System Identification
Spring 2009

Final Exam Solutions

Problem 1

a) The PEM model \( \hat{y}(t/t - 1; \theta) \) is obtained from the true system by setting
\[
\epsilon(t, \theta_0) = y(t) - \hat{y}(t/t - 1; \theta_0) = e(t),
\]
which for the output error system
\[
y(t) = \frac{b_0 q^{-1}}{1 + a_0 q^{-1}} u(t) + e(t),
\]
gives the model structure
\[
\hat{y}(t/t - 1; \theta) = \frac{b q^{-1}}{1 + a q^{-1}} u(t),
\]
where \( \theta = (a, b)^T \).
As this model structure includes the true system and the PEM estimator is consistent the PEM parameter estimates will asymptotically converge to the true system parameters, i.e. \( \lim_{N \to \infty} \hat{\theta}_N = \theta_0 = (a_0, b_0)^T \).

b) For large \( N \) the covariance of the PEM parameter estimate is given by
\[
\text{Cov}(\hat{\theta}_N) = \frac{\sigma^2}{N} R_{\Psi(\theta_0)},
\]
where
\[
R_{\Psi(\theta_0)} = E \left( \Psi(\theta_0) \Psi^T(\theta_0) \right) = E \left( \begin{pmatrix} -\frac{\partial \epsilon(t, \theta_0)}{\partial \theta} \end{pmatrix}^T \begin{pmatrix} -\frac{\partial \epsilon(t, \theta_0)}{\partial \theta} \end{pmatrix} \right).
\]

With
\[
\epsilon(t, \theta) = y(t) - \frac{b q^{-1}}{1 + a q^{-1}} u(t),
\]
the partial derivatives are evaluated as:

\[
\frac{\partial z(t, \theta)}{\partial a} = -\frac{-bq^{-2}}{(1+aq^{-1})^2}u(t)
\]

\[
\frac{\partial z(t, \theta)}{\partial b} = -\frac{q^{-1}}{1+aq^{-1}}u(t)
\]

Using the given series expansions for \(1/(1+x)\) and \(1/(1+x)^2\) with \(x = aq^{-1}\) we obtain

\[
\frac{\partial z(t, \theta)}{\partial a} = \left( \sum_{n=0}^{\infty} (n+1)(-aq^{-1})^n \right) bq^{-2}u(t)
\]

\[
\frac{\partial z(t, \theta)}{\partial b} = -\left( \sum_{n=0}^{\infty} (-aq^{-1})^n \right) q^{-1}u(t).
\]

\(R_\Psi(\theta_0)\) then becomes

\[
R_\Psi(\theta_0) = E \left( \begin{pmatrix} \sum_{n=0}^{\infty} (n+1)(-a_0q^{-1})^n \end{pmatrix} b_0q^{-2}u(t) \vphantom{\sum_{n=0}^{\infty} (n+1)(-a_0q^{-1})^n} \right) T
\]

\[
- \left( \sum_{n=0}^{\infty} (-a_0q^{-1})^n \right) q^{-1}u(t)
\]

\[
= \left( \sum_{n=0}^{\infty} (n+1)^2(-a_0)^{2n} \right) b_0^2\sigma_u^2 - \left( \sum_{n=0}^{\infty} (n+1)(-a_0)^{2n+1} \right) b_0\sigma_u^2
\]

\[
- \left( \sum_{n=0}^{\infty} (n+1)(-a_0)^{2n+1} \right) b_0^2
\]

\[
= \sum_{n=0}^{\infty} (n+1)^2(a_0^2)^n \vphantom{\sum_{n=0}^{\infty} (n+1)(a_0^2)^n}
\]

\[
\sum_{n=0}^{\infty} (n+1)(a_0^2)^n \vphantom{\sum_{n=0}^{\infty} (n+1)(a_0^2)^n}
\]

\[
= \sigma_u^2 \begin{pmatrix} 35.1509 \ast 0.04 & -7.7160 \ast 0.8 \ast 0.2 \\ -7.7160 \ast 0.8 \ast 0.2 & 2.7778 \end{pmatrix} = \sigma_u^2 \begin{pmatrix} 1.4060 & -1.2346 \\ -1.2346 & 2.7778 \end{pmatrix}.
\]

Finally this gives the covariance matrix

\[
Cov(\hat{\theta}_N) = \frac{\sigma_u^2}{N\sigma_u^2} \left( \begin{array}{cc} 1.4060 & -1.2346 \\ -1.2346 & 2.7778 \end{array} \right)^{-1}
\]

\[
= \frac{\sigma_u^2}{N\sigma_u^2} \frac{1}{2.3813} \left( \begin{array}{cc} 2.7778 & 1.2346 \\ 1.2346 & 1.4060 \end{array} \right) = \frac{\sigma_u^2}{N\sigma_u^2} \left( \begin{array}{cc} 1.1665 & 0.5185 \\ 0.5185 & 0.5904 \end{array} \right).
\]

**Problem 2**

We have that

\[
y(t) = -\sum_{n=1}^{na} a_n y(t-n) + \sum_{m=0}^{nc} c_m e(t-m).
\]
It is clear from this that \( y(t) \) is correlated with \( e(t), e(t-1), \ldots, e(t-nc) \) and that if we are looking for \( y(t-k) \)s that are uncorrelated with \( e(t), e(t-1), \ldots, e(t-nc) \) we will need \( y(t-k) \)s with \( k \geq nc + 1 \). To get a solvable equation system for the \( a \)-parameters we will also need a \( z(t) \) vector of dimension \( na \). Choosing \( z(t) = (y(t - nc - 1), \ldots, y(t - nc - na))^T \) we obtain

\[
E(z(t)y(t)) = E \left( \begin{pmatrix} y(t - nc - 1) \\ \vdots \\ y(t - nc - na) \end{pmatrix} y(t) \right) = E \left( \begin{pmatrix} y(t - nc - 1) \\ \vdots \\ y(t - nc - na) \end{pmatrix} \left( - \sum_{n=1}^{na} a_n y(t - n) + \sum_{m=0}^{nc} c_m e(t - m) \right) \right) = E \left( \begin{pmatrix} y(t - nc - 1) \\ \vdots \\ y(t - nc - na) \end{pmatrix} \left( - \sum_{n=1}^{na} a_n y(t - n) \right) \right) = -E \left( \begin{pmatrix} y(t - nc - 1) \\ \vdots \\ y(t - nc - na) \end{pmatrix} (y(t - 1) \cdots y(t - na)) \right) \begin{pmatrix} a_1 \\ \vdots \\ a_{na} \end{pmatrix},
\]

which finally gives the solution for the \( a \)-parameters

\[
\begin{pmatrix} a_1 \\ \vdots \\ a_{na} \end{pmatrix} = - \begin{pmatrix} r_y(nc) & \cdots & r_y(nc-na+1) \\ \vdots & \ddots & \vdots \\ r_y(nc+na-1) & \cdots & r_y(nc) \end{pmatrix}^{-1} \begin{pmatrix} r_y(nc+1) \\ \vdots \\ r_y((nc+na) \end{pmatrix}
\]

where \( r_y(k) = E(y(t) y(t-k)) \). For finite data the true autocorrelation function \( r_y(k) \) is replaced with the approximated autocorrelation \( \hat{r}_y(k) = 1/N \sum_{n=k+1}^{N} y(n)y(n-k) \).

**Problem 3**

In the sliding window formulation of the RLS algorithm \( R(t) \) is given by

\[
R(t) = \sum_{s=t-m+1}^{t} \varphi(s) \varphi(s)^T.
\]

It has the time recursion

\[
R(t) = R(t-1) + \varphi(t) \varphi(t) - \varphi(t-m) \varphi(t-m)^T
\]

\[
= R(t-1) + (\varphi(t) \varphi(t-m)) \begin{pmatrix} \varphi(t) \\ -\varphi(t-m) \end{pmatrix}
\]

Making use of the matrix inversion lemma (page 511 in the book)

\[
[A + BCD]^{-1} = A^{-1} - A^{-1}B \left[C^{-1} + DA^{-1}B\right]^{-1} DA^{-1}
\]
with \( A = R(t-1), B = (\varphi(t) \varphi(t-m)), C = I \) and \( D = (\varphi(t) - \varphi(t-m))^T \) we obtain the requested time recursion for \( P(t) = R^{-1}(t) \)

\[
P(t) = P(t-1) - P(t-1)(\varphi(t) \varphi(t-m)) \left( I + \begin{pmatrix} \varphi(t) \\ -\varphi(t-m) \end{pmatrix} P(t-1)(\varphi(t) \varphi(t-m)) \right)^{-1} \begin{pmatrix} \varphi(t) \\ -\varphi(t-m) \end{pmatrix} P(t-1).
\]

Problem 4
Assuming that we have \( N \) data samples available we define the data vectors

\[
y_k = \begin{pmatrix} y(k-(N-1)) \\ \vdots \\ y(k) \end{pmatrix}, \quad u_k = \begin{pmatrix} u(k-(N-1)) \\ \vdots \\ u(k) \end{pmatrix}.
\]

The model to be identified is then given by

\[
\hat{y}_N(\hat{\theta}) = (-y_{N-1} \ u_{N-1}) \begin{pmatrix} \hat{a} \\ \hat{b} \end{pmatrix}
\]

For this model to be identifiable we need that the two model or basis vectors \(-y_{N-1}\) and \(u_{N-1}\) are linearly independent. With the feedback controller \( u(t) = -fy(t) \) they are linearly dependent \( u_{N-1} = -fy_{N-1} \) and there is no unique solution to the model parameters \( \hat{a} \) and \( \hat{b} \).

Three possible changes to the feedback controller that result in linearly independent model vectors are for example the following.

1. Use two different feedback coefficients, \( f_1 \) over the first \( K \) time samples and \( f_2 \) over the remaining time samples. This results in the model vectors

\[
-y_{N-1} = \begin{pmatrix} y(0) \\ \vdots \\ y(K-1) \\ y(K) \\ \vdots \\ y(N-1) \end{pmatrix}, \quad u_{N-1} = \begin{pmatrix} f_1 y(0) \\ \vdots \\ f_1 y(K-1) \\ f_2 y(K) \\ \vdots \\ f_2 y(N-1) \end{pmatrix},
\]

which are linearly independent for \( f_1 \neq f_2 \). \( K/N \) must not go to zero as \( N \to \infty \).

2. Introduce a timedelay in the feedback controller. Choosing \( u(t) = -fy(t-1) \) results in the model vectors \(-y_{N-1}\) and \(u_{N-1} = -fy_{N-2}\), which are linearly independent.

3. Introduce an extra reference signal \( r(t) \) in the feedback controller. Choosing \( u(t) = -fy(t) + kr(t) \) results in the model vectors \(-y_{N-1}\) and \(u_{N-1} = -fy_{N-1} + kr_{N-1}\), which are linearly independent.
Problem 5
For large $N$ the covariance matrix of $\hat{\theta} = (\hat{a}_1 \hat{a}_2)$ in the model $\hat{y}(t) = \hat{a}_1 y(t - 1) + \hat{a}_2 y(t - 2)$ is given by

$$\text{Cov}(\hat{\theta}_N) = \frac{\lambda^2}{N} \begin{pmatrix} r_y(0) & r_y(1) \\ r_y(1) & r_y(0) \end{pmatrix}^{-1} = \frac{\lambda^2}{N} \begin{pmatrix} r_y(0) & -r_y(1) \\ -r_y(1) & r_y(0) \end{pmatrix}$$

The variances are thus given by

$$\text{var}(\hat{a}_1) = \text{var}(\hat{a}_2) = \frac{\lambda^2}{N r_y(0) \left(1 - \left(1 + a_2\right) r_y(0)/r_y(1)\right)^2}$$

Problem 6 (Alternative to homework assignments)

The model is given by $y(t) = \varphi^T(t)\theta$, with $\varphi^T(t) = [u(t)u(t-1)]$.

a) The covariance matrix for $\hat{\theta}_N$ is given by equation (4.12) in the book as

$$\text{cov}(\hat{\theta}_N) = \sigma_e^2 \left( \sum_{t=1}^{N} \varphi\varphi^T(t) \right)^{-1} = \sigma_e^2 \left( \sum_{t=1}^{N} \begin{pmatrix} u^2(t) & u(t)u(t-1) \\ u(t)u(t-1) & u^2(t-1) \end{pmatrix} \right)^{-1} = \sigma_e^2 \left( \begin{pmatrix} N & N-1 \\ N-1 & N-1 \end{pmatrix} \right)^{-1} = \sigma_e^2 \left( \begin{pmatrix} 1 & -1/N \\ -1 & N-1 \end{pmatrix} \right),$$

which gives the asymptotic variances

$$\text{var}(\hat{b}_0) = \sigma_e^2$$

$$\text{var}(\hat{b}_1) = (1 + 1/N)\sigma_e^2 \rightarrow \sigma_e^2 \text{ as } N \rightarrow \infty$$

b) The variance of the static gain $G(1, \hat{\theta}) = \hat{b}_0 + \hat{b}_1$ is given by

$$\text{var} \left( \hat{b}_0 + \hat{b}_1 \right) = (1 1) \text{cov}(\hat{\theta}_N) \begin{pmatrix} 1 \\ 1 \end{pmatrix} = (1 1) \sigma_e^2 \begin{pmatrix} 1 & -1/N \\ -1 & N-1 \end{pmatrix} \begin{pmatrix} 1 \\ 1 \end{pmatrix} = \frac{\sigma_e^2}{N-1} \rightarrow 0 \text{ as } N \rightarrow \infty$$

A step function is a static signal with all of its energy concentrated at the zero frequency. Therefore it gives persistent excitation at that frequency, which makes the variance of the static gain go to zero as $N \rightarrow \infty$. Note that the variance of $\hat{b}_0 - \hat{b}_1$ does not go to zero as $N \rightarrow \infty$. 5