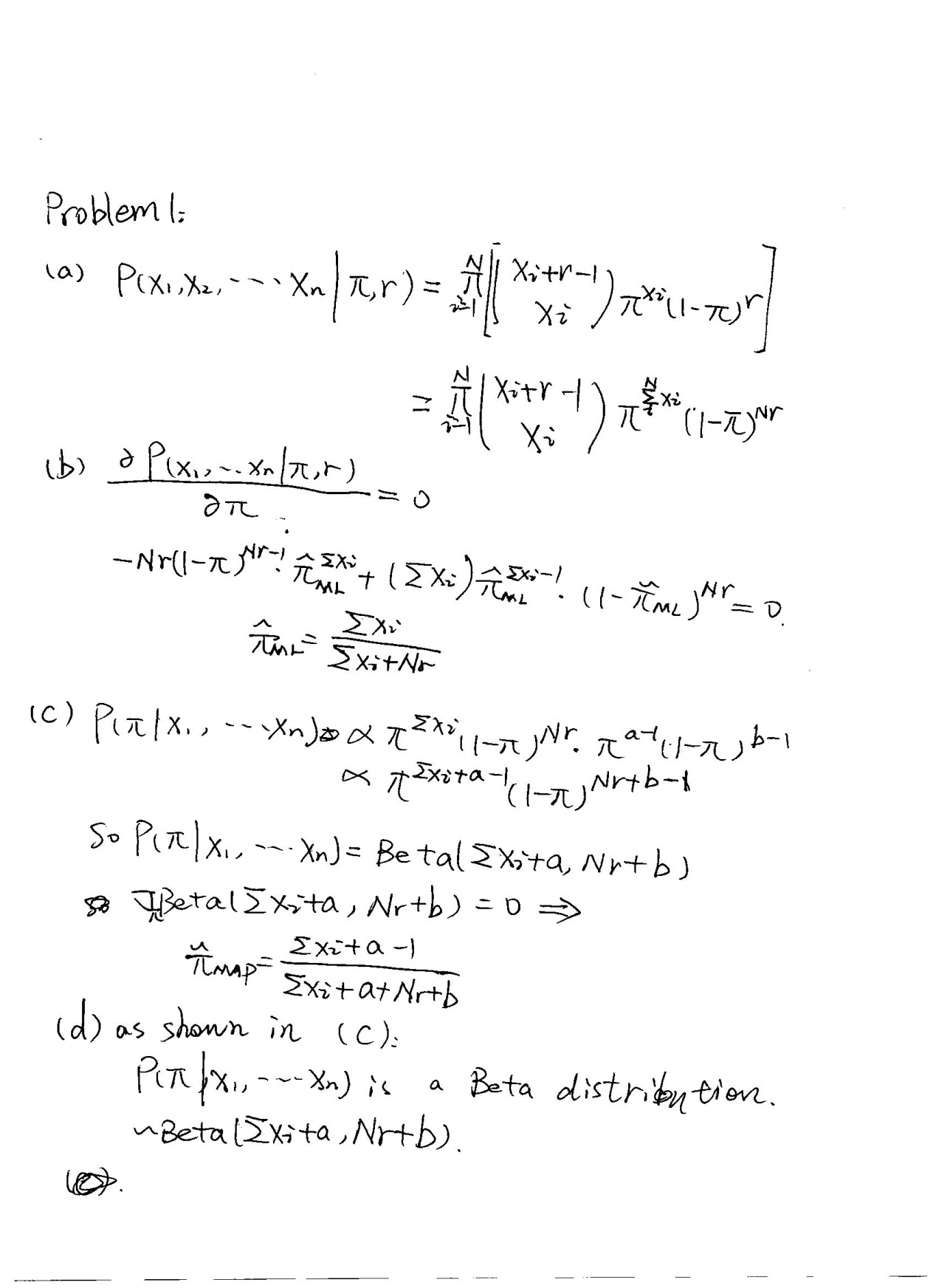
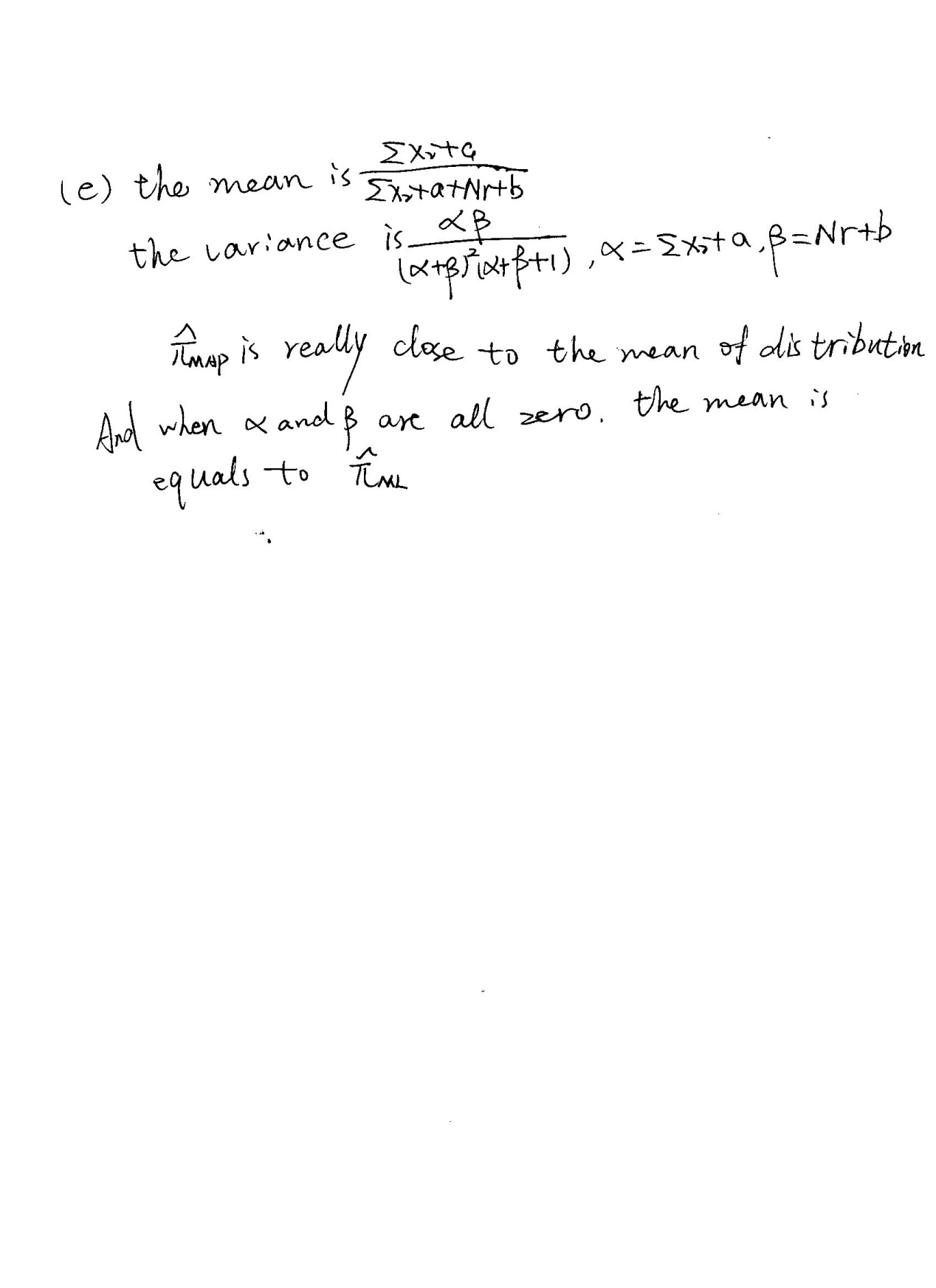
COMS4721\_hw1

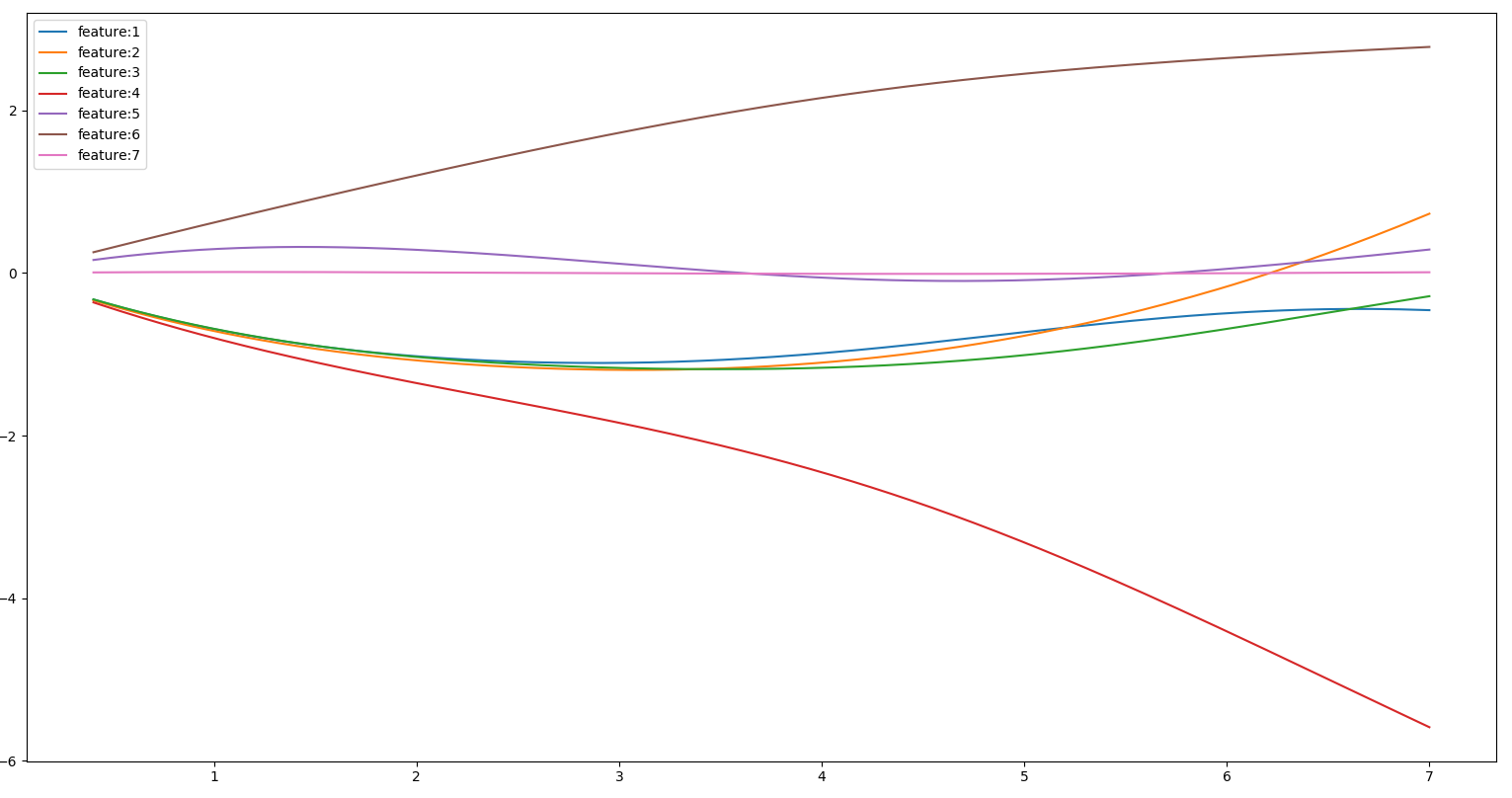
jc4609



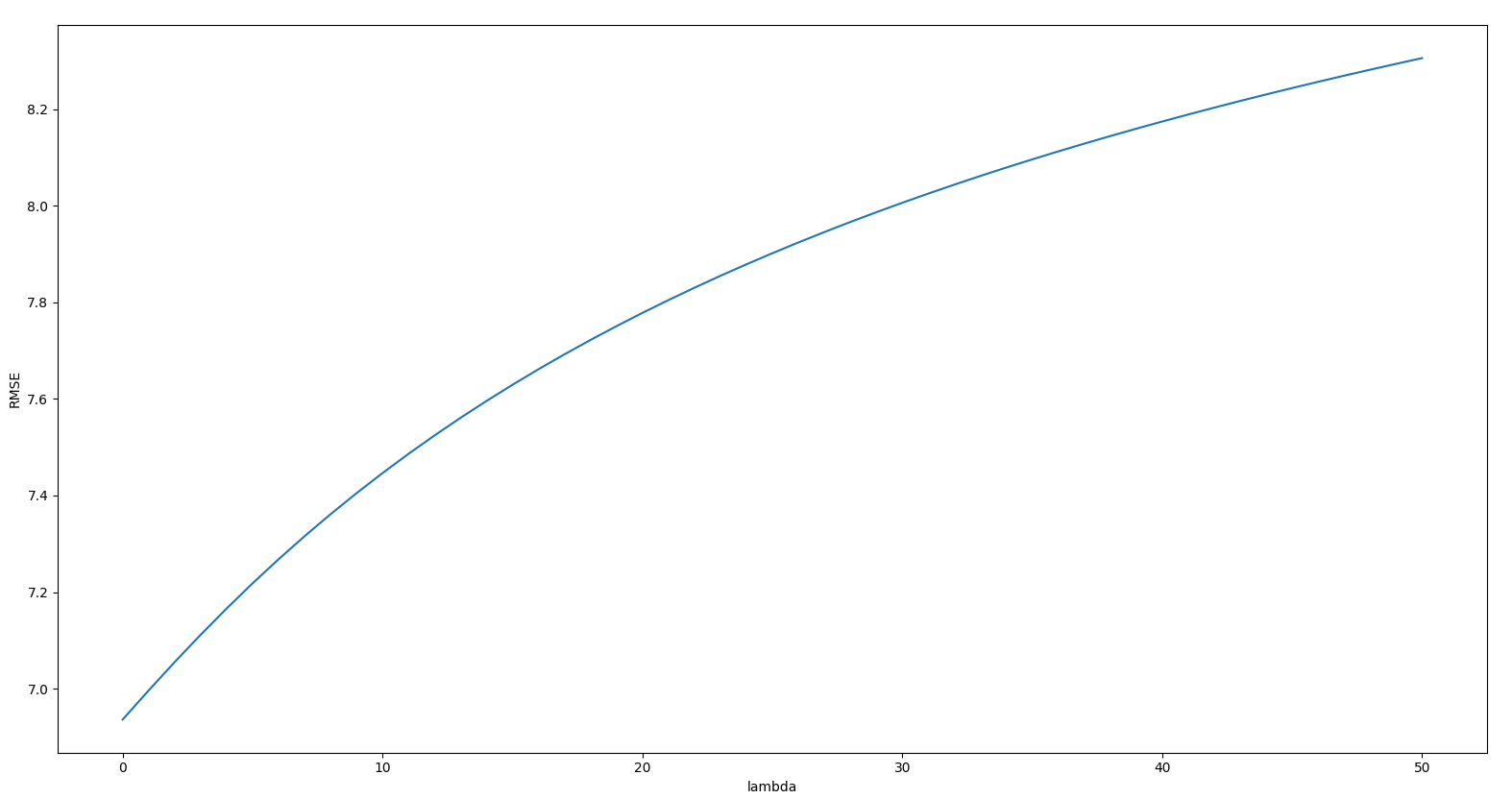


Problem2:

Part 1.(a)



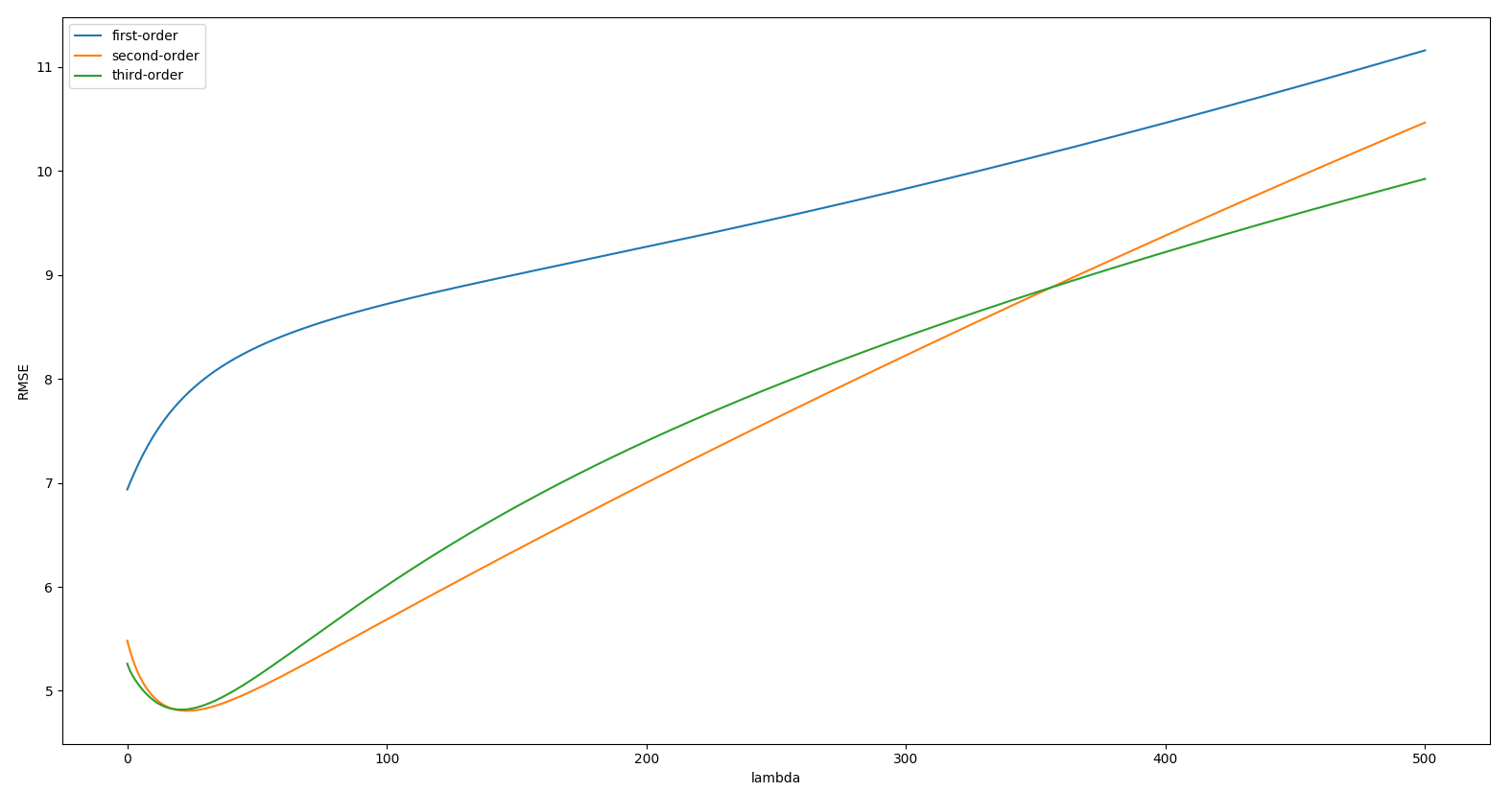
(b)It seems the forth and the sixth dimension have greater influence on “y” than other dimensions. When  is small , this two dimensions are much more influential than others. And it becomes less obvious when  increase.

(c). 

I would choose =0, in other words I would choose least square regression.

Part 2.

(d)



I think p = 2 is better.Because the RMSE of second order and third order is very similar, so I prefer to predict the result with less features since it can save calculation resource.

The primary goal of using  is preventing overfitting.In the case of first order regression, there is no issue about overfitting. So when  increase the RMSE just increace. And In the case of second and third order regression, the overfitting may affect our prediction. Thus we introduce a proper (about 25) to handle this ,and it works good. However , if  is too big it will still increase the RMSE.