STAT 510 Homework 1

Due Date: 11:00 A.M., Wednesday, January 22

1. Consider the matrix

$$\boldsymbol{X} = \begin{bmatrix} 1 & -3 & 0 & -3 \\ 1 & -2 & -1 & 2 \\ 2 & -5 & -1 & -1 \end{bmatrix}.$$

   (a) Show that the columns of $\boldsymbol{X}$ are linearly dependent.

   (b) Find the rank of $\boldsymbol{X}$.

   (c) Use the generalized inverse algorithm in slide set 1 to find a generalized inverse of $\boldsymbol{X}$.

   (d) Use the `R` function `ginv` in the `MASS` package to find a generalized inverse of $\boldsymbol{X}$. (To load the `MASS` package into your `R` workspace use the command `library(MASS)`. If the `MASS` package is not already installed, you will need to install it before loading. The command `install.packages(``MASS'')` should install the `MASS` package if necessary.)

   (e) Provide one matrix $\boldsymbol{X}^*$ that satisfies both of the following characteristics:

   - $\boldsymbol{X}^*$ has full-column rank (i.e., $\text{rank}(\boldsymbol{X}^*)$ is equal to the number of columns of $\boldsymbol{X}^*$), and
   - $\boldsymbol{X}^*$ has column space equal to the column space of $\boldsymbol{X}$; i.e., $\mathcal{C}(\boldsymbol{X}^*) = \mathcal{C}(\boldsymbol{X})$.

   (f) Describe a property that all the three-dimensional vectors in the column space of $\boldsymbol{X}$ share?

2. Prove that any list of n-dimensional vectors that contains the n-dimensional zero vector is linearly dependent.

3. Imagine extending a string from $(0, 0)$, the origin in $\mathbb{R}^2$, to a random point (x, y) in $\mathbb{R}^2$, where $x \sim N(4, 1)$ independent of $y \sim N(1, 1)$. Use `R` to find the probability that the string will need to be longer than 5 units to reach from $(0, 0)$ to (x, y). *Hint: Simulation is one way to solve the problem, but there is a better way.*

4. Suppose $z_1, z_2, z_3 \overset{iid}{\sim} N(0, 1)$. Find the distribution of the following random variables and prove that your answer is correct.

   (a) $(z_1 - z_2 + 2z_3)^2/6$

   (b) $(z_1 + z_2)/|z_1 - z_2|$

5. Suppose $y_1, \dots, y_n \overset{iid}{\sim} N(\mu, \sigma^2)$. Let $\boldsymbol{y} = [y_1, \dots, y_n]'$, and let $\bar{y}_{\cdot} = \frac{1}{n}\sum_{i=1}^n y_i$.

   (a) Show that $s^2 = \frac{1}{n-1}\sum_{i=1}^n (y_i - \bar{y}_{\cdot})^2$ can be written as $\boldsymbol{y}'\boldsymbol{B}\boldsymbol{y}$ for some matrix $\boldsymbol{B}$.

   (b) Prove that $(n-1)s^2/\sigma^2 \sim \chi^2_{n-1}$ using the "Important Distributional Result about Quadratic Forms" in slide set 1.

6. Prove that a matrix $\boldsymbol{A}$ is $\boldsymbol{0}$ if and only if $\boldsymbol{A}'\boldsymbol{A} = \boldsymbol{0}$. (Hint: What are the diagonal elements of $\boldsymbol{A}'\boldsymbol{A}$?)

7. Prove that $\boldsymbol{X}'\boldsymbol{X}\boldsymbol{A} = \boldsymbol{X}'\boldsymbol{X}\boldsymbol{B}$ if and only if $\boldsymbol{X}\boldsymbol{A} = \boldsymbol{X}\boldsymbol{B}$. Note that the "if" part of the proof, i.e.,

$$\boldsymbol{X}\boldsymbol{A} = \boldsymbol{X}\boldsymbol{B} \implies \boldsymbol{X}'\boldsymbol{X}\boldsymbol{A} = \boldsymbol{X}'\boldsymbol{X}\boldsymbol{B},$$

   holds trivially. Thus, proving the converse, i.e.,

$$\boldsymbol{X}'\boldsymbol{X}\boldsymbol{A} = \boldsymbol{X}'\boldsymbol{X}\boldsymbol{B} \implies \boldsymbol{X}\boldsymbol{A} = \boldsymbol{X}\boldsymbol{B},$$

   is the challenging part. One proof makes use of the result in problem 6.