Jerry Chen

CS541

<center>HW4</center>

Learning

$$F(U,V) = \frac{1}{2}\sum\left(M_{ij} - u_i v_j^T\right)^2 + \frac{\lambda}{2}\left(||U||_F^2 + ||V||_F^2\right)$$

1. $\frac{dF}{dU} = (M - uv)v^T - \lambda U, \frac{dF}{dV} = (M - uv)u^T - \lambda V$
2. For each non-zero element in the testing dataset, U and V are updated by subtracting its current value by its gradient multiplied by the learning rate. It is updated for 100 iterations in which the error should hopefully gradually approach 0.
3.
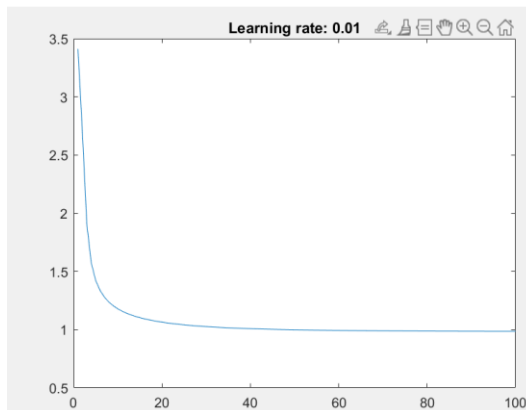
The average error of our model does decrease as the iterations go on but seems to converge at around 1 and does not quite reach 0.

Evaluation

1. RMSE for $\lambda = 1$ was 1.4895
2.
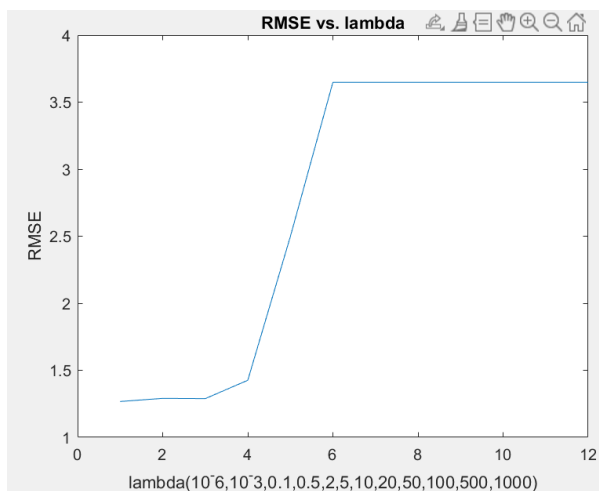
The RMSE is lower the smaller value lambda is and increases as lambda gets greater. It seems to even out at 3.65 after lamda is geater than 10.