

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

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

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:

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

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 .