

# CS 541 Artificial Intelligence: Homework 4

Tarquin Bennett

December 8, 2020

The gradient for  $\min_{U,V} F(U, V) := \frac{1}{2} \sum_{(i,j) \in \Omega_1} (M_{ij}^2 + u_i v_j^T)^2 + \frac{\lambda}{2} (\|U\|_F^2 + \|V\|_F^2)$   
with respect to U and V are:

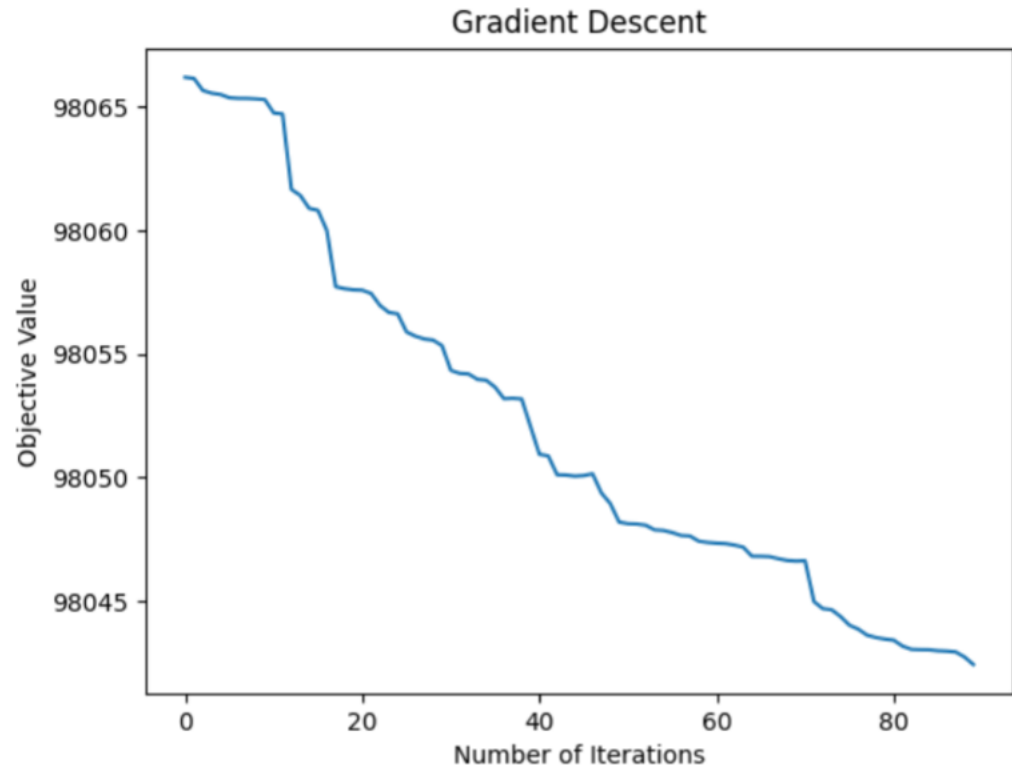
$$\frac{\partial F}{\partial U} = (M_{i,j} - u_i v_j^T)(-v_j) + \lambda u_i$$

$$\frac{\partial F}{\partial V} = (M_{i,j} - u_i v_j^T)(-u_i) + \lambda v_j$$

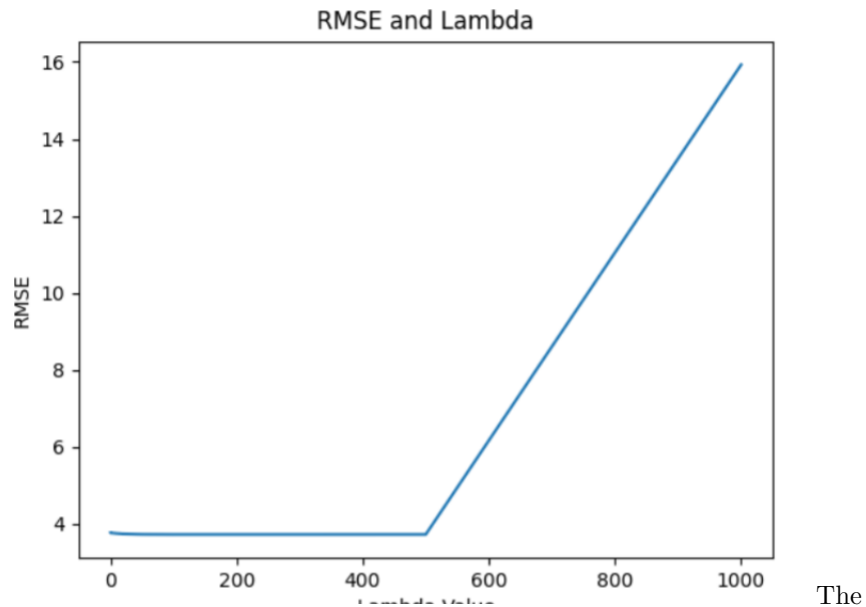
Since  $\lambda = 1$ , the updated gradients are:

$$\frac{\partial F}{\partial U} = (M_{i,j} - u_i v_j^T)(-v_j) + u_i$$

$$\frac{\partial F}{\partial V} = (M_{i,j} - u_i v_j^T)(-u_i) + v_j$$



This shows us that as the number of iterations goes up the objective value goes down. This is proof that the GD is working and getting us close to the true Objective Value. More iterations would get us closer to the true value.



The RMSE value for  $\lambda = 1$  is  $\approx 3.7861$ . The above graph shows that as the lambda value goes up the RMSE stays about the same till 1000 where the RMSE goes up to about 15. This means that around  $\lambda = 500$  the error shoots up, therefore the lambda value should be less than 500.