18.466 Midterm 1.

Let us consider the Normal Location problem. Suppose we observe sample X from n-dimensional multivariate normal distribution, i.e. $X \sim N(\mu, I_n)$, where $\mu \in R^n$ is the unknown mean vector and I_n is the identity matrix. Our goal is to estimate μ under quadratic loss

$$l(\mu, d) = \|\mu - d\|_2^2 = \sum_{i=1}^{n} (\mu_i - d_i)^2$$

1. Compute the risk function $R(\mu, \hat{\mu}_1)$ of the ordinary estimator $\hat{\mu}_1(X) = X$.

2. Suppose $\hat{\mu}_\pi$ is the Bayes estimator under some prior distribution π. Let $r(\hat{\mu}_\pi) = E_\pi R(\mu, \hat{\mu}_\pi)$ be the Bayes risk. Show that if there is another estimator $\hat{\mu}$ such that $\max\{\mu \in R^n : R(\mu, \hat{\mu})\} = r(\hat{\mu}_\pi)$, then $\hat{\mu}$ is a minimax estimator.

3. Moreover, suppose there is a sequence of prior $\pi_i, i = 1, 2, \cdots$ and the corresponding Bayes estimator $\hat{\mu}_{\pi_i}$. Show that if there is another estimator $\hat{\mu}$ such that $\max\{\mu \in R^n : R(\mu, \hat{\mu})\} = \lim_{i \to \infty} r(\hat{\mu}_{\pi_i})$, then $\hat{\mu}$ is a minimax estimator.

4. Show that the ordinary estimator $\hat{\mu}_1(X) = X$ is minimax.

5. Suppose $g : R^n \to R$ is a differentiable function. Show that

$$E((X_i - \mu_i) g(X)) = E(\frac{\partial}{\partial x_i} g(X)),$$

So long as the expectations exist.

6. Suppose $\lambda = (\lambda_1, \lambda_2, \cdots, \lambda_n) : R^n \to R^n$ such that each coordinate function of λ is differentiable. Show that

$$R(\mu, X + \lambda(X)) = n + E_\mu(2\nabla \cdot \lambda(X) + \|\lambda(X)\|_2^2),$$

where $\nabla \cdot \lambda(x) = \sum_{i=1}^{n} \frac{\partial}{\partial x_i} \lambda_i(x)$.

7. Suppose n>2. For any $C \in (0, 2(n-2))$, define estimator $\hat{\mu}_{JS} = (1 - \frac{C}{\|X\|_2^2})X$. Sow that $\hat{\mu}_{JS}$ is minimax and moreover, $R(\mu, \hat{\mu}_{JS}) < R(\mu, \hat{\mu}_1)$ for all μ. ($\hat{\mu}_1 = X$.)

8*. Suppose n>3, for any $C \in (0, 2(n-3))$, define estimator

$$\hat{\mu}_g = \overline{X} 1_n + (1 - \frac{C}{\|X - \overline{X} 1_n\|_2^2})(X - \overline{X} 1_n),$$

where $\overline{X}$ is the sample average and $1_n \in R^n$ is the vector whose coordinates are all 1. Show that $R(\mu, \hat{\mu}_g) < R(\mu, \hat{\mu}_1)$ for all μ.

9* Suppose n>2. For any $C \in (0, 2(n-2))$, define estimator positive part J-S estimator $\hat{\mu}_{JS}^{+} = (1 - \frac{C}{\|X\|_2^2})_{+} X$. Sow that $\hat{\mu}_{JS}$ is minimax and moreover, $R(\mu, \hat{\mu}_{JS}^{+}) < R(\mu, \hat{\mu}_{JS})$ for all μ. (Here $(1 - \frac{C}{\|X\|_2^2})_{+} = 1 - \frac{C}{\|X\|_2^2}$ when it is positive and zero otherwise.)