2.0 STATE VARIABLES

2.1 Introduction

Since the vast majority of processes or systems that might employ Kalman filters are dynamic, we will use differential or difference equations to describe them. We will use differential equations for continuous time, or analog, systems and difference equations for discrete time, or digital, systems.

2.2 Continuous_Time Systems

2.2.1 State Equations From A Physical Description of the System

Consider the simple spring-mass-damper system illustrated in Figure 2-1. The differential equation for this system is

\[ \ddot{z}(t) + \frac{B}{M} \dot{z}(t) + \frac{K}{M} z(t) = \frac{1}{M} f(t) \]  

(2-1)

where \( \dot{z}(t) = \frac{dz(t)}{dt} \) and \( \ddot{z}(t) = \frac{d^2z(t)}{dt^2} \).

Figure 2-1. Spring-Mass-Damper System

Recall, from mechanics, that the two independent quantities of interest in Equation 2-1 are the position, \( z(t) \), and velocity, \( \dot{z}(t) \), of the mass. These quantities we will call the states of the system. Thus, the spring-mass-damper system has two states: \( x_1(t) = z(t) \) and \( x_2(t) = \dot{z}(t) \). With a few manipulations we can rewrite Equation 2-1 in terms of two equation involving \( x_1(t) \) and \( x_2(t) \). Specifically

\[ x_1(t) = x_2(t) \]
\[ \dot{x}_2(t) = -\frac{B}{M} x_2(t) - \frac{K}{M} x_1(t) + \frac{1}{M} f(t) \]  

(2-2)
The two equations of Equation 2-2 are what is known as a state variable representation of the system of Figure 2-1. Some interesting points about this representation are:

- The differential equations are first order.
- The differential equations are written in terms of the system states and the input.
- There are as many differential equations as there are states.

It should be noted that the set of state variables selected in this example are not the only ones that could have been chosen. Another equally valid (although more complex) set of state variables could have been \( x_1(t) = z(t) - \dot{z}(t) \) and \( x_2(t) = z(t) + \dot{z}(t) \). As an interesting exercise, the reader should find the state variable representation with these state variables (see the exercises at the end of this chapter).

The particular set of states used in Equation 2-2 is appealing because the states represent physical parameters of the system. That is, \( x_1(t) \) represents position and \( x_2(t) \) represents velocity. Usually this is a desirable selection of states. However, as was indicated above this is not the only selection possible.

The significance of the state variable representation in Kalman filtering becomes evident when we realize that the Kalman filter will be used to estimate the states of the system. This further implies that the state variable representation should contain those states that will be estimated by the Kalman filter.

Equation 2-2 can be written in a more compact form using matrix notation. Specifically,

\[
\begin{bmatrix}
\dot{x}_1(t) \\
\dot{x}_2(t)
\end{bmatrix} = \begin{bmatrix}
0 & 1 \\
-\frac{K}{M} & -\frac{B}{M}
\end{bmatrix} \begin{bmatrix}
x_1(t) \\
x_2(t)
\end{bmatrix} + \begin{bmatrix}
0 \\
\frac{1}{M}
\end{bmatrix} f(t)
\]

or

\[
\dot{x}(t) = Ax(t) + bu(t)
\]

where

\[
\dot{x}(t) = \begin{bmatrix}
\dot{x}_1(t) \\
\dot{x}_2(t)
\end{bmatrix};
\quad x(t) = \begin{bmatrix}
x_1(t) \\
x_2(t)
\end{bmatrix};
\quad A = \begin{bmatrix}
0 & 1 \\
-\frac{K}{M} & -\frac{B}{M}
\end{bmatrix};
\quad b = \begin{bmatrix}
0 \\
\frac{1}{M}
\end{bmatrix};
\quad u(t) = f(t).
\]

Because of its compact form, we will generally use the matrix form of Equation 2-4.

Although we have devised a means of representing the system of Figure 2-1 we have not discussed how we observe the system. Let us assume that the only measurable output of the system is position, \( x_1(t) \). That is, we are capable...
of measuring the position of the mass at any time, but not its velocity. Thus, the system output (or measurable, or observable) is

\[ y(t) = x_1(t) \]  \hspace{1cm} (2-5)

or, in matrix form

\[ y(t) = \begin{bmatrix} 1 & 0 \end{bmatrix} \begin{bmatrix} x_1(t) \\ x_2(t) \end{bmatrix} = Cx(t) \]  \hspace{1cm} (2-6)

where the definitions of \( C \) and \( x(t) \) are obvious.

With Equations 2-4 and 2-5 we have constructed a model of the spring-mass-damper system (or process) and its measurement. At this point we have not included uncertainties; we will consider them later.

### 2.2.2 State Equations From Block or Flow Diagrams

The state variable equations above were derived from differential equations of the system. Another common source of these equations is the block, or flow, diagram of the system, when such a diagram is available. As an example of how to determine the state variable equation from a block diagram, consider Figure 2-2. In this figure \( a \) and \( b \) are constants and \( \int \) represents integration. This system has three states, \( x_1(t), x_2(t) \) and \( x_3(t) \); two inputs \( u_1(t) \) and \( u_2(t) \); and two outputs \( y_1(t) \) and \( y_2(t) \). The state variable equations are determined by expressing \( \dot{x}_1(t), \dot{x}_2(t) \) and \( \dot{x}_3(t) \) in terms of \( x_1(t), x_2(t), x_3(t) \), \( u_1(t) \) and \( u_2(t) \) as

\[
\begin{align*}
\dot{x}_1(t) &= -x_3(t) + u_1(t) \\
\dot{x}_2(t) &= x_1(t) - ax_2(t) + u_2(t) \\
\dot{x}_3(t) &= x_2(t) - bx_3(t)
\end{align*}
\]  \hspace{1cm} (2-7)

and

\[
\begin{align*}
y_1(t) &= x_2(t) \\
y_2(t) &= x_3(t)
\end{align*}
\]  \hspace{1cm} (2-8)

or, in matrix form

\[
\begin{align*}
\dot{x}(t) &= Ax(t) + Bu(t) \\
y(t) &= Cx(t)
\end{align*}
\]  \hspace{1cm} (2-9)

The reader should determine the form of the matrices \( A, B \) and \( C \).
2.2.3 Solving the Differential Equations

As a final note before leaving the continuous-time system state variable notation, we discuss the solution of Equation 2-9 for the states. Consider the first order differential equation

$$\dot{x}(t) = ax(t) + bu(t). \quad (2-10)$$

From elementary differential equations we can write the solution to Equation 2-10 as

$$x(t) = e^{at} x(t_0) + \int_{t_0}^{t} e^{a(t-\tau)}bu(\tau)\,d\tau. \quad (2-11)$$

In Equation 2-11 the input is applied at $t = t_0$ and $x(t_0)$ is the value of the state at $t = t_0$. To verify that Equation 2-11 is a solution to Equation 2-10 we differentiate Equation 2-11 using Leibniz’s rule to get

$$\dot{x}(t) = ae^{at} x(t_0) + a\int_{t_0}^{t} e^{a(t-\tau)}bu(\tau)\,d\tau + bu(t). \quad (2-12)$$

or

$$\dot{x}(t) = ax(t) + bu(t) \quad (2-13)$$

which is the same as Equation 2-10. Thus, Equation 2-11 is indeed a solution to Equation 2-10.

If we recognize that the matrix differential equation of Equation 2-9 is first order and similar to Equation 2-10, we can heuristically extend the form of Equation 2-11 to yield

$$x(t) = e^{At} x(t_0) + \int_{t_0}^{t} e^{A(t-\tau)}Bu(\tau)\,d\tau \quad (2-14)$$
where $e^{At}$ is a matrix rather than a scalar. There are several ways of determining $e^{At}$. However, it will suffice here to write it in terms of an infinite series as

$$e^{At} = I + At + A^2 t^2 / 2! + A^3 t^3 / 3! + \ldots = \sum_{k=0}^{\infty} A^k t^k / k!$$  \hspace{1cm} (2-15)

where $I$ is the identity matrix. The reader should differentiate Equation 2-14, using Equation 2-15, to show that it is indeed a solution to Equation 2-9.

As a more general case, we consider the time-varying differential equation

$$\dot{x}(t) = A(t) x(t) + B(t) u(t) .$$  \hspace{1cm} (2-16)

A solution to Equation 2-16 is

$$x(t) = \Phi(t,t_0) x(t_0) + \int_{t_0}^{t} \Phi(t,\tau) B(\tau) u(\tau) d\tau$$  \hspace{1cm} (2-17)

where $\Phi(t,t_0)$ satisfies the differential equation

$$\Phi(t,t_0) = A(t) \Phi(t,t_0)$$  \hspace{1cm} (2-18)

with $\Phi(t_0,t_0) = I$, the identity matrix.

2.3 Discrete-Time Systems

2.3.1 Introduction

The discussions of the previous section have been concerned with continuous-time systems. Another class of systems that arises in practical situations is discrete-time systems. Other names for this class of systems are sampled data systems and digital systems. Discrete-time systems are distinguished from continuous-time systems in that rather than have state and output information available at all times, this information is available only at discrete instants of time. Discrete-time systems can arise by design, e.g., processing by digital computer, or by sampling the input and output of a continuous-time system.

2.3.2 Discrete-time Equations as Solutions to Continuous-time Equations

To illustrate the derivation of discrete-time state variable equations from continuous-time state variable equations consider again the spring-mass-damper system of Figure 2-1. Rather than assume that the input $f(t)$ changes continuously, assume that it changes in time increments of $T$ as illustrated in Figure 2-3. Assume further that the output is observed only at time intervals $T$. Since the inputs change only every $T$ and the outputs are observed only at time intervals of $T$, we are only concerned with the values of the states at the times
\( t = kT \) where \( k \) is an integer. This process of changing the input and observing the output at intervals of \( T \) is called sampling the system.

Consider the state variable solution of Equation 2-14 with \( t_0 = 0 \) for convenience. Since the input \( u(t) = u(0) \) is constant over the interval \( t \in (0,T] \) we can write

\[
x(t) = e^{At}x(0) + \left[ \int_0^t e^{A(t-\tau)}d\tau \right] Bu(0) \quad t \in (0,T].
\]  

(2-19)

Since we are interested in the states only at \( t = T \) and not all \( t \in (0,T] \), we can further write

\[
x(T) = e^{AT}x(0) + \left[ \int_0^T e^{A(T-\tau)}d\tau \right] Bu(0).
\]  

(2-20)

Now let us find \( x(2T) \). For \( t \in (0,T] \) \( u(t) = u(0) \) and for \( t \in (T,2T] \) \( u(t) = u(T) \). Thus we can write (from Equation 2-14 with \( t = 2T \))

\[
x(2T) = e^{2AT}x(0) + \int_0^{2T} e^{A(2T-\tau)}Bu(\tau)d\tau
\]

\[
= e^{2AT}x(0) + \int_0^T e^{A(2T-\tau)}d\tau Bu(0) + \int_T^{2T} e^{A(2T-\tau)}d\tau Bu(T).
\]

(2-21)

Substituting Equation 2-20 into Equation 2-21 and manipulating results in

---

Figure 2-3. Continuous and Discrete Time Input Signals
\[ x(2T) = e^{AT}x(T) + \left[ \int_0^T e^{A(T-\tau)}d\tau \right]Bu(T). \] (2-22)

The interim steps are left as an exercise for the reader.

If one examines Equation 2-20 and 2-22 it will be noted that they are of the same form, except for the indexing of \( x(\cdot) \) and \( u(\cdot) \). Based on this we can extrapolate the results above to the general case as

\[ x((k+1)T) = Fx(kT) + Hu(kT) \] (2-23)

or, dropping the \( T \) for convenience,

\[ x(k+1) = Fx(k) + Hu(k). \] (2-24)

In the above \( F = e^{AT} \) and \( H = \left[ \int_0^T e^{A(T-\tau)}d\tau \right]B \).

Equation 2-24 is termed a difference equation and is the standard form for representing discrete-time systems. Note the absence of the derivative in Equation 2-24. It should be noted that discrete-time systems are very easily implemented on the digital computer since they require only multiplications and additions rather than the integrations required by continuous-time systems.

### 2.3.3 State Equations From Block or Flow Diagrams

As with continuous-time systems, the state equations for discrete-time systems can also be determined from a block diagram of the system. To illustrate this, consider the block diagram of Figure 2-4. In this figure, the blocks with \( z^{-1} \) in them are unit delays or data storage devices. Physically these could be flip-flops, delay lines or other memory devices. At time \( kT \), data (states \( x_1(k+1), x_2(k+1) \)) appear at the input to these devices. At time \( (k+1)T \) (i.e., one clock period later) this data is transferred to the output of the devices.

![Figure 2-4. Discrete Time System Block Diagram](image)
To formulate the state equations, we express \( x_1(k+1) \) and \( x_2(k+1) \) in terms \( x_1(k), \ x_2(k) \) and \( u(k) \) as
\[
\begin{align*}
x_1(k+1) &= u(k) + ax_1(k) + cx_2(k) \\
x_2(k+1) &= x_1(k) - bx_2(k)
\end{align*}
\] (2-25)

The output is
\[
y(k) = x_2(k). \tag{2-26}
\]

In matrix form we have
\[
\begin{align*}
x(k+1) &= Fx(k) + g u(k) \\
y(k) &= Hx(k)
\end{align*}
\] (2-27)

where \( x(k) = \begin{bmatrix} x_1(k) \\ x_2(k) \end{bmatrix}, \ F = \begin{bmatrix} a & c \\ 1 & -b \end{bmatrix}, \ g = \begin{bmatrix} 1 \\ 0 \end{bmatrix}, \) and \( H = \begin{bmatrix} 0 & 1 \end{bmatrix}. \)

### 2.3 An Alternate Method for Finding a System Model

In addition to the formal techniques discussed above for finding representations of discrete-time systems we introduce another technique often used when explicit information about a system is difficult to obtain. To illustrate the technique we consider the example of a radar that we wish to use to estimate range \( R \) and range-rate \( \dot{R} \) of an airborne target (aircraft, missile, etc.) In this case the system consists of the target, atmosphere (if it is in the atmosphere) and earth. It should be clear that it would be extremely difficult to develop an accurate model for this system. Furthermore, if such a model could be developed it would be very non-linear – an undesirable feature for systems to which we hope to apply a Kalman filter. As a result, we attempt a simplified approach to the problem. Since the states of interest are range and range-rate we will develop equations involving these two variables. From elementary calculus we can express \( R \) and \( \dot{R} \) as a Taylor series expansion about some time \( t_0 \) as
\[
\begin{align*}
R(t) &= R_0 + \dot{R}_0 (t-t_0) + \ddot{R}_0 \frac{(t-t_0)^2}{2!} + \dddot{R}_0 \frac{(t-t_0)^3}{3!} + \ldots \\
\dot{R}(t) &= \dot{R}_0 + \ddot{R}_0 (t-t_0) + \dddot{R}_0 \frac{(t-t_0)^2}{2!} + \ldots
\end{align*}
\] (2-28)

where \( R_0, \ \dot{R}_0 \) and so forth are \( R(t), \ \dot{R}(t) \) and so forth evaluated at \( t = t_0 \). In this system we don’t know anything about \( \dddot{R}(t) \) and higher derivatives (except, maybe, in a very broad sense). Therefore, for want of anything better to do, we
drop them (when we discuss Kalman filters we will include them as an uncertainty). Thus, letting $\Delta t = t - t_o$ we obtain

$$R(t) = R_0 + \dot{R}_0 \Delta t$$
$$\dot{R}(t) = \dot{R}_0 .$$ \hspace{1cm} (2-29)

If we further let $t = (k + 1)T$ and $t_0 = kT$ we get a sampled data version as

$$R(k + 1) = R(k) + T \dot{R}(k)$$
$$\dot{R}(k + 1) = \dot{R}(k) .$$ \hspace{1cm} (2-30)

In a typical radar we usually make the measurement at $t = (k + 1)T$. If we assume that we only measure range then we have

$$y(k + 1) = R(k + 1) .$$ \hspace{1cm} (2-31)

It will be noted that the index on the measurement is $k + 1$ rather than $k$ as in Equations 2-26 and 2-27. This is a subtle but important distinction between the “control theory” problem and the “tracking problem”. It results in somewhat different Kalman filter formulations as will be shown later.

Rewriting Equations 2-30 and 2-31 in matrix form results in

$$x(k + 1) = Fx(k)$$
$$y(k + 1) = Hx(k + 1)$$ \hspace{1cm} (2-32)

where $x(k) = \begin{bmatrix} R(k) \\ \dot{R}(k) \end{bmatrix}$, $F = \begin{bmatrix} 1 & T \\ 0 & 1 \end{bmatrix}$ and $H = \begin{bmatrix} 1 & 0 \end{bmatrix}$.

It should be noted that these equations may or may not be valid for the particular problem considered (when designing a Kalman filter). However, the only real test of this is to use the model, build the filter and test it. As we proceed to implement examples of Kalman filter realizations we will again visit the problem of formulating state variable representations for the systems we wish to examine. The primary intent of this section has been to introduce the concept of state variable notation for use in our later work.
2.4 Problems

1. Find the state variable representation of the system of Figure 2-1 for the state assignments $x_1(t) = z(t) - \dot{z}(t)$ and $x_2(t) = z(t) + \dot{z}(t)$.

2. Write explicit forms for the matrices of Equation 2-9.

3. Show that Equation 2-14 is a solution to the differential equation of Equation 2-9.

4. Show that Equation 2-17 is a solution to Equation 2-16.

5. Fill-in the steps necessary to derive Equation 2-22 from Equation 2-21.