ACM 104 Supplement Exercises

Eduardo da Veiga Beltrame

October 2019

This list of exercises is put together as a "supplement" to ACM104, so that my B- grade be counted towards the bioengineering course requirements, which typically requires a strict B.

On the exercises below, by "explain X" I mean "write a page or two about it". The references listed are ones I found while putting together this list and deciding what was important and that I should brush up on. If additional references are used I will mention them on the write-up.

1 Perron-Frobenius theorem

State and prove the Perron-Frobenius theorem. Make sure to define all terms and state all theorems used.

2 Spectral theorem

State and prove the Spectral theorem. Make sure to define all terms and state all theorems used.

3 Singular Value Decomposition (SVD)

Part a) Explain SVD and truncated SVD.
Part b) Implement SVD or truncated SVD in Python and compare results with sklearn truncatedSVD:
`https://scikit-learn.org/stable/modules/generated/sklearn.decomposition.TruncatedSVD.html#sklearn.decomposition.TruncatedSVD`

Some references:
`https://www.cs.cmu.edu/~venkatg/teaching/CStheory-infoage/book-chapter-4.pdf`

4 Principal Component Analysis (PCA)

Part a) Explain PCA
Part b) Implement PCA in Python. Test results on the iris dataset and compare with the sklearn PCA.

Iris dataset:
https://scikit-learn.org/stable/auto_examples/datasets/plot_iris_dataset.html#sphx-glr-auto-examples-datasets-plot-iris-dataset-py

sklearn PCA:
https://scikit-learn.org/stable/modules/decomposition.html#pca
https://scikit-learn.org/stable/modules/generated/sklearn.decomposition.PCA.html

5 Non-Negative Matrix Factorization (NMF)

Part a) Explain NMF
Part a) Implement NMF in Python and compare with sklearn

Some references:
https://doi.org/10.1016/B978-0-12-385022-5.00013-0 **sklearn NMF:** https://scikit-learn.org/stable/modules/generated/sklearn.decomposition.NMF.html#sklearn.decomposition.NMF

6 Canonical Correlation Analysis (CCA)

Part a) Explain CCA

Some references:
https://doi.org/10.1016/B978-0-12-385022-5.00013-0 https://www.cs.cmu.edu/~tom/10701_sp11/slides/CCA_tutorial.pdf

7 Multidimensional scaling (MDS)

Part a) Explain MDS

Some references:
http://www.cs.umd.edu/~djacobs/CMSC828/MDSexplain.pdf
https://www.stat.pitt.edu/sungkyu/course/2221Fall13/lec8_mds_combined.pdf
http://www.mathpsy.uni-tuebingen.de/wickelmaier/pubs/Wickelmaier2003SQRU.pdf

8 Page Rank

Part a) Convert the following dataset into a stochastic matrix:

`https://sites.google.com/site/cxnets/usairtransportationnetwork`
Undirected weighted network as obtained by considering the 500 US airports with the largest amount of traffic from publicly available data. Nodes represent US airports and edges represent air travel connections among them. The file reports the anonymized list of connected pairs of nodes and the weight associated to the edge, expressed in terms of number of available seats on the given connection on a yearly basis.

Part b) Implement Page Rank in Python and test it with the airport data above.

Some references:
`http://statweb.stanford.edu/~tibs/sta306bfiles/pagerank/ryan/01-24-pr.pdf`
`http://ilpubs.stanford.edu:8090/422/1/1999-66.pdf`
`http://home.ie.cuhk.edu.hk/~wkshum/papers/pagerank.pdf`
`https://www.cs.princeton.edu/~chazelle/courses/BIB/pagerank.htm`
`http://pi.math.cornell.edu/~mec/Winter2009/RalucaRemus/Lecture3/lecture3.html`