# Variational Inference - Assignment 2
# A Report

Ravi Regulagedda

## Outputs

```
❯ python homework.py
Please enter choice of dataset:

1. 500-500-3

2. 100-k

Please enter your choice: 1 (or) 2 [To Debug, type 0] - 1
Baseline rmse is:  1.8019448325127836
For Part 1:

Total Iterations:  25000

RMSE:  [1.781923883557573, 1.5937212241838992, 1.2321690119758026, 0.722144708569602, 0.3283120448423085, 0.2359814426974392
6, 0.21916363355676868, 0.2160546121083858
]
Time taken:  [0.48094797134399414, 0.8806238174438477, 1.3032078742980957, 1.7059125900268555, 2.105543613433838, 2.519871711730957, 2.923938751220703, 3.326377868652
3438]


For Part 2:
For the given ranks:  [1, 2, 3, 5, 10, 20]
The RMSE changes as follows:  [1.5068963472083776, 1.1600153878586528, 0.21661693422373104, 0.22146709229941794, 0.24548545758228488, 0.2738502473472803]
Please check your current directory for generated plots
```

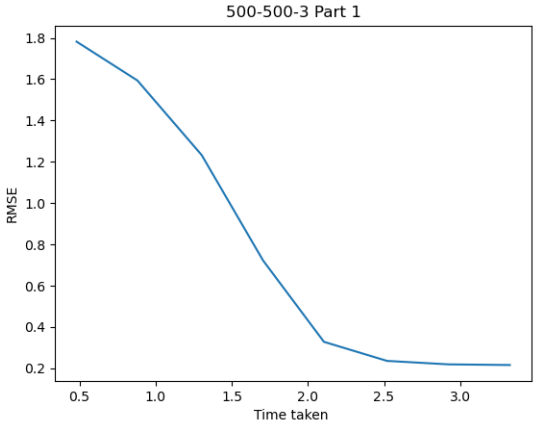
Figure 1: Output after running matrix factorization on 500-500 dataset

```
❯ python homework.py
Please enter choice of dataset:

1. 500-500-3

2. 100-k

Please enter your choice: 1 (or) 2 [To Debug, type 0] - 2
Baseline rmse is:  3.4397503673365066
For Part 1:

Total Iterations:  50000

RMSE:  [3.4112381981697038, 1.7459647590787626, 1.259336614373994, 1.0717570365475515, 1.0094145671801862, 0.9869535940433072, 0.9773663235991397, 0.9729829997699876,
 0.971110051528889, 0.9706030114922143, 0.970895231935521, 0.9716668351359257, 0.9727191894855078, 0.973922483582565, 0.9751909066372002, 0.9764687906820774, 0.977721
5207516838, 0.9789289840670536, 0.9800807532549812, 0.9811726218729699, 0.9822042020987551, 0.9831773211356803, 0.9840949888228743, 0.9849607561742121, 0.985778331668
2548, 0.9865513619941978, 0.9872833146365183, 0.9879774217741703, 0.9886366600621885, 0.9892637507258457, 0.9898611706150541, 0.9904311686869036, 0.9909757846923342,
0.9914968682259249, 0.9919960971235808, 0.9924749946841002, 0.9929349454815756, 0.9933772097052951, 0.9938029360595462, 0.9942131733075552, 0.9946088805674178, 0.9949
909364767786, 0.9953601473413681, 0.9957172543769532, 0.9960629401452848, 0.9963978342757025, 0.9967225185542518, 0.9970375314535498, 0.9973433721682833, 0.9976405042
139241, 0.9979293586393044, 0.9982103368982594, 0.9984838134197467, 0.9987501379115318, 0.9990096374283076]
Time taken:  [1.014092206954956, 1.9421393871307373, 2.8472652435302734, 3.7813050746917725, 4.687999963760376, 5.588575124740601, 6.562986135482788, 7.52633404731750
5, 8.519048929214478, 9.499677896499634, 10.39844298362732, 11.366978168487549, 12.267967224121094, 13.193376302719116, 14.053032159805298, 14.950924396514893, 16.002
52628326416, 16.864455461502075, 17.84455132484436, 18.75015616416931, 19.7137610912323, 20.627712965011597, 21.552350997924805, 22.5420880317688, 23.435638189315796,
 24.443799257278442, 25.36793327331543, 26.301979303359985, 27.26687240600586, 28.163235425949097, 29.044275283813477, 29.99522614479065, 30.884100437164307, 31.81834
84077453 6, 32.71967530250549, 33.687013149261475, 34.61783504486084, 35.56900691986084, 36.43928599357605, 37.2837769985199, 38.5317702293396, 39.508172273635864, 40.
4709312915802, 41.34211921691894 5, 42.29035234451294, 43.22833251953125, 44.223445415496826, 45.111032247543335, 46.00450420379639, 46.965280294418335, 47.98959326744
0796, 48.999062299728394, 50.0300931930542, 51.08395743370056, 52.17591118812561]


For Part 2:
For the given ranks:  [1, 2, 3, 5, 20, 100]
The RMSE changes as follows:  [0.9709810673837846, 0.9801935787643553, 0.9847149864732342, 1.0167913768950378, 1.1632308654873804, 1.1061426231943978]
Please check your current directory for generated plots
```

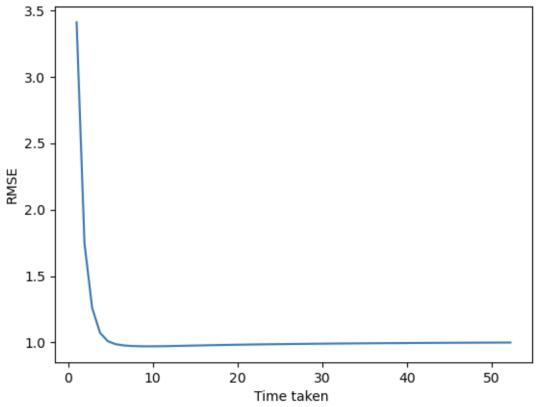
Figure 2: Output after running matrix factorization on 100-k dataset

## Results

From the graphs below, we can see that the RMSE for both the datasets goes down per iteration, showing that they converge to a minimum.

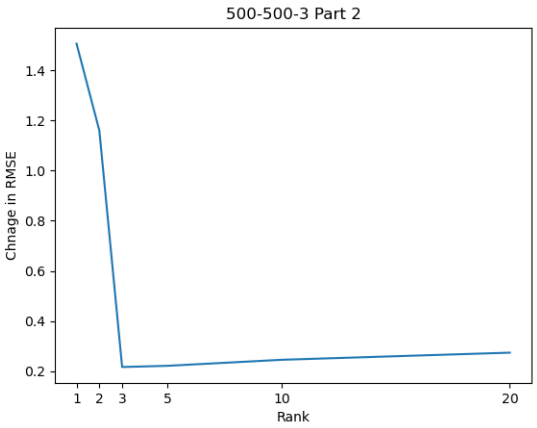



(a) Time v/s RMSE for 500-500 dataset

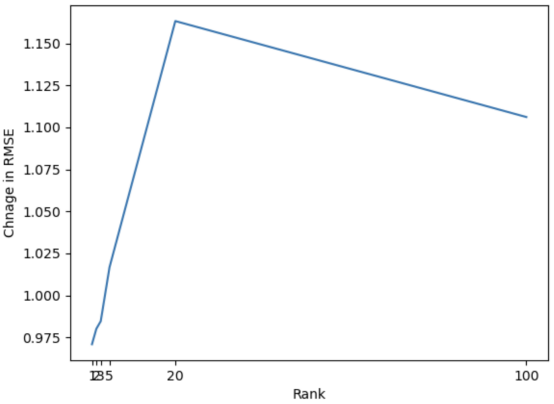



(b) Time v/s RMSE for 100k dataset

For the $500 - 500 - 3$ dataset, the RMSE falls sharply at rank 2 and 3 from rank 1. Following that the RMSE plateaus and rises only slightly. Meanwhile, in the 100k dataset, lowest RMSE is at rank 1 and rises sharply till rank 20 following which it falls as we move towards higher rank but not as low.




(a) Rank v/s RMSE for 500-500 dataset




(b) Rank v/s RMSE for 100k dataset

For both the datasets, the performance is hugely better than the baseline.