1 System Identification and Linear Control

A scalar discrete-time system has the following dynamics:

$$x(t+1) = \lambda x[t] + g(u[t]),$$

where $g : \mathbb{R} \to \mathbb{R}$ not necessarily linear.

a) If g is approximated to order 2 around the operating point $u^* = 0$, so that

$$x(t+1) \approx \lambda x[t] + \beta_0 + \beta_1 u[t] + \beta_2 u^2[t],$$

what should β_0, β_1, and β_2 be?

Answer

From the MacLaurin series,

$$g(u) = g[0] + g'[0]u + \frac{1}{2}g''[0]u^2 + \left(\text{higher-order terms}\right).$$

We assume the argument of g to be small enough that its higher powers vanish.

b) Suppose that $x[0] = 0$. We apply a sequence of inputs

$$\left(u[0], u[1], \ldots, u[N-1]\right) \tag{1}$$

and observe states $x[1], x[2], \ldots, x[N]$. Derive the least-squares estimates of λ, β_0, β_1, and β_2.

Answer

Notice that we may write $N-1$ system recurrences simultaneously as follows:

$$\begin{aligned}
\begin{bmatrix} x[1] \\ x[2] \\ \vdots \\ x[N] \end{bmatrix} &= \begin{bmatrix} x[0] \\ x[1] \\ \vdots \\ x[N-1] \end{bmatrix} \lambda + \begin{bmatrix} 1 & u[0] & u^2[0] \\ 1 & u[1] & u^2[1] \\ \vdots & \vdots & \vdots \\ 1 & u[N-1] & u^2[N-1] \end{bmatrix} \begin{bmatrix} \beta_0 \\ \beta_1 \\ \beta_2 \end{bmatrix} \\
&= \begin{bmatrix} x[0] & 1 & u[0] & u^2[0] \\ x[1] & 1 & u[1] & u^2[1] \\ \vdots & \vdots & \vdots & \vdots \\ x[N-1] & 1 & u[N-1] & u^2(N-1) \end{bmatrix} \begin{bmatrix} \lambda \\ \beta_0 \\ \beta_1 \\ \beta_2 \end{bmatrix}
\end{aligned}$$

Give these vectors and matrices new names.

$$\begin{aligned}
\vec{y} &= D\vec{p} \\
\hat{\vec{p}} &= \left(D^\top D\right)^{-1} D^\top \vec{y}
\end{aligned}$$

2 System Identification

Let's now look at how System Identification works in the vector case. Again you are given an unknown discrete-time system. We don't know its specifics but we know that it takes one scalar input and as two observable states.
We would like to find a linear model of the form

$$\vec{x}[t+1] = A\vec{x}[t] + Bu[t] + \vec{w}[t],$$

where $\vec{w}[t]$ is an error term due to unseen distrubances and noise, $u[t]$ is a scalar input, and

$$A = \begin{bmatrix} a_0 & a_1 \\ a_2 & a_3 \end{bmatrix}, \quad B = \begin{bmatrix} b_0 \\ b_1 \end{bmatrix}, \quad \vec{x}[t] = \begin{bmatrix} x_0[t] \\ x_1[t] \end{bmatrix}.$$

To identify the system parameters from measured data, we need to find the unknowns: a_0, a_1, a_2, a_3, b_0 and b_1, however, you can only interact with the system via a blackbox model. The model allows you to view the states $\vec{x}[t] = \begin{bmatrix} x_0[t] & x_1[t] \end{bmatrix}^T$ and it takes a scalar input $u[t]$ that allows the system to move to the next state $\vec{x}[t+1] = \begin{bmatrix} x_0[t+1] & x_1[t+1] \end{bmatrix}^T$.

a) **Write scalar equations for the new states, $x_0[t+1]$ and $x_1[t+1]$ in terms of a_i, b_i, the states $x_0[t]$, $x_1[t]$, and the input $u[t]$.** Here, assume that $\vec{w}[t] = \vec{0}$ (i.e. the model is perfect).

Answer

$$x_0[t+1] = a_0 x_0[t] + a_1 x_1[t] + b_0 u[t]$$
$$x_1[t+1] = a_2 x_0[t] + a_3 x_1[t] + b_1 u[t]$$

b) Now we want to identify the system parameters. We observe the system at the initial state $\vec{x}[0] = \begin{bmatrix} x_0[0] \\ x_1[0] \end{bmatrix}$, input $u[0]$ and observe the next state $\vec{x}[1] = \begin{bmatrix} x_0[1] \\ x_1[1] \end{bmatrix}$. We can continue this for an m long sequence of inputs.

What is the minimum value of m you need to identify the system parameters?

Answer

There are 6 unknowns so you need 6 equations to properly identify the system. We get two equations from each time step, so in order to uniquely solve the system, we will need to give the system $m = 3$ inputs.

Notice that the initial condition on its own gives us no equations because the unknowns we are interested in do not impact the initial condition. They govern the evolution of the system, and hence the states at times $1, 2, 3$.

c) Say we feed in a total of 4 inputs $u[0], u[1], u[2], u[3]$ into our blackbox. This allows us to observe $x_0[0], x_0[1], x_0[2], x_0[3], x_0[4]$ and $x_1[0], x_1[1], x_1[2], x_1[3], x_1[4]$, which we can use to identify the system.

To identify the system we need to set up an approximate (because of potential disturbances) matrix equation

$$D\vec{p} \approx \vec{y}$$

using the observed values above and the unknown parameters we want to find. Suppose you are given the form of D in terms of some of the observed data:

$$D = \begin{bmatrix} x_0[0] & x_1[0] & u[0] & 0 & 0 & 0 \\ x_0[1] & x_1[1] & u[1] & 0 & 0 & 0 \\ x_0[2] & x_1[2] & u[2] & 0 & 0 & 0 \\ x_0[3] & x_1[3] & u[3] & 0 & 0 & 0 \\ 0 & 0 & 0 & x_0[0] & x_1[0] & u[0] \\ 0 & 0 & 0 & x_0[1] & x_1[1] & u[1] \\ 0 & 0 & 0 & x_0[2] & x_1[2] & u[2] \\ 0 & 0 & 0 & x_0[3] & x_1[3] & u[3] \end{bmatrix}.$$

For this D, **what are $\vec{y}$ and the unknowns $\vec{p}$ so that $D\vec{p} \approx \vec{y}$ makes sense?** Tell us what the components of these vectors are, written in vector form.

Answer

There are two choices, both valid. Notice that the matrix D has a symmetry to it — the same block of data sits in two places.

Option 1:

$$\vec{y} = \begin{bmatrix} x_0[1] \\ x_0[2] \\ x_0[3] \\ x_0[4] \\ x_1[1] \\ x_1[2] \\ x_1[3] \\ x_1[4] \end{bmatrix} \qquad \vec{p} = \begin{bmatrix} a_0 \\ a_1 \\ b_0 \\ a_2 \\ a_3 \\ b_1 \end{bmatrix} \tag{2}$$

The flipped version is equally valid:

$$\vec{y} = \begin{bmatrix} x_1[1] \\ x_1[2] \\ x_1[3] \\ x_1[4] \\ x_0[1] \\ x_0[2] \\ x_0[3] \\ x_0[4] \end{bmatrix} \qquad \vec{p} = \begin{bmatrix} a_2 \\ a_3 \\ b_1 \\ a_0 \\ a_1 \\ b_0 \end{bmatrix} \tag{3}$$

This way, the equations come out exactly as they were in the first part of this problem.

d) Now that we have set up $D\vec{p} \approx \vec{y}$, **explain how you would use this approximate equation to estimate the unknown values a_0, a_1, a_2, a_3, b_0 and b_1 assuming the columns of D are linearly independent.**

Answer

Using the equation above we realize that we need to solve for $\vec{p}$ to learn the system. Since the matrix D is not invertible we can use the standard least squares formula from 16A

$$\vec{p}^{\,*} = (D^T D)^{-1} D^T \vec{y}$$

to find the unknown values.

This is valid because we have assumed that the columns of D are linearly independent. If the columns of D are linearly-independent, then D has no nontrivial nullspace by definition. Moreover, we know that $D^T D$ has the same nullspace as D, and thus we may invert $D^T D$.

e) **What could go wrong in the previous case? What kind of inputs would make least-squares fail to give you the parameters you want?**

Answer

In the previous parts we assumed that the columns of D were linearly independent. However, it is possible that $D^T D$ might not be invertible, which would cause our least-squares formulation to fail. This could happen if one or more of the outputs were directly proportional to the input, e.g. if $\vec{x}_0 = \alpha \vec{u}$. Some example parameters where this might happen:

$$a_0 = a_1 = 0, \ b_0 = 1$$

Hence,

$$x_0[t+1] = u[t]$$

Picking $\vec{u}[t] = 1$, then $\vec{x}_0[t] = 1$ also.