# Homework 1

Felipe Toledo
ft8bn

February 23, 2021

# Clarifications

- I discussed the homework with Dane Williamson (dw3zn).

# Problem 1

$$1. \quad \overset{M}{\begin{bmatrix} 0{,}1 & 0{,}3 & 0{,}7 & 0{,}4 & 0{,}2 \\ 0{,}8 & 1{,}0 & 0{,}6 & 0{,}1 & 0{,}9 \end{bmatrix}} \cdot \overset{x}{\begin{bmatrix} 1 \\ 0 \\ 2 \\ 1 \\ 1 \end{bmatrix}} = \overset{C}{\begin{bmatrix} c_1 \\ c_2 \end{bmatrix}} = \begin{bmatrix} 2{,}1 \\ 3{,}0 \end{bmatrix}$$

$$C_1 = 0{,}1 * 1 + \cancel{0{,}3 * 0} + 0{,}7 * 2 + 0{,}4 * 1 + 0{,}2 * 1$$
$$C_1 = 0{,}1 + 1{,}4 + 0{,}4 + 0{,}2$$
$$C_1 = 2{,}1$$

$$C_2 = 0{,}8 * 1 + \cancel{1{,}0 * 0} + 0{,}6 * 2 + 0{,}1 * 1 + 0{,}9 * 1$$
$$C_2 = 0{,}8 + 1{,}2 + 0{,}1 + 0{,}9$$
$$C_2 = 3{,}0$$

The code for this problem is in a python notebook called ”ft8bn-q1.ipynb”

# Problem 2

The code for this problem is in a python notebook called ”ft8bn-q2.ipynb”

# Problem 3

We know that $A \in \mathbb{R}^{n \times n}$ is an orthogonal matrix and $x \in \mathbb{R}^n$ is a vector. Therefore, $Ax \in \mathbb{R}^n$ is a vector.

$$\begin{aligned} \|Ax\|_2^2 &= (Ax)^T(Ax) \\ &= x^T A^T A x \\ &= x^T (A^T A) x \\ &= x^T I x \\ &= x^T x \\ &= \|x\|_2^2 \end{aligned}$$

If we remove the square in both terms of the above equation, we prove that:

$$\|Ax\|_2 = \|x\|_2$$

# Problem 4

A real-world NLP example could be a Text Language Identifier.

1. Event $X$ such that: the text is written in certain language.
2. Sample space of $X \in \{$ English, Spanish, German, French $\}$. The sample space could be greater since there are more than 4 languages in the world, but for simplicity, I just stated 4 languages.
3. Event $X$ can be described with a categorical distribution:

$$P(X = x) = \prod_{k=1}^{4} (\theta_k)^{x_k}$$

   Where:

   - $x_k \in \{0, 1\}$ is an indicator that takes value 1 when a text is in language k and value 0 if it is not in language k.
   - $\{\theta_k\}_{k=1}^{4}$ are the 4 possible languages, which is also the probability of a text of being in language k.

# Problem 5

Two random variables $X$ and $Y$ are independent if:

$$P(X, Y) = P(X) \cdot P(Y)$$

To prove whether these two variables are independent or not, we are going to plug the given values in the above equation:

$$\begin{aligned} P(X = 0, Y = 0) &= P(X = 0) \cdot P(Y = 0) \\ 0.12 &= 0.3 \cdot 0.4 \\ 0.12 &= 0.12 \end{aligned}$$

$$\begin{aligned} P(X = 0, Y = 1) &= P(X = 0) \cdot P(Y = 1) \\ 0.18 &= 0.3 \cdot 0.6 \\ 0.18 &= 0.18 \end{aligned}$$

$$\begin{aligned} P(X = 1, Y = 0) &= P(X = 1) \cdot P(Y = 0) \\ 0.28 &= 0.7 \cdot 0.4 \\ 0.28 &= 0.28 \end{aligned}$$

$$\begin{aligned} P(X = 1, Y = 1) &= P(X = 1) \cdot P(Y = 1) \\ 0.42 &= 0.7 \cdot 0.6 \\ 0.42 &= 0.42 \end{aligned}$$

As we can see, all the equations are valid, therefore variable $X$ and $Y$ are independent variables.

Another way of seeing if two variables are independent is to see if the following rule holds:

$$P(X|Y) = P(X)$$

Now, lets plug the given values in the above equation to see if the rule holds or not:

$$\begin{aligned}
P(X=0|Y=0) &= P(X=0)\\
\frac{P(X=0,Y=0)}{P(Y=0)} &= P(X=0)\\
\frac{0.12}{0.4} &= 0.3\\
0.3 &= 0.3
\end{aligned}$$

$$\begin{aligned}
P(X=0|Y=1) &= P(X=0)\\
\frac{P(X=0,Y=1)}{P(Y=1)} &= P(X=0)\\
\frac{0.18}{0.6} &= 0.3\\
0.3 &= 0.3
\end{aligned}$$

$$\begin{aligned}
P(X=1|Y=0) &= P(X=1)\\
\frac{P(X=1,Y=0)}{P(Y=0)} &= P(X=1)\\
\frac{0.28}{0.4} &= 0.7\\
0.7 &= 0.7
\end{aligned}$$

$$\begin{aligned}
P(X=1|Y=1) &= P(X=1)\\
\frac{P(X=1,Y=1)}{P(Y=1)} &= P(X=1)\\
\frac{0.42}{0.6} &= 0.7\\
0.7 &= 0.7
\end{aligned}$$

# Problem 6

$$\begin{aligned}
l(\theta) &= \sum_{i=1}^{n}\{x^{(i)}\log\theta + (1-x^{(i)})\log(1-\theta)\}\\
\frac{dl(\theta)}{d\theta} &= \sum_{i=1}^{n}\frac{x^{(i)}}{\theta} - \frac{n-x^{(i)}}{1-\theta}
\end{aligned}$$

For simplicity we are going to remove the summation for now, then we equal the derivative of theta to 0 and solve:

$$
\begin{aligned}
\frac{dl(\theta)}{d\theta} &= 0 \\
\frac{x}{\theta} - \frac{n-x}{1-\theta} &= 0 \\
\frac{x}{\theta} &= \frac{n-x}{1-\theta} \\
\frac{1-\theta}{\theta} &= \frac{n-x}{x} \\
\frac{1}{\theta} - 1 &= \frac{n}{x} - 1 \\
\frac{1}{\theta} &= \frac{n}{x} \\
\theta &= \frac{x}{n}
\end{aligned}
$$

Now we add the summation in the term where x is, and we get:

$$
\begin{aligned}
\theta &= \sum_{i=1}^{n} \frac{x^{(i)}}{n} \\
\theta &= \frac{1}{n} \sum_{i=1}^{n} x^{(i)}
\end{aligned}
$$