

## Homework 1

Due: April 10, 2020, 11:59PM PT

Student Name:

Instructor Name: John Lipor

**Problem 1 Multiclass Ridge Regression Classifier** (5 pts each)

In this problem, you will derive and implement a multiclass linear classifier based on ridge regression. Suppose you are given  $m$  training examples  $\{(x_i, y_i)\}$ , from  $K$  classes, where  $x_i \in \mathbb{R}^d$  and  $y_i \in \{e_1, \dots, e_K\}^1$  is a one-hot vector. One way to perform classification is to find the matrix  $W \in \mathbb{R}^{d \times K}$  that minimizes the least squares cost function given below. You can then estimate the *integer-valued* class labels as

$$\hat{y}_i = \arg \max_{k \in [K]} W^T x_i.$$

- (a) Derive a closed-form solution to the problem

$$\widehat{W} = \arg \min_{W \in \mathbb{R}^{d \times K}} \sum_{i=1}^m \|W^T x_i - y_i\|_2^2 + \lambda \|W\|_F^2,$$

where  $\lambda > 0$  is a regularization parameter.

- (b) Implement your trained classifier with  $\lambda = 10^{-4}$  on the full MNIST dataset included in the homework files. You may not use `sklearn` or other high-level libraries to perform the one-hot encoding, but you may use `numpy` and `scipy`. **Turn in** your code, as well as the classification error on the training and test sets.

**Problem 2 DSS: Using scikit-learn** (10 pts)

(*DSS rules apply.*) One important part of data science is fluency with popular libraries. One library that is useful for a variety of machine learning tasks (especially those not related to deep learning) is `scikit-learn`, a.k.a., `sklearn`. Use `sklearn` to perform multiclass ridge regression classification, taking special care to make sure the regularization parameter  $\lambda$  and the use of an offset are the same. **Turn in** your code, an explanation of what functions and options you used to perform the multiclass classification, as well as the classification error on the training and test sets.

**Problem 3 SLT** (5 pts)

(*SLT rules apply.*) UML, Ch. 2, Exercise 2. State how long you worked on the problem before looking at the solution.

**Problem 4 SLT** (5 pts each)

(*SLT rules apply.*) UML, Ch. 2, Exercise 3.1-3.2. State how long you worked on the problem before looking at the solution.

---

<sup>1</sup>The vector  $e_k$  is the  $k$ th standard basis vector, taking zero everywhere except in the  $k$ th element, where it takes the value one.