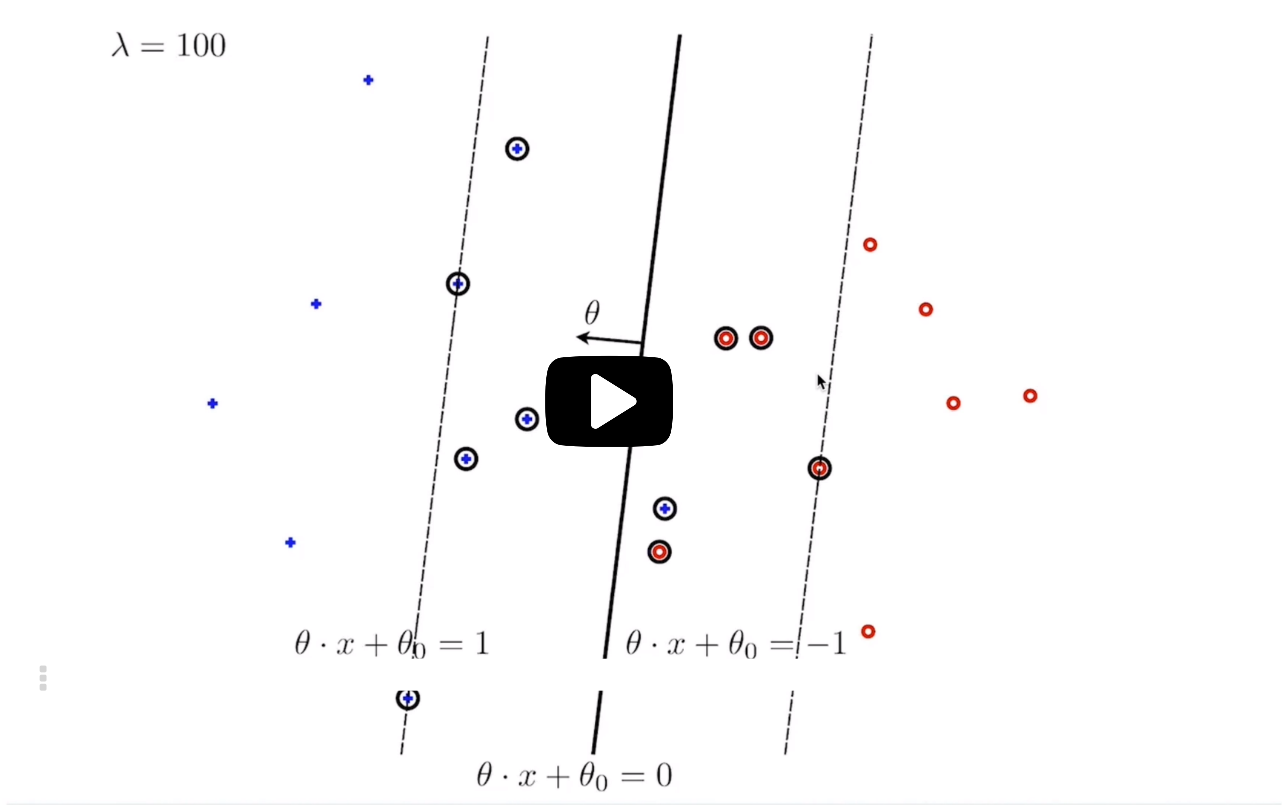
2. Review and the Lambda parameter

Introduction and Review



5:57 / 5:57 1.25x

so some of the margin constraints are violated.

So I will incur losses on some of those points already here.

As I start increasing the value of the regularization parameter

lambda, the solution starts changing as before,

being guided more by where the bulk of the points

are similar to before.

End of transcript. Skip to the start.

Video

Download video file

Transcripts

Download SubRip (.srt) file

Download Text (.txt) file

Distance from a line to a point in terms of components

1/1 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.

abs(a*x_0+b*y_0+c)/sqrt ✔ **Answer:** abs(a*x_0 + b*y_0 +c) / sqrt(a^2 + b^2)

STANDARD NOTATION

Solution:

Use the projection equation. Here θ is $[a, b]$, θ_0 is c and the point is $[x_0, y_0]$.

Submit You have used 1 of 3 attempts

Answers are displayed within the problem

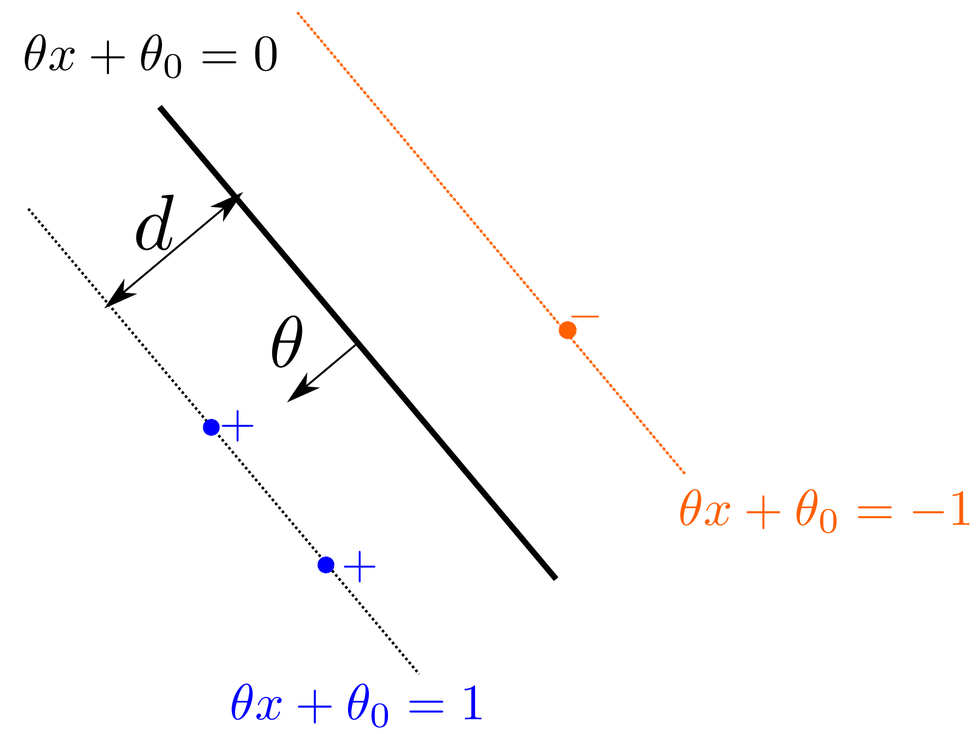
Varying Lambda in the Geometric Sense

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.

$$\begin{aligned} J &= \frac{1}{2}[(1-\theta+\theta_0)+(1-2\theta-\theta_0)]+\frac{\lambda}{2}\theta^2 \\ &= \frac{2-3\theta}{2}+\frac{\lambda}{2}\theta^2. \end{aligned}$$

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

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

Submit

You have used 1 of 2 attempts

Answers are displayed within the problem