Singular Value Decomposition (SVD)

What is Singular Value Decomposition? 1
Output 2
Positive & Negative Decomposition graphs 2
CSV 3
References 3

What is Singular Value Decomposition?

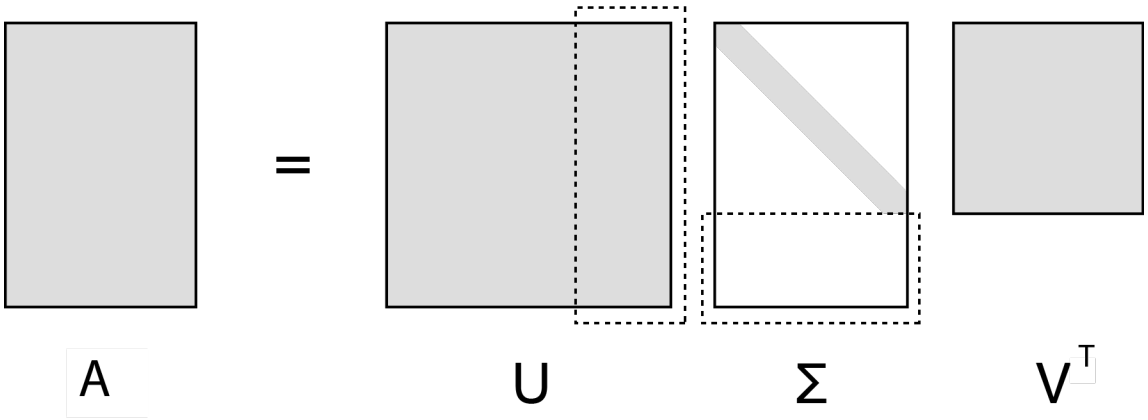
The Singular Value Decomposition (SVD) of a matrix is a factorization of that matrix into three matrices. It has some interesting algebraic properties and conveys important geometrical and theoretical insights about linear transformations.

The SVD of the m × n matrix A is given by the formula:

$$A = U\Sigma V^T$$

where:

- U: m × m orthogonal matrix
- Σ: m × n diagonal matrix with nonnegative real entries on the diagonal.
- V^T: n × n orthogonal matrix



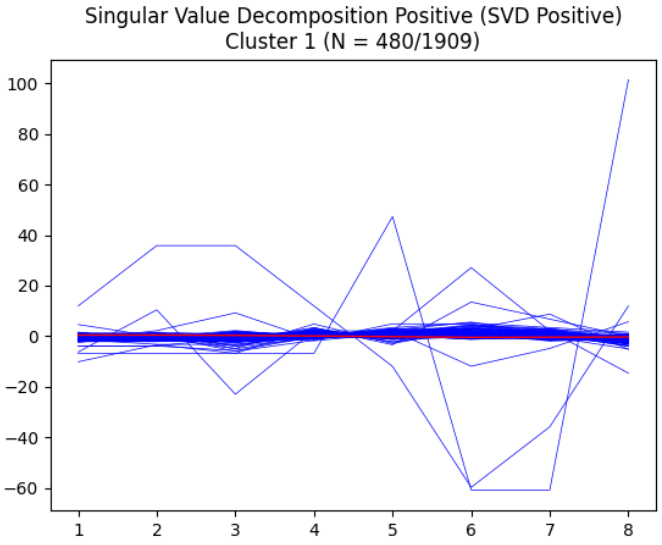
The goal of SVD is to reduce the dimension of your problem in order to get rid of any noise impacting your analysis. if A is a matrix representing some dataset, in our case a Stanford CoreNLP Sentiment Score table, and you provide a small integer k, then the singular value decomposition computes the rank k matrix B_k which best approximates A. And since now we're comfortable identifying a data matrix with the subspace defined by the data, this is the same thing as finding the k-dimensional subspace of the sentiments scores in A which is the best approximation of the data. Here k would be the number of clusters.

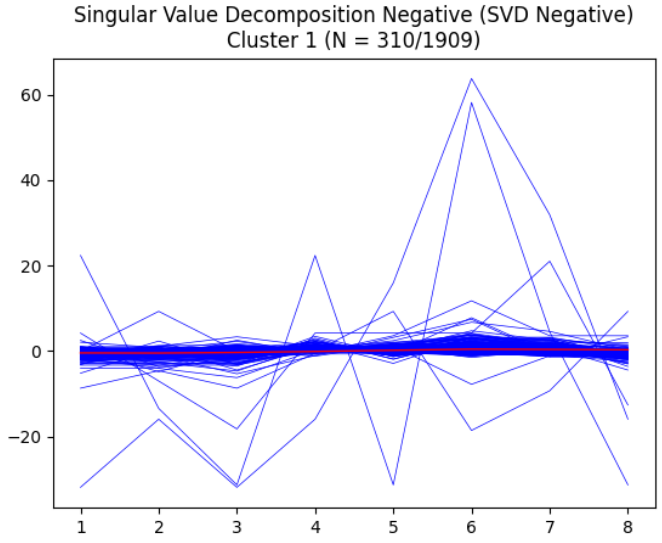
Output

Positive & Negative Decomposition graphs

In the NLP Suite, the Singular Value decomposition is displayed by two plots showing the Sentiment Score breakdown done by the decomposition. SVD has broken down sentiment scores into positive and negative sentiment scores for each cluster.

Here is a quick preview of the output of the Shape of Stories tool for a single cluster:

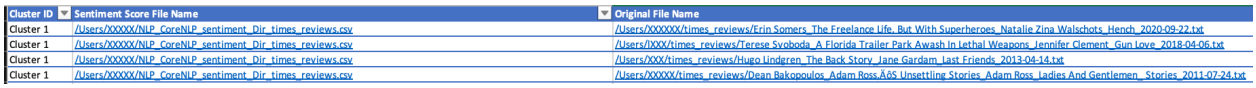




You will notice that unsurprisingly Negative and Positive decomposition seem to follow opposite symmetric path which is to be expected since they model opposite sentiments!

CSV

In addition to the 2 above graphs, you can use the csv file generated to find the content of each cluster in both the Positive and Negative Decompositions *SVD Positive Documents.csv* and *SVD Negative Documents.csv.*



Cluster ID	Sentiment Score File Name	Original File Name
Cluster 1	/Users/XXXXX/NLP_CoreNLP_sentiment_Dir_times_reviews.csv	/Users/XXXXXX/times_reviews/Erin Somers_The Freelance Life. But With Superheroes_Natalie Zina Walschots_Hench_2020-09-22.txt
Cluster 1	/Users/XXXXX/NLP_CoreNLP_sentiment_Dir_times_reviews.csv	/Users/IXXX/times_reviews/Terese Svoboda_A Florida Trailer Park Awash In Lethal Weapons_Jennifer Clement_Gun Love_2018-04-06.txt
Cluster 1	/Users/XXXXX/NLP_CoreNLP_sentiment_Dir_times_reviews.csv	/Users/XXX/times_reviews/Hugo Lindgren_The Back Story_Jane Gardam_Last Friends_2013-04-14.txt
Cluster 1	/Users/XXXXX/NLP_CoreNLP_sentiment_Dir_times_reviews.csv	/Users/XXXXX/times_reviews/Dean Bakopoulos_Adam Ross.Â©S Unsettling Stories_Adam Ross_Ladies And Gentlemen_ Stories_2011-07-24.txt

The CSV file contains the Cluster ID of each text as well as the directory of the original .txt file and the directory of the Sentiment Scores of the .txt file. You can use the csv file generated to find the content of each cluster (the file name and cluster number) for the SVD results.

References

Burrows, J.F. 1987. "Word-Patterns and Story-Shapes: The Statistical Analysis of Narrative Style." *Literary and Linguistic Computing*, Vol. 2, No. 2, pp. 61-70.

Franzosi, Roberto. 2004. *From Words to Numbers*. Cambridge, UK: Cambridge University Press.

Franzosi, Roberto. 2010. *Quantitative Narrative Analysis*. Thousand Oaks, CA: Sage.

Reagan, Andrew J., Lewis Mitchell, Dilan Kiley, Christopher M. Danforth, and Peter Sheridan Dodds. 2016. "The Emotional Arcs of Stories Are Dominated by Six Basic Shapes". *EPJ Data Science*, Vol. 5, No. 31, pp. 1-12.

Vonnegut, Kurt. https://www.youtube.com/watch?v=oP3c1h8v2ZQ

Vonnegut, Kurt. https://www.youtube.com/watch?v=GOGru_4z1Vc

Vonnegut, Kurt. 2005. "Here is a Lesson in Creative Writing." In: pp. 23-28, Kurt Vonnegut, *A Man Without a Country. Edited by Daniel Simon.* New York: Seven Stories Press.

Nakagome, S. (2019, February 10). *Linear Algebra 101-part 9: Singular value decomposition (SVD)*. Medium. Retrieved April 17, 2022, from https://medium.com/sho-jp/linear-algebra-101-part-9-singular-value-decomposition-svd-a6c53ed2319e