Observability matrix derivation for nonlinear systems

written by vishalcr on Functor Network
original link: https://functor.network/user/3019/entry/1130

1 Problem formulation

A nonlinear state space system given by

$$\begin{aligned}\dot{\mathbf{x}} &= f(\mathbf{x}, \mathbf{u}) \\ \mathbf{y} &= h(\mathbf{x})\end{aligned} \tag{1}$$

where $\mathbf{x} \in \mathbb{R}^n$ is the state vector, $\mathbf{u} \in \mathbb{R}^d$ is the *known* input vector and $\mathbf{y} \in \mathbb{R}^p$ is the measurement vector. The functions $f(\cdot)$ and $h(\cdot)$ are known as the state dynamics and measurement models, respectively. One definition of observability states that given sufficient measurements $\mathbf{y}$ and known inputs $\mathbf{u}$, it is possible to uniquely determine the initial state $\mathbf{x}_0 \equiv \mathbf{x}(t = 0)$. In order to arrive at the derivation for the nonlinear system observability condition, let us track back to obtain the observability condition for linear systems.

2 Observability of a linear state space system

If the state dynamics and measurement model functions in (1) are linear, then the system in (1) can be expressed as

$$\begin{aligned}\dot{\mathbf{x}} &= A\mathbf{x} + B\mathbf{u} \\ \mathbf{y} &= C\mathbf{x}\end{aligned} \tag{2}$$

where A, B and C are the state transition, input and measurement matrices, respectively. They are also assumed to be time invariant, i.e., their derivatives with respect to time are 0.

Applying the first derivative in time to the measurement equation in (2), we get

$$\dot{\mathbf{y}} = C\dot{\mathbf{x}} \tag{3}$$

Substituting for $\dot{\mathbf{x}}$ in (3) from (2)

$$\dot{\mathbf{y}} = CA\mathbf{x} + CB\mathbf{u} \tag{4}$$

Applying the second order derivative on both sides of (4), and again substituting from (2), we get

$$\begin{aligned}\ddot{\mathbf{y}} &= CA\dot{\mathbf{x}} + CB\dot{\mathbf{u}} \\ &= CA^2\mathbf{x} + CAB\mathbf{u} + CB\dot{\mathbf{u}}\end{aligned} \tag{5}$$

As $\mathbf{u}$ and its derivatives with respect to time are assumed known, absorbing the terms related to $\mathbf{u}$ and its derivates into the left hand side and simplifying the notations we get the following system

$$\begin{aligned}
\overline{\mathbf{y}}^{(0)} &= CA\mathbf{x} \\
\overline{\mathbf{y}}^{(1)} &= CA^2\mathbf{x} \\
&\vdots \\
\overline{\mathbf{y}}^{(n-1)} &= CA^{n-1}\mathbf{x}
\end{aligned} \tag{6}$$

where $\overline{\mathbf{y}}^{(n-1)}$ denotes the $(n-1)^{th}$ order derivative or $\dot{\overline{\mathbf{y}}}^{(n-2)}$.

Expressing (6) in matrix form, we have

$$\mathcal{Y} = \mathcal{O}\mathbf{x} \tag{7}$$

From (7) it is clear that, in order for a unique solution to exist for $\mathbf{x}$, the *observability matrix* $\mathcal{O} \in \mathbb{R}^{(np\times n)}$ needs to be full rank, i.e., rank needs to be n.

3 Observability of a nonlinear state space system

From the previous section, it is clear that in order to derive the observability condition of the nonlinear state space system in (1), we have to obtain the derivatives of $\mathbf{y}$. Consider the nonlinear measurement model

$$\mathbf{y} = h(\mathbf{x}) \tag{8}$$

Applying chain rule we get

$$\begin{aligned}
\dot{\mathbf{y}} &= \frac{\partial h}{\partial \mathbf{x}}\dot{x} \\
&= \frac{\partial h}{\partial \mathbf{x}} f(\mathbf{x}, \mathbf{u})
\end{aligned} \tag{9}$$

where the representation of $\dot{\mathbf{x}}$ has been substituted from (1). The second representation in (8) can also be described using the concept of Lie derivatives from differential geometry as follows

$$\mathcal{L}_f^{(1)} h = \frac{\partial h}{\partial \mathbf{x}} f(\mathbf{x}, \mathbf{u}) \tag{10}$$

where $\mathcal{L}_f^{(1)} h$ is the first order Lie derivative of the function h with respect to another function f. As a result the first order derivative of the measurement $\mathbf{y}$ in (9) can now be expressed as

$$\dot{\mathbf{y}} = \mathcal{L}_f^{(1)} h \tag{11}$$

Taking the second order derivative of the measurement vector $\mathbf{y}$ in (9), we get

$$\begin{aligned}\ddot{\mathbf{y}} &= \frac{\partial \dot{\mathbf{y}}}{\partial \mathbf{x}} \frac{\partial \mathbf{x}}{\partial t} \\ &= \frac{\partial}{\partial \mathbf{x}} \left\{ \mathcal{L}_f^{(1)} h \right\} f(\mathbf{x}, \mathbf{u}) \\ &= \mathcal{L}_f^{(2)} h \end{aligned} \tag{12}$$

Stacking up the derivatives of $\mathbf{y}$ up to the $(n-1)^{th}$ order we get

$$\begin{aligned}\mathbf{y}^{(0)} &= \mathcal{L}_f^{(0)} h \\ \mathbf{y}^{(1)} &= \mathcal{L}_f^{(1)} h \\ &\vdots \\ \mathbf{y}^{(n-1)} &= \mathcal{L}_f^{(n-1)} h \end{aligned} \tag{13}$$

where $\mathcal{L}_f^{(0)} h = h$ and $\mathbf{y}^{(0)} = \mathbf{y}$. Expressing (13) in matrix form, we get

$$\mathcal{Y} = \mathbf{L}_f h \tag{14}$$

It is straightforward to show that, if

$$\begin{aligned} f(\mathbf{x}, \mathbf{u}) &= A\mathbf{x} + B\mathbf{u} \\ h(\mathbf{x}) &= C\mathbf{x} \end{aligned} \tag{15}$$

then (14), reduces to (7). From (7), applying a partial derivative with respective to $\mathbf{x}$ to both sides, we get

$$\frac{\partial \mathcal{Y}}{\partial \mathbf{x}} = \mathcal{O} \tag{16}$$

Applying the same operation to both sides of (14) and using the definition of the observability matrix arrived at in (16), we get

$$\mathcal{O} = \frac{\partial \mathbf{L}_f h}{\partial \mathbf{x}} \tag{17}$$

Expressing (17) in vector form

$$\mathcal{O} = \begin{bmatrix} \frac{\partial \mathcal{L}_f^{(0)} h}{\partial \mathbf{x}} \\ \frac{\partial \mathcal{L}_f^{(1)} h}{\partial \mathbf{x}} \\ \vdots \\ \frac{\partial \mathcal{L}_f^{(n-1)} h}{\partial \mathbf{x}} \end{bmatrix} \tag{18}$$

where $\mathcal{O} \in \mathbb{R}^{np \times n}$. As has been shown in the section on linear observability, in a similar manner, for a nonlinear system (1) to be observable, the observability matrix defined in (18) has to be full rank.