Introduction to Deep Learning for Computer Vision Assignment 1: Preliminaries

Parm Johal
V00787710

January 14, 2019

Abstract

This report deals with preliminaries of the CSC486b course of Computer Vision. The first part deals with simple math while the second part deals with programming from a machine learning aspect.

2 Preliminaries: math

2.1 Basic Calculus

Question 2.1.1

$y = ax^2 + bx + c$

$\frac{dy}{dx} = 2ax + b$

Question 2.1.2

$y = sin(x)cos(x)$

$\frac{dy}{dx} = cos2(x) - sin2(x)$

Question 2.1.3

$y = \frac{1}{1 + e^{-x}} = (1 + e^{-x})^{-1}$

$\frac{dy}{dx} = -(1 + e^{-x})^{-2} * \frac{d}{dx}(1 + e^{-x})$

Question 2.1.4

$$y = \frac{e^x - e^{-x}}{e^x + e^{-x}} = (e^x - e^{-x})(e^x + e^{-x})^{-1}$$

$$\frac{dy}{dx} = -(e^x - e^{-x})(e^x + e^{-x})^{-2}(e^x - e^{-x}) + (e^x - e^{-x})^{-1}(e^x + e^{-x})$$

$$\frac{dy}{dx} = \frac{-(e^x - e^{-x})^2}{(e^x + e^{-x})^2} + 1$$

2.2 Taylor Expansion

$$note : y = f(x) = f(a) + f'(a)(x-a) + f''(a)\frac{(x-a)^2}{2}$$

Question 2.2.1

$$y = e^{ax+b}$$

$$f(x) = e^{ax+b} + be^{ax+b}(x-a) + b^2e^{ax+b}\frac{(x-a)^2}{2}$$

$$f(0) = e^b + be^b(-a) + \frac{b^2e^b(-a)^2}{2}$$

$$= e^b(1 - ab + \frac{(ab)^2}{2})$$

Question 2.2.2

$$y = cos(ax+b)$$

$$f(x) = cos(ax+b) - asin(ax+b)(x-a) - \frac{a^2}{2}cos(ax+b)(x-a)^2$$

$$f(0) = cos(b) + a^2sin(b) - \frac{a^4}{2}cos(b)$$

2.3 Matrix Multiplication

Question 2.3.1

$$M = \begin{bmatrix} 9 & 8 & 7 \\ 6 & 5 & 4 \\ 3 & 2 & 1 \end{bmatrix} \times \begin{bmatrix} -1 \\ 2 \\ -1 \end{bmatrix}$$

$$M = \begin{bmatrix} 9 \times -1 + 8 \times 2 + 7 \times -1 \\ 6 \times -1 + 5 \times 2 + 4 \times -1 \\ 3 \times -1 + 2 \times 2 + 1 \times -1 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix}$$

Question 2.3.2

$$M = \begin{bmatrix} -2 \\ 1 \\ -2 \end{bmatrix} \times \begin{bmatrix} 1 & -2 & 1 \end{bmatrix}$$

$$M = \begin{bmatrix} -2 \times 1 & -2 \times -2 & -2 \times 1 \\ 1 \times 1 & 1 \times -2 & 1 \times 1 \\ -2 \times 1 & -2 \times -2 & -2 \times 1 \end{bmatrix} = \begin{bmatrix} -2 & 4 & -2 \\ 1 & -2 & 1 \\ -2 & 4 & -2 \end{bmatrix}$$

2.4 Applying Chain Rule on Vectors and Matrices

$$Let A = \begin{bmatrix} a_1 & a_2 & a_3 \\ a_4 & a_5 & a_6 \\ a_7 & a_8 & a_9 \end{bmatrix}, x = \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix}, b = \begin{bmatrix} b_1 \\ b_2 \\ b_3 \end{bmatrix}$$

After combining the matrices, we end up with

$$\begin{bmatrix} a_1x_1 + a_2x_2 + a_3x_3 - b_1 \\ a_4x_1 + a_5x_2 + a_6x_3 - b_2 \\ a_7x_1 + a_8x_2 + a_9x_3 - b_3 \end{bmatrix}$$

Which we'll refer to as **Z**.

Since we know that **Z^TZ = ZZ** (dot product) we have

$$y = (a_1x_1 + a_2x_2 + a_3x_3 - b_1)^2 + (a_4x_1 + a_5x_2 + a_6x_3 - b_2)^2 + (a_7x_1 + a_8x_2 + a_9x_3 - b_3)^2$$

Which then we get

$$\frac{dy}{dx} = 2a_1(a_1x_1 + a_2x_2 + a_3x_3 - b_1) + 2a_2(a_4x_1 + a_5x_2 + a_6x_3 - b_2) + 2a_3(a_7x_1 + a_8x_2 + a_9x_3 - b_3)$$

3 Preliminaries: Programming

Refer to the Source code alongside this report in the directory.