IMPERIAL COLLEGE LONDON

E4.27
C2.3
ISE4.41

DEPARTMENT OF ELECTRICAL AND ELECTRONIC ENGINEERING
EXAMINATIONS 2005

MSc and EEE/ISE PART IV: MEng and ACGI

SYSTEM IDENTIFICATION

Friday, 13 May 10:00 am

Time allowed: 3:00 hours

Corrected Copy

There are FIVE questions on this paper.

Answer THREE questions.

All questions carry equal marks

Any special instructions for invigilators and information for candidates are on page 1.

Examiners responsible First Marker(s) : G. Weiss
Second Marker(s) : J.C. Allwright

Special information for invigilators: none

Information for candidates:

$$C(\tau) = E[(u(t)-\mu)(u(t+\tau)-\mu)]$$

$$f(x) = \frac{1}{\sqrt{2\pi}\,\sigma} e^{-\frac{x^2}{2\sigma^2}} \qquad S_{yy} = |G|^2 S_{uu}$$

$$Z_L = sL \qquad Z_c = \frac{1}{Cs}$$

$$\Phi^{\#} = (\Phi^*\Phi)^{-1}\Phi^* \qquad P = \Phi\Phi^{\#} \qquad S = \frac{1}{N-p}\|y - \Phi\widehat{\theta}\|^2$$

$$A^d = e^{Ah} \qquad B^d = \left(e^{Ah} - I\right)A^{-1}B \qquad G^d(z) \approx G\left(\frac{2}{h}\,\frac{z-1}{z+1}\right) \qquad G(s) \approx G^d\left(\frac{1+sh/2}{1-sh/2}\right)$$

$$C_k^{uu} g_0 + C_{k-1}^{uu} g_1 + C_{k-2}^{uu} g_2 + \ldots = C_k^{uy}$$

$$\mathrm{Cov}(X,Y) = E[(X-\mu_X)(Y-\mu_Y)]$$

$$E(X \cdot Y) = E(X)\cdot E(Y) + \mathrm{Cov}(X,Y)$$

$$\mathrm{Cov}(TX) = T\,\mathrm{Cov}(X)T^*$$

$$\widehat{v}(z) = \sum_{k=0}^{\infty} v_k z^{-k}$$

$$[(\Delta v)_k = v_{k+1}] \quad \Rightarrow \quad \Delta v\,(z) = z[\widehat{v}(z) - v_0]$$

$$[u_k = k v_k] \quad \Rightarrow \quad \widehat{u}(z) = -z\frac{d}{dz}\widehat{v}(z)$$

$$[v_k = \sin k\nu] \quad \Rightarrow \quad \widehat{v}(z) = \frac{z\sin\nu}{(z-e^{i\nu})(z-e^{-i\nu})}$$

$$\left[v_k = \rho^k\right] \quad \Rightarrow \quad \widehat{v}(z) = \frac{z}{z-\rho}$$

$$\left[v_k = \frac{1}{\rho} k \rho^k\right] \quad \Rightarrow \quad \widehat{v}(z) = \frac{z}{(z-\rho)^2}$$

$$P_n = \frac{1}{\lambda}\left[P_{n-1} - \frac{P_{n-1}\varphi_n^*\varphi_n P_{n-1}}{\lambda + \varphi_n P_{n-1}\varphi_n^*}\right]$$

$$\varepsilon_n = y_n - \varphi_n \hat{\theta}_{n-1}$$

$$\hat{\theta}_n = \hat{\theta}_{n-1} + P_n \varphi_n^* \varepsilon_n$$

$$\begin{aligned} y_k + a_1 y_{k-1} \cdots + a_n y_{k-n} &= b_0 u_k + b_1 u_{k-1} \cdots + b_n u_{k-n} \\ &\quad + e_k + c_1 e_{k-1} \cdots + c_n e_{k-n} \end{aligned}$$

$$C(z) = 1 + c_1 z^{-1} \ldots + c_n z^{-n}$$

$$\hat{u}^F = C^{-1}\hat{u}, \qquad \hat{y}^F = C^{-1}\hat{y}$$

$$\begin{aligned} \overline{y_k} &= (c_1 - a_1) y_{k-1}^F + (c_2 - a_2) y_{k-2}^F \cdots + (c_n - a_n) y_{k-n}^F \\ &\quad + b_0 u_k^F + b_1 u_{k-1}^F \cdots + b_n u_{k-n}^F \end{aligned}$$

1. A random signal (u_k) has the structure

$$u_k = A\sin(0.01k + \varphi) + w_k\,, \qquad k \in \mathbb{Z},$$

where $A > 0$ and $\varphi \in (-\pi, \pi]$ are unknown and (w_k) is a stationary ergodic Gaussian random signal. The measurements u_k are known for $k = 1, 2, 3, \ldots 10^7$. Our aim is to estimate $A, \varphi, E(w_k)$ and $\mathrm{Var}(w_k)$. In the first five parts below you are asked to prove certain statements. If you do not succeed to prove one of these statements, you can still use this statement to answer the other parts.

(a) Assume that (a_k) and (b_k) are independent normalized Gaussian white noise signals. Define the complex random sequence c_k by $c_k = \frac{1}{\sqrt{2}}(a_k + ib_k)$. Show that c_k is normalized Gaussian (we regard the complex plane as being equivalent to $\mathbb{R}^2$). [2]

(b) Let (ψ_k) be an arbitrary sequence of real numbers ($k \in \mathbb{Z}$) and let c_k be the random sequence from part (a). Show that the sequence $(e^{i\psi_k} c_k)$ is a normalized Gaussian white noise signal. [2]

(c) Show that for every $\nu \in \mathbb{R}$ we have

$$\lim_{N\to\infty} \frac{1}{N} \sum_{k=1}^{N} e^{i\nu k} c_k = 0,$$

with probability 1. Here, (c_k) is the sequence from part (a). [3]

(d) Let (a_k) be the sequence introduced in part (a). Show that we have

$$\lim_{N\to\infty} \frac{1}{N} \sum_{k=1}^{N} \cos(\nu k) a_k = 0, \qquad \lim_{N\to\infty} \frac{1}{N} \sum_{k=1}^{N} \sin(\nu k) a_k = 0,$$

with probability 1. Hint: justify and then use the fact that the result from part (c) is true also for the complex conjugate sequence $(\overline{c_k})$. [2]

(e) Let (a_k) be the sequence introduced in part (a). Assume for simplicity that (w_k) can be obtained from (a_k) by filtering it through a FIR filter with impulse response (g_k). Show that the two formulas from part (d) remain valid with w_k in place of a_k. [3]

(f) The statement from part (e) is true for $\nu = 0.01$ even without assuming that (g_k) is FIR. Using this fact, propose a method to estimate A and φ. Hint: think of the identification method that uses sinusoidal inputs to estimate the values of the frequency response function. [4]

(g) Propose a method to estimate $E(w_k)$ and $\mathrm{Var}(w_k)$. Hint: use the estimates for A and φ. [4]

2. The proposed mathematical model of a static system with two inputs, u and v, and with one output w is

$$w = \ln\left[\lambda + \left(\frac{u}{\alpha}\right)^2 + \left(\frac{v}{\beta}\right)^2\right]. \tag{1}$$

The variables u, v and w can be measured and α, β, λ are unknown positive parameters. We have 100 measurements available from experiments, $u_1, u_2, \ldots u_{100}$ and similarly for v and w. Because of measurement and modeling errors, the measurements do not fit any model of the form (1) exactly.

(a) By defining new variables if necessary, rewrite the model of the system in the form $y_k = \varphi_k\theta + e_k$, where y_k and φ_k are known, θ is the vector of unknown parameters and e_k are the equation errors. [3]

(b) State the condition under which a unique minimizing $\hat{\theta}$ exists for the cost $J(\theta) = e_1^2 + e_2^2 \ldots + e_{100}^2$. Assuming that this condition is satisfied, write the formula for the vector of estimated parameters $\hat{\theta}$ that minimizes $J(\theta)$. [3]

(c) Suppose our data are such that $u_k^2 + v_k^2 = 13$ for all $k \in \mathbb{N}$. Explain why in this case we cannot estimate α and β, but we can still estimate $\frac{1}{\alpha^2} - \frac{1}{\beta^2}$. Explain how to estimate $\frac{1}{\alpha^2} - \frac{1}{\beta^2}$. [4]

In the sequel, we assume that there is a unique minimizing $\hat{\theta}$ for the cost function $J(\theta)$ from part (b).

(d) Assume that e_k are independent and identically distributed random variables with $E(e_k) = 0$. Give a formula for an unbiased estimate of $Var(e_k)$ in terms of the values of φ_k and y_k from part (a). [3]

(e) Still assuming independent and identically distributed equation errors, give a formula for an unbiased estimate of $Cov(\hat{\theta})$, where $\hat{\theta}$ is the estimate from part (b). Note that $Var(e_k)$ is not known, but it can be estimated, as was required in part (d). [3]

(f) Suppose that each of the sequences u_k, v_k and e_k consists of independent and identically distributed random variables, and the three sequences are also independent of each other. Let $\hat{\theta}$ be the estimate from part (b). If, instead of 100 measurements, we have 500 measurements, approximately how many times do you expect $Cov(\hat{\theta})$ to decrease? Give, briefly, a reason for your answer. [4]

3. In this question, Σ is an unknown stable discrete-time LTI system with input signal u and output signal p. The measured output y is corrupted by the noise signal w, so that $y_k = p_k + w_k$. It is known that w is a stationary ergodic Gaussian signal with $E(w_k) = 0$ and the power spectral density of w, denoted by S^{ww}, satisfies $S^{ww}(e^{i\nu}) \geq 0.1$ for all $\nu \in (-\pi, \pi]$ (but otherwise S^{ww} is not known).

We have to identify Σ, based on the measurements of u_k and y_k. We would like to model Σ by a discrete-time transfer function of order 4:

$$p_k + a_1 p_{k-1} \dots + a_4 p_{k-4} = b_0 u_k + b_1 u_{k-1} \dots + b_4 u_{k-4} + v_k\,, \qquad (1)$$

where v_k is the equation error due to model mismatch. We assume that v is a stationary ergodic Gaussian random signal with $E(v_k) = 0$, and v is independent of w (i.e., v_k and w_j are independent for all integers k, j).

(a) Let α and β be two independent stationary random signals and $\gamma_k = \alpha_k + \beta_k$. How are the power spectral densities of α, β, γ related to each other? Give a short proof of your formula. [3]

(b) Consider the system with input u and output y. Describe this system by an ARMAX model with a white noise input denoted e. For this, introduce a new signal δ that accounts for the combined effect of w and v. Then represent δ as filtered white noise, where both the filter and its inverse are stable. Finally, approximate the filter by a FIR filter. Briefly, why is it possible to represent δ as described above, and why is it possible to approximate the filter by a FIR filter?

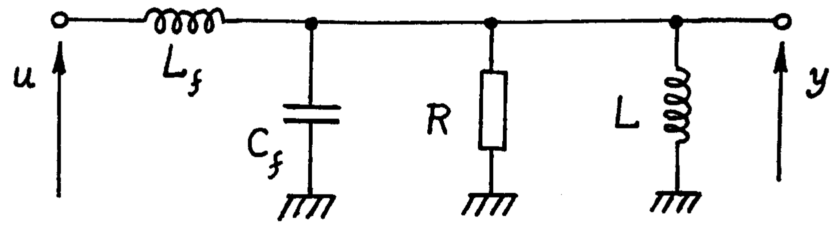
Hint: use the result from part (a). [5]

(c) Why is the ARMAX model of part (b) equivalent to an ARX model of very high order? Why do we need here that the inverse of the filter from (b) is stable? [4]

(d) Assuming that the measurements u_k and y_k are available for $k = 1, 2, \dots 20,000$, describe a least squares based method for estimating the unknown coefficients in the ARX model from part (c). [4]

(e) Use pseudolinear regression to explain how the unknown coefficients of the ARMAX model of part (b) can be estimated using the estimated coefficients of the ARX model of part (d). [4]

4. We want to model the output circuit of an inverter by the simplified circuit shown below, where the filter inductor $L_f > 0$ and the filter capacitor $C_f > 0$ are known, while the load resistor $R > 0$ and the load inductance $L > 0$ are unknown and should be estimated. We can choose the waveform of u, the output voltage of the inverter, and we can measure the load voltage y. We cannot expect a perfect match between the true circuit and this simplified circuit, but we would like to get a close match in a certain frequency range.



(a) Compute the transfer function $\mathbf{G}$ of the simplified circuit (from u to y), in terms of L_f, C_f, R and L. Is $\mathbf{G}$ stable? [4]

(b) Evaluate the order of a state space model for the simplified circuit. Briefly state your reasoning, but there is no need to actually construct such a state space model. What is the order of the transfer function $\mathbf{G}$ of part (a)? If there is a discrepancy between these orders, explain why. Hint: find an unobservable equilibrium state. [3]

(c) Suppose that by measurements using sinusoidal u, we have obtained estimates for $\mathbf{G}$ at 25 angular frequencies $\omega_1, \ldots \omega_{25}$, in the frequency range of interest. By defining new variables, rewrite the model of the system in the form $y_k = \varphi_k \theta + e_k$, where y_k and φ_k are known (possibly complex), θ is the vector of unknown parameters and e_k are the equation errors (possibly complex). Hint: think carefully about what is known and what has to be estimated. [3]

(d) For the model constructed in part (c), explain how to find the *real* vector $\hat{\theta}$ which minimizes $J(\theta) = \sum_{k=1}^{25} |e_k|^2$. Explain how we can estimate R and L using $\hat{\theta}$. [4]

(e) Construct a minimal realization of $\mathbf{G}$ (from part (a)), of the form $\dot{x} = Ax + Bu$, $y = Cx + Du$, where A, B, C and D are matrices. [3]

(f) We connect a hold device (D/A converter) at the input of our system and we connect a sampler (A/D converter) at its output, both converters working with the sampling period T. How can we compute the transfer function of the resulting discrete-time LTI system? Will this discrete-time system be stable? There is no need to perform any computations to answer this part. [3]

5. We want to estimate the impulse response g of a stable discrete-time LTI plant but we do not have the possibility to apply input signals of our choice, we can only observe the existing signals. The input signal is denoted by $u = (u_k)$ and it can be measured. The output signal v is corrupted by a noise signal w, such that the measured output signal is given by $y_k = v_k + w_k$. Both u and w are assumed to be stationary and ergodic (but not necessarily independent of each other). The measurements of u_k and y_k are available for $k = 1, 2, 3, \ldots 6000$.

(a) Describe a method for estimating the auto-correlation function C_τ^{uu} and the cross-correlation function C_τ^{yu} for $0 \leq \tau \leq 30$. Explain very briefly how this problem is related to the concept of ergodicity. [3]

(b) Express C^{yu} in terms of C^{uu}, C^{wu} and g. [4]

(c) Assume now that u_k and w_j are independent of each other, for all $k, j \in \mathbb{Z}$. Describe a method for estimating the terms $g_0, g_1, g_2, \ldots g_{30}$ from the results of part (a). Show briefly how this method can be derived from your answer to part (b). [4]

(d) What is the meaning of a random signal being "persistent of order N"? What is the significance of this concept in the context of part (c) above? Explain the following: if u is persistent of order 30, then it is also persistent of order 20. [3]

(e) If $e = (e_k)$ is normalized white noise and $u_k = e_k - 0.3e_{k-1}$ (for all $k \in \mathbb{Z}$), show that u is persistent of any order. [3]

(f) After having estimated the first N terms of the impulse response, $g_0, g_1, \ldots g_{N-1}$, how can we build a FIR filter whose transfer function is a good approximation to the true transfer function? Write the corresponding difference equation. [3]

[END]

SYSTEM IDENTIFICATION, Exam of May 2005, Solutions

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 $(e^{i\nu k} c_k)$ converge (with probability 1) to $E(e^{i\nu k} c_k) = e^{i\nu k} E(c_k) = 0$.

(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

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

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

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

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

$$\lim_{N\to\infty} \frac{1}{N}\sum_{k=1}^{N} \cos(\nu k)\, w_k = g_0 \lim_{N\to\infty} \frac{1}{N}\sum_{k=1}^{N} \cos(\nu k)\, a_k$$

$$+ g_1 \lim_{N\to\infty} \frac{1}{N}\sum_{k=1}^{N} \cos(\nu k)\, a_{k-1} \dots + g_n \lim_{N\to\infty} \frac{1}{N}\sum_{k=1}^{N} \cos(\nu k)\, a_{k-n}$$

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

$$c = \frac{1}{10^7}\sum_{k=1}^{10^7} u_k \cos(0.01k), \quad s = \frac{1}{10^7}\sum_{k=1}^{10^7} u_k \sin(0.01k).$$

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.

(g) Subtract $\hat{A}\sin(0.01k + \hat{\varphi})$ from u_k ($\hat{A}$ and $\hat{\varphi}$ are the estimates from part (f)), then use ergodicity to estimate $E(w_k)$ and $\mathrm{Var}(w_k)$ by averaging.

Question 2.

(a) $e^{w} = \lambda + \dfrac{u^2}{\alpha^2} + \dfrac{v^2}{\beta^2}$,

hence

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

(b) $J(\theta)$ has a unique minimum at $\theta = \hat{\theta}$ if and only if $\Phi^*\Phi$ is invertible, where $\Phi = \begin{bmatrix} \varphi_1 \\ \varphi_2 \\ \vdots \\ \varphi_{100} \end{bmatrix}$. Equivalently, Φ should have full column rank, i.e., 3 independent columns.

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

$$y = \begin{bmatrix} y_1 \\ y_2 \\ \vdots \\ y_{100} \end{bmatrix} \quad \text{and} \quad \Phi^{\#} = (\Phi^*\Phi)^{-1}\Phi^* .$$

(c) If $u_k^2 + v_k^2 = 13$ then the sum of the last two columns of Φ gives 13 times the first column, so that J has no unique minimum. We now have the model

$$e^{w_k} = \lambda + \frac{u_k^2}{\alpha^2} + \frac{13 - u_k^2}{\beta^2} + e_k$$

$$= \lambda + u_k^2\left(\frac{1}{\alpha^2} - \frac{1}{\beta^2}\right) + \frac{13}{\beta^2} + e_k = \begin{bmatrix} 1 & u_k^2 \end{bmatrix} \begin{bmatrix} \lambda + \frac{13}{\beta^2} \\ \frac{1}{\alpha^2} - \frac{1}{\beta^2} \end{bmatrix} + e_k .$$

From here, we can estimate the two numbers $\lambda + \frac{13}{\beta^2}$ and $\frac{1}{\alpha^2} - \frac{1}{\beta^2}$ in the standard way.

(d) $$\widehat{\mathrm{Var}(e_k)} = \frac{1}{97}\| y - \Phi\hat{\theta} \|^2 = \frac{1}{97} y^*(I - \Phi\Phi^{\#})y .$$

$(97 = 100 - 3)$

(e) $$\widehat{\mathrm{Cov}\,\hat{\theta}} = \widehat{\mathrm{Var}(e_k)}\,(\Phi^*\Phi)^{-1} .$$

(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

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

$\operatorname{Cov}\hat{\theta}$ is inverse proportional to N. Thus, for 500 measurements (instead of 100) we expect $\operatorname{Cov}\hat{\theta}$ to be 5 times smaller.

Question 3. (a) Denote $\mathring{\alpha}_k = \alpha_k - E(\alpha_k)$, and similarly for $\mathring{\beta}_k, \mathring{\gamma}_k$, so that $\mathring{\gamma}_k = \mathring{\alpha}_k + \mathring{\beta}_k$. We have $C_\tau^{\gamma\gamma} = E(\mathring{\gamma}_k \cdot \mathring{\gamma}_{k-\tau}) = E(\mathring{\alpha}_k \cdot \mathring{\alpha}_{k-\tau}) + E(\mathring{\alpha}_k \mathring{\beta}_{k-\tau}) + E(\mathring{\beta}_k \cdot \mathring{\alpha}_{k-\tau}) + E(\mathring{\beta}_k \mathring{\beta}_{k-\tau})$. 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

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

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

$$\hat{\delta} = \Xi\,\hat{e}, \quad \Xi, \Xi^{-1} \text{ stable}, \quad e \text{ white noise}.$$

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.

(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

$$y = \begin{bmatrix} y_1 \\ y_2 \\ \vdots \\ y_{20{,}000} \end{bmatrix}, \quad \phi = \begin{bmatrix} \varphi_1 \\ \varphi_2 \\ \vdots \\ \varphi_{20{,}000} \end{bmatrix}, \qquad \boxed{\phi^{\#} = (\phi^* \phi)^{-1} \phi^*}$$

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

$$\tilde{\varphi}_k = \begin{bmatrix} -y_{k-1} & -y_{k-2} & \dots & -y_{k-n} & u_k & u_{k-1} & \dots & u_{k-n} & e_{k-1} & e_{k-2} & \dots & e_{k-n} \end{bmatrix},$$

$$\tilde{\theta}^T = \begin{bmatrix} a_1 & a_2 & \dots & a_n & b_0 & b_1 & \dots & b_n & c_1 & c_2 & \dots & c_n \end{bmatrix},$$

and $\delta_k = e_k +$ new modeling error. From here we can estimate $\tilde{\theta}$ in the usual way.

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

$$\frac{1}{Z(s)} = C_f s + \frac{1}{R} + \frac{1}{Ls} = \frac{C_f R L s^2 + Ls + R}{RLs},$$

so that
$$Z(s) = \frac{RLs}{C_f RL\, s^2 + Ls + R}.$$

The transfer function from u to y is

$$G(s) = \frac{Z(s)}{Z(s) + L_f s} = \frac{RLs}{RLs + L_f s(C_f RLs^2 + Ls + R)}$$

$$= \frac{\frac{1}{L_f C_f}}{s^2 + \frac{1}{RC_f}s + \frac{L_f + L}{L_f C_f L}} = \frac{b_0}{s^2 + a_1 s + a_0}.$$

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 simplified circuit is of order 3, since it has 3 independent energy storage elements: L_f, C_f and L. But G is of order 2 only. This discrepancy is due to an unobservable eigenvector of the system matrix, which is actually an equilibrium point (i.e., it corresponds to the eigenvalue zero): all voltages are zero and a current $i_0 \neq 0$ is flowing through L_f (from left to right) and then down through L to the ground.

(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

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

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

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

(d) We are searching for the optimal real θ. We put $\tilde{y}_k = \mathrm{Re}\, y_k$, $\tilde{\varphi}_k = \mathrm{Re}\, \varphi_k$ for $k = 1, 2, \dots 25$, and $\tilde{y}_k = \mathrm{Im}\, y_{k-25}$, $\tilde{\varphi}_k = \mathrm{Im}\, \varphi_{k-25}$ 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.

(e)

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

(f)

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

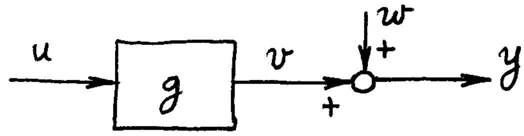
$$G^d(z) = C(zI - A^d)^{-1} B^d + D$$

(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 absolute values of the poles of G are much smaller than $2\pi/T$, 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, $E(u_k) \overset{a.s.}{=} \lim_{N\to\infty} \frac{1}{N}\sum_{j=1}^{N} u_j$, where the abreviation a.s. ("almost sure") means that the equality holds with probability 1. A similar formula holds for $E(y_k)$, obviously. Denote

$$\mathring{u}_k = u_k - E(u_k), \qquad \mathring{y}_k = y_k - E(y_k),$$

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

$$C^{uu}_\tau = E(\mathring{u}_k \cdot \mathring{u}_{k-\tau}) \overset{a.s.}{=} \lim_{N\to\infty} \frac{1}{N}\sum_{j=1}^{N} \mathring{u}_j\, \mathring{u}_{j-\tau},$$

$$C^{yu}_\tau = E(\mathring{y}_k \cdot \mathring{u}_{k-\tau}) \overset{a.s.}{=} \lim_{N\to\infty} \frac{1}{N}\sum_{j=1}^{N} \mathring{y}_j\, \mathring{u}_{j-\tau}.$$

In practice, we have only finitely many data, so that in all the above formulas, we have to replace $\lim_{N\to\infty} \frac{1}{N}\sum_{j=1}^{N}$ 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

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

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^{yu}_τ.

(b) $C^{yu}_\tau = E(\mathring{y}_k \cdot \mathring{u}_{k-\tau}) = E(\mathring{v}_k\, \mathring{u}_{k-\tau}) + E(\mathring{w}_k\, \mathring{u}_{k-\tau}) =$
$= C^{vu}_\tau + C^{wu}_\tau$, so that $C^{yu} = C^{vu} + C^{wu}$
$= g * C^{uu} + C^{wu}$.

(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:

$$\begin{bmatrix} C_0^{uu} & C_{-1}^{uu} & C_{-2}^{uu} & \cdots \\ C_1^{uu} & C_0^{uu} & C_{-1}^{uu} & \cdots \\ C_2^{uu} & C_1^{uu} & C_0^{uu} & \cdots \\ \vdots & \vdots & \vdots & \end{bmatrix} \cdot \begin{bmatrix} g_0 \\ g_1 \\ g_2 \\ \vdots \end{bmatrix} = \begin{bmatrix} C_0^{yu} \\ C_1^{yu} \\ C_2^{yu} \\ \vdots \end{bmatrix}. \qquad (**)$$

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

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

(f) The difference equation of the FIR filter is $\begin{pmatrix} u = \text{input}, \\ v = \text{output} \end{pmatrix}$

$$v_k = g_0 u_k + g_1 u_{k-1} + g_2 u_{k-2} \cdots + g_{N-1} u_{k-N+1}.$$