Take-home Quiz 476/676

Due: end of the day April 13, 2021

15+5 (bonus).

Submission: Directly write in this word document. No code is required.

Name: Yunting Chiu

Email: [yc6705a@american.edu](mailto:yc6705a@american.edu)

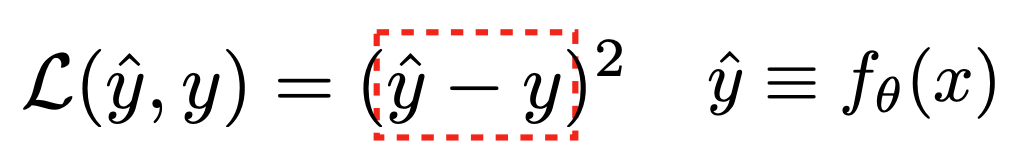
Basic machine learning quiz. The goal of the quiz is to push you read the assigned reading and slides.

You can find most of the answers in slides and readings.

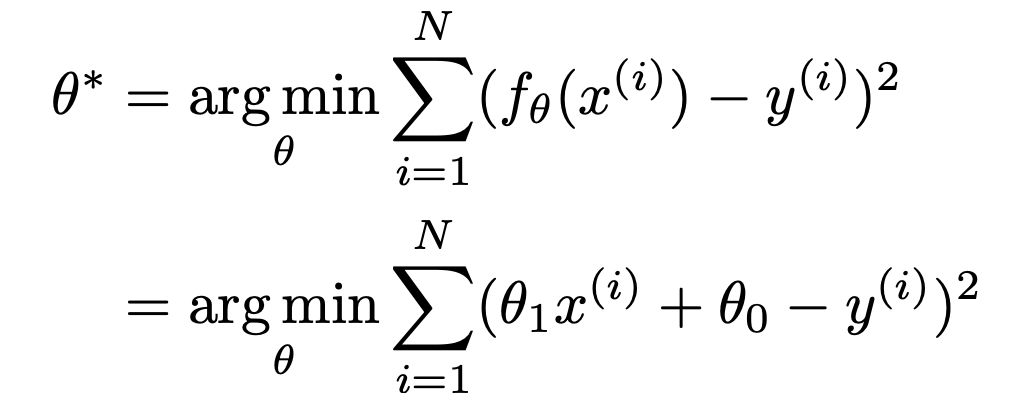
1. (5pts) What are the objective function, hypothesis space, and optimizer in the least-square algorithm? You can write as much equation as you wish or describe with graphs. Mathematical equations are more preferred than words. What is the purpose of regularization in a least-square algorithm? Give an example if possible.

**Objective Function:** the object of least-square algorithm is to search for the parameters, , that best fit the data, which is to minimize the squared error between predictions and values.

The lagrangian function is as follows:

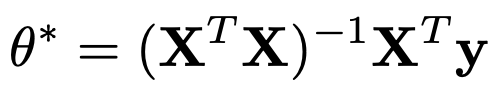


And we take the derivative of the function and set that to zero to find the optimal parameters:



**Hypothesis Space:** the relationship between X and Y is linear:

**Optimizer:** An optimizer helps minimize the objective function with respect to the parameters. The optimizer of the least-square algorithm is as follows:



**The purpose of regularization** in a least-square algorithm is to address overfitting problems. It is a constraint on the algorithm solution that it adds a penalty term to the error when we add an extra coefficient. In regularized least squares we try to minimize

regularized least squares

1. (5pts) True or False?
2. If X is a loss function, it should increase as our model get worse.

True

1. A model that is improving should assign higher and higher probability to the correct answer.

True

1. K in KNN model (K nearest Neighbor) is a hyper parameter that can be learned. True or False?

True

1. We choose hyperparameters using the validation set and run the test only one.

False

1. When we overfit a model, we tend to have lower training error and higher validation error.

True

(A model that is underfit will have high training and high testing error while an overfit model will have extremely low training error but a high testing error).

1. (5pts) We are performing a 3-way classification of images using linear classier with support vector machine loss (SVM loss). The following graph shows the scores for classification results and the equation of the loss function. Please compute the loss for each class and write down the results.

A picture containing graphical user interface

Description automatically generated

What happens to the loss if the score for car becomes larger or smaller?

The losses for each animal originally

Losses of cat = max(0, 5.1-3.2 +1) + max(0, -1.7-3.2 + 1) = 2.9

Losses of car = max(0, 1.3 - 4.9 +1) + max(0, 2.0-4.9 + 1) = 0

Losses of frog = max(0, 2.2 + 3.1 +1) + max(0, 2.5 + 3.1 + 1) = 12.9

Set the car score to 100. As the car's score becomes larger, we can see that the loss function for other animals becomes larger as well.

Losses of cat = max(0, 100-3.2 +1) + max(0, -1.7-3.2 + 1) = 97.8

Losses of car = max(0, 1.3 - 100 +1) + max(0, 2.0-100 + 1) = 0

Losses of frog = max(0, 2.2 + 3.1 +1) + max(0, 100 + 3.1 + 1) = 110.4

Set the car score to -100. We can see that as the car’s score decreases, so does the loss function for other animals. However, the losses of the car’s score become large.

Losses of cat = max(0, -100-3.2 +1) + max(0, -1.7-3.2 + 1) = 0

Losses of car = max(0, 1.3 + 100 +1) + max(0, 2.0 + 100 + 1) = 205.3

Losses of frog = max(0, 2.2 + 3.1 +1) + max(0, -100 + 3.1 + 1) = 6.3

What happens to the function if we use mean instead of sum?

Let we try to use mean instead of summation:

Losses of cat = mean(0, 5.1-3.2 +1, -1.7-3.2 + 1) = 0

Losses of car = mean(0, 1.3 - 4.9 +1, 2.0-4.9 + 1) = 0

Losses of frog = mean(0, 2.2 + 3.1 +1,2.5 + 3.1 + 1 ) = 0

The result shows that the loss function will be all zero if we use mean, meaning that model production will deteriorate because it is illogical.

What happens to the loss if it is changed from “1” to “10”? Write down the results

Higher loss is the bad prediction for any model so changing the loss from 1 to 10 will make it worse.

Describe in your own words, what is the objective function of the linear classifier? Write down a possible regularized loss for the linear classifier?

A linear classifier is one that maybe has a weight and a bias. The goal of a linear classifier is to compute model inputs to predicted outputs.

A possible regularized loss may be cross-entropy loss, softmax loss, logistic loss, focal loss….etc.

1. (5pts Bonus) In your own word, describe the **objective function** used in Zhang et al. “Colorful image colorization”( <https://arxiv.org/abs/1603.08511>) and explain the rationale of why it works.

The aim of this paper is to create a convincing colorization that could theoretically fool a human observer, rather than to recover the actual ground truth color. They rebalanced classes of rare color, taking more concernatreation to focus on a loss term.

The objective function used in Zhang et al. says that given an input lightness channel X ∈ RH×W×1 , the objective is to learn a mapping Yhat = F(X) to the two corresponding color channels Y ∈ RH×W×2 , where H, W are image dimensions. To solve the color ambiguity problem, add a new distribution Z as an input, mapping probability distribution Zhat to color values Yhat as a distribution output.