

Master - may 08

SYSTEM IDENTIFICATION, Exam of May 2008, Solutions

C2.2

Question 1. (a) We regard the complex number $x+iy$ as being equivalent to the vector $\begin{bmatrix} x \\ y \end{bmatrix} \in \mathbb{R}^2$.

Since a_k and b_k are independent, the probability density of the random vector $\begin{bmatrix} a_k \\ b_k \end{bmatrix}$ is $f(x,y) = f_a(x) \cdot f_b(y)$ where f_a and f_b are the densities of a_k and b_k , respectively. Since a_k and b_k are normalized Gaussian, we have $f_a(x) = f_b(x) = \frac{1}{\sqrt{2\pi}} e^{-\frac{1}{2}x^2}$, so that

$$f(x,y) = \frac{1}{2\pi} e^{-\frac{1}{2}(x^2+y^2)}.$$

(b) Since $e^{i\psi_k} c_k$ is a rotated version of c_k , and the density $f(x,y)$ is invariant under rotation (it only depends on the radius $r = \sqrt{x^2+y^2}$), it follows that $e^{i\psi_k} c_k$ has the same density as c_k (as computed in part (a)). Thus, $e^{i\psi_k} c_k$ is normalized Gaussian. Since a_j and b_j are independent of a_k and b_k (for $j \neq k$), any function of a_j, b_j is independent of any function of a_k, b_k . Thus, the terms of the sequence of random variables $e^{i\psi_k} c_k$ are independent of each other. Thus, by definition, this is normalized Gaussian white noise.

(c) White noise is ergodic. Hence, the averages of the white noise $(e^{i\psi_k} c_k)$ converge (with probability 1) to $E(e^{i\psi_k} c_k) = e^{i\psi_k} E(c_k) = 0$.

(d) The complex conjugate random variable corresponds to the random vector $\begin{bmatrix} a_k \\ -b_k \end{bmatrix}$, which is again normalized Gaussian. Thus, (\bar{c}_k) and also $(e^{i\nu k} \bar{c}_k)$ are normalized white noise signals, so that

$$\lim_{N \rightarrow \infty} \frac{1}{N} \sum_{k=1}^N e^{i\nu k} \bar{c}_k = 0, \text{ with prob. 1.}$$

Adding this to the result from part (c), we obtain

$$\lim_{N \rightarrow \infty} \frac{1}{N} \sum_{k=1}^N e^{i\nu k} a_k = 0, \text{ with prob. 1.}$$

Taking here real and imaginary parts, we obtain the desired statements.

(e) Assume that $w_k = g_0 a_k + g_1 a_{k-1} \dots + g_n a_{k-n}$.

Then

$$\begin{aligned} \lim_{N \rightarrow \infty} \frac{1}{N} \sum_{k=1}^N \cos(\nu k) w_k &= g_0 \lim_{N \rightarrow \infty} \frac{1}{N} \sum_{k=1}^N \cos(\nu k) a_k \\ &+ g_1 \lim_{N \rightarrow \infty} \frac{1}{N} \sum_{k=1}^N \cos(\nu k) a_{k-1} \dots + g_n \lim_{N \rightarrow \infty} \frac{1}{N} \sum_{k=1}^N \cos(\nu k) a_{k-n} \end{aligned}$$

where each limit on the right-hand side is zero.

For $\sin(\nu k)$ in place of $\cos(\nu k)$ the proof is similar.

(f) We compute

$$c = \frac{1}{10^6} \sum_{k=1}^{10^6} u_k \cos(0.01k), \quad s = \frac{1}{10^6} \sum_{k=1}^{10^6} u_k \sin(0.01k).$$

Since $u_k = A \cos(0.01k) \sin \varphi + A \sin(0.01k) \cos \varphi + w_k$, using the statement from (e) we obtain that

$$c \approx \frac{1}{2} A \sin \varphi, \quad s \approx \frac{1}{2} A \cos \varphi.$$

From here we can easily estimate A and φ .

Question 2. (2) $w^2 = \lambda + \frac{u^2}{\alpha^2} + \frac{v^2}{\beta^2}$,

hence

$$\underbrace{w_k^2}_{y_k} = \underbrace{\begin{bmatrix} 1 & u_k^2 & v_k^2 \end{bmatrix}}_{\phi_k} \underbrace{\begin{bmatrix} \lambda \\ 1/\alpha^2 \\ 1/\beta^2 \end{bmatrix}}_{\theta} + e_k.$$

(b) $J(\theta)$ has a unique minimum at $\theta = \hat{\theta}$ if and only if $\phi^* \phi$ is invertible, where $\phi = \begin{bmatrix} \phi_1 \\ \phi_2 \\ \vdots \\ \phi_{200} \end{bmatrix}$. Equivalently, ϕ should have full column rank, i.e., 3 independent columns.

If this is the case, then $\hat{\theta} = \phi^\# y$, where

$$y = \begin{bmatrix} y_1 \\ y_2 \\ \vdots \\ y_{200} \end{bmatrix} \quad \text{and} \quad \phi^\# = (\phi^* \phi)^{-1} \phi^*.$$

(c) If $u_k^2 - v_k^2 = 18$ of the last two columns of ϕ gives 18 times the first column, so that J has no unique minimum. We now have the model

$$\begin{aligned} w_k^2 &= \lambda + \frac{u_k^2}{\alpha^2} + \frac{u_k^2 - 18}{\beta^2} + e_k \\ &= \lambda + u_k^2 \left(\frac{1}{\alpha^2} + \frac{1}{\beta^2} \right) - \frac{18}{\beta^2} + e_k = \begin{bmatrix} 1 & u_k^2 \end{bmatrix} \begin{bmatrix} \lambda - \frac{18}{\beta^2} \\ \frac{1}{\alpha^2} + \frac{1}{\beta^2} \end{bmatrix} + e_k. \end{aligned}$$

From here, we can estimate the two numbers $\lambda - \frac{18}{\beta^2}$ and $\frac{1}{\alpha^2} + \frac{1}{\beta^2}$ in the standard way.

(d) $\widehat{\text{Var}}(e_k) = \frac{1}{197} \|y - \phi \hat{\theta}\|^2 = \frac{1}{197} y^* (I - \phi \phi^\#) y.$ (197 = 200 - 3)

(e) $\widehat{\text{Cov}} \hat{\theta} = \widehat{\text{Var}}(e_k) (\phi^* \phi)^{-1}.$

(f) For N measurements ($k=1,2,\dots,N$) we have

$$\Phi^* \Phi = \begin{bmatrix} N & \sum_{k=1}^N u_k^2 & \sum_{k=1}^N v_k^2 \\ \sum_{k=1}^N u_k^2 & \sum_{k=1}^N u_k^4 & \sum_{k=1}^N u_k^2 v_k^2 \\ \sum_{k=1}^N v_k^2 & \sum_{k=1}^N u_k^2 v_k^2 & \sum_{k=1}^N v_k^4 \end{bmatrix}.$$

Since u_k and v_k are independent white noise signals, they are jointly ergodic. Therefore

$$\Phi^* \Phi \approx N \begin{bmatrix} 1 & E(u_k^2) & E(v_k^2) \\ E(u_k^2) & E(u_k^4) & E(u_k^2 v_k^2) \\ E(v_k^2) & E(u_k^2 v_k^2) & E(v_k^4) \end{bmatrix}.$$

The 3×3 matrix on the right-hand side above is independent of N . Thus, $\Phi^* \Phi$ grows proportionally to N . According to our result at part (e) or, more precisely, because of

$$\text{Cov } \hat{\theta} = \text{Var}(e_k) (\Phi^* \Phi)^{-1},$$

$\text{Cov } \hat{\theta}$ is inverse proportional to N . Thus, for 800 measurements (instead of 200) we expect $\text{Cov } \hat{\theta}$ to be 4 times smaller.

Question 3. (a) Denote $\tilde{\alpha}_k = \alpha_k - E(\alpha_k)$, and similarly for $\tilde{\beta}_k, \tilde{\gamma}_k$, so that $\tilde{\gamma}_k = \tilde{\alpha}_k + \tilde{\beta}_k$. We have $C_{\tau}^{\gamma\gamma} = E(\tilde{\gamma}_k \cdot \tilde{\gamma}_{k-\tau}) = E(\tilde{\alpha}_k \cdot \tilde{\alpha}_{k-\tau}) + E(\tilde{\alpha}_k \tilde{\beta}_{k-\tau}) + E(\tilde{\beta}_k \tilde{\alpha}_{k-\tau}) + E(\tilde{\beta}_k \tilde{\beta}_{k-\tau})$. Since (α_k) and (β_k) are independent signals, the two middle terms are zero and we get $C_{\tau}^{\gamma\gamma} = C_{\tau}^{\alpha\alpha} + C_{\tau}^{\beta\beta}$.

Applying the \mathcal{Z} transformation, $S^{\gamma\gamma} = S^{\alpha\alpha} + S^{\beta\beta}$.

(b) Denote $A(z) = 1 + a_1 z^{-1} \dots + a_4 z^{-4}$, $B(z) = b_0 + b_1 z^{-1} \dots + b_4 z^{-4}$, then by (1) $A(z) \hat{p}(z) = B(z) \hat{u}(z) + \hat{v}(z)$. From $\hat{y} = \hat{p} + \hat{w}$ we get $A \hat{y} = A \hat{p} + A \hat{w} = B \hat{u} + \hat{v} + A \hat{w}$. According to the problem statement, A^{-1} is stable. Denoting $\hat{\delta} = \hat{w} + A^{-1} \hat{v}$, we obtain

$$A(z) \hat{y}(z) = B(z) \hat{u}(z) + A(z) \hat{\delta}(z). \quad (*)$$

According to our result from part (a) we have $S^{\delta\delta} = S^{ww} + |A^{-1}|^2 S^{vv}$. By the problem statement we have $S^{ww} \geq 0.1$, hence $S^{\delta\delta} \geq 0.1$. Since δ is Gaussian, this implies that δ can be represented as $\hat{\delta} = \Xi \hat{e}$, Ξ, Ξ^{-1} stable, e white noise.

Since Ξ is stable, its impulse response (ξ_k) tends to zero and we can approximate Ξ by truncating its impulse response:

$$\Xi(z) \approx 1 + \xi_1 z^{-1} + \xi_2 z^{-2} \dots + \xi_n z^{-n} = \Xi_n(z).$$

The coefficient ξ_0 has been taken $= 1$, which is possible by rescaling e . Now $(*)$ becomes

$$A(z) \hat{y}(z) = B(z) \hat{u}(z) + A(z) \Xi_n(z) \hat{e}(z),$$

which is the desired ARMAX model. — 5 —

(c) Denoting $C(z) = A(z) \Xi_n(z)$, the ARMAX equation from part (b) is $A \hat{y} = B \hat{u} + C \hat{e}$. By assumption, A^{-1} is stable. Since Ξ^{-1} from part (b) is stable, we may assume that also Ξ_n^{-1} is stable. This implies that C^{-1} is stable. Divide the ARMAX equation by C :
 $(A/C) \hat{y} = (B/C) \hat{u} + \hat{e}$, and introduce the impulse responses of A/C and B/C :

$$\frac{A(z)}{C(z)} = 1 + \alpha_1 z^{-1} + \alpha_2 z^{-2} + \dots$$

$$\frac{B(z)}{C(z)} = \beta_0 + \beta_1 z^{-1} + \beta_2 z^{-2} + \dots$$

Since C^{-1} is stable, the sequences (α_k) and (β_k) tend to zero. Hence, by truncating A/C and B/C to polynomials (in z^{-1}) of a high order m , we get good approximations of these functions, and the approximate ARX model

$$y_k + \alpha_1 y_{k-1} + \alpha_2 y_{k-2} \dots + \alpha_m y_{k-m} = \beta_0 u_k + \beta_1 u_{k-1} \dots + \beta_m u_{k-m} + e_k.$$

(d) We have

$$y_k = \underbrace{\begin{bmatrix} -y_{k-1} & -y_{k-2} & \dots & -y_{k-m} & u_k & u_{k-1} & \dots & u_{k-m} \end{bmatrix}}_{\varphi_k} \underbrace{\begin{bmatrix} \alpha_1 \\ \alpha_2 \\ \vdots \\ \alpha_m \\ \beta_0 \\ \beta_1 \\ \vdots \\ \beta_m \end{bmatrix}}_{\theta} + e_k.$$

Denoting $y = \begin{bmatrix} y_1 \\ y_2 \\ \vdots \\ y_{20,000} \end{bmatrix}$, $\phi = \begin{bmatrix} \varphi_1 \\ \varphi_2 \\ \vdots \\ \varphi_{20,000} \end{bmatrix}$,

$$\phi^\# = (\phi^* \phi)^{-1} \phi^*$$

the optimal least squares estimate of θ is $\hat{\theta} = \phi^\# y$.

(e) After having estimated θ from part (d), we can estimate (e_k) using the ARX equation. Now we rewrite the ARMAX equation from the top of this page as $y_k = \tilde{\varphi}_k \tilde{\theta} + \delta_k$, where

$$\tilde{\varphi}_k = [-y_{k-1} \ -y_{k-2} \ \dots \ -y_{k-4} \ u_k \ u_{k-1} \ \dots \ u_{k-4} \ e_{k-1} \ e_{k-2} \ \dots \ e_{k-q}],$$

$$\tilde{\theta}^T = [a_1 \ a_2 \ \dots \ a_4 \ b_0 \ b_1 \ \dots \ b_4 \ c_1 \ c_2 \ \dots \ c_q],$$

and $\delta_k = e_k + \text{new modeling error}$. From here we can estimate $\tilde{\theta}$ in the usual way ($q = n+4$). —6—

Question 4. (a) The impedance Z of the three components in parallel is given by

$$\frac{1}{Z(s)} = Cs + \frac{1}{R} + \frac{1}{L_1 s} = \frac{CRL_1 s^2 + L_1 s + R}{RL_1 s},$$

so that
$$Z(s) = \frac{RL_1 s}{CRL_1 s^2 + L_1 s + R}.$$

The transfer function from u to y is

$$G(s) = \frac{Z(s)}{Z(s) + Ls} = \frac{RL_1 s}{RL_1 s + Ls(CRL_1 s^2 + L_1 s + R)}$$

$$= \frac{\frac{1}{LC}}{s^2 + \frac{1}{RC}s + \frac{L+L_1}{LCL_1}} = \frac{b_0}{s^2 + a_1 s + a_0}.$$

Note that b_0 is known, while a_1, a_0 are unknown. G is stable, because a_1 and a_0 are > 0 .

(b) The sum of the two inductor voltages is u .

If u is a positive constant, then the inductor currents will grow to infinity, since L times the derivative of the current through the inductor L is the voltage of this inductor. Hence, the system is unstable. (More precisely, zero is an eigenvalue of the system, and the corresponding eigenvector is unobservable.)

(c) We denote by $G^e(i\omega_k)$ the values of the transfer function determined using a sinusoidal signal u (here, $k=1, 2, \dots, 25$). We have

$$b_0 = [(i\omega_k)^2 + a_1(i\omega_k) + a_0] G^e(i\omega_k) - e_k,$$

where e_k are the equation errors (due to measurement errors and model mismatch). - 7 -

Thus,
$$\underbrace{(i\omega_k)^2 G^e(i\omega_k) - b_0}_{y_k} = \underbrace{[-i\omega_k \ -1]}_{\varphi_k} G^e(i\omega_k) \underbrace{\begin{bmatrix} a_1 \\ a_0 \end{bmatrix}}_{\theta} + e_k.$$

(d) We are searching for the optimal real θ . We put $\tilde{y}_k = \text{Re } y_k$, $\tilde{\varphi}_k = \text{Re } \varphi_k$ for $k=1, 2, \dots, 25$, and $\tilde{y}_k = \text{Im } y_{k-25}$, $\tilde{\varphi}_k = \text{Im } \varphi_{k-25}$ for $k=26, 27, \dots, 50$. The new error terms \tilde{e}_k ($k=1, 2, \dots, 50$) are defined similarly. Then $\tilde{y}_k = \tilde{\varphi}_k \theta + \tilde{e}_k$ for $k=1, 2, \dots, 50$, and $\sum_{k=1}^{50} \tilde{e}_k^2 = \sum_{k=1}^{25} |e_k|^2$. The optimal θ (which minimizes $\sum_{k=1}^{50} \tilde{e}_k^2$) is given by $\hat{\theta} = \tilde{\Phi}^\# \tilde{y}$, where $\tilde{y} = [\tilde{y}_1, \dots, \tilde{y}_{50}]^T$, $\tilde{\Phi} = \begin{bmatrix} \tilde{\varphi}_1 \\ \vdots \\ \tilde{\varphi}_{50} \end{bmatrix}$, $\tilde{\Phi}^\# = (\tilde{\Phi}^* \tilde{\Phi})^{-1} \tilde{\Phi}^*$. From the estimated a_1 we estimate R , and then (from a_0) L_1 .

(e)
$$A = \begin{bmatrix} 0 & 1 \\ -a_0 & -a_1 \end{bmatrix}, \quad B = \begin{bmatrix} 0 \\ 1 \end{bmatrix}, \quad C = [b_0 \ 0], \quad D = 0.$$

(f)
$$A^d = e^{AT}, \quad B^d = (e^{AT} - I)A^{-1}B,$$

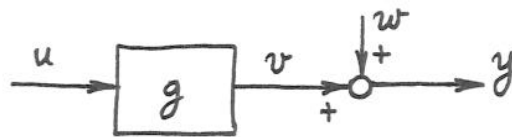
$$G^d(z) = C(zI - A^d)^{-1}B^d + D$$

(this is exact discretisation). Alternatively, we get a good approximation to G^d by Tustin's formula:

$$G^d(z) \approx G\left(\frac{2}{T} \cdot \frac{z-1}{z+1}\right),$$

valid if the poles of G are much smaller than absolute values of the $2\pi/T$, the sampling frequency in rad/sec. In the specific example, G is stable hence also G^d is stable.

Question 5.



- (a) If u and y are jointly ergodic, then the expectation of any function of u and y (which may depend on current and past values) can be approximated by averaging over a long time. Thus, for example, $E(u_k) = \stackrel{\text{a.s.}}{=} \lim_{N \rightarrow \infty} \frac{1}{N} \sum_{j=1}^N u_j$, where the abbreviation a.s. ("almost sure") means that the equality holds with probability 1. A similar formula holds for $E(y_k)$, obviously. Denote

$$\tilde{u}_k = u_k - E(u_k), \quad \tilde{y}_k = y_k - E(y_k),$$

then for any $\tau \in \mathbb{Z}$, ergodicity implies

$$C_{\tau}^{uu} = E(\tilde{u}_k \cdot \tilde{u}_{k-\tau}) \stackrel{\text{a.s.}}{=} \lim_{N \rightarrow \infty} \frac{1}{N} \sum_{j=1}^N \tilde{u}_j \tilde{u}_{j-\tau},$$

$$C_{\tau}^{yu} = E(\tilde{y}_k \cdot \tilde{u}_{k-\tau}) \stackrel{\text{a.s.}}{=} \lim_{N \rightarrow \infty} \frac{1}{N} \sum_{j=1}^N \tilde{y}_j \tilde{u}_{j-\tau}.$$

In practice, we have only finitely many data, so that in all the above formulas, we have to replace $\lim_{N \rightarrow \infty} \frac{1}{N} \sum_{j=1}^N$ with $\frac{1}{N} \sum_{j=a}^{a+N}$, where N is large (and the starting time a depends on the data that we have). In our specific case, when u_k and y_k are given for $k = 1, 2, \dots, 6000$ and $\tau = 0, 1, \dots, 30$, we approximate

$$C_{\tau}^{uu} \approx \frac{1}{6000 - \tau - 1} \sum_{j=\tau+1}^{6000} (u_j - \bar{u})(u_{j-\tau} - \bar{u}),$$

where \bar{u} is the average of all available u_j (so that $\bar{u} \approx E(u_k)$). A similar approximation can be used for C_{τ}^{yu} .

$$\begin{aligned} (b) \quad C_{\tau}^{yu} &= E(\tilde{y}_k \cdot \tilde{u}_{k-\tau}) = E(\tilde{v}_k \cdot \tilde{u}_{k-\tau}) + E(\tilde{w}_k \cdot \tilde{u}_{k-\tau}) = \\ &= C_{\tau}^{vu} + C_{\tau}^{wu}, \text{ so that } C^{yu} = C^{vu} + C^{wu} \\ &= g * C^{uu} + C^{wu}. \end{aligned}$$

- (c) If u and w are independent of each other, then $C^{wu} = 0$, so that (according to the result from part (b)), $C^{yu} = g * C^{uu}$. This can be written as an infinite matrix equation:

$$\begin{bmatrix} C_{0}^{uu} & C_{-1}^{uu} & C_{-2}^{uu} & \dots \\ C_{1}^{uu} & C_{0}^{uu} & C_{-1}^{uu} & \dots \\ C_{2}^{uu} & C_{1}^{uu} & C_{0}^{uu} & \dots \\ \vdots & \vdots & \vdots & \ddots \end{bmatrix} \cdot \begin{bmatrix} g_0 \\ g_1 \\ g_2 \\ \vdots \end{bmatrix} = \begin{bmatrix} C_{0}^{yu} \\ C_{1}^{yu} \\ C_{2}^{yu} \\ \vdots \end{bmatrix}. \quad (**)$$

Since $g_k \rightarrow 0$ (by stability), we can approximate $g_k \approx 0$ for $k > 30$. Looking only at the first 31 equations, we now get 31 equations with 31 unknowns g_0, g_1, \dots, g_{30} .

The coefficients C_{τ}^{uu} and C_{τ}^{yu} are not known exactly, but they have been estimated in (a). Recall that $C_{-\tau}^{uu} = C_{\tau}^{uu}$.

- (d) u is persistent of order N if the $N \times N$ truncation of the infinite matrix from (**) is invertible. If this is the case, and the coefficients in the equation have been estimated sufficiently accurately, then we can solve the truncated equation for g_0, g_1, \dots, g_{N-1} .

The matrix from (**) is ≥ 0 , hence any $N \times N$ truncation of it is also ≥ 0 . A matrix $P \geq 0$ is invertible if and only if $P > 0$, i.e., $x^* P x > 0$ for any vector $x \neq 0$ of matching dimension. This implies that if $P > 0$ and we truncate P , keeping definition of $P > 0$ only its first m rows and first m columns, then the truncated matrix is again > 0 (hence, invertible).

- (e) $\hat{u}(z) = (1 - 0.3z^{-1})\hat{e}(z)$, $S^{uu}(z) = |1 - 0.3z^{-1}|^2$, it is easy to see that $|1 - 0.3z^{-1}| \geq 0.7$ for all z with $|z| = 1$, hence the claim.

- (f) The difference equation of the FIR filter is
$$v_k = g_0 u_k + g_1 u_{k-1} + g_2 u_{k-2} + \dots + g_{N-1} u_{k-N+1}.$$
 (u=input,
v=output)