MATH524 – Spring 2025

Problem Set: Week 9

1. Let $X_1, \dots, X_n$ be an i.i.d. sample from $\mathcal{N}(\mu, \sigma^2)$ where σ is a known constant. Prove using Le Cam's two-points lemma that
$$\sup_{\mu \in \mathbb{R}} \mathbb{E}|\tilde{\mu} - \mu| \geq \frac{C}{\sqrt{n}},$$
for any estimator $\tilde{\mu}$ with some constant C.

   Solution: Let $d(\mu_1, \mu_2) := |\mu_1 - \mu_2|$, and note that this is a metric, so we can choose $A = 1$. Set $\mu_1 = 0$. Then $d(\mu_1, \mu_2) = |\mu_2|$ and
$$KL(\mathcal{N}(\mu_1, \sigma^2)^n, \mathcal{N}(\mu_2, \sigma^2)^n) = nKL(\mathcal{N}(\mu_1, \sigma^2), \mathcal{N}(\mu_2, \sigma^2)) = n\frac{\mu_2^2}{2\sigma^2}.$$
   To make this equal 1/4 choose $\mu_2 = \sigma/\sqrt{2n}$. By Le Cam's two points lemma, it follows that
$$\sup_{\mu \in \mathbb{R}} \mathbb{E}|\tilde{\mu} - \mu| \geq \frac{\sigma}{4\sqrt{2n}}.$$

2. (**Sparse mean-vector estimation**) In this exercise we will use a modified version of the Varshamov-Gilbert construction from the lecture notes to help identify minimax lower-bounds for sparse vector estimation.

   Lemma 1 (Sparse Varshamov-Gilbert)
   Conisder $\Omega = \{\omega \in \{0,1\}^d : \|\omega\|_0 \leq s\}$. Then, there exists $\Omega' \subseteq \Omega$ such that

   - $\|\omega\|_0 = s$ *for every* $\omega \in \Omega'$
   - $H(\omega_i, \omega_j) \geq s/2$
   - $|\Omega'| \geq c(de/s)^s$.

   Now, consider $X \sim \mathcal{N}(\theta, \sigma^2 I_d)$ with $\|\theta\|_0 \leq s$ (i.e. the mean vector is sparse). The following upper bound is well-known:
$$\mathbb{E}[\|\hat{\theta} - \theta\|^2] \precsim \frac{\sigma^2 s \log d}{n}$$
   where
$$\hat{\theta} = \begin{cases} \bar{X}_n & \text{if } \bar{X}_n \gtrsim \sigma\sqrt{\frac{\log d}{n}} \\ 0 & \text{otherwise} \end{cases}$$
   Using Lemma 1 to construct a reasonable packing set and Fano's inequality, show that a matching lower bound (up to log factors) can be established.

   Solution: To use Lemma 1, we first construct a packing set of size $N = (de/s)^s$ of s-sparse vectors that have Hamming distance of at least $s/2$.

   Start with $\theta_i = \omega_i \theta_{\min}$ where $\theta_{\min} = \min_{i:\theta_i \neq 0} |\theta_i|$. Then consider the collection of distributions $P_i = \mathbb{N}(\theta_i, \sigma^2 I_d)$. We know from previous exercises that the KL distance of any two such distributions is given by
$$KL(P_i, P_j) = \frac{n}{2\sigma^2}\|\theta_i - \theta_j\|_2^2 \leq \frac{ns\theta_{\min}^2}{\sigma^2}.$$

Then, by Lemma 1, we know that $\log |\Omega'| \gtrsim s \log(de/s)$. We can apply Fano if $KL(P_i, P_j) \precsim s \log(de/s)$, which means

$$\frac{ns\theta_{\min}^2}{\sigma^2} \precsim s \log \frac{de}{s} \Leftrightarrow \theta_{\min} \precsim \sigma \sqrt{\frac{\log(de/s)}{n}}.$$

If $\theta_{\min}$ satisfies this condition, then

$$\|\theta_i - \theta_j\|_2^2 \geq H(\omega_i, \omega_j)\theta_{\min}^2 \geq \frac{s}{2}\theta_{\min}^2 \geq \sigma^2 \frac{s \log(de/s)}{2n}.$$

This matches the upper bound up to log-factors.