E4.27
C2.2
ISE4.41

IMPERIAL COLLEGE LONDON

DEPARTMENT OF ELECTRICAL AND ELECTRONIC ENGINEERING
EXAMINATIONS 2007

MSc and EEE/ISE PART IV: MEng and ACGI

SYSTEM IDENTIFICATION

Corrected Copy

Tuesday, 8 May 10:00 am

Time allowed: 3:00 hours

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) : S. Evangelou

Special information for invigilators: none

Information for candidates:

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

1. Consider the system with input u and output y modeled by the ARMAX difference equation

$$y_k + a_1 y_{k-1} + a_2 y_{k-2} = b_1 u_{k-1} + e_k - 0.5 e_{k-1},$$

where a_1, a_2, b_1 are unknown real numbers. The signal e is white noise with unknown variance and such that $E(e_k) = 0$.

(a) Write the formula for the transfer function H from e to y. For which values of a_2 is it possible that H is stable? [3]

(b) Assuming that H is stable and $u = 0$, give explicit formulas for $E(y_k)$ (the expectation of y) and for $E(y_k^2)$ (the power of y) in terms of the impulse response of H, denoted $h = (h_0, h_1, h_2, \ldots)$, and in terms of the noise variance $\sigma^2 = E(e_k^2)$. [5]

(c) Introduce new input and output variables u^F and y^F such that (i) u^F and y^F can be computed from u and y, (ii) u and y can be computed from u^F and y^F, (iii) the relation between u^F, y^F and e is described by an ARX difference equation with the same unknown parameters a_1, a_2 and b_1. Write down this ARX equation. [4]

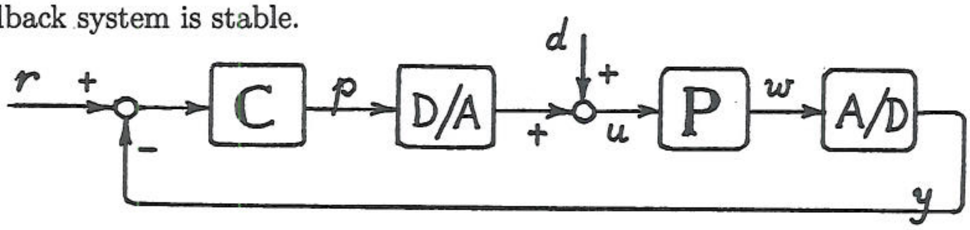
(d) Describe a least squares based method to estimate a_1, a_2 and b_1 from measurements of u_k and y_k for $k \leq 200$. Give a formula for an unbiased estimate of $Var(e_k)$. Hint: use the variables u^F and y^F introduced in part (c). [4]

(e) Assume that a_1, a_2 and b_1 have been estimated with great accuracy. Give a formula for an unbiased prediction of y_k, denoted $\overline{y_k}$, if the values $y_{k-1}, y_{k-2}, y_{k-3}, \ldots$ and $u_k, u_{k-1}, u_{k-2}, \ldots$ are known. [4]

2. Consider the feedback system shown in the figure below, where a continuous-time plant with transfer function $\mathbf{P}$ is controlled by a discrete-time controller with known clock period $h > 0$ and transfer function $\mathbf{C}$. Here, r is a variable reference signal and d is a *constant* disturbance signal. The signals r, p and y are discrete-time, while d, u and w are continuous-time. As usual, the D/A converter is a zero order hold of period h, while the A/D converter is a sampler of period h. We have

$$\mathbf{P}(s) = \frac{K}{1+Ts}, \qquad \mathbf{C}(z) = c_0 + c_1 z^{-1} + c_2 z^{-2} \ldots + c_9 z^{-9}.$$

The coefficients of $\mathbf{C}$ are known, while K, T and d are unknown and should be estimated. The true transfer function of the plant may be more complicated, but we would like to model it by the simple function given above. The whole feedback system is stable.



(a) Assuming that $\mathbf{P}$ is given by the simple formula above, compute (exactly) the tranfer function $\mathbf{P}^d$ of the discretized plant from p to y. For which values of K and T is $\mathbf{P}^d$ stable? [4]

(b) Suppose that the values r_k, y_k are available for $1 \leq k \leq 2000$. By defining new variables if necessary, find a model of the system of 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. [4]

(c) Describe a least squares based method for estimating K, T and d from the measurements of r_k and y_k, $1 \leq k \leq 2000$, using the model derived in part (b). [4]

(d) We denote by $\hat{\theta}$ the least squares estimate of θ for the model derived in part (b). Which of the three reference signals listed below will lead to the smallest covariance matrix Cov $\hat{\theta}$ (as measured by its norm)? Which will lead to the smallest value for $\widehat{\text{Var}}\, e_k$, the estimated variance of e_k? Give a brief reasoning.

$$\text{(i) } r_k = 1, \qquad \text{(ii) } r_k = \cos 0.2k,$$

$$\text{(iii) } r_k = \text{white noise with } E(r_k) = 0, \quad \text{Var } r_k = 1.$$ [4]

(e) If K and T have been found, how can we approximate the transfer function from p to y by a FIR transfer function of order 10? [4]

3. Suppose that $u = (u_k)$ is a stationary and ergodic Gaussian random signal in discrete time ($k \in \mathbb{Z}$). An LTI system with transfer function $\mathbf{G}$ has u as its input signal and $y = (y_k)$ as its output signal. We model $\mathbf{G}$ by

$$\mathbf{G}(z) = \frac{b}{z+a},$$

where the parameters $a \in (0,1)$ and b are unknown.

A company (your employer) requires you to produce formulas which give good predictions of u_n and y_n based on the measurements of u_k and y_k for $k = n-1, n-2, \ldots n-50$. The company supplies you with the measurements of u_k and y_k for $k = 1, 2, \ldots 5{,}000$, and you must design your prediction formulas using these data.

(a) Denote by C^{uu} the auto-correlation function of u. Explain how to estimate $\mu_u = E(u_k)$ and C^{uu}_j for $j = 0, 1, \ldots 50$. [3]

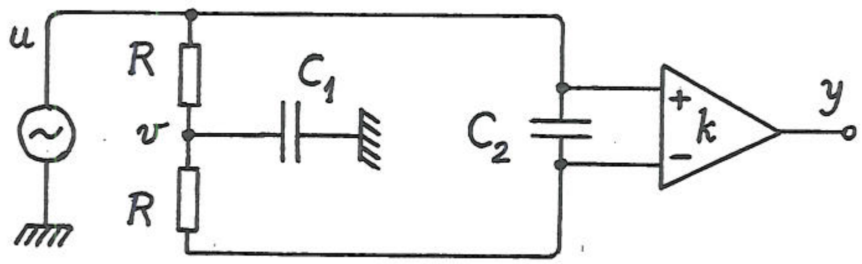
(b) Explain how to estimate the coefficients of a stable auto-regressive filter of order m, with transfer function denoted by Ξ, such that u can be regarded as the output function of this filter, when the input function of the filter is white noise $e = (e_k)$. Explain how to estimate $\mu_e = E(e_k)$ and $\sigma_e^2 = \mathrm{Var}(e_k)$ after Ξ has been identified. [5]

(c) Explain how to obtain the formula for the prediction of u_n, denoted by $\overline{u_n}$, using the results you obtained in part (b). Assuming that you have identified Ξ, μ_e and σ_e^2 with very high accuracy, give an estimate for the variance of the prediction error $u_n - \overline{u_n}$. [5]

(d) Outline a least squares based method to estimate a and b. Do not give any proofs. [3]

(e) Explain how to obtain the formula for the prediction of y_n, denoted by $\overline{y_n}$, using the results you obtained in the earlier parts. Give an estimate for the variance of the prediction error $y_n - \overline{y_n}$.
Hint: do not rush your answer, think carefully which data are available for predicting y_n. [4]

4. In the model circuit shown below, $R = 1k\Omega$ while C_1, C_2 and the gain k of the differential amplifier are unknown positive quantities. No current is flowing to the inputs of the amplifier. We can choose the waveform of u and we can measure the output voltage y. The true circuit is more complicated than the model circuit shown, and hence we cannot expect a perfect match between its response and the response of the model, but we would like to get a close match in a certain frequency range.



(a) Choose state variables and construct a state space representation of the model circuit, of the form $\dot{x} = Ax + Bu$, $y = Cx + Du$, where x is the state and A, B, C and D are matrices. Is this (model) system stable? [6]

(b) Compute the transfer function $\mathbf{G}$ of the model circuit (from u to y), in terms of R, C_1 and C_2. Evaluate the gain of $\mathbf{G}$ for very low and for very high frequencies (i.e., for $\omega \to 0$ and for $\omega \to \infty$). [6]

(c) Suppose that by measurements that use sinusoidal u, we have obtained estimates for $\mathbf{G}$ at 50 angular frequencies $\omega_1, \ldots \omega_{50}$, in the frequency range of interest. Using these data, how could we estimate C_1, C_2 and k using a least squares based algorithm? Write down the formulas which give the estimated C_1, C_2 and k, taking care to define all the symbols that you use. Take care to make sure that the estimates for C_1, C_2 and k are real. [8]

5. Assume that v is a stationary ergodic random signal with the expectation $E(v(t)) = 3$ and the auto-correlation function

$$C^{vv}(\tau) = e^{-100|\tau|}.$$

The signal v is sampled with a sampling period $h = 10^{-4}$, and the resulting discrete-time signal $u_k = v(hk)$ is applied to an unknown stable discrete-time LTI system Σ with the (unknown) transfer function $\mathbf{G}$. The output signal of Σ, denoted by y, is corrupted by measurement noise e, such that in terms of $\mathcal{Z}$-transforms

$$\hat{y} = \mathbf{G}\hat{u} + \hat{e}.$$

The measurement noise e is white noise with $E(e_k) = 0$, independent of v. The measurements u_k and y_k are available for $k = 1, 2, \ldots 5,000$.

(a) Are u and y jointly stationary? Explain very briefly what this concept means. Are u and y jointly ergodic? Again, explain very briefly what this means. [3]

(b) Compute the power and the power spectral density of the discrete time signal u. [3]

(c) Describe a method for estimating $E(y_k)$, $\mathrm{Var}(y_k)$ and C^{uy}_τ (the cross-correlation function of u and y), using the available measurements u_k and y_k. For what values of τ can we obtain reasonable estimates of C^{uy}_τ? [3]

(d) Assume that the system Σ is sufficiently stable so that its impulse response (g_k) is negligible for $k \geq 50$. Describe a method for estimating the first 50 terms $g_0, g_1, \ldots g_{49}$ from the results of part (c). [4]

(e) What is the meaning of a random signal being "persistent of order m"? What is the significance of this concept in the context of part (d) above? Is u persistent of order 50? Give a brief reasoning. Hint: use your result from part (b). [3]

(f) Compute $\mathrm{Var}(y_k)$ in terms of the impulse response (g_k), the cross-correlation function C^{uy} and the noise power $\mathrm{Var}(e_k)$ (the latter is not known). Hence, give a formula for estimating $\mathrm{Var}(e_k)$ in terms of quantities estimated earlier. [4]

[END]

SYSTEM IDENTIFICATION, May 2007
Solutions

Master copy
may 07

Question 1. (a) $H(z) = \dfrac{1-0.5z^{-1}}{1+a_1 z^{-1}+a_2 z^{-2}} =$

$$= \frac{z^2-0.5z}{z^2+a_1 z+a_2} = \frac{z(z-0.5)}{(z-\lambda_1)\cdot(z-\lambda_2)}, \quad a_2 = \lambda_1\lambda_2.$$

If H is stable, then $|\lambda_1|<1$, $|\lambda_2|<1$, hence $|a_2|<1$.

[3]

(b) $E(y_k) = H(1)\cdot E(e_k) = 0$. $E(y_k^2) =$

$$= E\big((h_0 e_k + h_1 e_{k-1} + h_2 e_{k-2} + \dots)(h_0 e_k + h_1 e_{k-1} + h_2 e_{k-2} + \dots)\big)$$

$$= E\Big(h_0^2 e_k^2 + h_1^2 e_{k-1}^2 + h_2^2 e_{k-2}^2 + \dots + \sum_{j\neq \ell} h_j h_\ell e_{k-j} e_{k-\ell}\Big)$$

$$= h_0^2 \underbrace{E(e_k^2)}_{\sigma^2} + h_1^2 E(e_{k-1}^2) + h_2^2 \underbrace{E(e_{k-2}^2)}_{\sigma^2} + \dots + \sum_{j\neq\ell} h_j h_\ell \underbrace{E(e_{k-j} e_{k-\ell})}_{0}$$

Thus,

$$E(y_k^2) = \sigma^2\left(h_0^2 + h_1^2 + h_2^2 + \dots\right).$$

The same result could be obtained from $C^{yy} = \check{h} * h * C^{ee}$, using $C^{ee} = \sigma^2\delta_0$ and computing $E(y_k^2) = C^{yy}(0)$. (Here, $\check{h}_k = h_{-k}$.) [5]

(c) $\hat{u}^F(z) = \dfrac{1}{1-0.5z^{-1}}\hat{u}(z)$, $\quad \hat{y}^F(z) = \dfrac{1}{1-0.5z^{-1}}\hat{u}(z)$, i.e., u^F and y^F are the solutions of

$$u^F_k - 0.5\, u^F_{k-1} = u_k \,, \qquad y^F_k - 0.5\, y^F_{k-1} = y_k$$

with initial conditions zero : $u^F_0 = y^F_0 = 0$ (the influence of the initial conditions is anyway negligible for large k). Then the ARX equation is

$$y^F_k + a_1 y^F_{k-1} + a_2 y^F_{k-2} = b_1 u^F_{k-1} + e_k \,. \qquad [4]$$

(d)

$$y^F_k = \underbrace{\begin{bmatrix} -y^F_{k-1} & -y^F_{k-2} & u^F_{k-1} \end{bmatrix}}_{\varphi_k} \underbrace{\begin{bmatrix} a_1 \\ a_2 \\ b_1 \end{bmatrix}}_{\theta} + e_k$$

$$\Phi = \begin{bmatrix} \varphi_1 \\ \varphi_2 \\ \vdots \\ \varphi_{200} \end{bmatrix}, \quad \Phi^{\#} = (\Phi^* \Phi)^{-1} \Phi^* ,$$

unbiased estimate of θ

$$\hat{\theta} = \Phi^{\#} y^F, \qquad y^F = \begin{bmatrix} y^F_1 \\ y^F_2 \\ \vdots \\ y^F_{200} \end{bmatrix}.$$

$$\mathrm{Var}(e_k) = \frac{1}{197} \| y^F - \Phi \hat{\theta} \|^2 \qquad \left(\begin{array}{c}\text{because} \\ 200-3=197\end{array}\right)$$

unbiased estimate of $\mathrm{Var}(e_k)$

$$= \frac{1}{197}\, y^{F*} (I - \Phi \Phi^{\#}) y^F . \qquad [4]$$

(e) According to the one step ahead prediction formula for systems described by an ARMAX equation, we have, denoting $\overline{y_k} = E(y_k \mid y_{k-1}, y_{k-2}, \dots)$ and assuming that the input signal u is known,

$$\overline{y_k} = (c_1 - a_1)\, y^F_{k-1} + (c_2 - a_2)\, y^F_{k-2} + b_0 u^F_k + b_1 u^F_{k-1} \,.$$

Here u^F_k and y^F_k are the signals introduced in our answer to part (c), $c_1 = -0.5$, $b_0 = 0$, $c_2 = 0$. [4]

Question 2. (a) A state space realization of $\mathbf{P}$ is $\dot{x} = -\frac{1}{T}x + \frac{1}{T}u$, $w = Kx$. This, together with the A/D and D/A blocks, gives the exact discretization of the plant, described by

$$x_{k+1} = e^{-h/T}x_k + (e^{-h/T} - 1)(-T)\cdot\frac{1}{T}(p_k + d)$$
$$= e^{-h/T}x_k + (1 - e^{-h/T})(p_k + d), \quad y_k = Kx_k,$$

which has the transfer function

$$\mathbf{P}^d(z) = K(z - e^{-h/T})^{-1}(1 - e^{-h/T}).$$

Denoting $a = e^{-h/T}$ and $b = K(1 - e^{-h/T})$, we obtain $\mathbf{P}^d(z) = b/(z-a)$. This is stable for $|a| < 1$, i.e., for all $T > 0$ and all $K \in \mathbb{R}$. [4]

(b) Since $\mathbf{C}$ is known and stable, we can compute

$$p_k = c_0(r_k - y_k) + c_1(r_{k-1} - y_{k-1}) \ldots + c_9(r_{k-9} - y_{k-9}).$$

The first nine values $p_1, \ldots p_9$ may be affected by the (unknown) initial state of the controller, afterwards the initial state has no influence. Since we have 2000 measurements, the effect of the initial state of the controller will be small. From $y_k - ay_{k-1} = b(p_{k-1} + d) + e_k$ we get, with $\theta^T = [a \;\; b \;\; bd]$,

$$y_k = [y_{k-1} \;\; p_{k-1} \;\; 1]\,\theta + e_k.$$

[4]

(c) Denoting $\varphi_k = [y_{k-1} \;\; p_{k-1} \;\; 1]$ and

$$\Phi = \begin{bmatrix} \varphi_1 \\ \vdots \\ \varphi_{2000} \end{bmatrix}, \quad \Phi^{\#} = (\Phi^* \Phi)^{-1} \Phi^*,$$

the vector θ which minimizes $\|e\|^2 = |e_1|^2 + |e_2|^2 \ldots + |e_{2000}|^2$ is given by $\hat{\theta} = \Phi^{\#} y$, where $y = [y_1 \;\; y_2 \;\; \ldots \;\; y_{2000}]^T$. $\hat{\theta}$ is an unbiased estimate of θ, if $E(e_k) = 0$. [4]

(d) The white noise r will lead to the smallest $\mathrm{Cov}\,\hat{\theta}$, because it subjects the plant to a wide range of frequencies. The possibilities (i) and (ii) test the system at one frequency only. (We remark that (ii) will work better than (i), i.e., will lead to smaller $\mathrm{Cov}\,\hat{\theta}$, because a sinusoidal signal has both amplitude and phase information, while a constant has only an amplitude. Moreover, for constant r, we cannot distinguish between the effect of r and the effect of d). For $\widehat{\mathrm{Var}}\, e_k$, the picture is very different: we expect it to be smallest for constant r_k. Indeed, for constant r_k, since the whole feedback system is stable, all signals will converge to constants. Then the equation $y_k = \varphi_k \theta + e_k$ can be satisfied with a suitable choice of θ such that practically $e_k = 0$. [4]

(e)

$$P^d(z) = \frac{b z^{-1}}{1 - a z^{-1}} = b z^{-1}\left(1 + a z^{-1} + a^2 z^{-2} + \ldots\right),$$

$$P^d(z) \approx b z^{-1} + b a z^{-2} + b a^2 z^{-3} \ldots + b a^9 z^{-10},$$

where $a = e^{-h/T}$, $b = K(1 - e^{-h/T})$. [4]

Question 3.

The block diagram corresponding to the problem statement and point (b):

$$e \longrightarrow \boxed{\Xi} \xrightarrow{\;u\;} \boxed{G} \longrightarrow y$$

(a) Since u is ergodic, we can estimate expectations by time averages:

$$\hat{\mu}_u = \frac{1}{N}\sum_{k=1}^{N} u_k, \qquad \hat{C}^{uu}_j = \frac{1}{N-j}\sum_{k=1}^{N-j} \tilde{u}_k \cdot \tilde{u}_{k+j},$$

where $\tilde{u}_k = u_k - \hat{\mu}_u$. This will work reasonably well since $j \leq 50$, which is much less than $N = 5{,}000$.

[3]

(b) We represent u as $\hat{u} = \Xi \hat{e}$, where $e = (e_k)$ is white noise and $\Xi^{-1}(z) = 1 + g_1 z^{-1} + g_2 z^{-2} + \ldots$, Ξ^{-1} is stable (like Ξ), so that $g_k \to 0$. We have

$$\begin{bmatrix} C^{uu}_0 & C^{uu}_1 & C^{uu}_2 & \ldots \\ C^{uu}_1 & C^{uu}_0 & C^{uu}_1 & \ldots \\ \vdots & \vdots & \vdots & \end{bmatrix} \cdot \begin{bmatrix} g_1 \\ g_2 \\ \vdots \end{bmatrix} = - \begin{bmatrix} C^{uu}_1 \\ C^{uu}_2 \\ \vdots \end{bmatrix},$$

which is an infinite sequence of equations in infinitely many unknowns $g_1, g_2, \ldots$. However, if we assume that g_k is negligible for $k > m$ then we can truncate this to a linear system of m equations with m unknowns $g_1, \ldots g_m$. The coefficients C^{uu}_k can be estimated, see part (a). If u is persistent of order m, then we can solve for $g_1, \ldots g_m$. This gives us an auto-regressive realization of the estimated Ξ.

We have $E(u) = \Xi(1) \cdot E(e)$. Since $\Xi(1) = (1 + g_1 + g_2 \dots + g_m)^{-1}$ has been estimated and also $E(u) = \mu_u$ (see part (a)), we can now estimate $\mu_e = E(e)$. To estimate $\mathrm{Var}(e)$, we use the Wiener-Lee formula for $\nu = 0$ ($e^{i\nu} = 1$):

$$\underbrace{S^{uu}(1)}_{\sum_{j=-\infty}^{\infty} C_j^{uu}} = \underbrace{|\Xi(1)|^2}_{(1+g_1+g_2\dots)^{-2}} \underbrace{S^{ee}(1)}_{\mathrm{Var}(e)}$$

We estimate $\widehat{S^{uu}(1)} = \sum_{j=-m+1}^{m-1} \widehat{C_j^{uu}}$ (recall that $C_{-j}^{uu} = C_j^{uu}$) and from here we obtain an estimate for $\mathrm{Var}(e)$:

$$\widehat{\mathrm{Var}(e)} = (1 + g_1 + g_2 \dots + g_m)^2 \sum_{j=-m+1}^{m-1} \widehat{C_j^{uu}}.$$

[5]

(c) We have (from $\Xi^{-1}\hat{u} = \hat{e}$)

$$u_n = -g_1 u_{n-1} - g_2 u_{n-2} \dots - g_m u_{n-m} + e_n. \quad (*)$$

$\bar{u}_n$ is the conditional expectation of u_n, given the measurements $u_{n-1}, u_{n-2}, \dots$. If we take conditional expectations of both sides of (*), we obtain

$$\bar{u}_n = -g_1 u_{n-1} - g_2 u_{n-2} \dots - g_m u_{n-m} + \mu_e.$$

Comparing this with (*), we see that the prediction error is $e_n - \mu_e$. Hence, the variance of the prediction error is $\mathrm{Var}(e)$ (which has been estimated in part (b)).

[5]

(d) We have $G(z) = \frac{bz^{-1}}{1+az^{-1}}$, hence

$y_k + a y_{k-1} = b u_{k-1}$, which we write as

$$y_k = \underbrace{\begin{bmatrix} -y_{k-1} & u_{k-1} \end{bmatrix}}_{\varphi_k} \underbrace{\begin{bmatrix} a \\ b \end{bmatrix}}_{\theta} + \varepsilon_k$$

where ε_k is the modeling error and θ is the vector of unknown parameters. Introducing

$$\Phi = \begin{bmatrix} \varphi_1 \\ \vdots \\ \varphi_N \end{bmatrix}, \quad \Phi^\# = (\Phi^* \Phi)^{-1} \Phi^*,$$

the estimate which minimizes $\varepsilon_1^2 + \varepsilon_2^2 \dots + \varepsilon_N^2$ is $\hat{\theta} = \Phi^\# y$, where $y = [y_1 \; y_2 \dots y_N]^T$.

[3]

(e) From $y_n = -a y_{n-1} + b u_{n-1} + \varepsilon_n$,

taking conditional expectations, we get

$$\overline{y_n} = -a y_{n-1} + b u_{n-1}.$$

Note that this has nothing to do with e_n and with predicting u_n, because we only need u_{n-1}. The prediction error is ε_n, and an estimate for $\mathrm{Var}(\varepsilon)$ is

$$\widehat{\mathrm{Var}(\varepsilon)} = \frac{1}{N-2} \| y - \Phi \hat{\theta} \|^2,$$

as we know from the least squares theory.

[4]

Question 4. (a) We denote by z the voltage on the capacitor C_2 (from top to bottom), so that $y = kz$. We denote by v the voltage on the capacitor C_1 (from left to right). Then the currents through the capacitors are

$$C_1 \dot{v} = \frac{u-v}{R} + \frac{u-z-v}{R},$$

$$C_2 \dot{z} = \frac{u-z-v}{R}.$$

The state vector should be $x = \begin{bmatrix} v \\ z \end{bmatrix}$. Then

$$\underbrace{\frac{d}{dt}\begin{bmatrix} v \\ z \end{bmatrix}}_{\dot{x}} = \underbrace{\begin{bmatrix} -\frac{2}{RC_1} & -\frac{1}{RC_1} \\ -\frac{1}{RC_2} & -\frac{1}{RC_2} \end{bmatrix}}_{A} \cdot \underbrace{\begin{bmatrix} v \\ z \end{bmatrix}}_{x} + \underbrace{\begin{bmatrix} \frac{2}{RC_1} \\ \frac{1}{RC_2} \end{bmatrix}}_{B} u,$$

$$y = \underbrace{\begin{bmatrix} 0 & k \end{bmatrix}}_{C} \cdot \begin{bmatrix} v \\ z \end{bmatrix} + \underbrace{\begin{bmatrix} 0 \end{bmatrix}}_{D} \cdot u$$

We have $\det A = \frac{2}{R^2 C_1 C_2} - \frac{1}{R^2 C_1 C_2}$ $= \frac{1}{R^2 C_1 C_2} > 0$. Since $\operatorname{trace} A < 0$ and $\det A > 0$, A is stable. [6]

(b) The characteristic polynomial of A is

$$p(s) = \det(sI - A) = s^2 - (\operatorname{trace} A)s + \det A$$

$$= s^2 + \underbrace{\frac{1}{R}\left(\frac{2}{C_1} + \frac{1}{C_2}\right)}_{a_1} s + \underbrace{\frac{1}{R^2 C_1 C_2}}_{a_0} .$$

We have $G(s) = C(sI - A)^{-1} B =$

$$= \begin{bmatrix} 0 & k \end{bmatrix} \frac{1}{p(s)} \begin{bmatrix} s + 1/RC_2 & -1/RC_1 \\ -1/RC_2 & s + 2/RC_1 \end{bmatrix} \begin{bmatrix} 2/RC_1 \\ 1/RC_2 \end{bmatrix}$$

$$= \frac{k}{p(s)} \begin{bmatrix} -\frac{1}{RC_2} & s + \frac{2}{RC_1} \end{bmatrix} \begin{bmatrix} \frac{2}{RC_1} \\ \frac{1}{RC_2} \end{bmatrix}$$

$$= \frac{k}{p(s)} \begin{bmatrix} -\nu_2 & s + \nu_1 \end{bmatrix} \begin{bmatrix} \nu_1 \\ \nu_2 \end{bmatrix} = \frac{k s \nu_2}{p(s)} ,$$

where we have used the notation

$$\nu_1 = \frac{2}{RC_1} , \qquad \nu_2 = \frac{1}{RC_2} .$$

Thus,

$$G(s) = \frac{\overbrace{k \nu_2}^{b_1} s}{s^2 + \underbrace{(\nu_1 + \nu_2)}_{a_1} s + \underbrace{0.5\, \nu_1 \nu_2}_{a_0}} .$$

We clearly have

$$\lim_{\omega \to 0} G(i\omega) = \lim_{\omega \to \infty} G(i\omega) = 0 . \qquad [6]$$

(c) We denote by $G^e(i\omega_k)$ the experimentally determined values of the true transfer function, $k=1,\dots 50$, which are subject to measurement errors. We have

$$b_1 i\omega_k = \left[(i\omega_k)^2 + a_1(i\omega_k) + a_0\right] G^e(i\omega_k) + e_k$$

where e_k is due to the combined effect of the measurement and modeling errors. We rewrite this:

$$\underbrace{\omega_k^2 G^e(i\omega_k)}_{y_k} = \underbrace{\left[i\omega_k G^e(i\omega_k) \quad G^e(i\omega_k) \quad -i\omega_k\right]}_{\varphi_k} \theta + e_k$$

where $\theta^T = \left[a_1 \ \ a_0 \ \ b_1\right]$ are the unknown parameters, while y_k and φ_k are known. Denoting

$$y = \begin{bmatrix} y_1 \\ \vdots \\ y_{50} \end{bmatrix}, \quad \Phi = \begin{bmatrix} \varphi_1 \\ \vdots \\ \varphi_{50} \end{bmatrix} \in \mathbb{C}^{50\times 3}, \quad e = \begin{bmatrix} e_1 \\ \vdots \\ e_{50} \end{bmatrix},$$

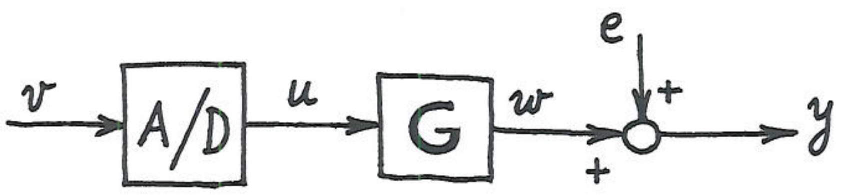
we obtain the usual equation $y = \Phi\theta + e$. However, Φ and y are complex, while we are searching for the optimal real θ. Thus we decompose into the 2×50 equations

$$\begin{cases} \mathrm{Re}\, y = (\mathrm{Re}\,\Phi)\theta + \mathrm{Re}\, e \\ \mathrm{Im}\, y = (\mathrm{Im}\,\Phi)\theta + \mathrm{Im}\, e \end{cases}$$

Denoting $\tilde{y} = \begin{bmatrix} \mathrm{Re}\, y \\ \mathrm{Im}\, y \end{bmatrix}$, $\tilde{\Phi} = \begin{bmatrix} \mathrm{Re}\,\Phi \\ \mathrm{Im}\,\Phi \end{bmatrix}$, $\tilde{e} = \begin{bmatrix} \mathrm{Re}\, e \\ \mathrm{Im}\, e \end{bmatrix}$, we arrive at the real equation $\tilde{y} = \tilde{\Phi}\theta + \tilde{e}$. Now we can find the optimal (real) θ by the least squares algorithm: $\hat{\theta} = \tilde{\Phi}^{\#} y$, where $\tilde{\Phi}^{\#} = (\tilde{\Phi}^*\tilde{\Phi})^{-1}\tilde{\Phi}^*$.

From a_1, a_0, b_1 we can compute ν_1, ν_2 and k, and hence also C_1, C_2. [8]

Question 5. The block diagram corresponding to the problem statement:



(e = white noise, independent of v)

(a) $u_k = v(hk)$ is stationary and ergodic, hence $\begin{bmatrix} u \\ w \end{bmatrix}$ is stationary and ergodic. Since e is independent of v, it is also independent of $\begin{bmatrix} u \\ w \end{bmatrix}$ (which is generated from v). Hence, $\begin{bmatrix} u \\ y \end{bmatrix}$ is stationary and ergodic. Stationarity of $x = \begin{bmatrix} u \\ y \end{bmatrix}$ (also called "joint stationarity of u, y") means that the distribution functions of

$$\begin{bmatrix} x_{m_1} \\ x_{m_2} \\ \vdots \\ x_{m_n} \end{bmatrix} \quad \text{and} \quad \begin{bmatrix} x_{m_1+\tau} \\ x_{m_2+\tau} \\ \vdots \\ x_{m_n+\tau} \end{bmatrix}$$

are equal, for any $n \in \mathbb{N}$, $m_1, \dots m_n \in \mathbb{Z}$ and $\tau \in \mathbb{Z}$. Ergodicity of $\begin{bmatrix} u \\ y \end{bmatrix}$ (also called "joint ergodicity") means that for any measurable function g of several variables,

$$E\left(g\left(\begin{bmatrix} u_{k_1} \\ y_{k_1} \end{bmatrix}, \dots \begin{bmatrix} u_{k_n} \\ y_{k_n} \end{bmatrix}\right)\right) = \lim_{N \to \infty} \frac{1}{2N} \sum_{j=-N}^{N} g\left(\begin{bmatrix} u_{k_1+j} \\ y_{k_1+j} \end{bmatrix}, \dots \begin{bmatrix} u_{k_n+j} \\ y_{k_n+j} \end{bmatrix}\right)$$

with probability 1 (i.e., expectations are equal to infinite time averages). [3]

(b) We have $C_k^{uu} = C^{vv}(hk) = e^{-100h|k|}$, where $h = 10^{-4}$. In particular, the power of u is $\mathrm{Var}(u) = C_0^{uu} = 1$. The power spectral density of u is (for $|z| = 1$)

$$S^{uu}(z) = \sum_{k=-\infty}^{\infty} z^{-k} C_k^{uu}$$

denote $\rho = e^{-100h}$

$$= \sum_{k=-\infty}^{-1} z^{-k} e^{-100h|k|} + \sum_{k=0}^{\infty} z^{-k} \left(e^{-100h}\right)^k$$

$$= \underbrace{\sum_{k=1}^{\infty} z^k \rho^k} + \underbrace{\sum_{k=0}^{\infty} z^{-k} \rho^k}$$

$$= \frac{\rho}{z^{-1} - \rho} + \frac{\rho}{z - \rho} + 1$$

$$= \frac{z - \rho + \bar{z} - \rho}{(\bar{z} - \rho)(z - \rho)} \cdot \rho + 1$$

because $z^{-1} = \bar{z}$

$$= \frac{(z + \bar{z}) - 2\rho}{1 - (z + \bar{z})\rho + \rho^2} \cdot \rho + 1 \qquad [3]$$

If we denote $z = e^{i\nu}$ $(\nu \in (-\pi, \pi])$ then we obtain

$$\boxed{\begin{aligned} S^{uu}(e^{i\nu}) &= 1 + 2\rho \frac{\cos\nu - \rho}{1 - 2\cos\nu\, \rho + \rho^2} \\ &= \frac{1 - \rho^2}{1 - 2\cos\nu\, \rho + \rho^2}. \end{aligned}}$$

(Any of the last 5 formulas is a good answer)

(c) We can estimate $E(y_k)$ and $C^{uy}_\tau = C^{yu}_{-\tau}$ using ergodicity, as follows:

$$\widehat{E(y_k)} = \frac{1}{N}\sum_{k=1}^{N} y_k \,, \quad \widehat{C^{yu}_\tau} = \frac{1}{N-\tau}\sum_{k=1}^{N-\tau} \tilde{u}_k \tilde{y}_{k+\tau}$$

where $\tilde{u}_k = u_k - 3$, $\tilde{y}_k = y_k - \widehat{E(y_k)}$, $N = 5{,}000$ and $\tau << N$. Similarly, $\widehat{Var(y_k)} = \frac{1}{N}\sum_{k=1}^{N} (\tilde{y}_k)^2$. [3]

(d) We can estimate $g_0, g_1, \dots g_{49}$ by solving

$$\begin{bmatrix} C^{uu}_0 & C^{uu}_1 & \dots & C^{uu}_{49} \\ C^{uu}_1 & C^{uu}_0 & \dots & C^{uu}_{48} \\ \vdots & \vdots & & \vdots \\ C^{uu}_{49} & C^{uu}_{48} & \dots & C^{uu}_0 \end{bmatrix} \cdot \begin{bmatrix} g_0 \\ g_1 \\ \vdots \\ g_{49} \end{bmatrix} = \begin{bmatrix} \widehat{C^{yu}_0} \\ \widehat{C^{yu}_1} \\ \vdots \\ \widehat{C^{yu}_{49}} \end{bmatrix} .$$

This is because $C^{wu}_k = C^{yu}_k$, since $y = w + e$ and e is independent of u. [4]

(e) u is persistent of order m if the $m \times m$
[3] Toeplitz matrix shown above for $m = 50$ is invertible. If S^{uu} is rational and $S^{uu}(e^{i\nu}) > 0$ for all $\nu \in (-\pi, \pi]$, then u is persistent of any order — this is the case for our u, as can be seen from part (b).

(f) $C^{ww} = \check{g} * C^{wu} = \check{g} * C^{yu}$, hence $C^{ww}_0 =$
[4] $= \sum_{k=0}^{\infty} g_k C^{yu}_k = Var(w)$. Since e is independent of u, we have $Var(y) = Var(w) + Var(e)$, hence $Var(e) = Var(y) - \sum_{k=0}^{\infty} g_k C^{yu}_k$.