Solutions to System Identification exam March 9, 2007.

1.

a. The model can be rewritten as

$$y_k = bu_{k-1} + \mu + v_k = \begin{pmatrix} u_{k-1} & 1 \end{pmatrix} \begin{pmatrix} b \\ \mu \end{pmatrix} + v_k = \phi_k \theta + v_k$$

where $v_k = e_k - \mu$ is white noise with expected value $E(v_k) = 0$ and $var(v_k) = \sigma_e^2$.

Collect data into matrices

$$\mathcal{Y} = \Phi\theta + \mathcal{V}$$

where

$$\mathcal{Y} = \begin{pmatrix} y_2 \\ \vdots \\ y_N \end{pmatrix} \qquad \Phi = \begin{pmatrix} u_1 & 1 \\ \vdots & \vdots \\ u_{N-1} & 1 \end{pmatrix}$$

The least-squares estimate is given by

$$\hat{\theta} = (\Phi^T\Phi)^{-1}\Phi^T\mathcal{Y} = \begin{pmatrix} \sum_{k=1}^{N-1} u_k^2 & \sum_{k=1}^{N-1} u_k \\ \sum_{k=1}^{N-1} u_k & \sum_{k=1}^{N-1} 1 \end{pmatrix}^{-1} \begin{pmatrix} \sum_{k=1}^{N-1} u_k y_{k+1} \\ \sum_{k=1}^{N-1} y_{k+1} \end{pmatrix}$$
$$= \begin{pmatrix} 51 & 210.4 \\ 210.4 & 999 \end{pmatrix}^{-1} \begin{pmatrix} 230.6 \\ 1054.9 \end{pmatrix} = \begin{pmatrix} 1.26 \\ 0.79 \end{pmatrix}$$

The parameter estimates are thus $\hat{b} = 1.26$, $\hat{\mu} = 0.79$.

An unbiased estimate of σ_e^2 is given by

$$\hat{\sigma_e^2} = \frac{2}{N-1} V(\hat{\theta})$$

$$V(\hat{\theta}) = \frac{1}{2}\sum_{k=2}^{N} \epsilon_k^2 = \frac{1}{2}\sum_{k=2}^{N} (y_k - \phi_k\hat{\theta})^2 =$$

$$\frac{1}{2}\sum_{k=2}^{N} (y_k^2 + \hat{b}u_{k-1}^2 + \hat{\mu}^2 - 2\hat{b}y_k u_{k-1} - 2\hat{\mu}y_k + 2\hat{b}\hat{\mu}u_{k-1}) = 76.7$$

Thus,

$$\hat{\sigma_e^2} = \frac{2}{999} 76.7 = 0.154$$

b. For consistency in the least-squares estimation algorithm, it must hold that $\lim_{N\to\infty} \frac{1}{N-1}\Phi^T\mathcal{V} = 0$ and that $\lim_{N\to\infty} \frac{1}{N-1}\Phi_N^T\Phi_N$ is invertible.

$$\lim_{N\to\infty} \frac{1}{N-1}\Phi^T\mathcal{V} = \lim_{N\to\infty} \frac{1}{N-1} \begin{pmatrix} \sum_{k=1}^{N-1} u_k v_{k+1} \\ \sum_{k=1}^{N-1} v_{k+1} \end{pmatrix}$$

Since $\{v_k\}$ is white noise, v_{k+1} must be uncorrelated with u_k for all choices of $\{u_k\}$.

$$\lim_{N\to\infty} \frac{1}{N-1}\Phi_N^T\Phi_N = \lim_{N\to\infty} \frac{1}{N-1}\begin{pmatrix} \sum_{k=1}^{N-1} u_k^2 & \sum_{k=1}^{N-1} u_k \\ \sum_{k=1}^{N-1} u_k & \sum_{k=1}^{N-1} 1 \end{pmatrix}$$

This matrix is invertible if its determinant is not zero. This gives the following condition on $\{u_k\}$,

$$\lim_{N\to\infty} \frac{1}{N-1}\sum_{k=1}^{N-1} u_k^2 \neq \left(\lim_{N\to\infty} \frac{1}{N-1}\sum_{k=1}^{N-1} u_k\right)^2$$

c. 1.

$$\lim_{N\to\infty} \frac{1}{N-1}\sum_{k=1}^{N-1} u_k^2 = 25$$

$$(\lim_{N\to\infty} \frac{1}{N-1}\sum_{k=1}^{N-1} u_k)^2 = 25$$

The estimate is not consistent

2. The signal is periodic with period $T = 4$, $\{u_k\} = \{0, 1, 0, -1, 0, \ldots\}$, so that

$$\lim_{N\to\infty} \frac{1}{N-1}\sum_{k=1}^{N-1} u_k^2 = 1/2$$

$$(\lim_{N\to\infty} \frac{1}{N-1}\sum_{k=1}^{N-1} u_k)^2 = 0$$

The estimate is consistent

3. As $N \to \infty$, the effect of the change of u_k at $k = 10$ will vanish, therefore

$$\lim_{N\to\infty} \frac{1}{N-1}\sum_{k=1}^{N-1} u_k^2 = 100$$

$$(\lim_{N\to\infty} \frac{1}{N-1}\sum_{k=1}^{N-1} u_k)^2 = 100$$

The estimate is not consistent

4. The closed-loop system is given by

$$y_k = 0.2by_{k-2} + \mu + e_k$$

Provided that the system is stable ($|0.2b| < 1$), $\{y_k\}$ will be a stochastic process, with expected value μ_y and non-zero variance σ_y. We then have $E(u_k) = 0.2\mu_y$ and $var(u_k) = 0.04\sigma_y^2$.

$$\lim_{N\to\infty} \frac{1}{N-1}\sum_{k=1}^{N-1} u_k^2 = E(u_k^2) = (E(u_k))^2 + var(u_k) = \mu_y^2 + 0.04\sigma_y^2$$

$$(\lim_{N\to\infty} \frac{1}{N-1}\sum_{k=1}^{N-1} u_k)^2 = 0.04\mu_y^2$$

Because $\sigma_y^2 \neq 0$, the estimate is consistent.

2. A PRBS reference signal is generally a good choice for closed-loop identification of a wide range of processes. A constant signal does not give sufficient excitation of the process to provide a good model. A gaussian white noise sequence provides good excitation but there is no limit on the amplitude of the signal, which may not be desirable for all processes. A sequence of sinusoidal signals can give good models at the given frequencies, but if they are not carefully selected there is a risk of not finding important system characteristics, for example resonances. The experiments may also take much time using sinusoidal signals.

3.

a. For a balanced realizaion, the observability Gramian Q is diagonal,

$$Q = \begin{pmatrix} q_1 & 0 \\ 0 & q_2 \end{pmatrix}$$

and can be computed from the Lyapunov equation

$$\Phi^T Q \Phi - Q + C^T C = 0$$

From the Lyapunov equation, we obtain two equations for computing q_1 and q_2

$$\begin{aligned}(\phi_{11}^2 - 1)q_1 + \phi_{12}^2 q_2 &= -c_1^2 \\ \phi_{11}\phi_{12}q_1 + \phi_{12}\phi_{22}q_2 q_2 &= -c_1 c_2\end{aligned}$$

Solving these equations give

$$Q = \begin{pmatrix} 1.37 & 0 \\ 0 & 0.07 \end{pmatrix}$$

Since $q_1 \gg q_2$, a reduced order model can accurately describe the process. The matrices Φ_0, Γ_0, C_0 and D_0 for the reduced order system are given by

$$\begin{aligned}\Phi_0 &= \phi_{11} + \frac{\phi_{12}^2}{1-\phi_{22}} = 0.5939 \\ \Gamma_0 &= \gamma_1 + \frac{\phi_{12}\gamma_2}{1-\phi_{22}} = -0.9406 \\ C_0 &= c_1 + \frac{c_2\phi_{12}}{1-\phi_{22}} = -0.9406 \\ D_0 &= \frac{c_2\gamma_2}{1-\phi_{22}} = 0.0442\end{aligned}$$

The reduced-order model is given by

$$\begin{aligned}z_{k+1} &= 0.5939 z_k - 0.9406 u_k \\ y_k &= -0.9406 z_k + 0.0442 u_k\end{aligned}$$

b. We can find a reduced order model by cancelling the pole in $z = 0.91$ and the zero in $z = 0.9$. The reduced-order model is then given by

$$H_{red1}(z) = \frac{K}{z - 0.5}$$

where K is chosen so that the static gain is preserved,

$$H_{red1}(1) = \frac{K}{0.5} = H(1) = 2.22$$

which gives $K = 1.11$. The reduced-order model from (a) has the transfer function

$$H_{red2}(z) = C_0(zI - \Phi_0)^{-1}\Gamma_0 + D_0 = \frac{0.044(z + 19.4)}{z - 0.5939}$$

The two reduced-order models are similar but not identical. An important difference is that the reduced-order model obtained through balanced realization model reduction has a direct term, even though the original model has not.

4. The input-output data suggests that the process is nonlinear. The gain from u to y is much larger for large u than for small u. We therefore cannot expect to find a single linear system that will provide a good model for the system over the entire operating range.

5.

a. Given a correct estimate θ, the residuals $\epsilon_k = y_k - \phi_k\theta = v_k$ have the distribution $f_v(\epsilon_k)$. The maximum-likelihood estimate is obtained by maximizing the likelihood function, which is defined as the joint probability function of the sequence $\{\epsilon_k\}$.

$$L(\theta) = f_v(\epsilon) = \prod_{k=n+1}^{N} f_v(\epsilon_k) = \prod_{k=n+1}^{N} f_v(y_k - \phi_k\theta)$$

Maximizing the likelihood function is equivalent to maximizing the log-likelihood function

$$\begin{aligned}\log L(\theta) &= \sum_{k=n+1}^{N} \log f_v(y_k - \phi_k\theta) = \sum_{k=n+1}^{N} \log\left(\frac{1}{\sqrt{2}\sigma} e^{-\sqrt{2}|y_k - \phi_k\theta|/\sigma}\right) \\ &= -(N-n)\log(\sqrt{2}\sigma) - \frac{\sqrt{2}}{\sigma}\sum_{k=n+1}^{N} |y_k - \phi_k\theta|\end{aligned}$$

We can see that

$$\max_{\theta} \log L(\theta)$$

is equivalent to

$$\min_{\theta} \sum_{k=n+1}^{N} |y_k - \phi_k\theta|$$

b. The maximum-likelihood estimate of σ is found by solving the optimization problem

$$\max_{\theta,\sigma} \log L(\theta, \sigma)$$

From the expression for $\log L(\theta, \sigma)$, we see that we can find σ by maximizing $f(\sigma) = \log L(\hat{\theta}_{ML}, \sigma) = -(N-n)\log(\sqrt{2}\sigma) - \frac{\sqrt{2}}{\sigma} V(\hat{\theta}_{ML})$.

Differentiating $f(\sigma)$ yields

$$\frac{df}{d\sigma} = \frac{1}{\sigma}\left(\frac{\sqrt{2}}{\sigma} V(\hat{\theta}_{ML}) - (N-n)\right)$$

The maximum of $f(\sigma)$ is given by $\frac{df}{d\sigma} = 0$, which gives

$$\hat{\sigma}_{ML} = \frac{\sqrt{2}}{N-n} V(\hat{\theta}_{ML}) = 0.86$$

6. We can expect to obtain a good model at frequencies where the coherence function γ_{uy} satisfies $\gamma_{uy}^2 \approx 1$, which corresponds to the excitation of the input being much larger than the excitation of the noise. We can see that the coherence function is low around the desired closed-loop bandwidth. The identified model can be improved if we increase the power of the input u at frequencies up to $\omega = 100$.

Estimating $H(e^{i\omega h})$ by dividing the discrete Fourier transforms of u and y has the great disadvantage that the variance of the estimate does not decrease as the number of data increases, i.e. the method is not consistent. To make the method consistent, we can divide the data into a number M of segments, estimate $H(e^{i\omega h})$ for each of these segments, and finally average the estimates.

7. The Markov parameters H_k are given by

$$H_k = CA^{k-1}B = y_k$$

where $\{y_k\}$ is the impulse response sequence.

In the Ho-Kalman algorithm, we find a state-space realization through the singular value decomposition of the matrix

$$\mathcal{H}_{rs}^{(0)} = \begin{pmatrix} H_1 & H_2 & \cdots & H_s \\ H_2 & H_3 & \cdots & H_{s+1} \\ \vdots & \vdots & \ddots & \vdots \\ H_r & H_{r+1} & \cdots & H_{r+s-1} \end{pmatrix}$$

We are given the singular value decomposition of $\mathcal{H}_{33}^{(0)}$, where $\text{rank}(\Sigma) = 2$. We can then find a second order state-space model for the system. Define

the matrices

$$\begin{aligned}
\Sigma_2 &= \text{diag}\{\sigma_1, \sigma_2\} \\
U_2 &= \text{matrix of first 2 columns of } U \\
V_2 &= \text{matrix of first 2 columns of } V \\
E_y &= (\,1 \quad 0 \quad 0\,)^T \\
E_u &= E_y \\
\mathcal{H}_{33}^{(1)} &= \begin{pmatrix} H_2 & H_3 & H_4 \\ H_3 & H_4 & H_5 \\ H_4 & H_5 & H_6 \end{pmatrix}
\end{aligned}$$

The state-space matrices are now given by

$$\begin{aligned}
A_2 &= \Sigma_2^{-1/2} U_2^T \mathcal{H}_{33}^{(1)} V_2 \Sigma_2^{-1/2} = \begin{pmatrix} 1.62 & 0.08 \\ -0.08 & -0.62 \end{pmatrix} \\
B_2 &= \Sigma_2^{1/2} V_2^T E_u = \begin{pmatrix} -1.07 \\ -0.37 \end{pmatrix} \\
C_2 &= E_y^T U_2 \Sigma_2^{1/2} = (\,-1.07 \quad -0.37\,) \\
D &= H_0 = y_0 = 1
\end{aligned}$$

We then have the state space model

$$\begin{aligned}
x_{k+1} &= \begin{pmatrix} 1.62 & 0.08 \\ -0.08 & -0.62 \end{pmatrix} x_k + \begin{pmatrix} -1.07 \\ -0.37 \end{pmatrix} u_k \\
y_k &= (\,-1.07 \quad -0.37\,)\, x_k + u_k
\end{aligned}$$

8.

a. The residual sequence is given by

$$\epsilon_k = y_k + \hat{a} y_{k-1} = (\hat{a} - a) y_{k-1} + e_k + c e_{k-i} = e_k + c e_{k-i}$$

The autocovariance function is defined as

$$\begin{aligned}
C_{\epsilon\epsilon}(\tau) &= E(\epsilon_k \epsilon_{k+\tau}) = E((e_k + c e_{k-i})(e_{k+\tau} + c e_{k-i\tau})) \\
&= E(e_k e_{k+\tau} + c e_k e_{k-i+\tau} + c e_{k-i} e_{k+\tau} + c^2 e_{k-i} e_{k-i+\tau}) \\
&= \begin{cases} (1 + c^2)\sigma_e^2 & \tau = 0 \\ c\sigma_e^2 & \tau = \pm i \\ 0 & \text{otherwise} \end{cases}
\end{aligned}$$

b. The number of zero-crossings is given by

$$\tau = \sum_{i=1}^{N-1} x_i$$

where

$$x_k = \begin{cases} 1 & \text{if } \epsilon_k \epsilon_{k+1} < 0 \\ 0 & \text{if} \epsilon_k \epsilon_{k+1} > 0 \end{cases}$$

Under $\mathcal{H}_0$, $\{\epsilon_k\}$ is white noise, and from the course book we know that τ has the asymptotic distribution

$$\tau \in \mathcal{N}(\frac{N}{2}, \frac{N}{4})$$

A confidence interval for τ given $\mathcal{H}_0$ on the level α is given by

$$\left[\frac{N}{2} - \lambda_{\alpha/2} \sqrt{\frac{N}{4}}, \frac{N}{2} + \lambda_{\alpha/2} \sqrt{\frac{N}{4}} \right]$$

For $N = 5000$, $\alpha = 0.01$, $\lambda_{\alpha/2} = 2.58$, we obtain the confidence interval

$$[2409, 2591]$$

Since the observed number of zero-crossings $\tau = 2642$ is not within this interval, we can reject $\mathcal{H}_0$ on the significance level $\alpha = 0.01$.