Control of Stochastic Systems
Lecture 2
Stochastic Processes

Jan H. van Schuppen

20 February 2025

Delft University of Technology
Delft, The Netherlands

Outline

Introduction

Concepts

Specific Stochastic Processes

Properties of Stochastic Processes

Conditional Independence

Markov Processes

Gaussian Processes

Finite-Valued Markov Processes

General Comment

Outline

Introduction

Concepts

Specific Stochastic Processes

Properties of Stochastic Processes

Conditional Independence

Markov Processes

Gaussian Processes

Finite-Valued Markov Processes

General Comment

Introduction

Time series of phenomena

Students may collect a time series of a phenomenon.
Examples of phenomena:

- Daily maximum temperature as recorded by a weather service.
- Energy use during a week
  of the house or of the room of a student.
- Daily price of rice or grain quoted at an agricultural market.
- Price of a bond or a stock quoted at a stock market.

Collection of a time series is not a required task of the course.

Introduction

Motivation of the use of stochastic processes for problems of control engineering

- Noise in communication channels during the 1920's.
- Noise in technical systems subject to process control, for example control of a paper machine, of a ship, etc.
- Fluctations in arrival rates of telephone networks and of communication networks.
- Noise in mechanical systems as cars, railway vehicles, and airplanes.
- Fluctuations in traffic flow on a motorway or on an urban road network.
- Fluctations in database entries.

Introduction

Learning goals of Lecture 02

- Properties of stochastic processes.
- Representations of:
  - a Gauss-Markov process and
  - a finite-valued Markov process.

Remark

A stochastic system is for a signal with fluctuations
a more realistic model and of lower complexity
than a deterministic system.

Outline

Introduction

Concepts

Specific Stochastic Processes

Properties of Stochastic Processes

Conditional Independence

Markov Processes

Gaussian Processes

Finite-Valued Markov Processes

General Comment

Notation

$$
\begin{aligned}
\mathbb{Z} &= \{\ldots,\ -1,\ 0,\ +1,\ldots\},\ \text{the integers},\\
\mathbb{Z}_+ &= \{1,\ 2,\ 3,\ \ldots\},\ \text{the positive integers},\\
\mathbb{N} &= \{0,\ 1,\ 2,\ 3,\ \ldots\},\ \text{the natural numbers},\\
\mathbb{Z}_n &= \{1,\ 2,\ 3,\ \ldots,\ n\},\ \text{the first } n \text{ integers},\ n \in \mathbb{Z}_+,\\
\mathbb{N}_n &= \{0,\ 1,\ 2,\ 3,\ \ldots,\ n\},\ \text{the first } n \text{ natural numbers},\\
\mathbb{R} &= (-\infty,\ +\infty),\ \text{the real numbers},\\
\mathbb{R}_+ &= [0,\ +\infty),\ \text{the positive real numbers},\\
\mathbb{R}^n &= \text{the vector space of } n\text{-tuples of the real numbers}.
\end{aligned}
$$

This slide is a recall of Lecture 1.

Def. Stochastic process

$$
\begin{aligned}
&(\Omega,\ F),\ (X,\ G),\ T \subseteq \mathbb{Z},\\
&x : \Omega \times T \to X,\ \text{is called a stochastic process},\\
&\text{if, } \forall\ t \in T,\ x(.,t) : \Omega \to X \text{ is a random variable}\\
&\text{(a measurable function)},\\
\Leftrightarrow\ &\forall\ t \in T,\ \forall\ A \in G,\ \{\omega \in \Omega |\ x(\omega,t) \in A\} \in F.
\end{aligned}
$$

Notation,

$$
\begin{aligned}
&x(t) = x_t = x_t(\omega) = x(\omega, t),\\
&\quad x = \{x(\omega,t) \in X,\ \forall\ t \in T,\ \forall\ \omega \in \Omega\};\\
&\qquad \forall\ \omega \in \Omega,\ x(\omega, .) : T \to X \text{ called a sample path of process;}\\
&\qquad \text{examples of discrete-time index sets are}\\
&\quad T_a = \{0,\ 1,\ 2,\ \ldots,\ t_e\},\ t_e \in \mathbb{Z}_+,\ \text{a finite horizon, or,}\\
&\quad T_b = \mathbb{N} = \{0,\ 1,\ 2,\ \ldots\},\ \text{a half-infinite forward horizon, or,}\\
&\quad T_c = \mathbb{Z} = \{\ldots,\ -1,\ 0,\ +1, \ldots\}\ \text{the horizon of all the integers.}
\end{aligned}
$$

Comments on definition of a stochastic process

- A time series of daily temperatures,
  may be modelled as a stochastic process.
- Observations of a phenomenon corresponds to
  a sample path of the stochastic process.
- System identification is the subject to go
  from the observations of a phenomenon
  to a model in the form of a stochastic system
  with its parameter values.

Def. Family of finite-dimensional probability distributions

Define the family of finite-dimensional probability distributions of a stochastic process as the collection,

$$x : \Omega \times T \to \mathbb{R}^n, \; T = \mathbb{N},$$
$$P_{fdpdf} = \left(\left\{ \begin{array}{l} pdf(.; \; x(t_1), \; x(t_2), \; \ldots, \; x(t_m)) \mid \\ \forall \; m \in \mathbb{Z}_+, \; \forall \; t_1, \; t_2, \; \ldots, t_m \in T, \; t_1 < t_2 < \ldots < t_m \end{array} \right\} \right);$$

$$\begin{aligned} & pdf((w_1, \; w_2, \; \ldots, w_m); \; (x(t_1), \; x(t_2), \; \ldots, \; x(t_m)) \\ = \; & P(\{x(t_1) \leq w_1, \; x(t_2) \leq w_2, \; \ldots, \; x(t_m) \leq w_m\}), \\ & \forall \; w_1, \; w_2, \; \ldots, \; w_m \in \mathbb{R}^n, \; \text{ is an element of } P_{fdpdf}. \end{aligned}$$

Consistency conditions of such a family, see book.

Remark

Family of finite-dimensional probability distributions
can in principle be approximated by observations of a phenomenon.
Statistics provides procedures for
the estimation of a pdf and of a probability density function.

Theorem. Existence stochastic process

Due to A.N. Kolmogorov (1950).
Consider a family P_{fdpdf}
of finite-dimensional probability distributions and
assume it satisfies the consistency conditions.

Then $\exists$ $(\Omega,\ F,\ P)$ and $\exists$ $x : \Omega \times T \rightarrow \mathbb{R}^n$ such that,

$$\begin{aligned} &\forall\ m \in \mathbb{Z}_+,\ \forall\ t_1,\ \ldots,\ t_m \in T,\ t_1 < t_2 < \ldots < t_m, \\ &\forall\ w_1,\ w_2, \ldots,\ w_m \in \mathbb{R}^n, \\ P&(\{\omega \in \Omega | x(\omega, t_1) \leq w_1,\ \ldots,\ x(\omega,\ t_m) \leq w_m, \}) \\ &= f((w_1,\ \ldots, w_m);\ (t_1,\ t_2, \ldots,\ t_m)) \in P_{fdpdf}. \end{aligned}$$

Kolmogorov proved theorem for $T = [0, 1]$.
Then also true for $T = \mathbb{R}_+$, and for $T = \mathbb{R}$.
Similarly true for $T = \mathbb{N}$.

Def. Equivalent processes

Consider T and $(X,\ G)$.
Two stochastic processes on these spaces
are called equivalent processes
if their families of finite-dimensional probability distribution functions
are identical.

Outline

Introduction

Concepts

Specific Stochastic Processes

Properties of Stochastic Processes

Conditional Independence

Markov Processes

Gaussian Processes

Finite-Valued Markov Processes

General Comment

Ways to define a stochastic process

1. Specify the family of
   all finite-dimensional probability distribution functions.
   Example. A Gaussian process, see below.
2. Specify that the stochastic process
   is a sequence of independent random variables
   of which one specifies
   the probability distribution function of the process at all times.
3. Specify the dependence relation of a stochastic process over time
   as a Markov process, and
   specify the transition function of a Markov process.

Below all three ways are used.

Def. Gaussian stochastic process

A stochastic process $x : \Omega \times T \to \mathbb{R}^n$ is called
a Gaussian process (or a Gaussian stochastic process)
if each member
of its family of finite-dimensional probability distribution functions
is a Gaussian probability distribution function.
In terms of notation,

$$\forall\, m \in \mathbb{Z}_+,\ \forall\, t_1,\ t_2,\ \ldots,\ t_m \in T,\ t_1 < t_2 < \ldots < t_m,$$
$$\{x(t_1),\ x(t_2),\ \ldots,\ x(t_m)\} \in G.$$

Examples. Gaussian processes used as approximate models

- Analog radio signals.
- Model of water fluctuations at sea.
- Model of vibrations of cars or airplanes.

Remark

The use of the Gaussian pdf is motivated by the central limit theorem: the scaled sum of a sequence of independent random variables converges to a random variable with a Gaussian probability distribution!

Def. Bernoulli process

A stochastic process is called a Bernoulli process if it satisfies,

$$
\begin{aligned}
& x : \Omega \times T \to \{0,\ 1\}, \\
& \{x(0),\ x(1),\ x(2),\ \ldots\}, \\
& \text{is a sequence of independent random variables,} \\
q(t) &= P(\{\omega \in \Omega |\ x(\omega, t) = 1\}), \\
1 - q(t) &= P(\{\omega \in \Omega |\ x(\omega, t) = 0\}), \quad q : T \to [0,\ 1].
\end{aligned}
$$

It is called an identically-distributed Bernoulli process
if, in addition, for all $t \in T$, $q(t) = q \in [0,\ 1]$.

Remark

A Bernoulli process is a model for
a stream of bits of a communication channel,
as used in information theory, which is part of electrical engineering.

Def. Discrete-time Poisson process

A stochastic process is called
a discrete-time Poisson process if,

$$x : \Omega \times T \to \mathbb{N} = \{0,\ 1,\ 2,\ \ldots,\}$$
$$\{x(t),\ \forall\ t \in T\} \text{ a sequence of independent rvs,}$$
$$\forall\ t \in T,\ x(.,t) : \Omega \to \mathbb{N} \text{ has a Poisson pdf with,}$$
$$P(\{\omega \in \Omega |\ x(\omega, t) = k\}) = \lambda(t)^k\ \exp(-\lambda(t))/k!,\ \forall\ k \in \mathbb{N};$$
$$\lambda : T \to (0, \infty) = \mathbb{R}_{s+},\ \text{called the rate of the Poisson process.}$$
$$\text{Recall } \sum_{k=0}^{\infty} \lambda(t)^k/k! = \exp(\lambda(t)) \ \Rightarrow \sum_{k=0}^{\infty} P(\{x(t) = k\}) = 1,$$

A discrete-time Poisson process is a model for
the arrivals of call requests at a telephone switch in short intervals,
say 6 seconds.

Remark

Named after S.D. Poisson (1781-1840, born in France).

Outline

Introduction

Concepts

Specific Stochastic Processes

Properties of Stochastic Processes

Conditional Independence

Markov Processes

Gaussian Processes

Finite-Valued Markov Processes

General Comment

Def. Integrability

Consider a stochastic process $x : \Omega \times T \to \mathbb{R}^{n_x}$ for $n_x \in \mathbb{Z}_+$.
The process is called integrable if,

$$\forall\, t \in T,\ \forall\, i \in \mathbb{Z}_{n_x},\ E|x_i(t)| < \infty;$$
$$m_x(t) = E[x(t)],\ m_x : T \to \mathbb{R}^{n_x}.$$

Call m_x the mean value function of the process.
The process is called square integrable if,

$$\forall\, t \in T,\ \forall\, i \in \mathbb{Z}_{n_x},\ E[x_i(t)^2] < \infty;$$

Cauchy-Schwartz inequality $\forall\, s,\ t \in T,\ \forall\, i,\ j \in \mathbb{Z}_{n_x}$,

$$E|x_i(t)x_j(s)| \leq (E[x_i(t)^2])^{1/2}(E[x_j(s)^2])^{1/2} < \infty;$$
$$C_x(t,s) = E[x(t)x(s)^T],\ C_x : T \times T \to \mathbb{R}^{n_x \times n_x},$$
$$W_x(t,s) = E[(x(t) - m_x(t))(x(s) - m_x(s))^T],$$
$$W_x : T \times T \to \mathbb{R}^{n_x \times n_x}.$$

C_x called the correlation function, W_x called the covariance function.

Def. Positive-definite function
A function $W : T \times T \to \mathbb{R}^{n\times n}$ is called
a positive-definite function if,

$$\begin{aligned}
&\forall\, m \in \mathbb{Z}_+,\\
&\forall\, t_1,\ t_2,\ \ldots,\ t_m \in T = \mathbb{N} = \{0,\ 1,\ \ldots\},\ t_1 < t_2 < \ldots < t_m,\\
&\forall\, c_1,\ c_2,\ \ldots,\ c_m \in \mathbb{R}^n,\\
0 \le &\sum_{i=1}^{m} \sum_{j=1}^{m} c_i^T\ W(t_i,\ t_j)\ c_j.
\end{aligned}$$

Remark
Condition of a positive-definite function
is an extension to an infinite sequence on $T = \mathbb{N}$
of the condition of a positive-definite symmetric matrix.

Proposition. Characterization of a covariance function

The function $W : T \times T \to \mathbb{R}^{n \times n}$ on $T = \mathbb{N} = \{0,\ 1,\ \ldots\}$
is a covariance function of a stochastic process
if and only if

1. $W(s, t) = W(t, s)^T$ for all $s,\ t \in T$
   called closed with respect to transposition; and
2. W is a positive-definite function.

Proof

($\Rightarrow$) Consider a square-integrable stochastic process with zero mean-value function and covariance function W. Then,

$$\sum_{i=1}^{m} \sum_{j=1}^{m} c_i^T \, W(t_i,\ t_j) \, c_j = \sum_{i=1}^{m} \sum_{j=1}^{m} E[c_i^T x(t_i) x(t_j)^T c_j]$$

$$= E[(\sum_{i=1}^{m} c_i^T \, x(t_i))^2] \geq 0.$$

($\Leftarrow$) For any $m \in \mathbb{Z}_+$, and $\forall\ t_1,\ t_2,\ \ldots,\ t_m \in T$,
the condition that W is a positive-definite function
implies that $Q_m \in \mathbb{R}^{nm \times nm}$ is a positive-definite matrix,
where $Q_{m,i,j} = W(t_i,\ t_j)$.
Hence $(0,\ Q_m)$ are the parameters of
a multivariate Gaussian probability distribution function.
By Kolmogorov's theorem there exists a Gaussian process
which has W as its covariance function.

Def. Stationarity

A stochastic process is called stationary
if any finite-dimensional probability distribution function
remains the same after any time shift.

$$
\begin{aligned}
& x : \Omega \times T \to \mathbb{R}^{n_x},\ T \subseteq \mathbb{Z}, \\
& \text{if } \forall\ m \in \mathbb{Z}_+,\ \forall\ t_1,\ t_2,\ \ldots,\ t_m \in T,\ \text{such that } t_1 < t_2 < \ldots < t_m, \\
& \forall\ s \in \mathbb{Z},\ \text{such that } t_1 + s,\ t_2 + s,\ \ldots t_m + s \in T, \\
& pdf(x(t_1),\ x(t_2),\ \ldots, x(t_m)) \\
= {} & pdf(x(t_1 + s),\ x(t_2 + s),\ \ldots,\ x(t_m + s)).
\end{aligned}
$$

Remark. For which engineering phenomena is a stationary process a realistic model?

- A stationary process is a modeling approximation.
- A modeling approach is often needed
  before one obtains a stationary process.
  Remove a trend, or remove the cycle of the seasons.
  A model may be needed, examples: a linear or a multiplicative relation.

Def. Time-Reversibility

A stochastic process is called a time-reversible process if any finite-dimensional probability distribution function remains the same after any time reversion.

$$
\begin{aligned}
& x : \Omega \times T \to \mathbb{R}^{n_x}, \\
& \text{if } \forall\, m \in \mathbb{Z}_+,\ \forall\, t_1,\ t_2,\ \ldots,\ t_m \in T,\ \text{such that } t_1 < t_2 < \ldots < t_m, \\
& \forall\, s \in \mathbb{Z}, \\
& \text{such that } s - t_1,\ s - t_2,\ \ldots,\ s - t_m \in T;\ \text{hence,} \\
& s - t_m < s - t_{m-1} < \ldots < s - t_2 < s - t_1; \\
& pdf(x(t_1),\ x(t_2),\ \ldots, x(t_m)) \\
= {} & pdf(x(s - t_1),\ x(s - t_2),\ \ldots,\ x(s - t_m)).
\end{aligned}
$$

Proposition. Time-Reversibility implies stationarity

A time-reversible process is a stationary process.

Remark. When is a time-reversible process a realistic model?

It is used as model in processes of communication networks.
It is used as model for particular phenomena of physics.

Outline

Introduction

Concepts

Specific Stochastic Processes

Properties of Stochastic Processes

Conditional Independence

Markov Processes

Gaussian Processes

Finite-Valued Markov Processes

General Comment

Conditional Independence

Definition. Independence (Recall from Lecture 1)

Consider

$$
\begin{aligned}
& (\Omega, F, P), \quad F_1,\ F_2 \subseteq F \text{ sub-}\sigma\text{-algebras.} \\
& \text{Call } F_1,\ F_2 \text{ independent if} \\
E[x_1\ x_2] = {} & E[x_1]\ E[x_2], \\
& \forall x_1 \in L(\Omega, F_1, \mathbb{R}_+),\ \forall x_2 \in L(\Omega, F_2, \mathbb{R}_+). \\
& \text{Notation } \ (F_1,\ F_2) \in I, \\
L(\Omega, F_1, \mathbb{R}_+) = {} & \left\{ \begin{array}{l} x_1 : \Omega \rightarrow \mathbb{R}_+ | \\ x_1 \text{ is a random variable, } F_1 \text{ measurable} \end{array} \right\}.
\end{aligned}
$$

Conditional Independence

Proposition

Consider (Ω, F, P), $F_1,\ F_2 \subseteq F$ sub-σ-algebras. Equivalence of:

(a) $(F_1,\ F_2)$ are independent sub-σ-algebras, see Lecture 1.

(b) $(F_1,\ F_2)$ are independent in terms of expectations, see below.

Proof

$$(b)\ \Rightarrow (a) \quad A_1 \in F_1,\ A_2 \in F_2, \text{ imply that,}$$

$$P(A_1 \cap A_2) = E[I_{A_1 \cap A_2}(\omega)] = E[I_{A_1}(\omega)\ I_{A_2}(\omega)] = E[I_{A_1}]\ E[I_{A_2}] = P(A_1)\ P(A_2);$$

$$(a)\ \Rightarrow (b) \quad E[I_{A_1 \cap A_2}] = P(A_1 \cap A_2) = P(A_1) \times P(A_2) = E[I_{A_1}]\ E[I_{A_2}],$$

$$E[x_1 \times x_2] = E\left[\left(\sum_{i=1}^{n_{i_{x_1}}} a_i\ I_{A_i}\right)\left(\sum_{j=1}^{n_{i_{x_2}}} b_j\ I_{B_j}\right)\right] = \sum_{i=1}^{n_{i_{x_1}}} \sum_{j=1}^{n_{i_{x_2}}} a_i\ b_j\ E[I_{A_i}\ I_{B_j}]$$

$$= \sum_{i=1}^{n_{i_{x_1}}} \sum_{j=1}^{n_{i_{x_2}}} a_i\ b_j\ E[I_{A_i}]\ E[I_{B_j}] = E[x_1] \times E[x_2].$$

Finally use the monotone class theorem for random variables.

Conditional Independence

Definition. Conditional independence relation

Define the relation,

$$
\begin{aligned}
& (\Omega,\ F,\ P),\ \ F_1,\ F_2,\ G \subseteq F,\ \text{sub-}\sigma\text{-algebras;} \\
E[x_1\ x_2|G] = {} & E[x_1|\ G]\ E[x_2|\ G], \\
& \forall\ x_1 \in L(\Omega, F_1, \mathbb{R}_+),\ \forall\ x_2 \in L(\Omega, F_2, \mathbb{R}_+). \\
& \text{Notation }\ (F_1,\ F_2|\ G) \in CI.
\end{aligned}
$$

Call $F_1,\ F_2$ conditionally independent given G if $(F_1,\ F_2|\ G) \in CI$.
One also says that G makes F_1 and F_2 conditionally independent.
Call CI the conditional independence relation.

Remarks

(1) Conditional independence is a generalization of independence.
(2) Conditional independence used:
in system theory of stochastic systems,
in Markov processes, and in statistics.

Conditional Independence

Theorem

Equivalence:

(a) $(F_1, F_2|G) \in CI$.

(b) $(F_2, F_1|G) \in CI$.

(c)

$$E[x_1|F_2 \vee G] = E[x_1|G], \ \forall \, x_1 \in L(\Omega, F_1, \mathbb{R}_+).$$

(d) $(F_1 \vee G, F_2 \vee G|G) \in CI$.

Proof in lecture notes and in book.

Conditional Independence

Proposition

$$(F_1, F_2 \vee G) \text{ independent } \Rightarrow (F_1, F_2|G) \in CI.$$

Proof

$$\forall\, x_1 \in L(\Omega, F_1, \mathbb{R}_+),$$

$$E[x_1|F_2 \vee G] = E[x_1],$$

because $(F_1,\ F_2 \vee G) \in I$ and by conditional expectation;

$$(F_1,\ F_2 \vee G) \in I,\quad G \subseteq F_2 \vee G \;\Rightarrow\; (F_1,\ G) \in I \;\Rightarrow$$

$$E[x_1|G] = E[x_1] \text{ by conditional expectation } \Rightarrow$$

$$E[x_1|F_2 \vee G] = E[x_1|G].$$

Conditional Independence

Theorem. Conditional independence of Gaussian rvs

Consider a triple of Gaussian random variables

$$(y_1,\ y_2,\ x) \in G(0,\ Q_{(y_1,y_2,x)}),$$
$$y_1 : \Omega \to \mathbb{R}^{n_{y_1}},\ y_2 : \Omega \to \mathbb{R}^{n_{y_2}},\ x : \Omega \to \mathbb{R}^{n_x},\ 0 \prec Q_x.$$

Then

$$(F^{y_1},\ F^{y_2} |\ F^x) \in CI \ \Leftrightarrow\ Q_{y_1,y_2} = Q_{y_1,x} Q_x^{-1} Q_{y_2,x}^T.$$

Proof

Conditional independence is equivalent with

$$\Leftrightarrow E[\exp(iw_1^T y_1 + iw_2^T y_2) |\ F^x]$$
$$= E[\exp(iw_1^T y_1) |\ F^x]\ E[\exp(iw_2^T y_2) |\ F^x],\ \forall\ (w_1,\ w_2) \in \mathbb{R}^{n_{y_1}} \times \mathbb{R}^{n_{y_2}}.$$

A calculation concludes the proof.

Outline

Introduction

Concepts

Specific Stochastic Processes

Properties of Stochastic Processes

Conditional Independence

Markov Processes

Gaussian Processes

Finite-Valued Markov Processes

General Comment

Def. Markov process

A stochastic process is called a Markov process
if, for all times, the future and the past of the process
are conditionally independent
when conditioned on the present of the process.
Equivalently,

$$
\begin{aligned}
&\forall\, t \in T,\ (F_t^{x+},\ F_{t-1}^{x-} |\ F^{x(t)}) \in CI;\ \text{where,} \\
&x : \Omega \times T \to X,\ (\Omega,\ F,\ P),\ (X,\ G), \\
F_t^{x+} &= \sigma(\{x(s),\ \forall\, s \geq t\}), \\
F_{t-1}^{x-} &= \sigma(\{x(s),\ \forall\, s \leq t-1\}).
\end{aligned}
$$

Remarks

A.A. Markov (1906, father) published definition.
There is a father and a son Markov, both were mathematicians.

Proposition. Equivalent conditions of a Markov process

Consider a stochastic process $x : \Omega \times T \to \mathbb{R}^{n_x}$.
The following statements are equivalent:

(a) x is a Markov process,

(b) $\forall\, t \in T,\ (F_t^{x+} \vee F^{x(t)},\ F_{t-1}^{x-} \vee F^{x(t)} |\ F^{x(t)}) \in CI,$

(c) $\forall\, s,\ t \in T,\ s < t,\ \forall\, w \in \mathbb{R}^{n_x},$

$$E[\exp(iw^T x(t))|\ F_s^{x-}] = E[\exp(iw^T x(t))|\ F^{x(s)}],$$

(d) $E[f(x(t))|\ F_s^{x-}] = E[f(x(t))|\ F^{x(s)}],$

$\forall\, s,\ t \in T,\ s < t,$

$\forall\, f : \mathbb{R}^{n_x} \to \mathbb{R}$ such that $E|f(x(t))| < \infty$.

Comments on definition Markov process

- Proof of proposition in book.
- Interpretation of (c) of proposition:
  Future conditioned on the past at time $s \in T$
  equals future condition on the present at s.
- Interpretation of a Markov process
  in terms of measurable map
  from a state to a conditional measure on a future state:

  $$x(s) \mapsto \text{cpdf}(x(t)|\ F_s^{x-}),\ \ \forall\ s,\ t \in T,\ s < t.$$

  cpdf denotes a conditional probability distribution function
  or a conditional measure.
  From this follows Proposition part (c) according to Exercise 4 of hset01.

Proposition. Recursive represenation of a Markov process

Consider an integrable Markov process.
There exists a recursive representation of the process of the form,

$$\begin{aligned} x &: \Omega \times T \rightarrow \mathbb{R}^{n_x}, \\ x(t+1) &= f(t,\ x(t)) + \Delta m(t),\ x(0) = x_0, \\ f(t,\ x(t)) &= E[x(t+1)|\ F^{x(t)}], \\ m(t) &= \sum_{s=1}^{t} \Delta m(s),\ m : \Omega \times T \rightarrow \mathbb{R}^{n_x}, \\ & \{m(t),\ F_t^x,\ t \in T\} \text{ is a martingale.} \end{aligned}$$

Proof

$$\begin{aligned} & E[\Delta m(t)|F_t^x] = E[x(t+1) - f(t,\ x(t))|F_t^x] = E[x(t+1)|\ F_t^x] - f(t,\ x(t)) \\ & = E[x(t+1)|\ F^{x(t)}] - f(t,\ x(t)), \text{ because } x \text{ is a Markov process,} \\ & = 0. \end{aligned}$$

Outline

Introduction

Concepts

Specific Stochastic Processes

Properties of Stochastic Processes

Conditional Independence

Markov Processes

Gaussian Processes

Finite-Valued Markov Processes

General Comment

Def. Gaussian process and notation

A stochastic process is called a Gaussian process
if every member of
its family of finite-dimensional probability distibution functions
is a Gaussian pdf. (Recall of definition.) Notation.

$$
\begin{aligned}
& x : \Omega \times T \to \mathbb{R}^{n_x}, \text{ a Gaussian process;} \\
& \forall\, t \in T,\ \forall\, i \in \mathbb{Z}_{n_x},\ E|x_i(t)|^2 < \infty, \\
& \text{hence a Gaussian process is square-integrable;} \\
m_x(t) &= E[x(t)],\ m_x : T \to \mathbb{R}^{n_x}, \\
& \text{mean value function}, \\
W_x(t,s) &= E[(x(t) - m_x(t))(x(s) - m_x(s))^T],\ W_x : T \times T \to \mathbb{R}^{n_x \times n_x}, \\
& \text{covariance function}, \\
Q_x(t) &= E[(x(t) - m_x(t))(x(t) - m_x(t))^T],\ Q_x : T \to \mathbb{R}^{n_x \times n_x}_{pds}, \\
& \text{variance function. Then} \\
Q_x(t) &= W_x(t,t),\ \forall\, t \in T.
\end{aligned}
$$

Def. Jointly Gaussian processes

Two stochastic processes are called jointly Gaussian
if each member of the family
of their joint finite-dimensional probability distribution functions
is Gaussian.

Proposition

(a) Each member of a tuple of jointly Gaussian process
is a Gaussian process.

(b) If two stochastic processes are independent and
if each of these processes is Gaussian
then the tuple of these processes is a jointly Gaussian process.

Proposition. A stationary Gaussian process

Consider a Gaussian process $x : \Omega \times T \to \mathbb{R}^{n_x}$ for $n_x \in \mathbb{Z}_+$.
This processs is stationary if and only if,

$$\begin{aligned} (1) \quad m_x(t) &= m_x(s), \ \forall\, s,\ t \in T; \\ (2) \quad W_x(t,s) &= W_x(t+r,\ s+r), \\ &\qquad \forall\, s,\ t \in T,\ \forall\, r \in \mathbb{Z} \text{ such that } s+r,\ t+r \in T. \end{aligned}$$

Notation. Stationary Gaussian process

$$m_x = m_x(0) = m_x(t),\ m_x \in \mathbb{R}^{n_x} \text{ called the mean value,}$$
$$W_x(t) = W_x(t,0) = W_x(t+r,\ r), \forall\, r \in \mathbb{Z} \text{ such that } r,\ t+r \in T,$$
$$W_x : T \to \mathbb{R}^{n_x \times n_x} \text{ called the covariance function.}$$

Note the abuse of notation for W_x.

Def. Gaussian white noise process

(a) A Gaussian white noise process $v : \Omega \times T \to \mathbb{R}^{n_v}$
is defined such that it is a Gaussian process and
$\{v(t),\ \forall\ t \in T\}$ is a sequence of independent random variables.
Then, for all $t \in T$, $v(t) \in G(m_v(t),\ Q_v(t))$.

(b) A stationary Gaussian white noise process
is a Gaussian white noise process which is also stationary.
Then there exist $m_v \in \mathbb{R}^{n_v}$ and $Q_v \in \mathbb{R}^{n_v \times n_v}_{pds}$ such that,
for all $t \in T$, $v(t) \in G(m_v,\ Q_v)$.

(c) It is called a standard stationary Gaussian white noise process
if it is a stationary Gaussian white noise process such that,
for all $t \in T$, $v(t) \in G(0, I_{n_v})$.

Proposition. Representation of a Gauss-Markov process

Consider a Gaussian process with the notation,

$$\begin{aligned} & x : \Omega \times T \to \mathbb{R}^{n_x},\ T = \mathbb{N}, \\ x(t) \in G(0, Q_x(t)), \quad & \text{assume } \forall\, t \in T,\ 0 \prec Q_x(t). \end{aligned}$$

Equivalence of:

(a) x is a Gauss-Markov process;

(b) x has the representation,

$$\begin{aligned} x(t+1) = A(t)\, x(t) + M(t)\, v(t),\ & x(0) = x_0, \\ & x_0 : \Omega \to \mathbb{R}^{n_x},\ x_0 \in G(0, Q_{x_0}),\ 0 \prec Q_{x_0}, \\ & v : \Omega \to \mathbb{R}^{n_v},\ \text{standard Gaussian white noise}, \\ & F^{x_0},\ F^{v}_{\infty},\ \text{are independent}, \\ & A : T \to \mathbb{R}^{n_x \times n_x},\ M : T \to \mathbb{R}^{n_x \times n_v}. \end{aligned}$$

Note the representation of a linear system
driven by standard Gaussian white noise!

Proof. Representation of a Gauss-Markov process

(a) $\Rightarrow$ (b) Fix $t \in T$. Gaussian process implies that $(x(t+1),\ x(t)) \in G$.

$$
\begin{aligned}
E[x(t+1)|\ F_t^x] &= E[x(t+1)|\ F^{x(t)}] = A(t)x(t),\\
A(t) &= E[x(t+1)x(t)^T]Q_x(t)^{-1},\ \text{Theorem 2.8.3(a);}\\
w(t) &= x(t+1) - A(t)x(t),\ (w(t),\ x(t+1),\ x(t)) \in G,\\
F_t &= F^{x_0} \vee F_t^w = F^{x_0} \vee \sigma(\{w(s),\ \forall\, s \leq t\});\\
&\quad E[\exp(iu^T w(t))|\ F_{t-1}] = E[E[\exp(iu^T w(t))|\ F_t^x]|\ F_{t-1}]\\
&= E[E[\exp(iu^T w(t))|\ F^{x(t)}]|\ F_{t-1}] = \exp(-u^T Q_w(t)u/2),\\
&\quad \text{by } E[w(t)|F^{x(t)}] = E[x(t+1)|\ F^{x(t)}] - A(t)x(t) = 0,\\
&= E[\exp(iu^T w(t))]\ \Rightarrow\ F^{w(t)},\ F_{t-1} \text{ independent,}\\
&\quad \Rightarrow\ w \text{ Gaussian white noise and } (F^{w(t)},\ F^{x_0}) \text{ independent;}\\
w(t) &= M(t)v(t),\ v(t) \in G(0, I_{n_v}),\\
&\quad v \text{ standard Gaussian white noise by Proposition 2.7.5;}\\
x(t+1) &= A(t)x(t) + M(t)v(t).
\end{aligned}
$$

Proposition. Representation of a stationary Gauss-Markov process

Consider a stationary Gaussian process with the notation,

$$x : \Omega \times T \to \mathbb{R}^{n_x},\ T = \mathbb{N},$$
$$x(t) \in G(0, Q_x);\ \text{assume } 0 \prec Q_x.$$

Equivalence of:

(a) x is a stationary Gauss-Markov process;

(b) x has the representation,

$$\begin{aligned} x(t+1) = A\,x(t) + M\,v(t),\ & x(0) = x_0, \\ & x_0 : \Omega \to \mathbb{R}^{n_x},\ x_0 \in G(0, Q_{x_0}),\ 0 \prec Q_{x_0}, \\ & v : \Omega \to \mathbb{R}^{n_v},\ \text{standard Gaussian white noise}, \\ & F^{x_0},\ F^v_\infty,\ \text{are independent}, \\ & A \in \mathbb{R}^{n_x \times n_x},\ M \in \mathbb{R}^{n_x \times n_v}. \end{aligned}$$

Proposition. When is a Gaussian process a Markov process?

Consider a Gaussian process with the notation,

$$x : \Omega \times T \rightarrow \mathbb{R}^{n_x},\ m_x(t), W_x(t,s);$$
$$\text{assume that } \forall\, t \in T,\ 0 \prec Q_x(t) = W_x(t,t).$$

This Gaussian process is a Markov process
if and only if
the covariance function satisfies,

$$W_x(t,s) = W_x(t,r) W_x(r,r)^{-1} W_x(r,s),$$
$$\forall\, s,\ r,\ t \in T,\ s < r < t.$$

Remark

Proof related to
characterization of conditional independence of Gaussian random variables.

Outline

Introduction

Concepts

Specific Stochastic Processes

Properties of Stochastic Processes

Conditional Independence

Markov Processes

Gaussian Processes

Finite-Valued Markov Processes

General Comment

Finite-Valued Processes

Def. Indicator representation of a finite-valued process

Consider a finite-valued stochastic process,

$$x : \Omega \times T \to \mathbb{Z}_{n_{i_x}} = \{1,\ 2,\ \ldots,\ n_{i_x}\} \subset \mathbb{Z},\ n_{i_x} \in \mathbb{Z}_+.$$

Define the indicator process of the finite valued process x according to,

$$\begin{aligned}
i_x(\omega,\ t) &= \begin{cases} +1, & \text{if } x(\omega, t) = j, \\ 0, & \text{else,} \end{cases} \quad \forall\, j \in \mathbb{Z}_{n_{i_x}}; \\
& i_x : \Omega \times T \to \mathbb{R}^{n_{i_x}}; \\
& \text{then,} \\
& i_x(t) \in \{e_1,\ e_2,\ \ldots,\ e_{n_{i_x}}\} \subset \mathbb{R}^{n_{i_x}}, \ \ \text{the set of unit vectors,} \\
x(t) &= C_x\ i_x(t), \\
C_x &= \begin{bmatrix} 1 & 2 & \ldots & n_{i_x} - 1 & n_{i_x} \end{bmatrix} \in \mathbb{R}^{n_x \times n_{i_x}}.
\end{aligned}$$

Finite-Valued Markov Processes

Proposition. Representation finite-valued Markov process

Consider a stationary finite-valued Markov process x
and its indicator process $i_x : \Omega \times T \to \mathbb{R}^{n_{i_x}}$.
Then there exists a system representation of the form,

$$
\begin{aligned}
i_x(t+1) &= A\, i_x(t) + \Delta m(t),\ i_x(0) = i_{x,0}, \\
x(t) &= C_x\ i_x(t), \\
&\quad \text{with } A \in \mathbb{R}_{st,+}^{n_{i_x} \times n_{i_x}} \text{ a stochastic matrix } (1_{n_{i_x}}^T A = 1_{n_{i_x}}^T), \\
&\quad \text{thus column sums of } A \text{ equal to one}, \\
A\, i_x(t) &= E[i_x(t+1)|\ F^{x(t)}] = E[i_x(t+1)|\ F^{i_x(t)}], \\
0 &= E[\Delta m(t)|\ F_t^x],\ \forall\ t \in T, \\
&\quad \Delta m : \Omega \times T \to \mathbb{R}^{n_{i_x}}.
\end{aligned}
$$

$\Delta m(t)$ is called a martingale increment at time $t \in T$.

Finite-Valued Markov Processes

Proof

$$\begin{aligned}
E[i_x(t+1)|\ F_t^x] &= E[i_x(t+1)|\ F^{x(t)}],\ \text{because } x \text{ is a Markov process,} \\
&= E[i_x(t+1)|\ F^{i_x(t)}] = A\, i_x(t),\ \text{by Thm. 2.8.4,} \\
&\quad \text{conditional expectation for finite-valued rvs and} \\
&\quad \text{because } x \text{ is a stationary process;} \\
\Delta m(t) &= i_x(t+1) - A\, i_x(t), \\
&\quad \text{then,} \\
E[\Delta m(t)|\ F_t^x] &= E[i_x(t+1)|\ F_t^x] - A\, i_x(t) \\
&= E[i_x(t+1)|\ F_t^{i_x}] - A\, i_x(t) = 0.
\end{aligned}$$

Finite-Valued Markov Processes

Example. Binary valued stochastic process

Define the binary-valued stationary Markov process according to the recursive representation,

$$\begin{aligned}
\mathbb{N}_1 &= \{0,\ 1\},\ n_x = 2,\\
& x : \Omega \times T \to \mathbb{N}_1,\\
& i_x : \Omega \times T \to \mathbb{R}^{n_{i_x}},\ \text{the indicator process of } x,\\
i_x(t) &= \begin{bmatrix} I_{\{x(t)=0\}} \\ I_{\{x(t)=1\}} \end{bmatrix},\\
A &= \begin{bmatrix} q_1 & 1-q_2 \\ 1-q_1 & q_2 \end{bmatrix} \in \mathbb{R}^{n_x \times n_x}_{st,+},\ q_1,\ q_2 \in [0,\ 1] \subset \mathbb{R},\\
i_x(t+1) &= A\, i_x(t) + \Delta m(t),\ i_x(0) = i_{x,0} = \begin{bmatrix} 1 \\ 0 \end{bmatrix},\\
x(t) &= C_x i_x(t) = i_{x,2}(t),\ C_x = \begin{bmatrix} 0 & 1 \end{bmatrix}.
\end{aligned}$$

Outline

Introduction

Concepts

Specific Stochastic Processes

Properties of Stochastic Processes

Conditional Independence

Markov Processes

Gaussian Processes

Finite-Valued Markov Processes

General Comment

General comment

Subsets of stochastic processes

- Stochastic processes consisting of
  a sequence of independent random variables.
  Example Gaussian white noise.
  Useful for generating a Markov process by a recursion.
- Martingales. See book, Section 20.2.
  Example a progressing sum of Gaussian white noise.
  Useful for convergence analysis.
- Markov processes.
  Example Gauss-Markov process.
  Example a stationary finite-valued Markov process.
  Useful as models of dynamic phenomena.