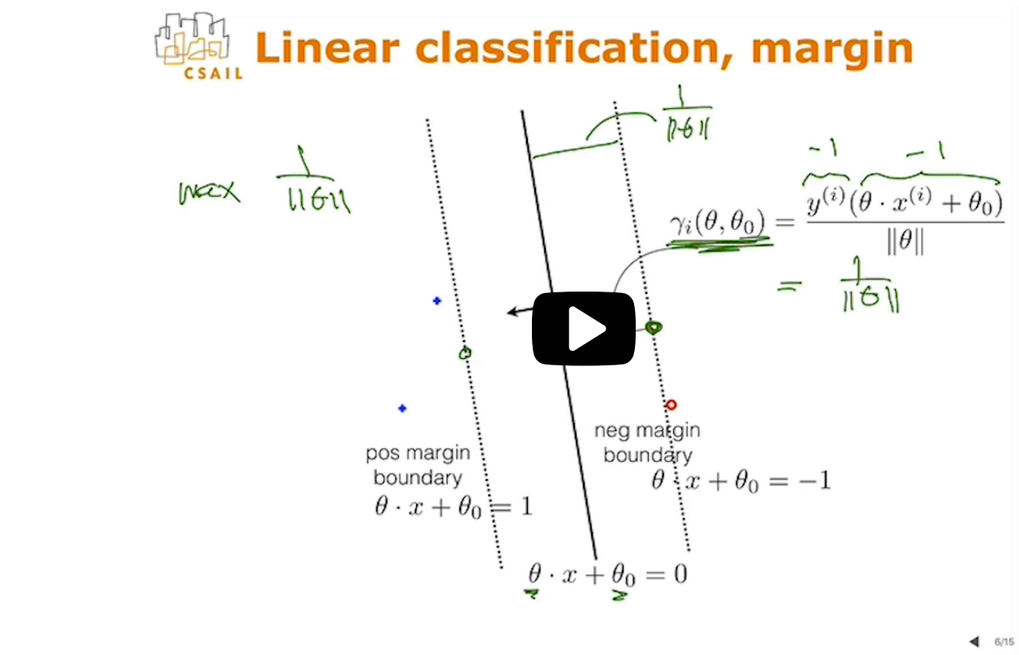
4. Hinge Loss and Objective Function
Hinge Loss and Objective Function



0:00 / 0:00 Speed 1.50x

Start of transcript. Skip to the end.

OK.

So now, our regularization goal here

is to maximize the distance that the margin boundaries are

from the decision boundaries.

This will be our regularization type, OK?

Now, we can proceed to define the objective function itself

for finding large margin decision boundaries.

It has two components, as we've already seen.

It has a loss

Video
Download video file

Transcripts
Download SubRip (.srt) file
Download Text (.txt) file

Hinge Loss Exercise 1

3 points possible (graded)
Compute the output of Hinge Loss function (as described in the video) for the following values:

$\text{Loss}_h\left(0\right) =$ **Answer:** 1

$\text{Loss}_h\left(0.2\right) =$ **Answer:** 0.8

$\text{Loss}_h\left(-10\right) =$ **Answer:** 11

Solution:

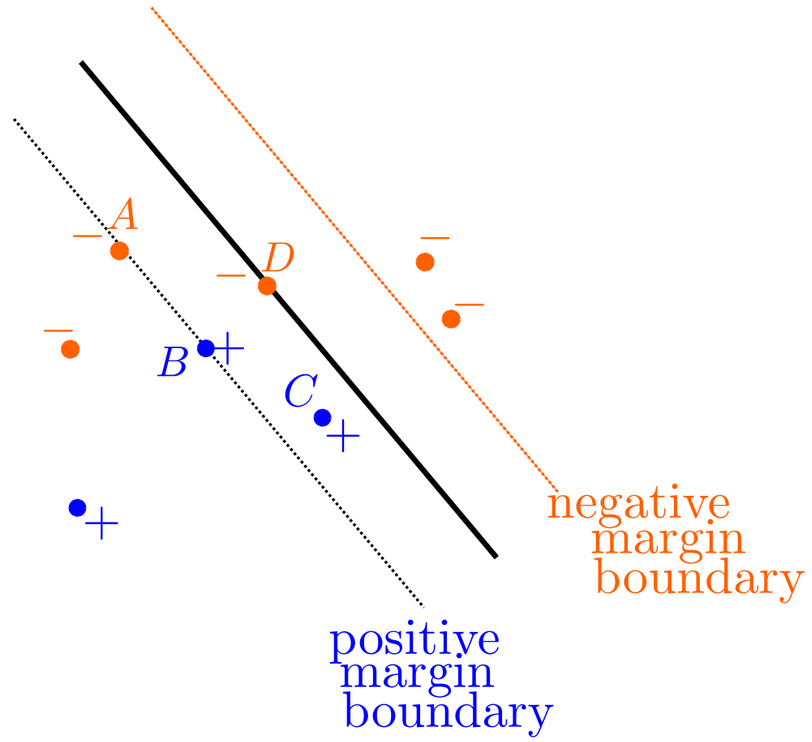
$$\text{Loss}_h\left(z\right) = \begin{cases} 0 \text{ if } z >= 1 \\ 1 - z \text{ otherwise} \end{cases}$$

Submit You have used 0 of 2 attempts

Answers are displayed within the problem

Hinge Loss Exercise 2

4 points possible (graded)
In a 2 dimensional space, there are points A, B, C, D as depicted below. Let $A = \left(x_a, y_a\right), B = \left(x_b, y_b\right), C = \left(x_c, y_c\right), D = \left(x_d, y_d\right)$



What is the hinge loss of point A, $\text{Loss}_h\left(y^{(a)}\left(\theta \cdot x^{(a)} + \theta_0\right)\right)$?

- 0
- between 0 and 1
- 1
- 2 ✔

What is the hinge loss of point B, $\text{Loss}_h\left(y^{(b)}\left(\theta \cdot x^{(b)} + \theta_0\right)\right)$?

- 0 ✔

- between 0 and 1
- 1

What is the hinge loss of point C, $\text{Loss}_h\left(y^{(c)}\left(\theta \cdot x^{(c)} + \theta_0\right)\right)$?

- 0
- between 0 and 1 ✔
- 1

What is the hinge loss of point D, $\text{Loss}_h\left(y^{(d)}\left(\theta \cdot x^{(d)} + \theta_0\right)\right)$?

- 0
- between 0 and 1
- 1 ✔

Solution:

A is on the positive margin boundary but with the label -1, so

$$y^{(a)}\left(\theta \cdot x^{(a)} + \theta_0\right) = -1.$$

Thus its hinge loss is 2. B is on the positive margin boundary and with the label $+1$, so

$$= y^{(b)}\left(\theta \cdot x^{(b)} + \theta_0\right) = 1.$$

Thus its hinge loss is 0. C lies between the decision boundary and the margin boundary. Thus

$$1 > y^{(c)} \left(\theta \cdot x^{(c)} + \theta_0\right) > 0.$$

Thus C's hinge loss is between 0 and 1. Similarly, because D is on the decision boundary,

$$y^{(d)} \left(\theta \cdot x^{(d)} + \theta_0\right) = 0.$$

Thus its hinge loss is 1.**Loss functions tell you in general how bad the prediction is.** The Hinge Loss tells us how undesirable a training example is, with regard to the margin and the correctness of its classification.

Submit You have used 0 of 3 attempts

Answers are displayed within the problem

Regularization

1 point possible (graded)

Remember that for points (x, y) on the boundary margin, the distance from the decision boundary to (x, y) is $\frac{1}{||\theta||}$. Thus

$$y^{(i)} \left(\theta \cdot x^{(i)} + \theta_0\right) = 1.$$

And

$$\frac{y^{(i)} \left(\theta \cdot x^{(i)} + \theta_0\right)}{|| \theta ||} = \frac{1}{|| \theta ||}.$$

Now our goal is to maximize the margin, that is to maximize $\frac{1}{||\theta||}$. Which of the following is **NOT** equivalent to maximizing $\frac{1}{||\theta||}$?

- maximizing $\frac{1}{||\theta||^2}$

- minimizing $|| \theta ||$
- maximizing $\sqrt{|| \theta ||}$ ✔

Solution:

Maximizing $\frac{1}{||\theta||}$ is equivalent to maximizing $\frac{1}{||\theta||^2}$. It is also equivalent to minimizing $|| \theta ||$.

Submit You have used 0 of 2 attempts

Answers are displayed within the problem

Objective

1 point possible (graded)
Remember that our objective is given as

$$J(\theta, \theta_0) = \frac{1}{n} \sum_{i=1}^{n} \text{Loss}_h \left(y^{(i)} \left(\theta \cdot x^{(i)} + \theta_0 \right) \right) + \frac{\lambda}{2} || \theta ||^2.$$

Our goal is to minimize this objective J. Now, which of the following is true if we have a large λ?

- We put more importance on maximizing the margin than minimizing errors ✔
- We put more importance on minimizing the margin than minimizing errors
- We put more importance on maximizing the margin than maximizing errors
- We put more importance on minimizing the margin than maximizing errors

Solution:

Remember that the first term

$$\frac{1}{n}\sum_{i=1}^{n}\text{Loss}_h\left(y^{(i)}\left(\theta\cdot x+\theta_0\right)\right)$$

corresponds to the sum of hinge losses on each training example, and the second term

$$\frac{\lambda}{2}\,||\,\theta\,||^2$$

corresponds to maximizing the margin. If we increase λ, we put more weight on maximizing the margin than minimizing the sum of losses.

Submit You have used 0 of 2 attempts

Answers are displayed within the problem

Discussion

Topic: Unit 1 Linear Classifiers and Generalizations (2 weeks):Lecture 3 Hinge loss, Margin boundaries and Regularization / 4. Hinge Loss and Objective Function

Hide Discussion

Add a Post

Show all posts ▼ by recent activity ▼

?	What's the difference between maximizing the margin than minimizing errors ? In most cases, do they lead to the same result ? In what cases we maximize the margin, in what cases we minimize errors and does it have anything to do with the distribu...	2
?	Confusion about the large margin linear classification as optimization	2
💬	How to evaluate Hinge Loss (memo)	21
☑	[Staff] Regularization Staff: My answer according the correct answer was ok, howerver, the grader evaluates as wrong. Please reset the answer. Thank you. Santiago	4
💬	[staff] problems with gradin Hi, i got all answers right from the lecture, but the progress tab says i got only 8 out of 12.	5
💬	[Staff]All questions marked correct but the progress bar shows 94%	2

As title have stated.

[Staff] Same answers in Objective
Two answers in Objective are exactly the same: "We put more importance on maximizing the margin than maximizing errors" I assume one of those should be: "We put m... 5

what does the first question even mean?
I don't know anything about this section of machine learning and have absolutely no idea for the first question.Can somebody pls help me out 7

Hinge Loss?
Can't Understand How is professor calculating Hinge Loss? 2

Why not to weight loss function. 6

Need Slides with Notes Posted under Course Resource
Is it possible to post lecture slides with notes??? 2

I can´t understand how loss function can assume a value greater than 2.
In the lecture he says that "if we go further it would be greater than two". I understand that " further" is a point after the margin boundary of the wrong side, that is, far fr... 4

Any suplementary material about this topic?
Does anyone knows sources to suplement the study of this topics? 2

[Staff]: Grader 2

Learn About Verified Certificates