

Math Camp - Homework 5

Question 1: Show that

$$k(\mathbf{X} + \mathbf{Y}) = k\mathbf{X} + k\mathbf{Y}$$

Question 2: As we discussed today, an important attribute of any matrix is its *dimensionality*: Given that any matrix is a quadrilateral, its size can be expressed as being $m \times n$ (said “ m by n ”), where m represents the number of rows and n represents the number of columns. In these terms, matrix \mathbf{A} is 3×1 , matrix \mathbf{C} is 2×3 , matrix \mathbf{E} is 3×3 , and so on.

We have seen that matrix addition or subtraction only works when the two matrices involved have the same dimensions. Based on your work in question 2, explicitly describe when matrices can or cannot be multiplied. (We touched on this in class, but make sure that it is solidly implanted into your mind.)

Question 3: Using the matrices below, calculate the following. Some may not be defined; if that is the case, say so. (Don’t be worried by the number of questions; most of these should go relatively quickly.)

$$\mathbf{A} = \begin{bmatrix} 3 \\ -2 \\ 9 \end{bmatrix} \quad \mathbf{B} = \begin{bmatrix} 8 \\ 0 \\ -1 \end{bmatrix} \quad \mathbf{C} = \begin{bmatrix} 7 & -1 & 5 \\ 0 & 2 & -4 \end{bmatrix} \quad \mathbf{D} = \begin{bmatrix} 3 & 1 \\ 3 & 4 \\ 3 & -7 \end{bmatrix} \quad \mathbf{E} = \begin{bmatrix} 5 & 2 & 3 \\ 1 & 0 & -4 \\ -2 & 1 & -6 \end{bmatrix}$$

$$\mathbf{F} = \begin{bmatrix} 4 & 1 & -5 \\ 0 & 7 & 7 \\ 2 & -3 & 0 \end{bmatrix} \quad \mathbf{G} = \begin{bmatrix} 2 & -8 & -5 \\ -3 & 7 & -4 \\ 1 & 0 & 3 \\ 1 & 2 & 6 \end{bmatrix} \quad \mathbf{K} = \begin{bmatrix} 9 \\ -2 \\ -1 \\ 0 \end{bmatrix} \quad \mathbf{L} = [5 \quad 0 \quad 3 \quad 1]$$

(Note: $\mathbf{A}' = \mathbf{A}^T$. Both are widely used notation for matrix transposes.)

- | | |
|-----------------------------------|-------------------------------------|
| (a) $\mathbf{A} + \mathbf{B}$ | (h) \mathbf{BC} |
| (b) $-\mathbf{G}$ | (i) \mathbf{FB} |
| (c) \mathbf{D}^T | (j) \mathbf{EF} |
| (d) $\mathbf{C} + \mathbf{D}$ | (k) $\mathbf{K} \cdot \mathbf{L}^T$ |
| (e) $3\mathbf{C} - 2\mathbf{D}^T$ | (l) $ \mathbf{K} $ |
| (f) $\mathbf{A} \cdot \mathbf{B}$ | (m) \mathbf{G}^T |
| (g) \mathbf{CB} | (n) $\mathbf{E} - 5\mathbf{I}_3$ |

Question 4: Prove the following statement, which was in the slides today: $(\mathbf{X} + \mathbf{Y})' = \mathbf{X}' + \mathbf{Y}'$.

Question 5: When it comes to real numbers, we know that if $xy = 0$, then either $x = 0$ or $y = 0$ or both. It is tempting to believe that a similar idea applies to matrices. However, this is not the case. Prove that if the matrix product $\mathbf{AB}=\mathbf{0}$ —by which we mean a matrix of appropriate dimensionality made up entirely of zeroes—then it is not necessarily true that either $\mathbf{A}=\mathbf{0}$ or $\mathbf{B}=\mathbf{0}$. (*Hint:* Remember that you are proving that it is not the case that something is always true. What’s the most straightforward way to do that?)

Question 6: Finding the inverse of matrices can be not only tedious, but difficult.¹ Sometimes, it’s not even clear how to start. Fortunately, for matrices of dimensions 2x2 or 3x3, there are formulas we can apply to find the inverse. This can still be tedious, but it’s no longer difficult. While Justin mentioned you can compute inverses in \mathbb{R} , knowing these formulas is helpful for many proofs where the elements of matrices are unknowns, not specific values. We’ll just focus on the 2x2 example here, but keep in mind the 3x3 formula also exists. The following formula gives the solution to \mathbf{A}^{-1} :

If

$$\mathbf{A} = \begin{bmatrix} a & b \\ c & d \end{bmatrix}$$

then

$$\mathbf{A}^{-1} = \frac{1}{\det(\mathbf{A})} \begin{bmatrix} d & -b \\ -c & a \end{bmatrix}$$

where

$$\frac{1}{\det(\mathbf{A})} = \frac{1}{ad - bc}$$

The term *det* denotes “the determinant” which is a special combination of the elements in a square matrix such that if the determinant is zero the matrix is singular (not invertible).²

¹In fact, there’s a large literature in math and computer science about how to efficiently compute matrix inverses. You’ll see later on in the methods sequence that if you have a lot of data, it can take \mathbb{R} a very long time to invert a matrix.

²For a more complete definition, see here: <http://mathworld.wolfram.com/Determinant.html>.

(You may sometimes see the determinant of \mathbf{A} denoted $|\mathbf{A}|$.) Use this formula to find the inverse of \mathbf{Q} . Then calculate $\mathbf{Q}\mathbf{Q}^{-1}$ to verify your solution is correct (i.e., how would we know if we really found the inverse?).

$$\mathbf{Q} = \begin{bmatrix} 1 & 3 \\ 2 & 7 \end{bmatrix}$$

Question 7: Often times practitioners will use a shorthand for complex matrices by partitioning them. For example, the following matrix,

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

can be written in partitioned form (also known as “block form”) like so:

$$\begin{bmatrix} \mathbf{A}_{11} & \mathbf{A}_{12} \\ \mathbf{A}_{21} & \mathbf{A}_{22} \end{bmatrix}$$

where:

$$\mathbf{A}_{11} = \begin{bmatrix} 1 & 4 \\ 2 & 9 \end{bmatrix}$$

$$\mathbf{A}_{12} = \begin{bmatrix} 0 \\ 0 \end{bmatrix}$$

$$\mathbf{A}_{21} = \begin{bmatrix} 0 & 0 \end{bmatrix}$$

and

$$\mathbf{A}_{22} = \begin{bmatrix} 6 \end{bmatrix}$$

Using the rules of matrix multiplication, show that:

$$\begin{bmatrix} \mathbf{A}_{11} & \mathbf{A}_{12} \\ \mathbf{A}_{21} & \mathbf{A}_{22} \end{bmatrix}' \begin{bmatrix} \mathbf{A}_{11} & \mathbf{A}_{12} \\ \mathbf{A}_{21} & \mathbf{A}_{22} \end{bmatrix} = \begin{bmatrix} \mathbf{A}_{11}'\mathbf{A}_{11} & \mathbf{0} \\ \mathbf{0} & \mathbf{A}_{22}'\mathbf{A}_{22} \end{bmatrix}$$

Question 8: :

- (a) Consider the following function, $\frac{1}{b-a}$, defined on the interval $[a,b]$ where $a=-9$ and $b=12$. Evaluate the integral over this interval.
- (b) Using integration, find the value q at which 95% of the area under the curve lies to the left of q and 5% lies to the right of q . In other words, calculate the 95th percentile of this function.
- (c) Consider a related function, $\frac{x}{b-a}$, defined over the same interval as the function in part a. Evaluate the integral of this function from -9 to q , the point you just identified in part b.

Question 9: We have been dealing with (infinite) series. Later in the methods sequence (especially 450C and 450D), you will encounter (infinite) products, which are notated with a capital pi, which looks like this: \prod . They work just like sums, except you are multiplying the terms rather than adding them. For example,

$$\prod_{n=1}^4 \frac{n}{2} = \frac{1}{2} \times \frac{2}{2} \times \frac{3}{2} \times \frac{4}{2} = 0.5 \times 1 \times 1.5 \times 2 = 1.5$$

$$\prod_{i=1}^n a_i = a_1 \times a_2 \times a_3 \times \dots \times a_n$$

$$\prod_{i=1}^{\infty} a_i = a_1 \times a_2 \times a_3 \times a_4 \times \dots \times a_{89} \times \dots \times a_{5,302} \times \dots$$

Justin briefly brought up the concept of maximum likelihood estimation, which we will spend plenty of time learning later – so don't worry if it's unclear now! Even though we haven't really discussed why this is the case, you probably noticed that maximum likelihood estimation starts with a (log of a) product.

$$\log \left[\prod_{i=1}^{\infty} f(x_i) \right]$$

As suggested in Justin's slides, we'd much rather prefer to deal with infinite series than with infinite products. (For one thing, it's much easier to find derivatives for an infinite sum of

terms than an infinite product of terms; imagine doing product rule on an infinite product!) We can make this happen. Let's do something similar to what Justin did in his slides on maximum likelihood estimation. Using your knowledge of logs and series, show how

$$\log \left[\prod_{i=1}^{\infty} f(x_i) \right] = \sum_{i=1}^{\infty} \log f(x_i)$$

(*Hint:* Write out the first few terms of the infinite product.)

If you understand the mathematical reasoning behind this equality now, it will be an absolute lifesaver in the future.