CSC 665: Artificial Intelligence

Homework 5

By turning in this assignment, I agree to abide by SFSU's academic integrity code and declare that all of my solutions are my own work.

1 Regularization

In this homework, you will reproduce the experimental results shown on the slides entitled "The sales pitch" from the 5/7 lecture.

First recall the setup. Our target function is $f(x) = \sin(\pi x)$ for $x \in [-1, 1]$, meaning that this is a regression problem. We assume our training dataset contains just $n = 2$ examples sampled uniformly at random from $[-1, 1]$. There is no noise in this problem, so training examples take the form $(x, f(x))$.

We will train a linear regression model using gradient descent on the sum-of-squared-errors cost function, both with and without regularization.

a. (10 points) The sum-of-squared-errors cost function is

$$\begin{aligned} C(w) &= \sum_{i=1}^{n} (y_i - (w_0 + w_1 x_i))^2 \\ &= (y_1 - (w_0 + w_1 x_1))^2 + (y_2 - (w_0 + w_1 x_2))^2, \end{aligned}$$

since $n = 2$. Compute $\frac{d}{dw_0} C(w)$ and $\frac{d}{dw_1} C(w)$.

b. (10 points) If we add an ℓ_2 regularization term to our cost function, we obtain the augmented cost function

$$\begin{aligned} \tilde{C}(w) &= \sum_{i=1}^{n} (y_i - (w_0 + w_1 x_i))^2 + \lambda \sum_{j=0}^{p} w_j^2 \\ &= (y_1 - (w_0 + w_1 x_1))^2 + (y_2 - (w_0 + w_1 x_2))^2 + \lambda(w_0^2 + w_1^2). \end{aligned}$$

Compute $\frac{d}{dw_0} \tilde{C}(w)$ and $\frac{d}{dw_1} \tilde{C}(w)$.

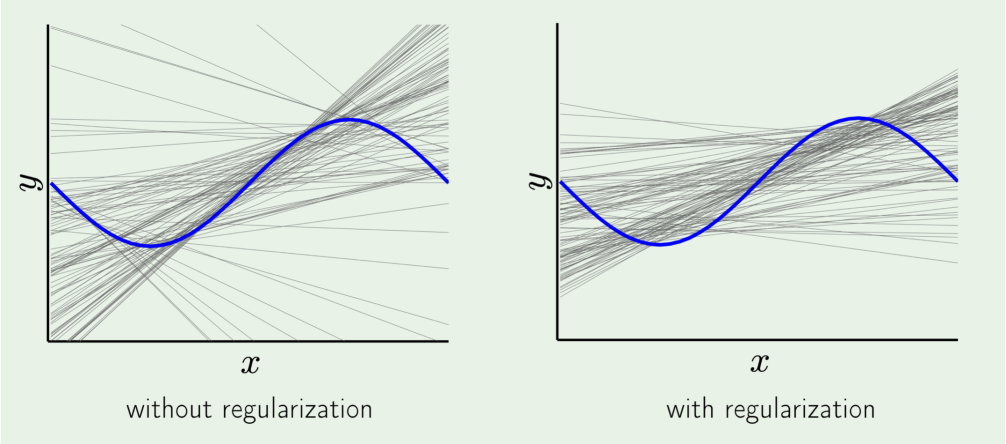
c. (10 points) Implement the function `fit_without_reg` which finds values of w_0 and w_1 that minimize $C(w)$. Note that there is a unique line that perfectly fits two given points, and you should find a formula for this line rather than perform gradient descent on $C(w)$.

d. (10 points) Implement the function `fit_with_reg` which finds values of w_0 and w_1 that minimize $\tilde{C}(w)$. You should use gradient descent with a step size of $\eta = 0.05$, setting $\lambda = 1$, and making 1000 gradient descent updates starting from $w_0 = w_1 = 0$.

e. (10 points) Perform the following experiment 1000 times. Sample two points from f using the provided function `generate_training_examples`. Call `fit_without_reg` and `fit_with_reg` (using $\lambda = 1$) on

these points to find values for w_0 and w_1 with and without regularization. Finally, call `test_error` to estimate the out-of-sample error of these two hypotheses. Report the two test errors you find, averaged over the 1000 trials of this experiment.

f. (10 points *extra credit*) Create a version of this figure from the slides by plotting the line (use a thin line width and low alpha value) corresponding the values you obtain for w_0 and w_1 in each of the 1000 trials you carry out in part (e). Include the target function f in your plots as shown below.



Submission

Submission is done on Canvas. You should submit two files: one containing your solutions to the written problems and one for the coding problems.

- Submit your written solutions in a single PDF file with your name at the top. Make sure to clearly indicate the number and letter of the problem corresponding to each solution. It is okay to hand-write your solutions and then scan them into a PDF, but *only if your handwriting is legible.*
- Submit your coding solutions in a file named `regularization.py`.