DATA 612 Project 3

Brian Liles

June 25, 2019

Matrix Factorization Methods

Singular Decomposition in R

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

According to Wikipedia, the **singular value decomposition** is a factorization of a real or complex matrix. In order to conduct this tasks the data must be complete.

```
# create a matrix entitled a
a <- matrix(c(4,8,3,8,7,
              3,1,1,0,2,
              1,1,1,0,6,
              0,0,1,7,0,
              3,7,6,8,0,
              2,4,3,5,3), nrow = 6, byrow = TRUE)

# create colnames
colnames(a) <- c("The Wire","Breaking Bad","The Sopranos","Game of Thrones","The Shield")

# create rownames
rownames(a) <- c("User1","User2","User3","User4","User5","User6")
a
```

```
##       The Wire Breaking Bad The Sopranos Game of Thrones The Shield
## User1        4            8            3               8          7
## User2        3            1            1               0          2
## User3        1            1            1               0          6
## User4        0            0            1               7          0
## User5        3            7            6               8          0
## User6        2            4            3               5          3
```

```
# obtain the rank of the matrix
qr(a)$rank
```

```
## [1] 5
```

```
# conduct the svd of the matrix a
svd_a <- svd(a)
```

Column **U** is the eigenvector of AA^T

```
# view column U of matrix a
svd_a$u
```

```
##            [,1]        [,2]        [,3]       [,4]        [,5]
## [1,] -0.6673607  0.34313946  0.12367932 -0.4922291 -0.39905344
## [2,] -0.1132145  0.26256139 -0.24545392  0.7342550 -0.56465371
## [3,] -0.1515994  0.63443197  0.20899580  0.2862133  0.60952621
## [4,] -0.2335687 -0.43412845  0.79548503  0.3186640 -0.08520961
## [5,] -0.5648297 -0.47044428 -0.49744774  0.1503104  0.30926814
## [6,] -0.3811075  0.03205561  0.02293173  0.1119018  0.21792860
```

Column **V** is the eigenvector of $A^T A$

```
# view column V of matrix a
svd_a$v
```

```
##            [,1]        [,2]       [,3]        [,4]        [,5]
## [1,] -0.2703775  0.17860233 -0.3193558  0.55372390 -0.69742305
## [2,] -0.5333983  0.05888475 -0.5262611 -0.65697418 -0.05876187
## [3,] -0.3385027 -0.15229087 -0.3855744  0.50979659  0.67354527
## [4,] -0.6448776 -0.48086139  0.5814204  0.02540000 -0.11920741
## [5,] -0.3345610  0.84274172  0.3665294  0.03517166  0.20560761
```

The diagonal matrix Σ are the square roots of the non-zero eigenvalues of AA^T and $A^T A$

```
# view the diagonal matrix
(d <- diag(svd_a$d))
```

```
##         [,1]     [,2]   [,3]     [,4]     [,5]
## [1,] 20.7761 0.000000 0.0000 0.000000 0.000000
## [2,]  0.0000 8.104331 0.0000 0.000000 0.000000
## [3,]  0.0000 0.000000 4.6316 0.000000 0.000000
## [4,]  0.0000 0.000000 0.0000 2.157748 0.000000
## [5,]  0.0000 0.000000 0.0000 0.000000 1.888362
```

Test and Training Data Sets

```
set.seed(50)
sample <- sample.int(n = nrow(a), size = floor(.75*nrow(a)), replace = F)
aTrain <- a[sample,]
aTest <- a[-sample,]
```

Conduct SVD for the training and test data sets

```r
# conduct the svd of the matrix a
(svd_aTrain <- svd(aTrain))
```

```
## $d
## [1] 20.0915803  7.3287615  2.0677045  0.5850251
##
## $u
##            [,1]        [,2]       [,3]       [,4]
## [1,] -0.5812976 -0.62271620  0.4142692 -0.3204664
## [2,] -0.1630037  0.70093821  0.6613893 -0.2113755
## [3,] -0.6931840  0.34726537 -0.6009625 -0.1942852
## [4,] -0.3937244  0.01780208  0.1725955  0.9027043
##
## $v
##            [,1]        [,2]        [,3]        [,4]
## [1,] -0.2721081  0.03512944 -0.07470228 -0.24700954
## [2,] -0.5650355 -0.11035206 -0.26891881 -0.68049133
## [3,] -0.3440005 -0.26473186  0.90046880 -0.01524136
## [4,] -0.6054514 -0.28853391 -0.30496101  0.67605440
## [5,] -0.3489762  0.91282724  0.13511833  0.13651371
```

```r
# conduct the svd of the matrix a
(svd_aTest <- svd(aTest))
```

```
## $d
## [1] 7.073086 3.869296
##
## $u
##            [,1]        [,2]
## [1,] 0.02853652  0.99959275
## [2,] 0.99959275 -0.02853652
##
## $v
##             [,1]        [,2]
## [1,] 0.012103564  0.77501908
## [2,] 0.004034521  0.25833969
## [3,] 0.145357944  0.25096457
## [4,] 0.989263961 -0.05162583
## [5,] 0.008069043  0.51667939
```