Classfication

Sigmoid function

hθ(x)=g(θTx) y {0, 1} so 0 <= h(x) <= 1.

z=θTx

g(z)=1/(1+e−z)

Decision boundary is the line that separates the area where y = 0 and where y = 1. It is created by our hypothesis function.

if y = 1 cost(h(x), y) = - log(h(x))

if y = 0 cost(h(x), y) = - log(1-h(x))

\*why we can’t use linear regression cost function with logistic?

Not convex, will have lots of optimal solutions.

cost(h(x), y) = - y log(h(x)) – (1-y)log(1-h(x))

J(theta) = −*m*1​*i*=1∑*m*​[*y*(*i*)log(*hθ*​(*x*(*i*)))+(1−*y*(*i*))log(1−*hθ*​(*x*(*i*)))]

partial derivative j(theta j) = 1/m ∑ (h(xi) – yi)thetaj

Overfitting: suit the training set too much, so that can’t perform well with new data.

two way to fit it:

1. reduce the number of features.

2. regularization.

when we have j(theta) and partial derivative j, we use function to calculate minimum of theta.