Math 562 Fall 2020

Renming Song

University of Illinois at Urbana-Champaign

August 26, 2020

Outline

Outline

The slides of the first lecture is available in the course webpage in my homepage.

I am trying to set up a Moodle page, where you can submit your homework. Recorded videos of the lectures will also be available there.

The slides of the first lecture is available in the course webpage in my homepage.

I am trying to set up a Moodle page, where you can submit your homework. Recorded videos of the lectures will also be available there.

Outline

Definition 1.4

A (centered) Gaussian space is a closed linear subspace of $L^2(\Omega, \mathcal{F}, \mathbb{P})$ which contains only centered Gaussian variables.

Definition 1.5

Let $(E, \mathcal{E})$ be a measurable space, and let T be an arbitrary index set. An E-valued stochastic process (indexed by T) is a collection $(X_t)_{t\in T}$ of random variables with values in E. If the measurable space $(E, \mathcal{E})$ is not specified, we will implicitly assume that $E = \mathbb{R}$ and $\mathcal{E} = \mathcal{B}(\mathbb{R})$.

Most of the time, the index set T will be $\mathbb{R}_+$ or an interval of the real line.

Definition 1.4

A (centered) Gaussian space is a closed linear subspace of $L^2(\Omega, \mathcal{F}, \mathbb{P})$ which contains only centered Gaussian variables.

Definition 1.5

Let $(E, \mathcal{E})$ be a measurable space, and let T be an arbitrary index set. An E-valued stochastic process (indexed by T) is a collection $(X_t)_{t \in T}$ of random variables with values in E. If the measurable space $(E, \mathcal{E})$ is not specified, we will implicitly assume that $E = \mathbb{R}$ and $\mathcal{E} = \mathcal{B}(\mathbb{R})$.

Most of the time, the index set T will be $\mathbb{R}_+$ or an interval of the real line.

Definition 1.4

A (centered) Gaussian space is a closed linear subspace of $L^2(\Omega, \mathcal{F}, \mathbb{P})$ which contains only centered Gaussian variables.

Definition 1.5

Let $(E, \mathcal{E})$ be a measurable space, and let T be an arbitrary index set. An E-valued stochastic process (indexed by T) is a collection $(X_t)_{t \in T}$ of random variables with values in E. If the measurable space $(E, \mathcal{E})$ is not specified, we will implicitly assume that $E = \mathbb{R}$ and $\mathcal{E} = \mathcal{B}(\mathbb{R})$.

Most of the time, the index set T will be $\mathbb{R}_+$ or an interval of the real line.

Definition 1.6

A (real-valued) stochastic process $(X_t)_{t \in T}$ is called a (centered) Gaussian process if any finite linear combination of the variables $X_t, t \in T$, is centered Gaussian.

Proposition 1.7

If $(X_t)_{t \in T}$ is a Gaussian process, the closed linear subspace of L^2 spanned by the variables $X_t, t \in T$, is a Gaussian space, which is called the Gaussian space generated by the process $(X_t)_{t \in T}$.

Proof of Proposition 1.7

It suffices to note that an L^2-limit of centered Gaussian variables is still centered Gaussian, by Proposition 1.1.

Definition 1.6

A (real-valued) stochastic process $(X_t)_{t\in T}$ is called a (centered) Gaussian process if any finite linear combination of the variables $X_t, t \in T$, is centered Gaussian.

Proposition 1.7

If $(X_t)_{t\in T}$ is a Gaussian process, the closed linear subspace of L^2 spanned by the variables $X_t, t \in T$, is a Gaussian space, which is called the Gaussian space generated by the process $(X_t)_{t\in T}$.

Proof of Proposition 1.7

It suffices to note that an L^2-limit of centered Gaussian variables is still centered Gaussian, by Proposition 1.1.

Definition 1.6

A (real-valued) stochastic process $(X_t)_{t \in T}$ is called a (centered) Gaussian process if any finite linear combination of the variables $X_t, t \in T$, is centered Gaussian.

Proposition 1.7

If $(X_t)_{t \in T}$ is a Gaussian process, the closed linear subspace of L^2 spanned by the variables $X_t, t \in T$, is a Gaussian space, which is called the Gaussian space generated by the process $(X_t)_{t \in T}$.

Proof of Proposition 1.7

It suffices to note that an L^2-limit of centered Gaussian variables is still centered Gaussian, by Proposition 1.1.

Definition 1.8

Let H be a collection of random variables defined on $(\Omega, \mathcal{F}, \mathbb{P})$. The σ-field generated by H, $\sigma(H)$, is the smallest σ-field on Ω such that all variables $\xi \in H$ are measurable with respect to this σ-field. If $\mathcal{C}$ is a collection of subsets of Ω, we write $\sigma(\mathcal{C})$ for the smallest σ-field containing $\mathcal{C}$.

The next theorem shows that, in some sense, independence is equivalent to orthogonality in a Gaussian space. This is a very particular property of the Gaussian distribution.

Definition 1.8

Let H be a collection of random variables defined on $(\Omega, \mathcal{F}, \mathbb{P})$. The σ-field generated by H, $\sigma(H)$, is the smallest σ-field on Ω such that all variables $\xi \in H$ are measurable with respect to this σ-field. If $\mathcal{C}$ is a collection of subsets of Ω, we write $\sigma(\mathcal{C})$ for the smallest σ-field containing $\mathcal{C}$.

The next theorem shows that, in some sense, independence is equivalent to orthogonality in a Gaussian space. This is a very particular property of the Gaussian distribution.

Theorem 1.9

Let H be a centered Gaussian space and let $(H_i)_{i \in I}$ be a collection of linear subspaces of H. Then the subspaces $H_i, i \in I$, are (pairwise) orthogonal in L^2 if and only the σ-fields $\sigma(H_i), i \in I$, are independent.

Proof of Theorem 1.9

Suppose that the σ-fields $\sigma(H_i), i \in I$, are independent. Then, if $i \neq j$, $X \in H_i$ and $Y \in H_j$,

$$\mathbb{E}[XY] = \mathbb{E}[X]\mathbb{E}[Y] = 0,$$

which implies that H_i and H_j are orthogonal.

Theorem 1.9

Let H be a centered Gaussian space and let $(H_i)_{i \in I}$ be a collection of linear subspaces of H. Then the subspaces $H_i, i \in I$, are (pairwise) orthogonal in L^2 if and only the σ-fields $\sigma(H_i), i \in I$, are independent.

Proof of Theorem 1.9

Suppose that the σ-fields $\sigma(H_i), i \in I$, are independent. Then, if $i \neq j$, $X \in H_i$ and $Y \in H_j$,

$$\mathbb{E}[XY] = \mathbb{E}[X]\mathbb{E}[Y] = 0,$$

which implies that H_i and H_j are orthogonal.

Proof of Theorem 1.9 (cont)

Conversely, suppose that the linear spaces H_i are pairwise orthogonal. From the definition of the independence of an infinite collection of σ-fields, it suffices to prove that, if $i_1, \ldots, i_p \in I$ are distinct, the $\sigma(H_1), \ldots, \sigma(H_p)$ are independent. For this, it suffices to verify that, if $\xi_1^1, \ldots, \xi_{n_1}^1 \in H_{i_1}, \ldots \xi_1^p, \ldots, \xi_{n_p}^p \in H_p$, the vectors $(\xi_1^1, \ldots, \xi_{n_1}^1), \ldots (\xi_1^p, \ldots, \xi_{n_p}^p)$ are independent. However, for every $j \in \{1, \ldots, p\}$ we can find an orthonormal basis $(\eta_1^j, \ldots, \eta_{m_j}^j)$ of the linear subspace of L^2 spanned by $\{\xi_1^j, \ldots, \xi_{n_j}^j\}$. The covariance matrix of the vector

$$(\eta_1^1, \ldots, \eta_{m_1}^1, \eta_1^2, \ldots, \eta_{m_2}^2, \ldots, \eta_1^p, \ldots, \eta_{m_p}^p) \tag{1}$$

is then the identity matrix. Moreover, this vector is Gaussian because its components belong to the Gaussian space H. By Proposition 1.2, $\eta_1^1, \ldots, \eta_{m_1}^1, \eta_1^2, \ldots, \eta_{m_2}^2, \ldots, \eta_1^p, \ldots, \eta_{m_p}^p$ are independent. This implies that $(\eta_1^1, \ldots, \eta_{m_1}^1), (\eta_1^2, \ldots, \eta_{m_2}^2), \ldots, (\eta_1^p, \ldots, \eta_{m_p}^p)$ are independent.

As an application of the previous theorem, we now discuss conditional expectations in the Gaussian framework. Again, the fact that these conditional expectations can be computed as orthogonal projections (as shown in the next corollary) is very particular to the Gaussian setting.

Corollary 1.10

Let H be a (centered) Gaussian space and let K be a closed linear subspace of H. Let p_K denote the orthogonal projection onto K in the Hilbert space L^2 and let $X \in H$.

(i) We have

$$\mathbb{E}[X|\sigma(K)] = p_K(X).$$

(ii) Let $\sigma^2 = \mathbb{E}[(X - p_K(X))^2]$. Then, for every Borel subset A of $\mathbb{R}$, the random variable $\mathbb{P}(X \in A|\sigma(K)]$ is given by

$$\mathbb{P}(X \in A|\sigma(K)](\omega) = Q(\omega, A),$$

where $Q(\omega, \cdot)$ denotes the $\mathcal{N}(p_K(X)(\omega), \sigma^2)$-distribution

$$Q(\omega, A) = \frac{1}{\sqrt{2\pi}} \int_A dy \exp\left(-\frac{(y - p_K(X)(\omega))^2}{2\sigma^2}\right).$$

Remarks

(i) For a general random variable X in L^2,

$$\mathbb{E}[X|\sigma(K)] = p_{L^2(\Omega,\sigma(K),\mathbb{P})}(X).$$

K is much smaller than $L^2(\Omega, \sigma(K), \mathbb{P})$.

(ii) Part (ii) of the Corollary can be interpreted by saying that the conditional distribution of X given $\sigma(K)$ is $\mathcal{N}(p_K(X), \sigma^2)$.

Proof of Corollary 1.10

((i) Let $Y = X - p_K(X)$. Then Y is orthogonal to K and, by Theorem 1.6, Y is independent of $\sigma(K)$. Then

$$\mathbb{E}[X|\sigma(K)] = \mathbb{E}[p_K(X)|\sigma(K)] + \mathbb{E}[Y|\sigma(K)] = p_K(X) + \mathbb{E}[Y] = p_K(X).$$

(ii) For every nonnegative measurable function f on $\mathbb{R}_+$,

$$\mathbb{E}[f(X)|\sigma(K)] = \mathbb{E}[f(p_K(X) + Y)|\sigma(K)] = \int P_Y(dy) f(p_K(X) + y),$$

where P_Y is the law of Y, which is $\mathcal{N}(0, \sigma^2)$ since Y centered Gaussian with variance σ^2. In the 2nd equality we used the following fact: if X is a $\mathcal{G}$-measurable random variable and if Y is independent of $\mathcal{G}$, then, for every nonnegative measurable function g, $\mathbb{E}[g(Y, Z)|\mathcal{G}] = \int g(y, Z) P_Y(dy)$. The conclusion of (ii) follows immediately.

Let $(X_t)_{t\in T}$ be a (centered) Gaussian process. The covariance function of $(X_t)_{t\in T}$ is the function $\Gamma : T \times T \mapsto \mathbb{R}$ defined by

$$\Gamma(s,t) = \mathrm{Cov}(X_s, X_t) = \mathbb{E}[X_s X_t].$$

This function characterizes the collection of finite-dimensional marginal distributions of the process $(X_t)_{t\in T}$, that is, the collection consisting for every choice of the distinct indices $t_1, \ldots, t_p$ in T, $p \geq 1$, of the law of the vector $(X_{t_1}, \ldots X_{t_p})$. Indeed this vector is a centered R^p-valued Gaussian vector in with covariance matrix $(\Gamma(t_i, t_j))_{1\leq i,j\leq p}$.

Let $(X_t)_{t\in T}$ be a (centered) Gaussian process. The covariance function of $(X_t)_{t\in T}$ is the function $\Gamma : T \times T \mapsto \mathbb{R}$ defined by

$$\Gamma(s,t) = \mathrm{Cov}(X_s, X_t) = \mathbb{E}[X_s X_t].$$

This function characterizes the collection of finite-dimensional marginal distributions of the process $(X_t)_{t\in T}$, that is, the collection consisting for every choice of the distinct indices $t_1, \dots, t_p$ in T, $p \geq 1$, of the law of the vector $(X_{t_1}, \dots X_{t_p})$. Indeed this vector is a centered R^p-valued Gaussian vector in with covariance matrix $(\Gamma(t_i, t_j))_{1\leq i,j\leq p}$.

Given a function Γ on $T \times T$, one may ask whether there exists a Gaussian process $(X_t)_{t\in T}$ whose covariance function in Γ. The function Γ must be symmetric ($\Gamma(s,t) = \Gamma(t,s)$) and of positive type in the following sense: if c is a real-valued function on T with finite support, then

$$\sum_{T\times T} c(s)c(t)\Gamma(s,t) \geq 0.$$

Indeed, if Γ is the covariance function of the process $(X_t)_{t\in T}$, we must have

$$\sum_{T\times T} c(s)c(t)\Gamma(s,t) = \mathrm{Var}\left(\sum_{T} c(s)X_s\right) \geq 0.$$

The next theorem solves the existence problem in the general case. This theorem is a direct consequence of the Kolmogorov extension theorem. We omit the proof.

Given a function Γ on $T \times T$, one may ask whether there exists a Gaussian process $(X_t)_{t\in T}$ whose covariance function in Γ. The function Γ must be symmetric ($\Gamma(s,t) = \Gamma(t,s)$) and of positive type in the following sense: if c is a real-valued function on T with finite support, then

$$\sum_{T\times T} c(s)c(t)\Gamma(s,t) \geq 0.$$

Indeed, if Γ is the covariance function of the process $(X_t)_{t\in T}$, we must have

$$\sum_{T\times T} c(s)c(t)\Gamma(s,t) = \mathrm{Var}\left(\sum_{T} c(s)X_s\right) \geq 0.$$

The next theorem solves the existence problem in the general case. This theorem is a direct consequence of the Kolmogorov extension theorem. We omit the proof.

Theorem 1.11

Let Γ be a symmetric function of positive type on $T \times T$. There exists, on an appropriate probability space $\Omega, \mathcal{F}, \mathbb{P})$, a centered Gaussian process whose covariance function is Γ.

Example

Consider the case $T = \mathbb{R}$ and let μ be a finite symmetric Borel measure on $\mathbb{R}$ ($\mu(-A) = \mu(A)$ for any Borel subset A of $\mathbb{R}$). Define, for $s, t \in \mathbb{R}$,

$$\Gamma(s, t) = \int e^{i\xi(t-s)} \mu(d\xi) \in \mathbb{R}.$$

Γ is symmetric. If c is a real-valued function on $\mathbb{R}$ with finite support,

$$\sum_{\mathbb{R}\times\mathbb{R}} c(s)c(T)\Gamma(s, t) = \int \Big|\sum_{\mathbb{R}} c(s)e^{i\xi s}\Big|^2 \mu(d\xi) \geq 0.$$

Theorem 1.11

Let Γ be a symmetric function of positive type on $T \times T$. There exists, on an appropriate probability space $\Omega, \mathcal{F}, \mathbb{P})$, a centered Gaussian process whose covariance function is Γ.

Example

Consider the case $T = \mathbb{R}$ and let μ be a finite symmetric Borel measure on $\mathbb{R}$ ($\mu(-A) = \mu(A)$ for any Borel subset A of $\mathbb{R}$). Define, for $s, t \in \mathbb{R}$,

$$\Gamma(s,t) = \int e^{i\xi(t-s)}\mu(d\xi) \in \mathbb{R}.$$

Γ is symmetric. If c is a real-valued function on $\mathbb{R}$ with finite support,

$$\sum_{\mathbb{R}\times\mathbb{R}} c(s)c(T)\Gamma(s,t) = \int \Big|\sum_{\mathbb{R}} c(s)e^{i\xi s}\Big|^2 \mu(d\xi) \geq 0.$$

Example (cont)

The function Γ has the additional property that $\Gamma(s,t)$ only depends on $t - s$. It immediately follows that any (centered) Gaussian process $(X_t)_{t\in\mathbb{R}}$ with covariance function is stationary (in a strong sense), meaning that

$$(X_{t_1+t}, X_{t_2+t}, \ldots, X_{tn1+t}) \stackrel{d}{=} (X_{t_1}, X_{t_2}, \ldots, X_{tn1})$$

for any choice of $t_1, t_2, t_n, t \in \mathbb{R}$. Conversely, any stationary Gaussian process $(X_t)_{t\in\mathbb{R}}$ has a covariance function of the preceding type, and the measure μ is called the spectral measure of $(X_t)_{t\in\mathbb{R}}$.

Outline

Definition 1.12

Let $(E, \mathcal{E})$ be a measurable space, and let μ be a σ-finite measure on $(E, \mathcal{E})$. A Gaussian white noise with intensity μ is an isometry G from $L^2(E, \mathcal{E}, \mu)$ into a (centered) Gaussian space.

Hence, if $f \in L^2(E, \mathcal{E}, \mu)$, $G(f)$ is centered Gaussian with variance

$$\mathbb{E}[G(f)^2] = \|G(f)\|^2_{L^2(E,\mathcal{E},\mu)} = \|f\|^2_{L^2(E,\mathcal{E},\mu)} = \int f^2 d\mu.$$

If $f, g \in L^2(E, \mathcal{E}, \mu)$, the covariance of $G(f)$ and $G(g)$ is

$$\mathbb{E}[G(f)G(g)] = \langle f, g\rangle_{L^2(E,\mathcal{E},\mu)} = \int fg d\mu.$$

In particular, if $f = 1_A$ with $\mu(A) < \infty$, then $G(1_A)$ is a $\mathcal{N}(0, \mu(A)$ variable, we will write $G(A) = G(1_A)$.

Definition 1.12

Let $(E, \mathcal{E})$ be a measurable space, and let μ be a σ-finite measure on $(E, \mathcal{E})$. A Gaussian white noise with intensity μ is an isometry G from $L^2(E, \mathcal{E}, \mu)$ into a (centered) Gaussian space.

Hence, if $f \in L^2(E, \mathcal{E}, \mu)$, $G(f)$ is centered Gaussian with variance

$$\mathbb{E}[G(f)^2] = \|G(f)\|^2_{L^2(E,\mathcal{E},\mu)} = \|f\|^2_{L^2(E,\mathcal{E},\mu)} = \int f^2 d\mu.$$

If $f, g \in L^2(E, \mathcal{E}, \mu)$, the covariance of $G(f)$ and $G(g)$ is

$$\mathbb{E}[G(f)G(g)] = \langle f, g \rangle_{L^2(E,\mathcal{E},\mu)} = \int fg d\mu.$$

In particular, if $f = 1_A$ with $\mu(A) < \infty$, then $G(1_A)$ is a $\mathcal{N}(0, \mu(A)$ variable, we will write $G(A) = G(1_A)$.

Let $A_1, \dots, A_n \in \mathcal{E}$ be disjoint with $\mu(A_j) < \infty$ for all $j = 1, \dots, n$. Then the vector

$$(G(A_1), \dots, G(A_n))$$

is a Gaussian vector in $\mathbb{R}^n$ and its covariance matrix is diagonal since,if $i \neq j$,

$$\mathbb{E}[G(A_i)G(A_j)] = \langle 1_{A_i}, 1_{A_j} \rangle_{L^2(E,\mathcal{E},\mu)} = 0.$$

Suppose $A \in \mathcal{E}$ with $\mu(A) < \infty$ and that A is the disjoint union of a countable collection $A_1, A_2, \dots$ of elemenst of $\mathcal{E}$. Then $1_A = \sum_{j=1}^{\infty} 1_{A_j}$ in $L^2(E, \mathcal{E}, \mu)$ and so by the isometry property

$$G(A) = \sum_{j=1}^{\infty} G(A_j), \text{ in } L^2(E, \mathcal{E}, \mu).$$

Since the random variables $G(A_j)$ are independent, the martingale convergence theorem implies that the series above also converges a.s.

Let $A_1, \ldots, A_n \in \mathcal{E}$ be disjoint with $\mu(A_j) < \infty$ for all $j = 1, \ldots, n$. Then the vector

$$(G(A_1), \ldots, G(A_n))$$

is a Gaussian vector in $\mathbb{R}^n$ and its covariance matrix is diagonal since,if $i \neq j$,

$$\mathbb{E}[G(A_i)G(A_j)] = \langle 1_{A_i}, 1_{A_j} \rangle_{L^2(E,\mathcal{E},\mu)} = 0.$$

Suppose $A \in \mathcal{E}$ with $\mu(A) < \infty$ and that A is the disjoint union of a countable collection $A_1, A_2, \ldots$ of elemenst of $\mathcal{E}$. Then $1_A = \sum_{j=1}^{\infty} 1_{A_j}$ in $L^2(E, \mathcal{E}, \mu)$ and so by the isometry property

$$G(A) = \sum_{j=1}^{\infty} G(A_j), \text{ in } L^2(E, \mathcal{E}, \mu).$$

Since the random variables $G(A_j)$ are independent, the martingale convergence theorem implies that the series above also converges a.s.

Properties of the mapping $A \mapsto G(A)$ are therefore very similar to those of a measure depending on the parameter $\omega \in \Omega$. However, one can show that, if ω is fixed, the mapping $A \mapsto G(A)(\omega)$ does not (in general) define a measure.