

Math Camp - Homework 5

Question 1: Show that

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

Solution: First, let's settle on some generic and common notation. Let's define an $m \times n$ matrices \mathbf{X} and \mathbf{Y} as the following:

$$\mathbf{X} = \begin{bmatrix} x_{11} & \cdots & x_{1n} \\ \vdots & \ddots & \vdots \\ x_{m1} & \cdots & x_{mn} \end{bmatrix} \quad \mathbf{Y} = \begin{bmatrix} y_{11} & \cdots & y_{1n} \\ \vdots & \ddots & \vdots \\ y_{m1} & \cdots & y_{mn} \end{bmatrix}$$

These are very generic descriptions of matrices that will help prove broad statements about matrices. Make sure that you understand how this notation works, and how the $m \times n$ dimensionality of the matrix lines up with the uses of m and n to describe the entries inside this generic matrix.

Let's first address $k(\mathbf{X} + \mathbf{Y})$. Using this notation, we can write

$$\mathbf{X} + \mathbf{Y} = \begin{bmatrix} x_{11} + y_{11} & \cdots & x_{1n} + y_{1n} \\ \vdots & \ddots & \vdots \\ x_{m1} + y_{m1} & \cdots & x_{mn} + y_{mn} \end{bmatrix}$$

We saw in class that scalar multiplication also works on matrices. So,

$$k(\mathbf{X} + \mathbf{Y}) = \begin{bmatrix} k(x_{11} + y_{11}) & \cdots & k(x_{1n} + y_{1n}) \\ \vdots & \ddots & \vdots \\ k(x_{m1} + y_{m1}) & \cdots & k(x_{mn} + y_{mn}) \end{bmatrix}$$

We know that distribution applies to scalars. We can distribute the k inside every term of this matrix.

$$k(\mathbf{X} + \mathbf{Y}) = \begin{bmatrix} kx_{11} + ky_{11} & \cdots & kx_{1n} + ky_{1n} \\ \vdots & \ddots & \vdots \\ kx_{m1} + ky_{m1} & \cdots & kx_{mn} + ky_{mn} \end{bmatrix}$$

Now, let's go back to the expression $k\mathbf{X} + k\mathbf{Y}$. Let's distribute the k into each of the matrices separately.

$$k\mathbf{X} = \begin{bmatrix} kx_{11} & \cdots & kx_{1n} \\ \vdots & \ddots & \vdots \\ kx_{m1} & \cdots & kx_{mn} \end{bmatrix} \quad k\mathbf{Y} = \begin{bmatrix} ky_{11} & \cdots & ky_{1n} \\ \vdots & \ddots & \vdots \\ ky_{m1} & \cdots & ky_{mn} \end{bmatrix}$$

Now, let's add these two matrices together to get $k\mathbf{X} + k\mathbf{Y}$.

$$k\mathbf{X} + k\mathbf{Y} = \begin{bmatrix} kx_{11} + ky_{11} & \dots & kx_{1n} + ky_{1n} \\ & \ddots & \\ kx_{m1} + ky_{m1} & \dots & kx_{mn} + ky_{mn} \end{bmatrix}$$

That matrix is identical to the last thing we got with $k(\mathbf{X} + \mathbf{Y})$. Therefore, $k(\mathbf{X} + \mathbf{Y}) = k\mathbf{X} + k\mathbf{Y}$. We could have also taken the last term we got with $k(\mathbf{X} + \mathbf{Y})$ and just split it into the two matrices that produced the sum. That would have also shown the equality.

Question 2: 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.)

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

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

(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) $\mathbf{E} - 5\mathbf{I}_3$ |

Solutions

(a)

$$\mathbf{A} + \mathbf{B} = \begin{bmatrix} 3 + 8 \\ -2 + 0 \\ 9 + (-1) \end{bmatrix} = \begin{bmatrix} 11 \\ -2 \\ 8 \end{bmatrix}$$

(b)

$$-\mathbf{G} = (-1) \begin{bmatrix} 2 & -8 & -5 \\ -3 & 7 & -4 \\ 1 & 0 & 3 \\ 1 & 2 & 6 \end{bmatrix} = \begin{bmatrix} -2 & 8 & 5 \\ 3 & -7 & 4 \\ -1 & 0 & -3 \\ -1 & -2 & -6 \end{bmatrix}$$

(c)

$$\mathbf{D}^T = \begin{bmatrix} 3 & 3 & 3 \\ 1 & 4 & -7 \end{bmatrix}$$

(d) $\mathbf{C} + \mathbf{D}$ does not exist.

(e)

$$\begin{aligned} 3\mathbf{C} - 2\mathbf{D}^T &= (3) \begin{bmatrix} 7 & -1 & 5 \\ 0 & 2 & -4 \end{bmatrix} - (2) \begin{bmatrix} 3 & 3 & 3 \\ 1 & 4 & -7 \end{bmatrix} \\ &= \begin{bmatrix} 21 & -3 & 15 \\ 0 & 6 & -12 \end{bmatrix} - \begin{bmatrix} 6 & 6 & 6 \\ 2 & 8 & -14 \end{bmatrix} \\ &= \begin{bmatrix} 15 & & 9 \\ -2 & -2 & 2 \end{bmatrix} \end{aligned}$$

(f)

$$\mathbf{A} \cdot \mathbf{B} = 3(8) + (-2)(0) + 9(-1) = 24 + 0 - 9 = 15$$

(g)

$$\begin{aligned} \mathbf{CB} &= \begin{bmatrix} 7 & -1 & 5 \\ 0 & 2 & -4 \end{bmatrix} \begin{bmatrix} 8 \\ 0 \\ -1 \end{bmatrix} \\ &= \begin{bmatrix} 7(8) + (-1)(0) + 5(-1) \\ 0(8) + 2(0) + (-4)(-1) \end{bmatrix} \\ &= \begin{bmatrix} 56 + 0 - 5 \\ 0 + 0 + 4 \end{bmatrix} \\ &= \begin{bmatrix} 51 \\ 4 \end{bmatrix} \end{aligned}$$

(h) \mathbf{BC} does not exist.

(i)

$$\begin{aligned}
\mathbf{FB} &= \begin{bmatrix} 4 & 1 & -5 \\ 0 & 7 & 7 \\ 2 & -3 & 0 \end{bmatrix} \begin{bmatrix} 8 \\ 0 \\ -1 \end{bmatrix} \\
&= \begin{bmatrix} 4(8) + 1(0) + (-5)(-1) \\ 0(8) + 7(0) + 7(-1) \\ 2(8) + (-3)(0) + 0(-1) \end{bmatrix} \\
&= \begin{bmatrix} 32 + 0 + 5 \\ 0 + 0 - 7 \\ 16 + 0 + 0 \end{bmatrix} \\
&= \begin{bmatrix} 37 \\ -7 \\ 16 \end{bmatrix}
\end{aligned}$$

(j) (Vertical and horizontal lines are added in the work here to clearly delineate each term.)

$$\begin{aligned}
\mathbf{EF} &= \left[\begin{array}{ccc|ccc|ccc} 5(4) & + & 2(0) & + & 3(2) & 5(1) & + & 2(7) & + & 3(-3) & 5(-5) & + & 2(7) & + & 3(0) \\ 1(4) & + & 0(0) & + & (-4)(2) & 1(1) & + & 0(7) & + & (-4)(-3) & 1(-5) & + & 0(7) & + & (-4)(0) \\ -2(4) & + & 1(0) & + & (-6)(2) & -2(1) & + & 1(7) & + & (-6)(-3) & -2(-5) & + & 1(7) & + & (-6)(0) \end{array} \right] \\
&= \begin{bmatrix} 20 + 0 + 6 & 5 + 14 - 9 & -25 + 14 + 0 \\ 4 + 0 - 8 & 1 + 0 + 12 & -5 + 0 + 0 \\ -8 + 0 - 12 & -2 + 7 + 18 & 10 + 7 + 0 \end{bmatrix} \\
&= \begin{bmatrix} 26 & 10 & -11 \\ -4 & 13 & -5 \\ -20 & 23 & 17 \end{bmatrix}
\end{aligned}$$

(k)

$$\mathbf{K} \cdot \mathbf{L}^T = \begin{bmatrix} 9 \\ -2 \\ -1 \\ 0 \end{bmatrix} \cdot \begin{bmatrix} 5 \\ 0 \\ 3 \\ 1 \end{bmatrix} = 9(5) + (-2)(0) + (-1)(3) + 0(1) = 45 + 0 - 3 + 0 = 42$$

(l)

$$\|\mathbf{K}\| = \sqrt{\mathbf{K} \cdot \mathbf{K}} = \sqrt{9^2 + (-2)^2 + (-1)^2 + 0^2} = \sqrt{81 + 4 + 1 + 0} = \sqrt{86}$$

(m)

$$\mathbf{G}^T = \begin{bmatrix} 2 & -8 & -5 \\ -3 & 7 & -4 \\ 1 & 0 & 3 \\ 1 & 2 & 6 \end{bmatrix}^T = \begin{bmatrix} 2 & -3 & 1 & 1 \\ -8 & 7 & 0 & 2 \\ -5 & -4 & 3 & 6 \end{bmatrix}$$

(n)

$$\begin{aligned}\mathbf{E} - 5\mathbf{I}_3 &= \begin{bmatrix} 5 & 2 & 3 \\ 1 & 0 & -4 \\ -2 & 1 & -6 \end{bmatrix} - (5) \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} \\ &= \begin{bmatrix} 5 & 2 & 3 \\ 1 & 0 & -4 \\ -2 & 1 & -6 \end{bmatrix} - \begin{bmatrix} 5 & 0 & 0 \\ 0 & 5 & 0 \\ 0 & 0 & 5 \end{bmatrix} \\ &= \begin{bmatrix} 0 & 2 & 3 \\ 1 & -5 & -4 \\ -2 & 1 & -11 \end{bmatrix}\end{aligned}$$

Question 3: 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.)

Solution: We talked about this, but to be clear, the number of columns in the first matrix must equal the number of rows in the second matrix for matrix multiplication to work. That is, given some matrix product \mathbf{AB} , it must be that \mathbf{A} is an $m \times n$ matrix and that \mathbf{B} is an $n \times p$ matrix. The m and p can be anything, but the two n ’s must match. Matrices must be *conformable* in this manner in order to be multiplied. Furthermore, if the matrices are conformable, the dimension of their product will be $m \times p$. Review the matrix multiplication problems in question 2 to verify that this is the case.

Always check whether matrices are conformable before attempting to do anything with them; you could save yourself a lot of time and hassle.

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

Solution: This is a relatively simple proof, and there are a lot of different ways to do it. The most straightforward way is simply to write out each of the two matrices and observe that they are identical. Without loss of generality, suppose \mathbf{X} and \mathbf{Y} are $m \times n$ matrices. We know that

$$\mathbf{X} + \mathbf{Y} = \begin{bmatrix} x_{11} + y_{11} & \dots & x_{1n} + y_{1n} \\ \vdots & \ddots & \vdots \\ x_{m1} + y_{m1} & \dots & x_{mn} + y_{mn} \end{bmatrix}$$

Then it must be that

$$(\mathbf{X} + \mathbf{Y})' = \begin{bmatrix} x_{11} + y_{11} & \dots & x_{m1} + y_{m1} \\ \vdots & \ddots & \vdots \\ x_{1n} + y_{1n} & \dots & x_{mn} + y_{mn} \end{bmatrix}$$

Now let's consider the right hand side of the equation. Since

$$\mathbf{X}' = \begin{bmatrix} x_{11} & \dots & x_{m1} \\ \vdots & \ddots & \vdots \\ x_{1n} & \dots & x_{mn} \end{bmatrix} \quad \mathbf{Y} = \begin{bmatrix} y_{11} & \dots & y_{m1} \\ \vdots & \ddots & \vdots \\ y_{1n} & \dots & y_{mn} \end{bmatrix}$$

we know that

$$\mathbf{X}' + \mathbf{Y}' = \begin{bmatrix} x_{11} + y_{11} & \dots & x_{m1} + y_{m1} \\ \vdots & \ddots & \vdots \\ x_{1n} + y_{1n} & \dots & x_{mn} + y_{mn} \end{bmatrix}$$

which is the same as $(\mathbf{X} + \mathbf{Y})'$.

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?)

Solution: Generally speaking, it is easy to show that something is *not* necessarily true. All that is needed is a single counterexample! And in this case, there are infinitely many counterexamples. Here's one:

$$\mathbf{A} = \begin{bmatrix} 1 & -1 \\ -1 & 1 \end{bmatrix} \quad \mathbf{B} = \begin{bmatrix} 1 & 1 \\ 1 & 1 \end{bmatrix}$$

$$\mathbf{AB} = \begin{bmatrix} 1(1) + 1(-1) & 1(1) + 1(-1) \\ 1(-1) + 1(1) & 1(-1) + 1(1) \end{bmatrix} = \begin{bmatrix} 0 & 0 \\ 0 & 0 \end{bmatrix}$$

Question 6: Finding the inverse of matrices can be not only difficult, but tedious. 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 \mathbf{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). For a more complete definition, see here: <http://mathworld.wolfram.com/Determinant.html>. Use this formula to find the inverse of \mathbf{Q} . Then solve $\mathbf{Q}\mathbf{Q}^{-1}$ to verify your solution is correct (if it is, you should get back the Identity matrix).

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

Answer:

$$\begin{aligned} \mathbf{Q}^{-1} &= \frac{1}{1 * 7 - 3 * 2} * \begin{bmatrix} 7 & -3 \\ -2 & 1 \end{bmatrix} \\ &= \frac{1}{1} * \begin{bmatrix} 7 & -3 \\ -2 & 1 \end{bmatrix} \\ &= \begin{bmatrix} 7 & -3 \\ -2 & 1 \end{bmatrix} \end{aligned}$$

Let's confirm this is an inverse:

$$\begin{aligned}
\mathbf{Q}\mathbf{Q}^{-1} &= \begin{bmatrix} 7 & -3 \\ -2 & 1 \end{bmatrix} \begin{bmatrix} 1 & 3 \\ 2 & 7 \end{bmatrix} \\
&= \begin{bmatrix} (7 * 1) + (-3 * 2) & (7 * 3) + (-3 * 7) \\ (-2 * 1) + (1 * 2) & (-2 * 3) + (1 * 7) \end{bmatrix} \\
&= \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}
\end{aligned}$$

Done.

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

$$\mathbf{A}_{22} = [6]$$

Using the rules of matrix multiplication, show that:

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

Answer:

$$\mathbf{A}'_{11}\mathbf{A}_{11} = \begin{bmatrix} 1 & 2 \\ 4 & 9 \end{bmatrix} \begin{bmatrix} 1 & 4 \\ 2 & 9 \end{bmatrix} = \begin{bmatrix} 5 & 22 \\ 22 & 97 \end{bmatrix}$$

$$\mathbf{A}'_{22}\mathbf{A}_{22} = [6] [6] = [36]$$

$$\begin{aligned} \begin{bmatrix} \mathbf{A}_{11} & \mathbf{A}_{12} \\ \mathbf{A}_{21} & \mathbf{A}_{22} \end{bmatrix}' \begin{bmatrix} \mathbf{A}_{11} & \mathbf{A}_{12} \\ \mathbf{A}_{21} & \mathbf{A}_{22} \end{bmatrix} &= \begin{bmatrix} 1 & 4 & 0 \\ 2 & 9 & 0 \\ 0 & 0 & 6 \end{bmatrix}' \begin{bmatrix} 1 & 4 & 0 \\ 2 & 9 & 0 \\ 0 & 0 & 6 \end{bmatrix} \\ &= \begin{bmatrix} 1 & 2 & 0 \\ 4 & 9 & 0 \\ 0 & 0 & 6 \end{bmatrix} \begin{bmatrix} 1 & 4 & 0 \\ 2 & 9 & 0 \\ 0 & 0 & 6 \end{bmatrix} \\ &= \begin{bmatrix} 5 & 22 & 0 \\ 22 & 97 & 0 \\ 0 & 0 & 36 \end{bmatrix} \\ &= \begin{bmatrix} \mathbf{A}'_{11}\mathbf{A}_{11} & \mathbf{0} \\ \mathbf{0} & \mathbf{A}'_{22}\mathbf{A}_{22} \end{bmatrix} \end{aligned}$$

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.

$$\begin{aligned}
\int_{-9}^{12} \frac{1}{12 - (-9)} dx &= \left. \frac{x}{21} \right|_{-9}^{12} \\
&= \frac{12}{21} - -\frac{9}{21} \\
&= \frac{21}{21} = 1
\end{aligned}$$

This function integrates out to 1. Later in Math Camp we'll talk about probability distribution functions, which have the requirement that they integrate out to 1. In this case you've just integrated a uniform probability distribution function with parameters $a = -9$ and $b = 12$. If you google image "uniform pdf" there are several great examples of what this looks like, its basically a flat line that goes from a to b on the x-axis at $1/(b-a)$ on the y-axis.

- (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.

$$\begin{aligned}
.95 &= \int_{-9}^q \frac{1}{21} dx \\
&= \left. \frac{x}{21} \right|_{-9}^q \\
&= \frac{q}{21} - -\frac{9}{21} \\
&= \frac{q + 9}{21} \\
.95 * 21 &= q + 9 \\
.95 * 21 - 9 &= 10.95
\end{aligned}$$

95% of the area under the function is to the left of 10.95 and 5% of the area under the function is to the right.

- (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 identified in part b.

$$\begin{aligned}
\int_{-9}^{10.95} \frac{x}{21} dx &= \frac{x^2}{42} \Big|_{-9}^{10.95} \\
&= \frac{(10.95)^2}{42} - \frac{(-9)^2}{42} \\
&= \frac{119.9025 - 81}{42} \\
&= 0.92625
\end{aligned}$$

Question 9: We have been dealing with (infinite) series. Later in 350A (and in 350B and 350C), 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,

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

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

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

Today, Justin 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.

Solution

Hopefully you will find that the solution is not that hard too understand if you take it step by step. As suggested in the hint, let's take the term with the product and write out the first few terms. Notice that this is all taking place within the log.

$$\log \left[\prod_{i=1}^{\infty} f(x_i) \right] = \log [f(x_1) \times f(x_2) \times f(x_3) \times \dots]$$

We're doing a lot of multiplication inside the log. Haven't we seen another way to write a log of products? Look back at your notes on logs: $\log(ab) = \log a + \log b$. We can extend this idea to logs of products that involve more than two terms. (Try it out yourself if you don't believe it!) If we apply that to our problem, we get that

$$\log[f(x_1) \times f(x_2) \times f(x_3) \times \dots] = \log f(x_1) + \log f(x_2) + \log f(x_3) + \dots$$

Now we're dealing with the sum of many logs! We can make this expression more compact using summation notation:

$$\log f(x_1) + \log f(x_2) + \log f(x_3) + \dots = \sum_{i=1}^{\infty} \log f(x_i)$$

We've reached our final destination and proved the equality.

It might help to see the whole process in one go. Talk yourself through it to make sure you understand what's happening at each step.

$$\begin{aligned} \log \left[\prod_{i=1}^{\infty} f(x_i) \right] &= \log [f(x_1) \times f(x_2) \times f(x_3) \times \dots] \\ &= \log f(x_1) + \log f(x_2) + \log f(x_3) + \dots \\ &= \sum_{i=1}^{\infty} \log f(x_i) \end{aligned}$$

This is the power of logs. Like we said in the homework, if you understand this math now, you won't have to worry about it when we do it over and over as we learn about maximum likelihood estimation. That's a good thing – there will be plenty of other things to worry about at that point!