227-0689-00L System Identification **05 December 2018**

Solution 9: Instrumental variable methods

We will not provide solutions for the MATLAB parts of the exercises. These will instead be discussed in the exercises sessions.

Background reading

The background material for this exercise is Sections 7.5 and 7.6 of Ljung (*System Identification; Theory for the User*, 2nd Ed., Prentice-Hall, 1999).

Problem 1:

Suppose that a true description of a certain system is given by

$$y(k) + a_1 y(k-1) + ... + a_{n_a} y(k-n_a) = b_1 u(k-1) + ... + b_{n_b} u(k-n_b) + e(k)$$

where $\{e(k)\}$ is white noise independent of the input. Let the regressor $\phi(k)$ be defined as usual as

$$\phi(k) = [-y(k-1) \ldots - y(k-n_a) \quad u(k-1) \ldots u(k-n_b)]^T .$$

Moreover, let $\tilde{\phi}(k)$ be given by

$$\tilde{\phi}(k) = [-y_0(k-1) \ldots - y_0(k-n_a) \quad u(k-1) \ldots u(k-n_b)]^T ,$$

where

$$y_0(k) + a_1 y_0(k-1) + ... + a_{n_a} y_0(k-n_a) = b_1 u(k-1) + ... + b_{n_b} u(k-n_b) .$$

Note that y_0 is the noise-free response of the true system.

Prove that for the vector of instrumental variables $z(k) = [u(k-1)\ u(k-2)]^T$ we have

$$\mathbb{E}[z(k)\phi(k)^T] = \mathbb{E}[z(k)\tilde{\phi}(k)^T] .$$

Solution

One has :

$$\mathbb{E}[z(k)\phi(k)^T] =$$
$$\mathbb{E}\left\{\begin{bmatrix} -u(k-1)y(k-1) & \cdots & -u(k-1)y(k-n_a) & u(k-1)u(k-1) & \cdots & u(k-1)u(k-n_b) \\ -u(k-2)y(k-1) & \cdots & -u(k-1)y(k-n_a) & u(k-2)u(k-1) & \cdots & u(k-2)u(k-n_b) \end{bmatrix}\right\}$$

and

$$\mathbb{E}[z(k)\tilde{\phi}(k)^T] =$$
$$\mathbb{E}\left\{\begin{bmatrix} -u(k-1)y_0(k-1) & \cdots & -u(k-1)y_0(k-n_a) & u(k-1)u(k-1) & \cdots & u(k-1)u(k-n_b) \\ -u(k-2)y_0(k-1) & \cdots & -u(k-1)y_0(k-n_a) & u(k-2)u(k-1) & \cdots & u(k-2)u(k-n_b) \end{bmatrix}\right\}.$$

At this point, note that for $k_2 \leq k_1$,

$$\mathbb{E}[u(k_1)y(k_2)] = \mathbb{E}[u(k_1)y_0(k_2)] = 0\,, \tag{9.1}$$

since both $y(k_2)$ and $y_0(k_2)$ do not depend on $u(k_1)$.

Hence, to prove that $\mathbb{E}[z(k)\phi(k)^T] = \mathbb{E}[z(k)\tilde{\phi}(k)^T]$ we are left with showing that $\mathbb{E}[u(k-2)y(k-1)] = \mathbb{E}[u(k-2)y_0(k-1)]$.

Expanding $\mathbb{E}[u(k-2)y_0(k-1)]$ we have

$$\begin{aligned} \mathbb{E}[u(k-2)y_0(k-1)] &= \mathbb{E}[-a_1u(k-2)y_0(k-2) - \cdots - a_{n_a}u(k-2)y_0(k-n_a-1) \\ &\quad + b_1u(k-2)u(k-2) + \cdots + b_{n_b}u(k-2)u(k-n_b-1)] \\ &= \mathbb{E}[b_1u(k-2)u(k-2) + \cdots + b_{n_b}u(k-2)u(k-n_b-1)]\,, \end{aligned}$$

where the last equality follows from (9.1).

Expanding $\mathbb{E}[u(k-2)y(k-1)]$ we have

$$\begin{aligned} \mathbb{E}[u(k-2)y(k-1)] &= \mathbb{E}[-a_1u(k-2)y(k-2) - \cdots - a_{n_a}u(k-2)y(k-n_a-1) \\ &\quad + b_1u(k-2)u(k-2) + \cdots + b_{n_b}u(k-2)u(k-n_b-1) + u(k-2)e(k-1)] \\ &= \mathbb{E}[b_1u(k-2)u(k-2) + \cdots + b_{n_b}u(k-2)u(k-n_b-1) + u(k-2)e(k-1)] \\ &= \mathbb{E}[b_1u(k-2)u(k-2) + \cdots + b_{n_b}u(k-2)u(k-n_b-1)]\,, \end{aligned}$$

where the second equality follows from (9.1) and the last equality since $\mathbb{E}[u(k-2)e(k-1)] = 0$, being the sequence $\{e(k)\}$ white noise and independent of the input.

Hence, $\mathbb{E}[u(k-2)y_0(k-1)] = \mathbb{E}[u(k-2)y_0(k-1)]$ as wanted.

Problem 2:

Consider the following system:

$$y(k) + ay(k-1) = b_1u(k-1) + b_2u(k-2) + v(k).$$

Parameters of the system should be estimated by using the instrumental variables method. To this end, it has been decided to use delayed inputs as instruments:

$$z(k) = [u(k-1)\, u(k-2)\, u(k-3)]^T.$$

Assuming that $u(k)$ is white noise with zero mean and unit variance and that it is uncorrelated with $v(k)$, find for which values of parameters a, b_1 and b_2, the instrumental variables are correlated with the regression variables (i.e. $\mathbb{E}[\{z(k)\varphi^T(k)\}$ is nonsingular).

Solution

We have that:

$$z(k) = [u(k-1), u(k-2), u(k-3)]^T,$$

and

$$\varphi(k) = [-y(k-1), u(k-1), u(k-2)]^T,$$

and therefore it follows that:

$$R = \mathbb{E}[\{z(k)\varphi^T(k)\} = \mathbb{E}[\left\{ \begin{bmatrix} -u(k-1)y(k-1) & u(k-1)u(k-1) & u(k-1)u(k-2) \\ -u(k-2)y(k-1) & u(k-2)u(k-1) & u(k-2)u(k-2) \\ -u(k-3)y(k-1) & u(k-3)u(k-1) & u(k-3)u(k-2) \end{bmatrix} \right\}$$

Since the signal u(k) is white noise signal with variance equal to 1, it holds that:

$$\mathbb{E}[\{u(k-1)u(k-2)\} = \mathbb{E}[\{u(k-2)u(k-1)\} = E\{u(k-3)u(k-2)\} = 0$$

$$\mathbb{E}[\{u(k-1)u(k-1)\} = \mathbb{E}[\{u(k-2)u(k-2)\} = 1.$$

In addition since $y(k-1)$ does not depend on $u(k-1)$, we have that:

$$\mathbb{E}[\{u(k-1)y(k-1)\} = 0.$$

Using the same rules, we also get:

$$\mathbb{E}[\{u(k-2)y(k-1)\} = \mathbb{E}[\{-au(k-2)y(k-2)+b_1u(k-2)u(k-2)+b_2u(k-2)u(k-3)+u(k-2)v(k-1)\} = b_1$$

$$\mathbb{E}[\{u(k-3)y(k-1)\} =$$
$$\mathbb{E}[\{-au(k-3)y(k-2) + b_1u(k-3)u(k-2) + b_2u(k-3)u(k-3) + u(k-3)v(k-1)\} =$$
$$\mathbb{E}[\{a^2u(k-3)y(k-3) - ab_1u(k-3)u(k-3) - ab_2u(k-3)u(k-4) - au(k-3)v(k-2)\} + b_2 = b_2 - ab_1.$$

Therefore we have that:

$$R = \begin{bmatrix} 0 & 1 & 0 \\ -b_1 & 0 & 1 \\ -(b_2 - ab_1) & 0 & 0 \end{bmatrix}.$$

Clearly R is nonsingular for $b_2 \neq ab_1$.

Matlab exercise:

The dataset 'Data_ex9.mat' needed for this exercise can be downloaded from the resources page on Piazza.

Consider data, $y(k)$ and $u(k)$, collected from the system,

$$y(k) = \frac{B(z)}{A(z)}u(k) + C(z)e(k), \quad e(k) \sim \mathcal{N}(0, \lambda).$$

The polynomials are of the form:

$$A(z) = 1 + a_1 z^{-1} + a_2 z^{-2}\,,$$
$$B(z) = b_1 z^{-1}\,,$$
$$C(z) = 1 + c_1 z^{-1}\,.$$

For the questions below, the following experimental data is provided:

- `ex9_u`: the input signal, $u(k)$, for $k = 1, 2, \ldots, K$,
- `ex9_y`: the corresponding output signal, $y(k)$, for $k = 1, 2, \ldots, K$.

The system is at rest and that there is no noise for $k \leq 0$,

$$u(k) = y(k) = e(k) = 0 \text{ for all } k \leq 0\,.$$

1. Using pseudo-linear regression (PLR) over the entire data, estimate the problem parameters $\hat{\theta}_{\text{PLR}} = [\hat{a}_1\ \hat{a}_2\ \hat{b}_1\ \hat{c}_1]^T$.
   (a) Formulate a pseudo-linear regression that, when solved, estimates the parameters $\hat{\theta}_{\text{PLR}}$. Express the one-step-ahead prediction estimator in terms of the solution to the pseudo-linear regression.
   (b) For the given input and output sequence, `ex9_u` and `ex9_y`, solve the above problem for the estimated parameters $\hat{\theta}_{\text{PLR}}$. Generate the resulting vector of prediction errors, $\epsilon(k) = y(k) - \hat{y}(k|\hat{\theta}_{\text{PLR}})$, $k = 1, \ldots, K$.
2. Use an instrumental variable (IV) method to estimate $\hat{\theta}_{\text{IV}} = [\hat{a}_1\ \hat{a}_2\ \hat{b}_1]^T$. Starting from a least-squares initialization, compute suitable instruments $\zeta(k)$.
   (a) Describe your choice of instrumental variables, $\zeta(k)$.
   (b) Estimate $\hat{\theta}_{\text{IV}}$ using your chosen $\zeta(k)$.

Solution Hints

1. (a) Estimate $\hat{\theta}_{PLR} = [a_1\ a_2\ b_1\ c_1]^T$ using a pseudo-linear regression A pseudo-linear regression should be of the form $y(k|k-1) = \varphi(\theta)\theta$. Let $w(k,\theta) = \frac{B(z)}{A(z)}u(k)$, with A and B of the form given in the problem statement. Then $w(k,\theta)(1 + a_1 z^{-1} + a_2 z^{-2}) = u(k) b_1 z^{-1}$. Refactoring, we get $w(k,\theta) = -a_1 w(k-1,\theta) - a_2 w(k-2,\theta) + b_1 u(k-1)$. Looking at the system output, we have $y(k) = w(k,\theta) + e(k)(1 + c_1 z^{-1})$. Letting $\hat{y}(k|k-1) = y(k) - e(k)$, the one-step-ahead prediction becomes:

$$\hat{y}(k|k-1) = w(k,\theta) + c_1 e(k-1) = [-w(k-1,\theta)\ \ -w(k-2,\theta)\ \ u(k-1)\ \ e(k-1)]\,[a_1\ a_2\ b_1\ c_1]^T$$

(b) This nonlinear equation can be solved for $[a_1\ a_2\ b_1\ c_1]^T$ using fmincon. Given a particular choice of θ_{PLR}, we can recursively compute $w(k-i,\theta) = -a_1 w(k-i-1,\theta) - a_2 w(k-i-2,\theta) + b_1 u(k-i)$, and let $e(k) = y(k) - y(k|k-1)$. The solution minimizes the norm of e, with the constraint that $y(k) - y(k|k-1) - e(k) = 0$. For this particular problem instance, $\hat{\theta}_{PLR} = [0.35601,\ \ 0.078556,\ \ 1.1147,\ \ 0.55409]^T$ The resulting error is plotted in Figure 1.

2. (a) Estimate $[a_1\ a_2\ b_1]^T$ using an instrumental variable approach: First, compute the least squares estimate of $[a_1\ a_2\ b_1]^T$ using linear regression: Assuming $y(k) = \frac{B(z)}{A(z)}u(k) + e(k)$, solve for the coefficients of $A(z) = 1 + a_1 z^{-1} + a_2 z^{-2}$ and $B(z) = b_1 z^{-1}$. Given y and u, estimate a_1, a_2, and $b_1 : y = [-y(k-1)\ \ -y(k-2)\ u(k-1)]\,[a_1\ a_2\ b_1]^T$ We find the least squares estimate for $[a_1\ a_2\ b_1]^T$ is $[-0.10333,\ \ -0.099327,\ \ 1.024]^T$. Next, create the instrumental variables by filtering prior input data using the least-squares. estimate of A(z) and B(z). Following the approach in the lecture, define the auxiliary signal $x(k) = \frac{P(z)}{Q(z)}u(k)$, where $P(z) = b_1 z^{-1}$ and $Q(z) = 1 + a_1 z^{-1} + a_2 z^{-2}$, with the coefficients a_1, a_2, and b_1 found using least squares regression in the previous step. Choose the instrumental variables as $\zeta(k) = [-x(k-1)\ \ -x(k-2)\ u(k-1)]$ and the regressor $\varphi(k) = [-y(k-1)\ \ -y(k-2)\ u(k-1)]$.

(b) We estimate the true parameters as:

$$\hat{\theta}^{IV} = \left(\frac{1}{K}\sum_{k=1}^{K}\xi(t)\varphi^T(t)\right)^{-1}\frac{1}{K}\sum_{k=1}^{K}(\xi(t)y(t))$$

Using the given data, we find $\hat{\theta}^{IV} = [0.31044,\ \ -0.03292,\ \ 1.0729]^T$.

Note that since we have used the same data to estimate the polynomials P and Q, as well as to obtain $\hat{\theta}^{IV}$, the resulting estimate may not be unbiased. This is because P and Q depend on y and therefore also on the noise, e. Hence, there can be a correlation between the instruments, x and e. In practice, however, this is probably very small.

Figure 1: One-step-ahead prediction error for pseudo-linear regression