MATH524 – Spring 2025

Problem Set: Week 12

1. (**Covering numbers for sums of functions**) Let $\mathcal{F}$ and $\mathcal{G}$ be two function classes on $\mathbb{R}^m$ Define $\mathcal{F} \oplus \mathcal{G} = \{f + g : f \in \mathcal{F}, g \in \mathcal{G}\}$. Then, for any data, z_1^n, and $\varepsilon, \delta > 0$, we have

$$N(\varepsilon + \delta, \mathcal{F} \oplus \mathcal{G}, L_1(z_1^n)) \leq N(\varepsilon, \mathcal{F}, L_1(z_1^n)) N(\delta, \mathcal{G}, L_1(z_1^n))$$

   Solution: Let $\{f_1, \dots, f_K\}$ and $\{g_1, \dots, g_L\}$ be and ε and δ cover of $\mathcal{F}$ and $\mathcal{G}$ respectively. For every $f \in \mathcal{F}$ and $g \in \mathcal{G}$, there exists a $k \in [K]$ and $l \in [L]$ such that

$$\frac{1}{n} \sum_{i=1}^{n} |f(z_i) - f_k(z_i)| < \varepsilon$$
$$\frac{1}{n} \sum_{i=1}^{n} |g(z_i) - g_l(z_i)| < \delta.$$

   By the triangle inequality,

$$\begin{aligned}
&\frac{1}{n} \sum_{i=1}^{n} |f(z_i) + g(z_i) - (f_k(z_i) + g_l(z_i))| \\
&\leq \frac{1}{n} \sum_{i=1}^{n} |f(z_i) - f_k(z_i)| + \frac{1}{n} \sum_{i=1}^{n} |g(z_i) - g_l(z_i)| \leq \varepsilon + \delta.
\end{aligned}$$

   Thus, $\{f_k + g_l : k \in [K], l \in [L]\}$ is an $\varepsilon + \delta$ cover of $\mathcal{F} \oplus \mathcal{G}$ on z_1^n.

2. (**KL divergence chain rule**) Let $p(x, y)$ and $q(x, y)$ be two distributions for a pair of random variables x, y. Show the following statement holds:

$$\mathrm{KL}(p(x, y), q(x, y)) = \mathrm{KL}(p(x), q(x)) + \mathrm{KL}(p(y|x), q(y|x)),$$

   where $p(x)$ and $q(x)$ represent the marginal distributions for x and $p(y|x)$ and $q(y|x)$ represent the conditional distributions of y given x.

   Solution: We start by plugging in for the definition of KL divergence for the LHS of the equation above:

$$\begin{aligned}
\mathrm{KL}(p(x, y), q(x, y)) &= \sum_{x,y} p(x, y) \log \frac{p(x, y)}{q(x, y)} \\
&= \sum_{x,y} p(x)p(y|x) \log \frac{p(x)p(y|x)}{q(x)q(y|x)} \\
&= \sum_{x} p(x) \log \frac{p(x)}{q(x)} \sum_{y} p(y|x) + \sum_{x} p(x) \sum_{y} p(y|x) \log \frac{p(y|x)}{q(y|x)} \\
&= \mathrm{KL}(p(x), q(x)) + \sum_{x} p(x) \mathrm{KL}(p(y|x = x), q(y|x = x)) \\
&= \mathrm{KL}(p(x), q(x)) + \mathrm{KL}(p(y|x), q(y|x)).
\end{aligned}$$