Close

Logistic Regression

5 questions

1.

Suppose that you have trained a logistic regression classifier, and it outputs on a new example *x* a prediction *hθ*(*x*) = 0.7. This means (check all that apply):



Our estimate for *P*(*y*=1|*x*;*θ*) is 0.3.



Our estimate for *P*(*y*=0|*x*;*θ*) is 0.3.



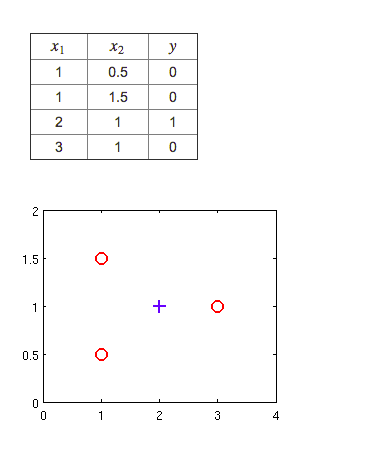
Our estimate for *P*(*y*=1|*x*;*θ*) is 0.7.



Our estimate for *P*(*y*=0|*x*;*θ*) is 0.7.

2.

Suppose you have the following training set, and fit a logistic regression classifier *hθ*(*x*)=*g*(*θ*0+*θ*1*x*1+*θ*2*x*2).



Which of the following are true? Check all that apply.



*J*(*θ*) will be a convex function, so gradient descent should converge to the global minimum.



Adding polynomial features (e.g., instead using *hθ*(*x*)=*g*(*θ*0+*θ*1*x*1+*θ*2*x*2+*θ*3*x*21+*θ*4*x*1*x*2+*θ*5*x*22) ) could increase how well we can fit the training data.



The positive and negative examples cannot be separated using a straight line. So, gradient descent will fail to converge.



Because the positive and negative examples cannot be separated using a straight line, linear regression will perform as well as logistic regression on this data.

3.

For logistic regression, the gradient is given by ∂∂*θjJ*(*θ*)=∑*mi*=1(*hθ*(*x*(*i*))−*y*(*i*))*x*(*i*)*j*. Which of these is a correct gradient descent update for logistic regression with a learning rate of*α*? Check all that apply.



*θ*:=*θ*−*α*1*m*∑*mi*=1(*θTx*−*y*(*i*))*x*(*i*).



*θj*:=*θj*−*α*1*m*∑*mi*=1(11+*e*−*θTx*(*i*)−*y*(*i*))*x*(*i*)*j* (simultaneously update for all *j*).



*θj*:=*θj*−*α*1*m*∑*mi*=1(*hθ*(*x*(*i*))−*y*(*i*))*x*(*i*) (simultaneously update for all *j*).



*θj*:=*θj*−*α*1*m*∑*mi*=1(*hθ*(*x*(*i*))−*y*(*i*))*x*(*i*)*j* (simultaneously update for all *j*).

4.

Which of the following statements are true? Check all that apply.



The cost function *J*(*θ*) for logistic regression trained with *m*≥1 examples is always greater than or equal to zero.



The sigmoid function *g*(*z*)=11+*e*−*z* is never greater than one (>1).



Linear regression always works well for classification if you classify by using a threshold on the prediction made by linear regression.



For logistic regression, sometimes gradient descent will converge to a local minimum (and fail to find the global minimum). This is the reason we prefer more advanced optimization algorithms such as fminunc (conjugate gradient/BFGS/L-BFGS/etc).

5.

Suppose you train a logistic classifier *hθ*(*x*)=*g*(*θ*0+*θ*1*x*1+*θ*2*x*2). Suppose *θ*0=6,*θ*1=0,*θ*2=−1. Which of the following figures represents the decision boundary found by your classifier?



Figure:

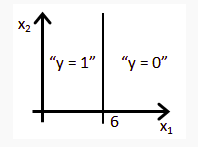




Figure:

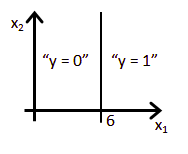




Figure:

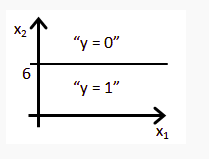
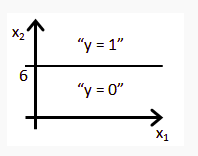




Figure:



1 question unanswered

Submit Quiz