8. Regularization

Exercises due Oct 14, 2020 05:29 IST *Completed*

Ridge Regression

Start of transcript. Skip to the end.

So now, let's just start seeing how

our intuition about what regularization

should be doing translates into a specific objective.

So now, it will have a slightly different notation.

And it's called reach regression.

Video
Download video file

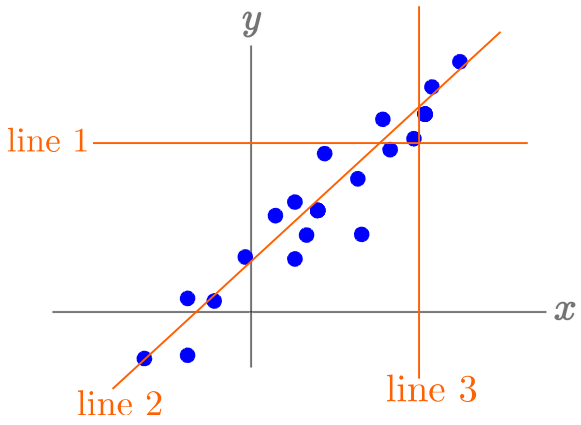
Transcripts
Download SubRip (.srt) file
Download Text (.txt) file

Regularization: extreme case 1

1/1 point (graded)
As in the video above, define the loss function

$$J_{n,\lambda}(\theta,\theta_0) = \frac{1}{n}\sum_{t=1}^{n}\frac{\left(y^{(t)} - \theta\cdot x^{(t)} - \theta_0\right)^2}{2} + \frac{\lambda}{2}\|\theta\|^2$$

where λ is the regularization factor.



In the figure above, the blue dots are the training examples. If we increase λ to ∞, where does $f(x) = \theta\cdot x + \theta_0$ converge to?

line 1

line 2

line 3

Solution:

If we increase λ to ∞, minimizing J is equivalent to minimizing $||\theta||$. Thus θ will have to be a zero vector. Thus $f(x) = \theta \cdot x + \theta_0$ becomes $f(x) = \theta_0$, a horizontal line. Thus f converges to line 1.

Submit You have used 2 of 2 attempts

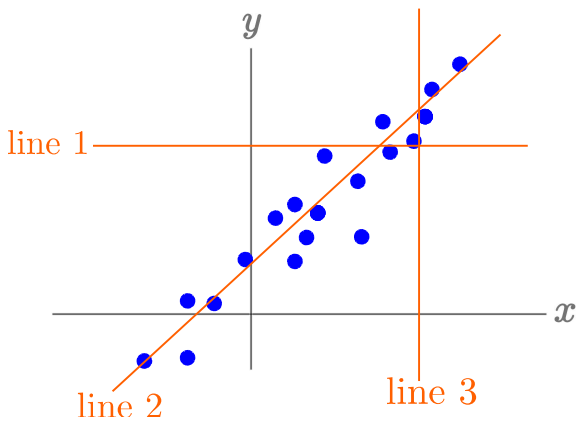
Answers are displayed within the problem

Regularization: Extreme case 2

1/1 point (graded)
As in the problem above,

$$J_{n,\lambda}(\theta, \theta_0) = \frac{1}{n}\sum_{t=1}^{n} \frac{\left(y^{(t)} - \theta \cdot x^{(t)} - \theta_0\right)^2}{2} + \frac{\lambda}{2}\|\theta\|^2$$

where λ is the regularization factor.



In the figure above, the blue dots are the training examples. If we decrease λ to 0, where does $f(x) = \theta \cdot x + \theta_0$ converge to?

line 1

line 2

line 3

Solution:

If we decrease λ to zero, minimizing J is equivalent to minimizing $\frac{1}{n}\sum^{n} \frac{\left(y^{(t)} - \theta \cdot x^{(t)} - \theta_0\right)^2}{2}$ which is the "fit." Thus f

converges to line 2.

Submit You have used 1 of 2 attempts

Answers are displayed within the problem

Discussion

Hide Discussion

Topic: Unit 2 Nonlinear Classification, Linear regression, Collaborative Filtering (2 weeks):Lecture 5. Linear Regression / 8. Regularization

Add a Post

Show all posts by recent activity

	@Staff How does adding the regularization term, ground the theta to a specific region? I understand the notion of regularization in classification, but its not entirely clear in regression. What does the professor mean by theta...	4
	Intuition behind adding the norm of theta as the penalty In classification, it was clear to me why norm of theta was introduced as the penalty but here it isn't. Can someone explain how a large n...	3
?	STAFF: cannot add a post to the next video, '9. Closing Comment' I can't add a post under the next video entitled '9. Closing Comment' and more specifically the text '(Optional) Equivalance of regularizati...	7
?	grad (J) at 5:00 Hi! Would you please explain why we omit 1/n and summation when we differentiate R term of J?	2
	James-Stein Estimation I recall from the Statistics class about the Stein paradox - in a multi-dimensional estimation, you need to deliberately shrink your least s...	1
	A good video on Ridge Regression and Regularization I really like his channel, where he explains the basics of ML in an easy tone. https://www.youtube.com/watch?v=Q81RR3yKn30	8
	Value of theta_0	3
	How to choose lambda I understood that better value of lambda increases regularization and the model fits better to test data but how do we get the value of l...	4
?	why I submitted answers on time but it shows past due Dear Professors, I submitted my answers for lecture 2 and lecture 3 on time, on Tuesday, before the deadline, But I just found it out it sa...	2

Previous Next

Legal

Connect