LSA is SVD

Salfo Bikienga

October 03, 2018

Table 1: Example matrix of words counts

	college	education	family	health	medicaid
document.1	4	6	0	2	2
document.2	0	0	4	8	12
document.3	6	9	1	5	6
document.4	2	3	3	7	10
document.5	0	0	3	6	9
document.6	4	6	1	4	5

```
svd_W = svd(W) # Singular Value Decomposition applied to the data
svd_W
```

```
## $d
## [1] 26.449198194252684146 12.427385681676595297  0.000000000000001342
## [4]  0.000000000000000636  0.000000000000000150
##
## $u
##         [,1]   [,2]    [,3]    [,4]   [,5]
## [1,] -0.201  0.454 -0.5762  0.5340  0.165
## [2,] -0.521 -0.471 -0.6059 -0.3134 -0.138
## [3,] -0.431  0.564  0.2008 -0.3494 -0.556
## [4,] -0.491 -0.126  0.3540 -0.1690  0.616
## [5,] -0.391 -0.353  0.3544  0.6785 -0.367
## [6,] -0.331  0.336  0.0981  0.0746  0.361
##
## $v
##         [,1]   [,2]   [,3]    [,4]    [,5]
## [1,] -0.215  0.506 -0.158 -0.1736 -0.8014
## [2,] -0.323  0.760  0.125 -0.0352  0.5495
## [3,] -0.208 -0.195 -0.460 -0.8169  0.2001
## [4,] -0.523 -0.137 -0.641  0.5416  0.0628
## [5,] -0.730 -0.332  0.581 -0.0888 -0.1086
```

```
library(lsa) # Needed to access the lsa function
lsa_W = lsa(W)
lsa_W # print the output
```

```
## $tk
##               [,1]   [,2]    [,3]    [,4]   [,5]
## document.1 -0.201  0.454 -0.5762  0.5340  0.165
## document.2 -0.521 -0.471 -0.6059 -0.3134 -0.138
## document.3 -0.431  0.564  0.2008 -0.3494 -0.556
## document.4 -0.491 -0.126  0.3540 -0.1690  0.616
## document.5 -0.391 -0.353  0.3544  0.6785 -0.367
## document.6 -0.331  0.336  0.0981  0.0746  0.361
```

```
##
## $dk
##              [,1]   [,2]   [,3]    [,4]    [,5]
## college   -0.215  0.506 -0.158 -0.1736 -0.8014
## education -0.323  0.760  0.125 -0.0352  0.5495
## family    -0.208 -0.195 -0.460 -0.8169  0.2001
## health    -0.523 -0.137 -0.641  0.5416  0.0628
## medicaid  -0.730 -0.332  0.581 -0.0888 -0.1086
##
## $sk
## [1] 26.449198194252684146 12.427385681676595297  0.000000000000001342
## [4]  0.000000000000000636  0.000000000000000150
##
## attr(,"class")
## [1] "LSAspace"
```

```
# lsa_W$tk%*%diag(lsa_W$sk) # to check equality of pc-scores
```