tau > m$?

3. What is the auto-covariance for $0 < \tau \leq m$?

4. When is the resulting auto-covariance matrix for $\tau = 0, \ldots, m$ indefinite?

Ex.3 (6/30 pt)

Given the predictor

$$f_{t+1|t}(\theta) = \frac{1-\alpha}{1-\alpha q^{-1}} u_t,$$

where $|\alpha| < 1$ is a constant.

1. For which model is this the optimal predictor, and what assumptions do we make on the noise process so that this is L_2 optimal?

2. What is the optimal two-step ahead predictor $f_{t+2|t}(\theta)$?

3. If using a numerical optimization procedure for implementing PEM, one wants to use the gradient of $\epsilon(t+1|t) = y_t - f_{t|t-1}(\theta)$ with respect to α. Give a closed form expression of this gradient.

Ex.4 (6/30 pt):

Consider the following system

$$y_t + a y_{t-1} = b(t) U_{t-1} + e_t,\ t = 1, \ldots, 300$$

where u_t and e_t are independent samples from the random variables $\{U_t\}_t$ and $\{D_t\}$ which are independent white noise sequences with mean zero. The noise has a small variance of $E[D_t^2] = 0.1$, and $E[U_t^2] = 1$. The parameter $a = -0.7$ is constant, while the parameter $b(t)$ varies over time. The identification experiment considers the following model

$$y_t = \varphi_t^T \theta + \epsilon_t$$

where $\varphi_t = (-y_{t-1}, u_{t-1})^T$. The parameters $\theta = (a, b)$ are estimated as follows:

- (a) The following 4 recursive estimators are applied to estimate θ of the above model.
  - (1) A Recursive Least Squares (RLS) estimate with forgetting-factor $\lambda = 0.93$.
  - (2) A Kalman filter based on the parameter time-varying model

    $$\begin{cases} \hat{\theta}(t) = \hat{\theta}(t-1) + v_t \\ y_t = \varphi_t^T \hat{\theta}(t) \end{cases}$$

    with the following covariance matrix for the zero mean, white process modeling the samples $\{v_t\}_t$

    $$R_v = \begin{bmatrix} 0 & 0 \\ 0 & 0.01 \end{bmatrix}$$

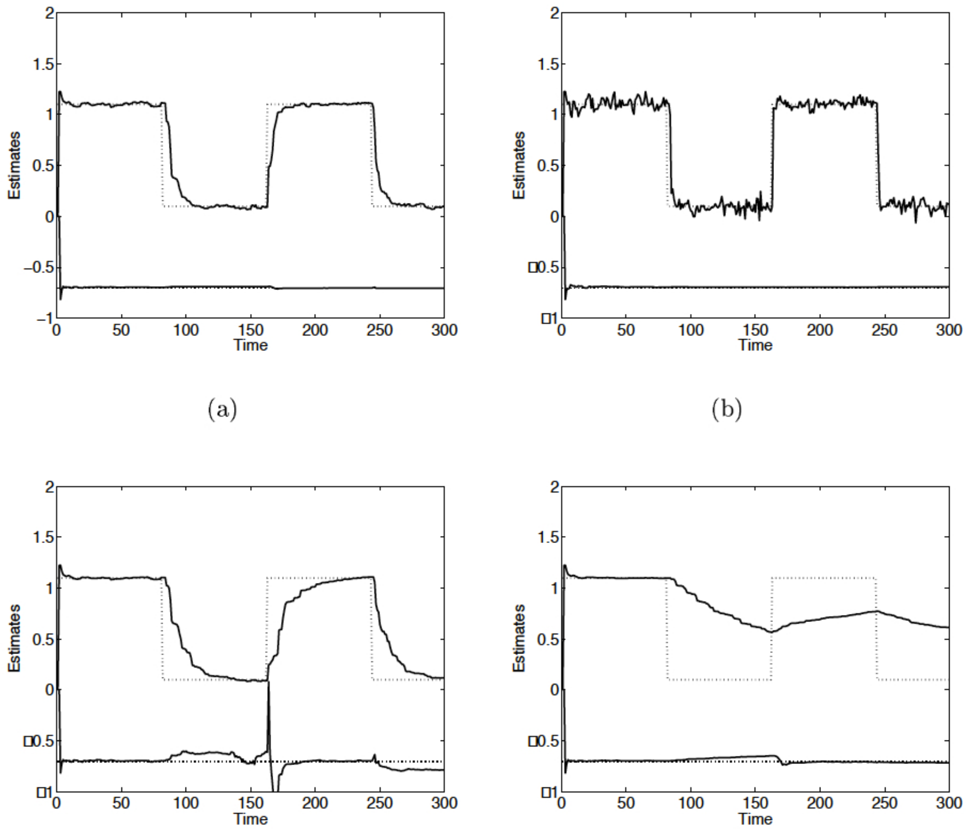
  - (3) The same approach as before with covariance matrix

    $$R_v = \begin{bmatrix} 0 & 0 \\ 0 & 1 \end{bmatrix}$$

  - (4) A Recursive Least Squares (RLS) approach with forgetting-factor $\lambda = 1$.

  The results from the experiments are displayed in the Figures (a)-(d) below. The dashed lines are the 'true' parameter values, and the solid lines indicate the respective estimates. Panel (a) corresponds with approach (a.1), and so on. Explain what you see here, i.e. why/in which case is each of the 4 methods a good choice, and which one is preferable in this particular case.

- (b) Consider panel (c). Explain the effect you see at $t = 160$. Why is this effect not apparent with the other techniques?



Ex.5 (6/30 pt)

Answer the following questions with a short answer (i.e. True/False + One sentence concisely motivating why you make that choice).

1. Given n i.i.d samples from a Gaussian distribution, is the ML estimator of the variance based on the samples equivalent to the LS estimate of this one based on the same samples?

2. Using the Recursive Pseudo-Least Squares algorithm for estimating the parameters of a ARMAX model gives always consistent estimates (1pt).

3. Does asymptotic unbiasedness of an estimator imply consistency of that estimator?

4. Given a random variable Z with mean μ and variance 0, and let X be a random variable possibly depending on Z. Is $\mathbb{E}[ZX] = \mu\mathbb{E}[X]$?

5. The presence of feedback from y_t to u_{t+1} always leads to bad LS estimates.

6. Is a stochastic stationary process necessarily ergodic?