Gaussian processes

A real random process $X = (X_t; t \in T)$ is a collection of real random variables X_t defined on the same probability space $(\Omega, \mathcal{A}, \mathbb{P})$.

Definition 1. *The process X is Gaussian if all linear combination of its coordinates is a Gaussian random variable.*

1 Mean and covariance

Thus, for all $n \geq 1, t_i \in T, u_i \in \mathbb{R} (i = 1, \dots n)$ the real random variable $\sum_{i=1}^n u_i X(t_i)$ is a Gaussian random variable. This is in particular the case for the coordinates $X_t \equiv X(t)$, which are square integrable and we consider their mean and variance.

Definition 2. *For a square-integrable process X, we define its mean function $m : T \to \mathbb{R}$ and covariance function $R : T \times T \to \mathbb{R}$ by*

$$m(t) = \mathbb{E}X(t), \qquad R(s,t) = \mathrm{Cov}(X(s), X(t)).$$

If X is Gaussian, its law is uniquely determind by these two functions, and is denoted by $\mathcal{N}(m, R)$.

Indeed, its law is determined by all the finite dimensional marginal, which are Gaussian vectors by definition: they themselves are determined by mean and covariance matrix.

Covariance functions have a remarkable property: a function $\rho : T \times T \to \mathbb{R}$ is called *semi-definite positive* if it is symmetric (1) and for all $n \geq 1, t_i \in T, u_i \in \mathbb{R} (i = 1, \dots n)$,

$$\sum_{i,j=1}^n u_i u_j \rho(t_i, t_j) \geq 0. \tag{1}$$

Note that $\rho(s,t)^2 \leq \rho(s,s)\rho(t,t)$.

Covariance functions are semi-definite positive :

$$\sum_{i,j=1}^n u_i u_j R(t_i, t_j) = \mathbb{E}\big[\big(\sum_{i=1}^n u_i X(t_i) \big)^2 \big] \geq 0$$

Given a arbitrary function $m : T \to \mathbb{R}$ and a semi-definite positive function ρ, there exists a Gaussian process X with law $\mathcal{N}(m, \rho)$, and the process is unique in law.

(1) Corresponding to complex r.v.'s, a more general setup is for complex valued $\rho : T \times T \to \mathbb{C}$, hermitian ($\rho(s,t) = \overline{\rho}(t,s)$), semi-definite positive : $\sum_{i,j=1}^n u_i \overline{u_j} \rho(t_i, t_j) \geq 0$ for $u_i \in \mathbb{C}$ (see Remark 6). We consider the symmetric case only, because we focus on real r.v.'s

□ We have already seen uniqueness. Existence follows from Kolmogorov extension theorem. Indeed, for all $I \subset\subset T$ denote by Q_I the Gaussian law on $\mathbb{R}^{|I|}$ with mean and covariance matrix given by the restrictions of m and ρ on I. Since subvectors of a Gaussian vector are still Gaussian, we see that the family $(Q_I; I \subset\subset T)$ is consistent. By Kolmogorov extension theorem there exists a probability measure on $(\mathbb{R}^T, \mathcal{B}(\mathbb{R})^{\otimes T})$ which projection on I is Q_I. ∎

2 Stationary process, independent increments process

Definition 3. *When $T = \mathbb{Z}$ or $\mathbb{R}$, the process X is called stationary if for all $t_0 \in T$,*

$$(X(t_0 + t); t \in T) \overset{\text{law}}{=} (X(t); t \in T)$$

Then, m is constant, and $\mathbb{E}[X(t)X(s)] = \mathbb{E}[X(t-s)X(0)]$ (take $t_0 = -s$), and

$$R(s,t) = r(s-t)$$

with a symmetric function r (i.e., $r(u) = R(u,0)$). This function r has the properties

$$r(-s) = r(s), \qquad \sum_{i,j=1}^{n} u_i u_j r(t_i - t_j) \geq 0.$$

Such a function is called semi-definite positive .

Definition 4. *A process on $T = \mathbb{Z}$ or $\mathbb{R}$ has independent increments if for all $n \geq 1, t_1 < t_2 < \ldots < t_n$, the increments $X(t_2) - X(t_1), \ldots, X(t_n) - X(t_{n-1})$ are independent.*

Then, $R(s,t) = \bar{r}(s \wedge t)$, with $\bar{r}(s) = R(s,s)$. Indeed, if $s \leq t$,

$$\begin{aligned} R(s,t) &= \operatorname{Cov}\big(X(s), X(s) + [X(t) - X(s)]\big) \\ &= \operatorname{Cov}(X(s), X(s)) + \operatorname{Cov}(X(s), X(t) - X(s)) \\ &= R(s,s) + 0. \end{aligned}$$

3 Stationary Gaussian process

Recall that the covariance function of a stationary process has the form $\operatorname{Cov}\big(X(s), X(t)\big) = r(s-t)$ with a semi-definite positive function.

Theorem 5 (Bochner)**.** *let $T = \mathbb{R}$. Continuous semi-definite positive function are Fourier transforms of finite, symmetric positive measures, and reciprocally. The symmetric measure μ with finite mass such that*

$$r(t) = \int_{\mathbb{R}} e^{itx} d\mu(x) \tag{2}$$

is called the spectral measure of the process X.

Remark 6. *Bochner theorem gives in fact the full characterization of characteristic functions. It is needed to define semi-definite positive function with complex values in the LHS of (2), and to consider non-symmetric measures μ in the RHS of the same equation.*

4 Main examples of centered Gaussian processes

Example I. *Brownian motion (Bm): $R(s,t) = s \wedge t$ on $T = \mathbb{R}^+$.*

Increments are independent, and also stationary: for $0 \le s \le t$,

$$\mathrm{Var}\big(X(t) - X(s)\big) = t - s$$

depends only on the length of the time interval.

Example II. *Stationary Ornstein-Uhlenbeck process (OU): $R(s,t) = \exp\{-|s-t|\}$ on $T = \mathbb{R}$.*

Since the covariance only depends on the difference $t - s$, the process is stationary, and we compute its spectral measure, see Bochner's theorem, which density exists and is given by Fourier inversion formula:

$$\begin{aligned}
\frac{d\mu(x)}{dx} &= \frac{1}{2\pi} \int_{\mathbb{R}} e^{-itx} e^{-|t|} dt \\
&= \frac{1}{2\pi} \left(\int_{-\infty}^{0} e^{-itx} e^{+t} dt + \int_{0}^{\infty} e^{-itx} e^{-t} dt \right) \\
&= \frac{1}{2\pi} \left(\frac{-1}{it-1} + \frac{1}{it+1} \right) \\
&= \frac{1}{\pi(1+t^2)},
\end{aligned}$$

the Cauchy density on $\mathbb{R}$.

Example III. *Brownian bridge (Bb): $R(s,t) = s \wedge t - st$ on $T = [0,1]$.*

It starts and ends at same point $X(0) = X(1) = 0$, and it behaves roughly like Brownian motion in between. To stress on the effect of pinning the process at value 0 at times $t = 0$ and $t = 1$, one could write the covariance function for $0 \le s \le t \le 1$ as $R(s,t) = s \times (1-t)$.

Example IV. *Fractional Brownian motion (fBm): on* $T = \mathbb{R}^+$,

$$R(s,t) = \frac{1}{2}\big(t^{2H} + s^{2H} - |t-s|^{2H}\big),$$

with Hurst index $H \in (0,1)$.

(i) The fractional Brownian motion has a nice scaling property: For any positive constant a, it is invariant under space-time rescaling:

$$Y = X \quad \text{in law, with} \quad Y(t) = a^{-H} X(at).$$

Indeed, Y is Gaussian, centered, with covariance R.

(ii) Increments of fBm are stationary: precisely[(2)], the law of $X(t) - X(s)$ is $\mathcal{N}(0, |t-s|^{2H})$

(iii) Particular cases. The increments of fBm are not independent, except for $H = 1/2$ where X is the Brownian motion. The case $H = 1$ is trivial, since $X(t) = tZ, \forall t \geq 0$ where the r.v. Z is given by $X(1)$[(3)].

(iv) Correlations of increments of fBm: we measure them by

$$\rho(n) = \mathrm{Cov}\big(X(1){-}X(0), X(n{+}1){-}X(n)\big) = \frac{1}{2}\big\{(n{+}1)^{2H}{+}(n{-}1)^{2H}{-}2n^{2H}\big\},$$

which behaves like $\rho(n) \sim H(2H-1)n^{-2(1-H)}$ as $n \to \infty$. (We assume $H \neq 1/2$.) We distinguish 2 cases:

- When $H > 1/2$, $\sum_n |\rho(n)| = \infty$: there is long-range dependence;
- If $H < 1/2$, $\sum_n |\rho(n)| < \infty$: fBm has short memory.

Exercise 1. *Let* $B = (B(t); t \geq 0)$ *be a Brownian motion.*

1. *Show that* $X(t) = e^{-t} B\big(e^{2t}\big), t \in \mathbb{R}$ *is an Ornstein-Uhlenbeck process. For that, check that* X *is a Gaussian process with the right mean and covariance.*
2. *Show that* $Y(t) = (1-t)B\big(\frac{t}{1-t}\big), Z(t) = tB\big(\frac{1}{t} - 1\big), t \in [0,1)$, *are Brownian bridges.*

Remark 7. *White noise is well-known to engineers and physicists, it would correspond to taking* $R(s,t) = r(s-t)$ *with*

$$r(u) = \delta_0(u) = \lim_{\varepsilon \to 0} \frac{1}{2\varepsilon} \mathbf{1}\{|u| \leq \varepsilon\},$$

(2) Check the computations !

(3) Compute $\|X(t) - tZ\|_2^2$ to prove the claim.

which is not a function but a measure. Hence white noise does not fit in our framework. White noise is not a collection of finite random variables, since the variance $R(t,t)$ is infinite. From the identity

$$\delta_0(t) = \frac{1}{\sqrt{2\pi}} \int_{\mathbb{R}} e^{itx} dx,$$

we note that the spectral measure is proportional to Lebesgue measure. All frequencies are present, explaining the name "white noise".

5 Elements of Gaussian calculus

Proposition 8 (Integration by parts)**.** *For $F : \mathbb{R} \to \mathbb{R}$ differentiable with $\lim_{|x|\to\infty} F(x)e^{-x^2/2\sigma^2} = 0$ and $g \sim \mathcal{N}(0, \sigma^2)$,*

$$\mathbb{E}gF(g) = \sigma^2 \mathbb{E}F'(g).$$

□ With $p(x) = (\sigma\sqrt{2\pi})^{-1} e^{-x^2/(2\sigma^2)}$,

$$\mathbb{E}gF(g) = \int gF(g)p(g)dg = \Big[- \sigma^2 p(g)F(g)\Big]_{-\infty}^{+\infty} + \sigma^2 \int F'(g)p(g)dg$$

by integration by parts. ■

Exercise 2. *Prove that $\mathbb{E}\left(g^2 e^{ag^2}\right) = \frac{\sigma^2}{\sqrt{1-2a\sigma^2}}$ for $2a\sigma^2 < 1$.*

□ By Proposition 8

$$\mathbb{E}\left(g^2 e^{ag^2}\right) = \mathbb{E}\left(g \times g e^{ag^2}\right) = \sigma^2 \mathbb{E}\left((1 + 2ag^2)e^{ag^2}\right),$$

though $\mathbb{E}e^{ag^2} = \frac{1}{\sqrt{1-2a\sigma^2}}$. ■

Proposition 9 (Integration by parts)**.** *For $F : \mathbb{R}^n \to \mathbb{R}$ smooth with moderate increase and $(g, z_1, \ldots, z_n)$ a centered Gaussian vector,*

$$\mathbb{E}gF(z_1, \ldots, z_n) = \sum_{\ell \le n} \mathbb{E}(gz_\ell)\mathbb{E}\left(\frac{\partial F}{\partial z_\ell}(z_1, \ldots, z_n)\right).$$

□ With $z'_\ell = z_\ell - \alpha_\ell g$ and $\alpha_\ell = \frac{\mathbb{E}(gz_\ell)}{\mathbb{E}(g^2)}$, the vector $(g, z'_1, \ldots, z'_n)$ is Gaussian with g independent of $(z'_1, \ldots, z'_n)$. Then the LHS is equal to

$$\mathbb{E}gF(z'_1+\alpha_1 g, \ldots, z'_n+\alpha_n g) = \sum_{\ell \le n} \mathbb{E}(g^2)\alpha_\ell \mathbb{E}\left(\frac{\partial F}{\partial z_\ell}(z'_1 + \alpha_1 g, \ldots, z'_n + \alpha_n g)\right)$$

by using Proposition 8 for fixed $(z'_1, \ldots, z'_n)$. ■

Theorem 10. *Let* $\mathbf{X}$ *and* $\mathbf{Y}$ *denote two centered Gaussian vectors in* $\mathbb{R}^n$*, with covariance matrices* Γ^X, Γ^Y*. Assume there exist subsets* A, B *of* $T \times T$ *(with* $T = \{1, \ldots, n\}$*) such that*

$$\begin{cases} \Gamma^X(i,j) \le \Gamma^Y(i,j), & (i,j) \in A, \\ \Gamma^X(i,j) \ge \Gamma^Y(i,j), & (i,j) \in B, \\ \Gamma^X(i,j) = \Gamma^Y(i,j), & (i,j) \notin A \cup B. \end{cases}$$

If $F : \mathbb{R}^n \to \mathbb{R}$ *is smooth, with moderate growth as well as its derivatives, and if*

$$\begin{cases} \frac{\partial^2 F}{\partial x_i \partial x_j} \ge 0, & (i,j) \in A, \\ \frac{\partial^2 F}{\partial x_i \partial x_j} \le 0, & (i,j) \in B, \end{cases}$$

then

$$\mathbb{E}F(\mathbf{X}) \le \mathbb{E}F(\mathbf{Y}).$$

□ With independent copies $\mathbf{X}, \mathbf{Y}$, consider the "linear" combination (in the Gaussian sense)

$$\mathbf{X}(t) = (1-t)^{1/2}\mathbf{X} + t^{1/2}\mathbf{Y}.$$

Differentiating in t we get

$$2X_i'(t) = -(1-t)^{-1/2}X_i + t^{-1/2}Y_i.$$

With $\phi(t) = \mathbb{E}F(\mathbf{X}(t))$,

$$\phi'(t) = \sum_{i=1}^n \mathbb{E}\frac{\partial F}{\partial x_i}(\mathbf{X}(t))X_i'(t). \tag{3}$$

By independence of $\mathbf{X}$ and $\mathbf{Y}$,

$$\mathbb{E}X_i(t)'X_j(t) = \frac{1}{2}\big(\Gamma^Y(i,j) - \Gamma^X(i,j)\big).$$

Using Proposition 9 in (3), we get

$$\phi'(t) = \frac{1}{2}\sum_{i,j=1}^n \left[\Gamma^Y(i,j) - \Gamma^X(i,j)\right]\mathbb{E}\frac{\partial^2 F}{\partial x_i \partial x_j}(\mathbf{X}(t)).$$

By assumption, the product in the above sum is non-negative for all choice of i, j. So, $\phi(0) \le \phi(1)$. ■

Corollary 11. *Let* $\mathbf{X}$ *and* $\mathbf{Y}$ *denote two centered Gaussian vectors in* $\mathbb{R}^n$*, with covariance matrices* Γ^X, Γ^Y*. Assume*

$$\Gamma^X(i,i) = \Gamma^Y(i,i), \qquad \Gamma^X(i,j) \ge \Gamma^Y(i,j) \quad \forall i \ne j.$$

Then, $\max_i X_i \le_{st} \max_i Y_i$*, i.e.,*

$$\mathbb{P}(\max_i X_i > x) \le \mathbb{P}(\max_i Y_i > x) \quad \forall x \in \mathbb{R}.$$

In particular, $\mathbb{E}\max_i X_i \le \mathbb{E}\max_i Y_i$*.*

□ Let $f_k : \mathbb{R} \to [0,1]$ be a sequence of smooth non-increasing functions, converging to $f(y) = \mathbf{1}_{y \le x}$, and $F_k(\mathbf{y}) = \prod_{i \le n} f_k(y_i)$. Since $\frac{\partial^2 F_k}{\partial y_i \partial y_j} \ge 0$ for $i \ne j$, we can apply Theorem 10 to $1 - F_k$ with B the complement of the diagonal and $A = \emptyset$, so that

$$\mathbb{E}F_k(\mathbf{X}) \ge \mathbb{E}F_k(\mathbf{Y}).$$

Now, we can further assume that $f_k \nearrow f$, so we obtain, by monotone convergence as $k \to \infty$,

$$1 - \mathbb{P}(\max_i X_i > x) \ge 1 - \mathbb{P}(\max_i Y_i > x).$$

The last clam follows by integration, since

$$\mathbb{E}X^+ = \int_0^\infty \mathbb{P}(X \ge y)dy, \qquad \mathbb{E}X^- = \int_0^\infty \mathbb{P}(X \le -y)dy,$$

for a r.v. X. ∎

Theorem 12 (Gaussian Lipschitz concentration inequality)**.** *Let $\mathbf{X} \sim \mathcal{N}(0, I_n)$, and $F : \mathbb{R}^n \to \mathbb{R}$ a Lipschitz function. Then, for all $r > 0$,*

$$\mathbb{P}\big(|F(\mathbf{X}) - \mathbb{E}F(\mathbf{X})| \ge r\big) \le 2\exp\big\{ -\frac{r^2}{2L_F^2}\big\},$$

with $L_F = \sup_{\mathbf{x} \ne \mathbf{y}} |F(\mathbf{x}) - F(\mathbf{y})|/\|\mathbf{x} - \mathbf{y}\|$.