

Homework 2

Handed Out: September 13

Due: September 27, 8:00 p.m.

- You are encouraged to format your solutions using \LaTeX . Handwritten solutions are permitted, but remember that you bear the risk that we may not be able to read your work and grade it properly — we will not accept post hoc explanations for illegible work. You will submit your solution manuscript for written HW 2 as a single PDF file.
- The homework is **due at 8:00 PM** on the due date. We will be using Gradescope for collecting the homework assignments. Please submit your solution manuscript as a PDF file via Gradescope. Post on Ed Discussion and contact the TAs if you are having technical difficulties in submitting the assignment.

1 Multiple Choice & Written Questions

Note: You do not need to show work for multiple choice questions. If formatting your answer in \LaTeX , use our LaTeX template [hw2_template.tex](#) (This is a read-only link. You'll need to make a copy before you can edit. Make sure you make only private copies.).

1. [Bias-Variance Tradeoff] (5pts) Suppose we have an L_2 -regularized linear regression model, which has loss $L(\beta) = \frac{1}{n} \sum_{i=1}^n (f_\beta(x_i) - y_i)^2 + \lambda \|\beta\|_2^2$. For each of the following, indicate whether it tends to increase bias, decrease bias, or keep bias the same, and similarly for variance:
 - A) Increase the number of training examples n
 - B) Increase the regularization parameter λ
 - C) Increase the dimension d of the features $\phi(x) \in \mathbb{R}^d$
 - D) Increase c , where we replace the features $\phi(x)$ with $c \cdot \phi(x)$, for some $c \in \mathbb{R}_{>0}$ (for this part, assume no regularization, i.e., $\lambda = 0$)
 - E) Increase the gradient descent step size α (but not so much that gradient descent diverges)

In addition, suppose you fit a model and find that it has low loss on the training data but high loss on the test data; for each of the above five values n , λ , d , c , and α , indicate whether you should increase or decrease it to reduce the test loss, or it has no impact on the test loss.

2. [Regularization/Sparsity] (4 pts) In class, we demonstrated the intuition behind L_1 and L_2 regularization through two dimensional ellipsoid visualizations. In this problem we will take a different angle, and try to see why L_1 regularization helps create sparsity from the perspective of gradient descent.

- A) (2pts) Write down the partial derivative of the L_1 and L_2 regularization terms with respect to some weight β_j (you can ignore the case $\beta_j = 0$ where the gradient may be undefined). [Recall that the L_1 regularization term is $R_1(\beta) = \lambda \sum_{j'=1}^d |\beta_{j'}|$, and the L_2 regularization term is $R_2(\beta) = \lambda \sum_{j'=1}^d \beta_{j'}^2$.]
- B) (2pts) Based on these results, which regularizer does a better job “pushing” β_j to zero? [Hint: Consider what happens when β_j is already small. For simplicity, you can assume that the partial derivative $\frac{\partial}{\partial \beta_j}$ of the MSE term is zero.]
3. [Linear Regression] (4 pts) Suppose the true function we are trying to approximate is $y = |x|$. Furthermore, suppose we use unregularized linear regression with the MSE loss function to fit a function $y = ax + b$.
- A) (2pts) Suppose we train on points $x \sim \text{Uniform}([0, 1])$ sampled uniformly at random from $[0, 1]$. Then, what is the learned model? Write out the coefficients a and b . In addition, what is the MSE on points $x \sim \text{Uniform}([-1, 0])$? [Hint: To generalize the MSE to a distribution of samples, you will need to compute an expectation, which involves an integral.]
- B) (2pts) Suppose we train on points $x \in \text{Uniform}([-1, 0])$. Then, what is the learned model? Write out the coefficients a and b . Also, what is the MSE on points $x \sim \text{Uniform}([0, 1])$?

2 Python Programming Questions

A IPython notebook is linked on the class website. It will tell you everything you need to do, and provide starter code. Remember to include the plots and answer the questions in your written homework submission!