Mathematical exercise: Multilayer-Perceptrons

June 5, 2020

1 Compute the gradient of the following composite function and draw a computational graph of the forward propagation

Given a single-hidden-layer neural network regression model f(x) and a squared loss l,

$$f(\mathbf{x}) = \mathbf{w}_2^\top \pi(\mathbf{W}_1^\top \mathbf{x}),$$

where $\mathbf{x} \in \mathbb{R}^d$ is a feature vector, $\mathbb{W}_1 \in \mathbb{R}^{d \times k}$ matrix of weights, and $\precsim_2 \in \urcorner$ is a weight vector. Further, $\pi : \mathbb{R}^k \to \mathbb{R}^k$ is the logistic sigmoid function applied element-wise as

$$\pi(x) = \frac{1}{1 + e^{-x}}.$$

The loss l of the network is a squared error

$$l = \frac{1}{2}(y - f(\mathbf{x}))^2.$$

a) Draw a computational graph of all operations in this function and
b) compute the gradient matrix $\mathbf{G}_1$ of l with respect to $\mathbf{W}_1$.
c) compute the gradient vector $\mathbf{g}_2$ of l with respect to $\mathbf{w}_2$.
d) If we add ℓ_2 regularization on $\mathbf{w_2}$, the Loss L becomes

$$l + \|\mathbf{w}_2\|_2^2$$

For this loss, the gradient vector with respect to $\mathbf{w}_2$ becomes $\tilde{\mathbf{g}}_2$. Note that you can use $\mathbf{g}_2$ as the solution of c), in case you cannot derive the solution. Calculate the gradient.