MATH888: High-dimensional probability and statistics — Fall 2021

Lecture 18 — October 18, 2021

Sebastien Roch, UW-Madison — *Scribe: Cheuk Yin (Eric) Lin, Sebastien Roch*

1 Suprema of Processes : ϵ-net

Let us first recall the definition of an ϵ-net of K, then we introduce a new quantity called the covering number of a set K. See Figure 3 for a visual illustration.

Definition 1 (ϵ-net, Definition 4.2.1 in [3])**.** *Let (T, d) be a metric space, $\epsilon > 0$ and a set $K \subseteq T$. A subset $N \subseteq K$ is an ϵ-net of K if for any $\boldsymbol{x} \in K$, there exists a $\boldsymbol{x}_0 \in N$ such that $d(\boldsymbol{x}, \boldsymbol{x}_0) \leq \epsilon$. Equivalently, N is an ϵ-net of K if and only if K can be covered by balls with centers in N and radii ϵ.*

Definition 2 (Covering number, Definition 4.2.2 in [3])**.** *The smallest cardinality of an ϵ-net of K is called the covering number and is denoted $\mathcal{N}(K, \epsilon)$. Equivalently, $\mathcal{N}(K, \epsilon)$ is the smallest number of closed balls with centers in K and radii ϵ whose union covers K.*

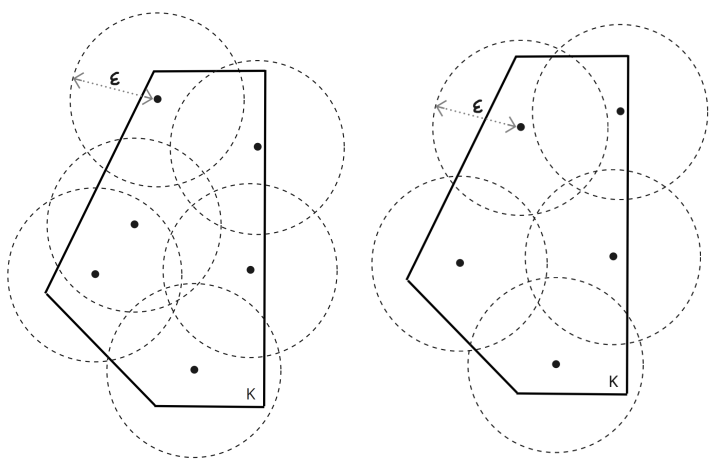

Figure 1: On the left figure, the pentagon K is covered by an ϵ-net of 6 balls. On the right figure, the pentagon K is covered by another ϵ-net of 5 and is the smallest cardinality of any ϵ-net of K, hence the covering number of K is 5.

Remark 3. *The closure of K is a compact set if and only if $\mathcal{N}(K, \epsilon)$ is finite for every $\epsilon > 0$.*

Definition 4. *The process $(X_t)_{t\in T}$ is Lipschitz if there exists a random variable L such that $|X_t - X_s| \leq L d(s,t)$ for all $s, t \in T$ almost surely.*

Let us introduce a theorem which bounds the suprema of a random process given some niceness.

Theorem 5. *Suppose that a random process $(X_t)_{t\in T}$ is Lipschitz, mean zero, and that $\|X_t\|_{\psi_2} \leq \sigma$ for all $t \in T$. Then*

$$\mathbb{E}\left(\sup_{t\in T} X_t\right) \leq \inf_{\epsilon>0}\left\{\epsilon\mathbb{E}[L] + \sqrt{C\sigma^2 \log \mathcal{N}(T, \epsilon)}\right\}.$$

Proof. For $t \in T$, let $\pi(t)$ be the closet point in an ϵ-net of T of smallest cardinality. Then we have

$$\begin{aligned}\sup_{t \in T} X_t &= \sup_{t \in T} \left(X_t - X_{\pi(t)} + X_{\pi(t)} \right) \\ &\le \sup_{t \in T} \left(X_t - X_{\pi(t)} \right) + \sup_{t \in T} X_{\pi(t)} \\ &\le \epsilon L + \sup_{s \in N} X_s \quad \text{by } L\text{-Lipschitz.}\end{aligned}$$

Taking expectations on both sides, we get

$$\mathbb{E}\left(\sup_{t \in T} X_t \right) \le \epsilon \mathbb{E}(L) + \sup_{s \in N} \mathbb{E}(X_s).$$

Result follows immediately by applying the bound for $\sup_{s \in N} \mathbb{E}(X_s)$ as shown in the previous lecture and taking the infimum of the R.H.S. over $\epsilon > 0$. □

Definition 6 (ϵ-separated and packing number, Definition 4.2.4 in [3]). *A subset $N \subseteq K \subseteq T$ is ϵ-separated if $d(\boldsymbol{x}, \boldsymbol{y}) > \epsilon$ for all $\boldsymbol{x}, \boldsymbol{y} \in N$. The largest cardinality of an ϵ-separated set of K is called the packing number and is denoted $\mathcal{P}(K, \epsilon)$.*

Lemma 7 (Nets from separated sets, Lemma 4.2.6 in [3]). *Let N be a maximal ϵ-separated subset of K. Then N is an ϵ-net of K.*

Proof. Suppose $\boldsymbol{x} \notin N$, then $N \cup \{\boldsymbol{x}\}$ is not ϵ-separated by definition, hence a contradiction. □

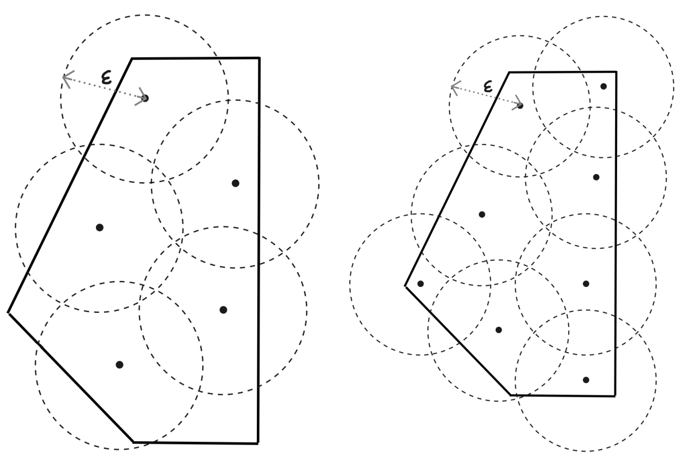

Figure 2: On the left figure, let N be the set of the all dotted points and by definition N is ϵ-separated. On the right figure, N is a maximal ϵ-separated subset of K and thus is an ϵ-net of K.

Lemma 8 (Equivalence of covering and packing number, Lemma 4.2.8 in [3]). *For any set $K \subset T$ and for any $\epsilon > 0$, we have $\mathcal{P}(K, 2\epsilon) \le \mathcal{N}(K, \epsilon) \le \mathcal{P}(K, \epsilon)$.*

Proof. The upper bound of $\mathcal{N}(K, \epsilon)$ follows directly from Lemma 7. To prove the lower bound, choose an 2ϵ-separated subset $\mathcal{P} = \{\boldsymbol{x}_i\}$ in K and an ϵ-net $\mathcal{N} = \{\boldsymbol{y}_i\}$ of K. By the definition of a net, each point $\boldsymbol{x}_i$ belongs to a closed ϵ-ball centered at some point $\boldsymbol{y}_i$. Moreover, since any closed ϵ-ball cannot contain a pair of 2ϵ-separated points, each ϵ-ball centered at $\boldsymbol{y}_i$ may contain at most one point $\boldsymbol{x}_i$. We can then conclude that $|\mathcal{P}| \le |\mathcal{N}|$ by the pigeonhole principle. □

Definition 9 (Minkowski sum)**.** *The Minkowski sum of sets A and $B \subseteq \mathbb{R}^n$ is*

$$A + B = \{a + b \mid a \in A, b \in B\}.$$

Lemma 10 (Covering number and volume, Lemma 4.2.12 in [3])**.** *Let K be a subset of $\mathbb{R}^n$ and $\epsilon > 0$. Then*

$$\frac{|K|}{|\epsilon B_2^n|} \le \mathcal{N}(K, \epsilon) \le \mathcal{P}(K, \epsilon) \le \frac{|K + (\epsilon/2)B_2^n)|}{|(\epsilon/2)B_2^n|},$$

where $|\cdot|$ denotes the volume in $\mathbb{R}^n$, B_2^n denotes the unit Euclidean ball in $\mathbb{R}^n$, so $\epsilon\mathbb{R}^n$ is an Euclidean ball with radius ϵ.

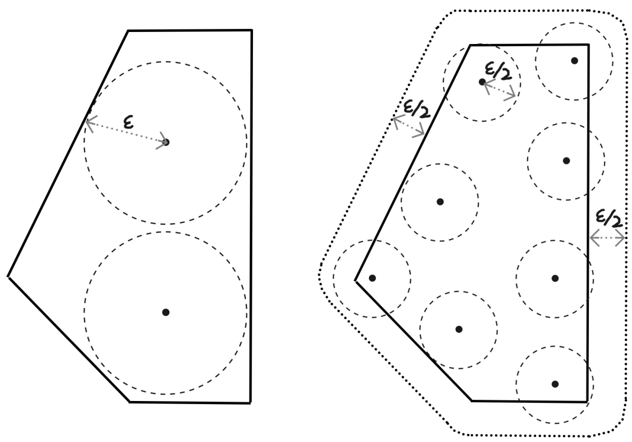

Figure 3: On the left figure, we can see that the covering number must be lower bounded by the number of ϵ-balls that can be fitted into K. On the right figure, we can see that the packing number must be upper bounded by the number of $(\epsilon/2)$-ball that can be fitted into the $(\epsilon/2)$-padded K.

References

[1] Rick Durrett, *Probability—theory and examples (fifth edition)*, Cambridge University Press, 2019.

[2] Ramon van Handel, *APC 550: Probability in High Dimension*, Lecture Notes, 2016. `https://web.math.princeton.edu/~rvan/APC550.pdf`

[3] Roman Vershynin, *High-dimensional probability: An introduction with applications in data science*, Cambridge University Press, 2018.