

This exercise is taken from the article “An Introduction to Latent Semantic Analysis” by T. Landauer et al..

Nine documents are given by:

Example of text data: Titles of Some Technical Memos	
c1:	<i>Human machine interface for ABC computer applications</i>
c2:	<i>A survey of user opinion of computer system response time</i>
c3:	<i>The EPS user interface management system</i>
c4:	<i>System and human system engineering testing of EPS</i>
c5:	<i>Relation of user perceived response time to error measurement</i>
m1:	<i>The generation of random, binary, ordered trees</i>
m2:	<i>The intersection graph of paths in trees</i>
m3:	<i>Graph minors IV: Widths of trees and well-quasi-ordering</i>
m4:	<i>Graph minors: A survey</i>

This results in the following word-document matrix:

	c1	c2	c3	c4	c5	m1	m2	m3	m4
human	1	0	0	1	0	0	0	0	0
interface	1	0	1	0	0	0	0	0	0
computer	1	1	0	0	0	0	0	0	0
user	0	1	1	0	1	0	0	0	0
system	0	1	1	2	0	0	0	0	0
response	0	1	0	0	1	0	0	0	0
time	0	1	0	0	1	0	0	0	0
EPS	0	0	1	1	0	0	0	0	0
survey	0	1	0	0	0	0	0	0	1
trees	0	0	0	0	0	1	1	1	0
graph	0	0	0	0	0	0	1	1	1
minors	0	0	0	0	0	0	0	1	1

Please calculate the SVD using numpy. Calculate the document-document correlations for a two dimensional subspace and compare it to the document-document correlations from the original word-document matrix.