Integral Entropy

> *Since there are some problems about mathjax, I also upload the pdf version /Large Sample / 6 Integral Entropy.*

Basic Definition

- **Stochastic Process** $\{X_t\}_{t\in\mathcal{T}}$ indexed by $\mathcal{T}$: A collection of real valued random variables.
- **Sub-Gaussian Process** $\{X_t\}_{t\in\mathcal{T}}$ indexed by $(\mathcal{T}, \rho)$:
  $\mathbb{E}\left[\exp\{\lambda\left(X_s - X_t\right)\}\right] \leq \exp\left\{\frac{\lambda^2\rho(s,t)^2}{2}\right\} \quad \forall\lambda > 0, s, t \in \mathcal{T}$.
  - e.g.1. Let $\mathcal{T} = \mathbb{R}^d$ and $Z \sim N\left(0, \sigma^2 I_d\right)$, then $X_t = t^T Z$ is a sub-gaussian process;
  - e.g.2. Consider $(T, \|\|)$, $X_i \in \mathcal{X},\ l: T\times\mathcal{X}\to\mathbb{R}$ and $|l(t,x) - l(s,x)| \leq \|t-s\|$. For a sequence of i.i.d Rademacher random variables ϵ_i, then $Z_t = \sum_{i=1}^n \epsilon_i l\left(t, X_i\right)$ is a sub-Gaussian process with $\rho(s,t) = \sqrt{n}\|s-t\|$.

Dudley's integral entropy

Goal: Bound $\mathbb{E}\left(\sup_{t\in\mathcal{T}} X_t\right)$ where $\{X_t\}_{t\in\mathcal{T}}$ is a separable sub-Gaussian and mean-zero process w.r.t the metric ρ;

Tool: Chaining;

Conclusion: $\mathbb{E}\left(\sup_t X_t\right) \leq 4\sqrt{2}\int_0^{D/2} \sqrt{\log N(u,\mathcal{T},\rho)}\mathrm{d}u$

Sketch Proof:

- Preparation:
  - Let $\epsilon_k = 2^{-k}D$, where $D = \sup_{s,t\in\mathcal{T}}\rho(s,t)$;
  - Let $\mathcal{T}_0 = \{t_0\}, \mathcal{T}_1, \mathcal{T}_2, \ldots$ be a sequence of ϵ_k-net of $\mathcal{T}$;
  - Define $\pi_k : \mathcal{T}\to\mathcal{T}_k$ that $\rho\left(r, \pi_k(r)\right) \leq \epsilon_k$.
- Decomposition for $t \in \mathcal{T}_k$:
  - $X_t - X_{t_0} = \sum_{i=1}^k \left(X_{\pi_i\circ\ldots\circ\pi_k(t)} - X_{\pi_{i-1}\circ\pi_i\cdots\circ\pi_k(t)}\right)$

- Bound the maximum for $t \in \mathcal{T}_k$:

$$\begin{aligned}\max_{t\in\mathcal{T}_k}\left(X_t - X_{t_0}\right) &\leq \sum_{i=1}^{k}\max_{t\in\mathcal{T}_k}\left(X_{\pi_i\circ\ldots\circ\pi_k(t)} - X_{\pi_{i-1}\circ\pi_i^{\cdots}\circ\pi_k(t)}\right) \\ &\leq \sum_{i=1}^{k}\max_{t\in T_i}\left(X_t - X_{\pi_{i-1}(t)}\right)\end{aligned}$$

  - $\max_{t\in T_i}\left(X_t - X_{\pi_{i-1}(t)}\right)$ is a finite maximum $|\mathcal{T}_i|$ of $\left(2^{1-i}D\right)^2$ -sub-Gaussian variables。

- Bound $\mathbb{E}\left\{\max_{t\in\mathcal{T}_k}\left(X_t - X_{t_0}\right)\right\}$ for $t \in \mathcal{T}_k$:
  - $\mathbb{E}\left\{\max_{t\in T_i}\left(X_t - X_{\pi i-1t)}\right)\right\} \leq \sqrt{2\cdot 2^{2(1-i)}D^2\log|\mathcal{T}_i|}$
  - Then

$$\begin{aligned}\mathbb{E}\left\{\max_{t\in\mathcal{T}_k}\left(X_t - X_{t_0}\right)\right\} &\leq \sum_{i=1}^{k}\mathbb{E}\left\{\max_{t\in\mathcal{T}_i}\left(X_t - X_{\pi_{i-1}(t)}\right)\right\} \\ &\leq \sum_{i=1}^{k}\sqrt{2\cdot 2^{2(1-i)}D^2\log|\mathcal{T}_i|} \\ &= \sum_{i=1}^{k}2\sqrt{2}D2^{-i}\sqrt{\log N\left(D2^{-i},\mathcal{T},\rho\right)} \\ &\leq 4\sqrt{2}\sum_{i=1}^{k}\int_{D2^{-(i+1)}}^{D2^{-i}}\sqrt{\log N(u,\mathcal{T},\rho)}\mathrm{d}u \\ &= 4\sqrt{2}\int_{D2^{-(k+1)}}^{D/2}\sqrt{\log N(u,\mathcal{T},\rho)}\mathrm{d}u\end{aligned}$$

- Bound $\mathbb{E}\left(\sup_t X_t\right)$ for $t \in \mathcal{T}$:

$$\begin{aligned}\mathbb{E}\left(\sup_t X_t\right) &= \mathbb{E}\left\{\lim_k \sup_{t\in\mathcal{T}_k}\left(X_t - X_{t_0}\right) + X_{t_0}\right\} \\ &\leq \liminf_k \mathbb{E}\left\{\sup_{t\in\mathcal{T}_k}\left(X_t - X_{t_0}\right)\right\} \\ &\leq 4\sqrt{2}\int_0^{D/2}\sqrt{\log N(u,\mathcal{T},\rho)}\mathrm{d}u\end{aligned}$$

Reference

- Integral Entropy