

Topics on High Dimensional Data Analytics

Final Exam

Q3

Ian Dover

Part 1:

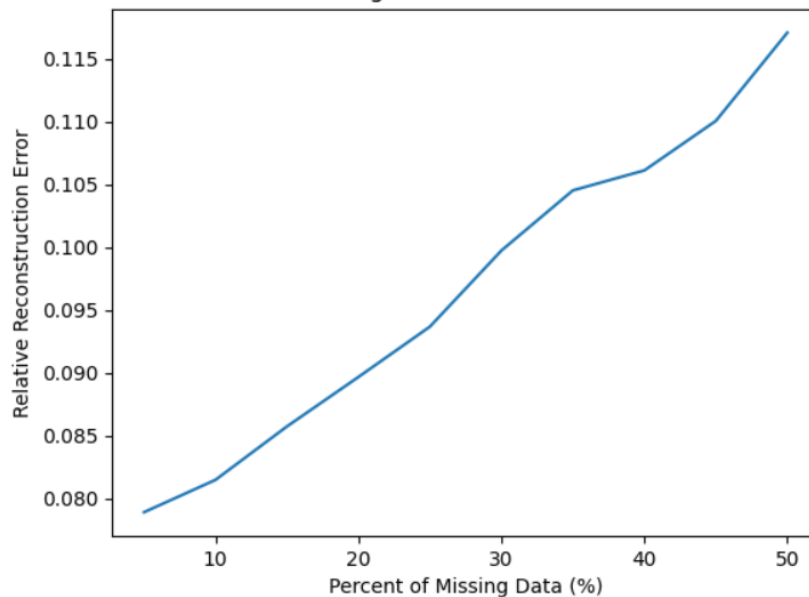
```
#####  
The matrix with missing values:  
0 appears 29958 times.  
1 appears 1415 times.  
2 appears 15673 times.  
3 appears 11400 times.  
4 appears 1524 times.  
5 appears 30 times.  
#####  
The recovered matrix:  
0 appears 4 times.  
1 appears 2940 times.  
2 appears 31098 times.  
3 appears 23315 times.  
4 appears 2626 times.  
5 appears 17 times.  
#####  
The original matrix:  
1 appears 2851 times.  
2 appears 31298 times.  
3 appears 22792 times.  
4 appears 3011 times.  
5 appears 48 times.  
#####
```



Part 2:

```
#####  
The matrix with missing values:  
0 appears 29923 times.  
1 appears 814 times.  
2 appears 9448 times.  
3 appears 17446 times.  
4 appears 2321 times.  
5 appears 48 times.  
#####  
The recovered matrix:  
1 appears 2942 times.  
2 appears 28658 times.  
3 appears 25613 times.  
4 appears 2758 times.  
5 appears 29 times.  
#####  
The original matrix:  
1 appears 2851 times.  
2 appears 31298 times.  
3 appears 22792 times.  
4 appears 3011 times.  
5 appears 48 times.  
#####
```

Part 2: Percent of Missing Data VS. Relative Reconstruction Error



Comparison with results of Part 1): “Missing not at random” and “missing at random” generate similar results on the relative reconstruction error. It appears that the “missing not at random” results in a reconstructed matrix with no zeros as opposed to the reconstructed matrix in “missing at random” which had 4 zeros.