

Homework 0: Math Review and Why NLP?

Due Feb 15 before class

CS 685, UMass Amherst, Spring 2021

Note

Wikipedia is a useful resource for basic probability, linear algebra, and calculus.

Please write answer in LaTeX (e.g. with Overleaf) and upload a PDF to Gradescope. We will only accept PDF format.

1 Domain of a joint distribution

Say we have a sequence of n binary random variables A_1, A_2, \dots, A_n . How many possible outcomes does the joint distribution $P(A_1, A_2, \dots, A_n)$ define probabilities for?

2 Independence versus Basic Definitions

Say we have three discrete random variables A and B and C . Note that we're using standard probability theory notation where $P(A, B) = P(B, A)$, which simply means the joint probability of both A and B occurring.

2.1

Which of the following statements is always true?

1. $P(A|B) = P(B|A)$
2. $P(A, B) = P(A|B)P(B)$
3. $P(A, B) = P(A)P(B)$
4. $P(A|B) = P(A)$
5. $P(A, B, C) = P(A)P(C)$
6. $P(A, B, C) = P(A)P(B)P(C)$
7. $P(A, B, C) = P(A)P(B|A)P(C|A, B)$
8. $P(A) = \sum_{b \in \text{domain}(B)} P(A, B = b)$
9. $P(A) = \sum_{b \in \text{domain}(B)} P(A|B = b)P(B = b)$
10. $\log(P(A)P(B)) = \log P(A) + \log P(B)$

2.2

Now assume that A , B , and C are all independent of each other. Which of the above statements are now true?

3 Logarithms

The logarithm base doesn't matter for these questions, but in general in this course, assume $\log(x)$ is a natural logarithm.

3.1 Log-probs

Let p be a probability, so it is bounded to $[0, 1]$ (between 0 and 1, inclusive). What is the range of possible values for $\log(p)$? Please be specific about open versus closed intervals.

3.2 Prob ratios

Let p and q both be probabilities. What is the range of possible values for p/q ?

3.3 Log prob ratios

What is the range of possible values for $\log(p/q)$?

4 Linear algebra review

\mathbf{x} is a 10-d real-valued vector (i.e., $\mathbf{x} \in \mathbb{R}^{10}$). \mathbf{y} is another vector of the same dimensionality ($\mathbf{y} \in \mathbb{R}^{10}$). \mathbf{W}_1 is a real-valued matrix of dimensionality 10×10 , and \mathbf{W}_2 is a real-valued matrix of dimensionality 20×10 .

Answer the following questions. Feel free to look at online resources such as Wikipedia for help, and/or additionally test out your answers programmatically using libraries such as numpy.

1. What is the dimensionality of the element-wise product $\mathbf{x} * \mathbf{y}$?
2. What is the dimensionality of the dot product of \mathbf{x} and \mathbf{y} (i.e., $\mathbf{x} \cdot \mathbf{y}$, or $\mathbf{x}^\top \mathbf{y}$, or $\mathbf{x}' \mathbf{y}$ in matrix notation)?
3. What is the dimensionality of the matrix-vector product $\mathbf{W}_1 \mathbf{x}$?
4. What is the dimensionality of $\mathbf{W}_2 \mathbf{y}$?
5. Assume the magnitude of \mathbf{x} is 1 (i.e., $\|\mathbf{x}\| = 1$). What is $\mathbf{x} \cdot \mathbf{x}$?
6. Assume \mathbf{x} and \mathbf{y} are orthogonal, and $\|\mathbf{x}\| = \|\mathbf{y}\| = 1$. What is $\mathbf{x} \cdot \mathbf{y}$?

5 Calculus review

For this problem, assume all variables are scalars (i.e., not vectors or matrices). If you're lost, read up on basic derivatives and the chain rule of calculus!

1. Let's say $L = 0.5(a - b)^2$. What are $\frac{\partial L}{\partial a}$ and $\frac{\partial L}{\partial b}$? In other words, what are the partial derivatives of L with respect to a and b ?
2. Let's say $a = \log(3mn)$. What is $\frac{\partial L}{\partial m}$?

6 Why NLP?

Why are you interested in natural language processing? Write as much as you wish.