STAT 608, Spring 2022 - Assignment 0 (Review - not for a grade)
SOLUTIONS

1. Matrix Algebra Review

   Define matrices $\mathbf{A}$, $\mathbf{B}$, and $\mathbf{C}$ as follows:

$$\mathbf{A} = \begin{bmatrix} 1 & 0 & 2 & 3 \\ -1 & 2 & 0 & -2 \end{bmatrix}, \ \mathbf{B} = \begin{bmatrix} 0 & -1 \\ 3 & 0 \\ 2 & 1 \\ 0 & -2 \end{bmatrix}, \text{ and } \mathbf{C} = \begin{bmatrix} 1 & 0 \\ 0 & 1 \\ 0 & 0 \end{bmatrix}$$

   (a) Calculate $\mathbf{A}'$, the transpose of $\mathbf{A}$.

$$\mathbf{A}' = \begin{bmatrix} 1 & -1 \\ 0 & 2 \\ 2 & 0 \\ 3 & -2 \end{bmatrix}$$

   (b) Calculate $\mathbf{A}' + \mathbf{B}$.

$$\mathbf{A}' + \mathbf{B} = \begin{bmatrix} 1 & -2 \\ 3 & 2 \\ 4 & 1 \\ 3 & -4 \end{bmatrix}$$

   (c) Calculate $\mathbf{AB}$, the matrix product of $\mathbf{A}$ and $\mathbf{B}$.

$$\mathbf{AB} = \begin{bmatrix} 4 & -5 \\ 6 & 5 \end{bmatrix}$$

   (d) Calculate $\mathbf{BA}$. Is $\mathbf{AB} = \mathbf{BA}$?

$$\mathbf{BA} = \begin{bmatrix} 1 & -2 & 0 & 2 \\ 3 & 0 & 6 & 9 \\ 1 & 2 & 4 & 4 \\ 2 & -4 & 0 & 4 \end{bmatrix}$$

   NOTE THAT $\mathbf{AB} \neq \mathbf{BA}$.

   (e) Is the matrix $\mathbf{AB}$ singular? Why or why not? (Invertible means nonsingular; see the Wikipedia page for invertible matrices for a review.)

   NO. THE EASIEST WAY TO SEE THIS IS VIA THE FACT THAT $\text{DET}(\mathbf{AB}) = 20 - (-30) = 50 \neq 0$. WE COULD ALSO SHOW THAT THE COLUMNS OF $\mathbf{AB}$ ARE LINEARLY INDEPENDENT: THE ONLY SET OF a_1, a_2 SUCH THAT $a_1\mathbf{v}_1 + a_2\mathbf{v}_2 = 0$, WHERE $\mathbf{v}_1$ AND $\mathbf{v}_2$ ARE THE COLUMN VECTORS OF $\mathbf{AB}$, ARE $a_1 = a_2 = 0$.

   (f) Calculate the trace of $\mathbf{AB}$.

   THE TRACE OF A MATRIX IS THE SUM OF ITS DIAGONAL ELEMENTS. HERE, $\text{TR}(\mathbf{AB}) = 4 + 5 = 9$.

(g) Write $(\mathbf{AB})'$ in another form algebraically: remove the parentheses.

DISTRIBUTE THE TRANSPOSE: $(\mathbf{AB})' = \mathbf{B}'\mathbf{A}'$.

(h) Calculate $(\mathbf{AB})^{-1}$.

$$(\mathbf{AB})^{-1} = \frac{1}{20-(-30)} \begin{bmatrix} 5 & 5 \\ -6 & 4 \end{bmatrix} = \begin{bmatrix} 1/10 & 1/10 \\ -3/25 & 2/25 \end{bmatrix}$$

(i) Write $\mathbf{I}_2$, the 2×2 identity matrix.

$$\mathbf{I}_2 = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}$$

(j) What is $\mathbf{I}_2\mathbf{A}$? Why?

$$\mathbf{I}_2\mathbf{A} = \mathbf{A}$$

$\mathbf{I}$ IS THE IDENTITY MATRIX. WHEN WE MULTIPLY ANY MATRIX $\mathbf{M}$ BY THE IDENTITY, WE GET BACK THE ORIGINAL MATRIX $\mathbf{M}$.

(k) Describe geometrically the space spanned by $\mathbf{C}$. That is, the space spanned by the two column vectors in the matrix $\mathbf{C}$. Assume we're working in three-dimensional space defined by axes xyz.

THE SPACE SPANNED BY $\mathbf{C}$ IS THE x, y PLANE: THE "FLOOR," IF YOU WILL.

(l) Calculate the projection matrix for $\mathbf{C}$. That is, what is the matrix that projects a vector in x, y, z space onto the x, y plane?

$$\begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 0 \end{bmatrix}$$

(m) Project the vector $\mathbf{d} = [2\ 2\ 2]'$ onto the space spanned by $\mathbf{C}$.

WE MULTIPLY $\mathbf{d}$ BY THE PROJECTION MATRIX OF $\mathbf{C}$: $[2\ 2\ 0]'$.

(n) Describe geometrically what you did in the previous step.

THE SPACE SPANNED BY $\mathbf{C}$ IS THE FLOOR OF THE xyz SPACE, SO WE FOUND THE SHADOW OF A VECTOR SHOOTING OUT OF THE CORNER AT A 45-DEGREE ANGLE. THE SHADOW IS STILL 45 DEGREES FROM EITHER WALL (THE xz PLANE OR THE yz PLANE), BUT ON THE FLOOR.

(o) Are the vectors $\mathbf{d}$ and $\mathbf{f} = [1\ 0\ 0]'$ orthogonal? Why or why not? (Talk about a dot product in your answer.)

TWO VECTORS ARE ORTHOGONAL IF THEIR DOT PRODUCT EQUALS ZERO. HERE, WE HAVE THAT $\mathbf{d}$ AND $\mathbf{f}$ ARE NOT ORTHOGONAL, SINCE

$$\mathbf{d} \cdot \mathbf{f} = 2 \neq 0$$

GEOMETRICALLY, THE VECTOR $\mathbf{f}$ IS THE CORNER BETWEEN THE FLOOR AND THE xz WALL, WHILE THE VECTOR $\mathbf{d}$ IS NOT ANY OF THE CORNERS; IT STICKS OUT OF THE xyz AXIS AT A 45-DEGREE ANGLE, SO IT'S NOT 90 DEGREES FROM ANY OF THE CORNERS.

(p) Calculate the dot product $\mathbf{1} \cdot \mathbf{1}$, where the vector $\mathbf{1} = [1\ 1\ \ldots 1]'$ is of length n.

THE DOT PRODUCT BETWEEN TWO VECTORS IS THE SUM OF THEIR COMPONENT-WISE PRODUCTS. IN THIS CASE, SINCE ALL COMPONENTS OF BOTH VECTORS EQUAL 1 AND THERE ARE n COMPONENTS, THE DOT PRODUCT EQUALS n.

(q) Calculate the dot product $\mathbf{1} \cdot \mathbf{x}$, where $\mathbf{1}$ is defined as above and $\mathbf{x} = [x_1\ x_2\ \ldots x_n]'$.

THIS IS THE SUM OF THE COMPONENTS OF $\mathbf{x}$: $\mathbf{1} \cdot \mathbf{x} = \sum_i x_i$.

(r) Calculate the dot product $\mathbf{x} \cdot \mathbf{x}$, where $\mathbf{x}$ is defined as above.

THIS IS THE SUM OF THE SQUARED COMPONENTS OF $\mathbf{x}$: $\mathbf{x} \cdot \mathbf{x} = \sum_i x_i^2$.

(s) Describe geometrically what the first eigenvector (sorted in order from highest eigenvalue to lowest) would tell you about the vector space.

THE EIGENVECTORS GIVE US AN ORTHOGONAL BASIS FOR A VECTOR SPACE. THE FIRST EIGENVECTOR GIVES THE MOST "INFORMATION" ABOUT THE ORIGINAL MATRIX; LATER IN THE COURSE, WE WILL SAY THAT THE FIRST EIGENVECTOR EXPLAINS THE MOST VARIABILITY IN A DATASET.

2. Calculus Review

Define

$$f(x, y) = 3x^2 + 2xy^2 - y$$

(a) Calculate $\frac{\partial}{\partial x} f(x, y)$.

THIS IS THE PARTIAL DERIVATIVE OF f WITH RESPECT TO x:

$$\frac{\partial}{\partial x} f(x, y) = 6x + 2y^2$$

(b) Calculate $\frac{\partial}{\partial y} f(x, y)$.

THIS IS THE PARTIAL DERIVATIVE OF f WITH RESPECT TO y:

$$\frac{\partial}{\partial y} f(x, y) = 4xy - 1$$

3. Log Review

(a) Calculate $\log(e)$. (Note that statisticians usually write "log" instead of "ln" when they mean log base e.)

WE HAVE THAT $\log(e) = 1$ AND $e^{\log(1)} = 1$. NOTE THAT "LOG" IS HOW R DENOTES LOG BASE e.

(b) Rewrite $\log\left(\frac{x}{y}\right)$ in terms of a difference.

THE LOG OF A FRACTION IS THE DIFFERENCE OF LOGS: $\log\left(\frac{x}{y}\right) = \log(x) - \log(y)$.

(c) Rewrite $\log(x^n)$ in terms of a product.

THE LOG OF AN EXPONENTIAL EXPRESSION IS A PRODUCT: $\log(x^n) = n \log(x)$.

(d) Solve $\log(x) = y$ for x.

$$x = e^y$$

4. Statistics and Linear Regression Review

After regressing eight patients' weights (in kg) on their height (in cm), a doctor found the following output.

Coefficient	Estimate	Std. Error	t-value	$Pr(>\|t\|)$
Intercept	-129.1667	24.3610	-5.302	0.001826
Height	1.1667	0.1521	????	0.000257

(a) Write down the least squares regression line using $\hat{y}$ = predicted weight and x = height.

$$\hat{y} = -129.1667 + 1.1667x$$

(b) What weight does the model predict for someone who is 160 cm tall?

$$-129.1667 + 1.1667 \times 160 = 57.5053\text{kg}$$

(c) Interpret the slope of the line in the context of the model.

According to our model, a one-unit increase in height is expected to be accompanied by a 1.1667 kg increase in weight. Note that our model does not say that weight always increases by 1.1667 kg when height increases by 1 cm. The association is "on average."

(d) Interpret the standard error of the slope in the context of the model.

The standard error of the slope (0.1521) is the estimated standard deviation of the slope estimate that we would see from sample to sample. That is, if we repeated this study many times, the standard deviation of the resulting slope estimates would be estimated to be 0.1521.

(e) Calculate the t-statistic for testing whether the slope is statistically significant.

The t-statistic is the estimate divided by its standard error: $t = \frac{1.1667}{0.1521} = 7.671$.

(f) Are height and weight linearly associated? Explain. (Assume assumptions are met.)

Yes, we have statistically significant evidence that height and weight are linearly associated, since the p-value for the slope is so small. We will talk later in class about how to check the model assumptions.

(g) A journal might report that height is a *significant* predictor of weight. Explain what this means in context, as if to someone with no statistical background.

The word "significant" in statistics actually does not mean important or large. Rather, it means "unlikely to have occurred by chance." In the context of our problem, it is unlikely that we would collect a random sample of patients that resulted in a sample slope of 1.1667 or even larger in absolute value, just by chance, if the true population slope equalled 0.

(h) Calculate a 95% confidence interval for the slope.

YOU WOULD CALCULATE THIS AS $\hat{\beta}_1 \pm t_{n-2}\mathrm{SE}_{\hat{\beta}_1}$. I DIDN'T TELL YOU THE SAMPLE SIZE, SO YOU COULDN'T COMPUTE THE ACTUAL INTERVAL.

(i) Interpret your interval above in context.

I AM 95% CONFIDENT THAT AS HEIGHT INCREASES BY 1 CM IN THE GENERAL POPULATION OF PATIENTS, WEIGHT INCREASES BY AN AMOUNT BETWEEN THE LOWER AND UPPER BOUNDS OF MY INTERVAL.