# MATH524 – Spring 2025

Problem Set: Week 7

1. (**Mean-absolute deviation forms GC class**) Consider the M-estimation example from Chapter 5 with the meand absolute deviation estimator:

$$m_\theta(x) = |x - \theta|$$

Assume further that $\mathbb{P}(x^2) < \infty$, $\Theta = [-1, 1]$, $d(x, y) = |x - y|$ and define $\mathcal{F} = \{m_\theta : \theta \in [-1, 1]\}$. In this exercise we will show that $\mathcal{F}$ forms a GC class.

(a) Show that $m_\theta(x)$ is Lipschitz and identify the lipschitz constant.

**Solution**: The lipschitz property follows directly from the definition:

$$|m_\theta(x) - m_{\theta'}(x)| \leq |\theta - \theta'|.$$

The lipschitz constant is 1.

(b) By showing that the $L_1$ bracketing number can be bounded as

$$N_{[]}(\varepsilon, \mathcal{F}, L_1) < 2\frac{\text{diam}(\Theta)}{\varepsilon},$$

show that $\mathcal{F}$ forms a GC class.

***Hint:*** *It may be helpful to identify a relationship between the bracketing and covering numbers for* $\Theta$.

**Solution**: We will first show that the bracketing number is bounded as

$$N_{[]}(\varepsilon, \mathcal{F}, L_p) < 2\frac{\text{diam}(\Theta)}{\varepsilon}$$

for all $p \geq 1$, where $\text{diam}(\cdot)$ is the diameter of a set.

Let $\Theta_\varepsilon$ be the finite $\varepsilon$-cover of $\Theta$ with distance function $d(x, y) = |x - y|$. Define the collection of brackets $\{[l_j, u_j]\}_{j=1}^K$ as

$$\begin{aligned} l_j &= m_{\theta_j} - \varepsilon \\ u_j &= m_{\theta_j} + \varepsilon \end{aligned}$$

for each $\theta_j \in \Theta_\varepsilon$. For any $m_\theta \in \mathcal{F}$, and some $\theta_j \in \Theta_\varepsilon$

$$|m_\theta - m_{\theta_j}| < \varepsilon,$$

by the observation that $m_\theta$ is 1-lipschitz and selecting the $\theta_j$ that is the element of $\Theta_\varepsilon$ that is $\varepsilon$-close to $\theta$. Then,

$$\mathbb{E}[(u_j - l_j)^2]^{1/2} \leq \mathbb{E}[(2\varepsilon)^2]^{1/2} = 2\varepsilon < \infty.$$

Now, we notice that for each element in $\Theta_\varepsilon$, we have at most 2 elements in the collection of $\varepsilon$-brackets of $\Theta$. Thus,

$$\begin{aligned} N_{[]}(\varepsilon, \Theta, L_p) &< 2N(\varepsilon, \Theta, L_p) \\ &< 2\frac{\text{Vol}(\Theta)}{\varepsilon} \\ &< 2\frac{\text{diam}(\Theta)}{\varepsilon}, \end{aligned}$$

by the fact that $\Theta$ is compact in $\mathbb{R}$.

Therefore, the $L_1$ bracketing number is finite, for every $\varepsilon > 0$ and so $\mathcal{F}$ is GC.

2. (**KL for Gaussians**) If $P = N(\theta, \sigma^2)$ and $Q = N(\mu, \sigma^2)$, show that

$$\mathrm{KL}(P, Q) = \frac{(\theta-\mu)^2}{2\sigma^2}.$$

**Solution**: Plugging into the definition of the KL divergence and simplifying the normal density,

$$\begin{aligned} KL(N(\theta,\sigma^2), N(\mu,\sigma^2)) &= \frac{1}{\sigma}\int \phi(\frac{x-\theta}{\sigma}) \log\left(\frac{\phi(\frac{x-\theta}{\sigma})}{\phi(\frac{x-\mu}{\sigma})}\right) \mathrm{d}x \\ &= \frac{\theta-\mu}{2\sigma^3}\int \phi(\frac{x-\theta}{\sigma})(2x-\theta-\mu)\,\mathrm{d}x = \frac{(\theta-\mu)^2}{2\sigma^2}. \end{aligned}$$

3. (**Scheffe's Theorem**) Define the total variation distance function as

$$TV(P,Q) = \sup_{A\in\mathcal{A}} |\textstyle\int_A (p-q)\,\mathrm{d}\nu|$$

Show the following equality between total variation and affinity

$$TV(P,Q) = \frac{1}{2}\int |p-q|\,\mathrm{d}\nu = 1 - \int \min(p,q)\,\mathrm{d}\nu.$$

***Hint:*** *Consider bounding the integral on sets of type* $\{x \in \mathcal{X} : q(x) \geq p(x)\}$

**Solution**: Observe that for the set $A_0 = \{x \in \mathcal{X} : q(x) \geq p(x)\}$

$$\int |p(\nu) - q(\nu)|\,\mathrm{d}\nu = 2\int_{A_0} (q-p)\,\mathrm{d}\nu.$$

Thus,

$$TV(P,Q) \geq Q(A_0) - P(A_0) = \frac{1}{2}\int |p-q|\,\mathrm{d}\nu = 1 - \int \min(p,q)\,\mathrm{d}\nu.$$

For the reverse inequality, notice that for all $A \in \mathcal{A}$,

$$\begin{aligned} |\int_A (q-p)\,\mathrm{d}\nu| &= |\int_{A\cap A_0} (q-p)\,\mathrm{d}\nu + \int_{A\cap A_0^c} (q-p)\,\mathrm{d}\nu| \\ &\leq \max\left\{\int_{A_0} (q-p)\,\mathrm{d}\nu, \int_{A_0^c} (p-q)\,\mathrm{d}\nu\right\} \\ &= \frac{1}{2}\int |p-q|\,\mathrm{d}\nu. \end{aligned}$$