

## ECE 805 – Machine Learning

### Tutorial 2, Monday, January 30, 2023

---

1. In this method, we will minimize  $J$  by explicitly taking its derivatives with respect to the  $\theta_j$ 's and setting them to zero. Let's introduce some notation for doing calculus with matrices.

$$\text{Training set: } x = \begin{bmatrix} (x^1)^T \\ (x^2)^T \\ \vdots \\ (x^N)^T \end{bmatrix} \quad \text{Target values: } y = \begin{bmatrix} y^1 \\ y^2 \\ \vdots \\ y^N \end{bmatrix}$$

- a. We aim to minimize  $J(\theta)$  with respect to  $\theta$ ,

$$\text{where } \theta = \begin{bmatrix} \theta_1 \\ \theta_2 \\ \vdots \\ \theta_N \end{bmatrix} \text{ and } J(\theta) = \frac{1}{2} \sum_{i=1}^m (h_{\theta}(x^i) - y^i)^2.$$

- b. Using the normal equations that you found in part (a), please find  $\theta'$  s for minimizing

$$J(\theta) \text{ for the following points: } x = \begin{bmatrix} 1 & 1 \\ 1 & 2 \\ 2 & 2 \\ 2 & 3 \end{bmatrix}, y = \begin{bmatrix} 6 \\ 8 \\ 9 \\ 11 \end{bmatrix} \text{ which are part of the line } y = x_0 + 2x_1 + 3.$$

2.

- A. Linear Regression example (in Python) using the diabetes dataset from python datasets, which includes 10 features and 442 samples. You can use only one feature and linear regression model.
  - a. without regularization
  - b. using L2 regularization – Ridge
  - c. using L1 regularization – Lasso
- B. Using the wine dataset which consists of 178 samples of 13 features, extract the mean square error (MSE) for:
  - a. Linear regression (without regularization)
  - b. Linear Regression with L2 regularization – Ridge
  - c. Linear Regression with L1 regularization – Lasso
  - d. SGD Regressor – L2 regularization
  - e. SGD Regressor – L1 regularization.

Compare the results between Linear Regression model and SGDRegressor in Python and discuss the differences.