Análisis de Datos Multivariantes
Práctica: Análisis de Componentes Principales (ACP)
Conjunto de datos: Ginecología

Contents

1 Datos. **2**

2 Análisis de Componentes Principales con R (princomp(...)) **4**
- 2.1 Puntuaciones de los países respecto de las CP 5
- 2.2 Distancias en el espacio de CP - espacio original 5
  - 2.2.1 Usando 2 CP 5

3 Correlación entre variables originales y CP **7**

4 ACP sobre los datos estandarizados **7**
- 4.1 Análisis de Componentes Principales con R (princomp(...)) 8
- 4.2 Puntuaciones de los países respecto de las CP 9
  - 4.2.1 Usando 2 CP 9

5 Correlación entre variables originales y CP **11**

1 Datos.

```
library(openxlsx)
```

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

```
Datos=read.xlsx("Ejemplo_ginecologia.xlsx")
Paises=Datos[,1]

Datos[,c(2:11)]
```

```
##    Etiqueta ISR1 ISR2  ISR3 ISR4 IAPP1 IAPP2    IG1  IG2  IG3
## 1        Ar  2.4 18.8  43.5 75.0  83.9  98.7  697.0  9.1 54.0
## 2        Ba  1.5 12.0  80.5 55.0 100.0 100.0  505.9  6.6 66.4
## 3        Be  3.1 26.6  68.4 47.0  98.0  97.0  126.0  4.8 51.2
## 4        Bo  3.7 28.7 390.0 48.0  55.0  65.0   54.0  5.5 69.1
## 5        Br  2.2 19.4  44.9 77.0  46.0  96.6  271.0  8.4 40.4
## 6        Ca  1.5 10.1   2.4 80.0 100.0  99.0 2151.0  9.3 72.5
## 7        Ch  2.3 18.0  18.7 60.6  83.1  99.8  331.0  7.1 65.9
## 8        Co  2.6 21.9 104.9 77.0  91.0  86.4  211.0  9.3 54.8
## 9        CR  2.2 18.8  38.0 80.0  83.0  96.0  285.0  9.1 75.5
## 10       Cu  1.6 11.5  41.8 82.0 100.0 100.0  139.0  6.7 82.5
## 11       Ec  2.7 22.6  97.0 66.0  83.0  69.2   65.0  3.9 51.0
## 12       ES  2.8 24.6 120.0 60.0  76.0  58.0  161.0  8.2 40.3
## 13       EU  2.1 14.5   9.8 76.4  98.9  99.4 4432.6 13.1 45.3
## 14       Gu  4.3 33.6 153.0 38.0  37.0  41.4   91.0  5.4 26.4
## 15      Guy  2.3 21.4 133.3 31.0  95.3  93.8   39.0  4.6 80.6
## 16       Ha  3.9 30.1 523.0 28.0  78.8  24.2   24.0  5.9 41.9
## 17       Ho  3.6 29.4 108.0 50.0  85.3  42.2   59.0  7.2 36.9
## 18       Ja  2.3 20.2 106.2 66.0  70.1  98.5  140.0  4.9 44.7
## 19       Me  2.5 22.0  76.9 68.0  89.2  86.7  221.0  5.3 47.2
## 20       Ni  3.7 30.9  97.0 60.0  86.2  79.3   43.0  9.7 59.6
## 21       Pa  2.7 22.4  71.3 58.0  81.6  92.1  231.0  6.8 72.3
## 22       Pa  3.8 29.3 160.7 57.0  69.3  84.8  128.0  7.2 36.8
## 23       Pe  2.8 22.9 185.0 50.3  83.8  59.3   91.0  4.4 57.2
## 24       RD  2.7 23.0  82.0 64.0  99.0  97.0  112.0  6.5 29.0
## 25       TT  1.6 13.8  44.7 53.0  64.2  99.0  214.0  4.5 48.5
## 26       Ur  2.3 16.7  11.1 84.0  93.1  99.7  618.0 10.3 46.5
## 27       Ve  2.7 22.5  67.2 43.0  25.5  95.3  437.0  8.7 51.0
```

```
X=Datos[,c(3:11)]
X
```

```
##    ISR1 ISR2  ISR3 ISR4 IAPP1 IAPP2    IG1  IG2  IG3
## 1   2.4 18.8  43.5 75.0  83.9  98.7  697.0  9.1 54.0
## 2   1.5 12.0  80.5 55.0 100.0 100.0  505.9  6.6 66.4
## 3   3.1 26.6  68.4 47.0  98.0  97.0  126.0  4.8 51.2
## 4   3.7 28.7 390.0 48.0  55.0  65.0   54.0  5.5 69.1
## 5   2.2 19.4  44.9 77.0  46.0  96.6  271.0  8.4 40.4
## 6   1.5 10.1   2.4 80.0 100.0  99.0 2151.0  9.3 72.5
## 7   2.3 18.0  18.7 60.6  83.1  99.8  331.0  7.1 65.9
## 8   2.6 21.9 104.9 77.0  91.0  86.4  211.0  9.3 54.8
## 9   2.2 18.8  38.0 80.0  83.0  96.0  285.0  9.1 75.5
## 10  1.6 11.5  41.8 82.0 100.0 100.0  139.0  6.7 82.5
## 11  2.7 22.6  97.0 66.0  83.0  69.2   65.0  3.9 51.0
```

```
## 12  2.8 24.6 120.0 60.0  76.0  58.0  161.0  8.2 40.3
## 13  2.1 14.5   9.8 76.4  98.9  99.4 4432.6 13.1 45.3
## 14  4.3 33.6 153.0 38.0  37.0  41.4   91.0  5.4 26.4
## 15  2.3 21.4 133.3 31.0  95.3  93.8   39.0  4.6 80.6
## 16  3.9 30.1 523.0 28.0  78.8  24.2   24.0  5.9 41.9
## 17  3.6 29.4 108.0 50.0  85.3  42.2   59.0  7.2 36.9
## 18  2.3 20.2 106.2 66.0  70.1  98.5  140.0  4.9 44.7
## 19  2.5 22.0  76.9 68.0  89.2  86.7  221.0  5.3 47.2
## 20  3.7 30.9  97.0 60.0  86.2  79.3   43.0  9.7 59.6
## 21  2.7 22.4  71.3 58.0  81.6  92.1  231.0  6.8 72.3
## 22  3.8 29.3 160.7 57.0  69.3  84.8  128.0  7.2 36.8
## 23  2.8 22.9 185.0 50.3  83.8  59.3   91.0  4.4 57.2
## 24  2.7 23.0  82.0 64.0  99.0  97.0  112.0  6.5 29.0
## 25  1.6 13.8  44.7 53.0  64.2  99.0  214.0  4.5 48.5
## 26  2.3 16.7  11.1 84.0  93.1  99.7  618.0 10.3 46.5
## 27  2.7 22.5  67.2 43.0  25.5  95.3  437.0  8.7 51.0
```

VARIABLES:

- ISR1 : Tasa global de fecundidad (hijos/mujer) 2003
- ISR2 : Tasa cruda de natalidad (por 1.000 habitantes)
- ISR3 : Razón de mortalidad materna reportada (por 100.000 nv)
- ISR4 : Prevalencia de uso de métodos anticonceptivos en mujeres en edad fértil (%)
- IAPP1 : % población gestante atendida por personal capacitado durante embarazo
- IAPP2 : Proporción de partos atendidos por personal capacitado (%)
- IG1 : Gasto nacional en salud por año per cápita (US$ corrientes) ($ per capita)
- IG2 :Gasto nacional en salud por año como proporción del PIB (%)
- IG3 : Gasto público en salud por año como % del gasto nacional en salud

2 Análisis de Componentes Principales con R (princomp(. . .))

```
CP=princomp(X)
CP
```

```
## Call:
## princomp(x = X)
##
## Standard deviations:
##      Comp.1      Comp.2      Comp.3      Comp.4      Comp.5      Comp.6
## 881.9151999 105.9558193  20.7363090  15.3541628  11.5093857   9.8795147
##      Comp.7      Comp.8      Comp.9
##   3.4734634   1.2819322   0.1263344
##
##  9  variables and  27 observations.
```

```
summary(CP,loading=T)
```

```
## Importance of components:
##                             Comp.1       Comp.2       Comp.3       Comp.4
## Standard deviation     881.915200 105.95581927 2.073631e+01 1.535416e+01
## Proportion of Variance   0.984636   0.01421254 5.443585e-04 2.984519e-04
## Cumulative Proportion    0.984636   0.99884855 9.993929e-01 9.996914e-01
##                             Comp.5       Comp.6       Comp.7       Comp.8
## Standard deviation     1.150939e+01 9.8795147326 3.473463e+00 1.281932e+00
## Proportion of Variance 1.676974e-04 0.0001235644 1.527381e-05 2.080425e-06
## Cumulative Proportion  9.998591e-01 0.9999826256 9.999979e-01 1.000000e+00
##                             Comp.9
## Standard deviation     1.263344e-01
## Proportion of Variance 2.020529e-08
## Cumulative Proportion  1.000000e+00
##
## Loadings:
##       Comp.1 Comp.2 Comp.3 Comp.4 Comp.5 Comp.6 Comp.7 Comp.8 Comp.9
## ISR1                                                -0.115         0.993
## ISR2                  0.122  0.106                -0.960 -0.181 -0.117
## ISR3         -0.985                -0.100
## ISR4                -0.182        -0.812  0.538
## IAPP1               -0.750  0.612        -0.244
## IAPP2         0.138 -0.338 -0.642 -0.313 -0.586 -0.108
## IG1    0.999
## IG2                                                 -0.187  0.980
## IG3                 -0.514 -0.440  0.476  0.550 -0.110
```

2.1 Puntuaciones de los países respecto de las CP

```
CP_X=CP$scores
rownames(CP_X)=Paises

# Solo las de las 6 primeras CP
round(CP_X[,1:6],2)
```

```
##                    Comp.1  Comp.2 Comp.3 Comp.4 Comp.5 Comp.6
## Argentina          259.69   54.76  -3.88  -2.64  -9.21   0.62
## Barbados            67.24   25.93 -24.52  -3.06   8.51  -9.16
## Belice            -311.95   52.12 -11.39   7.37   9.44 -19.95
## Bolivia           -397.60 -267.03  -9.80 -31.77  -6.71   8.65
## Brasil            -166.20   69.85  29.96 -17.66 -19.38   4.49
## Canadá            1714.21   35.14 -17.10  -1.43   4.37  11.50
## Chile             -105.03   93.95  -6.38  -6.91   9.14  -0.34
## Colombia          -228.43   13.48 -13.69   5.64 -13.96   5.27
## Costa Rica        -151.68   78.07 -15.50  -9.53  -3.00  17.02
## Cuba              -297.55   82.09 -35.51  -5.52  -3.07  15.47
## Ecuador           -374.22   23.84   2.30  14.13  -1.10   9.25
## El Salvador       -279.42   -5.39  16.32  19.73  -0.15   7.78
## Estados Unidos    3993.37  -69.81   8.02   4.94   0.86  -5.05
## Guatemala         -351.24  -40.72  59.87  10.42  12.03   5.91
## Guyana            -401.66   -9.19 -27.98 -11.29  30.29 -11.88
## Haití             -433.35 -403.88  -9.46   9.85  -3.25  -3.55
## Honduras          -380.98    7.85  19.60  38.44  12.77   7.66
## Jamaica           -299.53   15.20   4.42 -10.96 -14.56  -8.39
## México            -217.36   39.74  -4.07   9.66  -7.38  -2.72
## Nicaragua         -396.17   25.75  -5.89   6.55   4.87   3.98
## Panamá            -207.20   45.05 -10.61  -9.35  11.28   4.65
## Paraguay          -313.95  -41.03  11.01  -2.84 -11.92 -11.48
## Perú              -352.04  -66.44  -3.90  10.64   8.97   6.84
## R. Dominicana     -326.38   40.36  -5.69  16.84 -16.57 -23.50
## Trinidad y Tobago -223.16   71.76  14.75 -12.85   3.57  -9.77
## Uruguay            182.23   91.07  -6.28   8.47 -17.26  -0.27
## Venezuela           -1.62   37.46  45.41 -36.88  11.43  -3.02
```

2.2 Distancias en el espacio de CP - espacio original

2.2.1 Usando 2 CP

```
Y_r2=CP_X[,c(1,2)]
round(Y_r2[c(1:10),],3)
```

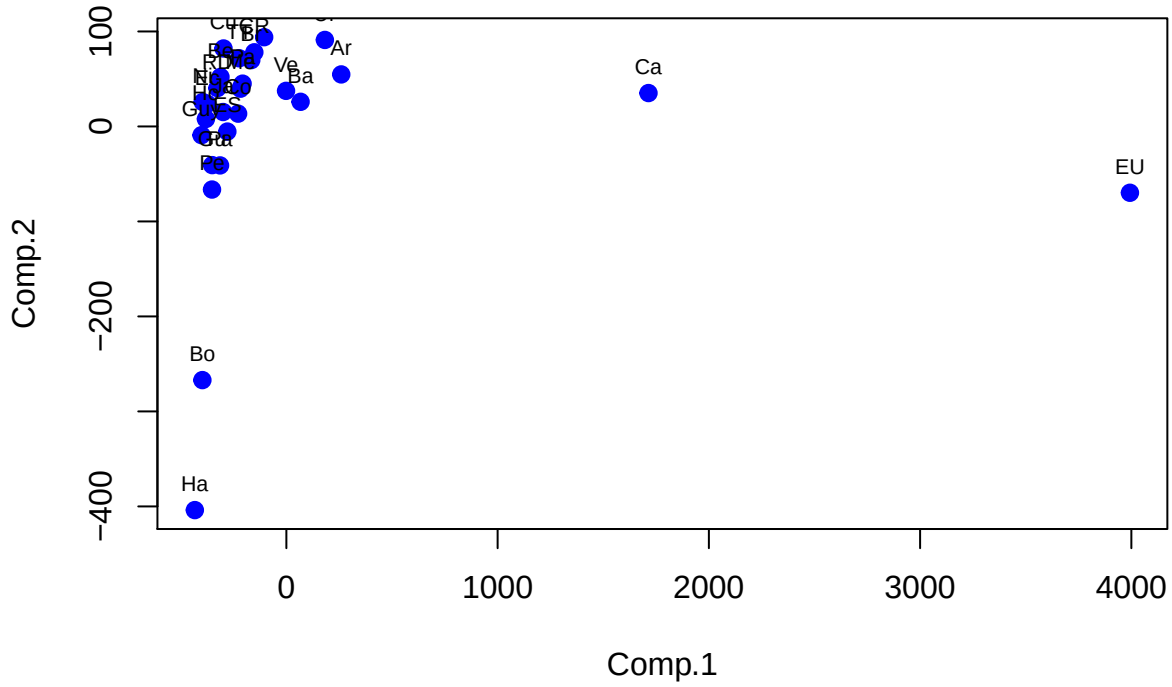
```
##              Comp.1   Comp.2
## Argentina   259.695   54.757
## Barbados     67.240   25.927
## Belice     -311.954   52.123
## Bolivia    -397.602 -267.032
## Brasil     -166.199   69.853
## Canadá     1714.215   35.144
## Chile      -105.034   93.948
## Colombia   -228.432   13.479
```

```
## Costa Rica -151.684   78.073
## Cuba       -297.545   82.091
```

```
round(var(Y_r2),3)
```

```
##          Comp.1   Comp.2
## Comp.1 807688.8     0.00
## Comp.2      0.0 11658.43
```

```
# Representación gráfica
{plot(Y_r2, col= "blue",  cex=1, pch = 19, lty = "solid", lwd = 2)
text(Y_r2, labels=Datos[,2], cex= 0.7, pos=3)}
```



3 Correlación entre variables originales y CP

```
round(cor(X,CP_X),3)
```

```
##       Comp.1 Comp.2 Comp.3 Comp.4 Comp.5 Comp.6 Comp.7 Comp.8 Comp.9
## ISR1  -0.355 -0.557  0.415  0.269  0.059 -0.015 -0.535 -0.017  0.168
## ISR2  -0.465 -0.479  0.414  0.266  0.093 -0.067 -0.545 -0.038 -0.002
## ISR3  -0.329 -0.944 -0.019 -0.011 -0.010 -0.003  0.000  0.000  0.000
## ISR4   0.375  0.550 -0.247  0.031 -0.611  0.347 -0.016 -0.007  0.000
## IAPP1  0.262  0.169 -0.806  0.487  0.020 -0.125 -0.003  0.001  0.000
## IAPP2  0.296  0.693 -0.332 -0.466 -0.170 -0.274 -0.018  0.000  0.000
## IG1    1.000 -0.005  0.000  0.000  0.000  0.000  0.000  0.000  0.000
## IG2    0.656  0.171  0.046 -0.003 -0.258  0.221 -0.299  0.577 -0.001
## IG3    0.023  0.142 -0.712 -0.452  0.367  0.363 -0.026 -0.002  0.000
```

4 ACP sobre los datos estandarizados

```
summary(X)
```

```
##       ISR1            ISR2            ISR3            ISR4
##  Min.   :1.500   Min.   :10.10   Min.   :  2.4   Min.   :28.00
##  1st Qu.:2.250   1st Qu.:18.40   1st Qu.: 44.1   1st Qu.:50.15
##  Median :2.600   Median :22.00   Median : 80.5   Median :60.00
##  Mean   :2.663   Mean   :21.69   Mean   :106.6   Mean   :60.53
##  3rd Qu.:2.950   3rd Qu.:25.60   3rd Qu.:114.0   3rd Qu.:75.70
##  Max.   :4.300   Max.   :33.60   Max.   :523.0   Max.   :84.00
##      IAPP1            IAPP2             IG1              IG2
##  Min.   : 25.50   Min.   : 24.20   Min.   :  24.0   Min.   : 3.90
##  1st Qu.: 73.05   1st Qu.: 74.25   1st Qu.:  91.0   1st Qu.: 5.35
##  Median : 83.80   Median : 95.30   Median : 161.0   Median : 6.80
##  Mean   : 79.86   Mean   : 83.64   Mean   : 439.9   Mean   : 7.13
##  3rd Qu.: 94.20   3rd Qu.: 98.85   3rd Qu.: 308.0   3rd Qu.: 8.90
##  Max.   :100.00   Max.   :100.00   Max.   :4432.6   Max.   :13.10
##       IG3
##  Min.   :26.40
##  1st Qu.:43.30
##  Median :51.00
##  Mean   :53.61
##  3rd Qu.:66.15
##  Max.   :82.50
```

```
summary(scale(X))
```

```
##       ISR1               ISR2               ISR3
##  Min.   :-1.52588   Min.   :-1.85946   Min.   :-0.92521
##  1st Qu.:-0.54183   1st Qu.:-0.52813   1st Qu.:-0.55510
##  Median :-0.08261   Median : 0.04931   Median :-0.23202
##  Mean   : 0.00000   Mean   : 0.00000   Mean   : 0.00000
##  3rd Qu.: 0.37661   3rd Qu.: 0.62675   3rd Qu.: 0.06532
##  Max.   : 2.14790   Max.   : 1.90996   Max.   : 3.69550
##       ISR4              IAPP1              IAPP2               IG1
##  Min.   :-2.0883   Min.   :-2.7638   Min.   :-2.7569   Min.   :-0.4632
```

```
##  1st Qu.:-0.6663   1st Qu.:-0.3464   1st Qu.:-0.4357   1st Qu.:-0.3886
##  Median :-0.0340   Median : 0.2002   Median : 0.5406   Median :-0.3106
##  Mean   : 0.0000   Mean   : 0.0000   Mean   : 0.0000   Mean   : 0.0000
##  3rd Qu.: 0.9739   3rd Qu.: 0.7289   3rd Qu.: 0.7052   3rd Qu.:-0.1469
##  Max.   : 1.5067   Max.   : 1.0238   Max.   : 0.7585   Max.   : 4.4467
##       IG2               IG3
##  Min.   :-1.4567   Min.   :-1.7860
##  1st Qu.:-0.8027   1st Qu.:-0.6768
##  Median :-0.1487   Median :-0.1714
##  Mean   : 0.0000   Mean   : 0.0000
##  3rd Qu.: 0.7985   3rd Qu.: 0.8230
##  Max.   : 2.6930   Max.   : 1.8961
```

```
round(var(scale(X)),3)
```

```
##         ISR1   ISR2   ISR3   ISR4  IAPP1  IAPP2    IG1    IG2    IG3
## ISR1   1.000  0.974  0.631 -0.566 -0.385 -0.751 -0.352 -0.179 -0.475
## ISR2   0.974  1.000  0.593 -0.603 -0.395 -0.719 -0.462 -0.267 -0.471
## ISR3   0.631  0.593  1.000 -0.633 -0.236 -0.738 -0.324 -0.376 -0.128
## ISR4  -0.566 -0.603 -0.633  1.000  0.350  0.569  0.372  0.564  0.151
## IAPP1 -0.385 -0.395 -0.236  0.350  1.000  0.266  0.261  0.131  0.346
## IAPP2 -0.751 -0.719 -0.738  0.569  0.266  1.000  0.293  0.288  0.391
## IG1   -0.352 -0.462 -0.324  0.372  0.261  0.293  1.000  0.655  0.022
## IG2   -0.179 -0.267 -0.376  0.564  0.131  0.288  0.655  1.000  0.000
## IG3   -0.475 -0.471 -0.128  0.151  0.346  0.391  0.022  0.000  1.000
```

4.1 Análisis de Componentes Principales con R (princomp(...))

```
CP=princomp(X, cor=T)
CP
```

```
## Call:
## princomp(x = X, cor = T)
##
## Standard deviations:
##    Comp.1    Comp.2    Comp.3    Comp.4    Comp.5    Comp.6    Comp.7
## 2.1294523 1.2284322 0.9813885 0.8211551 0.7470139 0.6202462 0.4512843
##    Comp.8    Comp.9
## 0.3977368 0.1199203
##
##  9  variables and  27 observations.
```

```
summary(CP,loading=T)
```

```
## Importance of components:
##                           Comp.1    Comp.2    Comp.3     Comp.4     Comp.5
## Standard deviation     2.1294523 1.2284322 0.9813885 0.82115507 0.74701390
## Proportion of Variance 0.5038408 0.1676718 0.1070137 0.07492174 0.06200331
## Cumulative Proportion  0.5038408 0.6715125 0.7785263 0.85344799 0.91545130
##                           Comp.6     Comp.7     Comp.8      Comp.9
## Standard deviation     0.62024619 0.45128429 0.39773685 0.119920313
## Proportion of Variance 0.04274504 0.02262861 0.01757718 0.001597876
## Cumulative Proportion  0.95819633 0.98082495 0.99840212 1.000000000
##
## Loadings:
```

```
##       Comp.1 Comp.2 Comp.3 Comp.4 Comp.5 Comp.6 Comp.7 Comp.8 Comp.9
## ISR1  -0.416  0.251  0.101  0.135 -0.298 -0.257         0.309  0.694
## ISR2  -0.425  0.175         0.202 -0.302 -0.333         0.206 -0.708
## ISR3  -0.366         0.422 -0.262         0.470 -0.622
## ISR4   0.366  0.209         0.359 -0.339  0.582         0.474
## IAPP1  0.232 -0.162  0.650  0.647  0.133 -0.191        -0.135
## IAPP2  0.392 -0.129 -0.312                -0.417 -0.733
## IG1    0.264  0.459  0.358 -0.374  0.413 -0.196         0.482
## IG2    0.240  0.592  0.178 -0.160 -0.439                -0.582
## IG3    0.209 -0.501  0.349 -0.398 -0.560 -0.118  0.238  0.199
```

4.2 Puntuaciones de los países respecto de las CP

```
CP_X=CP$scores
rownames(CP_X)=Paises
round(CP_X[1:10,],2)
```

```
##            Comp.1 Comp.2 Comp.3 Comp.4 Comp.5 Comp.6 Comp.7 Comp.8 Comp.9
## Argentina    1.53   0.61  -0.19   0.19  -0.42   0.11  -0.21  -0.09   0.06
## Barbados     1.96  -1.53   0.48  -0.28   0.78  -0.03  -0.10  -0.69   0.10
## Belice      -0.70  -0.82   0.05   0.97   0.35  -1.43  -0.36   0.29  -0.01
## Bolivia     -3.02  -0.69   0.83  -1.48  -0.94   0.58  -0.68   0.75   0.01
## Brasil       0.75   0.96  -2.01  -0.45  -0.21   0.86  -0.20  -0.06  -0.17
## Canadá       3.81   0.21   1.06  -0.63   0.31   0.28   0.44   0.29   0.08
## Chile        1.23  -0.76  -0.28  -0.18  -0.25  -0.47   0.17  -0.09   0.16
## Colombia     0.79   0.53   0.32   0.66  -0.88   0.49  -0.22  -0.26  -0.11
## Costa Rica   1.90  -0.30   0.07  -0.14  -1.44   0.33   0.24   0.04  -0.13
## Cuba         2.84  -1.81   0.40   0.17  -0.63   0.78   0.27   0.16   0.19
```

4.2.1 Usando 2 CP

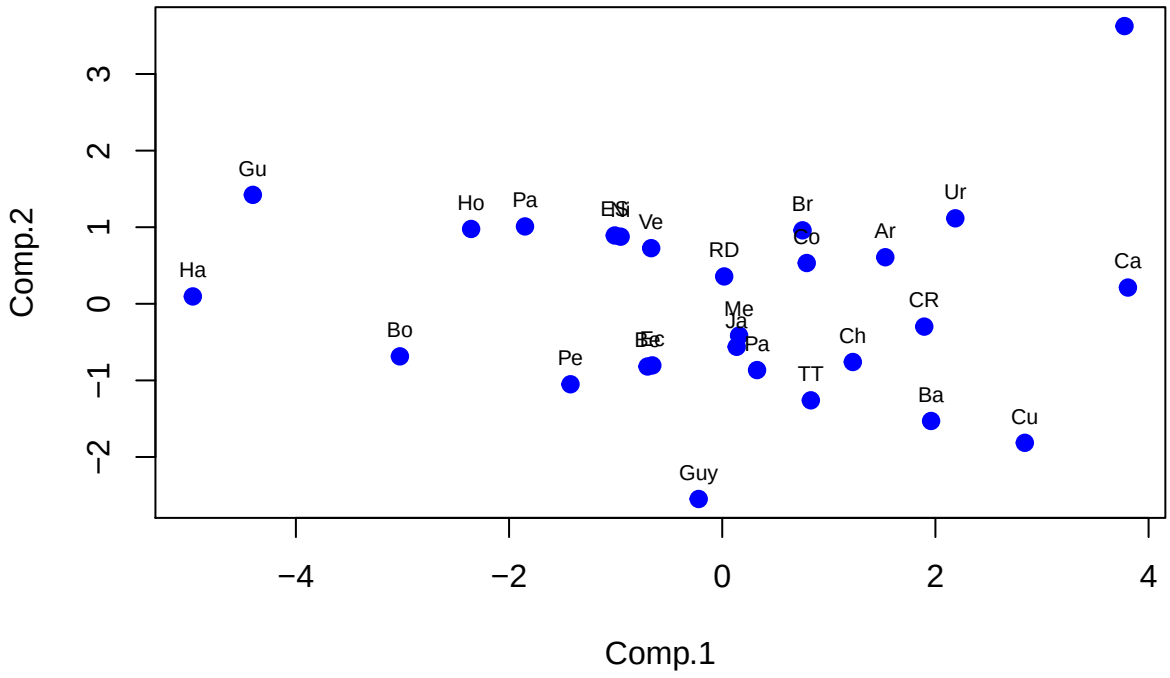
```
Y_r2=CP_X[,c(1,2)]
round(Y_r2[c(1:10),],3)
```

```
##            Comp.1 Comp.2
## Argentina   1.530  0.608
## Barbados    1.959 -1.530
## Belice     -0.701 -0.819
## Bolivia    -3.024 -0.686
## Brasil      0.753  0.959
## Canadá      3.806  0.212
## Chile       1.225 -0.760
## Colombia    0.793  0.532
## Costa Rica  1.895 -0.298
## Cuba        2.839 -1.815
```

```
round(var(Y_r2),3)
```

```
##        Comp.1 Comp.2
## Comp.1  4.709  0.000
## Comp.2  0.000  1.567
```

```r
# Representación gráfica
{plot(Y_r2, col= "blue",  cex=1, pch = 19, lty = "solid", lwd = 2)
text(Y_r2, labels=Datos[,2], cex= 0.7, pos=3)}
```



5 Correlación entre variables originales y CP

```
round(cor(X,CP_X),3)
```

```
##       Comp.1 Comp.2 Comp.3 Comp.4 Comp.5 Comp.6 Comp.7 Comp.8 Comp.9
## ISR1  -0.886  0.308  0.099  0.111 -0.223 -0.159 -0.028  0.123  0.083
## ISR2  -0.904  0.215  0.013  0.165 -0.226 -0.207 -0.021  0.082 -0.085
## ISR3  -0.779 -0.106  0.414 -0.215  0.017  0.291 -0.281 -0.006 -0.008
## ISR4   0.779  0.257 -0.098  0.295 -0.253  0.361 -0.019  0.188 -0.007
## IAPP1  0.494 -0.199  0.637  0.531  0.100 -0.119 -0.038 -0.054  0.002
## IAPP2  0.835 -0.158 -0.306 -0.031 -0.072 -0.259 -0.331  0.039  0.004
## IG1    0.563  0.564  0.352 -0.307  0.308 -0.121  0.021  0.192 -0.011
## IG2    0.511  0.727  0.175 -0.132 -0.328 -0.013 -0.023 -0.231 -0.001
## IG3    0.445 -0.616  0.343 -0.327 -0.418 -0.073  0.108  0.079 -0.002
```