

## REPORT

---

- 1) The value of the objective function at different epochs is shown below:

```
Value of Objective Function at Epoch 1: -4.633974371464905E7
Value of Objective Function at Epoch 2: -4648858.464687863
Value of Objective Function at Epoch 3: -229910.176737504
Value of Objective Function at Epoch 4: -99816.6579503876
Value of Objective Function at Epoch 5: -91618.15985194322
Value of Objective Function at Epoch 6: -91391.57164093175
Value of Objective Function at Epoch 7: -91385.77989984387
Value of Objective Function at Epoch 8: -91385.65537992776
Value of Objective Function at Epoch 9: -91385.65342605243
Value of Objective Function at Epoch 10: -91385.65340059137
Value of Objective Function at Epoch 11: -91385.65340028696
Value of Objective Function at Epoch 12: -91385.65340028565
Value of Objective Function at Epoch 13: -91385.6534002848
Value of Objective Function at Epoch 14: -91385.65340028609
Value of Objective Function at Epoch 15: -91385.65340028524
Value of Objective Function at Epoch 16: -91385.65340028482
Value of Objective Function at Epoch 17: -91385.65340028396
Value of Objective Function at Epoch 18: -91385.65340028258
Value of Objective Function at Epoch 19: -91385.65340028459
Value of Objective Function at Epoch 20: -91385.65340028456
```

As seen above, value of the objective function keeps on increasing at each epoch (becomes less negative). Furthermore, it starts converging to some maxima as the epoch increases.

- 2) TODO
- 3) TODO
- 4) TODO
- 5) No, we are not guaranteed to get the same  $W$  and  $H$ , every time we run DSGD with a new random initialization. The loss function for the matrix factorization is a non – convex function. Hence, there a number of local minima. Hence, if one uses different a random initialization, it may converge to different local minima.
- 6) The decompositions represented by the options **a**, **c** and **d** are valid decompositions. A decomposition is valid if and only if no two blocks share any row or column. This is a required condition for DSGD to work efficiently. In option b, many of the blocks have some

or the other row/column in common.

- 7) The option **d** represents a valid decomposition. However, it is a suboptimal decomposition as not all rows/columns are distributed in one pass.

8)

- **Did you receive any help whatsoever from anyone in solving this assignment?**  
No
- **Did you give any help whatsoever to anyone in solving this assignment?**  
No