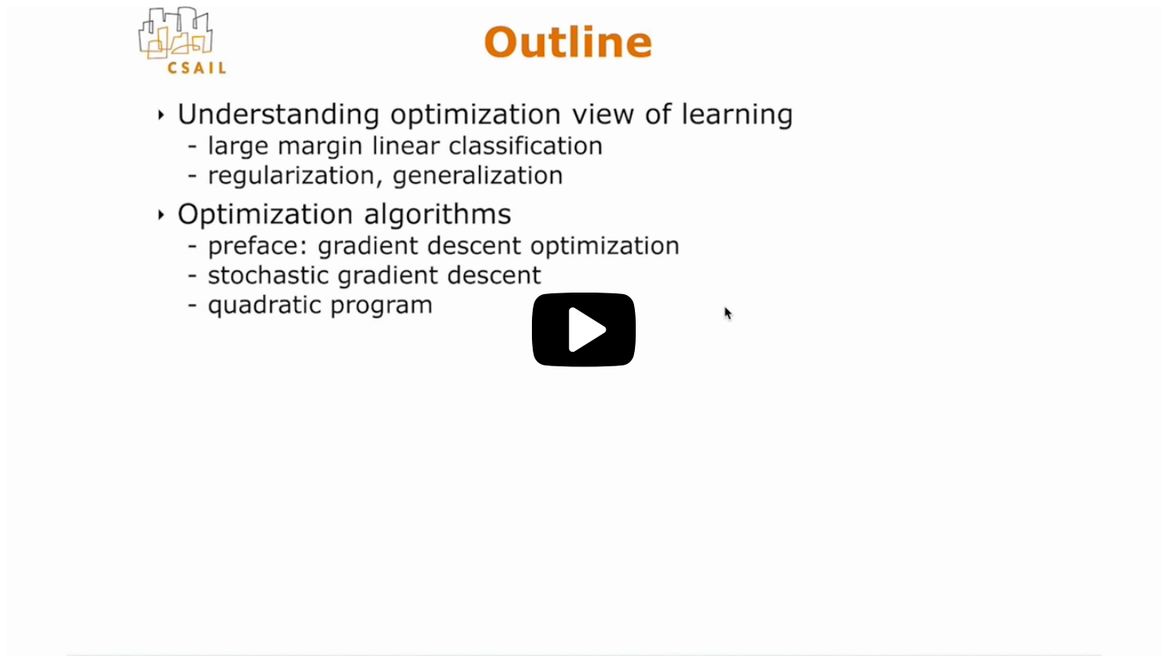
2. Review and the Lambda parameter

Introduction and Review

Start of transcript. Skip to the end.



Welcome back.

This is machine learning lecture number four.

Last time, we talked about how to formulate maximum margin

linear classification as an optimization problem.

Today, we're going to try to understand

the solutions to that optimization problem

and how to find those solutions.

If you recall, our objective function

for a learning problem, the objective function,

Video

Download video file

Transcripts

Download SubRip (.srt) file

Download Text (.txt) file

Distance from a line to a point in terms of components

1.0/1 point (graded)
In a 2 dimensional space, a line L is given by $L: ax+by+c=0$, and a point P is given by $P=(x_0,y_0)$. What is d, the shortest distance between L and P? Express d in terms of a, b, c, x_0, y_0.

(a*x_0+b*y_0+c)/sqrt(a^2

STANDARD NOTATION

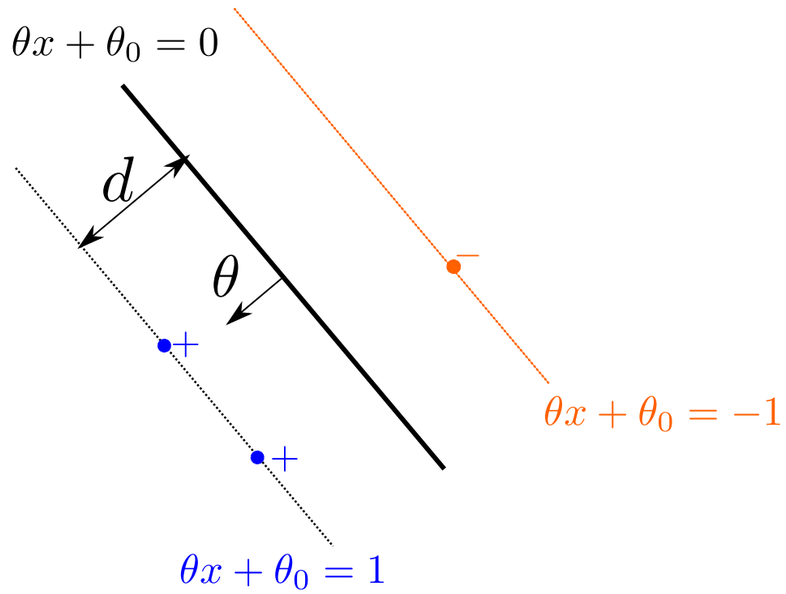
Submit You have used 1 of 3 attempts

Varying Lambda in the Geometric Sense

1/1 point (graded)
Remember that the objective

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

In the picture below, what happens to d, the distance between the decision boundary and the margin boundary, as we increase λ?



- d decreases
- d increases ✔
- d converges to λ

Hint: You can answer with your intuition in this question. To see whether d converges to λ, think of a simple setting where we are working in 1 dimension with just two points with labels $x_1 = -1, x_2 = 2, y_1 = -1, y_2 = 1$ and assume that λ is large enough where it dominates the loss function and pushes θ close enough to 0 where all points are margin violators.

Solution:

Increasing λ means we put more weight on maximizing the margin. Thus d increases.
It is not true that d always converges to λ as λ increases. Here is a counter example:
Consider a simple setting where we are working in 1 dimension with just two points with labels $x_1 = -1, x_2 = 2, y_1 = -1, y_2 = 1$ and assume that λ is large enough where it dominates the loss function and pushes θ close enough to 0 where all points are margin violators.

$$J = \frac{1}{2}[(1 - \theta + \theta_0) + (1 - 2\theta - \theta_0)] + \frac{\lambda}{2}\theta^2$$

$$= \frac{2-3\theta}{2} + \frac{\lambda}{2}\theta^2.$$

Solve this explicitly by taking $\frac{\partial J}{\partial \theta} = 0$:

$$\begin{aligned} \frac{-3}{2} + \lambda\theta &= 0 \\ \theta &= \frac{3}{2\lambda} \\ d &= \frac{1}{\theta} = \frac{2}{3}\lambda. \end{aligned}$$

Submit You have used 2 of 2 attempts

Answers are displayed within the problem

Discussion

Hide Discussion

Topic: Unit 1 Linear Classifiers and Generalizations (2 weeks):Lecture 4. Linear Classification and Generalization / 2. Review and the Lambda parameter

Add a Post

Show all posts ▼ by recent activity ▼

?	Why do we assign the points exactly on the margin boundary (assuming they are on the correct side) as having 0 hinge loss? Why do we assign the points exactly on the margin boundary (assuming they are on the correct side) as having 0 hinge loss? I know this is so by the definition of the loss f...	5
💬	Recomputing Decision Boundary Based on Larger Lambda Values As the decision boundary is pushed out it better represents the average of the data, not just the point closes to the decision boundary. In the video example, the decision ...	1
?	What is margin boundary for ?	17
?	Varying Lambda in the Geometric Sense - Is the same that varying the regularization parameter? [Staff] Hi!, In the video, (00:02:18), the transcription says: "- So when we minimize the regularization term here, {pointing to the term lambda*norm(theta)/2} - we are push...	6
☑	How to write subscript in standard notation? How to write subscript in standard notation?	2
?	lambda and boundary rotation It is shown that by changing lambda the boundary rotates? How is it possible to affect thetas by manipulating lambda only? Isn't regularization part of the objective functi...	2

?	Margin violators. Hi, Am I right that margin violators is points between dec. boundary and marginal boundary? (Even if not, I will call points between dec bou. and mar. bou as margin violat...	7
💬	Why would we want to push the margin boundaries out? Is this a means of adjusting the sensitivity of the classifier, sacrificing certain events within the margin for the sake of increasing the MB?	2
💬	How is question 1 related to this particular lecture segment? Just curious, why this question was put here.	1
💬	[Staff] - parsing error Hi Staff, I think I have given the correct formula for distance.. But it gave a an error like this.. **Could not parse \'|.....|/|sqrt(....)|\' as a formula.** Is it because of the **...	3
?	Varying Lambda in the Geometric Sense: x missing superscript	2
💬	[staff] Typos in solution for "Varying Lambda in the Geometric Sense" Community TA	2
?	Distance from a point to line Is this distance the signed distance or unsigned distance? My answer with the unsigned distance was marked wrong and I do not know why,	5

Learn About Verified Certificates