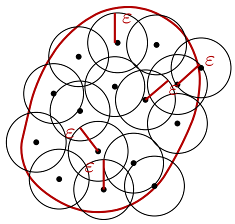
An ε-cover is a set of points that approximates every point in a given set up to an error ε, where error is measured by a metric or norm.



Let (X, d) be a metric space and let $S \subseteq X$. A set $\mathcal{N} \subseteq X$ is called an ε-cover of S with respect to d if for every $x \in S$, there exists some $\bar{x} \in \mathcal{N}$ such that $d(x, \bar{x}) \leq \varepsilon$. Equivalently,

$$S \subseteq \bigcup_{x \in \mathcal{N}} B_d(x, \varepsilon),$$

where

$$B_d(x, \varepsilon) := \{\bar{x} \in X : d(\bar{x}, x) \leq \varepsilon\}.$$

Remark: If $\mathcal{N} \subseteq S$, it is called internal ε-cover.

Definition: The covering number is the smallest possible size of an ε-cover, i.e.,

$$N(S, d, \varepsilon) = \min \left[|C| : C \subseteq X,\ S \subseteq \bigcup_{x \in C} B_d(x, \varepsilon) \right].$$

In statistical learning theory, one often covers a function class $\mathcal{F}$. Suppose $\mathcal{F} = \{f : X \to \mathbb{R}\}$. To define an ε-cover, we need a metric on functions. For example, under the uniform metric,

$$d_\infty(f, g) = \sup_{x \in X} |f(x) - g(x)|.$$

Then $\mathcal{N} \subseteq \mathcal{F}$ is an ε-cover of $\mathcal{F}$ if

$$\forall f \in \mathcal{F} \quad \exists g \in \mathcal{N} \quad : \quad \sup_{x \in X} |f(x) - g(x)| \leq \varepsilon.$$

That means every function in $\mathcal{F}$ is uniformly approximated by some function in the finite set $\mathcal{N}$.