GAUSSIAN PROCESSES – EXAMPLES SHEET 1

Part III, Lent 2017, RN (email: r.nickl@statslab.cam.ac.uk)

1. Let T be any non-empty set and let Φ be a *covariance* mapping defined on it, that is, $\Phi : T \times T \mapsto \mathbb{R}$ is such that for all $n \in \mathbb{N}, t_1, \dots, t_n \in T$, the matrix $(\Phi(t_i, t_j))_{i,j=1}^n$ is symmetric and and non-negative definite. Use Kolmogorov's consistency theorem to show that there exists a centred Gaussian process $(X(t) : t \in T)$ such that $E[X(s)X(t)] = \Phi(s,t)$.

2. Use the previous exercise to show the existence of Gaussian processes with the following index sets and covariances, and find their intrinsic covariance metrics d_X.
i) $T = [0,1]$, $\Phi(s,t) = \min(s,t)$ (the Wiener or Brownian motion process)
ii) $T = [0,1]$ and $\Phi(s,t) = \min(s,t) - st$ (the Brownian bridge process).
iii) $T = [0,1]$ and $\Phi(s,t) = s^{2H} + t^{2H} - |s-t|^{2H}$ for some $H \in (0,1)$ (the fractional Brownian motion process with Hurst index H)
iv) $T = [0,1]^n$ and $\Phi(s,t) = \min(s_1,t_1) \times \dots \times \min(s_n,t_n)$ (the Brownian sheet process).

3.* Let $(B, \|\cdot\|_B)$ be a separable Banach space with topological dual space $B^* = \{f : B \to \mathbb{R}$ linear and continuous$\}$ normed by $\|f\|_{B^*} = \sup_{\|x\|_B \le 1} |f(x)|$. Use the Hahn-Banach theorem from functional analysis to show that there exists a countable subset D of the unit ball of B^* such that $\|x\|_B = \sup_{f \in D} |f(x)|$. Show further that any Gaussian random variable X taking values in B defines a Gaussian process $(\tilde{X}(f) : f \in B^*)$ through the map $f \mapsto \tilde{X}(f) \equiv f(X), f \in B^*$ and deduce that $\|X\|_B = \sup_{f \in D} |\tilde{X}(f)|$.

4. Let $K : \mathbb{R}^n \to \mathbb{R}$ be a function and consider a mapping $\Phi(s,t) = K(s-t), s,t \in \mathbb{R}^n$. Show that if $K(x) = \int_{\mathbb{R}^n} e^{iux} f(u) du$ for some non-negative, symmetric and integrable function f, then Φ defines a covariance on $T = \mathbb{R}^n$, and deduce the existence of a Gaussian process that has Φ as a covariance.

5. Let X and Y be independent identically distributed centred normal random vectors in $\mathbb{R}^n$. For $\theta \in [0, 2\pi]$, define $X(\theta) = X \sin\theta + Y \cos\theta$ and let $X'(\theta) = (d/d\theta) X(\theta)$. Show that the random vectors (X, Y) and $(X(\theta), X'(\theta))$ have the same distribution in $\mathbb{R}^n \times \mathbb{R}^n$

6. Let $(X_n : n = 1, 2, \dots)$ be an infinite sequence of jointly normal random variables such that $EX_n = 0, EX_n^2 = 1$ for all n. Show that for some $\alpha > 0$,

$$E \exp\left\{ \alpha \left[\sup_n \left(\frac{|X_n|}{\sqrt{\log(1+n)}} \right) \right]^2 \right\} < \infty.$$

7. Let $(X(t) : t \in T)$ be a separable centred Gaussian process and denote its norm by $\|X\|_T = \sup_{t \in T} |X(t)|$. Assume $\Pr(\|X\|_T < \infty) > 0$. Show that for all $p \ge 1$ there exists a finite positive constant K_p that depends only on p such that

$$(E\|X\|_T^p)^{1/p} \le K_p E\|X\|_T.$$

8. Prove the any finite Borel measure μ on a complete separable metric space S is *tight* (or *Radon*), that is, for every $\varepsilon > 0$ there exists a compact set K_ε such that $\mu(S \setminus K_\varepsilon) < \varepsilon$.

9. Prove that fractional Brownian motion with Hurst index $H > 0$ is almost surely sample-continuous on $[0,1]$ for the usual distance $d(s,t) = |s-t|$, and deduce that it defines a tight Gaussian random variable on the Banach space $C([0,1])$.

10. Prove the sample-continuity of the Brownian sheet in dimension n for its intrinsic covariance metric on $[0,1]^n$.

11. Let ε be a standard normal vector in $\mathbb{R}^p, p \geq 2$. Let

$$B_0(k) = \{\theta \in \mathbb{R}^p : \theta_j \neq 0 \text{ for at most } k \text{ vector components}\}, k \leq p,$$

be the set of k-sparse vectors in $\mathbb{R}^p$. Show that for any $1 \leq k \leq p$,

$$E \sup_{\theta \in B_0(k), \|\theta\| \leq 1} |\varepsilon^T \theta| \leq 2\sqrt{k \log p}.$$

12. [Dudley's metric entropy bound extends to *sub*-Gaussian and *Rademacher* processes.] Say that a stochastic process $(Y(t) : t \in T)$ is sub-Gaussian for a metric d on T if

$$Ee^{\lambda(Y(s)-Y(t))} \leq e^{\lambda^2 d^2(s,t)/2} \quad \forall \lambda > 0, \forall s, t \in T.$$

a) Let $\varepsilon_1, \dots, \varepsilon_n$ be i.i.d. Rademacher random variables such that $\Pr(\varepsilon = \pm 1) = 1/2$, and let T be a bounded subset of $\mathbb{R}^n$. Show that the Rademacher process

$$Y(t) = \sum_{i=1}^n \varepsilon_i t_i, \; t = (t_1, \dots, t_n) \in \mathbb{R}^n$$

is sub-Gaussian for the distance $d(s,t) = \|s - t\|$ where $\| \cdot \|$ is the usual Euclidean norm on $\mathbb{R}^n$.

b) Show that the 'chaining lemma' for Gaussian processes from lectures remains true for 'sub-Gaussian' processes.

c) Deduce that if there exists $t_0 \in T$ such that $Y(t_0) = 0$, and T is countable, then for some numerical constant $C > 0$,

$$E \sup_{t \in T} |Y(t)| \leq C \int_0^D \sqrt{H(T, d, \varepsilon)} d\varepsilon,$$

where D is the d-diameter of T and $H(T, d, \epsilon)$ is the d-metric entropy of T.