Properties of stochastic systems

Previously, we have seen how stochastic systems are defined. Now we look at what properties such systems can have.

1 Properties of Gaussian system representations

1.1 Definitions

Gaussian system representations have several properties. But before we can examine these properties, we need to make some definitions. First, we define the **state transition function** $\Phi : T \times T \to \mathbb{R}^{n \times n}$, associated with $A(t)$, recursively as the function

$$\Phi(t,s) = \begin{cases} A(t)\Phi(t-1,s) & \text{if } s < t, \\ I & \text{if } s = t, \\ 0 & \text{if } s < t. \end{cases} \tag{1.1}$$

For time-invariant systems, this can be reduced to

$$\Phi(t,s) = \begin{cases} A^{t-s} & \text{if } s \leq t, \\ 0 & \text{if } s > t. \end{cases} \tag{1.2}$$

Second, we define the following notations.

$$F_t^{x+} = \sigma\left(\{x(s), \forall s > t\}\right), \tag{1.3}$$

$$F_t^{x(t)} = \sigma\left(\{x(s), \forall s = t\}\right) = \sigma(\{x(t)\}), \tag{1.4}$$

$$F_t^x = F_t^{x-} = \sigma\left(\{x(s), \forall s \leq t\}\right). \tag{1.5}$$

So, whereas $F_t^{x(t)}$ is the σ-algebra generated by $x(t)$, F_t^x is the σ-algebra generated by all $x(s)$ with $s \leq t$.

We also have definitions for Gaussian processes and Markov processes. We'll examine them.

- A **Gaussian process** is a process $x(t)$ (with $t \in T$) such that all finite linear combinations of $x(t)$ is normally distributed as well. Thus, any variable $z = c_1 x(t_1) + \ldots + c_n x(t_n)$ is Gaussian.
- A **Markov process** is a process that satisfies the **Markov property**. This property requires that, given the current state $x(t)$ of a system, the future does not depend on the past. In an equation, this property/requirement can be written as

$$E\left[e^{iu^T x(t+1)} | F_t^{x(t)}\right] = E\left[e^{iu^T x(t+1)} | F_t^x\right]. \tag{1.6}$$

  In other words, the distribution of $x(t+1)$ depends on the distribution of $x(t)$. Knowing the distribution of $x(s)$ for $s < t$ doesn't influence this in any way.
- A process which is both a Gaussian process and a Markov process is called a **Gauss-Markov process**.

1.2 Properties

Let's examine a Gaussian system representation without any input $u(t)$. This system representation now has several properties. We'll list a couple.

1. For all $t \in T$, the σ-algebras F_t^{v+} and $(F_t^x \cup F_{t-1}^y)$ are independent. In other words, there is absolutely no relation between $v(s)$ for $s > t$ and either $x(s)$ for $s \leq t$ or $y(s)$ for $s \leq t-1$. Of course, there is a link between $v(t)$ and $y(t)$, just as there is a link between $v(t)$ and $x(t+1)$.
2. Let's suppose that we know the state $x(s)$ at some time $s \in T$. We can then find the state $x(t)$ and the output $y(t)$ at some time $t \in T$, using

$$x(t) = \Phi(t,s)x(s) + \sum_{u=s}^{t-1} \Phi(t-1,u)M(u)v(u), \tag{1.7}$$

$$y(t) = C(t)\left(\Phi(t,s)x(s) + \sum_{u=s}^{t-1} \Phi(t-1,u)M(u)v(u)\right) + N(t)v(t). \tag{1.8}$$

3. The process (x,y) is a jointly Gaussian process.
4. The state process $x(t)$ is a **Gauss-Markov** process with $x(t) \in G(m_x(t), Q(t))$. Here, we have

$$m_x(t+1) = A(t)m_x(t) \qquad \text{with} \qquad m_x(t_0) = m_0, \tag{1.9}$$

$$Q(t+1) = A(t)Q(t)A(t)^T + M(t)Q_v(t)M(t)^T \qquad \text{with} \qquad Q(t_0) = Q_0, \tag{1.10}$$

$$W_x(t,s) = E\left[(x(t)-m_x(t))(x(s)-m_x(s))^T\right] = \Phi(t,s)Q(s) \qquad (\text{for } t \geq s). \tag{1.11}$$

5. The output process $y(t)$ is a Gaussian process with $y(t) \in G(m_y(t), Q_y(t))$. Now we have

$$m_y(t) = C(t)m_y(t), \tag{1.12}$$

$$Q_y(t) = C(t)Q(t)C(t)^T + N(t)Q_v(t)N(t)^T, \tag{1.13}$$

$$W_y(t,s) = \begin{cases} Q_y(t) = C(t)Q(t)C(t)^T + N(t)Q_v(t)N(t)^T & \text{if } s = t, \\ C(t)\Phi(t,s)Q(s)C(s)^T + C(t)\Phi(t-1,s)M(s)Q_v(s)N(s)^T & \text{if } s < t. \end{cases} \tag{1.14}$$

2 Properties of time-invariant system representations

2.1 The impulse response function

Let's examine a time-invariant system without any noise. So, we have

$$x(t+1) = Ax(t) + Bu(t) \qquad \text{and} \qquad y(t) = Cx(t) + Du(t), \tag{2.1}$$

with $x(t_0) = x_0$. The state and the output of the system can now be determined using

$$x(t) = A^{t-t_0}x_0 + \sum_{s=t_0}^{t-1} A^{t-1-s}Bu(s) \qquad \text{and} \qquad y(t) = CA^{t-t_0}x_0 + \sum_{s=t_0}^{t-1} CA^{t-1-s}Bu(s) + Du(t). \tag{2.2}$$

We can also define the **impulse response function** $H(t)$ according to

$$H(t) = \begin{cases} D, & \text{if } t = 0, \\ CA^{t-1}B, & \text{if } t = 1, 2, \ldots. \end{cases} \tag{2.3}$$

Now, if $x_0 = 0$, we have

$$y(t) = \sum_{s=t_0}^{t} H(t-s)u(s). \tag{2.4}$$

2.2 Controllability

An important concept for systems is the concept of controllability. We say that a system is **controllable** if there is a time $t_1 \in T$ such that from every initial state $x_0 \in X$, any final state $x(t_1) = x_1 \in X$ can be reached. With 'can be reached', we mean that there is an input u such that, if $x(t_0) = x_0$, we have $x(t_1) = x_1$.

So how do we check if a system is controllable? For that, we can examine the **controllability matrix**, defined as

$$\text{conmat}(A, B) = \begin{bmatrix} B & AB & A^2B & \dots & A^{n-1}B \end{bmatrix}. \tag{2.5}$$

The system is controllable if and only if this controllability matrix is of full rank (i.e. it has rank n). In this case, we say that the pair of matrices (A, B) is a controllable pair. (The controllability only depends on the system matrices A and B, and not on C or D.) For discrete systems, it can be shown that if the system is controllable, then any state x_1 can be reached within a time t_1 for any t_1 satisfying $t_1 - t_0 \geq n$.

Now let's examine the case where (A, B) is not controllable. In this case, there is a state-space transformation S (with $\det S \neq 0$) such that $\bar{x}(t) = Sx(t)$. With respect to this new basis, the system representation takes the so-called **Kalman controllability form**

$$\bar{x}(t+1) = \begin{bmatrix} A_{11} & A_{12} \\ 0 & A_{22} \end{bmatrix} \bar{x}(t) + \begin{bmatrix} B_1 \\ 0 \end{bmatrix} u(t) \qquad \text{and} \qquad y(t) = \begin{bmatrix} C_1 & C_2 \end{bmatrix} \bar{x}(t) + Du(t). \tag{2.6}$$

In the above equation, (A_{11}, B_1) is a controllable pair. So, we have split up the system into a fully controllable part and a fully uncontrollable part.

2.3 Observability

A concept very similar to controllability is observability. Let's say that we don't know the state $x(t)$ of a system, but we do know the system matrices A, B, C and D. We say that the system is observable if there is a time t_1 such that, after t_1, we can always uniquely determine the state x of the system.

To find whether a system is observable, we can examine the **observability matrix**, defined as

$$\text{obsm(A,C)} = \begin{bmatrix} C \\ CA \\ CA^2 \\ \vdots \\ CA^{n-1} \end{bmatrix}. \tag{2.7}$$

The pair (A, C) is controllable if the above matrix is of full rank. If a discrete system is controllable, then it can be shown that the state x can always be uniquely determined within a time $t_1 - t_0 \geq n$.

Just like with controllability, we can split a system also up in an observable part and an unobservable part. This time, we find that

$$\bar{x}(t+1) = \begin{bmatrix} A_{11} & 0 \\ A_{21} & A_{22} \end{bmatrix} \bar{x}(t) + \begin{bmatrix} B_1 \\ B_2 \end{bmatrix} u(t) \qquad \text{and} \qquad y(t) = \begin{bmatrix} C_1 & 0 \end{bmatrix} \bar{x}(t) + Du(t). \tag{2.8}$$

Now, (A_{11}, C_1) is an observable pair.

We can also split up a system in both controllable and observable parts. We then find that

$$\bar{x}(t+1) = \begin{bmatrix} A_{11} & 0 & A_{13} & 0 \\ A_{21} & A_{22} & A_{23} & A_{24} \\ 0 & 0 & A_{33} & 0 \\ 0 & 0 & A_{43} & A_{44} \end{bmatrix} \bar{x}(t) + \begin{bmatrix} B_1 \\ B_2 \\ 0 \\ 0 \end{bmatrix} u(t), \tag{2.9}$$

$$y(t) = \begin{bmatrix} C_1 & 0 & C_3 & 0 \end{bmatrix} \bar{x}(t) + Du(t). \tag{2.10}$$

2.4 Stabilizability and detectability

The concepts of stabilizability and detectability are similar to controllability and observability, respectively. To test on **stabilizability** (i.e. whether a pair (A, B) is stabilizable), we can split up the system as in equation (2.6). If the noncontrollable part A_{22} is exponentially stable, then the pair (A, B) is stabilizable. Otherwise it is not. (With **exponentially stable**, we mean that the set of eigenvalues of A, denoted by $\text{spec}(A)$, falls within the unit circle, denoted by $\mathbb{D}_o$. Thus, $\text{spec}(A) \subset \mathbb{D}_o$.)

Alternatively, also the Hautus test can be used. Examine the matrix

$$\begin{bmatrix} (sI - A) & B \end{bmatrix}. \tag{2.11}$$

If the above matrix has rank n for all unstable eigenvalues $\lambda \in \text{spec}(A)$, then the pair (A, B) is stabilizable. Otherwise it is not.

The test for **detectability** is similar. The pair (A, C) is detectable if and only if the nonobservable part A_{22} in equation (2.8) is stable. Alternatively, the Hautus test can again be used. Now examine the matrix

$$\begin{bmatrix} (sI - A) \\ C \end{bmatrix}. \tag{2.12}$$

If the above matrix has rank n for all unstable eigenvalues $\lambda \in \text{spec}(A)$, then the pair (A, B) is detectable. Otherwise it is not.

2.5 Invariant measures

Let's consider an exponentially stable time-invariant Gaussian system. So, the matrices A, B, C, D and Q_v are constant in time. It can now be shown that there is an **invariant measure** $x(t) = G(0, Q_x)$. In other words, its distribution Q_x is constant. The corresponding stationary output is denoted by $y(t) = G(0, Q_y)$. The matrices Q_x and Q_y and also Q_{xy} have to be found by solving

$$Q_x = AQ_xA^T + MQ_vM^T, \tag{2.13}$$

$$Q_y = CQ_xC^T + NQ_vN^T, \tag{2.14}$$

$$Q_{xy} = AQ_xC^T + MQ_vN^T. \tag{2.15}$$

The equation for Q_x is known as the **Lyapunov equation**. If A is exponentially stable, then it always has a unique solution $Q_x = Q_x^T \geq 0$. We'll give the Lyapunov equation a closer look in a moment. But first, we mention that the covariance functions are given by

$$W_x(t) = A^tQ_x, \tag{2.16}$$

$$W_y(t) = \begin{cases} CA^{t-1}\left(AQ_xC^T + MQ_vN^T\right) & \text{if } t > 0, \\ Q_y & \text{if } t = 0, \end{cases} \tag{2.17}$$

$$W_{xy}(t) = CA^tQ_x. \tag{2.18}$$

Let's suppose that we have some process $x(t)$ which is not equal to the invariant measure. If the system is exponentially stable, then it can be shown that $x(t)$ will always converge to the invariant measure. So, $\lim_{t \to \infty} Q(t) = Q_x$, with Q_x the solution to the above equation.

2.6 The Lyapunov equation

The equation

$$Q = AQA^T + MQ_vM^T \tag{2.19}$$

is known as the **Lyapunov equation**. It should be solved for Q. Mostly, numerical methods are employed here, like the Matlab function `dlyap`. But this equation is also subject to a few theories. Let's suppose that there is some G satisfying $GG^T = MQ_vM^T$. If (A, G) is stabilizable and if there is some $Q = Q^T \geq 0$ satisfying the Lyapunov equation, then A is exponentially stable.

Next to this, consider the following three statements. When two of these statements hold, then the third must hold as well. (Or equivalently, when one doesn't hold, then at least one of the others doesn't hold either.)

- A is an exponentially stable matrix. (So, $\text{spec}(A) \subset \mathbb{D}_o$.)
- (A, G) is a reachable pair. (Reachable is another word for controllable.)
- Q is positive definite. (So, $Q > 0$.)