PCA Analysis

Xiuyang Guan 4/2/2017

All locations

```
library(ggplot2)
library(reshape2)
require(gdata)
## Loading required package: gdata
## gdata: read.xls support for 'XLS' (Excel 97-2004) files ENABLED.
##
## gdata: read.xls support for 'XLSX' (Excel 2007+) files ENABLED.
##
## Attaching package: 'gdata'
## The following object is masked from 'package:stats':
##
##
       nobs
## The following object is masked from 'package:utils':
##
##
       object.size
## The following object is masked from 'package:base':
##
##
       startsWith
library(ggmap)
library(leaflet)
library(lattice)
setwd("~/Google Drive/Xunyang_Fall16Spring17/Hilsa fish research/3.Analysis/Rscript/shinyapp")
load("data/Overview_DATA.RData")
getSvdMostInfluential <- function(mtrx,</pre>
                                   quantile,
                                   similarity_threshold,
                                                   = TRUE,
= 0.05,
                                   plot_selection
                                   plot_threshold
                                                     = NULL){
                                   varnames
  if(any(is.na(mtrx)))
    stop("Missing values are not allowed in this function. Impute prior to calling this function (try w
  if (quantile < 0 || quantile > 1)
    stop("The quantile mus be between 0-1")
  if (similarity threshold < 0 || similarity threshold > 1)
    stop("The similarity_threshold mus be between 0-1")
```

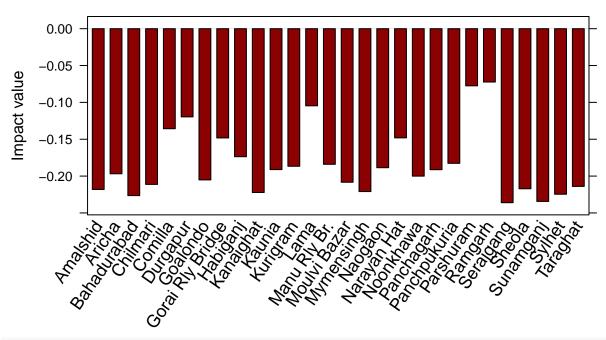
```
svd_out <- svd(scale(mtrx))</pre>
perc_explained <- svd_out$d^2/sum(svd_out$d^2)</pre>
# Select the columns that we want to look at
cols_expl <- which(cumsum(perc_explained) <= quantile)</pre>
# Select the variables of interest
getMostInfluentialVars <- function(){</pre>
 vars <- list()</pre>
 require("Hmisc")
 for (i in 1:length(perc_explained)){
    v_abs <- abs(svd_out$v[,i])</pre>
    maxContributor <- which.max(v_abs)</pre>
    similarSizedContributors <- which(v_abs >= v_abs[maxContributor]*similarity_threshold)
    if (any(similarSizedContributors %nin% maxContributor)){
      maxContributor <- similarSizedContributors[order(v_abs[similarSizedContributors], decreasing=TR</pre>
    vars[[length(vars) + 1]] <- maxContributor</pre>
 return(vars)
vars <- getMostInfluentialVars()</pre>
plotSvdSelection <- function(){</pre>
 require(lattice)
  if (plot_threshold < 0 || plot_threshold > 1)
    stop("The plot_threshold mus be between 0-1")
  if (plot_threshold > similarity_threshold)
    stop(paste0("You can't plot less that you've chosen - it makes no sense",
            " - the plot (", plot_threshold, ")",
             " similarity (", similarity_threshold, ")"))
  # Show all the bars that are at least at the threshold
  # level and group the rest into a "other category"
 bar_count <- length(perc_explained[perc_explained >= plot_threshold])+1
  if (bar count > length(perc explained)){
    bar_count <- length(perc_explained)</pre>
    plot_percent <- perc_explained</pre>
 }else{
    plot_percent <- rep(NA, times=bar_count)</pre>
    plot_percent <- perc_explained[perc_explained >= plot_threshold]
    plot_percent[bar_count] <- sum(perc_explained[perc_explained < plot_threshold])</pre>
  # Create transition colors
  selected_colors <- colorRampPalette(c("darkgreen", "#FFFFFF"))(bar_count+2)[cols_expl]</pre>
 nonselected_colors <- colorRampPalette(c("darkgrey", "#FFFFFF"))(bar_count+2)[(max(cols_expl)+1):ba
 max_no_print <- 10</pre>
 names <- unlist(lapply(vars[1:bar_count], FUN=function(x){</pre>
            if (is.null(varnames)){
```

```
varnames <- colnames(mtrx)</pre>
              }
              if (length(x) > max_no_print)
                 ret <- paste(c(varnames[1:(max_no_print-1)],</pre>
                         sprintf("+ %d other", length(x) + 1 - max_no_print)), collapse="\n")
               else
                 ret <- paste(varnames[x], collapse="\n")</pre>
              return(ret)
    rotation <- 45 + (max(unlist(lapply(vars[1:bar_count],</pre>
                       function(x) {
                         min(length(x), max_no_print)
                       })))-1)*(45/max_no_print)
    if (bar_count < length(perc_explained)){</pre>
      names[bar_count] <- "Other"</pre>
      nonselected_colors[length(nonselected_colors)] <- grey(.5)</pre>
    }
    las <- 2
    m \leftarrow par(mar=c(8.1, 4.1, 4.1, 2.1))
    on.exit(par(mar=m))
    p1 <- barchart(plot_percent * 100 ~ 1:bar_count,
        horiz=FALSE,
        ylab="Percentage explained (%D)",
        main="SVD - the maximum contributors defined by V column",
        xlab="Pattern contributing variables",
        col=c(selected_colors, nonselected_colors),
        key=list(text=list(c("Selected", "Not selected")),
            rectangles=list(col=c("darkgreen", "#777777"))),
        scales=list(x=list(rot=rotation, labels=names)))
    print(p1)
  if (plot_selection)
    plotSvdSelection()
  ret <- list(most_influential = unique(unlist(vars[cols_expl])),</pre>
      svd = svd_out)
 return(ret)
}
Names_NoHuman <- c("Bahadurabad", "Chilmari", "Kanaighat", "Sarighat", "Sherpur")
PCA_Input <- Overview_DATA[Overview_DATA$Types == "Monsoon",]</pre>
PCA_Input <- dcast(PCA_Input, Year ~ variable)</pre>
PCA_Input <- PCA_Input[,colSums(is.na(PCA_Input))<(nrow(PCA_Input)-30)]
## 1984 - 2014
PCA_Output <- Overview_DATA[Overview_DATA$variable == "Bangladesh_Inland",]$value
PCA_Input <- na.omit(PCA_Input[15:45,][-1])</pre>
```

```
SVD <- svd(scale(PCA_Input))</pre>
b_clr <- c("steelblue", "darkred")</pre>
key <- simpleKey(rectangles = TRUE, space = "top", points=FALSE,</pre>
 text=c("Positive", "Negative"))
key$rectangles$col <- b_clr</pre>
perc_explained <- SVD$d^2/sum(SVD$d^2)</pre>
b1 <- barchart(as.table(SVD$v[,1]),</pre>
  main="First column",
  horizontal=FALSE, col=ifelse(SVD$v[,1] > 0,
      b_clr[1], b_clr[2]),
  ylab="Impact value",
  scales=list(x=list(rot=55, labels=colnames(PCA_Input), cex=1.1)),
  key = key)
b2 <- barchart(as.table(SVD$v[,2]),</pre>
  main="Second column",
  horizontal=FALSE, col=ifelse(SVD$v[,2] > 0,
      b_clr[1], b_clr[2]),
  ylab="Impact value",
  scales=list(x=list(rot=55, labels=colnames(PCA_Input), cex=1.1)),
  key = key)
b3 <- barchart(as.table(SVD$v[,3]),
  main="Third column",
  horizontal=FALSE, col=ifelse(SVD$v[,3] > 0,
      b_clr[1], b_clr[2]),
  ylab="Impact value",
  scales=list(x=list(rot=55, labels=colnames(PCA_Input), cex=1.1)),
  key = key)
b1
```



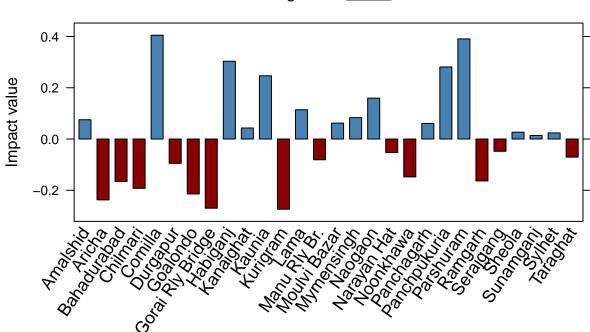
Positive Negative



b2

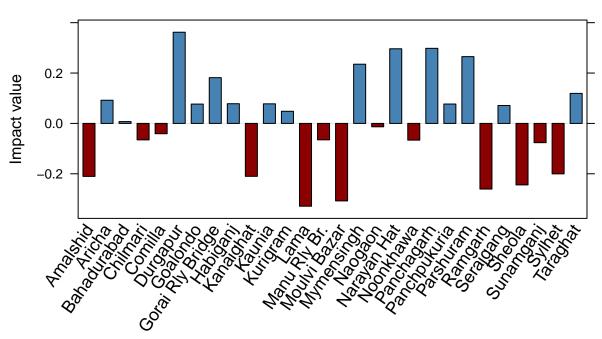
Second column

Positive Negative





Positive Negative



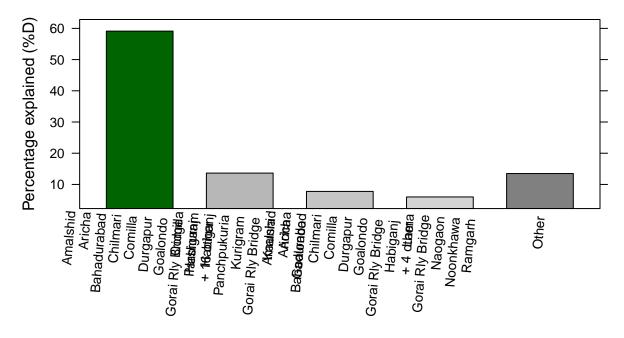
getSvdMostInfluential(PCA_Input, quantile = 0.7, similarity_threshold = 0.5)

```
## Loading required package: Hmisc
## Loading required package: survival
## Loading required package: Formula
##
## Attaching package: 'Hmisc'
## The following object is masked from 'package:gdata':
##
## combine
## The following objects are masked from 'package:base':
##
## format.pval, round.POSIXt, trunc.POSIXt, units
```

SVD – the maximum contributors defined by V column

Selected Not selected





Pattern contributing variables

```
## $most_influential
    [1] 24 26
               3 27 10 16 1 25 28 4 15 7 19 2 20 11 17 12 14 21
##
   [24]
        5
##
## $svd
## $svd$d
    [1] 1.575807e+01 7.567138e+00 5.709377e+00 5.015247e+00 4.477649e+00
##
    [6] 3.479306e+00 2.952253e+00 2.191458e+00 1.857450e+00 1.694009e+00
##
   [11] 1.265082e+00 1.173697e+00 8.608345e-01 7.788636e-01 5.944618e-01
   [16] 2.891041e-14
##
##
##
   $svd$u
##
                [,1]
                             [,2]
                                          [,3]
                                                                     [,5]
##
    [1,] -0.14874626
                      0.599249434 -0.409410482
                                                0.0006075540
                                                              0.19271239
##
    [2,]
         0.13345235
                      0.202518859
                                   0.159617245
                                                0.3055013889 -0.25049822
                      0.309399708 -0.006128197
                                                0.0046796219 0.17398449
##
    [3,] -0.30244315
    [4,] -0.19454823 -0.009269804
                                   0.324240572 -0.0003575427 -0.25219035
                                   0.291980622 -0.5376773986 -0.34842434
##
    [5,] 0.11072799
                      0.092539995
##
    [6,] -0.11489884
                      0.022218112
                                   0.072299554 0.5775849385 -0.33170920
##
    [7,] -0.06801220
                      0.128581356
                                   0.097815019 -0.1810326701 -0.07072962
##
    [8,] -0.17178781
                      0.115066141
                                   0.092502444 -0.1991462470
                                                              0.09465565
    [9,] -0.26216394 -0.304693679
                                   0.476984243
                                                0.1067260171
  [10,] 0.70249890 0.131438864
                                   0.214484899
                                                0.2034943513
                                                              0.15549208
  [11,] 0.14919114 -0.095821256 -0.135011372 -0.1496088516
  [12,] -0.22024772 -0.258608960 -0.347305842 0.1799744063 -0.11432518
   [13,] -0.01489023 -0.190096635 -0.159559032 -0.3093469292 -0.07379207
## [14,] -0.07118684 -0.325817361 -0.197645482 0.0697502419 -0.07761834
```

```
## [15,] 0.13414677 -0.366389731 -0.195637836 -0.0561585898 -0.11578612
  [16,] 0.33890807 -0.050315044 -0.279226356 -0.0149902908 0.31095940
##
             [,6]
                       [,7]
                                 [,8]
                                          [,9]
##
   ##
   [3,] -0.196650522 -0.690848211 -0.03188923 0.08756972 0.30550333
##
   [4.] -0.209342742 0.173228189 0.46841071 0.21912064 0.24339546
   [5,] -0.168319337 -0.071097237 -0.23633010 0.18729725 -0.20894401
##
   [6,] 0.301632704 -0.067821604 -0.25075953 0.09191080 -0.02359378
   [7,] 0.338297401 0.090273789 -0.25414947 0.02605841 0.05608086
   [8,] 0.576651866 -0.043046910 0.20743713 -0.29726821 0.10499675
   ## [12,] -0.385501382 -0.109183674 -0.21506437 -0.20537782 -0.04128902
## [13,] -0.041884947 -0.058739102 -0.20560369 -0.34363146 -0.30176011
  [14,] 0.262877503 0.120840295 0.02639515 -0.08704176 0.12304720
  [15,] 0.005951547 0.110631931 0.01011696 0.53754774 0.20234009
  [16,] 0.185701943 -0.325642232 0.32943747 0.32099459 -0.37399087
             [,11]
                       [,12]
                                [,13]
                                          [,14]
                                                   [,15] [,16]
##
   [1,] 0.1261318902 -0.05088871 0.06374393 -0.04255816 0.06176862 0.25
   [2,] 0.0011432565 0.29517958 -0.06306769 -0.24572868 -0.33034629
   [3,] 0.1243935799 0.03550306 0.11262274 -0.09493615 -0.25214793
##
   [4,] 0.0004600858 0.04935514 0.26584325 0.44261997 0.24761988
##
##
   [5,] 0.2374298563 -0.23511024 -0.33369624 -0.12405683 0.12065421
                                                        0.25
   [6,] -0.1251749145 -0.50967057 0.09010254 -0.09578406 0.12144738
   [7,] -0.5011559259  0.22660452 -0.19591081  0.39851986 -0.42223377
                                                        0.25
   [8,] 0.0230473797 0.22924083 -0.06075398 -0.31216059 0.45447993
  [9,] -0.0030145666 -0.07029146 -0.04797231 -0.09462319 -0.05007507
## [10,] 0.1291585778 0.19434514 0.10488492 0.10110983 0.16106775
## [11,] -0.2727351911 -0.51095271 -0.18670844 -0.14818618 -0.18209498
                                                        0.25
  [12,] -0.1513848400 0.28313227 -0.37248763 0.13803620 0.38010233
                                                        0.25
  [13,] -0.0439562884 -0.05340700 0.71244682 0.07081121 -0.02330812
## [14,] 0.7030468095 0.01451397 -0.14255319 0.21360465 -0.32979944
                                                        0.25
  [15,] -0.1688868858  0.26960696  0.15088267 -0.50325083 -0.10447429
  [16,] -0.0785028233 -0.16716077 -0.09737657 0.29658296 0.14733979
                                                        0.25
##
## $svd$v
                      [,2]
                                [,3]
                                          [,4]
                                                   [,5]
##
            [,1]
   ##
   [2,] -0.19685551 -0.23746706 0.092009159 -0.21355768 -0.16437804
   [3,] -0.22649170 -0.16588829 0.007297176 0.10340296 -0.09732951
##
   [4,] -0.21127439 -0.19249156 -0.065331492 0.19033791 0.03944019
##
   [5,] -0.13573013  0.40510017 -0.040892722 -0.07829892  0.13741839
   [6,] -0.11964887 -0.09491172 0.362061653 0.06186301 0.45055462
   [7,] -0.20515720 -0.21438758   0.076777386 -0.20942346 -0.13795119
   [8,] -0.14824455 -0.27002820 0.181680146 -0.37881065 -0.09942459
  [9,] -0.17369157  0.30352447  0.078222186 -0.10529074 -0.01808419
## [12,] -0.18668430 -0.27391230 0.048239279 0.18064972 0.18392975
## [14,] -0.18400010 -0.08097736 -0.065194876 -0.06626877 0.16041970
```

```
## [18,] -0.14813629 -0.05281692 0.296354422 -0.16991875 0.16088846
## [19,] -0.20004671 -0.14772533 -0.066328476 0.26872729 0.07173000
## [23,] -0.07237807 -0.16349190 -0.260640293 -0.23263213 0.61514724
## [24,] -0.23607935 -0.04792280 0.071160118 0.01674713 -0.14954619
[26,] -0.23438184  0.01387995 -0.076444302  0.15489312  0.08494009
  [28,] -0.21400407 -0.07060134 0.119201698 -0.20738873 -0.19972757
##
            [,6]
                      [,7]
                                [,8]
                                          [,9]
##
   [1,] 0.001036507 0.031848917 -0.3990265950 0.088966790 0.025720875
##
   [3,] 0.110320976 -0.012915461 0.0323650592 0.016271269 -0.056818571
##
   [4,] -0.095658341 -0.035017889 0.3044272060 -0.062999655 0.089417508
  [5,] -0.071200632  0.042145870  0.0862383220  0.154063873  0.148438672
   [6,] -0.248649052 0.265277018 -0.2966153419 -0.306709033 -0.317878215
##
  [7,] -0.055063319  0.020660109 -0.0338510703  0.117719482  0.100141891
  [8,] -0.013832819  0.084487782 -0.1805643463  0.150781932 -0.022142615
  [9,] -0.214613782 -0.181014298 -0.1963796501 -0.299220160 0.156115037
## [10,] 0.046546067 0.038519038 -0.2446732898 0.029577527 0.021456442
 [11,] 0.107623089 -0.080994301 0.1476914185 0.227241209 -0.467945524
 [12,] 0.095522217 0.057813254 0.1220643892 -0.009118315 -0.206752215
## [13,] 0.305215626 0.288517712 0.1005036368 -0.481424737 0.003434547
## [14,] -0.407473894 -0.603808742 0.1137659752 -0.176380856 0.102627419
## [15,] -0.168268557 -0.120355513 0.0396426248 0.024817544 -0.048608970
## [18,] 0.506443676 -0.470016855 -0.1063527465 -0.126457565 -0.030891966
## [19,] 0.253436364 -0.045368985 0.2958093975 0.050349608 0.198072471
## [20,] 0.089520179 0.221647244 0.0756693740 -0.267266636 0.511711239
## [21,] 0.304722338 -0.203095087 -0.0251201446 0.015245034 -0.123676464
## [22,] 0.023050514 0.118202955 -0.0212916093 0.370930129 0.213865497
## [23,] 0.148371365 0.091256795 -0.0819993740 0.258220647 0.206425468
[26,] 0.065701223 0.111590663 0.0816656348 -0.039048265 -0.213021142
  [27,] 0.063753246 0.060496633 -0.0898151738 -0.004270587 -0.020216813
  [28,] -0.185961472 0.033872844 0.0751537964 0.292588431 0.141041391
           [,11]
                     [,12]
                              [,13]
                                       [,14]
                                                [.15]
   [1,] -0.258062521 -0.028931937 -0.08006254 0.22694270 -0.43828906
##
  [2,] 0.029671908 -0.086616998 0.02923517 0.04399644 0.06807306
  [3,] 0.191227136 0.009400349 -0.05148530 -0.02225701 0.03542229
##
   [4,] -0.043529060 0.075088255 -0.35272217 0.05910109 0.11105000
  [5,] 0.118415497 -0.203946138 -0.10524290 -0.34161161 -0.26393329
  [6,] 0.200929083 -0.047918030 0.13753674 0.06258507 -0.14011506
##
  [7,] -0.143879646 -0.219661569 -0.10202105 0.01841510 0.16595338
  [8,] -0.007326514   0.184531139 -0.05489648 -0.43964139 -0.21746757
##
  [9,] 0.186633241 0.403657328 -0.13643664 0.23436433 0.19814703
## [10,] -0.050722057 -0.122546808 -0.02610021 0.09332303 0.15562285
```

```
## [12,] -0.026862741 -0.010350888 -0.32887971 -0.20296140 0.31099385
## [13,] 0.206453416 0.057773262 -0.07004262 -0.18095590 0.01917866
## [15,] 0.057131493 -0.180256366 0.38627337 -0.22372313
                                                    0.22371839
## [16,] 0.016192754 0.014786330 -0.16024965
                                         0.02972384
                                                    0.05249047
## [17,] -0.251450073 -0.255321852 0.15001040 0.34616369
                                                    0.06051806
## [18,] -0.039422989 -0.235101880 0.27423464 0.09407035
                                                    0.04559309
## [19,] 0.452542665 0.100466579 0.23569758 0.10708750 -0.25677023
0.05962095
## [21,] -0.023648670 -0.117540839 -0.41420631 0.15545715 -0.02145110
## [22,] 0.171632603 -0.071532751 -0.09093432 -0.15263192 0.19170293
## [24,] 0.063850661 0.151773887 0.06152752 0.15896063 -0.08590387
## [25,] 0.103214010 -0.043853316 0.14162093 -0.14415647 0.35404714
## [26,] -0.104679865 -0.071844403 -0.09906412 -0.11921585 -0.33764153
## [27,] -0.043041425 -0.102553204 0.09669711
                                         0.03432023 -0.07842698
  [28,] 0.303355954 -0.011782184 0.09178535 0.23897597 -0.05027109
##
              [,16]
##
   [1,] -0.027857783
##
   [2,] 0.109830956
##
   [3,] 0.165131232
   [4,] -0.233422316
   [5,] -0.035292227
##
   [6,] 0.133713282
##
   [7,] -0.164320802
   [8,] -0.151472040
   [9,] 0.001698074
## [10,] -0.122043286
## [11,] 0.008125813
## [12,] 0.188785288
## [13,] -0.013535949
## [14,] 0.053597525
## [15,] 0.086497668
## [16,] -0.263991081
## [17,] -0.069100079
## [18,] -0.252591796
## [19,] 0.012818926
## [20,] 0.335812448
## [21,] 0.405095819
## [22,] -0.202913357
## [23,] 0.054029391
## [24,] -0.406588889
## [25,] 0.005653668
## [26,] 0.024159539
## [27,] -0.038461996
## [28,] 0.385211130
```

No-human impacted