Homework #2

September 29, 2017

This is due at the beginning of class on **Tuesday, October 10**. Working in groups is fine but the write-up must be entirely your own and you should *list all collaborators* on the cover sheet of your submission.

1 Estimators

1. Recall from lecture we had the *expected (squared) prediction error*

$$EPE(f) = \mathbb{E}[(Y - f(X))^2]$$

for decision function (or *hypothesis*) $f(x)$. This is also called the MSE, mean-squared error, and is used to estimate how well a hypothesis will do on new data. Suppose we have data $D = \{(\mathbf{x}^{(1)}, y^{(1)}), ..., (\mathbf{x}^{(m)}, y^{(m)})\}$ which we put into a *design matrix* $X = \begin{bmatrix} \mathbf{x}^{(1)} & \cdots & \mathbf{x}^{(m)} \end{bmatrix}$. Suppose we assume a linear relationship between the *features* $\mathbf{x}^{(j)}$ and *targets* $Y = \begin{bmatrix} y^{(1)} \\ \vdots \\ y^{(m)} \end{bmatrix}$ i.e.

$$y^{(i)} = f(\mathbf{x}^{(i)}) = \mathbf{x}^{(i)T}\boldsymbol{\beta}$$

Prove that $\boldsymbol{\beta}^* = \arg\min_{\boldsymbol{\beta}} EPE(f) = \arg\min_{\boldsymbol{\beta}} \mathbb{E}[\|Y - X^T\boldsymbol{\beta}\|^2]$ is given by

$$\boldsymbol{\beta}^* = \mathbb{E}[XX^T]^{-1}\mathbb{E}[XY]$$

2. Prove that if an MVU estimator exists then it is unique. *Hint: Consider* $\frac{1}{2}(d_1 \pm d_2)$ *for MVUs* d_1, d_2.

3. Given an iid sample $\{X_i\} \sim \mathcal{N}(0, \theta)$ show that $\frac{1}{n}\sum_i X_i^2$ is an unbiased estimator of $\theta = \sigma^2$.

4. Consider N iid observations of a $\mathcal{U}([\theta, \theta + 1])$ variable. Find a sufficient statistic for the data set. *Remark: your answer should be a pair even though there is only a single parameter to estimate.*

5. Consider N iid observations of a $\mathcal{U}([a, b])$ variable. Find a sufficient statistic for the data set.

6. Consider N iid observations of an $Exp(\theta)$ variable. Find a sufficient statistic for the data set.
7. Prove that invertible functions of sufficient statistics are sufficient statistics. Give an example of a situation in which this kind of result might be useful.
8. For $\{X_j\} \sim \mathcal{U}(0, \theta)$ find the MLE of θ.
9. Give a sufficient condition on the log-likelihood function which ensures that the MLE yields a unique global maximum.
10. For $\{X_j\} \sim \mathcal{N}(\mu, \sigma^2)$ verify that $(\sum_j X_j, \sum_j X_j^2)$ is a complete and sufficient statistic for (μ, σ).
11. Let f, g be two pdfs. Use Jensen's inequality to show that $D_{KL}(f \parallel g) \geq 0$ with equality $f - a.e.$

2 Programming

1. This is to be done in the Python programming language. Print and attach both your sample code and plots. Select a stock of your choice and download a list of at least 100 quotes. Transform the quotes into returns via

$$\mu_{t_i} = \log(\frac{S_{t_{i+1}}}{S_{t_i}}), \qquad i = 1, ..., N \geq 100$$

Normalize your sample by subtracting off the sample mean and dividing by the sample variance. Create a histogram of your normalized data set and overlay a standard normal distribution curve. As well create a Q-Q plot comparing the quantiles of the distribution of returns versus a standard normal (see this link for details on quantile plots if you've not seen them before). Comment on the types of differences you see between your histogram and the standard normal.

Next, two higher-order moments are skewness and kurtosis, given by

$$skew[X] = \mathbb{E}[(\frac{X-\mu}{\sigma})^3], \quad kurt[X] = \mathbb{E}[(\frac{X-\mu}{\sigma})^4]$$

which we may estimate using

$$\widehat{skew}[X] = \frac{\frac{1}{n}\sum_j (x_j - \bar{x})^3}{(\sqrt{\frac{1}{n-1}\sum_j (x_j - \bar{x})^2})^3}, \quad \widehat{kurt}[X] = \frac{\frac{1}{n}\sum_j (x_j - \bar{x})^4}{(\sqrt{\frac{1}{n-1}\sum_j (x_j - \bar{x})^2})^4}$$

Compare also the higher-order sample moments. Compare the higher-order statistics from your data sample with a standard Gaussian and comment on differences/similarities.