Problem 2.1. a) 1/1

Q: Is $f(w) = \frac{1}{N} \sum_{i \in [N]} f_i(w) + \lambda \|w\|_2^2$ Lipschitz cont?
If so, find the smallest B.

$f_i(w) = \log\left(1 + e^{-y_i x_i^T w}\right)$.

A: Solution: We will show that $\|\nabla f(w)\| \le B$ for $\|w\| \le D$

- $\nabla f(w) = \frac{1}{N} \sum_i \nabla f_i(w) + 2\lambda w$.
- $\nabla f_i(w) = \left(1 + e^{-y_i x_i^T w}\right)^{-1} e^{-y_i x_i^T w} \cdot (-y_i x_i)$

$$= \left(e^{y_i x_i^T w} + 1\right)(-y_i x_i) = -\frac{y_i}{1 + e^{y_i x_i^T w}} x_i$$

$$\Rightarrow \nabla f(w) = \frac{1}{N} \sum_{i \in [N]} -\frac{y_i}{1 + e^{y_i x_i^T w}} x_i + 2\lambda w.$$

$$\|\nabla f(w)\| = \left\| \frac{1}{N} \sum_i -\frac{y_i}{1 + e^{y_i x_i^T w}} x_i + 2\lambda w \right\|$$

triangular ineq. $\rightarrow$

$$\le \left\| \frac{1}{N} \sum_i \frac{-y_i}{1 + e^{y_i x_i^T w}} x_i \right\| + 2\lambda \|w\|$$

$$\le \frac{1}{N} \sum_i \left\| -\frac{y_i}{1 + e^{y_i x_i^T w}} x_i \right\| + 2\lambda \|w\|$$

$$= \frac{1}{N} \sum_i \frac{|y_i| \|x_i\|}{1 + e^{y_i x_i^T w}} + 2\lambda \|w\|$$

{using that $-|y_i| \|x_i\| \|w\| \le x_i^T w$ from Cauchy-Schwarz, (By $|y_i x_i^T w| \le |y_i| \|x_i\| \|w\|$) we get}

$$\le \frac{1}{N} \sum \frac{|y_i| \|x_i\|}{1 + e^{-|y_i| \|x_i\| \|w\|}} + 2\lambda \|w\|$$

{For $\|w\| \le D$, we have}

$$\le \underbrace{\frac{1}{N} \sum \frac{|y_i| \|x_i\|}{1 + e^{-|y_i| \|x_i\| D}} + 2\lambda D}_{\text{is the smallest } B \text{ such that } \|\nabla f(w)\| \le B,\ \|w\| \le D.} \le \frac{1}{N} \sum_i |y_i| \|x_i\| + 2\lambda D$$