Exercise Set 7

1. *The Ho-Kalman realization algorithm*
Consider the samples $\{h_k\}_{k=1}^{\infty}$ of the Kronecker pulse response of a strictly proper discrete-time MIMO LTI system (also known as its *Markov parameters*), and take $\alpha, \beta \in \mathbb{N}$. The Hankel matrix of the system is defined as:

$$H(k-1) := \begin{bmatrix} h_k & h_{k+1} & \cdots & h_{k+\beta-1} \\ h_{k+1} & h_{k+2} & \cdots & h_{k+\beta} \\ \vdots & \vdots & \ddots & \vdots \\ h_{k+\alpha-1} & h_{k+\alpha} & \cdots & h_{k+\alpha+\beta-2} \end{bmatrix}; \qquad k = 0,1,\ldots$$

(a) Prove that if the system has minimal order n then $\operatorname{rank}\{H(k-1)\} \le n$ for all k. Also show that for α, β we have $\operatorname{rank}\{H(k-1)\} = n$. Notice that we are considering the noise free case here.
Hint. Write h_k in terms of the state space matrices of the system, and use the Cayley-Hamilton theorem to relate the columns of $H(k-1)$.

(b) Let $H(0) = R\Sigma S^T$ be the singular value decomposition (SVD) of $H(0)$, i.e. R and S are square and unitary ($RR^T = SS^T = I$) and Σ is diagonal (but of the same dimensions as $H(0)$). Suppose that

$$\Sigma = \begin{bmatrix} \Sigma_n & 0 \\ 0 & 0 \end{bmatrix}$$

where $\Sigma_n \in \mathbb{R}^{n\times n}$ is diagonal and nonsingular (because of part (a)). Then we can rewrite $H(0) = R_n \Sigma_n S_n^T$, where R_n and S_n are formed by the first n columns of R and S, respectively. Prove that a state space realization of the system is given by:

$$\begin{aligned} \hat{A} &= \Sigma_n^{-1/2} R_n^T H(1) S_n \Sigma_n^{-1/2} \\ \hat{B} &= \Sigma_n^{1/2} S_n^T E_m \\ \hat{C} &= E_p^T R_n \Sigma_n^{1/2} \end{aligned} \qquad (1)$$

where m and p are the number of inputs and outputs of the system, respectively, and

$$E_j^T = [I_j \quad 0_j \quad \cdots \quad 0_j]$$

where I_j is the identity matrix of order j, and O_j is the zero matrix of order j.
Hint. Based on the hint from part (a), show that there is a matrix T (independent of k) such that $H(k) = H(k-1)T$ for every $k \in \mathbb{N}$. Then, using this relation between $H(1)$ and $H(0)$, show that the expressions (1) produce a system which has a Kronecker pulse response given by $\{h_k\}$.

2. *A subspace identification method based on the Ho-Kalman algorithm*
Consider a system $y_t = G(q)u_t + w_t$, where

$$G(q) = \frac{q^{-1}}{1-0.6q^{-1}}$$

$\{w_t\}$ is Gaussian white noise of zero mean and variance 1, and

$$u_t = \sum_{k=0}^{3} \cos(k\pi/4)$$

To obtain a transfer function estimate of this system, we can use the Ho-Kalman algorithm described in Problem 1. An estimate of the Markov parameters can be obtained by correlation analysis [1, Section 6.1], since

$$R_{yu}(\tau) = \sum_{k=1}^{\infty} h_k R_u(k-\tau)$$

(a) Simulate the system, and obtain an estimate of the Markov parameters using the Matlab command *cra*.
(b) Use the Ho-Kalman algorithm of Problem 1 to estimate a low order transfer function. Use your judgement in the truncation of Σ.
(c) Use the Matlab command *spa*, and compare the results with part (b).

3. *An indirect PEM* (Problem 7.14 of [2])
Consider the system

$$y_t = b_0 u_{t-1} + \frac{1}{1+a_0 q^{-1}} w_t$$

where $\{u_t\}$ and $\{w_t\}$ are mutually independent white noise sequences with zero means and variances σ^2 and λ^2, respectively.

(a) Consider two ways of identifying the system.
Case (i): The model structure is given by

$$\mathcal{M}_1: \quad y_t = b u_{t-1} + \frac{1}{1+aq^{-1}} \varepsilon_t, \qquad \theta_1 = [a \quad b]^T$$

and a prediction error method is used.
Case (ii): The model structure is given by

$$\mathcal{M}_2: \quad y_t + a y_{t-1} = b_1 u_{t-1} + b_2 u_{t-2} + \varepsilon_t, \qquad \theta_1 = [a \quad b_1 \quad b_2]^T$$

and the least squares method is used.
Determine the asymptotic variances of the estimates. Compare the variances of $\hat{a}$ and $\hat{b}$ obtained by using $\mathcal{M}_1$ with the variances of $\hat{a}$ and $\hat{b}_1$ obtained by using $\mathcal{M}_2$.

(b) **(Advanced)** Case (i) gives better accuracy but requires much more computation than case (ii). One can therefore think of the following approach. First compute $\hat{\theta}_2$ as in case (ii). As a second step the parameters of $\mathcal{M}_1$ are estimated from

$$f(\hat{\theta}_1^*) = \hat{\theta}_2$$

where (compare $\mathcal{M}_1$ and $\mathcal{M}_2$)

$$f(\theta_1) = [a \quad b \quad ab]^T$$

Since this is an overdetermined system (3 equations, 2 unknowns), an exact solution is in general not possible. To overcome this difficulty the estimate $\hat{\theta}_1^*$ can be defined as the minimum point of

$$V(\theta_1) = [\hat{\theta}_2 - f(\theta_1)]^T Q [\hat{\theta}_2 - f(\theta_1)]$$

where Q is a positive weighting matrix. (Note that $V(\theta_1)$ does not depend explicitly on the data, so the associated minimization problem should require much less computation than that of case (i).)
Show that the asymptotic covariance matrix of $\hat{\theta}_1^*$ is given by

$$P_{\theta_1^*} = (F^T QF)^{-1} F^T QP_{\hat{\theta}_2} QF(F^T QF)^{-1}$$

where $P_{\hat{\theta}_2}$ is the covariance matrix of $\hat{\theta}_2$ and

$$F = \left. \frac{df(\theta_1)}{d\theta_1} \right|_{\theta_1 = \theta_{10}}$$

Hint. Let θ_{10}, θ_{20} denote the true parameter vectors. Then by a Taylor series expansion

$$\frac{1}{2} \left. \frac{dV(\theta_1)}{d\theta_1} \right|_{\theta_1 = \hat{\theta}_1^*} \approx -[\hat{\theta}_2 - f(\hat{\theta}_1^*)]^T QF$$

and $\hat{\theta}_2 - f(\hat{\theta}_1^*) \approx (\hat{\theta}_2 - \theta_{20}) - F(\hat{\theta}_1^* - \theta_{10})$. From these equations an asymptotically valid expression for $\hat{\theta}_1^* - \theta_{10}$ can be obtained.

(c) **(Advanced)** Show that the covariance matrix $P_{\theta_1^*}$ is minimized with respect to Q by the choice $Q = P_{\hat{\theta}_2}^{-1}$ in the sense that

$$P_{\theta_1^*} \geq (F^T P_{\hat{\theta}_2}^{-1} F)^{-1}$$

Evaluate the right-hand side explicitly for the system above. Compare with the covariance matrix for $\hat{\theta}_1$.

Remark. The choice $Q = P_{\hat{\theta}_2}^{-1}$ is not realistic in practice. Instead $Q = \hat{P}_{\hat{\theta}_2}^{-1}$ can be used. ($\hat{P}_{\hat{\theta}_2}$ is an estimate of $P_{\hat{\theta}_2}$ obtained by replacing $E\{\cdot\}$ by $1/N \sum_{t=1}^{N} \{\cdot\}$ in the expression for $P_{\hat{\theta}_2}$.)

4. *Numerical comparison of identification methods and prior information*
Consider the following system:

$$A_0(q) y_t = B_0(q) u_t + H_0(q) w_t$$

where

$$A_0(q) = 1 + 0.5q^{-1}$$
$$B_0(q) = q^{-1}$$
$$H_0(q) = \frac{1 + 0.2q^{-1}}{1 + 0.6q^{-1}}$$

Simulate the system with both $\{u_t\}$ and $\{w_t\}$ being mutually independent Gaussian white noise sequences of zero mean and variance 1.

(a) Use the Matlab command *spa* to obtain an estimate of the frequency response of $G_0(q) = B_0(q)/A_0(q)$, using the simulated data. Explain the results.

(b) Use the Matlab command *arx* to estimate the parameters of $A(q)$ and $B(q)$ assuming that you knew the corresponding orders and the time delay of the true system. Are the estimates asymptotically biased? Verify the results theoretically.

(c) Can your results be improved if you knew $H_0(q)$ *a priori*? How can you use this information in your estimation algorithm? Verify your claims both numerically and theoretically.

(d) Do your estimates change if you assumed that the system has no time delays (again using *spa* and *arx*).

5. **(Advanced)** *Nonsingularity condition of instruments used in subspace methods* (Problem 10G.6 of [1])
Consider the notation used in [1]. Show that, in open-loop operation, the matrix $\tilde{T}$ defined in (10.115) in [1] has full rank n provided that the following conditions are simultaneously satisfied:

a) $\bar{E}\{\varphi_s(t)\varphi_s^T(t)\}$ is positive definite.

b) $E\left\{\begin{bmatrix} \hat{x}(t) \\ U_r(t) \end{bmatrix}[\hat{x}^T(t) \quad U_r^T(t)]\right\}$ is positive definite. This means that the r future inputs should not be linearly dependent on the current prediction of the state.

c) s_1 and s_2 are sufficiently large so that $\hat{x}(t) \approx L_x\varphi_s(t)$ for some L_x. (See (10.123) in [1].)

Hint. Use that $E\{x(t)\varphi_s^T(t)\} = E\{\hat{x}(t)\varphi_s^T(t)\}$, where $\hat{x}(t)$ is that part of the state that can be reconstructed from past input-outputs. Similarly $E\{x(t)U_r^T(t)\} = E\{\hat{x}(t)U_r^T(t)\}$.

References

[1] L. Ljung. *System Identification: Theory for the User*, 2nd Edition. Prentice-Hall, 1999.
[2] T. Söderström and P. Stoica. *System Identification*. Prentice-Hall, 1989.