Assignment 1: CS 224d (Jerrick Hoang)

Due: Wednesday

1. Softmax is translational invariant $\rho(\boldsymbol{x}+c)_j = \frac{\exp(x_j+c)}{\sum_{k=1}^{\boldsymbol{K}} \exp(x_k+c)} = \frac{\exp(x_j)\exp(c)}{\exp(c)\sum_{k=1}^{\boldsymbol{K}} \exp(x_k)} = \frac{\exp(x_j)}{\sum_{k=1}^{\boldsymbol{K}} \exp(x_k)} = \rho(\boldsymbol{x})$
2. (a) Let $\rho(\boldsymbol{x}) = \frac{1}{1+e^{-\boldsymbol{x}}}$ Then $\frac{\partial \rho}{\partial \boldsymbol{x}} = \frac{-1}{(1+e^{-\boldsymbol{x}})^2} \frac{d(1+e^{-\boldsymbol{x}})}{d\boldsymbol{x}} = \frac{e^{-\boldsymbol{x}}}{(1+e^{-\boldsymbol{x}})^2} = (1-\rho(\boldsymbol{x})\rho(\boldsymbol{x})$