# Control of Discrete-Time Stochastic Systems Summary

# 1. Probability theory

In this summary, we will be examining a lot of stochastic systems. Stochastic systems deal with probabilities. So, let's dive into the probability theory first.

## 1.1 The probability distribution function

### 1.1.1 Definition of the probability distribution function

Very important in probability theory is the **probability distribution function** (PDF) $f(u)$. This function has as limits $f(-\infty = 0)$ and $f(\infty) = 1$. It also increases: if $u < v$, then $f(u) \leq f(v)$. Finally, PDFs are also **right continuous**. To find out what this means, we examine some discontinuity $v$ in the graph of $f(u)$. Now let's approach this point $v$ from the right. The value which we get is $f(u+)$. Right continuous functions now must have $f(u) = f(u+)$.

We can make a distinction between **continuous** and **discrete** PDFs. Continuous PDFs usually have a continuous shape: the value of $f(u)$ more or less gradually increases from 0 to 1. For **continuous PDFs** we also have

$$f(u) = \int_{-\infty}^{u} p(v)\,dv, \qquad \text{where} \qquad f(\infty) = \int_{-\infty}^{\infty} p(v)\,dv = 1. \tag{1.1.1}$$

In the above equation, $p(u)$ is the **probability density function**.

**Discrete PDFs** are rather different. The graph of $f(u)$ now takes the shape of a staircase. The points where $f(u)$ jumps up are denoted by $u_n$.

$$f(u) = \sum_{u_n < u} p(n), \qquad \text{where} \qquad f(\infty) = \sum p(n) = 1. \tag{1.1.2}$$

Now, $p(n)$ is called the **probability frequency function**.

### 1.1.2 Examples of probability distribution functions

Several examples of PDFs exist. We'll examine a few now. The **Bernoulli distribution** has a discrete PDF. Given the parameter $q$ (satisfying $0 \leq q \leq 1$), the distribution is defined by

$$p(1) = q \qquad \text{and} \qquad p(0) = 1 - q. \tag{1.1.3}$$

The **Poisson distribution** is discrete as well. Given the parameter $\lambda$ (satisfying $\lambda \in \mathbb{R}_+$), it is defined by

$$p(k) = \lambda^k \frac{e^{-\lambda}}{k!}. \tag{1.1.4}$$

In this equation, we must have $k \in \mathbb{N} = \{0, 1, \ldots\}$.

The **gamma distribution** is continuous. Its parameters are $\lambda$ and $r$ and satisfy $\lambda, r \in \mathbb{R}_+$. The distribution is defined by

$$p(v) = \frac{v^{r-1}}{\lambda^r} \frac{e^{-\frac{v}{\lambda}}}{\Gamma(r)}, \qquad \text{where} \qquad \Gamma(r) := \int_0^{\infty} v^{r-1} e^{-v}\,dv. \tag{1.1.5}$$

The function $\Gamma(r)$ is known as the **gamma function**. (By the way, the ':=' means 'is per definition'.)

However, the most important distribution is the **Gaussian distribution**, also known as the **normal distribution**. This continuous distribution has as parameters a mean vector $m$ and a variance matrix $Q$. $m$ satisfies $m \in \mathbb{R}^n$ while $Q$ satisfies both $Q \in \mathbb{R}^{n \times n}$ and $Q = Q^T > 0$. (With $Q > 0$ we mean that $Q$ is **strictly positive definite**, which in turn demands that $x^t Q x > 0$ for all vectors $x$. This, in turn, demands that all eigenvectors of $Q$ are positive.) The distribution is now defined by

$$p(v_1, v_2, \ldots, v_n) = \frac{1}{\sqrt{(2\pi)^n \det(Q)}} e^{-\frac{1}{2}(v-m)^T Q^{-1}(v-m)}. \tag{1.1.6}$$

But why is this distribution so important? Well, let's suppose that we have a number of independent distributions. If we add these distributions up and normalize them, then the **central limit theorem** claims that the resulting distribution will converge to a Gaussian distribution. The more distributions are added up, the close the result will be a to a Gaussian distribution. And since many phenomena in real life are the result of sums of distributions, we can use the Gaussian distribution to approximate them.

## 1.2 Measurable spaces and probability spaces

### 1.2.1 $\sigma$-algebras

Let's examine a set $\Omega$. A $\sigma$-algebra $F$ on $\Omega$ is a collection of subsets of $\Omega$, satisfying three important rules.

1. If the set $A$ is in $F$ ($A \in F$), then the complement $A^c$ is also in $F$ ($A^c \in F$). (In other words, $F$ is closed with respect to complementation.)
2. Let's examine a set of sets $\{A_1, A_2, \ldots, A_n\}$ such that all $A_i$ are in $F$. Now let's take the union of all these sets. This union must now also be in $F$. In an equation, we have $A_1 \cup A_2 \cup \ldots \cup A_n \in F$.
3. The set $\Omega$ is in $F$. (And thus, due to rule 1, also the empty set $\emptyset$ is in $F$.)

Examples of $\sigma$-algebras include $\{\emptyset, \Omega\}$ and $\{\emptyset, A, A^c, \Omega\}$ for every set $A \in \Omega$.

A tuple $(\Omega, F)$, consisting of a set $\Omega$ and a $\sigma$-algebra $F$ on $\Omega$, is called a **measurable space**. A $\sigma$-algebra $G$ on $\Omega$ consisting of subsets of the $\sigma$-algebra $F$ (thus satisfying $G \subseteq F$) is called a sub-$\sigma$-algebra.

### 1.2.2 Probability measures

Suppose we have a measurable space $(\Omega, F)$. Let's examine a function $P : F \to \mathbb{R}_+$. (In other words, the function $P$ takes as input elements of $F$ and as output it gives elements of $\mathbb{R}_+$.) Also examine any disjoint set of sets $\{A_1, A_2, \ldots, A_n\}$ such that all $A_i$ are in $F$. (With disjoint, we mean that $A_i$ and $A_j$ (with $i \neq j$) have no elements in common: $A_i \cap A_j = \emptyset$.) We now say that $P$ is $\sigma$-additive if

$$P(A_1 \cup A_2 \cup \ldots \cup A_n) = \sum_{i=1}^{n} P(A_n). \tag{1.2.1}$$

If we also have that $P(\Omega) = 1$, then we say that $P$ is a **probability measure**. We also say that the triple $(\Omega, F, P)$ is a **probability space**. Such a probability space has several interesting properties.

1. $P(\emptyset) = 0$.
2. If $A_1 \subseteq A_2$, then $P(A_1) \leq P(A_2)$.
3. $P(A_1 \cup A_2 \cup \ldots \cup A_n) \leq \sum_{i=1}^{n} P(A_i)$ for any combination of sets $A_1, \ldots, A_n$.
4. For any $A \in F$, we have $0 \leq P(A) \leq 1$.

## 1.3 Random variables

### 1.3.1 What is a random variable?

Let's suppose we have some experiment, but we don't know its outcome $x$ yet. We can then define $x$ as a **random variable**. If some **event** $\omega_i$ occurs, $x$ will have the value $x(\omega_i)$, whereas if some other event $\omega_j$ occurs, $x$ will have the value $x(\omega_j)$. The events $\omega_i$ are part of the **event space** $\Omega$. $x$ is thus a function from $\Omega$ to $\mathbb{R}$ ($x : \Omega \to \mathbb{R}$).

A rather basic example of a random variable is the **indicator function**. The indicator function $I_A(\omega)$ of a subset $A \in \Omega$ is defined as

$$I_A(\omega) = \begin{cases} 1, & \text{if } \omega \in A \\ 0, & \text{if } \omega \notin A. \end{cases} \tag{1.3.1}$$

A **simple** random variable is a finite linear combination of indicator functions of measurable sets. In other words, if we have a certain combination of sets $A_1, \ldots, A_n \in F$, then the random variable

$$x = \sum_{i=1}^{n} c_i I_{A_i} \tag{1.3.2}$$

is a simple random variable.

### 1.3.2 PDFs and $\sigma$-algebras of a random variable

Every random variable $x$ has a PDF $f_x(u)$ attached to it. Generally speaking, the PDF $f_x(u)$ is the probability that $x(\omega) < u$. If we combine this with our knowledge on probability spaces, we find that

$$f(u) = P(\{\omega \in \Omega | x(\omega) \leq u\}) = P(A) \qquad \text{with} \qquad A = \{\omega \in \Omega | x(\omega) \leq u\}. \tag{1.3.3}$$

What does the above equation mean? Well, we first look at all events $\omega \in \Omega$ for which $x(\omega) \leq u$. We denote the set of all these events by $A$. The value of $f(u)$ now equals the value of the probability measure $P(A)$.

Let's examine a random variable defined on the measurable space $(\Omega, F)$. We denote the set of all possible values of $x$ by $X$. We now say that $x$ takes values in the measurable space $(X, G)$. Here, the set $G$ has a relationship with $F$. In fact, for every set $A \in G$, we have

$$x^{-1}(A) := \{\omega \in \Omega | x(\omega) \in A\} \in F. \tag{1.3.4}$$

We can now also define $x^{-1}(G)$, according to

$$x^{-1}(G) := \{x^{-1}(A) | \forall A \in G\}. \tag{1.3.5}$$

Note that we now must have $x^{-1}(G) \subseteq F$. However, it is not necessarily true that $x^{-1}(G) = F$. But it can be shown that $x^{-1}(G)$ is a $\sigma$-algebra. We define this $\sigma$-algebra as $F^x := F(x) := x^{-1}(G)$. We say that $F^x$ is the **$\sigma$-algebra generated by** $x$.

### 1.3.3 The characteristic function

Consider a random variable $x$ with PDF $f_x(u)$. The **expectation** $E[x]$ of this random variable can now be found using

$$E[x] = \int_{-\infty}^{\infty} v p_x(v)\, dv \quad \text{(for continuous)} \qquad \text{and} \qquad E[x] = \sum v_n p_x(v_n) \quad \text{(for discrete).} \tag{1.3.6}$$

The function $E[.]$ is called the **expectation function**. We use it to define the **characteristic function** $c_x : \mathbb{R}^n \to \mathbb{C}$ of a random variable $x$, according to

$$c_x(u) = E[e^{iu^T x}] = \int_{-\infty}^{\infty} e^{iuv} p(v)\, dv. \tag{1.3.7}$$

In the above equation, $i = \sqrt{-1}$ denotes the complex variable. The characteristic function is quite convenient. If you have it, you can find the corresponding PDF, and vice versa.

### 1.3.4 Gaussian random variables

Previously, we have seen the PDF of a Gaussian distribution. Any random variable $x : \Omega \to \mathbb{R}^n$ with such a PDF is called a **Gaussian random variable** with parameters $m$ and $Q$. This is denoted by $x \in G(m, Q)$. The characteristic function of $x$ has the form

$$c_x(u) = E[e^{iu^T x}] = e^{iu^T m - \frac{1}{2} u^T Q u}. \tag{1.3.8}$$

Let's examine several Gaussian random variables $x_1, \ldots, x_n$. We can put them together in a vector $x^T = [x_1^T \ \ldots \ x_n^T]$. If the new random vector $x$ is also Gaussian (thus satisfying $x \in G(m, Q)$ for some $m, Q$), then we say that $x_1, \ldots, x_n$ are **jointly Gaussian**.

Gaussian random variables have several nice properties. Let's examine a few.

- Every linear combination $y = Ax + b$ of a Gaussian random variable is also a Gaussian random variable. In fact, if $x \in G(m, Q)$, then $y \in G(Am + b, AQA^T)$.
- Let's examine two jointly Gaussian random variables $x$ and $y$. We now have

$$\begin{bmatrix} x \\ y \end{bmatrix} \in G\left( \begin{bmatrix} m_x \\ m_y \end{bmatrix}, \begin{bmatrix} Q_x & Q_{xy} \\ Q_{xy}^T & Q_y \end{bmatrix} \right), \qquad \text{where} \qquad Q_{xy} = E\left[(x - m_x)(y - m_y)^T\right] = Q_{yx}^T. \tag{1.3.9}$$

  If $Q_{xy} = 0$, then $F^x$ and $F^y$ are independent, and vice versa. In other words, when Gaussian random variables are uncorrelated, they are also independent, and vice versa.
- Independent Gaussian random variables are always jointly Gaussian. (The converse is of course not always true.)
- If $y \in G(m, Q)$ and $S = S^T$, then $E[y^T S y] = \mathrm{tr}(SQ) + m^T S m$. (The function tr(.) is the trace of the matrix: the sum of the diagonal elements.)

## 1.4 Conditional expectation

### 1.4.1 Properties of conditional expectation

Let's examine a measurable space $(\Omega, F)$. Also examine a sub-$\sigma$-algebra $G$ of $F$. We now define the **conditional expectation** of $x$ given $G$, denoted by $E[x|G]$, as the random variable $E[x|G]$ that is both $G$ measurable and satisfies

$$E[xI_A] = E[E[x|G]I_A] \tag{1.4.1}$$

for every set $A \in G$. By the way, the random variable $E[x|G](\omega)$ is $G$ measurable if

$$\{\omega \in \Omega | E[x|G](\omega) \leq r\} \in G \quad \text{for all } r \in \mathbb{R}. \tag{1.4.2}$$

There are several properties of the conditional expectation. We will examine a few.

- Let's examine two random variables $x$ and $y$ that are integrable. (This means that $E[|x|]$ and $E[|y|]$ are finite.) Also suppose that we can write $y$ as

$$y = \sum_{k=1}^{n} c_k I_{A_k}, \tag{1.4.3}$$

where $A_1, \ldots, A_n$ is a finite partition of $\Omega$. (In other words, the sets $A_1, \ldots, A_n$ are disjoint, but their union equals $\Omega$.) It can now be shown that

$$E[x|F^y] = \sum_{k=1}^{n} d_k I_{A_k} \qquad \text{where} \qquad d_k = \frac{E[xI_{A_k}]}{E[I_{A_k}]}. \tag{1.4.4}$$

- Let's examine two jointly Gaussian random variables $x$ and $y$. Assume that $Q_y > 0$. We now have

$$E[x|F^y] = m_x + Q_{xy}Q_y^{-1}(y - m_y), \tag{1.4.5}$$

$$E[(x - E[x|F^y])(x - E[x|F^y])^T|F^y] = E[(x - E[x|F^y])(x - E[x|F^y])^T] = Q_x - Q_{xy}Q_y^{-1}Q_{xy}^T, \tag{1.4.6}$$

$$E[e^{iu^T x}|F^y] = e^{iu^T E[x|F^y] - \frac{1}{2}u^T \tilde{Q} u} \qquad \text{for all } u \in \mathbb{R}^n, \tag{1.4.7}$$

$$E[e^{iu^T E[x|F^y]}] = e^{iu^T m_x - \frac{1}{2}u^T Q_{xy}Q_y^{-1}Q_{xy}^T u} \qquad \text{for all } u \in \mathbb{R}^n. \tag{1.4.8}$$

In the above equations, we have used the definition $\tilde{Q} := Q_x - Q_{xy}Q_y^{-1}Q_{xy}^T$.

- Conditional expectation is linear. So,

$$E[c_1 x_1 + x_2 x_2|G] = c_1 E[x_1|G] + c_2 E[x_2|G]. \tag{1.4.9}$$

- If $x \leq y$ for all $\omega \in \Omega$, then $E[x|G] \leq E[y|G]$.
- If $y$ is $G$ measurable, then $E[y|G] = y$.
- If $G_1 \subseteq G_2$, then $E[x|G_1] = E[E[x|G_2]|G_1]$. In particular, if we set $G_1 = \{\emptyset, \Omega\}$ and simply write $G_2 = G$, then this reduces to $E[E[x|G]] = E[x]$.
- If $F^x$ and $G$ are independent sub-$\sigma$-algebras (with respect to $P$), then $E[x|G] = E[x]$. Also, $F^x$ and $G$ are independent if and only if for all $u \in \mathbb{R}$, we have $E[e^{iu^T x}|G] = E[e^{iu^T x}]$.

### 1.4.2 Independence and conditional independence

Let's consider two $\sigma$-algebras $F_1$ and $F_2$. We say that $F_1$ and $F_2$ are independent if $E[x_1 x_2] = E[x_1]E[x_2]$ for all $x_1, x_2 : \Omega \to \mathbb{R}$ for which $F_1$ and $F_2$ are $\sigma$-algebras, respectively.

We can extend this idea to conditional expectations. We say that $F_1$ and $F_2$ are **conditionally independent**, given a sub-$\sigma$-algebra $G$, if

$$E[x_1 x_2|G] = E[x_1|G]E[x_2|G] \tag{1.4.10}$$

for all $x_1, x_2$ with the same conditions as stated earlier. We generally denote this conditional independence by $(F_1, F_2|G) \in CI$. Conditional independence has several properties. In fact, the following four statements are equivalent:

$$(F_1, F_2|G) \in CI, \qquad (F_2, F_1|G) \in CI, \qquad (F_1 \vee G, F_2 \vee G|G) \in CI, \tag{1.4.11}$$

$$E[x_1|F_2 \vee G] = E[x_1|G] \quad \text{for all } x_1 \text{ with } F_1 \text{ as } \sigma\text{-algebra.} \tag{1.4.12}$$

Also, if $F_1$ and $(F_2 \vee G)$ are independent, then also $(F_1, F_2|G) \in CI$.

We can ask ourselves, when are Gaussian random variables conditionally independent? Well, let's consider Gaussian random variables $x$, $y_1$ and $y_2$ with $Q_x > 0$. It can be shown that $(F^{y_1}, F^{y_2}|F^x) \in CI$ if and only if

$$Q_{y_1 y_2} = Q_{y_1 x}Q_x^{-1}Q_{x y_2}. \tag{1.4.13}$$

# 2. 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?

## 2.1 Stochastic processes

### 2.1.1 Definitions of stochastic processes

Consider a probability space $(\Omega, F, P)$ 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 $\omega$) 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.

### 2.1.2 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**

$$W_x(t, s) = E[(x(t) - m_x(t))(x(s) - m_x(s))^T]. \tag{2.1.1}$$

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

$$\sum_{i=1}^{m} \sum_{j=1}^{m} c_i^T W(t_i, t_j) c_j \geq 0. \tag{2.1.2}$$

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.

### 2.1.3 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.

## 2.2 Representing stochastic systems

### 2.2.1 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

$$y(t) = \sum_{i=1}^{n} a_i y(t-i) + \sum_{i=0}^{n} b_i u(t-i-k) + \sum_{i=0}^{n} c_i v(t-i). \tag{2.2.1}$$

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

$$\begin{aligned} x(t+1) &= A(t)x(t) + B(t)u(t) + M(t)v(t), & (2.2.2)\\ y(t) &= C(t)x(t) + D(t)u(t) + N(t)v(t). & (2.2.3) \end{aligned}$$

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.

### 2.2.2 Interconnecting systems

Let's consider figure 2.1. 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

$$\begin{aligned} x_1(t+1) &= A_1x_1(t) + B_1u_1(t), & y_1(t) &= C_1x_1(t) + D_1u_1(t), & (2.2.4)\\ x_2(t+1) &= A_2x_2(t) + B_2u_2(t), & y_2(t) &= C_2x_2(t) + D_2u_2(t), & (2.2.5)\\ x_3(t+1) &= A_3x_3(t) + B_3u_3(t), & y_3(t) &= C_3x_3(t) + D_3u_3(t). & (2.2.6) \end{aligned}$$

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

$$x = \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} \qquad \text{and} \qquad v = \begin{bmatrix} v_2 \\ v_3 \end{bmatrix}. \tag{2.2.7}$$

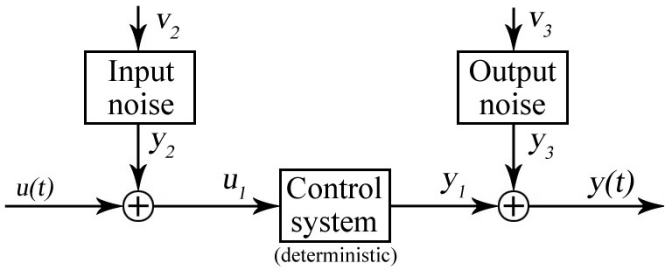


Figure 2.1: The interconnection of Gaussian systems.

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

$$x(t+1) = \begin{bmatrix} A_1 & B_1C_3 & 0 \\ 0 & A_2 & 0 \\ 0 & 0 & A_3 \end{bmatrix} x(t) + \begin{bmatrix} B_1 \\ 0 \\ 0 \end{bmatrix} u(t) + \begin{bmatrix} B_1N_2 & 0 \\ M_2 & 0 \\ 0 & M_3 \end{bmatrix} v(t), \tag{2.2.8}$$

$$y(t) = \begin{bmatrix} C_1 & D_1C_2 & C_3 \end{bmatrix} x(t) + D_1u(t) + \begin{bmatrix} D_1N_2 & N_3 \end{bmatrix} v(t). \tag{2.2.9}$$

### 2.2.3 Stochastic systems in literature

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

$$x(t+1) = A(t)x(t) + M_1(t)r(t), \qquad y(t) = C(t)x(t) + N_1(t)w(t). \tag{2.2.10}$$

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

$$v(t) = \begin{bmatrix} r(t) \\ w(t) \end{bmatrix}, \qquad M(t) = \begin{bmatrix} M_1(t) & 0 \end{bmatrix} \quad \text{and} \quad N(t) = \begin{bmatrix} 0 & N_1(t) \end{bmatrix}. \tag{2.2.11}$$

Our new noise signal is now given by

$$v(t) \in G\left(0, \begin{bmatrix} Q_r(t) & 0 \\ 0 & Q_w(t) \end{bmatrix}\right). \tag{2.2.12}$$

### 2.2.4 Forward and backward representation

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

$$x(t+1) = A^f x(t) + Mv^f(t) \qquad \text{and} \qquad y(t) = C^f(t)x(t) + Nv^f(t). \tag{2.2.13}$$

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

$$A^f(t) = E[x(t+1)x(t)^T]Q(t)^{-1}, \tag{2.2.14}$$

$$C^f(t) = E[y(t)x(t)^T]Q(t)^{-1}, \tag{2.2.15}$$

$$Q_v^f(t) = \begin{bmatrix} Q(t+1) & E[x(t+1)y(t)^T] \\ E[y(t)x(t+1)^T] & E[y(t)y(t)^T] \end{bmatrix} - \begin{bmatrix} A^f(t) \\ C^f(t) \end{bmatrix} Q(t)^{-1} \begin{bmatrix} A^f(t)^T & C^f(t)^T \end{bmatrix}. \tag{2.2.16}$$

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

$$x(t-1) = A^b x(t) + M v^b(t) \qquad \text{and} \qquad y(t-1) = C^b(t)x(t) + N v^b(t). \tag{2.2.17}$$

This time, the system matrices satisfy

$$\begin{aligned}
A^b(t) &= E[x(t-1)x(t)^T]Q(t)^{-1}, && (2.2.18)\\
C^b(t) &= E[y(t-1)x(t)^T]Q(t)^{-1}, && (2.2.19)\\
Q_v^b(t) &= \begin{bmatrix} Q(t-1) & E[x(t-1)y(t)^T] \\ E[y(t)x(t-1)^T] & E[y(t)y(t)^T] \end{bmatrix} - \begin{bmatrix} A^b(t) \\ C^b(t) \end{bmatrix} Q(t)^{-1} \begin{bmatrix} A^b(t)^T & C^b(t)^T \end{bmatrix}. && (2.2.20)
\end{aligned}$$

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

$$\begin{aligned}
A^f(t)Q(t) &= Q(t+1)A^b(t+1)^T, && (2.2.21)\\
A^b(t)Q(t) &= Q(t-1)A^f(t-1)^T, && (2.2.22)\\
C^f(t)Q(t) &= C^b(t+1)Q(t+1)A^b(t+1)^T + NQ_v^b(t+1)M^T, && (2.2.23)\\
C^b(t)Q(t) &= C^f(t-1)Q(t-1)A^f(t-1)^T + NQ_v^f(t-1)M^T. && (2.2.24)\\
& && (2.2.25)
\end{aligned}$$

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

# 3. Properties of stochastic systems

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

## 3.1 Properties of Gaussian system representations

### 3.1.1 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

$$\Phi(t,s) = \begin{cases} A(t)\Phi(t-1,s) & \text{if } s < t, \\ I & \text{if } s = t, \\ 0 & \text{if } s < t. \end{cases} \tag{3.1.1}$$

For time-invariant systems, this can be reduced to

$$\Phi(t,s) = \begin{cases} A^{t-s} & \text{if } s \leq t, \\ 0 & \text{if } s > t. \end{cases} \tag{3.1.2}$$

Second, we define the following notations.

$$F_t^{x+} = \sigma\left(\{x(s), \forall s > t\}\right), \tag{3.1.3}$$

$$F_t^{x(t)} = \sigma\left(\{x(s), \forall s = t\}\right) = \sigma(\{x(t)\}), \tag{3.1.4}$$

$$F_t^x = F_t^{x-} = \sigma\left(\{x(s), \forall s \leq t\}\right). \tag{3.1.5}$$

So, whereas $F_t^{x(t)}$ is the $\sigma$-algebra generated by $x(t)$, $F_t^x$ is the $\sigma$-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

  $$E\left[e^{iu^T x(t+1)} | F_t^{x(t)}\right] = E\left[e^{iu^T x(t+1)} | F_t^x\right]. \tag{3.1.6}$$

  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**.

### 3.1.2 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 $\sigma$-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

$$x(t) = \Phi(t,s)x(s) + \sum_{u=s}^{t-1} \Phi(t-1,u)M(u)v(u), \tag{3.1.7}$$

$$y(t) = C(t)\left(\Phi(t,s)x(s) + \sum_{u=s}^{t-1} \Phi(t-1,u)M(u)v(u)\right) + N(t)v(t). \tag{3.1.8}$$

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

$$m_x(t+1) = A(t)m_x(t) \quad \text{with} \quad m_x(t_0) = m_0, \tag{3.1.9}$$

$$Q(t+1) = A(t)Q(t)A(t)^T + M(t)Q_v(t)M(t)^T \quad \text{with} \quad Q(t_0) = Q_0, \tag{3.1.10}$$

$$W_x(t,s) = E\left[(x(t) - m_x(t))(x(s) - m_x(s))^T\right] = \Phi(t,s)Q(s) \quad (\text{for } t \geq s). \tag{3.1.11}$$

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

$$m_y(t) = C(t)m_y(t), \tag{3.1.12}$$

$$Q_y(t) = C(t)Q(t)C(t)^T + N(t)Q_v(t)N(t)^T, \tag{3.1.13}$$

$$W_y(t,s) = \begin{cases} Q_y(t) = C(t)Q(t)C(t)^T + N(t)Q_v(t)N(t)^T & \text{if } s = t, \\ C(t)\Phi(t,s)Q(s)C(s)^T + C(t)\Phi(t-1,s)M(s)Q_v(s)N(s)^T & \text{if } s < t. \end{cases} \tag{3.1.14}$$

## 3.2 Properties of time-invariant system representations

### 3.2.1 The impulse response function

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

$$x(t+1) = Ax(t) + Bu(t) \quad \text{and} \quad y(t) = Cx(t) + Du(t), \tag{3.2.1}$$

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

$$x(t) = A^{t-t_0}x_0 + \sum_{s=t_0}^{t-1} A^{t-1-s}Bu(s) \quad \text{and} \quad y(t) = CA^{t-t_0}x_0 + \sum_{s=t_0}^{t-1} CA^{t-1-s}Bu(s) + Du(t). \tag{3.2.2}$$

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

$$H(t) = \begin{cases} D, & \text{if } t = 0, \\ CA^{t-1}B, & \text{if } t = 1, 2, \ldots. \end{cases} \tag{3.2.3}$$

Now, if $x_0 = 0$, we have

$$y(t) = \sum_{s=t_0}^{t} H(t-s)u(s). \tag{3.2.4}$$

### 3.2.2 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

$$\text{conmat}(A, B) = \begin{bmatrix} B & AB & A^2B & \dots & A^{n-1}B \end{bmatrix}. \tag{3.2.5}$$

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**

$$\bar{x}(t+1) = \begin{bmatrix} A_{11} & A_{12} \\ 0 & A_{22} \end{bmatrix} \bar{x}(t) + \begin{bmatrix} B_1 \\ 0 \end{bmatrix} u(t) \quad \text{and} \quad y(t) = \begin{bmatrix} C_1 & C_2 \end{bmatrix} \bar{x}(t) + Du(t). \tag{3.2.6}$$

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.

### 3.2.3 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

$$\text{obsm(A,C)} = \begin{bmatrix} C \\ CA \\ CA^2 \\ \vdots \\ CA^{n-1} \end{bmatrix}. \tag{3.2.7}$$

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

$$\bar{x}(t+1) = \begin{bmatrix} A_{11} & 0 \\ A_{21} & A_{22} \end{bmatrix} \bar{x}(t) + \begin{bmatrix} B_1 \\ B_2 \end{bmatrix} u(t) \quad \text{and} \quad y(t) = \begin{bmatrix} C_1 & 0 \end{bmatrix} \bar{x}(t) + Du(t). \tag{3.2.8}$$

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

$$\begin{aligned} \bar{x}(t+1) &= \begin{bmatrix} A_{11} & 0 & A_{13} & 0 \\ A_{21} & A_{22} & A_{23} & A_{24} \\ 0 & 0 & A_{33} & 0 \\ 0 & 0 & A_{43} & A_{44} \end{bmatrix} \bar{x}(t) + \begin{bmatrix} B_1 \\ B_2 \\ 0 \\ 0 \end{bmatrix} u(t), && (3.2.9) \\ y(t) &= \begin{bmatrix} C_1 & 0 & C_3 & 0 \end{bmatrix} \bar{x}(t) + Du(t). && (3.2.10) \end{aligned}$$

### 3.2.4 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 (3.2.6). 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

$$\begin{bmatrix} (sI - A) & B \end{bmatrix}. \tag{3.2.11}$$

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 (3.2.8) is stable. Alternatively, the Hautus test can again be used. Now examine the matrix

$$\begin{bmatrix} (sI - A) \\ C \end{bmatrix}. \tag{3.2.12}$$

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.

### 3.2.5 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

$$\begin{aligned} Q_x &= AQ_xA^T + MQ_vM^T, && (3.2.13) \\ Q_y &= CQ_xC^T + NQ_vN^T, && (3.2.14) \\ Q_{xy} &= AQ_xC^T + MQ_vN^T. && (3.2.15) \end{aligned}$$

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

$$\begin{aligned} W_x(t) &= A^t Q_x, && (3.2.16) \\ W_y(t) &= \begin{cases} CA^{t-1}\left(AQ_xC^T + MQ_vN^T\right) & \text{if } t > 0, \\ Q_y & \text{if } t = 0, \end{cases} && (3.2.17) \\ W_{xy}(t) &= CA^tQ_x. && (3.2.18) \end{aligned}$$

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.

### 3.2.6 The Lyapunov equation

The equation

$$Q = AQA^T + MQ_vM^T \tag{3.2.19}$$

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, $Q > 0$.)

# 4. 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?

## 4.1 Basic ideas of stochastic realizations

### 4.1.1 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

$$z_a = \frac{1}{t_1}\sum_{s=1}^{t_1} z(t) \qquad \text{and} \qquad \hat{W}(t) = \frac{1}{t_1 - t}\sum_{s=1}^{t_1 - t}(z_{t+s} - z_a)(z_s - z_a)^T. \tag{4.1.1}$$

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**.

### 4.1.2 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.

### 4.1.3 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.

## 4.2 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

$$x(t+1) = Fx(t) + Gu(t) \qquad \text{and} \qquad y(t) = Hx(t) + Ju(t), \tag{4.2.1}$$

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.

### 4.2.1 Positive definite functions

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

$$\mathbf{U} = \left\{ u : T \to \mathbb{R}^p \,|\, ||u|| = \sqrt{\sum_{s\in T} u(s)^T u(s)} < \infty \right\}. \tag{4.2.2}$$

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

$$W(t) = \begin{cases} HF^{t-1}G & \text{if } t > 0, \\ 2J & \text{if } t = 0, \\ G^T(F^T)^{-t-1}H^T & \text{if } t < 0. \end{cases} \tag{4.2.3}$$

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

$$(\mathbf{W}(u))(t) = \sum_{s=-\infty}^{t-1} W(t,s)u(s) + \frac{1}{2}W(t,t)u(t). \tag{4.2.4}$$

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

$$u^T\mathbf{W}u = \sum_{s\in T}\sum_{t\in T} u(t)^T W(t,s)u(s) \geq 0. \tag{4.2.5}$$

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

### 4.2.2 Supply rates and storage functions

Let's again consider the system of the form of equation (4.2.1). We define the **supply rate** as

$$h(u(t),y(t)) = h(t) = u(t)^T y(t) = \frac{1}{2}\begin{bmatrix} u(t) \\ y(t) \end{bmatrix}^T J_s \begin{bmatrix} u(t) \\ y(t) \end{bmatrix}, \qquad \text{with} \qquad J_s = \begin{bmatrix} 0 & I_p \\ I_p & 0 \end{bmatrix}. \tag{4.2.6}$$

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

$$S(x(t)) \leq S(x(s)) + \sum_{\tau=s}^{t-1} h(u(\tau),y(\tau)). \tag{4.2.7}$$

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

$$\mathbf{U}(t_0, x_0, t_1, x_1) = \{u \in \mathbf{U} | \text{if } x(t_0) = x_0 \text{ and } u \text{ is applied as input, then } x(t_1) = x_1\}. \tag{4.2.8}$$

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

$$S^-(x) = \sup_{t>0, u \in \mathbf{U}(0,x,t,.)} \left( -\sum_{\tau=0}^{t-1} h(\tau) \right). \tag{4.2.9}$$

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

$$S^+(x) = \inf_{t<0, u \in \mathbf{U}(t,.,0,x)} \left( \sum_{\tau=t}^{-1} h(\tau) \right). \tag{4.2.10}$$

(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 $S(x)$.

### 4.2.3 Characterizing linear dissipative systems

Consider the system of equation (4.2.1). Its dual system is defined as

$$x(t+1) = F^T x(t) + H^T u(t) \qquad \text{and} \qquad y(t) = G^T x(t) + J u(t). \tag{4.2.11}$$

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

$$V_{lsp}(Q) = \begin{bmatrix} Q - F^T Q F & H^T - F^T Q G \\ H - G^T Q F & 2J - G^T Q G \end{bmatrix} \quad \text{and} \quad V_{lsdp}(Q) = \begin{bmatrix} Q - F Q F^T & G - F Q H^T \\ G^T - H Q F^T & 2J - H Q H^T \end{bmatrix}. \tag{4.2.12}$$

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

$$\frac{1}{2} x(t)^T Q x(t) - \frac{1}{2} x(s)^T Q x(s) - \sum_{\tau=s}^{t-1} h(\tau) = -\sum_{\tau=s}^{t-1} \frac{1}{2} \begin{bmatrix} u(t) \\ y(t) \end{bmatrix}^T V_{lsp}(Q) \begin{bmatrix} u(t) \\ y(t) \end{bmatrix}. \tag{4.2.13}$$

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

$$\mathbf{Q_{lsp}} = \left\{ Q \in \mathbb{R}^{n \times n} | Q = Q^T \geq 0 \text{ and } V_{lsp}(Q) \geq 0 \right\}, \tag{4.2.14}$$

$$\mathbf{Q_{lsdp}} = \left\{ Q \in \mathbb{R}^{n \times n} | Q = Q^T \geq 0 \text{ and } V_{lsdp}(Q) \geq 0 \right\}. \tag{4.2.15}$$

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, $S(x) = \frac{1}{2}x^T Q x$ 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**

$$D(Q) = \left(Q - F^T Q F\right) - \left(H^T - F^T Q G\right)\left(2J - G^T Q G\right)^{-1}\left(H - G^T Q F\right) = 0. \tag{4.2.16}$$

(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

$$S^-(x) = \frac{1}{2}x^T Q^- x \qquad \text{and} \qquad S^+(x) = \frac{1}{2}x^T Q^+ x. \tag{4.2.17}$$

## 4.3 Stochastic realizations

### 4.3.1 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?

$$x(t+1) = Ax(t) + Mv(t) \qquad \text{and} \qquad y(t) = Cx(t) + Nv(t). \tag{4.3.1}$$

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

$$H_W = \begin{bmatrix} W(1) & W(2) & \cdots & W(k_2) \\ W(2) & \ddots & \ddots & \vdots \\ \vdots & \ddots & \ddots & W(k_1+k_2-2) \\ W(k_1) & \cdots & W(k_1+k_2-2) & W(k_1+k_2-1) \end{bmatrix}, \tag{4.3.2}$$

$$\operatorname{rank}(H_W) = \sup_{k_1,k_2\in\mathbb{Z}} \operatorname{rank}(H_W(k_1,k_2)). \tag{4.3.3}$$

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

$$W(t) = \begin{cases} HF^{t-1}G & \text{if } t > 0, \\ 2J & \text{if } t = 0, \\ G^T(F^T)^{-t-1}H^T & \text{if } t < 0. \end{cases} \tag{4.3.4}$$

(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

$$x(t+1) = Ax(t) + Mv(t) \qquad \text{and} \qquad y(t) = Cx(t) + Nv(t). \tag{4.3.5}$$

### 4.3.2 Properties of Gaussian system representation

Let's examine a Gaussian system representation of the form of equation (4.3.5). 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 $\text{spec}(A) \subset \mathbb{D}_o$. 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 $\text{spec}(A - KC) \subset \mathbb{D}_o$.

### 4.3.3 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

$$x(t+1) = Ax(t) + Kw(t) \qquad \text{and} \qquad y(t) = Cx(t) + w(t), \tag{4.3.6}$$

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

# 5. Stochastic control

In this chapter, we're going to examine how we can control a system using stochastic control. First, we examine some basic principles of stochastic control. After that, we'll look at an example of a field where stochastic control theory can often be applied: statistical decision problems.

## 5.1 Basic principles of stochastic control

### 5.1.1 Information structures

Let's examine the system

$$x(t+1) = Ax(t) + Bu(t) + Mv(t), \tag{5.1.1}$$
$$y(t) = Cx(t) + Du(t) + Nv(t). \tag{5.1.2}$$

We can now use the input $u(t)$ to control the system. To control the system, we need information. But luckily, information is available. We say that, at every time $t \in T$, the $\sigma$-algebra $G_t$ specifies the available information. The family of all such $\sigma$-algebras $G_t$, being $\{G_t, t \in T\}$, is called the **information structure**.

There are different types of information structures. The type depends on which data is available.

- The **past-output information structure** is $\left\{F_{t-1}^{y-}, t \in T\right\}$. So, we have all previous outputs available as data. (But the states $x$ are not available.)
- The **current output information structure** is $\left\{F^{y(t)}, t \in T\right\}$. So, only the current output is available.
- The **past-state information structure** is $\left\{F_t^{x-}, t \in T\right\}$. So, all previous states are available.
- The **Markov information structure** (also called the **current state information structure**) is $\left\{F^{x(t)}, t \in T\right\}$. So, only the current state is available.

### 5.1.2 Control laws

Based on the information structure, we can make a control law. A **control law** (also called **control policy**) is a measurable mapping from the available data to the input space $U$. How the control law is called depends on the information structure that is used. We will examine the most important control laws now.

- The **past-output control law** uses the past-output information structure. So, for every $t \in T$, $g_t$ is a measurable map $g_t : Y^t \to U$. This implies that we can also see $g$ itself as a mapping $g : T \times Y^t \to U$.
- The **output control law** uses the current output information structure. So we now have $g : T \times Y \to U$.
- The **past-state control law** uses the past-state information structure. So now $g : T \times X^{t+1} \to U$.
- The **Markov control law** uses the Markov information structure. So now $g : T \times X \to U$. We denote the set of all possible Markov control laws by $G_M$.
- The **stationary Markov control law** also uses the Markov information structure. But now the control law $g$ does not depend on time. So, $g : X \to U$.

Let's suppose that we use a control law $g$. For example, we use the Markov control law. The resulting control system parameters are then written with the superscript $g$. So, we have

$$x^g(t+1) = Ax^g(t) + Bu^g(t) + Mv(t), \tag{5.1.3}$$
$$y^g(t) = Cx^g(t) + Du^g(t) + Nv(t). \tag{5.1.4}$$

In this equation, we have $u^g(t) = g(t, x^g(t))$. The above system representation is called a **closed-loop stochastic control system**.

### 5.1.3 Control objectives

The question remains, which control law do we use? We usually choose a control law such that control objectives are met. A **control objective** is a property of the closed-loop control system which we should strive to attain. Examples of control objectives are

- Making the system **stable**.
- **Suppressing noise**.
- Optimizing a **performance measure**. For example, we might want to choose $g$ such that a cost function $J$ is minimized. The cost function $j$ can then have a form like

$$J = E\left[\sum_{t=0}^{t_1} (c_1 x(t) + c_2 u(t))\right]. \tag{5.1.5}$$

- Making the system **robust**. Robustness means that, even when deviations are applied in the model, the system still has a satisfactory performance.

The **stochastic control problem** is now defined as the problem of finding a control law $g$ such that the control objectives are satisfied as well as possible. Solving this problem consists of two steps. First, in **control synthesis**, possible control laws $g$ need to be generated. Then, in **control design**, the best of these control laws $g$ needs to be chosen. In practice, this often means that the numerical parameters of the control law need to be chosen.

## 5.2 Statistical decision problems

### 5.2.1 Statistical decision problems

**Statistical decision problems** are often good examples of stochastic control problems. Let's suppose that we have $x$ money. We can invest this in 2 investment opportunities. (It works the same when there are more investment opportunities.) The two investment opportunities return $y_1 = r_1 x_1$ and $y_2 = r_2 x_2$, respectively, where $r_1$ and $r_2$ are random variables and $x_1$ and $x_2$ are the amount of money invested in $r_1$ and $r_2$, respectively. The total return which you get is thus $y = y_1 + y_2$. The question now is how to invest $x$. Which fraction $u$ should we invest in $r_1$ and which fraction $(1 - u)$ should we invest in $r_2$?

To solve this problem, we need a **utility function** $U(y)$. This utility function is a measure of how 'happy' you are with a return $y$. For most normal people, this is a concave function. (That is, $d^2U/dy^2 < 0$. Initially, people are very happy when they get more money. But, as people get richer, the extra happiness decreases if they get more money.) Of course, $y$ is a random variable as well. So, we need to select $u$ such that the expected utility $E[U(y)]$ is maximized.

### 5.2.2 An example

Let's demonstrate the above procedure with an example. Let's say that opportunity 1 is a 'sure' investment opportunity: $r_1$ is always 1.5. On the other hand, opportunity 2 is a 'risky' investment opportunity: there is a chance of 50% that $r_2 = 3$, but also a chance of 50% that $r_2 = 1$. We also define the utility function as $U(y) = 6y - y^2$. We now have

$$y = r_1 x_1 + r_2 x_2 = r_1 u x + r_2 (1 - u) x. \tag{5.2.1}$$

We should thus maximize

$$E[U(y)] = E\left[6(r_1 ux + r_2(1-u)x) - (r_1 ux + r_2(1-u)x)^2\right]. \tag{5.2.2}$$

Using the data given for $r_1$ and $r_2$ gives

$$E[U(y)] = \frac{1}{2}\left(6(1.5ux + (1-u)x) - (1.5ux + (1-u)x)^2\right) + \frac{1}{2}\left(6(1.5ux + 3(1-u)x) - (1.5ux + 3(1-u)x)^2\right). \tag{5.2.3}$$

We can find the maximum of this equation by differentiating for $u$. This then shows that the maximum occurs at $u = 2/5$.

### 5.2.3 Risk

Risk plays an important role in statistical decision problems. We say that a decision makes with a utility function $U(y)$ is **risk averse** if $E[U(y)] < U(E[y])$. That is, he prefers the certain pay-off $U(E[y])$ above the uncertain pay-off $E[U(y)]$. Similarly, the decision maker is **risk preferring** if $E[U(y)] > U(E[y])$ and **risk neutral** if $E[U(y)] = U(E[y])$.

We can also define the **index of absolute risk aversion** $r(y)$. Assuming that $U$ is twice differentiable, it is defined as

$$r(y) = -\frac{U''(y)}{U'(y)}. \tag{5.2.4}$$

The index $r(E[y])$ is roughly proportional to the amount of money one would pay to avoid risks. Thus, if $r(E[y]) > 0$, then the person would pay to avoid risks and is thus risk averse. Similarly, if $r(E[y]) < 0$, the person would pay to have risks and is thus risk preferring.

# 6. Dynamic programming

The most important method to find the optimal control law for a recursive state-observed system is dynamic programming. In this chapter, we'll look at how it works. In the first part, we look at finite horizons. That is, time is limited. In the second part, we examine what to do if time can run on indefinitely.

## 6.1 Dynamic programming on a finite horizon

### 6.1.1 The problem and the cost function

Let's consider a recursive state-observed stochastic system, described by

$$x(t+1) = f(t, x(t), u(t), v(t)). \tag{6.1.1}$$

We assume that the time horizon is finite. So, $T = \{0, \ldots, t_1\}$. To control this system, we use the input $u(t)$. This input is given by the control law $g$, according to

$$u^g(t) = g(t, x^g(t), x^g(t-1), \ldots, x^g(t_0)). \tag{6.1.2}$$

The resulting state of the system is denoted by $x^g(t)$.

The question arises: which input $u(t)$ should we select? Our goal is to control the system in such a way that a **cost function** $J(g)$ is minimized. An example of such a cost function is

$$J(g) = E_g \left[ \sum_{s=0}^{t_1-1} b(s, x^g(s), u^g(s)) + b_1(x^g(t_1)) \right]. \tag{6.1.3}$$

Here, the $E_g$ operator means the expectation, given that the control law $g$ is used. Also, $b_1$ is the **terminal cost** and $b$ is the **current cost**. We thus need to find the optimal control law $g^*$, such that $J^* = J(g^*) = \inf_{g \in G} J(g)$. (Note that $G$ is the set of all possible control laws $g$.) To find $g^*$, a helpful function is the **conditional cost-to-go** $J(g,t)$ at time $t$. It is defined as

$$J(g,t) = E_g \left[ \sum_{s=t}^{t_1-1} b(s, x^g(s), u^g(s)) + b_1(x^g(t_1)) | F_t^{x^g} \right]. \tag{6.1.4}$$

It is interesting to note that we have $J(g) = E_g[J(g, t_0)]$.

### 6.1.2 The dynamic programming procedure

To find $g^*$, we make use of a **value function** $V(t, x^g(t))$. This value function satisfies $V(t, x^g(t)) \leq J(g,t)$. In fact, if we use the optimal control law $g = g^*$, then we have an equality. In other words, we have

$$J^* = J(g^*) = E_{g^*}[J(g^*, t_0)] = E_{g^*}[V(t_0, x_0)]. \tag{6.1.5}$$

But how do we find $V$? The value function is generally derived by backward recursion. First, we define $V$ at time $t = t_1$. So, $V(t_1, x) = b_1(x)$ for all $x \in X$. We then find $V$ recursively using

$$V(t,x) = \inf_{u \in U} \left( b(t,x,u) + E[V(t+1, f(t,x,u,v(t)))|F^{x,u}] \right). \tag{6.1.6}$$

The above equation is known as the **dynamic programming equation**. Once we have the value function, we can find the optimal control law $g^*$. We simply select the $u$ for which the above relation is minimized. This value of $u$ is denoted by $u^*$. We thus have

$$g^*(t,x) = u^* = \arg \inf_{u \in U} \left( b(t,x,u) + E[V(t+1, f(t,x,u,v(t)))|F^{x,u}] \right). \tag{6.1.7}$$

It can now be shown that the value function must satisfy $V(t_1, x^{g^*}(t_1)) = b_1(x^{g^*}(t_1))$ and

$$V(t, x^{g^*}(t_1)) = b(t, x^{g^*}(t), g^*(t, x^{g^*}(t))) + E_{g^*}\left[V(t+1, f(t, x^{g^*}(t), g^*(t, x^{g^*}(t)), v(t)))|F_t^{x^{g^*}}\right]. \quad (6.1.8)$$

Note that in the above **dynamic programming (DP) procedure**, we don't directly try to find $g^*$ from the set $G$. Instead, we simply find the optimal value of $u^*$ from the set $U$ at every time-step. This is generally much easier.

### 6.1.3 Linear quadratic-Gaussian stochastic control

Let's consider a special case of the above problem. First of all, we examine a Gaussian stochastic control system

$$x(t+1) = A(t)x(t) + B(t)u(t) + M(t)v(t). \quad (6.1.9)$$

Also, we use the quadratic cost function

$$J(g) = E_g\left[\sum_{s=t}^{t_1-1}\begin{bmatrix}x(s)\\u(s)\end{bmatrix}^T L(s)\begin{bmatrix}x(s)\\u(s)\end{bmatrix} + x(t_1)^T Q_1 x(t_1)\right], \quad \text{with} \quad L(t) = \begin{bmatrix}Q(t) & S(t)\\S(t)^T & R(t)\end{bmatrix} \quad (6.1.10)$$

and with $Q_1$ being the terminal cost. Also, we have $L(t) = L(t)^T \geq 0$ and $R(t) > 0$. For this case, we can derive an analytic expression for $g^*(t)$. First we define $P(t+1) = Q_1$. Now we recursively define

$$\begin{aligned}
H_{11}(t) &= A(t)^T P(t+1)A(t) + Q(t), & (6.1.11)\\
H_{12}(t) &= A(t)^T P(t+1)B(t) + S(t), & (6.1.12)\\
H_{22}(t) &= B(t)^T P(t+1)B(t) + R(t), & (6.1.13)\\
F(t) &= -H_{22}^{-1}(t)H_{12}^T(t), & (6.1.14)\\
P(t) &= H_{11}(t) - H_{12}(t)H_{22}^{-1}(t)H_{12}^T(t). & (6.1.15)
\end{aligned}$$

In this way, the above matrices can be determined for every time $t$. The optimal control law $g^*(t, x)$ can now be found using

$$g^*(t, x) = F(t)x. \quad (6.1.16)$$

The corresponding value function and optimal cost function can be found using

$$\begin{aligned}
V(t, x) &= x^T P(t)x + r(t), & (6.1.17)\\
r(t) &= r(t+1) + \text{trace}(M(t)^T P(t+1)M(t)V(t), & (6.1.18)\\
J^* &= E_g\left[x_0^T P(t_0)x_0\right] + r(t_0). & (6.1.19)
\end{aligned}$$

## 6.2 Dynamic programming on an infinite horizon

### 6.2.1 Cost functions

This time, we consider the case where we have an **infinite horizon**. So, $T = \mathbb{N} = \{0, 1, \ldots\}$. For the infinite-horizon case, variations in the system dynamics are often negligible. So we assume that the system is time-invariant, implying that

$$x(t+1) = f(t, x(t), u(t), v(t)). \quad (6.2.1)$$

The problem with infinite-horizon problems is that the cost function $J(g)$ often becomes infinite. To solve this problem, we have to use a different cost function. An example of such a cost function is the

**discounted cost function**, being

$$J_d(g) = E_g\left[\sum_{s=0}^{\infty} r^s b(x^g(s), u^g(s))\right], \tag{6.2.2}$$

where $r$ is the **discount rate**. With this cost function, costs that occur in the future are weighted less. Another option for a cost function is the **average cost function**. This is defined as

$$J_{av}(g) = \lim_{t\to\infty}\left(\frac{1}{t}E_g\left[\sum_{s=0}^{t-1} b(x^g(s), u^g(s))\right]\right). \tag{6.2.3}$$

Which cost function is the most suitable depends on the problem. In the rest of this chapter, we will mainly consider the discounted cost function.

### 6.2.2 The transition measure

For simplicity, we will consider the case where the input space $U$ and the state space $X$ are both finite. We also assume that we can only use Markov control laws $u^g(t) = g(t, x^g(t))$. In this case, we can write the system dynamics in a completely different form. We don't use the function $f$ anymore. Instead, we denote the chance that $x(t+1) = i$, given that $x(t) = j$ and $u(t) = u$, by

$$P(i, j, u) = P(x(t+1) = i | x(t) = j, u(t) = u). \tag{6.2.4}$$

Here, $P$ is the **transition measure**. We can also use this transition measure for conditional expectation. We then get

$$E_g\left[V(t+1, x^g(t+1)) | F^{x^g(t)}\right] = \sum_{x_1\in X} V(t+1, x_1) P(x_1, x^g(t), u_g(t)). \tag{6.2.5}$$

### 6.2.3 The forwardly defined value function

For the problem we are considering, the backwardly defined cost-to-go function $W(t, x)$ is defined as $W(t_1, x) = 0$ and

$$W(t, x) = \min_{u\in U}\left(r^t b(x, u) + \sum_{x_1\in X} W(t+1, x_1) P(x_1, x, u)\right). \tag{6.2.6}$$

However, in our case, there is no ending time $t_1$. So we would better consider the forwardly defined value function $V(t, x)$. It is defined as $V(0, x) = 0$ and

$$V(t, x) = \min_{u\in U}\left(b(x, u) + r\sum_{x_1\in X} V(t-1, x_1) P(x_1, x, u)\right). \tag{6.2.7}$$

It can be noted that we have $V(t, x) = r^{t-t_1} W(t_1 - t, x)$. As time $t$ goes to infinity, it can be shown that $V(t, x)$ converges. In fact, if we write $V(\infty, x) = V(x)$, then we have

$$V(x) = DP(V)(x) = \min_{u\in U}\left(b(x, u) + r\sum_{x_1\in X} V(x_1) P(x_1, x, u)\right). \tag{6.2.8}$$

The above relation has a unique solution for the function $V(x)$. And it can also be shown that the resulting value function $V(x)$ minimizes the discounted cost function. So,

$$V(x_0) = \inf_{g\in G} E_g\left[\sum_{s=0}^{\infty} r^s b(x^g(s), u^g(s)) | F^{x_0}\right]. \tag{6.2.9}$$

Once we have the value function $V(x)$, the optimal control law can be found using

$$g^*(x) = \arg\min_{u \in U} \left( b(x,u) + r \sum_{x_1 \in X} V(x_1) P(x_1, x, u) \right). \tag{6.2.10}$$

Note that this control law does not depend on time $t$. So, it is a stationary control law.

### 6.2.4 Algorithms for finding the value function

The question remains, how can we find $V(x)$? There are multiple methods. The first one we consider is **value iteration**. In this method, we make use of a function $h(m,x)$ which is similar to $V(t,x)$. We initialize $h(0,x) = 0$ for all $x$. We then keep on updating $h(t,x)$, according to

$$h(m+1,x) = \min_{u \in U} \left( b(x,u) + r \sum_{x_1 \in X} h(m, x_1) P(x_1, x, u) \right). \tag{6.2.11}$$

We know that, as $m \to \infty$, then $V(x) = h(m,x)$. The downside of this method is that we need infinitely many computations before convergence occurs.

The method of **policy improvement** solves the above problem. Before we discuss this algorithm, we make some definitions though. Let's suppose that the state space $X$ has $n$ elements $x_1, \dots x_n$. We define the vectors $V$ and $b(g)$ as

$$V = \begin{bmatrix} V(x_1) \\ V(x_2) \\ \vdots \\ V(x_n) \end{bmatrix} \qquad \text{and} \qquad b(g) = \begin{bmatrix} b(x_1, g(x_1)) \\ b(x_2, g(x_2)) \\ \vdots \\ b(x_n, g(x_n)) \end{bmatrix}. \tag{6.2.12}$$

Also, let's define the matrix $P(g)$ as the matrix with elements $P_{ij}(g) = P(x_j, x_i, g(x_i))$. In other words, if we are in a state $x_i$ and use an input $g(x_i)$, then the $i$th row gives us the chances that we reach a state $x_j$ in the next time step, with $j$ the column number. (Note that the sum of the elements in every row of $P(g)$ is 1.) We now need to find a control policy $g$ such that

$$V(g) = DP(V(g)) = b(g) + rP(g)V(g). \tag{6.2.13}$$

To do this, we first take a random control law $g_0$ and solve the equation $V(g_0) = b(g_0) + rP(g_0)V(g_0)$ for $V(g_0)$. (Note that this simply is a linear equation.) We then derive a new control law $g_m$ which satisfies

$$b(g_m) + rP(g_m)V(g_{m-1}) = \min_{g \in G} \left( b(g) + rP(g)V(g_{m-1}) \right). \tag{6.2.14}$$

Note that the set of possible control laws $G$ has a finite size, because the state space $X$ and the input space $U$ are finite. So the above relation can be solved in finite time. Once we have the new control law $g_m$, we can also update our value function $V(g_m)$ by solving

$$V(g_m) = b(g_m) + rP(g_m)V(g_m). \tag{6.2.15}$$

This algorithm continues as long as improvements are made to the policy. That is, as long as

$$\min_{g \in G} \left( b(g) + rP(g)V(g_{m-1}) \right) < V(g_{m_1}). \tag{6.2.16}$$

If the above condition does not hold anymore, then the algorithm has converged to the optimal policy (or one of the optimal policies, in case there are multiple ones). This optimal policy is then simply given by $g^* = g_m$.

# 7. Kalman filters

Kalman filters are very good at finding the state $x$ of a system. But what kinds of Kalman filters are there? How do they work? And what are their properties? That's what this chapter is about.

## 7.1 The time-varying Kalman filter

### 7.1.1 The Kalman filtering problem

Let's suppose that we are observing some stochastic process $y$. Based on this process, we should estimate another stochastic process $z$. In other words, we need to determine $z(t)$ given $F_{t-1}^y$. A special case of this problem is the **Kalman filtering problem**. Now $z$ (or $x$) equals the state of some system with output $y$. In other words, given the system

$$\begin{aligned} x(t+1) &= A(t)x(t) + M(t)v(t), \quad x(t_0) = x_0 \in G(m_0, Q_0)m & (7.1.1)\\ y(t) &= C(t)x(t) + N(t)v(t), \quad v(t) \in G(0, Q_v(t)), & (7.1.2) \end{aligned}$$

we need to determine the conditional distribution of $x(t)$ given $F_{t-1}^y$.

### 7.1.2 The time-varying Kalman filter

To solve this problem, we can use a **time-varying Kalman filter**. However, this only works if $N(t)Q_v(t)N(t)^T > 0$. (If this is not the case, then we can split up the system. The condition then holds for the first part of the system, while the second part is not affected by the noise.) In this case, the distribution of $x(t)|F_{t-1}^y$ is Gaussian and is specified by the characteristic function

$$E[\exp(iw^T x(t))|F_{t-1}^y] = \exp(iw^T \hat{x}(t) - \frac{1}{2}w^T Q_f(t)w), \ \forall w \in \mathbb{R}^n. \tag{7.1.3}$$

In the above equation, we call $\hat{x}$ the **conditional mean process** and $Q_f$ the **conditional variance process**.

But how do we find $\hat{x}$ and $Q_f$? For that, we can use the recursively defined equations

$$\begin{aligned} \hat{x}(t+1) &= A(t)\hat{x}(t) + K(t)[y(t) - C(t)\hat{x}(t)], \ \hat{x}(t_0) = E[x_0] = m_0, & (7.1.4)\\ H_{11}(t) &= A(t)Q_f(t)A(t)^T + M(t)Q_v(t)M(t)^T, & (7.1.5)\\ H_{12}(t) &= A(t)Q_f(t)C(t)^T + M(t)Q_v(t)N(t)^T, & (7.1.6)\\ H_{22}(t) &= C(t)Q_f(t)C(t)^T + N(t)Q_v(t)N(t)^T, & (7.1.7)\\ K(t) &= H_{12}(t)H_{22}^{-1}(t), & (7.1.8)\\ Q_f(t+1) &= H_{11}(t) - H_{12}(t)H_{22}^{-1}(t)H_{12}(t)^T, \ Q_f(t_0) = Q_0. & (7.1.9) \end{aligned}$$

If we use the above relations, then we have

$$\hat{x}(t) = E[x(t)|F_{t-1}^y] \qquad \text{and} \qquad Q_f(t) = E[(x(t) - \hat{x}(t))(x(t) - \hat{x}(t))^T|F_{t-1}^y]. \tag{7.1.10}$$

### 7.1.3 Related process

When applying the Kalman filter, we can define the **innovation process** $\bar{v}(t)$ as $\bar{v}(t) = y(t) - C(t)\hat{x}(t)$. Now $\bar{v}(t)$ is a Gaussian white noise process, with

$$\bar{v}(t) \in G(0, Q_{\bar{v}}(t)), \quad Q_{\bar{v}}(t) = H_{22}(t) = C(t)Q_f(t)C(t)^T + N(t)Q_v(t)N(t)^T. \tag{7.1.11}$$

Given this white noise process, the output $y(t)$ can be determined. In other words, we have $F_t^y = F_t^{\bar{v}}$. Next to this, the error process $e(t)$ is defined as $e(t) = x(t) - \hat{x}(t)$. Its recursive equation is

$$e(t+1) = (A(t) - K(t)C(t))e(t) + (M(t) - K(t)N(t))v(t). \tag{7.1.12}$$

## 7.2 The time-invariant Kalman filter

### 7.2.1 The Filter algebraic Ricatti equation

Previously, we considered the case where the state space matrices depended on time. Now let's assume that $A$, $C$, $M$, $N$ and $Q_v$ are constant in time. In this case, the time-varying Kalman filter may converge to a **time-invariant Kalman filter**. To see when this happens, we first make some assumptions.

- Assume that $NQ_vN^T > 0$.
- Define $F$ and $G$ such that

$$\begin{aligned} F &= A - MQ_vN^T(NQ_vN^T)^{-1}C, & (7.2.1)\\ GG^T &= MQ_vM^T - MQ_vN^T(NQ_vN^T)^{-1}(MQ_vN^T)^T. & (7.2.2) \end{aligned}$$

  Assume that $(F, G)$ is a stabilizable pair.
- Assume that $(A, C)$ is a detectable pair.

Let's define the function $f(Q)$ as

$$f(Q_f) = AQ_fA^T + MQ_vM^T - (AQ_fC^T + MQ_vN^T)(CQ_fC^T + NQ_vN^T)^{-1}(AQ_fC^T + MQ_vN^T)^T. \tag{7.2.3}$$

The above assumptions now imply that there exists a unique solution $Q_f = Q_f^T \geq 0$ to the **Filter algebraic Riccati equation** $Q_f = F(Q_f)$. This solution $Q_f$ is actually the limit $\lim_{t\to\infty} Q_f(t)$ of the time-varying Kalman filter. Also, if $(F, G)$ is controllable as well, then we have $Q_f = Q_f^T > 0$.

### 7.2.2 The time-invariant Kalman filter

Let's define the matrices $K(Q_f)$ and $A(Q_f)$ as

$$\begin{aligned} K(Q_f) &= (AQ_fC^T + MQ_vN^T)(CQ_fC^T + NQ_vN^T)^{-1}, & (7.2.4)\\ A(Q_f) &= A - K(Q_f)C. & (7.2.5) \end{aligned}$$

It can be shown that $\text{spec}(A(Q_f)) \subset \mathbb{D}_o$. The time-invariant Kalman filter is now given by

$$\bar{x}(t+1) = A\bar{x}(t) + K(y(t) - C\bar{x}(t)). \tag{7.2.6}$$

It is important to note that we don't denote this filter by $\hat{x}(t)$. This is because generally $\bar{x} \neq E[x(t)|F_{t-1}^y]$. But if this is the case, then why would we use the time-invariant Kalman filter, and not the time-varying Kalman filter.

The answer is that a filter with a time-varying matrix $K$ (like the time-varying Kalman filter) is quite complicated. A filter with a constant matrix $K$ is a lot simpler. And it can be shown that the time-invariant Kalman filter is better than any other filter with a constant matrix $K$. With 'better' we mean that the variance of the error $e(t) = x(t) - \bar{x}(t)$ of the Kalman filter is always better (or just as good) as any other filter.

# 8. Control using partial observations

Sometimes we may need to control a system of which we don't know the state. This sounds like a completely impossible problem. But, by using the output of the system as well as possible, it is still possible to do this. How it works is explained below.

## 8.1 Information states and the separation principle

Consider a stochastic control system of the form

$$\begin{aligned} x(t+1) &= f(t, x(t), u(t), v(t)), & (8.1.1)\\ y(t) &= h(t, x(t), u(t), v(t)). & (8.1.2) \end{aligned}$$

Our task now is to find a control law $g$ which minimizes a cost function like

$$J(g) = E_g \left[ \sum_{s=0}^{t_1-1} b(s, x^g(s), u^g(s)) + b_1(x^g(t_1)) \right]. \quad (8.1.3)$$

To determine the control law $g$, we need information. We can summarize this information in a so-called **information state**. An example of an information state is

$$z(t) = [y(t-1), \ldots, y(0), u(t-1), \ldots, u(0)]^T. \quad (8.1.4)$$

There is just one problem: this information state grows with time. An information state which does not grow with time is generally more convenient to use.

The information state is descibed by its recursive relation

$$z(t+1) = f_1(t, z(t), y(t), u(t)). \quad (8.1.5)$$

We say that an information state is a **sufficient information state for the state given past outputs and inputs** if it provides sufficient information about the state $x$ of the system. To be more precise, if

$$E[\exp(iw^T x(t))|F_{t-1}^y \vee F_{t-1}^u] = E[\exp(iw^T x(t))|F^{z(t)} \vee F^{u(t-1)}]. \quad (8.1.6)$$

We now say that a control law $h_1$ is a **separated control law** if it uses a sufficient information state $z$. That is, if

$$u(t) = h_1(t, z(t), y(t)). \quad (8.1.7)$$

The set of all separated control laws $g$ is denoted by $G_s \subset G$. An optimal stochastic control problem is said to have the **separation property** if the optimal control law is a separated control law.

It is nice if a problem has the separation property, because separated control laws are relatively easy to find. In fact, sometimes we only search for separated control laws, even when we are not dealing with an optimal stochastic control problem. This means that we do not find the most optimal control law, but we do find a control law close to the optimal one in a relatively easy way. This principle is calle the **separation principle**.

## 8.2 Control of partially observed systems

### 8.2.1 The problem statement

Consider the system

$$\begin{aligned} x(t+1) &= A(t)x(t) + B(t)u(t) + M(t)v(t), & (8.2.1)\\ y(t) &= C(t)x(t) + D(t)u(t) + N(t)v(t). & (8.2.2) \end{aligned}$$

Now also suppose that we don't know the state $x$ of the system. But our goal is to come up with a control law $g$ which minimizes a cost function $J(g)$. How do we do this?

It turns out that this problem has a very special property. We have

$$F_t^{y,g} \vee F_t^{u,g} = F_t^{y,0} \vee F_t^{u,0}. \tag{8.2.3}$$

In this equation $F_t^{y,0}$ is the $\sigma$-algebra obtained if we use the control law $u = 0$. In other words, it does not matter which control law we use, we always get the same information from the system.

So what does this mean? It means that we can simply use a Kalman filter to estimate the state. And this works for every control law $g$ which we might use. We then get a state estimate $\hat{x}^g(t)$. The Kalman filter used is

$$\begin{aligned}
\hat{x}^g(t+1) &= A(t)\hat{x}^g(t) + B(t)u^g(t) + K(t)\left(y(t) - C(t)\hat{x}^g(t) - D(t)u^g(t)\right), && (8.2.4)\\
I_{11}(t) &= A(t)Q_f(t)A(t)^T + M(t)Q_v(t)M(t)^T, && (8.2.5)\\
I_{12}(t) &= A(t)Q_f(t)C(t)^T + M(t)Q_v(t)N(t)^T, && (8.2.6)\\
I_{22}(t) &= C(t)Q_f(t)C(t)^T + N(t)Q_v(t)N(t)^T, && (8.2.7)\\
Q_f(t+1) &= I_{11}(t) - I_{12}(t)I_{22}^{-1}(t)I_{12}(t)^T, && (8.2.8)\\
K(t) &= I_{12}(t)I_{22}^{-1}(t). && (8.2.9)
\end{aligned}$$

Now all that is left to do is find a good control law $g$. And for that, we can simply use the state estimate $\hat{x}$.

### 8.2.2 Finding the value function

In this problem, we are using Gaussian stochastic parameters. So, given the mean $m$ and the variance $Q$, these Gaussian parameters have a PDF of

$$G(w; m, Q) = \frac{1}{\sqrt{2\pi \det(Q)}} \exp\left(-\frac{1}{2}(m-w)^T Q^{-1}(m-w)\right), \tag{8.2.10}$$

where $w$ is the variable of the PDF. We need to use this function to find the value function $V(t, x)$ for our problem. This function is defined recursively. At the final time $t_1$, it can be found using

$$V(t_1, x) = \int_X b_1(w) G(w; x, Q(t_1))\, dw. \tag{8.2.11}$$

The subscript $X$ means we integrate over the entire state space $X$. To find the value function at earlier times, we use the recursive relation

$$V(t, x) = \inf_{u \in U}\left(\int_X b(t, w, u) G(w; x, Q(t))\, dw + \int_X V(t+1, w) G(w; A(t)x + B(t)u, Q_u(t+1))\right). \tag{8.2.12}$$

The parameter $Q_u(t)$ is the variance matrix obtained at time $t+1$ if we choose an input $u$. Now, for every control law $g \in G$, we have $V(t, \hat{x}^g(t)) \leq J(g, t)$.

The question remains, how do we find the optimal control law? Given that we are in a time $t$ and an estimated state $\hat{x}$, we simply need to select the input $u$ which minimizes the above relation. If, for all $x \in X$, there is an input $u^* \in U$ which minimizes the above relation, then the map $h_1(t, x) = u^*$ forms the optimal control law.

## 8.3 Special cost functions

### 8.3.1 The quadratic cost function

Let's examine some special cases. In fact, let's examine the quadratic cost function

$$J(g) = E_g \left[ \sum_{s=0}^{t_1-1} \begin{bmatrix} x^g(s) \\ u^g(s) \end{bmatrix}^T L(s) \begin{bmatrix} x^g(s) \\ u^g(s) \end{bmatrix} + x^g(t_1)^T L_1 x^g(t_1) \right]. \tag{8.3.1}$$

Here, the matrix $L(t)$ can be split up according to

$$L(t) = \begin{bmatrix} L_{11}(s) & L_{12}(s) \\ L_{12}(s)^T & L_{22}(s) \end{bmatrix}, \tag{8.3.2}$$

where it is assumed that $L(t) = L(t)^T \geq 0$ and $L_{22}(t) = L_{22}(t)^T > 0$. The optimal control law is now given by

$$u(t) = F(t)\hat{x}(t), \tag{8.3.3}$$

where $\hat{x}(t)$ is given by the Kalman filter as defined in equation (8.2.4). The matrix $F(t)$ is recursively defined as

$$\begin{aligned} H_{11}(t) &= A^T(t)Q_c(t)A(t) + L_{11}(t), & (8.3.4)\\ H_{12}(t) &= A^T(t)Q_c(t)B(t) + L_{12}(t), & (8.3.5)\\ H_{22}(t) &= B^T(t)Q_c(t)B(t) + L_{22}(t), & (8.3.6)\\ Q_c(t+1) &= H_{11}(t) - H_{12}(t)H_{22}^{-1}(t)H_{12}(t)^T, & (8.3.7)\\ F(t) &= -H_{22}^{-1}(t)H_{12}^T(t). & (8.3.8) \end{aligned}$$

This solution can be proven using the value function. This function is given by

$$V(t,x) = x^T P(t)x + r(x). \tag{8.3.9}$$

Here, the function $r(x)$ is recursively defined such that $r(t_1) = \mathrm{tr}(L_1 Q(t_1))$ and

$$r(t) = r(t+1) + \mathrm{tr}(L_{11}Q(t) + Q(t+1)P(t+1)) + \mathrm{tr}(K(t)I_{22}(t)K(t)^T P(t+1)). \tag{8.3.10}$$

By the way, tr(.) is the trace function, being the sum of the diagonal elements of the matrix.

### 8.3.2 Infinite horizon cost functions

Let's examine a time-invariant system. Previously, there always was some final time $t_1$. What happens if $t_1 = \infty$? In this case, our cost function will become infinite, which is not convenient. So we need a different cost function. Options are the **discounted quadratic cost function** and the **infinite-horizon average quadratic cost function**, respectively defined as

$$\begin{aligned} J_{dc}(g) &= E_g \left[ \sum_{s=0}^{\infty} r^s \begin{bmatrix} x^g(s) \\ u^g(s) \end{bmatrix}^T L \begin{bmatrix} x^g(s) \\ u^g(s) \end{bmatrix} \right], & (8.3.11)\\ J_{av}(g) &= \lim_{t\to\infty} \frac{1}{t} E_g \left[ \begin{bmatrix} x^g(s) \\ u^g(s) \end{bmatrix}^T L \begin{bmatrix} x^g(s) \\ u^g(s) \end{bmatrix} \right]. & (8.3.12) \end{aligned}$$

Another cost function is the **minimum variance cost function**. It is used when one wants to minimize the variance of the state. This cost function is defined as

$$J_{mv}(g) = \lim_{t\to\infty} \frac{1}{t} E_g \left[ \sum_{s=0}^{t-1} x^g(s)^T L x^g(s) \right]. \tag{8.3.13}$$

### 8.3.3 Solving the infinite horizon problem

In literature, a certain procedure is often used to solve this problem. However, there is no proof yet that this procedure actually minimizes the cost functions. But we will discuss it anyway.

The idea is that we simply take the limit case of $t \to \infty$ of the previous problem. In this case, $Q$ satisfies the algebraic Riccati equation of Kalman filtering

$$Q = AQA^T + MVM^T - (AQC^T + MVN^T)(CQC^T + NVN^T)^{-1}(AQC^T + MVN^T)^T. \tag{8.3.14}$$

The matrix $P$ satisfies the algebraic Riccati equation of control

$$P = A^T PA + L_{11} - (A^T PB + L_{12})(B^T PB + L_{22})^{-1}(A^T PB + L_{12})^T. \tag{8.3.15}$$

Using $Q$ and $P$, we can find the matrices $K$ and $F$, according to

$$\begin{aligned} K &= (AQC^T + MVN^T)(CQC^T + NVN^T)^{-1}, & (8.3.16)\\ F &= (B^T PB + L_{22})(A^T PB + L_{12})^T. & (8.3.17) \end{aligned}$$

It can now be shown that we have

$$\operatorname{spec}(A - KC) \subset \mathbb{D}_o, \qquad \operatorname{spec}(A + BF) \subset \mathbb{D}_o \qquad \text{and} \qquad \operatorname{spec}\begin{pmatrix} A & BF \\ KC & A + BF - KC \end{pmatrix} \subset \mathbb{D}_o. \tag{8.3.18}$$

The control law which supposedly minimizes the infinite horizon cost functions is

$$u(t) = Fz(t), \qquad \text{with} \qquad z(t+1) = Az(t) + Bu(t) + K(y(t) - Cz(t) - Du(t)). \tag{8.3.19}$$