CSE 325/425 (Spring 2021) Homework 1

Due on 11:55pm, Feb 17, 2021

Grading: All questions have the same points (25 each). We will randomly grade some of the questions.

Submitting: Only electronic submissions on Coursesite are accepted. You can handwrite your answers on papers and then scan them to images. If you need to plot figures using a computer, the plotted files should be saved and included in the submitted pdf file. Submit a single pdf file named

<Your LIN>HW1.pdf

Other format will not be accepted.

Questions:

1. Verify that the Laplacian smoothing for a unigram language model leads to a probability distribution over the words in the vocabulary V.

   [[[Laplacian smoothing for a unigram language model uses <==

   $$\Pr(w_i) = \frac{C(w_i)+1}{\sum_{w\in V} C(w) + |V|} > 0. \tag{1}$$

   Summing over all words in the vocabulary, we have

   $$\sum_{w_i} \Pr(w_i) = \frac{\sum_{w_i\in V}[C(w_i)+1]}{\sum_{w\in V} C(w)+|V|} = \frac{\sum_{w\in V} C(w)+|V|}{\sum_{w\in V} C(w)+|V|} = 1. \tag{2}$$

   ]]]

2. Using the bi-gram language model, write down the likelihood of a corpus, considered as a single long sequence $[w_1, \ldots, w_n]$ of n words sampled from V.

   [[[<==

   $$\Pr([w_1, \ldots, w_n]) = \Pr(w_1|\emptyset)\Pr(w_2|w_1)\ldots\Pr(w_n|w_{n-1}). \tag{3}$$

   ]]]

3. Find the gradient of the multi-variate objective function $f(x, y) = x^2 - y^2$ with respect to the vector $[x, y]$.

   [[[<==

   $$\frac{\partial}{\partial x} = 2x. \tag{4}$$

   $$\frac{\partial}{\partial y} = -2y. \tag{5}$$

   Therefore, the gradient is the vector

   $$\begin{bmatrix} 2x, \\ -2y \end{bmatrix} \tag{6}$$

   ]]]

4. Continuing the above question, use the gradient at the current location $[2, 1]$ to move the parameter to the next location using learning rate 0.1. Is the function value increased or decreased?

   [[[The gradient at the location $[2, 1]$ is $[4, -2]$, after plugging the location in the gradient. <==

   Using the learning rate 0.1, the new location is

$$[2, 1] + 0.1 \times [4, -2] = [2.4, 0.8]. \tag{7}$$

   (The answer $[2, 1] - 0.1 \times [4, -2] = [1.6, 1.2]$ is also accepted.)

   ]]]