Basics of stochastic systems

Stochastic systems are systems with stochastic processes: there is uncertainty. How do we deal with this uncertainty? That is what we will look at in this chapter. To be more precise, we'll examine the basics of stochastic systems. How are they defined and written down?

1 Stochastic processes

1.1 Definitions of stochastic processes

Consider a probability space (Ω, F, P) and a measurable space (X, G). Also, we have an **index set** T. This index set usually denotes time. So, $t \in T$ with either $T = \mathbb{N}$ or $T = \mathbb{Z}$. (The system is discrete in time.) A **stochastic process** is a function $x : \Omega \times T \to X$. In other words, for all t, the parameter $x(., t)$ is a random variable. It is sometimes also denoted as x_t or $x(t)$. On the other hand, the function $x(\omega, .) : T \to X$ (for fixed ω) is called a **sample path** of the process x.

A stochastic process x is called a **Gaussian process** if every subset of random variables $(x_{t_1}, x_{t_2}, \ldots, x_{t_m})$ (with $t_i \in T$) is jointly Gaussian. Similarly, two stochastic processes x and y are called **jointly Gaussian** if every subset of random variable $(x_{t_1}, \ldots, x_{t_m}, y_{s_1}, \ldots, y_{s_n})$ is jointly Gaussian. Two independent Gaussian processes are always jointly Gaussian.

An example of a stochastic process is a **Gaussian white noise process** $v_t \in G(0, V(t))$. So, it is a Gaussian process with mean 0 and variance matrix $V(t)$. Here, $V(t)$ satisfies $V(t) = V(t)^T \geq 0$ for all t. Furthermore, all v_t are independent with respect to each other.

1.2 Properties of stochastic processes

Let's consider a stochastic process x. We define the **mean** of the process as $m_x(t) = E[x(t)]$. Similarly, we have the **joint moment function** or **correlation function** $C_x(t, s) = E[x(t)x(s)^T]$ and the **covariance function**

$$W_x(t, s) = E[(x(t) - m_x(t))(x(s) - m_x(s))^T]. \tag{1.1}$$

The covariance function has as property that that $W(t, s) = W(s, t)^T$ for all $s, t \in T$. Also, W is positive definite ($W \geq 0$). For a function $W : T \times T \to \mathbb{R}^{n \times n}$ this means that, for every set $t_1, \ldots, t_m$ and constant vectors $c_1, \ldots, c_m$, we have

$$\sum_{i=1}^{m} \sum_{j=1}^{m} c_i^T W(t_i, t_j) c_j \geq 0. \tag{1.2}$$

Let's examine a subset $(x_{t_1}, x_{t_2}, \ldots, x_{t_m})$ with $t_i \in T$. We time-shift this subset by a time s such that also $t_i + s \in T$. Now also consider the subset $(x_{t_1+s}, x_{t_2+s}, \ldots, x_{t_m+s})$. We say that the process is **stationary** if these two subsets have the same joint distribution for all subsets t_i and all time-shifts s.

The concept of time-reversibility is defined similarly. Now examine the subsets $(x_{t_1}, x_{t_2}, \ldots, x_{t_m})$ and $(x_{t-t_1}, x_{t-t_2}, \ldots, x_{t-t_m})$ for some time $t \in T$. We say that the process is **time-reversible** if these two subsets have the same joint distribution for all subsets t_i and times t. Also, it can be shown that a time-reversible process is always time-invariant. The converse doesn't always hold.

1.3 Properties of Gaussian stochastic processes

Let's examine a Gaussian process x on the time index $T = \mathbb{Z}$. It can be shown that x is stationary if $m(t) = m(0)$ for all $t \in T$ and if $W(t, s) = W(t + u, s + u)$ for all $t, s, u \in T$. If this is the case, then we can define a new covariance function $W_1(t) = W(t, 0) = W(t + s, s)$. In other words, the covariance function only depends on one argument. This new function is **para-symmetric**: $W(t) = W(-t)^T$ for all $t \in T$.

Let's examine a stationary Gaussian process x with zero mean. This process is also time-reversible if $W(t) = W(-t)$ or, equivalently, $W(t) = W(t)^T$. This also implies that a scalar stationary Gaussian process is always time-reversible.

2 Representing stochastic systems

2.1 Modeling a stochastic system

Let's examine a **stochastic system**. This system has an **output** $y(t)$, an **input** $u(t)$ and a **noise** $v(t)$. We can usually control the input $u(t)$. However, the noise $v(t)$ is uncontrollable. In fact, it is assumed to be Gaussian white noise. So, $v(t) \in G(0, Q_v(t))$. We can now model the system with an **ARMAX representation**, being

$$y(t) = \sum_{i=1}^{n} a_i y(t-i) + \sum_{i=0}^{n} b_i u(t-i-k) + \sum_{i=0}^{n} c_i v(t-i). \tag{2.1}$$

In the above equation, k is the **input delay** and $t \in T$. Also, we generally have $c_0 = 1$. Do note that, since the noise $v(t)$ is a stochastic process, also the output $y(t)$ will be a stochastic process.

Working with the above representation has disadvantages. Luckily, it can be rewritten to the **state space representation** of stochastic systems, also known as the **Gaussian system representation**. This is

$$x(t+1) = A(t)x(t) + B(t)u(t) + M(t)v(t), \tag{2.2}$$
$$y(t) = C(t)x(t) + D(t)u(t) + N(t)v(t). \tag{2.3}$$

Here, the stochastic process $x(t)$ is the **state** of the system. It is assumed that $x_0 = x(t_0) \in G(m_0, Q_0)$ is known.

Sometimes, we assume that the system is **stationary/time invariant**. This means that the matrices A, B, M, C, D, N and Q_v don't depend on time. This significantly simplifies the system.

2.2 Interconnecting systems

Let's consider figure 1. In this figure, three Gaussian systems are connected. There are the control system (1), the input noise (2) and the output noise (3). These three systems can be modeled by

$$x_1(t+1) = A_1 x_1(t) + B_1 u_1(t), \qquad y_1(t) = C_1 x_1(t) + D_1 u_1(t), \tag{2.4}$$
$$x_2(t+1) = A_2 x_2(t) + B_2 u_2(t), \qquad y_2(t) = C_2 x_2(t) + D_2 u_2(t), \tag{2.5}$$
$$x_3(t+1) = A_3 x_3(t) + B_3 u_3(t), \qquad y_3(t) = C_3 x_3(t) + D_3 u_3(t). \tag{2.6}$$

Also note that $u_1(t) = u(t) + y_2(t)$ and $y(t) = y_1(t) + y_3(t)$. It may seem complicated to deal with this system. But luckily, we can write the whole system in state space form as well. If we are to do this, we first ought to define

$$x = \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} \qquad \text{and} \qquad v = \begin{bmatrix} v_2 \\ v_3 \end{bmatrix}. \tag{2.7}$$

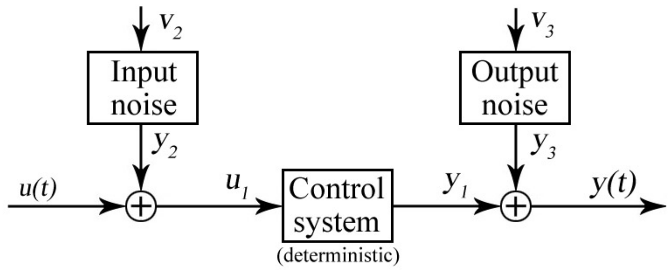

Figure 1: The interconnection of Gaussian systems.

Now, by using the equations above, it can be derived that

$$x(t+1) = \begin{bmatrix} A_1 & B_1C_3 & 0 \\ 0 & A_2 & 0 \\ 0 & 0 & A_3 \end{bmatrix} x(t) + \begin{bmatrix} B_1 \\ 0 \\ 0 \end{bmatrix} u(t) + \begin{bmatrix} B_1N_2 & 0 \\ M_2 & 0 \\ 0 & M_3 \end{bmatrix} v(t), \tag{2.8}$$

$$y(t) = \begin{bmatrix} C_1 & D_1C_2 & C_3 \end{bmatrix} x(t) + D_1u(t) + \begin{bmatrix} D_1N_2 & N_3 \end{bmatrix} v(t). \tag{2.9}$$

2.3 Stochastic systems in literature

In the literature, you often find a representation of the form

$$x(t+1) = A(t)x(t) + M_1(t)r(t), \qquad y(t) = C(t)x(t) + N_1(t)w(t). \tag{2.10}$$

Here, $r(t)$ and $w(t)$ are independent white noise processes. This is, however, only a special case of our previous representation. In fact, we can put the above representation into our own form, using

$$v(t) = \begin{bmatrix} r(t) \\ w(t) \end{bmatrix}, \qquad M(t) = \begin{bmatrix} M_1(t) & 0 \end{bmatrix} \quad \text{and} \quad N(t) = \begin{bmatrix} 0 & N_1(t) \end{bmatrix}. \tag{2.11}$$

Our new noise signal is now given by

$$v(t) \in G\left(0, \begin{bmatrix} Q_r(t) & 0 \\ 0 & Q_w(t) \end{bmatrix}\right). \tag{2.12}$$

2.4 Forward and backward representation

Previously, we have considered systems in the **forward representation**. It was written as

$$x(t+1) = A^f x(t) + Mv^f(t) \qquad \text{and} \qquad y(t) = C^f(t)x(t) + Nv^f(t). \tag{2.13}$$

Though, if it is clear that we are using the forward representation, the superscript f is not written. If we have $Q(t) = E[x(t)x(t)^T] > 0$, then it can be shown that

$$A^f(t) = E[x(t+1)x(t)^T]Q(t)^{-1}, \tag{2.14}$$

$$C^f(t) = E[y(t)x(t)^T]Q(t)^{-1}, \tag{2.15}$$

$$Q_v^f(t) = \begin{bmatrix} Q(t+1) & E[x(t+1)y(t)^T] \\ E[y(t)x(t+1)^T] & E[y(t)y(t)^T] \end{bmatrix} - \begin{bmatrix} A^f(t) \\ C^f(t) \end{bmatrix} Q(t)^{-1} \begin{bmatrix} A^f(t)^T & C^f(t)^T \end{bmatrix}. \tag{2.16}$$

We also have $M = [I_n \; 0]$ and $N = [0 \; I_p]$. However, we could also use the backward representation of the system. It is then written as

$$x(t-1) = A^b x(t) + M v^b(t) \qquad \text{and} \qquad y(t-1) = C^b(t)x(t) + N v^b(t). \tag{2.17}$$

This time, the system matrices satisfy

$$\begin{aligned} A^b(t) &= E[x(t-1)x(t)^T]Q(t)^{-1}, & (2.18)\\ C^b(t) &= E[y(t-1)x(t)^T]Q(t)^{-1}, & (2.19)\\ Q_v^b(t) &= \begin{bmatrix} Q(t-1) & E[x(t-1)y(t)^T] \\ E[y(t)x(t-1)^T] & E[y(t)y(t)^T] \end{bmatrix} - \begin{bmatrix} A^b(t) \\ C^b(t) \end{bmatrix} Q(t)^{-1} \begin{bmatrix} A^b(t)^T & C^b(t)^T \end{bmatrix}. & (2.20) \end{aligned}$$

Based on the above equations, we can also find the relation between te forward and the backward representations. It is given by

$$\begin{aligned} A^f(t)Q(t) &= Q(t+1)A^b(t+1)^T, & (2.21)\\ A^b(t)Q(t) &= Q(t-1)A^f(t-1)^T, & (2.22) \end{aligned}$$

$$C^f(t)Q(t) = C^b(t+1)Q(t+1)A^b(t+1)^T + NQ_v^b(t+1)M^T, \tag{2.23}$$

$$C^b(t)Q(t) = C^f(t-1)Q(t-1)A^f(t-1)^T + NQ_v^f(t-1)M^T. \tag{2.24}$$

$$\tag{2.25}$$

If the system is stationary and the matrices are thus constant in time, then the above equations can be simplified somewhat.