Tsinghua-Berkeley Shenzhen Institute
Learning from Data
Fall 2018

Problem Set 5

Issued: Tuesday 18th December, 2018 **Due:** Monday 24th December, 2018

5.1. The covariance matrix of a random vector $X \in \mathbb{R}^d$ is defined as

$$\mathrm{Cov}(X) \triangleq \mathbb{E}[(X - \mathbb{E}[X])(X - \mathbb{E}[X])^{\mathrm{T}}] \in \mathbb{R}^{d \times d},$$

where $\mathbb{E}[\cdot]$ is the mathematical expectation. Given m i.i.d. random variables $X_1, \ldots, X_m$ with the same distribution as X, show that

$$\hat{C} \triangleq \frac{1}{m-1} \sum_{i=1}^{m} (X_i - \hat{\mu})(X_i - \hat{\mu})^{\mathrm{T}}$$

is an unbiased estimation of $\mathrm{Cov}(X)$, i.e., $\mathbb{E}[\hat{C}] = \mathrm{Cov}(X)$, where

$$\hat{\mu} \triangleq \frac{1}{m} \sum_{i=1}^{m} X_i.$$

5.2. Suppose X is a random variable taking values from a finite set $\mathcal{X}$, and there are two hypotheses H_0 and H_1

$$H = H_0 : X \sim P_X,$$
$$H = H_1 : X \sim Q_X,$$

with prior probabilities $\mathbb{P}(H = H_i) = p_i, i = 0, 1$.

Given m independent observations of X: $x_1, \ldots, x_m$, decide the decision rule $\hat{H}(x_1, \ldots, x_m)$, such that the error probability $P_e \triangleq \mathbb{P}(\hat{H} \neq H)$ is minimized.

5.3. Suppose $X_1, \cdots, X_m \in \mathbb{R}^d$ are m independent random vectors with $\mathbb{E}[X_i] = \mu_i, \mathrm{Cov}(X_i) = \Sigma_i (i = 1, \cdots, m)$, Y is a categorical distribution taking values from $\{1, \ldots, m\}$ with $\mathbb{P}(Y = i) = p_i$, and Y is independent of all X_i's. Evaluate $\mathbb{E}[Z]$ and $\mathrm{Cov}(Z)$, where

$$Z \triangleq \sum_{i=1}^{m} \mathbb{1}\{Y = i\} \cdot X_i.$$