## 1. GAUSSIAN PROCESSES

A *Gaussian process* on a set $T$ is a collection of random variables $X = (X_t)_{t\in T}$ on a common probability space such that for any $n \geq 1$ and any $t_1, \ldots, t_n \in T$, the vector $(X(t_1), \ldots, X(t_n))$ has a normal distribution. The joint distribution of $X$ is determined by the *mean function* $t \to \mathbf{E}[X(t)]$ and the covariance kernel $K(t,s) := \mathrm{Cov}(X(t), X(s))$.

The index set $T$ is arbitrary for now, but we shall usually take it to be a separable metric space $(T,d)$, and in many discussions that the metric space is compact or at least totally bounded.

**In terms of measures:** Consider the space $\mathbb{R}^T$ of all real-valued functions on $T$. A subset of the form $\{f : f(t_i) \in A_i,\ 1 \leq i \leq n\}$ for some $n \geq 1$, $t_i \in T$ and some Borel sets $A_i \subseteq \mathbb{R}$ is called a *cylinder set*. Let $\mathcal{G}$ be the sigma-algebra generated by all cylinder sets. Equivalently, we may consider the product topology on $\mathbb{R}^T$ (by definition the smallest topology that makes the projection maps $\Pi_{t_1,\ldots,t_n}(f) = (f(t_1), \ldots, f(t_n))$ from $\mathbb{R}^T$ to $\mathbb{R}^n$ continuous) and define $\mathcal{G}$ as the Borel sigma-algebra of this topology.

Then, a measure $\mu$ on $(\mathbb{R}^T, \mathcal{G})$ is called a Gaussian measure if for every $n \geq 1$ and every $t_1, \ldots, t_n \in T$, the push-forward measure $\mu \circ \Pi_{t_1,\ldots,t_n}^{-1}$ on $\mathbb{R}^n$ is a Gaussian measure (with some mean vector and some covariance matrix).

The relationship between the language of random variables and the language of measures is this. If $X = (X_t)_{t\in T}$ is a Gaussian process, then $X : \Omega \to \mathbb{R}^T$ and the push-forward $\mathbf{P} \circ X^{-1}$ is a Gaussian measure on $\mathbb{R}^T$. Conversely, if $\mu$ is a Gaussian measure on $(\mathbb{R}^T, \mathcal{G})$, then on the probability space $(\mathbb{R}^T, \mathcal{G}, \mu)$, the co-ordinate random variables $\Pi = (\Pi_t)_{t\in T}$ form a Gaussian process.

**Exercise 1.** If $A \in \mathcal{G}$, then there exists a countable set $F \subseteq T$ such that $A$ is in the sigma-algebra generated by the projections $\Pi_t$, $t \in F$. Therefore, if $X, Y$ are two $\mathbb{R}^T$ valued random variables with the same finite-dimensional distributions, then $X$ and $Y$ have the same distribution.

**Existence of a Gaussian process:** An $n \times 1$ Gaussian vector can have any mean vector but its covariance matrix must be p.s.d. Therefore, if $X = (X_t)_{t\in T}$ is a Gaussian process with $K(t,s) = \mathrm{Cov}(X_t, X_s)$, then the matrix $K[t_1, \ldots, t_n] := (K(t_i, t_j))_{i,j\leq n}$ must be p.s..d for any $n \geq 1$ and any $t_1, \ldots, t_n \in T$. It is a pleasant and fundamental fact that no other conditions are needed.

**Theorem 2.** *Let $T$ be any index set and let $M : T \to \mathbb{R}$ be any function. Let $K : T \times T \to \mathbb{R}$ be a p.s.d. kernel, i.e., $K$ is symmetric and $(K(t_i, t_j))_{i,j\leq n}$ is a p.s.d. matrix for every $n \geq 1$ and every $t_1, \ldots, t_n \in T$.*

*Then, there exists a probability space $(\Omega, \mathcal{F}, \mathbf{P})$ and random variables $X_t : \Omega \to \mathbb{R}$ for each $t \in T$ such that $X = (X_t)_{t\in T}$ is a centered Gaussian process with mean function $M$ and covariance kernel $K$.*

*Proof.* The projection of Gaussian distribution on $\mathbb{R}^n$ with mean $(M(t_1), \ldots, M(t_n))$ and covariance matrix $K[t_1, \ldots, t_n]$, to the first $n-1$ co-ordinates, is precisely the Gaussian distribution with mean $(M(t_1), \ldots, M(t_{n-1}))$ and covariance matrix $K[t_1, \ldots, t_{n-1}]$. By the Daniell-Kolmogorov theorem (more familiar as Kolmogorov's existence/consistency theorem) we get the existence of a measure $\mu$ with the required finite dimensional distributions. ∎

**Remark 3.** In case $T$ is countable, there is a more straightforward way to see this from the existence of i.i.d. random variables. Without loss of generality let $T = \mathbb{N}$ and let $\xi_i$ be i.i.d. $N(0,1)$ random variables on some probability space (for eg., on $[0,1]$ with the Lebesgue measure on Borel sets). Then set $X_i = \sum_{j=1}^{i} \alpha_{i,j}\xi_j$ where the infinite lower-triangular matrix $A = (\alpha_{i,j})_{i,j}$ is defined by $AA^t = (K(i,j))_{i,j}$. More explicitly, the equations here are such that inductively we can write $\alpha_{i,j}$ in terms of $K(i',j')$, $j' \leq i' \leq i$.

**Exercise 4.** Let $T = [0,1]$ and let $K(t,s) = t \wedge s$. Show that there exists a centered Gaussian process with covariance kernel $K$.

**Asking for more?** Once a Gaussian process is defined, we want to calculate probabilities of interesting events. Consider the Gaussian process in Exercise **??**. The event $\{f \in \mathbb{R}^{[0,1]} : \lim f(1/n) \text{ exists}\}$ is an event in $\mathcal{G}$ and one can calculate its probability. But the following more interesting events all lie outside $\mathcal{G}$ (by Exercise **??**) and hence we cannot ask their probability.

- $\{f : f \text{ is continuous at } 0\}$.
- $C(T) = \{f : f \text{ is continuous on } T\}$.
- $B(T) = \{f : f \text{ is bounded on } T\}$.
- $\{f : f \text{ has no zeros}\}$.

Analogously, $\sup_t X_t$, $\int_0^1 X_t dt$ etc. are all non-measurable. In simpler language, although we can talk about $X_t(\omega)$ for a fixed $t$ (or countably many of them), we cannot talk about the trajectories $t \to X_t(\omega)$.

This makes the Gaussian process as defined quite uninteresting mathematically and totally inadequate to model physical situations (like trajectories of a Brownian particle, signals with noise, graph of stock prices etc.). Clearly these problems arise when $T$ is uncountable. We state two ways to deal with it. One uses measure theoretical language, the other is more down to earth.

**Random continuous functions:** Let $T$ be a metric space and let $C(T)$ be the space of continuous functions on $T$. Recall the following facts.

(1) When $T$ is compact, $C(T)$ becomes a normed-linear space with $\|f\|_T = \sup_{t\in T} |f(t)|$.

(2) When $T$ is sigma-compact, $C(T)$ becomes a metric space with $d(f,g) = \sup_n (\|f-g\|_{K_n} \wedge 1)$ where $K_n$ is any fixed sequence of compact sets that exhausts $T$.

(3) When $T$ is not sigma-compact, $C(T)$ becomes merely a topological space given by the semi-norms $\|f\|_K$ for comapct sets $K$. We do not care about this case.

In any case, $C(T)$ has a Borel sigma-algebra $\mathcal{B}(C(T))$. What we would like is random variables $X : \Omega \to C(T)$ that are measurable in the Borel sigma-algebra (then we call $X$ a $C(T)$-valued random variable).

Note that $\mathcal{B}(C(T))$ is also generated by finite-dimensional cylinder sets and hence Exercise **??** continues to be valid. However, many more interesting events become available. For example, when $T$ is compact, the set $\{f : \|f\| < 1\}$ is a Borel set in $C(T)$ and we can talk about its probability.

However, given a consistent family of finite dimensional distributions, it is not trivial to check whether a $C(T)$ valued random variable with the given finite dimensional distributions even exists. In fact the following question is a deep and fundamental (and fully solved) question that we shall study later.

**Question:** Let $T$ be a compact metric space. Given $M : T \to \mathbb{R}$ and a $K : T \times T \to \mathbb{R}$, does there exist a $C(T)$-valued random variable $X = (X_t)_{t\in T}$ such that $(X_{t_1}, \ldots, X_{t_n})$ has Gaussian distribution with mean vector $(M(t_1), \ldots, M(t_n))$ and covariance matrix $K[t_1, \ldots, t_n]$?

Show the following necessary condition.

**Exercise 5.** Show that for a positive answer to the above question, it is necessary that that $M$ and $K$ must be continuous on $T$ and $T \times T$ respectively.

There are two ways in which to rephrase the basic question.

**In the language of measure theory:** Given $M$ and $K$, we have constructed a Gaussian measure $\mu$ on $(\mathbb{R}^T, \mathcal{G}_T)$ with mean $M$ and covariance $K$. Complete the sigma-algebra under the measure $\mu$. Recall that this means considering all $A \subseteq \mathbb{R}^T$ such that there are sets $B, C \in \mathcal{G}_T$ with $B \subseteq A \subseteq C$ and such that $\mu(B) = \mu(C)$. The set

of all such subsets $A$ is a larger sigma-algebra denote $\mathcal{G}_T^\mu$. If we define $\mu(A)$ to be the common value of $\mu(B)$ and $\mu(C)$, then $\mu$ extends as a probability measure to $\mathcal{G}_T^\mu$. The completion depends on $\mu$, of course.

Now, if it happens that $\mathcal{G}_T^\mu$ contains $\mu(C(T))$ and $\mu(C(T)) = 1$, then by simply restricting it to $C(T)$ we would have answered the above question positively. Thus, the question really is, when does the completion give full measure to $C(T)$.

**Less sophisticated language:** In all cases of interest to us, $T$ will have a countable dense subset $T'$. If we also assume that $T$ is compact, then continuous functions on $T$ are in one to one correspondence with uniformly continuous functions on $T'$.

Note that there is no ambiguity of versions of a Gaussian process on $T'$, since it is countable (therefore $\mathbb{R}^{T'}$ is metrizable and $\mathcal{G}_{T'}$ is its Borel sigma-algebra). Therefore, we may simply consider $M$ and $K$ restricted to $T'$ and $T' \times T'$ respectively, construct the Gaussian process on $T'$ (which can be done b enumerating $T'$ and using a countable sequence of i.i.d. standard normals). Then the question is whether the resulting process $X'$ on $T'$ is uniformly continuous (w.p.1).

If the answer is yes, then sample by sample, it extends to a random function in $C(T)$. Further, but properties of Gaussians, then extended process with also be Gaussian. Since $M$ and $K$ are continuous on $T$ and $T \times T$, the process $X$ has this mean and covariance.

## 1. Examples of Gaussian processes

We give many examples. In all of them we take the mean function to be zero.

**Example 1.** For an arbitrary $T$, let $f_1, \ldots, f_n : T \to \mathbb{R}$ be fixed (non-random) functions. Let $\xi_i$ be i.i.d. standard normal random variables. Then define $X_t(\omega) = \sum_{i=1}^n \xi_i(\omega) f_i(t)$. Then, $X$ is a Gaussian process with covariance kernel $K(t,s) = \sum_{i=1}^n f_i(t) f_i(s)$. We study some special examples later.

(1) $T = \{1, 2 \ldots, n\}$, in which case $f_i$ may be identified with the vector $\mathbf{v}_{i=}(f_i(1), \ldots, f_i(n))^t$. Then $X$ may be written as vector $X = \xi_1 \mathbf{v}_1 + \ldots + \xi_n \mathbf{v}_n$ which is just $N_n(0, \Sigma)$ with $\Sigma = \mathbf{v}_1 \mathbf{v}_1^t + \ldots + \mathbf{v}_n \mathbf{v}_n^t$. In the notation above $K(i,j) = \sigma_{i,j}$.

(2) $T = \mathbb{R}$ or $\mathbb{C}$ and $f_i(z) = z^i$, $0 \le i \le n-1$. Then $X(t) = \sum_{k=0}^{n-1} \xi_k t^k$ is a Gaussian polynomial.

(3) $T = S^1$ or $\mathbb{R}$ and $f_k(t) = e^{i\lambda_k t}$ where $\lambda_k \in \mathbb{R}$ (if $T = S^1$ we take $\lambda_k \in \mathbb{Z}$). Then $X(t)$ is a Gaussian trigonometric polynomial.

**Example 2.** Gaussian series. Let $\xi_n$ be i.i.d. $N(0,1)$ random variables and define the following.

(1) $X(z) = \sum_{n=0}^\infty c_n \xi_n z^n$ for $z \in \mathbb{C}$ and a specified sequence $c_n$. It is a random analytic function.

(2) $X(t) = \sum_{n \in \mathbb{Z}} c_n \xi_n e^{int}$ for $t \in [0, 2\pi)$ and a specified sequence $c_n$.

For appropriate choices of $c_n$s these series will converge (for all $z$ or for all $t$, w.p.1) and then defines a Gaussian process on $\mathbb{C}$ or $[0, 2\pi)$, respectively. Similarly, whenever the random series $\sum_n \xi_n f_n(t)$ converges w.p.1. for all $t \in T$ for some deterministic functions $f_n$, we get a Gaussian process on $T$.

We shall see later that all Gaussian processes are essentially of this form, for an appropriate choice of the functions $f_n$.

**Example 3.** Let $X \sim N_n(0, \Sigma)$. We already said that this is a centered Gaussian process on $[n]$ with covariance $K(i,j) = \sigma_{i,j}$. We can expand the index set by defining $X_{\mathbf{v}} = \mathbf{v}^t X$ for all $\mathbf{v} \in \mathbb{R}^n$. Then, the collection $(X_{\mathbf{v}})_{\mathbf{v} \in \mathbb{R}^n}$ is also Gaussian process wth covariance kernel $K(\mathbf{v}, \mathbf{u}) = \mathbf{v}^t \Sigma \mathbf{u}$.

Note that $(\mathbf{v}, \mathbf{u}) := \mathbf{v}^t \Sigma \mathbf{u}$ is an inner product on $\mathbb{R}^n$ whenever $\Sigma$ is a p.d. matrix. Then, $\mathbf{v} \to X_{\mathbf{v}}$ is an isometry from this inner product space into $L^2(\Omega, \mathcal{F}, \mathbf{P})$. In other words, the collection $\{X_{\mathbf{v}} : \mathbf{v} \in \mathbb{R}^n\}$ is a "concrete realization" of the given inner product space[9].

**Example 4.** Let $\xi_n$, $n \in \mathbb{Z}$ be i.id. $N(0,1)$ random variables on a probability space $(\Omega, \mathcal{F}, \mathbf{P})$. For each $\mathbf{v} = (v_n) \in \ell^2$, define $X_{\mathbf{v}} := \sum_{n \in \mathbb{Z}} \xi_n v_n$. It is a standard fact from a first course in probability (we shall prove similar and more general things later) that for each fixed $\mathbf{v}$, the series defining $X_{\mathbf{v}}$ converges (i.e., $\lim_{m,n \to \infty} \sum_{k=-m}^n \xi_k v_k$ exists a.s.) and has $N(0, \|\mathbf{v}\|^2)$ distribution. Further, any finite number of them have a joint normal distribution and $\mathbf{E}[X_{\mathbf{v}} X_{\mathbf{u}}] = \sum_n v_n u_n$. Thus, $X = (X_{\mathbf{v}})_{\mathbf{v} \in \ell^2}$ is a Gaussian process with covariance kernel $K(\mathbf{v}, \mathbf{u}) = \langle \mathbf{v}, \mathbf{u} \rangle_{\ell^2}$.

There is a subtlety here when compared to the previous example. For every $\mathbf{v}$, let $E_{\mathbf{v}} \subseteq \Omega$ be the set of $\omega$ for which the series defining $X_{\mathbf{v}}$ converges. Then $\mathbf{P}(E_{\mathbf{v}}) = 1$ but $\mathbf{P}(\cap_{\mathbf{v}} E_{\mathbf{v}}) = 0$ (see Exercise **??**). This means that for a.e. $\omega$, there are some (in fact many!) $\mathbf{v}$ for which $X_{\mathbf{v}}(\omega)$ does not make sense!

**Exercise 5.** (1) Suppose $\sum_n u_n v_n$ converges for every $\mathbf{v} \in \ell^2$. Show that $\mathbf{u} \in \ell^2$.

(2) Deduce that $\mathbf{P}(\cap_{\mathbf{v}} E_{\mathbf{v}}) = 0$ in the above example.

**Example 6.** Continuation of the previous example. Let $V = \{\mathbf{v} \in \ell^2 : \sum_n n^2 v_n^2 < \infty\}$. It can be shown that $\mathbf{P}(E_{\mathbf{v}}) = 1$. Hence, if we consider the restricted Gaussian process $(X_{\mathbf{v}})_{\mathbf{v} \in V}$, then the "trouble" in the previous example does not arise. Not only is $X$ a Gaussian process, for a.e. $\omega$, we can consider all the values of $X_{\mathbf{v}}$ together. Indeed, $\mathbf{v} \to X_{\mathbf{v}}$ defines a random bounded linear functional of $V$.

[9] This looks like an empty exercise but here is a thing to think about. Random variables can be multiplied and hence the collection of linear combinations of $X_{\mathbf{v}_1} \ldots X_{\mathbf{v}_k}$ where $k \ge 1$ and $\mathbf{v}_i \in \mathbb{R}^n$ gives a new subspace $H$ of $L^2(\Omega, \mathcal{F}, \mathbf{P})$ which contains the original vector space $V = \{X_{\mathbf{v}} : \mathbf{v} \in \mathbb{R}^n\}$. If we stayed with $\mathbb{R}^n$ with the given inner product, can you figure out how to get this new vector space?

**Exercise 7.** Show that $\mathbf{P}(\bigcap_{\mathbf{v}\in V} E_{\mathbf{v}}) = 1$ in the above example.

**Example 8.** As a continuation of this theme, consider a weight sequence $w = (w_n)$. Let $\ell^2_{\mathbf{w}} = \{\mathbf{v} : \sum_n v_n^2 w_n < \infty\}$. Define $X_{\mathbf{v}} = \sum_n w_n v_n X_n$.

**Example 9.** Let $T = [0,1]$. Let $K(t,s) = \delta_{t-s}$. The corresponding Gaussian process is called *white noise*.

**Example 10.** Let $T = [0,1]$. Let $K(t,s) = t \wedge s$. Exercise **??** shows that $K$ is p.s.d. Hence it defines a Gaussian process. It will be shown that one can construct a probability space and random variables $X_t$ such that $t \to X_t(\omega)$ is a continuous random function for a.e. $\omega$ and then it will be called *Brownian motion*.

**Example 11.** Let $T = [0,1]^2$ and let $K((t_1,t_2),(s_1,s_2)) = (t_1 \wedge s_1) + (t_2 \wedge s_2)$. The corresponding Gaussian process can again be realized as a random continuous function and will thereafter be called *Brownian sheet*.

**Example 12.** Let $G = (V,E)$ be a finite connected graph with a specified subset $B \subseteq V$ that will be called the boundary. Let $\mathcal{V} := \{(x_v)_{v\in V} : x_v = 0 \text{ for } v \in B\}$. Then define $Q(x,x) = \sum_{u\sim v}(x_u - x_v)^2$ where the sum is over adjacent pairs of vertices in $V$. Then $Q$ is a non-degenerate quadratic form on $\mathcal{V}$ and we can talk of the Gaussian measure on $V$ with density proportional to $\exp\{-\frac{1}{2}Q(x,x)\}$ (w.r.t. $\prod_{v\notin B} dx_v$). The corresponding Gaussian process $X$ is called the *Gaussian free field on $G$ with zero boundary conditions on $B$*.

What is the covariance of $X$? It is a good exercise to work out that $\mathbf{E}[X_u X_v] = \mathcal{G}(u,v)$ where $\mathcal{G}$ is the Green's function on $V$ with Dirichlet boundary condition on $B$. One way of defining it is this. Start a simple random walk on $G$ at $u$ and stop it when it hits a vertex in $B$ for the first time. The number of visits to $v$ is a random variable whose expected value is defined to be $G(u,v)$.

**Exercise 13.** Let $T$ be a finite rooted tree and let $B = \{\text{root}\}$ be a singleton. In this case, describe the GFF on $T$ (if the question sounds vague, assume that you have a supply of i.i.d. $N(0,1)$ random variables and use them to construct the GFF on $T$).

**Example 14.** Let $T = \mathbb{R}^2$ and let $K(\mathbf{u},\mathbf{v}) = J_0(\|\mathbf{u}-\mathbf{v}\|)$ where $J_0$ is the Bessel function. It can be shown that $J_0(\|\mathbf{u}\|)$ is the Fourier transform of the uniform measure on $S^1$ and hence $K$ is a p.s.d. kernel. The corresponding Gaussian process (which can be shown to have smooth sample paths) is called the *random plane wave*.

**Stationary Gaussian process:** Let $G$ be a group that acts on an index set $T$ (a group of bijections of $T$ under composition). A stochastic process $X = (X_t)_{t\in T}$ is said to be invariant or stationary (under the action of $G$) if $(X_{g(t)})_{t\in T} \overset{d}{=} (X_t)_{t\in T}$.

Since finite dimensional distributions determine the distribution of a process, the condition is equivalent to asking that $(X_{g(t_1)},\dots,X_{g(t_n)}) \overset{d}{=} (X_{t_1},\dots,X_{t_n})$ for all $t_i \in T$ and for all $g \in G$. In case of a Gaussian process the condition may be stated in terms of the mean and covariance alone, i.e., $M(g(t)) = M(t)$ and $K(g(t),g(s)) = K(t,s)$ for all $t,s \in T$ and all $g \in G$.

Among the example stated earlier, the random plane wave is a stationary process (under translations and rotations of the plane). Brownian motion is not, since $\text{Var}(X_t) = t$ depends on $t$. Here is a variant of it that is.

**Example 15.** Let $T = \mathbb{R}$ and let $K(t,s) = e^{-|t-s|}$. This is a valid covariance kernel and the corresponding Gaussian process (which can be made to have continuous sample paths) is a stationary process (under translations of the line). It is known as the *Ornstein-Uhlenbeck process*. It is related to Brownian motion in a simple way. If $(W_t)_{t\in(0,\infty)}$ is a Brownian motion (covariance is $t \wedge s$), then $X_t := e^{-t/2} W_{e^t}$ is an Ornstein-Uhlenbeck process.