STATE-SPACE MODELLING

Introduction: Over the past decade or so there has been an ever increasing use of state-space modelling in control books and applications. Although state-space is often considered to be a complex representation, it is in the end simply a method for describing the behaviour of dynamic systems.

The growing popularity of state-space modelling can be attributed to the fact that it has a very general form and is able to represent linear, non-linear, time-invariant and time-varying dynamics in a relatively compact form. Furthermore, in many application areas the state-space approach is very intuitive, allowing rapid development of dynamic models from first principles.

This article will illustrate the ideas that underpin state-space analysis and illustrate them with a straightforward example.

The Basics, and an Example: In the process industries, control theory has typically been based on dynamic models that are represented as Ordinary Differential Equations (ODEs) and/or transfer functions. As an example, consider the system illustrated in Figure 1. This system is used in many text-books and concerns the relationship between the flow of liquid into the tank, \( q_{in} \), and the level of liquid, \( h \). From this model a feedback control system can be designed to maintain the level of liquid in the tank.

![Figure 1: Liquid Level System](image)

The flow of liquid out of the tank is proportional to level of liquid \( h \) if the flow is laminar:

\[ q_{out}(t) = kh(t) \]

where \( k \) is a constant parameter.

A mass balance across the tank gives:

\[ q_{in}(t) - q_{out}(t) = \frac{dV(t)}{dt} \]

i.e. the rate of change of liquid volume, \( V \), is equal to the volumetric flow rate of liquid in minus the flow rate of liquid out.

The mass balance can be re-arranged to give:

\[ q_{in}(t) - q_{out}(t) = \frac{dA_r h(t)}{dt} \]

and since \( A_r \) is constant:

\[ q_{in}(t) - q_{out}(t) = A_r \frac{dh(t)}{dt} . \]

Replacing \( q_{out}(t) \) by \( kh(t) \) gives:

\[ q_{in}(t) - kh(t) = A_r \frac{dh(t)}{dt} \]

and more re-arrangement gives:

\[ \frac{dh(t)}{dt} + \frac{k}{A_r} h(t) = \frac{1}{A_r} q_{in}(t) \]  

Equation (1) describes a linear, first order relationship between the flow into the tank and the liquid level. If zero initial conditions are assumed, then this equation can be transformed into the Laplace domain as follows:

\[ sH(s) + \frac{k}{A_r} H(s) = \frac{1}{A_r} Q_{in}(s) \]

where \( H(s) \) and \( Q_{in}(s) \) are the Laplace transforms of the time trends \( h(t) \) and \( q_{in}(t) \).

For instance, if \( q_{in}(t) \) is a step change of magnitude \( 0 \), then \( Q_{in}(s) = q_0 / s \).

After some algebra, the transfer function for the tank is:

\[ \frac{H(s)}{Q_{in}(s)} = \frac{1}{A_r s + k} = \frac{1/k}{(A_r/k)s + 1} \]  

(2)

The generic form of a first order transfer function is \( \frac{K_p}{s \tau + 1} \) where \( K_p \) is the system gain and \( \tau \) is the time constant, hence by comparison, the gain of the tank system is \( 1/k \) and its time constant is \( A_r/k \).

So far, the analysis has involved an ODE with one dependent variable \( h(t) \), and one independent variable \( q_{in}(t) \). Application of the Laplace transform to the ODE gave a single-input-single-output transfer function showing the gain and time constant. Next, in anticipation of more complicated systems with more than one dependent variable, let’s examine a more general formulation of the same equation. Equation (1) is re-considered and a new
variable, \( x_1(t) \) is now defined, such that \( x_1(t) = h(t) \).

Equation (1) can be re-written as:

\[
\frac{dx_1(t)}{dt} = -\frac{k}{A_T}x_1(t) + \frac{1}{A_T}q_{in}(t)
\]

which in a more general form can be expressed as a dynamic state equation:

\[
\frac{dx(t)}{dt} = \dot{x}(t) = Ax(t) + Bu(t)
\]

where:
\[
x(t) = x_1(t), \quad u(t) = q_{in}(t), \quad A = -\frac{k}{A_T} \quad \text{and} \quad B = \frac{1}{A_T}
\]

Equation (3) is the general, linear state equation, where the vector \( x \) represents the state variables of the process. It anticipates more complex systems by use of vectors and matrices. In the tank example, however, there is a single state variable \( h(t) \) and a single input variable \( q_{in}(t) \) so the \( A \) and \( B \) matrices are scalars.

State variables are the smallest subset of system variables that describe the entire dynamic characteristics of the system. They can be thought of as internal elements of the system that are related to, or in some case are actually equal to, the output variables. Although it can sometimes be beneficial, it is not necessary for the state variables to be measured or even have physical meaning.

In the even more general non-linear case the state equation is defined as:

\[
\frac{dx(t)}{dt} = f(x,u,t)
\]

where \( f \) is a non-linear function.

The coupled two tank system has two physical states, \( h_1 \) and \( h_2 \) and, if the tanks each have unit cross sectional area (i.e. \( A_T = 1 \)), the state equation is:

\[
\begin{bmatrix}
\frac{dh_1}{dt} \\
\frac{dh_2}{dt}
\end{bmatrix} = \begin{bmatrix}
k_2 & k_2 \\
-(k_2 + k_3) & k_2
\end{bmatrix} \begin{bmatrix} h_1(t) \\
q_{in}
\end{bmatrix} + \begin{bmatrix} 1 \\
0
\end{bmatrix} q_{in}
\]

Expanding out the matrices and rearranging shows that the state equation is nothing more than a volumetric balance for each tank.

\[
\begin{align*}
\frac{dh_1}{dt} &= -k_2(h_1 - h_2) + q_{in} \\
\frac{dh_2}{dt} &= k_2(h_1 - h_2) - k_3 h_2
\end{align*}
\]

Thus, there is no new physics in a state equation. The reason why control engineers use it is that is a convenient way of studying the mathematical properties of the physical equations. Using state-space leads to some very powerful controller designs that would be too cumbersome to consider without the compact matrix formulation.

Referring back to the general state equation:

\[
\frac{dx(t)}{dt} = \dot{x}(t) = Ax(t) + Bu(t)
\]

In the two-tank example, the state vector \( x(t) \) is \( \begin{bmatrix} h_1 \\
h_2
\end{bmatrix} \), the input \( u(t) \) is \( q_{in} \) and the constant-coefficient \( A \) and \( B \) matrices are:

\[
A = \begin{bmatrix}
k_2 & k_2 \\
-(k_2 + k_3) & k_2
\end{bmatrix} \quad \text{and} \quad B = \begin{bmatrix} 1 \\
0
\end{bmatrix}
\]

Observer Equation: The state equation (3) describes how the states vary with time. In the case of the examples above, the output variables were the liquid levels, which were also defined as the state variables. However, in many examples, the state variables will be different from the output variables. It is therefore necessary to have another equation which describes the relationship between the process outputs and state variables. For general systems, this equation is defined as follows:

\[
y(t) = g(x,u,t)
\]

where \( g \) is a non-linear function. For a linear, time-invariant system this equation always takes the form:

\[
y(t) = Cx(t) + Du(t)
\]

where \( y(t) \) is a vector of output variables. In the two-tank example:

\[
C = \begin{bmatrix} 1 \\
0 
\end{bmatrix} \quad \text{and} \quad D = \begin{bmatrix} 0 \\
0
\end{bmatrix}
\]
The fact \( C \) is a unit matrix shows the output variables are equal to the state variables in the two tank example.

The above equation for \( y(t) \) is referred to as the **observer equation**. The combination of the state equation (3) and observer equation (4) is called a state-space model.

**Another state-space form:** For any dynamic system, there is no unique state-space model. In fact there is an infinite number of state-space equations that can be developed, depending on how the states are defined. As an example, the physical balance equations for the two tank system naturally yielded two simultaneous first order differential equations. But another way to rearrange the equations is to eliminate either \( h_1 \) or \( h_2 \) to give a single second order differential equation. For instance, \( h_1 \) may be eliminated to give an equation for \( h_2 \) as follows:

Adding the two mass balances:

\[
\frac{dh_1}{dt} = -k_2 (h_1 - h_2) + q_{in}
\]
\[
\frac{dh_2}{dt} = k_2 (h_1 - h_2) - k_3 h_2
\]

gives

\[
\frac{dh_1}{dt} + \frac{dh_2}{dt} = q_{in} - k_3 h_2
\]

and, by rearrangement:

\[
\frac{dh_2}{dt} = q_{in} - k_3 h_2 - \frac{dh_1}{dt}
\]

The second mass balance can be differentiated to give:

\[
\frac{d^2 h_2}{dt^2} = k_2 \left( \frac{dh_1}{dt} \right) - k_3 \frac{dh_2}{dt}
\]

and now \( \frac{dh_1}{dt} \) can be replaced to give:

\[
\frac{d^2 h_2}{dt^2} = k_2 \left( q_{in} - k_3 h_2 - \frac{dh_2}{dt} \right) - k_3 \frac{dh_2}{dt}
\]

and finally:

\[
\frac{d^2 h_2}{dt^2} + (k_3 + 2k_2) \frac{dh_2}{dt} + k_2 k_3 h_2 = k_2 q_{in} \quad (5)
\]

Again, there is no new physics, but now the focus is on \( h_2 \) which obeys a second order differential equation. Now, the states can be represented not as \( h_1 \) and \( h_2 \), but as \( h_2 \) and \( \frac{dh_2}{dt} \). Therefore the state vector would be:

\[
x(t) = \begin{pmatrix} x_1(t) \\ x_2(t) \end{pmatrix} = \begin{pmatrix} h_2(t) \\ \frac{dh_2}{dt} \end{pmatrix}
\]

Clearly (by definition) one of the state equations is

\[
\frac{dx_1}{dt} = x_2(t)
\]

and the second comes from the rearrangement of equation (5):

\[
\frac{d^2 h_2}{dt^2} = -(k_3 + 2k_2) \frac{dh_2}{dt} - k_2 k_3 h_2 + k_2 q_{in} \quad (6)
\]

Using \( h_2(t) = x_1(t) \) and \( \frac{dh_2}{dt} = x_2(t) \) gives:

\[
\frac{dx_2}{dt} = -(k_3 + 2k_2) x_2(t) - k_2 k_3 x_1(t) + k_2 q_{in}
\]

Hence the state equations are:

\[
\begin{align*}
\frac{dx_1}{dt} &= x_2(t) \\
\frac{dx_2}{dt} &= -k_2 k_3 x_1(t) - (k_3 + 2k_2) x_2(t) + k_2 q_{in}
\end{align*}
\]

and they can be expressed in matrix form as:

\[
\begin{pmatrix} \frac{dx_1}{dt} \\ \frac{dx_2}{dt} \end{pmatrix} = \begin{pmatrix} 0 & 1 \\ -k_2 k_3 & -(k_3 + 2k_2) \end{pmatrix} \begin{pmatrix} x_1(t) \\ x_2(t) \end{pmatrix} + \begin{pmatrix} k_2 \\ 0 \end{pmatrix} q_{in}
\]

This way of describing a state-space is referred to as the **controllable canonical form**. There are however many other ways in which the \( A, B, C \) and \( D \) matrices can be re-arranged to produce, for example, the **observable canonical form**.

**Relationship Between State-Space, ODEs and Transfer Functions:** Process control engineers tend to be familiar with ODEs and the transfer function derived from the Laplace transform of the ODE. As shown in the previous examples state-space does not describe anything new, it is simply a re-arrangement of the dynamic equations into a form that is more compact and convenient for control system development and analysis. Therefore it is useful to see how the coefficients in the ODE relate to the parameters of the transfer function and the elements in the state-space \( A, B, C \) and \( D \) matrices.

The ODE in the general case is a single input, single output, high order dynamic system described by the following differential equation:

\[
\frac{d^n y(t)}{dt^n} + a_{n-1} \frac{d^{n-1} y(t)}{dt^{n-1}} + \cdots + a_1 \frac{dy(t)}{dt} + a_0 y(t) = b_m \frac{d^m u(t)}{dt^m} + \cdots + b_0 u(t)
\]
The transfer function for this system, derived by applying Laplace transforms with zero initial condition is:
\[ G(s) = \frac{Y(s)}{U(s)} = \frac{s^m + b_{m-1}s^{m-1} + \ldots + b_1s + b_0}{s^n + a_{n-1}s^{n-1} + \ldots + a_1s + a_0} \]
and the equivalent state-space equation in controllable canonical form is as follows:
\[
A = \begin{bmatrix}
0 & 1 & 0 & \ldots & 0 \\
0 & 0 & 1 & \ldots & 0 \\
\vdots & \vdots & \vdots & \ddots & \vdots \\
0 & 0 & 0 & \ldots & 1 \\
-a_0 & -a_1 & -a_2 & \ldots & -a_{n-1}
\end{bmatrix},
B = \begin{bmatrix}
0 \\
0 \\
\vdots \\
0 \\
1
\end{bmatrix}
\]
\[
C = \begin{bmatrix}
b_0 & b_1 & \ldots & b_{m-1}
\end{bmatrix} \text{ and } D = 0
\]
As can be seen, the same coefficients appear in all three forms indicating that they are three different ways of looking at the same equations.

**Relationship with step-response models:** The dynamics for electrical and mechanical systems are often reasonably well defined and can be derived from a physical understanding of the system. In contrast, process systems may not be understood with sufficient accuracy to develop a model from first principles and it is typical for step or other tests to be performed to better understand the dynamics. This section will demonstrate that the results of step response tests can be used to provide both transfer function models and a state-space model.

Figure 3 shows step tests for a two-tank system in which the inlet flow rate was doubled at time 20 sec. The levels \( h_1 \) and \( h_2 \) both changed. Level \( h_1 \) shows a response with a steep initial gradient that is characteristic of a first order system. Figure 3 indicates that the gain is 12 because the unit step in \( q_{in} \) produces a change in level of 12 on the \( h_1 \) scale. The time constant is approximately 12 seconds, and hence:
\[ \frac{H_1(s)}{Q_{in}(s)} \approx \frac{12}{12s + 1} \]

This transfer function is approximate, because although the step response looks first order there is in practice a back-pressure on the first tank from the second tank. The dynamics are actually of a higher order, however the effect is small and was not detected by the step test.

The level in the second tank, \( h_2 \), has the S-shaped response that is characteristic of a second order system. It requires a numerical method to fit a transfer function model to such step response data, for instance by using the ARX command in Matlab. The results from system identification (to one s.f.) were:
\[ \frac{H_2(s)}{Q_{in}(s)} = \frac{0.1}{s^2 + 0.7s + 0.05} \]

If both levels are now considered, the complete transfer function model for this system is:
\[ \frac{H_1(s)}{Q_{in}(s)} = \frac{12}{12s + 1}, \quad \frac{H_2(s)}{Q_{in}(s)} = \frac{0.1}{s^2 + 0.7s + 0.05} \]

An equivalent state-space equation is as follows:
\[
\dot{x}(t) = Ax(t) + Bu(t),
\]
\[ y(t) = Cx(t) + Du(t) \]

where:
\[
x(t) = \begin{bmatrix} h_1(t) \\ h_2(t) \end{bmatrix}, \quad y(t) = \begin{bmatrix} h_1(t) \\ h_2(t) \end{bmatrix}, \quad u(t) = q_{in}(t)
\]

and
\[
A = \begin{bmatrix}
-1/12 & 0 & 0 \\
0 & 0 & 1 \\
0 & -0.05 & -0.7
\end{bmatrix}, \quad B = \begin{bmatrix} 1 \\ 0 \\ 0.1
\end{bmatrix}, \quad C = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \end{bmatrix}, \quad D = \begin{bmatrix} 0 \\ 0 \end{bmatrix}
\]

The transfer functions and state-space models have the same input-output behaviour, again illustrating that the state-space form does not include any new physical information. However, as is beginning to become apparent from this example, with more input and output variables the transfer function model quickly becomes unwieldy. In contrast, the complexity of the state-space representation does not really change. The state-space model always consists of two compact equations, one is the dynamic state equation and the other is the algebraic observer.
equation. The only change as the system becomes more complex is that the dimensions of the matrices increase.

There are three states in the above model, which effectively combine the states $h_1(t)$ and $h_2(t)$ from the model based on volumetric balancing with the states $h_2(t)$ and $dh_2/dt$ that were present in the controllable canonical form. It can take some judgement to select the best states for a given purpose, however the ability to make good choices grows with experience. There are also some useful algorithms such as sub-space identification which can help.

Comments and Conclusions: The article has shown how to explore and investigate the physical equations of a system by means of its internal states. The states may have a direct physical meaning, or they may be abstract ones such as when the controllable canonical form is used. There are many reasons for making the states of a system explicit in the mathematical formulation. For instance, if the states can be estimated from process measurements, then there is an opportunity to use state feedback. Rate feedback, where the rate of change of the controlled variable is used as a feedback signal, is an example of state feedback. Here is a short list of some of the other things that working with state-space makes possible:

1. Provides a model structure from which feedback control systems can be designed with relative ease.
2. Enables closed-loop characteristics, such as robustness and stability, to be analysed and considered during control design.

The complexity of the mathematics involved for this is such that it would not be practical to apply it to other modelling formats.

3. In situations where the output variables are not directly measured, the observer equation can be used to estimate these measurements from the state variables. Despite their advantages, there is a very serious weakness to state-space models. When using transfer models, particularly first and second order, it is straightforward to visualize the step-response of the system. Unfortunately, this is not the case with state-space models which are more complicated to interpret and typically require the use of a computer to simulate step responses and evaluate the system time constants.

On the other hand, it is quite easy to obtain a state-space model from a step-response, as shown in this article. The example considered was straightforward, however more realistic cases require care and judgement in selection of the states, and optimal state-space models may not result. Methods for obtaining good state-space models for complex systems using pseudo random binary input sequences or step-response data will be considered in a future article on sub-space identification.

Further reading. Those who are interested to dig deeper might like to visit the URL below:
http://en.wikipedia.org/wiki/State_space_(controls)