HOMEWORK 1.   GRADIENT DESCENT

Jinghua Huang   Kun Yuan

September 18, 2023

Attention: Turn in your homework at the beginning of our lecture on Sep. 26, 2023

1   Linear regression

Consider the following linear regression problem

$$\min_{x \in \mathbb{R}^d} \quad f(x) = \frac{1}{2}\|Ax - b\|^2 \tag{1}$$

where $A \in \mathbb{R}^{n \times d}$ has full column rank. Please answer the following questions:

- Prove $f(x)$ is a convex function.

- Calculate $\nabla f(x)$.

- Prove $f(x)$ is L-smooth, and calculate the value of L.

- Prove $f(x)$ is μ-strongly convex, and calculate the value of μ.

- Write down the gradient descent recursion to solve problem (1).

2   Logistic regression

Consider the following linear regression problem

$$\min_{x \in \mathbb{R}^d} \quad f(x) = \frac{1}{N}\sum_{i=1}^{N} \log(1 + \exp(-b_i a_i^T x)) \tag{2}$$

where $b_i \in \{+1, -1\}$ and $a_i \in \mathbb{R}^d$ are given data samples for any $i \in \{1, \cdots, N\}$.

- Prove $f(x)$ is a convex function.

- Calculate $\nabla f(x)$.

- Write down the gradient descent recursion to solve problem (2).

3 Forward backward propagation

Consider the following 3 layer linear neural network

$$\min_{\{W_1, W_2, W_3\}} \quad f(W_1, W_2, W_3; a, b) = \frac{1}{2}\|W_3 W_2 W_1 a - b\|^2 \tag{3}$$

where $a \in \mathbb{R}^d$, $b \in \mathbb{R}^m$, $W_1 \in \mathbb{R}^{q \times d}$, $W_2 \in \mathbb{R}^{p \times q}$, and $W_3 \in \mathbb{R}^{m \times p}$. Calculate $\frac{\partial f}{\partial W_1}$, $\frac{\partial f}{\partial W_2}$, and $\frac{\partial f}{\partial W_3}$ respectively.