Exercise Set 8

1. *Persistence of excitation for ARX models* (Problem 13E.2 of [1])
Consider the ARX structure

$$A(q)y_t = B(q)u_t + w_t$$

where the degree of $B(q)$ is n_b. Show that an open-loop input that is persistently exciting of order n_b is sufficiently informative with respect to this model set, regardless of the order of $A(q)$, provided that the process noise is persistently exciting.

2. *Influence of the input on the asymptotic variance of the noise model* (Problem 13E.3 of [1])
Consider a model structure

$$y_t = G(q,\rho)u_t + H(q,\eta)w_t$$

with independent parameterization of G and H. Show that it is not possible, if G can be correctly modeled, to affect the asymptotic accuracy of $\hat{\eta}$ (using PEM) by changing the input signal in an open-loop experiment.

3. *Effect of the input signal on the estimation of the static gain* (Exercise 10 of [2])
Consider the system

$$y_t + a_0 y_{t-1} = b_0 u_{t-1} + w_t, \qquad E\{w_t w_s\} = \lambda_0 \delta_{t,s}$$

identified with the LS method in the model structure

$$y_t + a y_{t-1} = b u_{t-1} + \varepsilon_t$$

The static gain $S = b_0/(1+a_0)$ can be estimated as

$$\hat{S}_N = \frac{\hat{b}_N}{1+\hat{a}_N}$$

The variance of $\hat{S}_N$ can be (approximately) evaluated from

$$\hat{S}_N - S = \frac{\hat{b}_N}{1+\hat{a}_N} - \frac{b_0}{1+a_0} = \frac{-b_0(\hat{a}_N - a_0) + (\hat{b}_N - b_0)(1+a_0)}{(1+\hat{a}_N)(1+a_0)} \approx \frac{-b_0(\hat{a}_N - a_0) + (\hat{b}_N - b_0)(1+a_0)}{(1+a_0)^2}$$

Thus we have expressed $\hat{S}_N - S$ as a *linear* combination of $\hat{\theta}_N - \theta_0$. Then the variance of $\hat{S}_N$ can be easily found from the covariance matrix of $\hat{\theta}_N$.

(a) Compute the variance of $\hat{S}_N$ for two experimental conditions:

(a.1) $\{u_t\}$ is zero mean white noise of variance σ^2

(a.2) $\{u_t\}$ is a step signal of size σ

(Note that in both cases we have $\bar{E}\{u_t^2\} = \sigma^2$). Which case will give the smallest variance? Evaluate the variance numerically when $a_0 = -0.9$, $b_0 = 1$, $\lambda_0 = 1$ and $\sigma^2 = 1$.

(b) Compute the variance for cases (a.1) and (a.2) using Ljung's asymptotic (in model order) variance formulas, and compare the results with those obtained in part (a).

(c) Simulate the system with the parameter values given in (a) for different sample sizes N. For which values of N the variance results of (a) hold reasonably well?

4. (**Advanced**) *Convergence of the ML estimator of sinusoids with Gaussian noise* (Problem 9.12 of [3]) Consider a (real) signal given by

$$y_t = \sum_{m=1}^{M} A_m \cos(\omega_m t - \phi_m) + w_t$$

where $\{w_t\}$ is zero mean Gaussian white noise of (known) variance σ^2. We want to estimate the parameters $\theta := \{A_1, \omega_1, \phi_1, \ldots, A_M, \omega_M, \phi_M\}$ via maximum likelihood, given the data $Y^N := \{y_1, \ldots, y_N\}$. Assuming that θ_0 belongs to a (known) compact set $\Theta \subset \mathbb{R}^{3M}$, show that the ML estimator $\hat{\theta}_N$ converges almost surely to the true parameters θ_0 by carrying the following steps:

(a) Let

$$x_t(\theta) := \sum_{m=1}^{M} A_m \cos(\omega_m t - \phi_m)$$

Prove that the negative of the normalized log-likelihood can be expressed as

$$-\frac{2}{N}\log P\{Y^N \mid \theta\} = \log 2\pi + \log \sigma^2 + \frac{1}{N\sigma^2}\left\{\sum_{t=1}^{N}[x_t(\theta) - x_t(\theta_0)]^2 + 2\sum_{t=1}^{N} w_t[x_t(\theta) - x_t(\theta_0)] + \sum_{t=1}^{N} w_t^2\right\}$$

(b) Using the Borel-Cantelli lemma and Chebyshev's inequality, prove that $N^{-1}\sum_{t=1}^{N} w_t[x_t(\theta) - x_t(\theta_0)]$ converges to zero almost surely. Also, show that $N^{-1}\sum_{t=1}^{N} w_t^2$ converges to σ^2 almost surely. From these results, conclude that $-2N^{-1}\log P\{Y^N \mid \theta\}$ converges almost surely as $N \to \infty$ (to what?). Is the convergence uniform in $\theta \in \Theta$?
How can you deduce from (b) about the convergence of $\hat{\theta}_N$?
Hint. Since $\{w_t\}$ is Gaussian, it has bounded moments of every order. Therefore, it is possible to avoid using the full argument of the method of subsequences by considering moments of sufficiently high order in Chebyshev's inequality.

(c) The term $N^{-1}\sum_{t=1}^{N}[x_t(\theta) - x_t(\theta_0)]^2$ is obviously zero for $\theta = \theta_0$, and non-negative otherwise. However, to prove the consistency of $\hat{\theta}_N$, it is necessary to show that it is strictly positive in the limit as $N \to \infty$. Prove that $\lim_{N\to\infty} N^{-1}\sum_{t=1}^{N}[x_t(\theta) - x_t(\theta_0)]^2$ is strictly positive for all $\theta \neq \theta_0$.

References

[1] L. Ljung. *System Identification: Theory for the User*, 2nd Edition. Prentice-Hall, 1999.
[2] T. Söderström. *Lecture Notes in Identification*. Uppsala Universitet, 1984.
[3] B. Porat. *Digital Processing of Random Signals: Theory and Methods*. Prentice-Hall, 1994.