

Análisis Multivariado. Tarea 5

José Antonio García Ramírez

April 27, 2018

1. Muestra que la matriz de covarianza

$$\rho = \begin{pmatrix} 1.0 & .63 & .45 \\ .63 & 1.0 & .35 \\ .45 & .35 & 1.0 \end{pmatrix}$$

Para las $p = 3$ v.a. estandarizadas Z_1, Z_2 y Z_3 pueden ser generadas por el modelo de factores con $m = 1$

$$Z_1 = 0.9F_1 + \epsilon_1$$

$$Z_2 = 0.7F_1 + \epsilon_2$$

$$Z_3 = 0.5F_1 + \epsilon_3$$

Donde $\text{Var}(F_1) = 1, \text{Cov}(\epsilon, F_1) = 0$ y

$$\Psi = \text{Cov}(\epsilon) = \begin{pmatrix} 0.19 & 0 & 0 \\ 0 & 0.51 & 0 \\ 0 & 0 & 0.75 \end{pmatrix}$$

Esto es, escribe ρ de la forma $\rho = LL^t + \Psi$

En vista de que de manera a priori tenemos que los supuestos del modelo de análisis de factores se cumplen y como sabemos que las cargas de cada factor son las covarianzas del factor con cada variable Z_1 , tenemos que

$$L = \begin{pmatrix} 0.9 \\ 0.7 \\ 0.5 \end{pmatrix}$$

Lo que nos permite escribir

$$\rho = LL^t + \Psi = \begin{pmatrix} 0.9 \\ 0.7 \\ 0.5 \end{pmatrix} (0.9, 0.7, 0.5)^t + \Psi = \begin{pmatrix} 0.81 & 0.63 & 0.45 \\ 0.63 & 0.49 & 0.35 \\ 0.45 & 0.35 & 0.25 \end{pmatrix} + \begin{pmatrix} 0.19 & 0.00 & 0.00 \\ 0.00 & 0.51 & 0.00 \\ 0.00 & 0.00 & 0.75 \end{pmatrix} = \begin{pmatrix} 1.0 & .63 & .45 \\ .63 & 1.0 & .35 \\ .45 & .35 & 1.0 \end{pmatrix}$$

2. Usa la información del ejercicio anterior

a. Calcula las comunales $h_i^2, i, i = 1, 2, 3$ e interpreta estas cantidades.

Los valores h_i^2 son, $h_1^2 = 0.81, h_2^2 = 0.49$ y $h_3^2 = 0.25$. De manera general estos coeficientes h_i^2 son las comunales de cada variable con el factor F_1 es decir que este factor explica h_i^2 de la varianza de cada variable Z_i , es importante notar que el factor F_1 explica gran parte de la varianza de Z_1 mientras que solo explica la mitad de la varianza de Z_2 y apenas la cuarta parte de la varianza de Z_3 .

b. Calcula $\text{Corr}(Z_i, F_1)$ para $i = 1, 2, 3$. ¿Cual variable podría llevar el mayor peso en la interpretación del factor común? ¿Porqué?

Sabemos que la correlación de la variable i -ésima con el factor j -ésimo es justamente el coeficiente (i, j) de la matriz L , por lo que $\text{Corr}(Z_1, F_1) = 0.9, \text{Corr}(Z_2, F_1) = 0.7$ y $\text{Corr}(Z_3, F_1) = 0.5$. Entonces la variable que lleva el mayor peso en la interpretación del factor F_1 es Z_1 puesto que su correlación es la más alta en comparación con las otras dos variables.

3. Los valores y vectores propios de la matriz de correlaciones ρ en el ejercicio 1 son

$$\begin{aligned}\lambda_1 &= 1.96, e_1^t = [0.625, 0.593, 0.507] \\ \lambda_2 &= 0.68, e_2^t = [-0.219, -0.491, 0.843] \\ \lambda_3 &= 0.36, e_3^t = [0.749, -0.638, -0.177]\end{aligned}$$

a) *Asumiendo un modelo de factores con $m = 1$, calcula la matriz de cargas L y la matriz de varianzas específicas Ψ usando el método por componentes principales. Compara los resultados con los del ejercicio 1.*

En vista de que conocemos la descomposición espectral de la matriz de correlaciones podemos utilizar la estimación que emplea PCA, por lo que la matriz de *loadings* L al usar este método de estimación es la siguiente:

$$\hat{L} = \sqrt{1.96}e_1^t = 1.4 \begin{pmatrix} 0.625 \\ 0.539 \\ 0.507 \end{pmatrix} = \sqrt{1.96}e_1 = 1.4 (0.625, 0.539, 0.507) =$$

Así que la matriz de especificadas esta dada por :

```
L <- c(0.625, 0.539, 0.507)*1.4
L <- as.matrix(L)
R <- matrix(c(0.19, 0, 0,
              0, 0.51, 0,
              0, 0, 0.75), byrow = TRUE, ncol = 3)
Psi <- R-L%*%t(L)
R

##      [,1] [,2] [,3]
## [1,] 0.19 0.00 0.00
## [2,] 0.00 0.51 0.00
## [3,] 0.00 0.00 0.75
```

De donde podemos ver que la manera de estimar el primer y único factor del inciso anterior fue por PCA, pues las estimaciones son idénticas.

b) *¿Qué proporción de la varianza poblacional total es explicada por el primer factor común?*

Es fácil saberlo ya que conocemos la matriz con la que estamos trabajando y sus vectores propios son dados, la proporción de la varianza poblacional total explicada por el primer factor esta dada por: $\hat{\lambda}_1/3 = 1.96/3 \approx 0.6533$

4. *(Solución única pero impropia: caso Heywood). Considere un modelo factorial con $m = 1$ para la población con matriz de covarianza*

$$\Sigma = \begin{pmatrix} 1 & 0.4 & 0.9 \\ 0.4 & 1 & 0.7 \\ 0.9 & 0.7 & 1 \end{pmatrix}$$

Muestra que existe una única elección de L y Ψ con $\Sigma = LL^t + \Psi$ pero que $\psi_3 < 0$, por lo que la elección no es admisible.

Realicemos los cálculos, para estimar Ψ , por el método de PCA y observaremos que el fenómeno también se observa:

```

Sigma <- matrix(c(1,.4,.9,
                 .4,1,.7,
                 .9,.7,1), byrow = TRUE, ncol = 3)
valores <- eigen(Sigma)$values
L <- eigen(Sigma)$vectors
L1 <- (valores[1]**.5)**as.matrix(L[, 1], ncol=1)
Psi <- Sigma - (L1 %*%t(L1))
Psi

```

```

##           [,1]      [,2]      [,3]
## [1,]  0.3895611 -0.22992068 0.30681391
## [2,] -0.2299207  0.34997578 0.08788274
## [3,]  0.3068139  0.08788274 0.42357911

```

De la salida anterior podemos ver que la matriz de Ψ tiene números fuera de la diagonal lo que no debe ser pues. Por lo que el modelo esta mal planteado y requiere más factores.

Y ahora realicemos la estimación por máxima verosimilitud para poder reproducir la falla o bien replicar lo que el ejercicio pide:

```

library(psych)
valores <- factanal( rotation = 'none', factors = 1, covmat = Sigma)
L <- valores$loadings
L1 <- as.matrix(L[, 1], ncol=1)
Psi <- diag(valores$uniquenesses)
Sigma - (L1 %*%t(L1) + Psi)

```

```

##           [,1]      [,2]      [,3]
## [1,]  2.494100e-07 -2.271320e-01  1.538661e-03
## [2,] -2.271320e-01  1.796799e-06  5.431971e-03
## [3,]  1.538661e-03  5.431971e-03 -7.359586e-05

```

Donde de manera análoga al ejercicio anterior, la estimación con PCA, vemos que existen números en particular en la posición (2,1) y (1,2) no cero en la diferencia $\Sigma - (LL^t + \Psi)$ por lo que el modelo está mal planteado pues no cumple los supuestos.

5. *El Proyecto de Evaluación de la Apertura Sintética de la Personalidad (SAPA) es una colección de datos psicológicos basada en la web. Un subconjunto de los datos está disponible en R como bfi en la biblioteca ‘psych’.*

Este subconjunto contiene datos en tres variables demográficas y 25 ítems de personalidad de 2800 voluntarios. Como ejemplos de estos ítems, tenemos:

- Sé cómo consolar a los demás.
- Desperdicio mi tiempo.
- Hago amigos con facilidad.

Cada ítem es clasificado en una escala de 1-7, en si el encuestado siente que él o ella está de acuerdo con la declaración mucho, no está de acuerdo mucho o cae en algún lugar intermedio. Consulta el archivo de ayuda de bfi para obtener más detalles.

- a. *Utilice el comando `complete.cases()` para eliminar individuos en bfi con cualquier valor faltantes*

El código se encuentre en el archivo `tarea5.Rmd`, pero despues de realizar el procedimiento anterior el conjunto de datos se llama ‘datos’, no se realizó ninguna rotación de los factores. Es importante mencionar que este conjunto de datos ‘toy’ es fácil de analizar pues todas sus variables son numericas.

- b. *Utilice el análisis de factores para agrupar elementos de naturaleza similar. Trata de interpretar la naturaleza de los ítems que se agrupan. Este es un ejercicio útil en psicología. El test de chi cuadrado para el número de factores puede no ser apropiado con una muestra tan grande.*

Como la documentación del conjunto de datos original ‘bfi’ dice explícitamente que se representan 5 factores, utilizare una estimación por máxima verosimilitud para encontrar 5 factores a pesar de que el modelo no pasa la prueba de Bartlett, obtenemos los siguientes 5 factores:

```
f5$loadings
```

```
##
## Loadings:
##      Factor1 Factor2 Factor3 Factor4 Factor5
## A1      0.218      0.129      -0.345
## A2     -0.384     0.358    -0.160     0.136     0.315
## A3     -0.454     0.399    -0.227      0.280     0.286
## A4     -0.372     0.208    -0.111     0.280     0.174
## A5     -0.534     0.290    -0.244      0.192
## C1     -0.273     0.187     0.458
## C2     -0.252     0.239     0.472     0.222     0.117
## C3     -0.284     0.126     0.374     0.281
## C4       0.442      -0.488    -0.238
## C5       0.482      -0.332    -0.293     0.123
## E1       0.362    -0.320     0.221      0.250
## E2       0.590    -0.237     0.166      0.320
## E3     -0.461     0.436      -0.216
## E4     -0.563     0.333    -0.287     -0.150
## E5     -0.423     0.426     0.143     -0.179
## N1       0.588     0.572      -0.183
## N2       0.569     0.554     -0.137
## N3       0.525     0.496
## N4       0.595     0.226      0.291
## N5       0.425     0.312     0.188     0.196
## O1     -0.273     0.217     0.218    -0.402
## O2       0.194      -0.308     0.356
## O3     -0.324     0.328     0.157    -0.495
## O4       0.131     0.158     0.136    -0.268     0.332
## O5       0.174      -0.262     0.396    -0.113
## gender      0.204      0.241
## education    -0.110     0.105
## age     -0.144
##
##      Factor1 Factor2 Factor3 Factor4 Factor5
## SS loadings    4.429    2.423    1.562    1.331    0.966
## Proportion Var    0.158    0.087    0.056    0.048    0.034
## Cumulative Var    0.158    0.245    0.301    0.348    0.383
```

Del diccionario de datos disponible para el conjunto de datos ‘bfi’ podemos ver que las variables N1 a N5 en su mayoría tienen cargas mayores en comparación a las demás variables (aunque también las variables A5, E2 y E4 tienen cargas grandes). Las variables N1 a N5 corresponden a variables que miden la estabilidad de la persona por lo que este factor “puede ser considerado” un factor de estabilidad emocional. Por otro lado, las variables A2, E1, E3 y E5 tienen pesos sobresalientes en el segundo factor y corresponden a preguntas de empatía por lo que así se puede interpretar este factor. En comparación el tercer factor tiene loadings altos en las variables C1 a C5 que reflejan escrupulosidad y orden por lo que podemos definir a este factor como una variable que refleja el nivel de detalle en las actividades del individuo (pues los pesos para C4 y C5 son negativos pero grandes) además. Para nuestra sorpresa el cuarto componente se ve cargado en sus cargas por las variables O1 a O5 que reflejan actitudes intelectuales por lo que este sería un factor que refleja la “inteligencia” en el sentido popular y finalmente el quinto factor es difícil de interpretar que podría considerarse como un factor aparte.

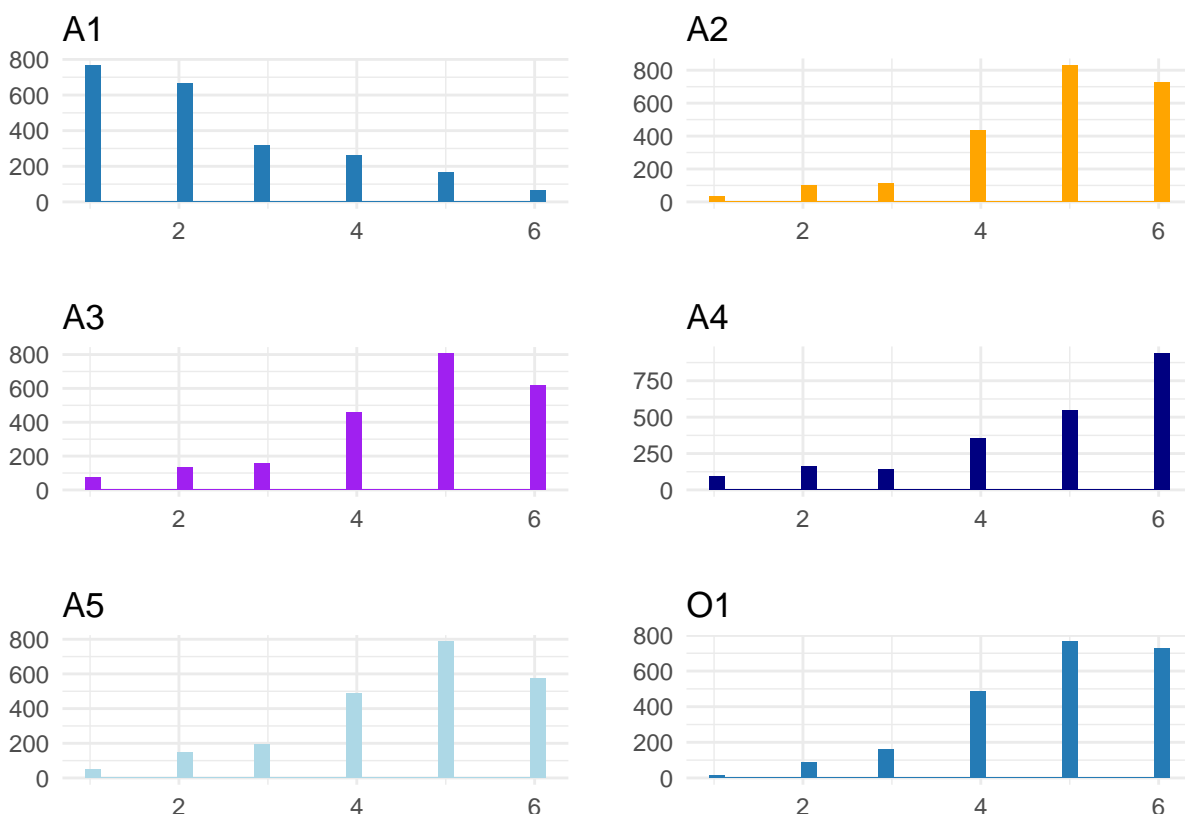
Pese a que la muestra es de tamaño grande $n = 2236$ y se cumple la condición de que $m < \frac{1}{2}(2p+1 - \sqrt{(8p+1)})$, podemos aplicar el test de Bartlett, teniendo reservas en su resultado. La salida del test es la siguiente:

```
## objective
## 1.797042e-254
```

Por lo que rechazamos la hipótesis nula y cinco factores no son buenos para representar a las 28 variables. Es importante recalcar que ente conjunto de datos aunque todas las variables son continuas, la edad y el genero tambien lo son, valdría mucho la pena dar una codificación correcta a estas variables en su naturaleza ordinal pero eso lo posponemos para el tema de analisis de datos categóricos que veremos en otras clases.

c. Identifica las preguntas que tienen una preponderancia de acuerdo extremo y / o en desacuerdo las respuestas. Del mismo modo, identifica casos atípicos tales como personas que parecen responder de manera extrema. Es decir, las personas que tienden a estar totalmente de acuerdo o en desacuerdo con la mayoría de las preguntas.

Para identificar las respuestas que tienen una preponderancia extrema se graficaron una por una las distribuciones de las respuestas y se identificó gráficamente que las variables A1, A2, A3, A4, A5, O1 y O4 son las que presentan preponderancias extremas como se puede ver en la siguiente figura:



En vista de que el modelo que se tiene con 5 factores no cumple el test de Bartlett, para identificar a los individuos que parecen responder de manera atípica se filtro el conjunto de registros manteniendo solo a los que estan totalmente de acuerdo o en desacuerdo en la mayoría de las preguntas que no conciernen a la edad ni al sexo ni a la educación. Se pudo identificar la distribución muestral del numero de veces que las personas contestaban de manera extrema la mayoría de las veces (es importante remarcar que el rango de esta distribución es $[0, 25]$ con mediana de 7.5) por lo que en la siguiente tabla mostramos los registros de las personas que presentaron ese comportamiento (superiores al último centil). Los cuales son 28 individuos (Con fines ilustrativos la tabla se muestra transpuesta y solo mostramos los primeros 12, los más “tranposos”).

	5	9	10	11	14	4	7	8	16	18	20	1
A1	6.00	1.00	1.00	1.00	1.00	6.00	6.00	1.00	6.00	6.00	1.00	1.00
A2	1.00	1.00	6.00	6.00	1.00	6.00	6.00	6.00	1.00	6.00	6.00	6.00
A3	1.00	1.00	6.00	6.00	1.00	6.00	6.00	6.00	1.00	6.00	6.00	6.00
A4	1.00	1.00	6.00	6.00	1.00	6.00	6.00	6.00	6.00	6.00	6.00	6.00
A5	1.00	1.00	6.00	6.00	1.00	6.00	6.00	6.00	1.00	6.00	6.00	6.00
C1	6.00	1.00	6.00	6.00	1.00	5.00	6.00	6.00	6.00	6.00	6.00	6.00
C2	6.00	1.00	6.00	1.00	1.00	6.00	5.00	6.00	5.00	6.00	6.00	6.00
C3	6.00	1.00	6.00	6.00	1.00	6.00	1.00	6.00	6.00	1.00	6.00	6.00
C4	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00
C5	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00
E1	6.00	1.00	1.00	1.00	1.00	1.00	6.00	4.00	6.00	6.00	1.00	5.00
E2	6.00	1.00	1.00	1.00	1.00	1.00	6.00	1.00	6.00	6.00	1.00	1.00
E3	1.00	1.00	6.00	6.00	1.00	6.00	1.00	6.00	1.00	6.00	6.00	6.00
E4	1.00	1.00	6.00	6.00	1.00	6.00	6.00	6.00	1.00	1.00	6.00	6.00
E5	6.00	1.00	6.00	6.00	1.00	6.00	6.00	6.00	1.00	6.00	6.00	6.00
N1	6.00	1.00	1.00	1.00	1.00	6.00	1.00	1.00	1.00	1.00	1.00	1.00
N2	6.00	1.00	1.00	1.00	1.00	6.00	6.00	1.00	1.00	1.00	1.00	6.00
N3	6.00	1.00	1.00	1.00	1.00	6.00	1.00	1.00	1.00	1.00	1.00	1.00
N4	6.00	1.00	1.00	1.00	1.00	1.00	6.00	1.00	1.00	1.00	1.00	1.00
N5	1.00	1.00	1.00	1.00	1.00	6.00	1.00	1.00	1.00	6.00	1.00	1.00
O1	6.00	1.00	6.00	6.00	1.00	6.00	6.00	6.00	6.00	6.00	6.00	6.00
O2	1.00	1.00	1.00	1.00	1.00	6.00	6.00	1.00	1.00	5.00	3.00	1.00
O3	6.00	1.00	6.00	1.00	1.00	6.00	1.00	6.00	1.00	6.00	6.00	3.00
O4	6.00	1.00	6.00	6.00	1.00	6.00	6.00	1.00	6.00	6.00	6.00	6.00
O5	1.00	1.00	1.00	6.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00
marca	25.00	25.00	25.00	25.00	25.00	24.00	24.00	24.00	24.00	24.00	24.00	23.00

6. El conjunto de datos *Harmon23.cor* en el paquete “*datasets*” es una matriz de correlación de ocho mediciones físicas realizadas en 305 niñas entre las edades de 7 y 17 años.

a. Realiza un análisis factorial de estos datos.

En vista de que solo disponemos de los datos, no podemos estimar verificar la normalidad de los datos. Por lo que damos un salto de fe y suponemos que existen factores con distribución normal con especificidades también normales (ambas en sentido multivariado) y realizamos una estimación por máxima verosimilitud de los factores sin realizar ninguna rotación.

En vista de que no leí con detenimiento la documentación de la implementación de que estamos usando construí a pie la prueba de verosimilitud (esto es posible porque aunque estamos haciendo la estimación máximo verosímil sobre la matriz de correlaciones el cociente de verosimilitudes entre la varianza muestral sujeta a los parámetro que estimamos es el mismo independientemente de la escala, puesto que la matriz de correlaciones es semejante a la matriz de varianzas por una transformación lineal como lo vimos en clase), por lo que decidí que 4 factores son adecuados pues el tamaño de muestra es razonablemente $n = 305$ y el número de factores no es tan reducido confió en la potencia de la prueba de Bartlett (y justo con 4 factores es cuando la prueba no rechaza la hipótesis nula, además con una rotación varimax los 4 factores son interpretables)

```
##
## Call:
## factanal(factors = m, covmat = ninas, rotation = "none")
##
## Uniquenesses:
##      height      arm.span      forearm      lower.leg      weight
##      0.137      0.005      0.191      0.116      0.138
## bitro.diameter  chest.girth  chest.width
```

```
##          0.283          0.178          0.488
##
## Loadings:
##          Factor1 Factor2 Factor3 Factor4
## height      0.861   0.135   0.316
## arm.span     0.997
## forearm      0.888          0.140
## lower.leg    0.843   0.122   0.384 -0.103
## weight       0.406   0.827          0.110
## bitro.diameter 0.350   0.716          0.279
## chest.girth  0.301   0.779 -0.144 -0.322
## chest.width  0.430   0.538 -0.173
##
##          Factor1 Factor2 Factor3 Factor4
## SS loadings    3.797   2.127   0.323   0.217
## Proportion Var  0.475   0.266   0.040   0.027
## Cumulative Var  0.475   0.740   0.781   0.808
##
## The degrees of freedom for the model is 2 and the fit was 0.0155
## [1] FALSE
```

b. *Varía el número de factores para encontrar un ajuste adecuado del modelo e interprete las cargas factoriales resultantes.*

Esto lo realice en el inciso anterior aprovecho este espacio para mostrar los factores rotados con una rotación varimax

```
f.ninas <- factanal(covmat = ninas, factors = m, rotation = 'varimax', n.obs = 305 )
L <- f.ninas$loadings
L
```

```
##
## Loadings:
##          Factor1 Factor2 Factor3 Factor4
## height      0.879   0.277          -0.115
## arm.span     0.937   0.194          0.277
## forearm      0.875   0.191
## lower.leg    0.887   0.209   0.135 -0.188
## weight       0.246   0.882   0.111 -0.109
## bitro.diameter 0.187   0.822
## chest.girth  0.117   0.729   0.526
## chest.width  0.263   0.644          0.141
##
##          Factor1 Factor2 Factor3 Factor4
## SS loadings    3.382   2.595   0.323   0.165
## Proportion Var  0.423   0.324   0.040   0.021
## Cumulative Var  0.423   0.747   0.787   0.808
```

Donde podemos ver en el primer factor que la altura de la niña se relaciona con la longitud de sus extremidades, con lo que este factor puede ser llamarse ‘crecimiento’, el segundo factor tiene que ver con el pecho y las mediciones de la cintura y pecho (una practica usual de la medicina familiar) con lo que este factor de que tan ‘obesa’ y siendo gentiles que tan delgada es la chica, por otra parte el tercer factor refleja que el peso se relaciona con la longitud de las extremidades lo que puede interpretarse como un factor de ‘complexión’ siendo bajo si el peso es bajo y finalmente el último factor puede considerarse como un complemento del tercero (puede sonar a información redundante y que tal vez el test de Bartlett fallo por el número de muestra).

7. La matriz de correlación dada a continuación proviene de las puntuaciones de 220 chicos en seis asignaturas escolares: 1) Francés, 2) Inglés, 3) Historia, 4) Aritmética, 5) Álgebra y 6) Geometría.

$$R = \begin{matrix} & \begin{matrix} \text{Frances} \\ \text{Ingles} \\ \text{Historia} \\ \text{Aritmetica} \\ \text{Algebra} \\ \text{Geometria} \end{matrix} & \begin{pmatrix} 1 & & & & & \\ 0.44 & 1 & & & & \\ 0.41 & 0.35 & 1.0 & & & \\ 0.29 & 0.35 & 0.16 & 1.0 & & \\ 0.33 & 0.32 & 0.19 & 0.59 & 1.0 & \\ 0.25 & 0.33 & 0.18 & 0.47 & 0.46 & 1 \end{pmatrix} \end{matrix}$$

- a. Encuentre la solución de dos factores de un análisis de factor de máxima verosimilitud.

Los factores son los siguientes:

```
mal.ajuste

##
## Call:
## factanal(factors = 2, covmat = R, n.obs = 220, rotation = "none")
##
## Uniquenesses:
## [1] 0.508 0.595 0.644 0.377 0.440 0.628
##
## Loadings:
##      Factor1 Factor2
## [1,] 0.558 0.425
## [2,] 0.569 0.286
## [3,] 0.392 0.450
## [4,] 0.738 -0.279
## [5,] 0.718 -0.209
## [6,] 0.595 -0.133
##
##      Factor1 Factor2
## SS loadings 2.204 0.603
## Proportion Var 0.367 0.101
## Cumulative Var 0.367 0.468
##
## Test of the hypothesis that 2 factors are sufficient.
## The chi square statistic is 2.18 on 4 degrees of freedom.
## The p-value is 0.703
```

De la salida anterior se pueden apreciar muchas cosas: dos factores bastan para explicar las 6 observaciones, la diferencia entre el tamaño de muestra con respecto al número de factores y la dimensión del vector aleatorio que se esta observando es aun mejor que en el caso de las niñas por lo existen elementos para confiar en el test de Bartlett, pese a que estos solo explican poco menos de la mitad de la varianza.

2. Mediante una inspección de las cargas, encuentre una rotación ortogonal que permite una interpretación más fácil de los resultados.

A continuación, se presentan los dos factores citados anteriormente rotados con la transformación varimax.

```
mal.ajuste <- factanal(covmat = R, factors = 2, rotation = 'varimax', n.obs = 220 )
mal.ajuste$loadings
```

```
##
## Loadings:
##      Factor1 Factor2
## [1,] 0.233 0.661
```



```
## [2,] 0.319 0.551
## [3,]      0.591
## [4,] 0.770 0.172
## [5,] 0.715 0.220
## [6,] 0.570 0.215
##
##               Factor1 Factor2
## SS loadings      1.593  1.215
## Proportion Var   0.265  0.202
## Cumulative Var   0.265  0.468
```

Donde confirmamos lo que esperábamos, y hemos revisado varias veces en el curso, las asignaturas relacionadas con un proceso de abstracción (justamente las de matemáticas) se agrupan en el primer factor que podemos llamar ‘Morphos’ de relativo a la forma, y las otras tres materias ‘Sociales’ que es como se agrupan esas tres asignaturas en la prepa en la que yo asistí.