Matrix-Decompositions

Chase Enzweiler

9/11/2017

SVD on Raw Data

use svd() to compute the SVD of USArrests dataset in R

```
svd_arrests <- svd(USArrests)
```

take the output of svd() and create the matrices U, D, V

```
U <- svd_arrests$u
D <- svd_arrests$d
V <- svd_arrests$v
```

Confirm that the data of USArrests can be obtained as a product of: UDV^T

```
U %*% diag(D) %*% t(V) - USArrests
```

```
##                      Murder     Assault     UrbanPop        Rape
## Alabama        1.421085e-14 1.421085e-13 3.552714e-14 4.618528e-14
## Alaska         1.065814e-14 3.410605e-13 1.421085e-14 0.000000e+00
## Arizona        1.243450e-14 2.842171e-13 4.263256e-14 1.776357e-14
## Arkansas       1.243450e-14 2.842171e-13 6.394885e-14 2.842171e-14
## California     1.243450e-14 2.842171e-13 9.947598e-14 2.842171e-14
## Colorado       8.881784e-15 2.842171e-13 8.526513e-14 2.842171e-14
## Connecticut    2.664535e-15 1.136868e-13 7.105427e-14 2.486900e-14
## Delaware       9.769963e-15 3.126388e-13 5.684342e-14 3.552714e-14
## Florida        1.598721e-14 3.410605e-13 7.105427e-14 4.618528e-14
## Georgia        7.105427e-15 2.557954e-13 9.947598e-14 4.263256e-14
## Hawaii         0.000000e+00 8.526513e-14 1.136868e-13 2.486900e-14
## Idaho          3.552714e-15 1.136868e-13 3.552714e-14 1.421085e-14
## Illinois       8.881784e-15 2.557954e-13 5.684342e-14 3.907985e-14
## Indiana        4.440892e-15 1.563194e-13 7.105427e-14 2.842171e-14
## Iowa           8.881784e-16 6.394885e-14 6.394885e-14 1.598721e-14
## Kansas         4.440892e-15 1.278977e-13 8.526513e-14 2.131628e-14
## Kentucky       3.552714e-15 1.421085e-13 7.815970e-14 2.842171e-14
## Louisiana      1.598721e-14 2.842171e-13 7.105427e-14 4.973799e-14
## Maine          2.220446e-15 7.105427e-14 4.973799e-14 1.687539e-14
## Maryland       1.776357e-14 3.979039e-13 7.105427e-14 3.197442e-14
## Massachusetts  5.329071e-15 1.705303e-13 7.105427e-14 2.842171e-14
## Michigan       1.065814e-14 2.557954e-13 4.263256e-14 2.842171e-14
## Minnesota      8.881784e-16 8.526513e-14 5.684342e-14 2.131628e-14
## Mississippi    1.421085e-14 3.410605e-13 3.552714e-14 4.618528e-14
## Missouri       8.881784e-15 1.989520e-13 8.526513e-14 2.486900e-14
## Montana        3.552714e-15 1.136868e-13 5.684342e-14 2.486900e-14
## Nebraska       3.552714e-15 9.947598e-14 5.684342e-14 1.776357e-14
## Nevada         8.881784e-15 2.842171e-13 7.105427e-14 2.131628e-14
## New Hampshire  4.440892e-16 4.263256e-14 4.263256e-14 1.598721e-14
## New Jersey     3.552714e-15 1.705303e-13 9.947598e-14 3.197442e-14
## New Mexico     1.065814e-14 2.842171e-13 4.263256e-14 2.842171e-14
## New York       1.065814e-14 2.842171e-13 7.105427e-14 3.552714e-14
```

```
## North Carolina 2.131628e-14 4.547474e-13 2.131628e-14 3.197442e-14
## North Dakota   7.771561e-16 4.973799e-14 4.973799e-14 1.332268e-14
## Ohio           5.329071e-15 1.563194e-13 9.947598e-14 3.197442e-14
## Oklahoma       7.105427e-15 1.705303e-13 7.105427e-14 2.842171e-14
## Oregon         7.105427e-15 1.705303e-13 7.105427e-14 2.131628e-14
## Pennsylvania   3.552714e-15 1.278977e-13 9.947598e-14 3.197442e-14
## Rhode Island   5.329071e-15 1.705303e-13 8.526513e-14 3.375078e-14
## South Carolina 1.598721e-14 3.410605e-13 4.263256e-14 3.552714e-14
## South Dakota   2.664535e-15 9.947598e-14 4.263256e-14 1.598721e-14
## Tennessee      8.881784e-15 1.989520e-13 7.105427e-14 2.842171e-14
## Texas          1.065814e-14 2.273737e-13 9.947598e-14 3.907985e-14
## Utah           3.108624e-15 1.563194e-13 8.526513e-14 1.776357e-14
## Vermont        2.220446e-15 6.394885e-14 3.552714e-14 8.881784e-15
## Virginia       5.329071e-15 1.705303e-13 5.684342e-14 2.486900e-14
## Washington     1.776357e-15 1.421085e-13 5.684342e-14 1.421085e-14
## West Virginia  2.664535e-15 9.947598e-14 4.263256e-14 1.776357e-14
## Wisconsin      4.440892e-16 5.684342e-14 8.526513e-14 2.131628e-14
## Wyoming        7.105427e-15 2.273737e-13 5.684342e-14 3.019807e-14
```

confirm that the sum equals USArrests

```
X <- D[1] * U[,1] %*% t(V[,1])

for(i in 2:4){

  X <- X + D[i] * U[,i] %*% t(V[,i])
}

X - U %*% diag(D) %*% t(V)
```

```
##               [,1]          [,2]          [,3]          [,4]
##  [1,]  0.000000e+00 -2.842171e-14 -7.105427e-15  0.000000e+00
##  [2,]  3.552714e-15  0.000000e+00 -1.421085e-14  0.000000e+00
##  [3,]  3.552714e-15  0.000000e+00 -1.421085e-14  0.000000e+00
##  [4,]  0.000000e+00  0.000000e+00  0.000000e+00  3.552714e-15
##  [5,]  0.000000e+00  0.000000e+00  0.000000e+00  0.000000e+00
##  [6,]  2.664535e-15  0.000000e+00  0.000000e+00  0.000000e+00
##  [7,] -8.881784e-16  0.000000e+00  0.000000e+00  0.000000e+00
##  [8,]  0.000000e+00  0.000000e+00  0.000000e+00 -1.776357e-15
##  [9,] -1.776357e-15  0.000000e+00  0.000000e+00  0.000000e+00
## [10,]  0.000000e+00  2.842171e-14  0.000000e+00  0.000000e+00
## [11,]  0.000000e+00  0.000000e+00  0.000000e+00  0.000000e+00
## [12,]  4.440892e-16  0.000000e+00  0.000000e+00  3.552714e-15
## [13,]  1.776357e-15  0.000000e+00  0.000000e+00  0.000000e+00
## [14,]  0.000000e+00 -1.421085e-14  0.000000e+00 -3.552714e-15
## [15,]  4.440892e-16  1.421085e-14  0.000000e+00  1.776357e-15
## [16,]  0.000000e+00  1.421085e-14  0.000000e+00  3.552714e-15
## [17,]  1.776357e-15  1.421085e-14  0.000000e+00  3.552714e-15
## [18,]  0.000000e+00 -2.842171e-14  0.000000e+00 -7.105427e-15
## [19,]  0.000000e+00  0.000000e+00  0.000000e+00  0.000000e+00
## [20,]  0.000000e+00 -5.684342e-14 -1.421085e-14  0.000000e+00
## [21,]  1.776357e-15  2.842171e-14  0.000000e+00  0.000000e+00
## [22,]  1.776357e-15  0.000000e+00  1.421085e-14  7.105427e-15
## [23,]  4.440892e-16  0.000000e+00  1.421085e-14  0.000000e+00
## [24,]  0.000000e+00  0.000000e+00  1.421085e-14  0.000000e+00
```

```
## [25,]  0.000000e+00  2.842171e-14  1.421085e-14  3.552714e-15
## [26,]  0.000000e+00  0.000000e+00  0.000000e+00  0.000000e+00
## [27,] -8.881784e-16  1.421085e-14  0.000000e+00  0.000000e+00
## [28,]  1.776357e-15  2.842171e-14  1.421085e-14  7.105427e-15
## [29,]  8.881784e-16  2.131628e-14  7.105427e-15  0.000000e+00
## [30,]  1.776357e-15  0.000000e+00 -1.421085e-14  0.000000e+00
## [31,]  0.000000e+00  5.684342e-14  1.421085e-14  0.000000e+00
## [32,] -1.776357e-15 -2.842171e-14  0.000000e+00 -3.552714e-15
## [33,] -1.776357e-15 -5.684342e-14  1.421085e-14  0.000000e+00
## [34,] -1.110223e-16  0.000000e+00  0.000000e+00  1.776357e-15
## [35,] -8.881784e-16 -1.421085e-14  0.000000e+00  0.000000e+00
## [36,] -8.881784e-16  0.000000e+00 -1.421085e-14 -7.105427e-15
## [37,]  0.000000e+00  0.000000e+00  0.000000e+00  0.000000e+00
## [38,]  0.000000e+00  2.842171e-14  0.000000e+00  0.000000e+00
## [39,]  1.776357e-15  0.000000e+00  0.000000e+00  3.552714e-15
## [40,]  0.000000e+00  0.000000e+00  7.105427e-15  0.000000e+00
## [41,] -8.881784e-16  0.000000e+00  7.105427e-15  0.000000e+00
## [42,]  0.000000e+00 -2.842171e-14  0.000000e+00  0.000000e+00
## [43,] -1.776357e-15 -2.842171e-14  0.000000e+00 -3.552714e-15
## [44,]  0.000000e+00  0.000000e+00  0.000000e+00  0.000000e+00
## [45,]  0.000000e+00  7.105427e-15  0.000000e+00  0.000000e+00
## [46,]  0.000000e+00  2.842171e-14  0.000000e+00  0.000000e+00
## [47,]  1.776357e-15 -2.842171e-14  0.000000e+00  7.105427e-15
## [48,]  0.000000e+00  0.000000e+00 -7.105427e-15  0.000000e+00
## [49,]  4.440892e-16  0.000000e+00  0.000000e+00 -1.776357e-15
## [50,] -8.881784e-16 -2.842171e-14 -7.105427e-15  0.000000e+00
```

create a new variable MA by by adding murder + assault

```r
MA <- USArrests$Murder + USArrests$Assault
```

create a new data frame Arrests2

```r
Arrests2 <- data.frame(USArrests, MA)
```

compute the SVD of Arrests2

```r
svd_arrests2 <- svd(Arrests2)

#compare the values
D
```

```
## [1] 1419.06140  194.82585   45.66134   18.06956
```

```r
svd_arrests2$d
```

```
## [1] 1.993780e+03 2.003163e+02 4.566134e+01 2.166518e+01 2.140724e-13
```

Arrests2 has 5 singular values instead of 4. d in arrests2 are greater than USArrsts for the values except for the third value and the fifth value is very small

what is the rank of arrests2

```r
library(matlib)

R(as.matrix(Arrests2))
```

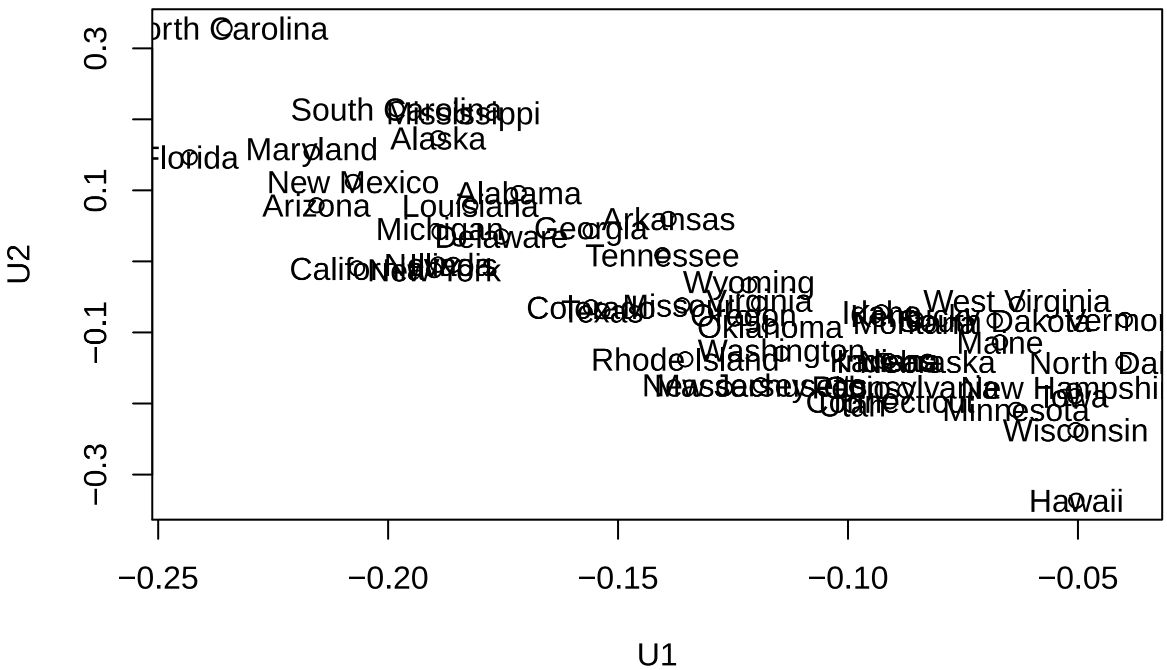
```
## [1] 4
```

Data Visualization

```r
library(ggplot2)
```

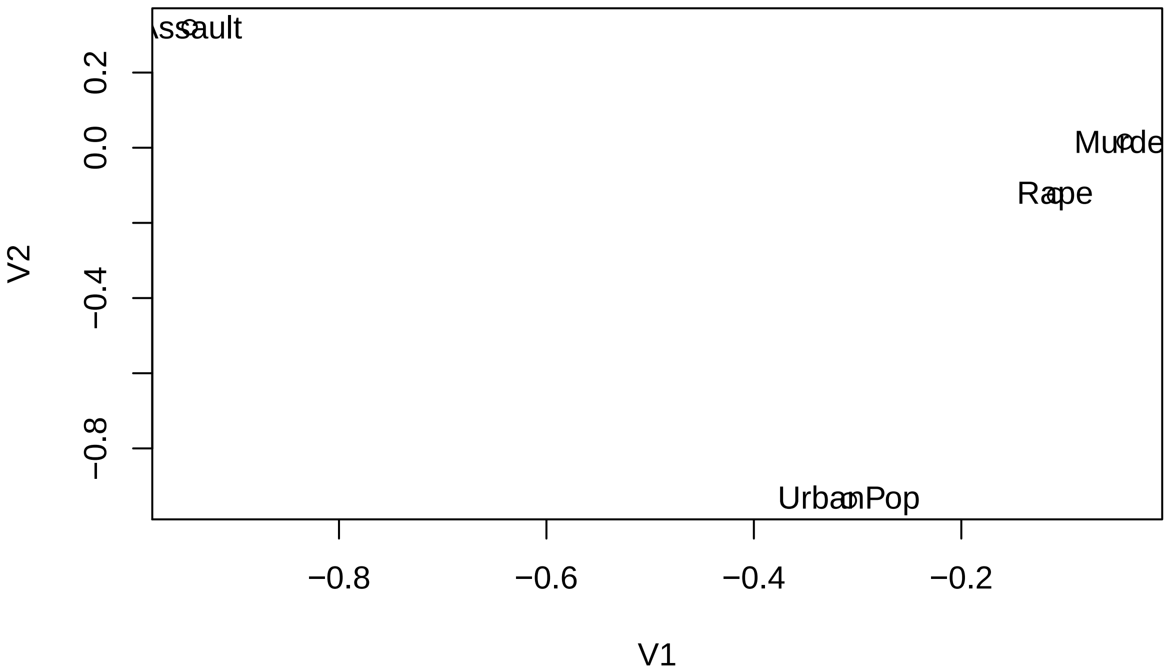
```
## Warning: package 'ggplot2' was built under R version 3.3.2
```

```r
U1 <- U[,1]
U2 <- U[,2]
plot(U1, U2)
text(U1, U2, labels = row.names(USArrests))
```




```r
V1 <- V[,1]
V2 <- V[,2]

plot(V1, V2)
text(V1, V2, labels = colnames(USArrests))
```



Eigenvalue decomposition

without using scale output mean centered data of USArrests

```
mean_matrix <- cbind(rep(mean(USArrests[,1]), dim(USArrests)[1]), rep(mean(USArrests[,2]), dim(USArrests

X <- as.matrix(USArrests) - mean_matrix
```

Calculate the sum of squares and cross product matrix

```
#sum of squares

S <- t(X) %*% X

cov_x <- 1/ (dim(X)[1] - 1) * t(X) %*% X
# equivalence to covariance
cov(X) - cov_x
```

```
##                 Murder      Assault     UrbanPop          Rape
## Murder   3.552714e-15  5.684342e-14 8.881784e-16  3.552714e-15
## Assault  5.684342e-14 -1.818989e-12 0.000000e+00 -1.136868e-13
## UrbanPop 8.881784e-16  0.000000e+00 5.684342e-14  2.131628e-14
## Rape     3.552714e-15 -1.136868e-13 2.131628e-14  1.421085e-14
```

use solve to compute the inverse

```
inverse_S <- solve(S)

inverse_S
```

```
##                 Murder      Assault      UrbanPop          Rape
## Murder    0.0032804923 -1.304887e-04  1.794220e-04 -2.014203e-04
## Assault  -0.0001304887  1.046419e-05 -6.597122e-06 -2.354634e-05
```

```
## UrbanPop  0.0001794220 -6.597122e-06  1.271111e-04 -8.877578e-05
## Rape     -0.0002014203 -2.354634e-05 -8.877578e-05  4.812179e-04
```

```r
evd_s <- eigen(S)

A <- evd_s$values
V <- evd_s$vectors

inverse_S - (V %*% solve(diag(A) %*% t(V)))
```

```
##                 Murder      Assault      UrbanPop          Rape
## Murder    0.0032808455 -1.025046e-04 -2.793823e-04  1.645757e-04
## Assault  -0.0001304569  7.163709e-06  4.567156e-05  3.265896e-03
## UrbanPop  0.0001765722 -1.920141e-05  4.996886e-05 -9.530235e-06
## Rape     -0.0002018954  7.265691e-05  3.636980e-05  4.979897e-04
```

Power Method

```r
# the power method
# returns dominant eigenvalue and corresponding eigenvector
Pow_Method <- function(vect, mat, iter){

  old <- vect

  for(k in 1:iter){

    newr <- (mat %*% old)
    scaled <- lp_norm(newr, p = "max")

    newr <- newr / scaled

    print(paste('iteration = ', k))
    print(newr)
    print(scaled)
    old <- newr

  }


}
A <- matrix(c(5, -4, 3, -14, 4, 6, 11, -4, -3), ncol = 3)
w <- c(1,1,1)

Pow_Method(w, A, 10)
```

```
## [1] "iteration =  1"
##            [,1]
## [1,]  0.3333333
## [2,] -0.6666667
## [3,]  1.0000000
## [1] 6
## [1] "iteration =  2"
##            [,1]
## [1,]  1.0000000
```

```
## [2,] -0.3636364
## [3,] -0.2727273
## [1] 22
## [1] "iteration =  3"
##            [,1]
## [1,]  1.0000000
## [2,] -0.6153846
## [3,]  0.2307692
## [1] 7.090909
## [1] "iteration =  4"
##             [,1]
## [1,]  1.00000000
## [2,] -0.45714286
## [3,] -0.08571429
## [1] 16.15385
## [1] "iteration =  5"
##             [,1]
## [1,]  1.00000000
## [2,] -0.52459016
## [3,]  0.04918033
## [1] 10.45714
## [1] "iteration =  6"
##             [,1]
## [1,]  1.00000000
## [2,] -0.48854962
## [3,] -0.02290076
## [1] 12.88525
## [1] "iteration =  7"
##             [,1]
## [1,]  1.00000000
## [2,] -0.50592885
## [3,]  0.01185771
## [1] 11.58779
## [1] "iteration =  8"
##              [,1]
## [1,]  1.000000000
## [2,] -0.497087379
## [3,] -0.005825243
## [1] 12.21344
## [1] "iteration =  9"
##              [,1]
## [1,]  1.000000000
## [2,] -0.501469148
## [3,]  0.002938296
## [1] 11.89515
## [1] "iteration =  10"
##              [,1]
## [1,]  1.000000000
## [2,] -0.499268649
## [3,] -0.001462701
## [1] 12.05289
```

```r
eigen(A)
```

```
## $values
```

```
## [1]  1.200000e+01 -6.000000e+00  4.930713e-16
##
## $vectors
##               [,1]          [,2]       [,3]
## [1,] -8.944272e-01  7.071068e-01 -0.2672612
## [2,]  4.472136e-01  1.040834e-16  0.5345225
## [3,] -5.945103e-17 -7.071068e-01  0.8017837
```

Deflation

```r
Q <- eigen(A)

v <- Q$vectors[,1]
lambda <- Q$values[1]

A_deflated <- A - lambda * v %*% t(v)


eigen(A_deflated)
```

```
## $values
## [1] -6+0.000000000000i  0+0.000000183473i  0-0.000000183473i
##
## $vectors
##               [,1]                          [,2]                          [,3]
## [1,]  0.8970852+0i  8.944272e-01+0.000000e+00i  8.944272e-01+0.000000e+00i
## [2,] -0.2760262+0i -4.472136e-01+0.000000e+00i -4.472136e-01-0.000000e+00i
## [3,] -0.3450328+0i  0.000000e+00+2.564114e-08i  0.000000e+00-2.564114e-08i
```