Andrew Elysee (aelys2)

CS446

Homework #4

1. 1. The empirical risk with squared loss is:  
      Since we know that our are represented by the standard basis vectors we can substitute our for   
      Since represents an standard basis vector would be a matrix with zeros everywhere except at where it would equal 1. This means we can say that . Keep in mind that is technically a length vector with 0’s everywhere except at position where it is equal to .  
      Since is a length standard basis vector and is a scalar, let , the resulting vector of multiplying the two. At this point we can also remove outer summation to focus on each individual component of .

This is what we wanted to show since both and are both 0 everywhere except , we can think of them as the scalar value at th component of each vector.

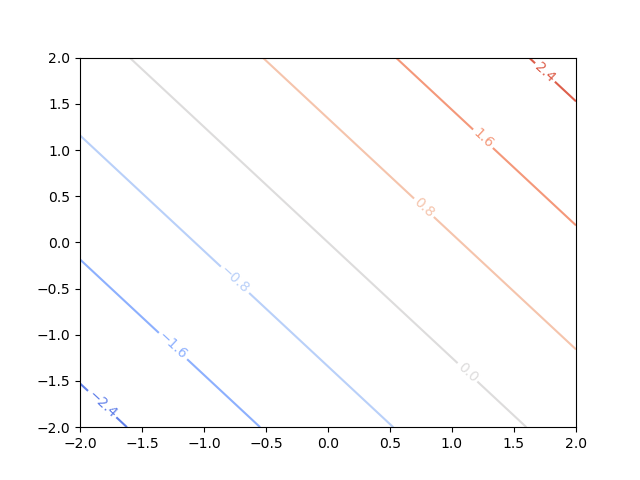
* 1. This is what we needed to show.
  2. Suppose that has a meaning that spans . That means that has non-zero singular values. We also know that the squares of the singular values of are the eigenvalues of . Consider that is an because is and is . Also, the determinant is equal to the product of the eigenvalues. Since all the eigenvalues of are all non-zero, we can come to the conclusion that is invertible because the determinant does not equal 0.   
       
     Now, let’s consider the case that does not span . This would mean that , so there would be less than non-zero singular values. So, would have at least 1 eigenvalue equal to 0, meaning that it’s determinant would be 0. Thus, is not invertible if does not span .
  3. Let X be:   
     where the SVD is:  
     As we can clearly see, since X has only 2 singular values, the meaning that it spans but not . For this reason alone, we can conclude that is invertible while is not. Furthermore, if we calculate and , we see that their eigenvalues are [4,1] and [4,1,0] respectively. This means the making invertible and the making not invertible.

1. 1. In hw4.py
   2. In hw4.py
   3. Chart, scatter chart

      Description automatically generated
2. 1. In hw4.py
   2. In hw4.py
   3. The Polynomial Regression gives a significantly better approximation of the data. As we can see below the data has a slight curve upwards, so the polynomial regression is more capable of staying withing the points and following that curve, while the linear regression is closer to the edge of the points in a lot of places.   
      Chart, scatter chart

      Description automatically generated
   4. As we can see in our plots below and our predictions, only the Polynomial Regression was able to correctly classify all our points. This makes sense because the polynomial is solving for the term (which is the XOR function) while the linear regression does not.  
      Linear Predictions: [-1.4901e-08, 1.4901e-08, -1.3411e-07, 1.3411e-07]  
      Polynomial Predictions: [-1.0000, -1.0000, 1.0000, 1.0000]

**Plots on next page**

Linear Prediction Plot  
  
  
Polynomial Prediction Plot  
