Multivariate Analysis: Continuous Assessment 2

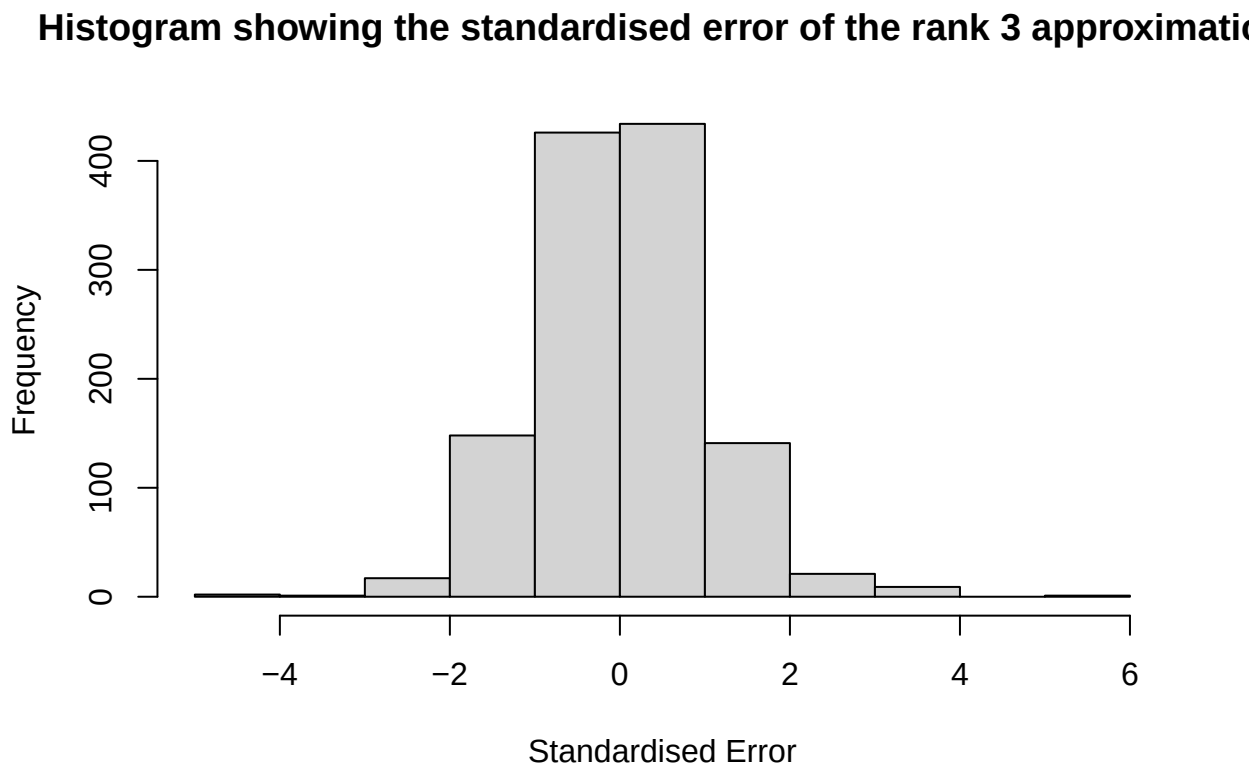
Molly Ryan (RYNERI002) Tinotenda Mutsemi (MTSTIN007)

2024-02-23

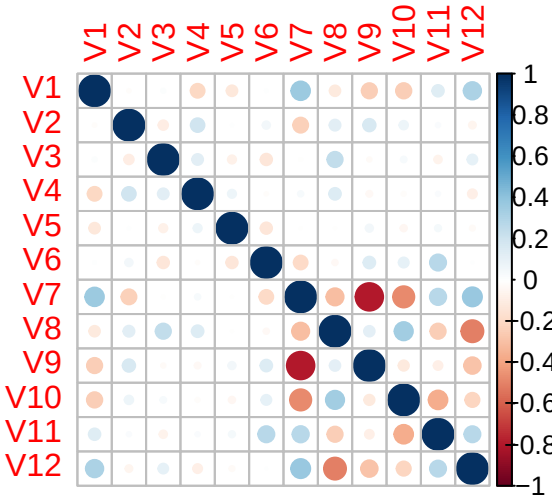
Singular Value Decomposition was used to find lower rank approximations of X, ranging from rank 12 (full rank) down to rank 1. Then the error for each approximation of X was calculated.

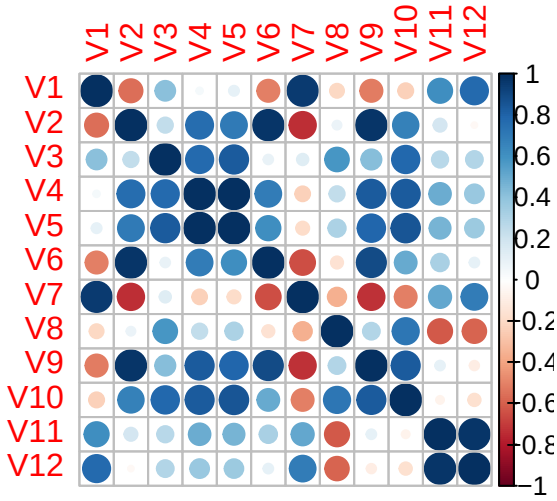
The mean vector of the approximation error for the rank 3 approximation can be seen below.

A histogram plotting the standardised error for the rank 3 approximation matrix follows. It can be seen that the error is approximately normally distributed.



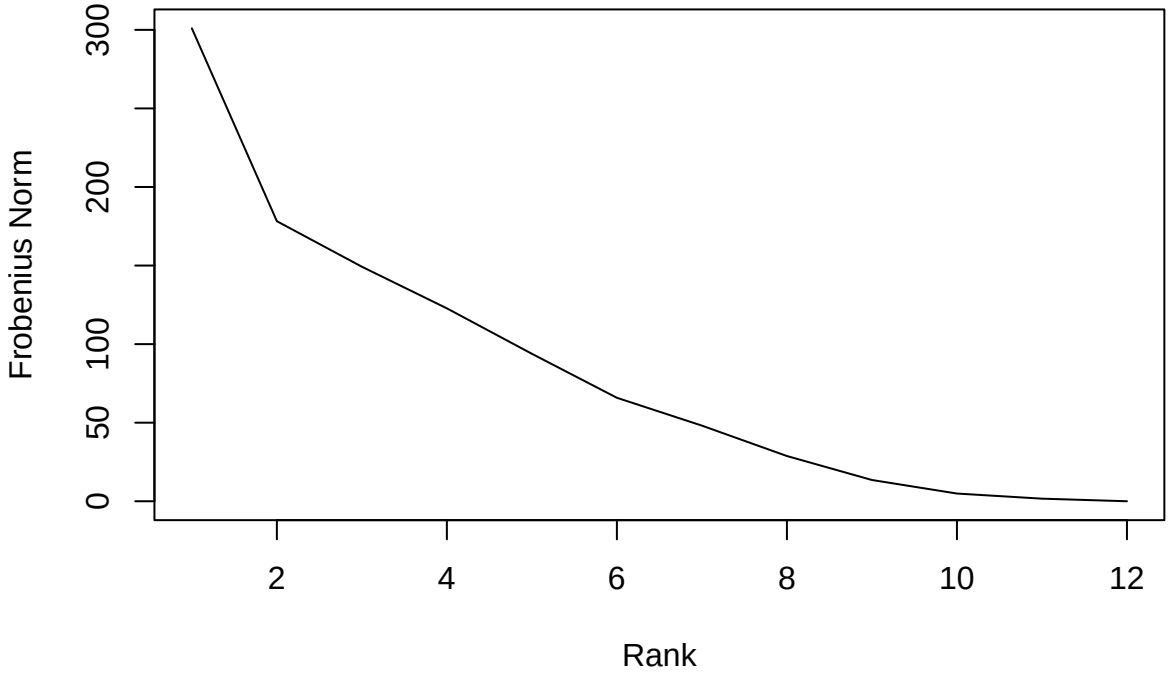
When comparing the correlations between the variables in X to the correlations between the variables in the rank 3 approximation, a clear difference can be seen. This is clear in the heatmaps plotted below. Indeed, it appears that the lower the rank of the approximation, the less correlation is preserved.






```
##  [1] 3.009667e+02 1.782370e+02 1.492686e+02 1.227721e+02 9.384240e+01
##  [6] 6.584375e+01 4.813117e+01 2.877578e+01 1.352022e+01 4.891382e+00
## [11] 1.646137e+00 1.836098e-11
```

Lastly, the Frobenius Norm of the matrix can be seen to decrease with increasing rank (k). This relationship is plotted below.


Frobenius Norm
0
50
100
200
300
2
4
6
8
10
12
Rank