Michael Lee

CSE 446

HW2

1 a) False, is not typically sparse. This is because by taking the 2-norm of the vector we don’t select sparse solutions optimally. For example:

If we calculate the L2 norm of the vectors:

Versus if we calculate the L1 norm:

In the L2 case we consider where as in the L1 case we consider

, hence we don’t optimize for solutions of sparsity.

b) False? By including the penalty we avoid this error.

c) True, with by increasing lamda our probabilities increase so

d) False,

e)

i)

ii) The relationship between is that is accounting for some other weight given to the same data point

iii) Would it change if using L2 regularization?

2) Boosting

a)

b) show:

c)

i)

= 0

..