Stat705 PS 1. Due in class September 15th

NOTE 1: It is fine to work with other students on problem sets, and that is encouraged. Each person's write up must be done separately though, and identical answers (to relatively complicated problems) from different students will not be graded.

NOTE 2: If you use a source other than the textbook to do these problems, you must list the source. Additionally, it is never OK to copy from another source verbatim! If we notice that, the problem won't be graded.

Please read chapter 1 in Plane Answers...

1. Prove that a covariance matrix is symmetric.

2. What is a correlation matrix?

3. A matrix $\mathbf{S}$ is non-negative-definite if $\mathbf{c}^{\mathrm{T}}\mathbf{S}\mathbf{c} \geq 0$ for any vector $\mathbf{c}$. Prove that all covariance matrices are non-negative-definite. If $\mathbf{c}^{\mathrm{T}}\mathbf{S}\mathbf{c} > 0$ the matrix is positive definite. Show that a positive definite matrix is invertible. Also, give an example of a covariance matrix that is not positive definite.

4. Let $\mathbf{z}$ be a random vector of length n that has a $MVN(\mathbf{0}_n, \mathbf{I}_n)$ distribution. Let $\mathbf{y} = \mathbf{A}\mathbf{z} + \mathbf{b}$. Derive $\mathbf{y}$'s moment generating function. What is $\mathbf{y}$'s distribution?

5. Let $\mathbf{z}$ be a random vector of length n that has a $MVN(\mathbf{0}_n, \mathbf{I}_n)$ distribution. Let $U = \mathbf{z}^{\mathrm{T}}\mathbf{z}$. Show that Us moment generating function is $(1-2t)^{-n/2}$ (for $1-2t > 0$). This is the MGF for a distribution that has two names. What are they, and what are their parameters?

6. Suppose $\mathbf{y} = \begin{pmatrix} \mathbf{y}_1 \\ \mathbf{y}_2 \end{pmatrix}$ where

$$\begin{pmatrix} \mathbf{y}_1 \\ \mathbf{y}_2 \end{pmatrix} \sim N\left\{\begin{pmatrix} \mu_1 \\ \mu_2 \end{pmatrix}, \begin{pmatrix} \Sigma_{11} & \Sigma_{12} \\ \Sigma_{21} & \Sigma_{22} \end{pmatrix}\right\}.$$

   You may assume that the covariance matrix has full rank. The dimensions of the moments correspond to the moments of the components of $\mathbf{y}$. Derive the conditional distribution of $\mathbf{y}_1|\mathbf{y}_2 = a$.

7. Suppose $\mathbf{z}$ is $MVN(\mathbf{0}_n, \mathbf{I}_n)$. Find $\mathbf{a}$ and $\mathbf{A}$ so that $\mathbf{a} + \mathbf{A}\mathbf{z}$ has a $MVN(\mu, \Sigma)$ distribution. Please use the singular value decomposition to find $\mathbf{A}$.

8. What is a basis? (Please do not include undefined notation in your definition. Reading Appendix A in Plane Answers is recommended too.)

9. Let $\mathbf{X}$ be an n by $(p+1)$ matrix with $n > (p+1)$ that has $K \leq (p+1)$ non-zero singular values. Denote the SVD of $\mathbf{X}$ as $\mathbf{U}\Sigma\mathbf{V}^{\mathrm{T}}$.

   (a) The column space (aka span aka range) of a matrix is the set $\{\widehat{\mathbf{y}}$ where $\widehat{\mathbf{y}} = \mathbf{X}\boldsymbol{\beta}$ for any$\boldsymbol{\beta} \in R^{p+1}\}$. Show that any element of this set is a linear combination of the first K columns of $\mathbf{U}$. Is that a basis? Why or why not?

   (b) What is $\mathbf{X}$'s rank?

   (c) The null space of a matrix is the set $\{\boldsymbol{\alpha} \in R^{p+1}$ where $\mathbf{X}\boldsymbol{\alpha} = \mathbf{0}\}$. Show that any element of this set is a linear combination of the last $(p+1) - K$ columns of $\mathbf{V}$. Is that a basis? Why or why not?

10. From another class, you have seen that the variance of a random variable can be estimated from an *iid* sample $(y_1, \ldots, y_n)$ by $S^2 = \frac{1}{n-1}\sum_{i=1}^{n}(y_i - \overline{y})^2$. Show that $(n -$

1)$S^2 = \mathbf{y}^{\mathrm{T}}(\mathbf{I}_n - \frac{1}{n}\mathbf{J}_n)\mathbf{y}$ where $\mathbf{I}_n$ is an n by n identity matrix, and $\mathbf{J}_n$ is an n by n matrix of all 1s.