

**CSE 404: Introduction to Machine Learning (Fall 2019)**

Homework #4

Due 10/4/2019 11:59PM

Note: (1) LFD refers to the textbook “Learning from Data”.

1. (20 points) Recall the objective function for linear regression can be expressed as

$$E(w) = \frac{1}{N} \|Xw - y\|^2,$$

as in (3.3) of LFD. Minimizing this function with respect to  $w$  leads to the optimal  $w$  as  $(X^T X)^{-1} X^T y$ . This solution holds only when  $X^T X$  is nonsingular. To overcome this problem, the following objective function is commonly minimized instead:

$$E_2(w) = \|Xw - y\|^2 + \lambda \|w\|^2,$$

where  $\lambda > 0$  is a user-specified parameter. Please do the following:

- (a) (10 points) Derive the optimal  $w$  that minimize  $E_2(w)$ .
  - (b) (10 points) Explain how this new objective function can overcome the singularity problem of  $X^T X$ .
2. (20 points) Exercise 3.3 in LFD (page 87).