Homework 2

Samuel Sledzieski

March 19, 2018

1 Gradient Derivation

$$LL(\mathbf{w}) = \sum_{i=1}^{n} y_i \mathbf{w}^T \mathbf{x}_i - \log\left(1 + \exp\left(\mathbf{w}^T \mathbf{x}_i\right)\right) \tag{1}$$

$$\frac{\delta LL}{\delta \mathbf{w}} = \frac{\delta}{\delta \mathbf{w}} \sum_{i=1}^{n} y_i \mathbf{w}^T \mathbf{x}_i - \log\left(1 + \exp\left(\mathbf{w}^T \mathbf{x}_i\right)\right) \tag{2}$$

$$\frac{\delta LL}{\delta \mathbf{w}} = \sum_{i=1}^{n} y_i \mathbf{x}_i - \frac{\delta}{\delta \mathbf{w}} \log\left(1 + \exp\left(\mathbf{w}^T \mathbf{x}_i\right)\right) \tag{3}$$

$$\frac{\delta LL}{\delta \mathbf{w}} = \sum_{i=1}^{n} y_i \mathbf{x}_i - \frac{\frac{\delta}{\delta \mathbf{w}} \exp\left(\mathbf{w}^T \mathbf{x}_i\right)}{\left(1 + \exp\left(\mathbf{w}^T \mathbf{x}_i\right)\right)} \tag{4}$$

$$\frac{\delta LL}{\delta \mathbf{w}} = \sum_{i=1}^{n} y_i \mathbf{x}_i - \frac{\mathbf{x}_i \exp\left(\mathbf{w}^T \mathbf{x}_i\right)}{\left(1 + \exp\left(\mathbf{w}^T \mathbf{x}_i\right)\right)} \tag{5}$$

2 Logistic Regression Classifier

Files The attached tarball contains 4 files.

- logreg_classifier.py
- roc.py
- main.py
- Runtime.ipynb

logreg_classifier.py contains the Logistic Regression Classifier. roc.py contains function for plotting ROC. main.py loads Wisconsin Breast Cancer dataset from sklearn and divides it into a training set and test set. It then fits the logistic regression classifier using the training set, predicts probabilities on the test set, and plots the ROC curve. Runtime.ipynb is a Jupyter notebook which performs the same functions as main.py, then fits an sklearn logistic regression classifier and plots an ROC curve on the same datasets.

ROC Curve Below is the ROC curve of my logistic regression classifier on the Wisconsin Breast Cancer dataset