Stochastic realizations

In this chapter, we're going to examine stochastic realizations. What are they? How can we find them? And how can we be sure that they are minimal?

1 Basic ideas of stochastic realizations

1.1 The weak Gaussian stochastic realization problem

Let's suppose that we're looking at some actual process. From it, we derive some output $z(t)$. After perhaps some filtering, we can derive the **average** z_a and the **covariance** $\hat{W}(t)$ of the data, according to

$$z_a = \frac{1}{t_1}\sum_{s=1}^{t_1} z(t) \qquad \text{and} \qquad \hat{W}(t) = \frac{1}{t_1 - t}\sum_{s=1}^{t_1-t}(z_{t+s} - z_a)(z_s - z_a)^T. \tag{1.1}$$

Now the question arises, can we find a time-invariant Gaussian system such that the output y of this system equals the considered process z? If there is, then we call such a system **realization** of the considered process. But if there is such a system, then the question also arises, is it minimal? (A realization is called **minimal** if there is no other realization with a smaller dimension.) And can we find all minimal realizations? This problem is actually known as the **weak Gaussian stochastic realization problem**. And a solution is known as a **weak Gaussian stochastic realization**.

1.2 Stochastic observability

When considering minimality of realizations, we will need the concepts of stochastic observability and stochastic reconstructibility. So we will consider those here. Let's examine a time-invariant Gaussian system. Suppose that we know the conditional distributions $\left(\{y(t), y(t+1), \ldots, y(t+t_1)\}|F^{x(t)}\right)$ of the future outputs. (So basically, we have experimental data on the output of the system.) If we can derive the distribution of $x(t)$ from this, then we call the system **stochastically observable**.

There is a relatively easy way to determine whether a system is stochastically observable. Let's assume that our time-invariant system is exponentially stable. Also assume that there is some G such that $GG^T = MQ_vM^T$ and that (A^f, G) is a reachable pair. (Reachable is another word for supportable or controllable.) Then the system is stochastically observable if and only if (A^f, C^f) is an observable pair.

1.3 Stochastic reconstructibility

Stochastic reconstructibility (also sometimes called **stochastic coobservability** is similar to stochastic observability. Again examine a time-invariant Gaussian system. But now suppose that we know the conditional distributions $\left(\{y(t), y(t-1), \ldots, y(t-t_1)\}|F^{x(t)}\right)$ of the past outputs. If we can derive the distribution of $x(t)$ from this, then we call the system **stochastically reconstructible**.

Again, there is a way to check stochastic reconstructibility. Let's assume that our time-invariant system is exponentially stable. Also assume that there is some G such that $GG^T = MQ_vM^T$ and that (A^b, G) is a reachable pair. Then the system is stochastically reconstructible if and only if (A^b, C^b) is an observable pair.

So, stochastic observability means that, given a series of conditional output distributions, you can find the distribution of the initial state. On the other hand, stochastic reconstructibility means that, given a series of conditional output distributions, you can find the distribution of the final state. Often, when a system is stochastically observable, it is also stochastically reconstructible and vice versa. But this is definitely not always the case.

2 Dissipative systems

Another relevant topic to the subject of stochastic realizations is the topic of dissipative systems. Let's consider a system of the form

$$x(t+1) = Fx(t) + Gu(t) \qquad \text{and} \qquad y(t) = Hx(t) + Ju(t), \tag{2.1}$$

where we assume that $J = J^T$ and thus that $u(t)$ and $y(t)$ are of the same size p. The matrices F, G, H and J (and their sizes) are the **linear system parameters** (LSP) corresponding to this system. We'll use this system when discussing dissipative systems.

2.1 Positive definite functions

Let's define $\mathbf{U}$ as a set of inputs to the system above, according to

$$\mathbf{U} = \left\{ u : T \to \mathbb{R}^p \,|\, ||u|| = \sqrt{\sum_{s \in T} u(s)^T u(s)} < \infty \right\}. \tag{2.2}$$

Now consider a function $W : T \times T \to \mathbb{R}^{p \times p}$. (It will be similar to the covariance function.) We say that W is **stationary** if $W(t,s) = W(t-s,0)$ for all $s,t \in T$. In this case, we simply write $W(t,0) = W_1(t)$. Also, W is called **parasymmetric** if $W(t,s) = W(s,t)^T$. For stationary functions this is equivalent to $W(t) = W(-t)^T$. And we say that W is **finite-dimensional** if there are linear system parameters such that

$$W(t) = \begin{cases} HF^{t-1}G & \text{if } t > 0, \\ 2J & \text{if } t = 0, \\ G^T (F^T)^{-t-1} H^T & \text{if } t < 0. \end{cases} \tag{2.3}$$

Now let's define the operator $\mathbf{W}$ according to

$$(\mathbf{W}(u))(t) = \sum_{s=-\infty}^{t-1} W(t,s)u(s) + \frac{1}{2} W(t,t)u(t). \tag{2.4}$$

We say that $\mathbf{W}$ is a **positive definite operator** and W is a **positive definite function** if for any $u \in \mathbf{U}$ we have

$$u^T \mathbf{W} u = \sum_{s \in T} \sum_{t \in T} u(t)^T W(t,s) u(s) \geq 0. \tag{2.5}$$

Similarly, $\mathbf{W}$ and W are called **strictly positive definite** if $u^T \mathbf{W} u > 0$ for all nonzero $u \in \mathbf{U}$.

2.2 Supply rates and storage functions

Let's again consider the system of the form of equation (2.1). We define the **supply rate** as

$$h(u(t), y(t)) = h(t) = u(t)^T y(t) = \frac{1}{2} \begin{bmatrix} u(t) \\ y(t) \end{bmatrix}^T J_s \begin{bmatrix} u(t) \\ y(t) \end{bmatrix}, \qquad \text{with} \qquad J_s = \begin{bmatrix} 0 & I_p \\ I_p & 0 \end{bmatrix}. \tag{2.6}$$

The system is called **dissipative** if there is a **storage function** $S(x(t))$ such that for all $s,t \in T$ and for all inputs $u \in \mathbf{U}$ we have

$$S(x(t)) \leq S(x(s)) + \sum_{\tau=s}^{t-1} h(u(\tau), y(\tau)). \tag{2.7}$$

Basically, S can be seen as the 'energy' in the system and h as the 'energy supply rate'. So for dissipative systems, energy is lost.

A special type of storage function is the **available storage** $S^-(x)$. Before we define it, we first define the set $\mathbf{U}(t_0, x_0, t_1, x_1)$ as

$$\mathbf{U}(t_0, x_0, t_1, x_1) = \{u \in \mathbf{U} | \text{if } x(t_0) = x_0 \text{ and } u \text{ is applied as input, then } x(t_1) = x_1\}. \tag{2.8}$$

Now, the available storage $S^-(x)$ is defined as

$$S^-(x) = \sup_{t>0, u \in \mathbf{U}(0,x,t,.)} \left(-\sum_{\tau=0}^{t-1} h(\tau) \right). \tag{2.9}$$

Basically, the available storage can be seen as the maximum amount of energy that can be extracted from a system without initial supply/energy. It can now be shown that the system is dissipative if and only if the available storage is finite. (So if for all $x \in \mathbb{R}^n$ we have $S^-(x) > \infty$.) Also, if the system is dissipative, then $S^-(x) \geq 0$ is a storage function. And every other storage function $S(x)$ will be at least as big. So, $0 \leq S^-(x) \leq S(x)$.

Similarly, we can define the **required supply** $S^+(x)$ as

$$S^+(x) = \inf_{t<0, u \in \mathbf{U}(t,.,0,x)} \left(\sum_{\tau=t}^{-1} h(\tau) \right). \tag{2.10}$$

(Just like the supremum can be seen as an upper bound, the infimum is like a lower bound.) The required supply $S^+(x)$ can be seen as the minimum supply that is necessary at a negative time t to bring the system to state x at $\tau = 0$ with zero supply. If the system is dissipative and controllable, then $S^+(x)$ exists, is finite, and we have $S(x) \leq S^+(x)$ for all other storage functions $S(x)$.

2.3 Characterizing linear dissipative systems

Consider the system of equation (2.1). Its dual system is defined as

$$x(t+1) = F^T x(t) + H^T u(t) \qquad \text{and} \qquad y(t) = G^T x(t) + J u(t). \tag{2.11}$$

For the system and its dual system, given the linear system parameters H, G, H and J, we can define the matrices

$$V_{lsp}(Q) = \begin{bmatrix} Q - F^T Q F & H^T - F^T Q G \\ H - G^T Q F & 2J - G^T Q G \end{bmatrix} \qquad \text{and} \qquad V_{lsdp}(Q) = \begin{bmatrix} Q - F Q F^T & G - F Q H^T \\ G^T - H Q F^T & 2J - H Q H^T \end{bmatrix}. \tag{2.12}$$

These matrices play an important role with supply rates. In fact, if $Q = Q^T$, we have

$$\frac{1}{2} x(t)^T Q x(t) - \frac{1}{2} x(s)^T Q x(s) - \sum_{\tau=s}^{t-1} h(\tau) = -\sum_{\tau=s}^{t-1} \frac{1}{2} \begin{bmatrix} u(t) \\ y(t) \end{bmatrix}^T V_{lsp}(Q) \begin{bmatrix} u(t) \\ y(t) \end{bmatrix}. \tag{2.13}$$

If $V_{lsp}(Q)$ is positive definite, then the right side of the equation is negative or zero. This means that $\frac{1}{2} x^T Q x$ is a storage function. So we would like to know for which Q the matrix $V_{lsp}(Q)$ is positive definite. We thus define the sets $\mathbf{Q_{lsp}}$ and $\mathbf{Q_{lsdp}}$ as

$$\mathbf{Q_{lsp}} = \left\{ Q \in \mathbb{R}^{n \times n} | Q = Q^T \geq 0 \text{ and } V_{lsp}(Q) \geq 0 \right\}, \tag{2.14}$$

$$\mathbf{Q_{lsdp}} = \left\{ Q \in \mathbb{R}^{n \times n} | Q = Q^T \geq 0 \text{ and } V_{lsdp}(Q) \geq 0 \right\}. \tag{2.15}$$

If we now assume that (F, G) is controllable and (F, H) is observable, then the following three statements are equivalent.

- The system is dissipative.
- W is a positive definite function.
- There exists a $Q \in \mathbf{Q_{lsp}}$.

Also, $S(x) = \frac{1}{2}x^T Qx$ with $Q = Q^T \geq 0$ is a storage function if and only if $Q \in \mathbf{Q_{lsp}}$. We can also define Q^- and Q^+ as the minimal and maximal solutions of the **algebraic Riccati equation**

$$D(Q) = \left(Q - F^T QF\right) - \left(H^T - F^T QG\right)\left(2J - G^T QG\right)^{-1}\left(H - G^T QF\right) = 0. \tag{2.16}$$

(It can be noted that $D(Q)$ is the Schur complement of the matrix $V_{lsp}(Q)$.) Now it can be shown that for any $Q \in \mathbf{Q_{lsp}}$ we have $Q^- \leq Q \leq Q^+$. Next to this, we also have

$$S^-(x) = \frac{1}{2}x^T Q^- x \qquad \text{and} \qquad S^+(x) = \frac{1}{2}x^T Q^+ x. \tag{2.17}$$

3 Stochastic realizations

3.1 The covariance realization

Now we can finally get back to stochastic realizations. Let's suppose we have some output signal with covariance matrix $W(t)$ with $W(0) > 0$ but $\lim_{t\to\infty} W(t) = 0$. Is there a weak Gaussian stochastic realization of the following form?

$$x(t+1) = Ax(t) + Mv(t) \qquad \text{and} \qquad y(t) = Cx(t) + Nv(t). \tag{3.1}$$

Well, it can be shown that such a realization exists if and only if the Hankel matrix H_W associated with W satisfies $\text{rank}(H_W) < \infty$. Here, the Hankel matrix $H_W(k_1, k_2)$ and $\text{rank}(H_W)$ are defined as

$$H_W = \begin{bmatrix} W(1) & W(2) & \cdots & W(k_2) \\ W(2) & \ddots & \ddots & \vdots \\ \vdots & \ddots & \ddots & W(k_1+k_2-2) \\ W(k_1) & \cdots & W(k_1+k_2-2) & W(k_1+k_2-1) \end{bmatrix}, \tag{3.2}$$

$$\text{rank}(H_W) = \sup_{k_1,k_2\in\mathbb{Z}} \text{rank}(H_W(k_1,k_2)). \tag{3.3}$$

In this case, there exist linear system parameters F, G, H and J such that

$$W(t) = \begin{cases} HF^{t-1}G & \text{if } t > 0, \\ 2J & \text{if } t = 0, \\ G^T(F^T)^{-t-1}H^T & \text{if } t < 0. \end{cases} \tag{3.4}$$

(Algorithms for finding F, G, H and J exist, but we will not discuss them here.) A system with the above form will be called a **covariance realization** of the covariance function W. We should also select a $Q \in \mathbf{Q_{ldsp}}$. The matrices A, B, M, N and Q_v are now given by $A = F$, $C = H$, $M = [I_n \ 0]$, $N = [0 \ I_p]$ and $Q_v = V_{lsdp}(Q)$. This gives us the well-known Gaussian system representation

$$x(t+1) = Ax(t) + Mv(t) \qquad \text{and} \qquad y(t) = Cx(t) + Nv(t). \tag{3.5}$$

3.2 Properties of Gaussian system representation

Let's examine a Gaussian system representation of the form of equation (3.5). In this equation A has size $n \times n$, M has size $n \times m$, C has size $p \times n$ and N has size $p \times m$. The representation can have several properties.

- The representation is **regular** if $\text{rank}(NVN^T) = p$.
- The representation is **square** if it is regular and $m = p$. It is non-square if it regular and $m > p$.
- Assume that $\text{spec}(A) \in \mathbb{D}_o$. The representation is **supported** on the full state space if $x(t) \in G(0, Q)$ with $Q = Q^T > 0$. Here, Q is the solution of the Lyapunov equation $Q = AQA^T + MVM^T$.
- The representation is an **output-based** stochastic realization of its output process if $\text{rank}(V) = \text{rank}(VNV^T)$. An output-based stochastic realization is sometimes also called an **internal stochastic realization**.
- The representation is a **Kalman realization** of the associated output process if the system is regular, output-based and satisfies $\text{spec}(A) \subset \mathbb{D}_o$ and $\text{spec}(A - KC) \subset \mathbb{D}_o$.

3.3 Minimal stochastic realizations

We have noted before that a realization is called minimal if there is no other realization with a smaller state. It can be shown that a realization is minimal if and only if it is stochastically observable, it is stochastically reconstructible and the support of the state process equals the state space. This latter condition is equivalent to $Q > 0$.

Minimal realizations are, however, not unique. Let's take a nonsingular matrix S. We now replace A by SAS^{-1}, C by CS^{-1} and M by SM. (N and Q_v are just left the same.) This gives us a completely new stochastic realization.

There is also another way to find other stochastic realizations. This time, we don't change A, C and M, but we change Q_v instead. To do this, we simply take another $Q \in \mathbf{Q_{lsdp}}$. And, by choosing specific Q, we can also vary the properties of the realization we get. For example, if we take $Q^- \in \mathbf{Q_{lsdp}}$, then we acquire the Kalman realization. As was mentioned before, this realization is regular, output-based and satisfies $\text{spec}(A) \subset \mathbb{D}_o$ and $\text{spec}(A - KC) \subset \mathbb{D}_o$. Furthermore, the realization can be written as

$$x(t+1) = Ax(t) + Kw(t) \qquad \text{and} \qquad y(t) = Cx(t) + w(t), \tag{3.6}$$

with $w(t) \in G(0, W)$ and $W = W^T > 0$.