Homework 0

Amy Pitts

1/23/19

1 Instructions

No deliverables

2 Course Setup

2.1 Python

After downloading all the required packages running this script shows that everything is install and their versions.

```
import sys
import numpy
import scipy
import sklearn
import matplotlib
import pandas

print (sys.version) #prints python version
print (numpy.__version__) #prints numpy s version
print (scipy.__version__) #print scipy s version
print (sklearn.__version__) #print sklearn's version
print (matplotlib.__version__) #print matplotlib version
print (pandas.__version__) #prints pandas  version
```

Output of the script file

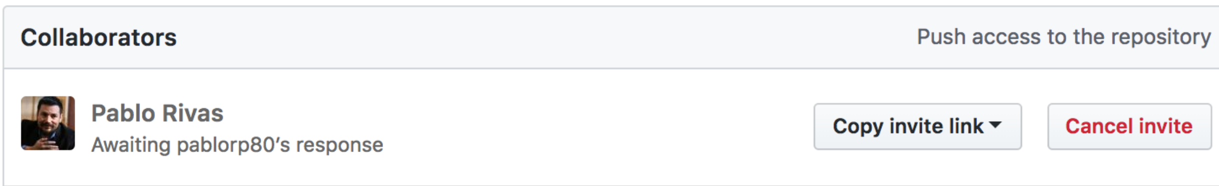
```
1. bash
148-100-152-90:Documents Amy$ python3 install_check.py
3.7.2 (v3.7.2:9a3ffc0492, Dec 24 2018, 02:44:43)
[Clang 6.0 (clang-600.0.57)]
1.16.0
1.2.0
0.20.2
3.0.2
0.23.4
148-100-152-90:Documents Amy$
```

2.2 GitHub Class Repository

- Github name and class repository:

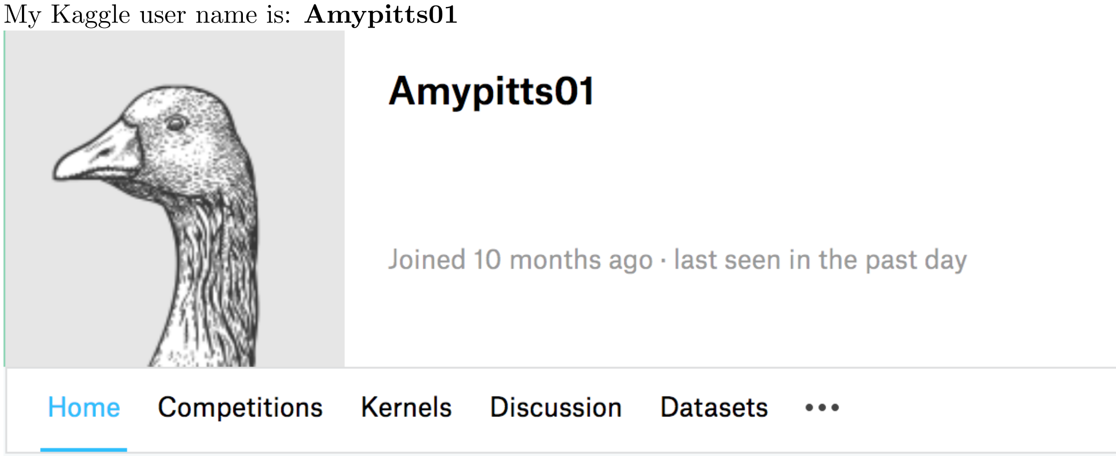
- Link to class repository: *https://github.com/amypitts01/data440/invitations*

- Proof that pablorp80 was asked to be a collaborator:



2.3 Kaggle Account

My Kaggle user name is: **Amypitts01**



3 Mathematics Requirements

No deliverables

4 Problems

1. For the function $g(x) = -3x^2 + 24x - 30$, find the value for x that maximizes $g(x)$.

Solution: To maximize this function the first step is to take the derivative such that $g'(x) = -6x + 24$. Setting the derivative equal to 0 to find the critical values we have:

$$
\begin{aligned}
-6x + 24 &= 0 \\
-6(x - 4) &= 0 \\
x - 4 &= 0 \\
x &= 4.
\end{aligned}
$$

The maximun occurs when the derivative is positive left of the critical value and negative right of the critical value. Since there is only one critical value we do not need to compare to other critical values to find the biggest y value. To show that 4 is in fact a max we see that:

Left: $g'(0) = -6(0) + 24 = 24$ *positive*

Right: $g'(10) = -6(10) + 24 = -36$ *negative*

Therefore, $x = 4$ maximizes the function.

2. Consider the following function:

$$f(x) = 3x_0^3 - 2x_0x_1^2 + 4x_1 - 8$$

what are the partial derivatives of $f(x)$ with respect to x_0 and x_1.

Solution:

$$\frac{\partial f}{\partial x_0} = 9x_0^2 - 2x_1^2$$
$$\frac{\partial f}{\partial x_1} = -4x_0x_1 + 4$$

3. Consider the matrix $A = \begin{bmatrix} 1 & 4 & -3 \\ 2 & -1 & 3 \end{bmatrix}$ and $B = \begin{bmatrix} -2 & 0 & 5 \\ 0 & -1 & 4 \end{bmatrix}$m then answer the following and verify your answers in Python:

(a) can you multiply the two matrices? ellaborate on your answer.

- **Solution:** No you can not multiply the two becuase A does not have the same number of collumns and B's rows. They are both $(2, 3)$ matrices and to be able to multiply A would need to be $(2, 3)$ and B dims $(3, 2)$ or A could be $(3, 2)$ and B $(2, 3)$.

(b) multiply A^T and B and give its *rank*.

- **Solution:** First we need to find the transopes of A. This results in $A^T = \begin{bmatrix} 1 & 2 \\ 4 & -1 \\ -3 & 3 \end{bmatrix}$. Now we can multiply A^TB. This gives us $A^TB = \begin{bmatrix} 1 & 2 \\ 4 & -1 \\ -3 & 3 \end{bmatrix} \begin{bmatrix} -2 & 0 & 5 \\ 0 & -1 & 4 \end{bmatrix}$
$= \begin{bmatrix} (1)(-2) + (2)(0) & (1)(0) + (2)(-1) & (1)(5) + (2)(4) \\ (4)(-2) + (-1)(0) & (4)(0) + (-1)(-1) & (4)(5) + (-1)(4) \\ (-3)(-2) + (3)(0) & (-3)(0) + (3)(-1) & (-3)(5) + (3)(4) \end{bmatrix}$
$= \begin{bmatrix} -2 & -2 & 13 \\ -8 & 1 & 16 \\ 6 & -3 & -3 \end{bmatrix}$.
- **Solution:** Next to find the rank we need to row reduce until we reach row echelon form. Then we have $\begin{bmatrix} -2 & -2 & 13 \\ -8 & 1 & 16 \\ 6 & -3 & -3 \end{bmatrix} \rightarrow \begin{bmatrix} 1 & 1 & \frac{-13}{2} \\ -8 & 1 & 16 \\ 6 & -3 & -3 \end{bmatrix} \rightarrow \begin{bmatrix} 1 & 1 & \frac{-13}{2} \\ 0 & 9 & -36 \\ 0 & -9 & 36 \end{bmatrix} \rightarrow \begin{bmatrix} 1 & 1 & \frac{-13}{2} \\ 0 & 1 & -4 \\ 0 & 0 & 0 \end{bmatrix}$. The row eachelon of A^TB shows two independent row vectors therefore the **rank is 2**.

(c) (extra credit) let $c = \begin{bmatrix} 1 & 0 \\ 0 & 2 \end{bmatrix}$ be a new matrix; what is the result of $AB^T + C^{-1}$?

- **Solution:** First we need to find c^{-1}. This gives us
$c^{-1} = \frac{1}{ad-bc} \begin{bmatrix} d & -b \\ -c & a \end{bmatrix} = \frac{1}{2} \begin{bmatrix} 2 & 0 \\ 0 & 1 \end{bmatrix} = \begin{bmatrix} 1 & 0 \\ 0 & \frac{1}{2} \end{bmatrix}$.
Now we need to find B^T. Then we have $B^T = \begin{bmatrix} -2 & 0 \\ 0 & -1 \\ 4 & 5 \end{bmatrix}$. Now we can find AB^T. Thus we have

$AB^T = \begin{bmatrix} 1 & 4 & -3 \\ 2 & -1 & 3 \end{bmatrix} \begin{bmatrix} -2 & 0 \\ 0 & -1 \\ 4 & 5 \end{bmatrix} = \begin{bmatrix} -14 & -19 \\ 8 & 16 \end{bmatrix}$. Then adding in C^{-1} we get
$AB^T + C^{-1} = \begin{bmatrix} -14 & -19 \\ 8 & 16 \end{bmatrix} + \begin{bmatrix} 2 & 0 \\ 0 & \frac{1}{2} \end{bmatrix} = \begin{bmatrix} -13 & -19 \\ 8 & 16.5 \end{bmatrix}$.

Solution: Here is the scipt I ran to check the matrices.

```
import numpy as np

#The two arrays given in question three
A = np.array([[1,4,-3],[2,-1,3]])
B = np.array([[-2,0,5],[0,-1,4]])

#Question 3.a
A.dot(B) #This shows that we can not multiply the two arrays
#Question 3.b
A.T.dot(B) #this is just the transpose
np.linalg.matrix_rank(A.T.dot(B))#this is the rank
```

Python answer for 3.a

```
148-100-152-90:hw0 Amy$ python3 matrix_checks.py
Traceback (most recent call last):
  File "matrix_checks.py", line 8, in <module>
    A.dot(B) #This shows that we can not multiply the two arrays
ValueError: shapes (2,3) and (2,3) not aligned: 3 (dim 1) != 2 (dim 0)
```

Python answer for 3.b

```
>>> A.T.dot(B)
array([[-2, -2, 13],
       [-8,  1, 16],
       [ 6, -3, -3]])
>>> np.linalg.matrix_rank(A.T.dot(B))
2
```

4. Give the mathematical definitions of the simple Gaussian, multivariate Gaussian, Bernoulli, binomial, and exponential distributions.

Solutions:

Gaussian: This is also known as the normal distibution and has a pdf of $f_X(x) = \frac{1}{\sigma\sqrt{2\pi}} exp\{-\frac{(x-\mu)^2}{2\sigma^2}\}$ and it is notation is $N(\mu, \sigma^2)$.

Multivariate Gaussian: This distrabution has a pdf of $f_X(x) = (2\pi)^{\frac{k}{\Sigma}} |\Sigma|^{\frac{-1}{2}} e^{\frac{-1}{2}(x-\mu)^T \Sigma^{-1}(x-\mu)}$ and it is notation is MVN(μ, Σ).

Bernoulli: This distrabution has a pdf of $f_X(x) = p^x(1-p)^{1-x}$ with notation Bern(p).

binomial: This distrabution has a pdf of $f_X(x) = \binom{n}{k} p^x(1-p)^{n-x}$ with notation Bin(n, p).

exponential distributions: This distrabution has a pdf of $f(x;\lambda) = \begin{cases} \lambda e^{-\lambda x} & x \geq 0 \\ 0 & x < 0 \end{cases}$.

5. (extra credit) What is the relationship between the Bernoulli and binomial.

Solution: The Bernoulli distributions is made up of Bernoulli trails where an even occurs p or the even does not occur $1-p$. There is only two options success or failure. When you have n independent Bernoulli trails those successes and failures are represented in a Binomial Distribution.

6. Suppoe that random variable $X \sim N(2,3)$. What is its expected value?

Solution: $N(2,3)$ is a normal distrabution it has a $E[x] = \mu$. Since $N(\mu, \sigma^2)$ we have that in our case $E[x] = 2$.

7. An *euclidean projection* of a d-dimensional point $y \in \mathbb{R}^d$ to a set $\mathcal{Z}$ is given by the following optimzation problem:

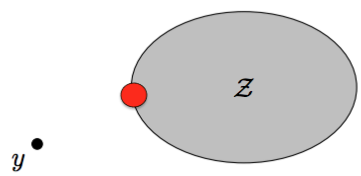
$$x^* = \arg\min_x ||x - y||_2^2, \text{ subject to } x \in \mathcal{Z}$$

where $\mathcal{Z}$ is the feasible set, $||\cdot||_2$ is the l_2-norm (euclidean) of a vector, and $x^* \in \mathbb{R}^d$ is the projected vector.

(a) What is x^* if $y = 1.1$ and $\mathcal{Z} = \mathbb{N}$?

- **Solution:** the $\arg\min_x ||x-y||_2^2$ is the value of x for which $||x-y||_2^2$ attains it's min value. Since $x \in \mathcal{Z} = \mathbb{N}$ we can use trial and errors to find the min.
  First lets start with $x = 0$ then we have: $||0 - 1.1||_2^2 = \sqrt{(0-1.1)\cdot(0-1.1)}^2 = 1.21$.
  Now looking at $x = 1$ we have: $||1 - 1.1||_2^2 = \sqrt{(1-1.1)\cdot(1-1.1)}^2 = 0.01$.
  Finally looking at $x = 2$ we have $||2 - 1.1||_2^2 = \sqrt{(2-1.1)\cdot(2-1.1)}^2 = .81$.
  Because the natural numbers only increase from here the smallest value is going to be made when $x = 1$ obtaining a $x^* = 0.01$.

(b) Locate x^* in the following picture:
The red dot should be the closest point on the perimeter of the $\mathcal{Z}$.



Solution:

8. Suppose that random variable Y has distribution:

$$p(Y = y) = \begin{cases} e^{-y} & y \geq 0 \\ 0 & otherwise \end{cases}$$

(a) Verify that $\int_{y=-\infty}^{\infty} p(Y = y)dy = 1$

(b) What is $\mu_y = E[Y] = \int_{y=-\infty}^{\infty} p(Y = y)ydy$; that is, the expected value of Y?

(c) What is $\sigma^2 = Var[Y] = \int_{y=-\infty}^{\infty} p(Y = y)(y - \mu_y)^2dy$; that is, the variance of Y?

(d) What is $E[Y|Y \geq 10]$; that is, the expected value of Y, given that (or conditioned on) $Y \geq 10$?

8.a Solution:

$$\begin{aligned} \int_{y=-\infty}^{\infty} p(Y = y)dy &= 1 \\ \int_{-\infty}^{0} 0dy + \int_{0}^{\infty} e^{-y}dy &= 1 \\ 0 - e^{-y}|_0^\infty &= 1 \\ 1 &= 1. \end{aligned}$$

8.b Solution:

$$\begin{aligned}
\mu_y = E[Y] &= \int_{y=-\infty}^{\infty} p(Y=y)y dy \\
&= \int_{-\infty}^{0} 0 dy + \int_{0}^{\infty} y e^{-y} dy \\
&= 0 + \int_{0}^{\infty} y e^{-y} dy
\end{aligned}$$

Now we need to do integration by parts. Using $u = y$, $du = 1dy$, $v = -e^{-y}$ and $dv = e^{-y}$ we have:

$$\begin{aligned}
\mu_y &= \int_{0}^{\infty} y e^{-y} dy \\
&= -ye^{-y}|_0^{\infty} + \int_{0}^{\infty} e^{-y} dy \\
&= 0 - e^{-y}|_0^{\infty} \\
&= 1.
\end{aligned}$$

8.c Solution:

$$\begin{aligned}
\sigma^2 = Var[Y] &= \int_{y=-\infty}^{\infty} p(Y=y)(y-\mu_y)^2 dy \\
&= \int_{-\infty}^{0} 0(y-\mu_y)^2 dy + \int_{0}^{\infty} e^{-y}(y-\mu_y)^2 dy \\
&= \int_{0}^{\infty} e^{-y}(y^2 - 2y\mu_y + \mu_y^2) dy \\
&= \int_{0}^{\infty} e^{-y} y^2 dy - \int_{0}^{\infty} 2e^{-y} y\mu_y dy + \int_{0}^{\infty} e^{-y}\mu_y^2 dy
\end{aligned}$$

Now looking at each integral individually we have

$$\begin{aligned}
\int_{0}^{\infty} e^{-y} y^2 dy &= -y^2 e^{-y}|_0^{\infty} - 2\int_{0}^{\infty} e^{-y} y dy \\
&= 0 + 2E[Y] \\
&= 2.
\end{aligned}$$

$$\begin{aligned}
-\int_{0}^{\infty} 2e^{-y} y\mu_y dy &= -\int_{0}^{\infty} 2e^{-y} y\mu_y dy \\
&= -2\mu_y \int_{0}^{\infty} e^{-y} y dy \\
&= -2\mu_y E[Y] \\
&= -2\mu_y.
\end{aligned}$$

$$\int_{0}^{\infty} e^{-y}\mu_y^2 dy = \mu_y^2 \int_{0}^{\infty} e^{-y} dy = \mu_y^2.$$

Putting this all together we have that

$$Var[Y] = \int_{0}^{\infty} e^{-y} y^2 dy - \int_{0}^{\infty} 2e^{-y} y\mu_y dy + \int_{0}^{\infty} e^{-y}\mu_y^2 dy = 2 - 2\mu_y + \mu_y^2.$$

8.d Solution:

$$
\begin{aligned}
E[Y|Y \geq 10] &= \int_{10}^{\infty} p(Y = y) y dy \\
&= \int_{10}^{\infty} y e^{-y} dy
\end{aligned}
$$

Now we need to do integration by parts. Using $u = y$, $du = 1dy$, $v = -e^{-y}$ and $dv = e^{-y}$ we have:

$$
\begin{aligned}
\int_{10}^{\infty} y e^{-y} dy &= -y e^{-y}|_{10}^{\infty} + \int_{10}^{\infty} e^{-y} dy \\
&= 0 + 10e^{-10} - e^{-y}|_{10}^{\infty} \\
&= 10e^{-10} + e^{-10} \\
&= 11e^{-10}.
\end{aligned}
$$