Problem Set 4

Bhargavi Ganesh

11/10/2019

Factor analysis

Question 1: How do CFA and EFA differ?

Factor Analysis is a set of techniques to figure out whether a latent factor (or factors) could account for covariation among a set of input features. In exploratory analysis, we go in without expectations on what the factor structure of the data would be, and find factor loadings for different numbers of factors. In confirmatory analysis, we have a pre-established theory on what number of factors there are and feature loadings should be. We see whether our results confirm these expectations.

Question 2: Fit three exploratory factor analysis models initialized at 2, 3, and 4 factors. Present the loadings from these solutions and discuss in substantive terms. How does each fit? What sense does this give you of the underlying dimensionality of the space? And so on.

```
countries <- read.csv("countries.csv", header = TRUE, row.names = 1)
```

```
factan.1 <- fa(countries[-1],
               nfactors = 2)

factan.2 <- fa(countries[-1],
               nfactors = 3)

factan.3 <- fa(countries[-1],
               nfactors = 4)
```

```
kable(factan.1$uniquenesses)
```

	x
polity	0.0287106
polity2	0.0287106
democ	0.0570816
autoc	0.1604508
unreg	0.8548967
physint	0.3651464
speech	0.4945365
new_empinx	0.1794688
wecon	0.6985714
wopol	0.7266392
wosoc	0.5719576
elecsd	0.2239443
gdp.pc.wdi	0.5317426
gdp.pc.un	0.5396301
pop.wdi	0.7928305
amnesty	0.3555881
statedept	0.2532221
milper	0.7932059
cinc	0.8672774
domestic9	0.7983253

```r
summary(factan.1)
```

```
##
## Factor analysis with Call: fa(r = countries[-1], nfactors = 2)
##
## Test of the hypothesis that 2 factors are sufficient.
## The degrees of freedom for the model is 151  and the objective function was  51.14
## The number of observations was  107  with Chi Square =  4968.68  with prob <  0
##
## The root mean square of the residuals (RMSA) is  0.13
## The df corrected root mean square of the residuals is  0.14
##
## Tucker Lewis Index of factoring reliability =  0.057
## RMSEA index =  0.574  and the 10 % confidence intervals are  0.536 NA
## BIC =  4263.08
##  With factor correlations of
##       MR1   MR2
## MR1  1.00 -0.39
## MR2 -0.39  1.00
```

Looking at the factor analysis above, we see a few interesting items that suggest whether or not 2 factors is the appropriate amount of factors to fit the data. First, we can look at the uniqueness (represented by u2). Uniqueness is measured on a scale from 0 to 1, with a high uniqueness generally suggesting that the data doesn't fit well to our factors. The uniqueness appears to be fairly high on all of the variables except for polity and democ. This is our first clue that 2 factors are probably not sufficient for the data. The results above also show a chi-squared test that two factors are sufficient. Given that the p-value is low, we can reject the null hypothesis that two factors are sufficient for the data. Below, we look specifically at the loadings on the variables.

```r
factan.1$loadings
```

```
##
## Loadings:
##            MR1    MR2
## polity      0.993
## polity2     0.993
## democ       0.930
## autoc      -0.965 -0.152
## unreg       0.412  0.124
## physint           -0.772
## speech      0.634 -0.157
## new_empinx  0.806 -0.204
## wecon             -0.507
## wopol       0.546
## wosoc       0.292 -0.482
## elecsd      0.855
## gdp.pc.wdi        -0.653
## gdp.pc.un         -0.652
## pop.wdi     0.203  0.494
## amnesty            0.812
## statedept          0.833
## milper      0.156  0.493
## cinc        0.210  0.391
## domestic9   0.278  0.478
##
##                  MR1   MR2
## SS loadings    6.322 4.282
## Proportion Var 0.316 0.214
## Cumulative Var 0.316 0.530
```

Looking at the loadings and resulting output above, we can see, for each variable, how much of its variance is explained by each factor. On the whole, 30% of the variance is explained by factor 1, and 22% of the variance is explained by factor 2. This means that only 52% of the variance is explained overall. These results further suggest that two factors are not sufficient for the data. As an aside, we can note the crossloadings on the variables indicating that some of the variables load onto more than one factor.

Below, we inspect the results when fitting the data to 3 factors.

```r
summary(factan.2)
```

```
##
## Factor analysis with Call: fa(r = countries[-1], nfactors = 3)
##
## Test of the hypothesis that 3 factors are sufficient.
## The degrees of freedom for the model is 133  and the objective function was  46.39
## The number of observations was  107  with Chi Square =  4476.29  with prob <  0
##
## The root mean square of the residuals (RMSA) is  0.06
## The df corrected root mean square of the residuals is  0.07
##
## Tucker Lewis Index of factoring reliability =  0.028
## RMSEA index =  0.583  and the 10 % confidence intervals are  0.541 NA
```

```
## BIC =  3854.8
##  With factor correlations of
##       MR1   MR3   MR2
## MR1  1.00  0.37 -0.06
## MR3  0.37  1.00 -0.12
## MR2 -0.06 -0.12  1.00
```

```
kable(factan.2$uniquenesses)
```

	x
polity	0.0271659
polity2	0.0271659
democ	0.0602876
autoc	0.1187822
unreg	0.8522293
physint	0.3747252
speech	0.4912632
new_empinx	0.1568091
wecon	0.6802854
wopol	0.7230259
wosoc	0.5365278
elecsd	0.2196630
gdp.pc.wdi	0.2824210
gdp.pc.un	0.2912335
pop.wdi	0.1848638
amnesty	0.3895778
statedept	0.2623136
milper	0.0830259
cinc	0.0170299
domestic9	0.8056749

We see that the extra factor seems to have brought down the uniqueness slightly for each of the variables. This is our first hint that three factors may be a better fit for the data. The chi square value has also come down slightly. Still, we see that the p-value is below 0.05, so we would reject the null hypothesis that 3 factors are sufficient for the data. Below, we again inspect the loadings on the variables.

```
factan.2$loadings
```

```
##
## Loadings:
##            MR1    MR3    MR2
## polity      0.990
## polity2     0.990
## democ       0.909  0.144
## autoc      -0.991  0.188
## unreg       0.412 -0.122
## physint            0.731 -0.137
## speech      0.648  0.132
## new_empinx  0.842  0.135 -0.123
## wecon              0.523
## wopol       0.548
## wosoc       0.269  0.537
## elecsd      0.861
```

```
## gdp.pc.wdi         0.858  0.159
## gdp.pc.un          0.856  0.159
## pop.wdi                   0.893
## amnesty           -0.705  0.245
## statedept         -0.792  0.146
## milper                    0.948
## cinc                      0.998
## domestic9   0.262 -0.443
##
##                  MR1   MR3   MR2
## SS loadings    6.276 4.017 2.882
## Proportion Var 0.314 0.201 0.144
## Cumulative Var 0.314 0.515 0.659
```

Inspecting the loadings with 3 factors, we see that the overall variance explained by the factors goes up from 52% to 65%. 30.8% of the variance is explained by factor 1, 20.4% of the variance is explained by factor 2, and 13.7% of the variance is explained by factor 3. This suggests that 3 factors might be a better fit for the data than 2 factors. Again, we note that there are crossloadings, and only 1 of the variables (new_empinx) loads onto all three of the variables.

Finally, we inspect the results from 4 factors below.

```r
summary(factan.3)
```

```
##
## Factor analysis with Call: fa(r = countries[-1], nfactors = 4)
##
## Test of the hypothesis that 4 factors are sufficient.
## The degrees of freedom for the model is 116  and the objective function was  43.23
## The number of observations was  107  with Chi Square =  4142.87  with prob <  0
##
## The root mean square of the residuals (RMSA) is  0.04
## The df corrected root mean square of the residuals is  0.06
##
## Tucker Lewis Index of factoring reliability =  -0.04
## RMSEA index =  0.603  and the 10 % confidence intervals are  0.557 NA
## BIC =  3600.82
##  With factor correlations of
##       MR1   MR3   MR4   MR2
## MR1  1.00 -0.07  0.31 -0.27
## MR3 -0.07  1.00 -0.01  0.24
## MR4  0.31 -0.01  1.00 -0.50
## MR2 -0.27  0.24 -0.50  1.00
```

```r
kable(factan.3$uniquenesses)
```

	x
polity	0.0283829
polity2	0.0283829
democ	0.0588486
autoc	0.1230394
unreg	0.8455780
physint	0.2586295
speech	0.4899532
new_empinx	0.1514040
wecon	0.6807477
wopol	0.7195888
wosoc	0.5478672
elecsd	0.2200265
gdp.pc.wdi	0.0250723
gdp.pc.un	0.0396671
pop.wdi	0.1461241
amnesty	0.3675082
statedept	0.1608203
milper	0.0616399
cinc	0.0284156
domestic9	0.5577865

Looking at the uniqueness of each variable, we see that adding the extra factor had mixed results. The uniqueness actually went up slightly for the idealpoint, polity, polity2, autoc, wosoc, and amnesty variables. This suggests that 4 factors may not be the best fit for these variables. Generally, the chi-squared value came down slightly, suggesting that the overall fit is slightly better than that of 3 factors. The p-value is still less than 0.05, so we reject the null hypothesis that 4 factors is sufficient.

Below, we look at the loadings for 4 factors.

```
factan.3$loadings
```

```
##
## Loadings:
##              MR1    MR3    MR4    MR2
## polity        0.992
## polity2       0.992
## democ         0.919         0.131
## autoc        -0.983         0.144
## unreg         0.403                0.156
## physint       0.132         0.104 -0.751
## speech        0.660                -0.115
## new_empinx    0.858                -0.150
## wecon         0.110         0.399 -0.167
## wopol         0.551
## wosoc         0.301         0.357 -0.224
## elecsd        0.868
## gdp.pc.wdi                  0.985
## gdp.pc.un                   0.979
## pop.wdi              0.927
## amnesty              0.178 -0.209  0.587
## statedept    -0.149        -0.158  0.761
## milper               0.963
## cinc                 0.981  0.113
```

```
## domestic9   0.233         0.198  0.767
##
##                   MR1   MR3   MR4   MR2
## SS loadings    6.373 2.812 2.400 2.223
## Proportion Var 0.319 0.141 0.120 0.111
## Cumulative Var 0.319 0.459 0.579 0.690
```

We can see that adding a fourth factor slightly increased the proportion of the variance explained by the factors, from 65% to 68%. This jump was not as drastic as the jump that happened when we increased from 2 to 3 factors, but illustrates a potential better fit nonetheless. Again, we can note that there are crossloadings, with 4 variables (idealpoint, wecon, wosoc, and domestic9) loading onto 3 of the 4 factors.

Question 3: Rotate the 3-factor solution using any oblique method you would like and present a visual of the unrotated and rotated versions side-by-side. How do these differ and why does this matter (or not)?

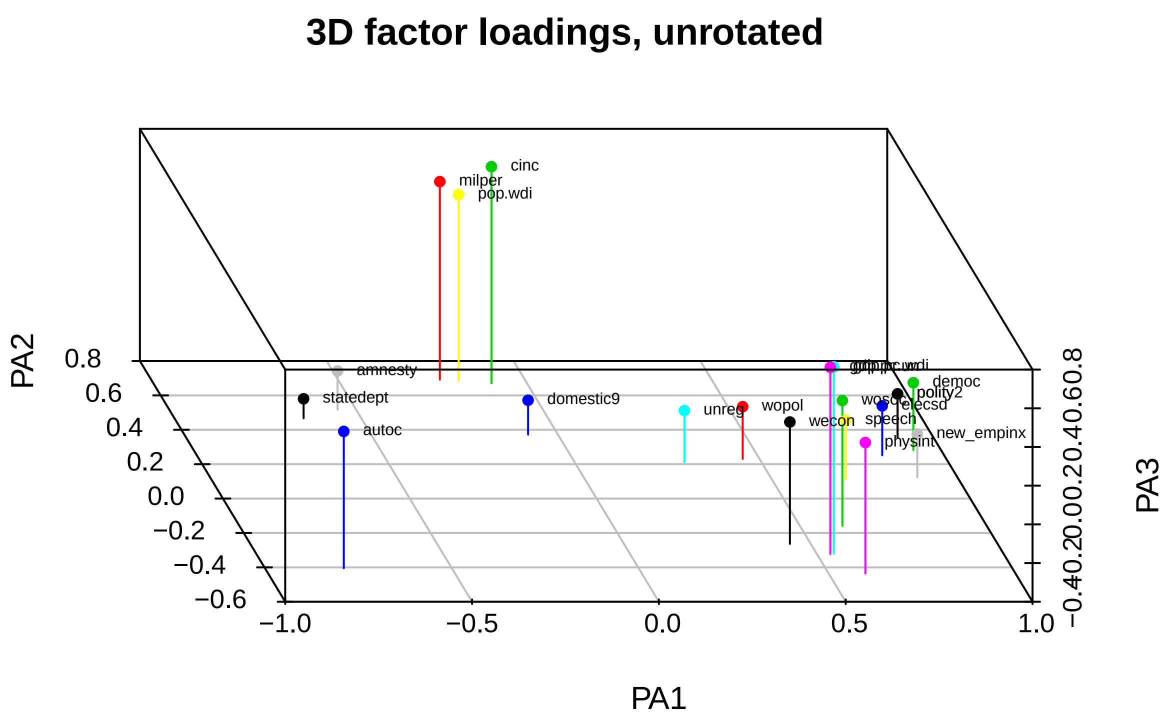
We can see the 3D factor loadings for the unrotated factor model below:

```r
nonrotated.factors <- fa(cor(countries[,-1]),
        fm = "pa", # communalities along the diagonal (total variation across features)
        nfactors = 3,
        rotate = "none",
        residuals = TRUE)

# Plot unrotated
s3d <- scatterplot3d(
        as.data.frame(unclass(nonrotated.factors$loadings)),
        main="3D factor loadings, unrotated",
        color=1:ncol(countries[-1]),
        pch=20,
        type="h",
        angle=100)

s3d.coords <- s3d$xyz.convert(
(as.data.frame(unclass(nonrotated.factors$loadings))$PA1),     (as.data.frame(unclass(nonrotated.factor
(as.data.frame(unclass(nonrotated.factors$loadings))$PA3))

text(s3d.coords$x, s3d.coords$y,
          labels=colnames(countries[-1]),
          pos=4, cex=.5)
```



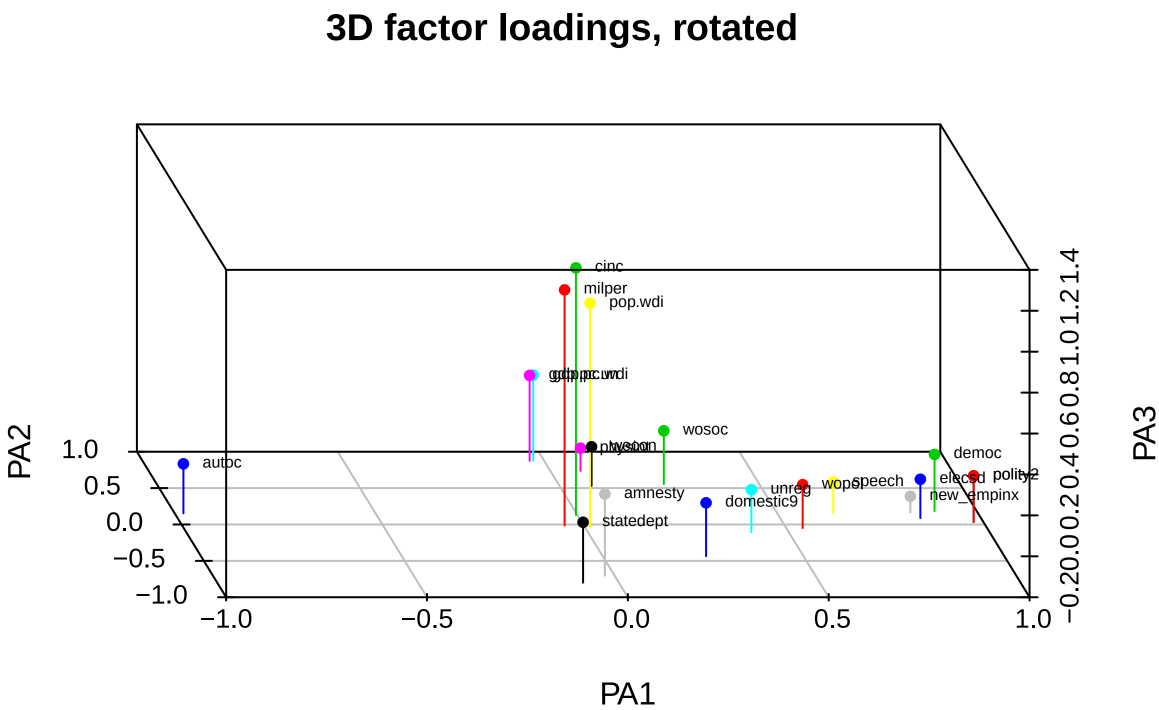
We can see the 3D factor loadings for the rotated factor model below. The rotation was done using the oblique method of "promax".

```
# Oblique rotated factor solution
promax.factors <- fa(cor(countries[,-1]),
                      fm = "pa",
                      nfactors = 3,
                      rotate = "promax")

# Plot rotated
s3d <- scatterplot3d(
        as.data.frame(unclass(promax.factors$loadings)),
        main="3D factor loadings, rotated",
        color=1:ncol(countries[-1]),
        pch=20,
        type="h",
        angle=100)

s3d.coords <- s3d$xyz.convert(
              (as.data.frame(unclass(promax.factors$loadings))$PA1),
              (as.data.frame(unclass(promax.factors$loadings))$PA2),
              (as.data.frame(unclass(promax.factors$loadings))$PA3))

text(s3d.coords$x, s3d.coords$y,
          labels=colnames(countries[-1]),
          pos=4, cex=.5)
```



When we use the oblique method of "promax", we see some noticeable differences in the plotting of the factor loadings. The variable "auto" appears to be more on its own, and farther to the left while statedept, amnesty, etc move more to the righthand side. We can represent this difference numerically below. The autoc variable clearly loads more onto factor 1 after the rotation, while the amnesty and statedept variables load more onto factors 2 and 3. The differences matter because they highlight one of the downsides of factor analysis which is that rotation could give us more interpretable results but these may not always be in line with the original model.

```
nonrotated.factors$loadings
```

```
##
## Loadings:
##             PA1    PA2    PA3
## polity      0.904  0.354 -0.174
## polity2     0.904  0.354 -0.174
## democ       0.926  0.282
## autoc      -0.789 -0.405  0.307
## unreg       0.293  0.210 -0.132
## physint     0.598 -0.436  0.277
## speech      0.698  0.111
## new_empinx  0.893  0.124 -0.176
## wecon       0.444 -0.264  0.230
## wopol       0.455  0.231 -0.130
## wosoc       0.614 -0.159  0.248
## elecsd      0.834  0.252 -0.145
## gdp.pc.wdi  0.546 -0.320  0.565
```

```
## gdp.pc.un    0.536 -0.323  0.566
## pop.wdi     -0.179  0.685  0.562
## amnesty     -0.549  0.517 -0.200
## statedept   -0.654  0.467 -0.300
## milper      -0.227  0.692  0.623
## cinc                0.670  0.720
## domestic9           0.372 -0.222
##
##                  PA1   PA2   PA3
## SS loadings    7.734 3.197 2.485
## Proportion Var 0.387 0.160 0.124
## Cumulative Var 0.387 0.547 0.671
```

```
promax.factors$loadings
```

```
##
## Loadings:
##            PA1    PA2    PA3
## polity      0.976
## polity2     0.976
## democ       0.895  0.183
## autoc      -0.978  0.153
## unreg       0.407 -0.107
## physint            0.738
## speech      0.640  0.156
## new_empinx  0.833  0.165 -0.121
## wecon              0.531
## wopol       0.541
## wosoc       0.262  0.555
## elecsd      0.849
## gdp.pc.wdi         0.877  0.219
## gdp.pc.un          0.874  0.218
## pop.wdi                   0.896
## amnesty           -0.705  0.200
## statedept         -0.799
## milper                    0.952
## cinc               0.131  1.008
## domestic9   0.257 -0.435
##
##                  PA1   PA2   PA3
## SS loadings    6.103 4.145 2.909
## Proportion Var 0.305 0.207 0.145
## Cumulative Var 0.305 0.512 0.658
```

PCA

Question 1: What is the statistical difference between PCA and FA? Describe the basic construction of each approach using equations and then point to differences that exist across these two widely used methods for reducing dimensionality.

In PCA, the components/factors are a linear combination of all of the variables, whereas in FA, each variable is loaded onto one factor based on correlation similarities.

Another way of saying this is that in FA, each observed indicator is made up of the factor it loads onto and a residual, shown by the equation below:

X1 = b1F + d1U, where X1 is a given observable characteristic, and b1 represents the loading onto a factor F and d1 represents the loading onto a residual U.

By contrast, in PCA, we try to first find the direction that provides the maximal variance, and then fit the residuals to other components. This can be shown by the equation below:

F = L1X1 + L2X2 + L3X3. . . .LkXk.

L1, L2, etc represent weights on a given observable characteristic. The weights determine how much each X1, X2, etc. contribute to a factor/component in a given direction.

There are a couple of statistical differences. The latent variables in FA are usually assumed to have a gaussian distribution, whereas PCA does not assume any specific distribution, allowing for PCA to have more statistical independence.

Question 2: Fit a PCA model. Present the proportion of explained variance across the first 10 components. What do these values tell you substantively (e.g., how many components likely characterize these data?)?

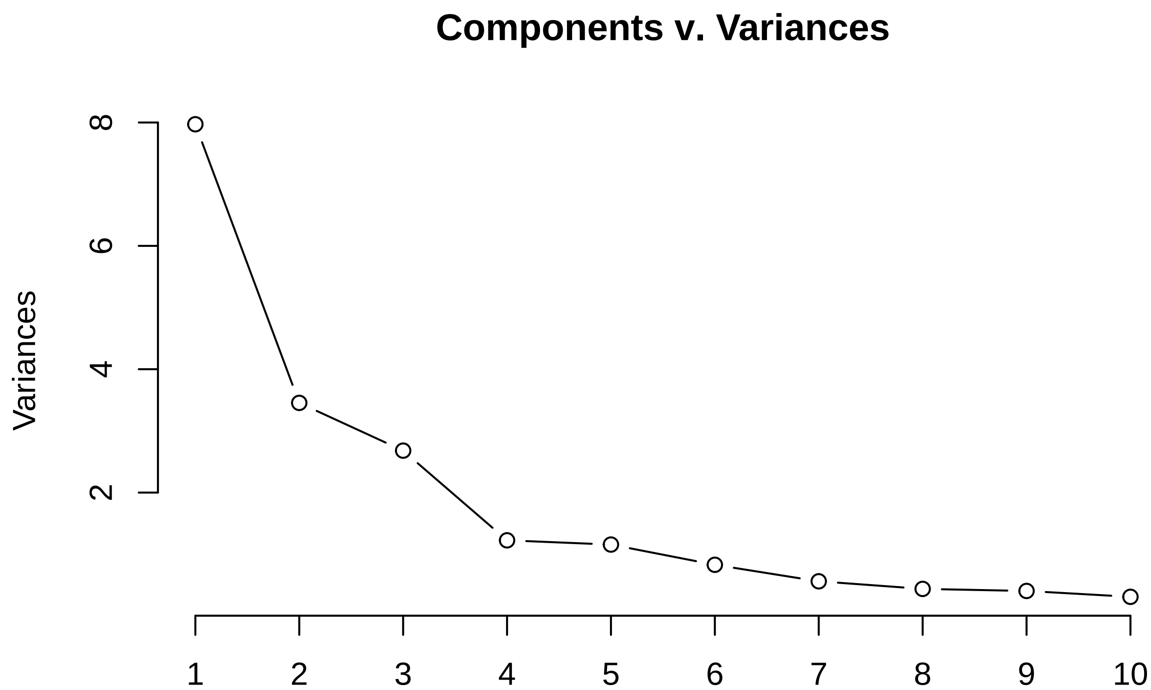
The PCA model is fit below:

```
pca1 <- prcomp(countries[-1],
               scale=TRUE,
               center = TRUE)
as.data.frame(unclass(summary(pca1)$importance))[1:10]
```

```
##                              PC1      PC2      PC3      PC4      PC5
## Standard deviation     2.823437 1.858524 1.637026 1.107098 1.075506
## Proportion of Variance 0.398590 0.172710 0.133990 0.061280 0.057840
## Cumulative Proportion  0.398590 0.571300 0.705290 0.766570 0.824410
##                              PC6       PC7       PC8       PC9      PC10
## Standard deviation     0.9103437 0.7491357 0.6617182 0.6360879 0.5564807
## Proportion of Variance 0.0414400 0.0280600 0.0218900 0.0202300 0.0154800
## Cumulative Proportion  0.8658400 0.8939000 0.9158000 0.9360300 0.9515100
```

The table above shows the proportion of variance explained by each component. By 5 components we have already accounted for approximately 80% of the variance. There is no systematic way to determine the number of components that should be chosen. A screeplot (shown below) is sometimes helpful for this purpose.

```
plot(pca1,
     type="l",
     main = "Components v. Variances")
```

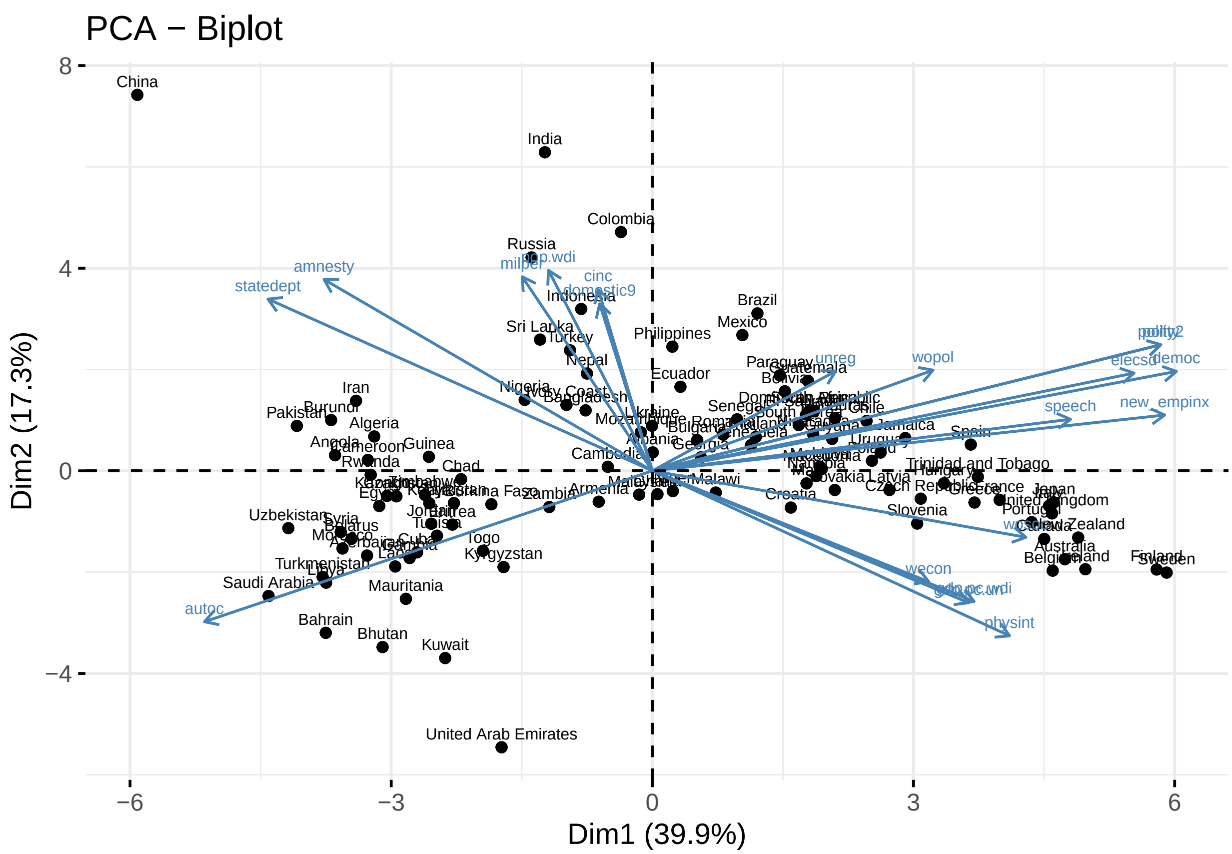


By visual inspection, we can see that by the 10th component, the variance approaches zero. It would therefore not make sense to have more than 10 components. Visually, it looks like the biggest dropoff occurs after component 6, so it may make sense to have 6 components.

Question 3: Present a biplot of the PCA fit from the previous question. Describe what you see (e.g., which countries are clustered together? Which input features are doing the bulk of the explaining? How do you know this?

The biplot can be seen below:

```
fviz_pca_biplot(pca1, labelsize=2, label="all")
```



The biplot above shows the scores and loadings on the principal components.

We can see that in the lefthand corner, countries in the Middle East are clustered together, whereas on the righthand side, European countries are clustered together. In the middle right, we have more Latin American and African countries clustered together and in the middle left, we see a clustering of Asian countries.

The input features doing the bulk of the explaining are the ones that have the maximal variance in a given direction. In this case, polity, polity2, democ, and autoc appear to explain the bulk of the variance. Polity and polity2 indicate regime type, as do the democ and autoc variables. This shows that most of the variance in these variables is due to regime type. The clustering of countries near these variables makes intuitive sense. For example, middle eastern countries appear to be clustered near the the autocracy component on the bottom left.

Bonus Question: Fit a sparse PCA model and a probabilistic PCA model. Compare these results substantively. What does each tell you and why do these distinctions matter in terms of inference (or not)?

A fitting of the sparse PCA is shown below. I could not figure out how to fit the probabilistic PCA model.

```
spca1 <- spca(countries[-1],
              scale=TRUE,
              center = TRUE)
```

```
## [1] "Iteration:    1, Objective: 5.77982e+00, Relative improvement Inf"
## [1] "Iteration:   11, Objective: 5.75359e+00, Relative improvement 3.69817e-04"
## [1] "Iteration:   21, Objective: 5.73503e+00, Relative improvement 2.95021e-04"
```

```
## [1] "Iteration:   31, Objective: 5.71929e+00, Relative improvement 2.60155e-04"
## [1] "Iteration:   41, Objective: 5.70516e+00, Relative improvement 2.38410e-04"
## [1] "Iteration:   51, Objective: 5.69209e+00, Relative improvement 2.23235e-04"
## [1] "Iteration:   61, Objective: 5.67970e+00, Relative improvement 2.13710e-04"
## [1] "Iteration:   71, Objective: 5.66794e+00, Relative improvement 2.03194e-04"
## [1] "Iteration:   81, Objective: 5.65667e+00, Relative improvement 1.97077e-04"
## [1] "Iteration:   91, Objective: 5.64567e+00, Relative improvement 1.92570e-04"
## [1] "Iteration:  101, Objective: 5.63489e+00, Relative improvement 1.89950e-04"
## [1] "Iteration:  111, Objective: 5.62433e+00, Relative improvement 1.86147e-04"
## [1] "Iteration:  121, Objective: 5.61396e+00, Relative improvement 1.83839e-04"
## [1] "Iteration:  131, Objective: 5.60369e+00, Relative improvement 1.82685e-04"
## [1] "Iteration:  141, Objective: 5.59349e+00, Relative improvement 1.81903e-04"
## [1] "Iteration:  151, Objective: 5.58334e+00, Relative improvement 1.81409e-04"
## [1] "Iteration:  161, Objective: 5.57333e+00, Relative improvement 1.76588e-04"
## [1] "Iteration:  171, Objective: 5.56366e+00, Relative improvement 1.70162e-04"
## [1] "Iteration:  181, Objective: 5.55423e+00, Relative improvement 1.69236e-04"
## [1] "Iteration:  191, Objective: 5.54498e+00, Relative improvement 1.66182e-04"
## [1] "Iteration:  201, Objective: 5.53586e+00, Relative improvement 1.64344e-04"
## [1] "Iteration:  211, Objective: 5.52677e+00, Relative improvement 1.64384e-04"
## [1] "Iteration:  221, Objective: 5.51776e+00, Relative improvement 1.61005e-04"
## [1] "Iteration:  231, Objective: 5.50891e+00, Relative improvement 1.60297e-04"
## [1] "Iteration:  241, Objective: 5.50009e+00, Relative improvement 1.60372e-04"
## [1] "Iteration:  251, Objective: 5.49132e+00, Relative improvement 1.59580e-04"
## [1] "Iteration:  261, Objective: 5.48256e+00, Relative improvement 1.59866e-04"
## [1] "Iteration:  271, Objective: 5.47391e+00, Relative improvement 1.57749e-04"
## [1] "Iteration:  281, Objective: 5.46527e+00, Relative improvement 1.58046e-04"
## [1] "Iteration:  291, Objective: 5.45663e+00, Relative improvement 1.57936e-04"
## [1] "Iteration:  301, Objective: 5.44801e+00, Relative improvement 1.58301e-04"
## [1] "Iteration:  311, Objective: 5.43938e+00, Relative improvement 1.58724e-04"
## [1] "Iteration:  321, Objective: 5.43076e+00, Relative improvement 1.57868e-04"
## [1] "Iteration:  331, Objective: 5.42219e+00, Relative improvement 1.58124e-04"
## [1] "Iteration:  341, Objective: 5.41364e+00, Relative improvement 1.56224e-04"
## [1] "Iteration:  351, Objective: 5.40525e+00, Relative improvement 1.54358e-04"
## [1] "Iteration:  361, Objective: 5.39691e+00, Relative improvement 1.54662e-04"
## [1] "Iteration:  371, Objective: 5.38863e+00, Relative improvement 1.53530e-04"
## [1] "Iteration:  381, Objective: 5.38036e+00, Relative improvement 1.53675e-04"
## [1] "Iteration:  391, Objective: 5.37210e+00, Relative improvement 1.53815e-04"
## [1] "Iteration:  401, Objective: 5.36383e+00, Relative improvement 1.54374e-04"
## [1] "Iteration:  411, Objective: 5.35553e+00, Relative improvement 1.55007e-04"
## [1] "Iteration:  421, Objective: 5.34722e+00, Relative improvement 1.55562e-04"
## [1] "Iteration:  431, Objective: 5.33891e+00, Relative improvement 1.55274e-04"
## [1] "Iteration:  441, Objective: 5.33062e+00, Relative improvement 1.55206e-04"
## [1] "Iteration:  451, Objective: 5.32234e+00, Relative improvement 1.55253e-04"
## [1] "Iteration:  461, Objective: 5.31407e+00, Relative improvement 1.55710e-04"
## [1] "Iteration:  471, Objective: 5.30581e+00, Relative improvement 1.55194e-04"
## [1] "Iteration:  481, Objective: 5.29758e+00, Relative improvement 1.55447e-04"
## [1] "Iteration:  491, Objective: 5.28935e+00, Relative improvement 1.55211e-04"
## [1] "Iteration:  501, Objective: 5.28115e+00, Relative improvement 1.54710e-04"
## [1] "Iteration:  511, Objective: 5.27298e+00, Relative improvement 1.55144e-04"
## [1] "Iteration:  521, Objective: 5.26479e+00, Relative improvement 1.55725e-04"
## [1] "Iteration:  531, Objective: 5.25659e+00, Relative improvement 1.54278e-04"
## [1] "Iteration:  541, Objective: 5.24849e+00, Relative improvement 1.54451e-04"
## [1] "Iteration:  551, Objective: 5.24037e+00, Relative improvement 1.55133e-04"
## [1] "Iteration:  561, Objective: 5.23226e+00, Relative improvement 1.54725e-04"
```

```
## [1] "Iteration:  571, Objective: 5.22416e+00, Relative improvement 1.55265e-04"
## [1] "Iteration:  581, Objective: 5.21612e+00, Relative improvement 1.53385e-04"
## [1] "Iteration:  591, Objective: 5.20812e+00, Relative improvement 1.53702e-04"
## [1] "Iteration:  601, Objective: 5.20011e+00, Relative improvement 1.54184e-04"
## [1] "Iteration:  611, Objective: 5.19208e+00, Relative improvement 1.54763e-04"
## [1] "Iteration:  621, Objective: 5.18403e+00, Relative improvement 1.55411e-04"
## [1] "Iteration:  631, Objective: 5.17596e+00, Relative improvement 1.56110e-04"
## [1] "Iteration:  641, Objective: 5.16787e+00, Relative improvement 1.56852e-04"
## [1] "Iteration:  651, Objective: 5.15975e+00, Relative improvement 1.57234e-04"
## [1] "Iteration:  661, Objective: 5.15168e+00, Relative improvement 1.55689e-04"
## [1] "Iteration:  671, Objective: 5.14364e+00, Relative improvement 1.56375e-04"
## [1] "Iteration:  681, Objective: 5.13559e+00, Relative improvement 1.57113e-04"
## [1] "Iteration:  691, Objective: 5.12751e+00, Relative improvement 1.56894e-04"
## [1] "Iteration:  701, Objective: 5.11949e+00, Relative improvement 1.56401e-04"
## [1] "Iteration:  711, Objective: 5.11147e+00, Relative improvement 1.57065e-04"
## [1] "Iteration:  721, Objective: 5.10346e+00, Relative improvement 1.55177e-04"
## [1] "Iteration:  731, Objective: 5.09555e+00, Relative improvement 1.55313e-04"
## [1] "Iteration:  741, Objective: 5.08763e+00, Relative improvement 1.55774e-04"
## [1] "Iteration:  751, Objective: 5.07970e+00, Relative improvement 1.56318e-04"
## [1] "Iteration:  761, Objective: 5.07175e+00, Relative improvement 1.56929e-04"
## [1] "Iteration:  771, Objective: 5.06378e+00, Relative improvement 1.57349e-04"
## [1] "Iteration:  781, Objective: 5.05585e+00, Relative improvement 1.56831e-04"
## [1] "Iteration:  791, Objective: 5.04792e+00, Relative improvement 1.57151e-04"
## [1] "Iteration:  801, Objective: 5.03998e+00, Relative improvement 1.57822e-04"
## [1] "Iteration:  811, Objective: 5.03201e+00, Relative improvement 1.58548e-04"
## [1] "Iteration:  821, Objective: 5.02402e+00, Relative improvement 1.59313e-04"
## [1] "Iteration:  831, Objective: 5.01600e+00, Relative improvement 1.60107e-04"
## [1] "Iteration:  841, Objective: 5.00797e+00, Relative improvement 1.60252e-04"
## [1] "Iteration:  851, Objective: 4.99993e+00, Relative improvement 1.60971e-04"
## [1] "Iteration:  861, Objective: 4.99188e+00, Relative improvement 1.60921e-04"
## [1] "Iteration:  871, Objective: 4.98395e+00, Relative improvement 1.57798e-04"
## [1] "Iteration:  881, Objective: 4.97610e+00, Relative improvement 1.57296e-04"
## [1] "Iteration:  891, Objective: 4.96828e+00, Relative improvement 1.57390e-04"
## [1] "Iteration:  901, Objective: 4.96045e+00, Relative improvement 1.57824e-04"
## [1] "Iteration:  911, Objective: 4.95262e+00, Relative improvement 1.58384e-04"
## [1] "Iteration:  921, Objective: 4.94476e+00, Relative improvement 1.59028e-04"
## [1] "Iteration:  931, Objective: 4.93689e+00, Relative improvement 1.59143e-04"
## [1] "Iteration:  941, Objective: 4.92903e+00, Relative improvement 1.59435e-04"
## [1] "Iteration:  951, Objective: 4.92120e+00, Relative improvement 1.58881e-04"
## [1] "Iteration:  961, Objective: 4.91337e+00, Relative improvement 1.59407e-04"
## [1] "Iteration:  971, Objective: 4.90553e+00, Relative improvement 1.60018e-04"
## [1] "Iteration:  981, Objective: 4.89767e+00, Relative improvement 1.60523e-04"
## [1] "Iteration:  991, Objective: 4.88985e+00, Relative improvement 1.60051e-04"
```

```r
summary(spca1)
```

```
##                           PC1   PC2   PC3   PC4   PC5   PC6   PC7   PC8
## Explained variance      7.968 3.451 2.677 1.222 1.154 0.826 0.558 0.435
## Standard deviations     2.823 1.858 1.636 1.106 1.074 0.909 0.747 0.660
## Proportion of variance 0.398 0.173 0.134 0.061 0.058 0.041 0.028 0.022
## Cumulative proportion  0.398 0.571 0.705 0.766 0.824 0.865 0.893 0.915
##                           PC9  PC10  PC11  PC12  PC13  PC14  PC15  PC16
## Explained variance      0.402 0.307 0.277 0.219 0.166 0.112 0.081 0.062
## Standard deviations     0.634 0.554 0.526 0.468 0.407 0.335 0.285 0.250
```

```
## Proportion of variance 0.020 0.015 0.014 0.011 0.008 0.006 0.004 0.003
## Cumulative proportion  0.935 0.950 0.964 0.975 0.983 0.989 0.993 0.996
##                         PC17  PC18  PC19  PC20
## Explained variance     0.032 0.000 0.000 0.000
## Standard deviations    0.180 0.018 0.000 0.000
## Proportion of variance 0.002 0.000 0.000 0.000
## Cumulative proportion  0.997 0.998 0.998 0.998
```