These notes have not received the scrutiny of publication. They could be missing important references, etc.

Title

1 Gaussian Process

Definition. Let T be an index set. A collection of random variables $\{g_t\}_{t\in T}$ is called a *Gaussian Process (GP)* if for every finite subset $S \subset T$, the random vector

$$g_S := \{g_t\}_{t\in S}$$

follows a joint Gaussian distribution, i.e.

$$p_{g_S}(x) = \frac{1}{(2\pi)^{d/2}|\Sigma|^{1/2}} \exp\bigl(-\tfrac{1}{2}(x-\mu)^\top \Sigma^{-1}(x-\mu)\bigr)\,.$$

Mean-zero GP. A Gaussian process is *mean-zero* if

$$\mathbb{E}[g_t] = 0, \quad \forall\, t \in T.$$

Example 1.1: Gaussian Random Walk

Let $T = \{0, 1, 2, \ldots\}$ and define

$$X_0 = 0, \quad X_t = \sum_{s=1}^{t} Z_s,$$

where $Z_1, Z_2, \ldots \overset{\text{i.i.d.}}{\sim} \mathcal{N}(0,1)$. Then

$$\mathbb{E}[X_t] = 0, \quad \mathrm{Cov}(X_t, X_s) = \mathrm{Cov}\left(\sum_{i=1}^{t} Z_i, \sum_{j=1}^{s} Z_j\right) = \sum_{i=1}^{s\wedge t} \mathrm{Var}(Z_i) = s \wedge t.$$

2 Brownian Motion (Wiener Process)

A Brownian motion $\{W_t\}_{t\geq 0}$ satisfies:

$$W_0 = 0, \quad \mathbb{E}[W_t] = 0, \quad \mathrm{Cov}(W_s, W_t) = s \wedge t,$$

and the sample paths $t \mapsto W_t$ are almost surely continuous.

3 Canonical GP over $T \subset \mathbb{R}^n$

Let $g \sim \mathcal{N}(0, I_n)$ and define

$$X_t = \langle g, t\rangle, \quad t \in T.$$

Define

$$W(T) = \mathbb{E}_g \sup_{t\in T} X_t.$$

Then

$$\mathrm{Var}(X_t) = \mathbb{E}\langle g, t\rangle^2 = \langle t, t\rangle = \|t\|^2, \quad \mathrm{Var}(X_t - X_s) = \|t - s\|^2.$$

Example 3.1: The Unit Sphere

Let $S_{N-1} := \{x \in \mathbb{R}^N : \|x\|_2 = 1\}$. Then

$$W(S_{N-1}) = \mathbb{E}_g \sup_{x \in S_{N-1}} \langle g, x \rangle = \mathbb{E}_g \frac{\langle g, g \rangle}{\|g\|_2}.$$

By Jensen's inequality,

$$\mathbb{E}\|g\|_2 \leq \sqrt{\mathbb{E}\|g\|_2^2} = \sqrt{\mathbb{E}\sum_{i=1}^N g_i^2} = \sqrt{N}.$$

Example 3.2: Probability Simplex

Let

$$\Delta_N := \{p \in \mathbb{R}_{\geq 0}^{N+1} : \sum_{i=1}^{N+1} p_i = 1\}.$$

Then

$$W(\Delta_N) = \mathbb{E}_{g \sim \mathcal{N}(0, I_n)} \sup_{p \in \Delta_N} \langle p, g \rangle = \mathbb{E}_{g \sim \mathcal{N}(0, I_n)} \max_{i \leq N} g_i.$$

4 Theorem

$$\lim_{N \to \infty} \frac{\mathbb{E} \sup_i g_i}{\sqrt{2 \log N}} = 1.$$

Proof:

Lower Bound. Let $\varepsilon > 0$ and $Z \sim \mathcal{N}(0, 1)$. Consider

$$P\Big\{Z \in [(\sqrt{2} - \varepsilon)\sqrt{\log N}, \sqrt{2 \log N}]\Big\}.$$

Since $e^{-x^2/2} \geq e^{-\log N} = \frac{1}{N}$ for $x \leq \sqrt{2 \log N}$,

$$P \geq \frac{1}{\sqrt{2\pi}} \int_{(\sqrt{2}-\varepsilon)\sqrt{\log N}}^{\sqrt{2 \log N}} e^{-x^2/2}\, dx \geq \frac{1}{\sqrt{2\pi}} \varepsilon \sqrt{\log N} \frac{1}{N} = \frac{\varepsilon}{\sqrt{2\pi}} \frac{\sqrt{\log N}}{N}.$$

Then since $\text{Binomial}\Big(N, \frac{\sqrt{\log N}}{N}\Big) \approx \text{Poisson}(\sqrt{\log N})$, so

$$P[\text{Poisson}(\sqrt{\log N}) = 0] \to 0 \quad (N \to \infty),$$

hence

$$P\Big\{\nexists g_i \in [(\sqrt{2} - \varepsilon)\sqrt{\log N}, \sqrt{2 \log N}]\Big\} \to 0.$$

Letting $\varepsilon \to 0$ completes the lower bound.

Upper Bound. We claim

$$\mathbb{E} \max_{i \leq N} g_i \leq \sqrt{2 \log N}.$$

Observation: For any $a_i \in \mathbb{R}$ and $\lambda > 0$,

$$\max_i a_i \leq \frac{1}{\lambda} \log \sum_i e^{\lambda a_i}.$$

Thus

$$\mathbb{E}\max_i g_i \leq \frac{1}{\lambda}\mathbb{E}\log\sum_{i=1}^{N} e^{\lambda g_i} \leq \frac{1}{\lambda}\log\sum_{i=1}^{N}\mathbb{E}e^{\lambda g_i} = \frac{1}{\lambda}\big(\log N + \tfrac{1}{2}\lambda^2\big).$$

Minimizing over $\lambda > 0$ gives

$$\inf_{\lambda>0}\left(\frac{\log N}{\lambda} + \frac{\lambda}{2}\right) = \sqrt{2\log N}.$$

5 Random Energy Model

Setting: $x \in \{\pm 1\}^n$, $N = 2^n$, and define random variables

$$E(x) \sim \mathcal{N}\big(0, \tfrac{n}{2}\big)\,.$$

Let $\beta > 0$ and define

$$p_\beta(x) = \frac{1}{Z_\beta}e^{-\beta E(x)}, \quad Z_\beta = \sum_{x\in\{\pm 1\}^n} e^{-\beta E(x)}.$$

Then

$$\mathbb{E}\log Z_\beta = \mathbb{E}\log\sum_{x\in\{\pm 1\}^n} e^{-\beta E(x)}.$$

To be continued next class.