Instructor: Vishnu Boddeti
CSE 891-001: Deep Learning
Homework 2
October, 11

1 Backpropagation (2pts)

1. $\frac{\partial \mathcal{L}}{\partial x} = (I_N + W^{(1)^T} diag(\sigma'(x)) W^{(2)^T})(y - s) + W^{(1)^T} diag(\sigma'(x)) r$

3 Linear Regression (3pts)

1. **Deriving the Gradient (0.5pt)** $\frac{\partial L}{\partial W} =$.
2. **Underparameterized Model (0.5pt)**
   (a) $w = (X^T X) X^T t$. Yes, the solution is unique.
3. **Overparameterized Model: 2D Example (1pt)**
   (a) $2w_1 + w_2 = 2$
4. **Overparameterized Model: General Case (1pt)**
   (a) $w = X^T (X X^T t)$

3 Linear Regression and Optimization (3pts)

1. **Stochastic Gradient Descent (1pt):** Yes
2. **Mini-Batch SGD (1pt):** Yes
3. **Adaptive Methods (1pt):** No