STA 302f16 Assignment Three[1]

Please bring your R printout from Question 3 to the quiz; you may be asked to hand it in, or maybe not. The other problems are preparation for the quiz in tutorial, and are not to be handed in.

This exercise set has an unusual feature. *Some of the questions ask you to prove things that are false.* That is, they are not true in general. In such cases, just write "The statement is false," and give a brief explanation to make it clear that you are not just guessing. The explanation is essential for full marks. Sometimes a small counter-example is desirable.

1. Please read pages 11-19 in the textbook. In Section 1.7, the measure of model fit R^2 is presented without much explanation. The idea is that $\sum_{i=1}^n (y_i - \overline{y})^2$ represents the sum of squared vertical distances of the points on a scatterplot from a horizontal line with slope zero and intercept $\overline{y}$. Mathematically this could be the best fitting line, but in practice the line $y = b_0 + b_1 x$ is going to do better. That is, $\sum_{i=1}^n (y_i - \widehat{y}_i)^2 = \sum_{i=1}^n e_i^2 \leq \sum_{i=1}^n (y_i - \overline{y})^2$, so that $0 \leq \frac{\sum_{i=1}^n e_i^2}{\sum_{i=1}^n (y_i - \overline{y})^2} \leq 1$. Small values of this ratio represent good performance of the least squares line relative to the horizontal line. It could be described as an index of "lack of fit," because big values indicate relatively poor performance. Thus, $R^2 = 1 - \frac{\sum_{i=1}^n e_i^2}{\sum_{i=1}^n (y_i - \overline{y})^2}$ is a measure of good fit. Why it's something *squared* will be taken up later.
   (a) Prove that the least squares line must always pass through the point $(\overline{x}, \overline{y})$, regardless of the data.
   (b) Show the work leading to (1.28). Use the formula sheet.
   (c) Do Exercise 1.2.
   (d) Do Exercise 1.4, but only the expected value, not the variance. The variance is much easier with matrices.
   (e) Do Exercise 1.5. Answer all three parts of the question.
   (f) Do both parts of Exercise 1.8. You don't have to do all the work of differentiating and solving again. Start with (1.11) and (1.14), and keep simplifying.

2. Let $\mathbf{A} = \begin{pmatrix} 1 & 2 \\ 2 & 4 \end{pmatrix}$ $\mathbf{B} = \begin{pmatrix} 0 & 2 \\ 2 & 1 \end{pmatrix}$ $\mathbf{C} = \begin{pmatrix} 2 & 0 \\ 1 & 2 \end{pmatrix}$
   (a) Calculate $\mathbf{AB}$ and $\mathbf{AC}$
   (b) Do we have $\mathbf{AB} = \mathbf{AC}$? Answer Yes or No.
   (c) Prove $\mathbf{B} = \mathbf{C}$.. Show your work.

[1]Copyright information is at the end of the last page.

3. Make up a your own 4×4 symmetric matrix that is not singular (that is, the inverse exists), and is *not a diagonal matrix*. If your first try is singular, try again. Call it $\mathbf{A}$. Enter it into R using `rbind` (see lecture slides). Make sure to display the input. Then,

   (a) Calculate $|\mathbf{A}^{-1}|$ and $1/|\mathbf{A}|$, verifying that they are equal.

   (b) Calculate $|\mathbf{A}^2|$ and $|\mathbf{A}|^2$, verifying that they are equal.

   (c) Calculate the eigenvalues and eigenvectors of $\mathbf{A}$.

   (d) Calculate $\mathbf{A}^{1/2}$.

   (e) Calculate $\mathbf{A}^{-1/2}$.

   Display both input and output for each part. Label the output with comments. Bring the printout to the quiz.

4. Recall the definition of linear independence. The columns of $\mathbf{X}$ are said to be *linearly dependent* if there exists a $p \times 1$ vector $\mathbf{v} \neq \mathbf{0}$ with $\mathbf{Xv} = \mathbf{0}$. We will say that the columns of $\mathbf{X}$ are linearly *independent* if $\mathbf{Xv} = \mathbf{0}$ implies $\mathbf{v} = \mathbf{0}$. Let $\mathbf{A}$ be a square matrix. Show that if the columns of $\mathbf{A}$ are linearly dependent, $\mathbf{A}^{-1}$ cannot exist. Hint: $\mathbf{v}$ cannot be both zero and not zero at the same time.

5. Let $\mathbf{a}$ be an $n \times 1$ matrix of real constants. How do you know $\mathbf{a}'\mathbf{a} \geq 0$?

6. Recall the *spectral decomposition* of a square symmetric matrix (For example, a variance-covariance matrix). Any such matrix $\mathbf{\Sigma}$ can be written as $\mathbf{\Sigma} = \mathbf{CDC}'$, where $\mathbf{C}$ is a matrix whose columns are the (orthonormal) eigenvectors of $\mathbf{\Sigma}$, $\mathbf{D}$ is a diagonal matrix of the corresponding eigenvalues, and $\mathbf{C}'\mathbf{C} = \mathbf{CC}' = \mathbf{I}$.

   (a) Let $\mathbf{\Sigma}$ be a square symmetric matrix with eigenvalues that are all strictly positive.

      i. What is $\mathbf{D}^{-1}$?

      ii. Show $\mathbf{\Sigma}^{-1} = \mathbf{CD}^{-1}\mathbf{C}'$

   (b) Let $\mathbf{\Sigma}$ be a square symmetric matrix, and this time some of the eigenvalues might be zero.

      i. What do you think $\mathbf{D}^{1/2}$ might be?

      ii. Define $\mathbf{\Sigma}^{1/2}$ as $\mathbf{CD}^{1/2}\mathbf{C}'$. Show $\mathbf{\Sigma}^{1/2}$ is symmetric.

      iii. Show $\mathbf{\Sigma}^{1/2}\mathbf{\Sigma}^{1/2} = \mathbf{\Sigma}$.

   (c) Now return to the situation where the eigenvalues of the square symmetric matrix $\mathbf{\Sigma}$ are all strictly positive. Define $\mathbf{\Sigma}^{-1/2}$ as $\mathbf{CD}^{-1/2}\mathbf{C}'$, where the elements of the diagonal matrix $\mathbf{D}^{-1/2}$ are the reciprocals of the corresponding elements of $\mathbf{D}^{1/2}$.

      i. Show that the inverse of $\mathbf{\Sigma}^{1/2}$ is $\mathbf{\Sigma}^{-1/2}$, justifying the notation.

      ii. Show $\mathbf{\Sigma}^{-1/2}\mathbf{\Sigma}^{-1/2} = \mathbf{\Sigma}^{-1}$.

(d) The (square) matrix $\boldsymbol{\Sigma}$ is said to be *positive definite* if $\mathbf{v}'\boldsymbol{\Sigma}\mathbf{v} > 0$ for all vectors $\mathbf{v} \neq \mathbf{0}$. Show that the eigenvalues of a positive definite matrix are all strictly positive.

(e) Let $\boldsymbol{\Sigma}$ be a symmetric, positive definite matrix. Putting together a couple of results you have proved above, establish that $\boldsymbol{\Sigma}^{-1}$ exists.

7. Prove that the diagonal elements of a positive definite matrix must be positive.

8. Using the Spectral Decomposition Theorem and $tr(\mathbf{AB}) = tr(\mathbf{BA})$, prove that the trace is the sum of the eigenvalues for a symmetric matrix.

9. Using the Spectral Decomposition Theorem and $|\mathbf{AB}| = |\mathbf{BA}|$, prove that the determinant of a symmetric matrix is the product of its eigenvalues.

10. Let $\mathbf{X} = [X_{ij}]$ be a random matrix. Show $E(\mathbf{X}') = E(\mathbf{X})'$.

11. Let $\mathbf{X}$ be a random matrix, and $\mathbf{B}$ be a matrix of constants. Show $E(\mathbf{XB}) = E(\mathbf{X})\mathbf{B}$. Recall the definition $\mathbf{AB} = [\sum_k a_{i,k}b_{k,j}]$.

12. Let the $p \times 1$ random vector $\mathbf{x}$ have expected value $\boldsymbol{\mu}$ and variance-covariance matrix $\boldsymbol{\Sigma}$, and let $\mathbf{A}$ be an $m \times p$ matrix of constants. Prove that the variance-covariance matrix of $\mathbf{Ax}$ is either

    - $\mathbf{A}\boldsymbol{\Sigma}\mathbf{A}'$, or
    - $\mathbf{A}^2\boldsymbol{\Sigma}$..

    Pick one and prove it. Start with the definition of a variance-covariance matrix on the formula sheet.

13. If the $p \times 1$ random vector $\mathbf{x}$ has mean $\boldsymbol{\mu}$ and variance-covariance matrix $\boldsymbol{\Sigma}$, show $\boldsymbol{\Sigma} = E(\mathbf{XX}') - \boldsymbol{\mu}\boldsymbol{\mu}'$.

14. Let $\mathbf{x}$ be a $p \times 1$ random vector. Starting with the definition on the formula sheet, prove $cov(\mathbf{x}) = \mathbf{0}$..

15. Let the $p \times 1$ random vector $\mathbf{x}$ have mean $\boldsymbol{\mu}$ and variance-covariance matrix $\boldsymbol{\Sigma}$, let $\mathbf{A}$ be an $r \times p$ matrix of constants, and let $\mathbf{c}$ be an $r \times 1$ vector of constants. Find $cov(\mathbf{Ax} + \mathbf{c})$. Show your work.

16. Let the scalar random variable $y = \mathbf{v}'\mathbf{x}$, where $\mathbf{x}$ is a $p \times 1$ random vector. What is $Var(y)$? Use this to prove that *any* variance-covariance matrix must be positive semi-definite.

17. The square matrix $\mathbf{A}$ has an eigenvalue equal to λ with corresponding eigenvector $\mathbf{x} \neq \mathbf{0}$ if $\mathbf{Ax} = \lambda\mathbf{x}$.

    (a) Show that the eigenvalues of a variance-covariance matrix cannot be negative.

    (b) How do you know that the determinant of a variance-covariance matrix must be greater than or equal to zero? The answer is one short sentence.

    (c) Let x and y be scalar random variables. Recall $Corr(x,y) = \frac{Cov(x,y)}{\sqrt{Var(x)Var(y)}}$. Using what you have shown about the determinant, show $-1 \leq Corr(x,y) \leq 1$. You have just proved the Cauchy-Schwarz inequality using probability tools.

18. Let $\mathbf{x}$ be a $p \times 1$ random vector with mean $\boldsymbol{\mu}_x$ and variance-covariance matrix $\boldsymbol{\Sigma}_x$, and let $\mathbf{y}$ be a $q \times 1$ random vector with mean $\boldsymbol{\mu}_y$ and variance-covariance matrix $\boldsymbol{\Sigma}_y$.

    (a) What is the (i,j) element of $cov(\mathbf{x},\mathbf{y})$? See the definition on the formula sheet.

    (b) Find an expression for $cov(\mathbf{x}+\mathbf{y})$ in terms of $\boldsymbol{\Sigma}_x$, $\boldsymbol{\Sigma}_y$ and $cov(\mathbf{x},\mathbf{y})$. Show your work.

    (c) Simplify further for the special case where $Cov(x_i, y_j) = 0$ for all i and j.

    (d) Let $\mathbf{c}$ be a $p \times 1$ vector of constants and $\mathbf{d}$ be a $q \times 1$ vector of constants. Find $cov(\mathbf{x}+\mathbf{c}, \mathbf{y}+\mathbf{d})$. Show your work.

    (e) Starting with the definition on the formula sheet, show $cov(\mathbf{x},\mathbf{y}) = cov(\mathbf{y},\mathbf{x})$..

    (f) Starting with the definition on the formula sheet, show $cov(\mathbf{x},\mathbf{y}) = \mathbf{0}$..
