**COMP5212: Machine Learning** **Fall 2023**

Homework 1: Due Sunday Oct. 1, 11:59 PM

**Instructions**: upload a PDF report using LaTeX containing your answers to Canvas (remember to include your name and ID number).

# Problem 1. Sigmoid function in logistic regression

Let $g(z) = \frac{1}{1+e^{-z}}$ be the sigmoid activation function

(a) (10 pt) Show that $\frac{\partial g}{\partial z} = g(z)(1 - g(z))$

(b) (10 pt) Show that $1 - g(z) = g(-z)$

# Problem 2. Convexity

(a) (15 pt) Assume that $f : \mathbb{R}^d \to \mathbb{R}$ can be written as $f(\boldsymbol{w}) = g(\langle \boldsymbol{w}, \boldsymbol{x} \rangle + y)$, for some $\boldsymbol{x} \in \mathbb{R}^d$, $y \in \mathbb{R}$, and $g : \mathbb{R} \to \mathbb{R}$. Prove $f$ is convex if $g$ is convex.

(b) (15 pt) For $i = 1, \ldots, r,$, let $f_i : \mathbb{R}^D \to \mathbb{R}$ be a convex function. Prove the $g(x) = \max_{i \in [r]} f_i(x)$ from $\mathbb{R}^d$ to $\mathbb{R}$ is also convex.

# Problem 3. Smoothness

A differential function $f$ is said to be $L$-smooth if its gradietns are Lipschitz continuous, that is

$$\|\nabla f(x) - \nabla f(y)\| \leq L\|x - y\|$$

let $f : \mathbb{R}^d \to \mathbb{R}$ be a twice differentiable function. If $f$ is $L$-smooth then prove the following inequality:

- (25 pt) Prove $\langle \nabla^2 f(x)v, v \rangle \leq L\|v\|_2^2, \quad \forall x, v \in \mathbb{R}^d$
- (25 pt) Prove $f(y) \leq f(x) + \langle \nabla f(x), y - x \rangle + \frac{L}{2}\|y - x\|_2^2$