Análisis Multivariado

Tarea 2

Duran Santiago Rafael

Garay Araujo Fernando Raul

Santillán Arriaga Diego Armando

1. Consideremos el vector aleatorio (X_1, X_2) con matriz de covarianza $\Sigma = \begin{pmatrix} 1 & \rho \\ \rho & 1 \end{pmatrix}$ con $0 < \rho < 1$. Por la forma de la matriz, sin pérdida de generalidad podemos considerar que las variables están estandarizadas de modo que $\mathbb{E}[(X_1, X_2)] = (0,0)^T$. Apliquemos análisis de componentes a este vector:

   el polinomio característico de Σ es

$$\begin{aligned} p(\lambda) &= det\Big(\Sigma - \lambda I\Big) \\ &= det\left(\begin{pmatrix} 1-\lambda & \rho \\ \rho & 1-\lambda \end{pmatrix}\right) \\ &= (1-\lambda)^2 - \rho^2 \\ &= \lambda^2 - 2\lambda + (1-\rho^2) \\ p(\lambda) &= \Big(\lambda - (1+\rho)\Big)\Big(\lambda - (1-\rho)\Big) \end{aligned}$$

   de donde se siguen que los eigenvalores de Σ son $\lambda_1 = 1+\rho$ y $\lambda_2 = 1-\rho$. Observemos que si $(x,y)^T$ es un eigenvector de Σ correspondiente a λ_1 entonces

$$\begin{pmatrix} -\rho & \rho \\ \rho & -\rho \end{pmatrix}\begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix}$$

   pero reduciendo $\Sigma - (1+\rho)I$

$$\Sigma - (1+\rho)I = \begin{pmatrix} -\rho & \rho \\ \rho & -\rho \end{pmatrix} \xrightarrow{-\frac{1}{\rho}R_1} \begin{pmatrix} 1 & -1 \\ \rho & -\rho \end{pmatrix} \xrightarrow{-\rho R_1 + R_2} \begin{pmatrix} 1 & -1 \\ 0 & 0 \end{pmatrix}$$

   el sistema anterior es equivalente al sistema:

$$\begin{pmatrix} 1 & -1 \\ 0 & 0 \end{pmatrix}\begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix}$$

   de donde se sigue que $x = y$ y las soluciones del sistema son de la forma $(r,r)^T$ con $r \in \mathbb{R}$. Similarmente un eigenvector $(x,y)^T$ de Σ correspondiente a λ_2 satisface que:

$$\begin{pmatrix} \rho & \rho \\ \rho & \rho \end{pmatrix}\begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix}$$

   y reduciendo $\Sigma - (1-\rho)I$

$$\Sigma - (1-\rho)I = \begin{pmatrix} \rho & \rho \\ \rho & \rho \end{pmatrix} \xrightarrow{\frac{1}{\rho}R_1} \begin{pmatrix} 1 & 1 \\ \rho & \rho \end{pmatrix} \xrightarrow{-\rho R_1 + R_2} \begin{pmatrix} 1 & 1 \\ 0 & 0 \end{pmatrix}$$

   de manera que el sistema anterior es equivalente a:

$$\begin{pmatrix} 1 & 1 \\ 0 & 0 \end{pmatrix}\begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix}$$

de donde se tiene que $x = -y$ y las soluciones son de la forma $(r, -r)$ con $r \in \mathbb{R}$. Escogiendo los eigenvectores unitarios $(\frac{1}{\sqrt{2}}, \frac{1}{\sqrt{2}}))^T$ y $(\frac{1}{\sqrt{2}}, -\frac{1}{\sqrt{2}}))^T$ de λ_1 y λ_2 respectivamente, por el teorema espectral tenemos que

$$\Sigma = \Gamma\Lambda\Gamma^T$$

donde $\Gamma = \begin{pmatrix} \frac{1}{\sqrt{2}} & \frac{1}{\sqrt{2}} \\ \frac{1}{\sqrt{2}} & -\frac{1}{\sqrt{2}} \end{pmatrix}$ y $\Lambda = \begin{pmatrix} 1+\rho & 0 \\ 0 & 1-\rho \end{pmatrix}$. De aquí se sigue que los componentes principales de $\mathbf{X} = (X_1, X_2)$ están dados por:

$$\begin{aligned} \mathbf{Y} &= \Gamma^T(\mathbf{X} - \mathbb{E}[\mathbf{X}]) \\ &= \Gamma\mathbf{X} \\ \mathbf{Y} &= \begin{pmatrix} \frac{1}{\sqrt{2}}X_1 + \frac{1}{\sqrt{2}}X_2 \\ \frac{1}{\sqrt{2}}X_1 - \frac{1}{\sqrt{2}}X_2 \end{pmatrix}. \end{aligned}$$

Por lo tanto en este caso la primer componente principal de $\mathbf{X}$ es $Y_1 = \frac{1}{\sqrt{2}}X_1 + \frac{1}{\sqrt{2}}X_2$ y $Y_2 =$ y la segunda es $Y_2 = \frac{1}{\sqrt{2}}X_1 - \frac{1}{\sqrt{2}}X_2$. Por su parte, la proporción de la varianza total $tr\Sigma = tr\Lambda = \lambda_1 + \lambda 2 = 2$ explicada por la primer componente principal es $\frac{Var(Y_1)}{tr\Sigma} = \frac{1}{2}(1+\rho)$ y la proporción explicada por la segunda es $\frac{Var(Y_2)}{tr\Sigma} = \frac{1}{2}(1-\rho)$. El diagrama de codo correspondiente es la recta l_1 entre los puntos $(1, \lambda_1) = (1, 1+\rho)$ y $(2, \lambda_2) = (2, 1-\rho)$ que tiene pendiente $m_1 = \frac{(1-\rho)-(1+\rho)}{2-1} = -2\rho$.
Ahora sea $c > 1$, consideremos el vector (cX_1, X_2). Observemos que $Var(cX_1) = c^2Var(X_1) = c^2$ y $Cov(cX_1, X_2) = c\,Cov(X_1, X_2) = c\rho$ por lo que su matriz de covarianza es $S = \begin{pmatrix} c^2 & c\rho \\ c\rho & 1 \end{pmatrix}$. El polinomio característico de esta matriz es

$$\begin{aligned} q(\mu) &= det(S - \mu I) \\ &= det\left(\begin{pmatrix} c^2-\mu & c\rho \\ c\rho & 1-\mu \end{pmatrix}\right) \\ &= (c^2-\mu)(1-\mu) - c^2\rho^2 \\ &= c^2 - c^2\mu - \mu + \mu^2 - c^2\rho^2 \\ q(\mu) &= \mu^2 - \mu(c^2+1) + c^2 - c^2\rho^2 \end{aligned}$$

que tiene como raíces

$$\begin{aligned} \mu &= \frac{1}{2}(c^2+1) \pm \frac{1}{2}\Big((c^2+1)^2 - 4(c^2 - c^2\rho^2)\Big)^{\frac{1}{2}} \\ &= \frac{1}{2}(c^2+1) \pm \frac{1}{2}\Big((c^2+1)^2 - 4(c^2 - c^2\rho^2)\Big)^{\frac{1}{2}} \\ &= \frac{1}{2}(c^2+1) \pm \frac{1}{2}\Big(c^4 + 2c^2 + 1 - 4c^2 + 4c^2\rho^2\Big)^{\frac{1}{2}} \\ &= \frac{1}{2}(c^2+1) \pm \frac{1}{2}\Big(c^4 - 2c^2 + 1 + 4c^2\rho^2\Big)^{\frac{1}{2}} \\ \lambda &= \frac{1}{2}(c^2+1) \pm \frac{1}{2}\Big((c^2-1)^2 + 4\rho^2c^2\Big)^{\frac{1}{2}}, \end{aligned}$$

esto es,

$$\mu_1 = \frac{1}{2}(c^2+1) + \frac{1}{2}\Big((c^2-1)^2 + 4\rho^2c^2\Big)^{\frac{1}{2}}$$
$$\mu_2 = \frac{1}{2}(c^2+1) - \frac{1}{2}\Big((c^2-1)^2 + 4\rho^2c^2\Big)^{\frac{1}{2}}.$$

Por lo tanto, en este caso la varianza de la primer y segunda componente principal es μ_1 y μ_2 respectivamente. Además el diagrama de codo corresponde a la recta l_2 entre los puntos $(1,\mu_1) = (1, \frac{1}{2}(c^2+1)+\frac{1}{2}\Big((c^2-1)^2+4\rho^2c^2\Big)^{\frac{1}{2}})$ y $(2,\mu_2) = (2, \frac{1}{2}(c^2+1)-\frac{1}{2}\Big((c^2-1)^2+4\rho^2c^2\Big)^{\frac{1}{2}})$ cuya pendiente es $m_2 = \frac{\mu_2-\mu_1}{2-1} = -\Big((c^2-1)^2+4\rho^2c^2\Big)^{\frac{1}{2}}$.
Observemos que como $c > 1$, $4\rho^2c^2 > 4\rho^2$ y $(c^2-1)^2+4\rho^2c^2 > 4\rho^2$ por lo que

$$|m_2| = \sqrt{(c^2-1)^2+4\rho^2c^2} > 2\rho = |m_1|.$$

Esto implica que $m_2 < m_1 < 0$, es decir, la recta l_2 está más inclinada que la recta l_1 o en otras palabras el ángulo entre l_2 y el eje horizontal es mayor al ángulo entre l_1 y este mismo eje. Adicionalmente, $c^2+1 > 2$ por lo que $\frac{1}{2}(c^2+1) > 1$ y $\frac{1}{2}\sqrt{(c^2-1)^2+4\rho^2c^2} > \rho$ de manera que $\mu_1 = \frac{1}{2}(c^2+1) + \frac{1}{2}\sqrt{(c^2-1)^2+4\rho^2c^2} > 1+\rho = \lambda_1$, es decir, el punto $(1,\mu_1)$ está arriba del punto $(1,\lambda_1)$.

2. Sea U una variable aleatoria uniforme en $[0,1]$ y $a = (a_1,a_2,a_3)^T \in \mathbb{R}^3$ un vector de constantes. Consideremos el vector aleatorio $X = (X_1,X_2,X_3)^T = Ua$; sea $\mathbb{E}[X] = \mu$ y Σ su matriz de varianzas y covarianzas con eigenvalores λ_1,λ_2,$\lambda_3 \in \mathbb{R}$, $\lambda_1 \geq \lambda_2 \geq \lambda_3$. También sean v_1,v_2,$v_3 \in \mathbb{R}^3$ eigenvectores estandarizados correspondientes a λ_1,λ_2 y λ_3. Entonces por lo visto en clase, la primer, segunda y tercer componentees principales de X son $Y_i = v_i^T(X-\mu) = v_i^T(Ua) - v_i^T\mu = U(v_i^Ta) - v_i^T\mu$ con $i=1$, $i=2$ e $i=3$ respectivamente. Observemos que v_i^Ta y $v_i^T\mu$ son constantes en $\mathbb{R}$. Luego, dados i,$j \in \{1,2,3\}$ tales que $i \neq j$, si $v_i^Ta \neq 0$ y $v_j^Ta \neq 0$ entonces $Var(Y_i) \neq 0$, $Var(Y_j) \neq 0$ y:

$$\begin{aligned}
\rho(Y_i,Y_j) &= \frac{Cov(Y_i,Y_j)}{\sqrt{Var(Y_i)Var(Y_j)}} \\
&= \frac{Cov(U(v_i^Ta)-v_i^T\mu, U(v_j^Ta)-v_j^T\mu)}{\sqrt{Var(U(v_i^Ta)-v_i^T\mu)Var(U(v_j^Ta)-v_j^T\mu)}} \\
&= \frac{Cov((v_i^Ta)U,(v_j^Ta)U)}{\sqrt{Var((v_i^Ta)U)Var((v_j^Ta)U)}} \\
&= \frac{(v_i^Ta)(v_j^Ta)Cov(U,U)}{\sqrt{\Big((v_i^Ta)(v_j^Ta)Var(U)\Big)^2}} \\
&= \frac{(v_i^Ta)(v_j^Ta)Var(U)}{|(v_i^Ta)(v_j^Ta)Var(U)|} \\
\rho(Y_i,Y_j) &= \pm 1,
\end{aligned}$$

lo cual indica que existe una dependencia lineal directa o inversa entre la i-ésima y la j-esima componente. Esto se puede corroborar en los siguientes diagramas de dispersión:

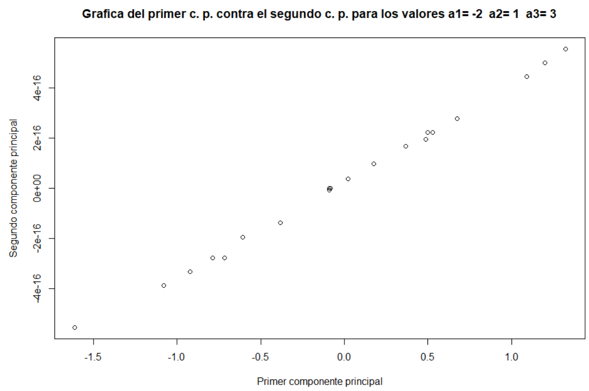


Figura 1: Caption

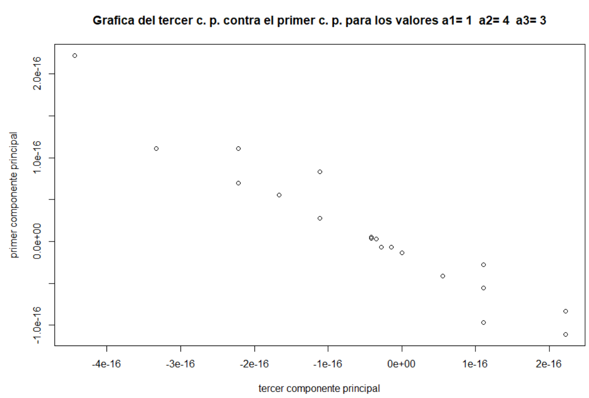


Figura 2: Caption

Si por lo contrario $v_i^T a = 0$ ó $v_j^T a = 0$, entonces una de las componentes es constante y los datos quedan dispersos en una recta de pendiente 0 ó ∞ como en el siguiente diagrama de dispersión:

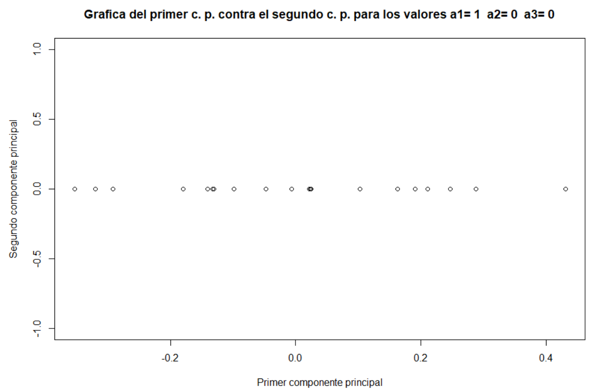


Figura 3: Caption

3. Sea $X = (X_1, X_2, ..., x_n)$ un vector aleatorio con $Var(X) = \Sigma$ y $\mathbb{E}[X] = \mu = (\mu_1, ..., \mu_n)^T$. Sin pérdida de generalidad supongamos que el n-ésimo eigenvalor de Σ, $\lambda_n \in \mathbb{R}$ es muy pequeño. Entonces por lo visto en clase, si $Y_n = \sum_{i=1}^n a_i(X_i - \mu_i) = \sum_{i=1}^n a_i X_i - \sum_{i=1}^n a_i\mu_i$ es la n-ésima componente de X, $Var(Y_n) = \lambda_n$. Luego, por la desigualdad de Chebyshev para todo $\epsilon > 0$ tenemos que

$$P(|Y - \mathbb{E}[Y]| \geq \epsilon) = P(|Y| \geq \epsilon) = P(|\sum_{i=1}^n a_i X_i - \sum_{i=1}^n a_i \mu_i| \geq \epsilon) \leqslant \frac{Var(Y_i)}{\epsilon^2} = \frac{\lambda_n}{\epsilon^2}$$

o equivalentemente

$$P(|\sum_{i=1}^n a_i X_i - \sum_{i=1}^n a_i \mu_i| < \epsilon) \geq 1 - \frac{\lambda_n}{\epsilon^2}\,.$$

Como por hipótesis λ_n es muy pequeño podríamos considerar que para valores de ϵ muy pequeños, por ejemplo $\epsilon_0 \approx \frac{1}{1000}$, $\frac{\lambda_n}{\epsilon_0^2} \leqslant \frac{1}{10}$ de modo que $1 - \frac{\lambda_n}{\epsilon_0^2} \geq 0.9$ y $0.9 \leqslant P(|\sum_{i=1}^n a_i X_i - \sum_{i=1}^n a_i \mu_i| < \epsilon_0) \leqslant 1$. Pero observemos que tomando $j \in \{1, ..., n\}$ fija

$$\begin{aligned}
P(|\sum_{i=1}^n a_i X_i - \sum_{i=1}^n a_i \mu_i| < \epsilon_0) &= P(-\epsilon_0 < \sum_{i=1}^n a_i X_i - \sum_{i=1}^n a_i \mu_i < \epsilon_0) \\
&= P(\sum_{i=1}^n a_i \mu_i - \epsilon_0 < \sum_{i=1}^n a_i X_i < \sum_{i=1}^n a_i \mu_i + \epsilon_0) \\
&= P(-\sum_{\substack{i=1\\ i\neq j}}^n a_i X_i + \sum_{i=1}^n a_i \mu_i - \epsilon_0 < a_j X_j < -\sum_{\substack{i=1\\ i\neq j}}^n a_i X_i + \sum_{i=1}^n a_i \mu_i + \epsilon_0) \\
&= P(\sum_{\substack{i=1\\ i\neq j}}^n \Big(-\frac{a_i}{a_j}\Big) X_i + \frac{\sum_{i=1}^n a_i \mu_i}{a_j} - \frac{\epsilon_0}{a_j} < X_j < \sum_{\substack{i=1\\ i\neq j}}^n \Big(-\frac{a_i}{a_j}\Big) X_i + \frac{\sum_{i=1}^n a_i \mu_i}{a_j} + \\
&= P(\sum_{\substack{i=1\\ i\neq j}}^n m_i X_i + b - \delta < X_j < \sum_{\substack{i=1\\ i\neq j}}^n m_i X_i + b + \delta)
\end{aligned}$$

$$P(|\sum_{i=1}^n a_i X_i - \sum_{i=1}^n a_i \mu_i| < \epsilon_0) = P(|X_j - (\sum_{\substack{i=1\\ i\neq j}}^n m_i X_i + b)| < \delta)\,,$$

suponiendo que $a_j > 0$ en la cuarta igualdad (el caso en que $a_j < 0$ es análogo) y haciendo $m_i = -\frac{a_i}{a_j}$ $\forall i \in \{1, ..., n\}$, $b = \frac{\sum_{i=1}^{n} a_i \mu_i}{a_j}$ y $\delta = \frac{\epsilon_0}{a_j}$ en la quinta igualdad. Esto significa que dada una observación $\vec{x} = (x_1, ..., x_n) \in \mathbb{R}^n$ del vector X es muy probable que se cumpla la relación $|x_j - (\sum_{\substack{i=1 \\ i \neq j}}^{n} m_i x_i + b)| < \delta$; de hecho, si se tienen varias observaciones de X se verá que en la mayoría de los casos se cumple. Por lo tanto, como δ es próximo a 0 (por ejemplo si $a_j > 1$, $\delta < 10^{-3}$) se sigue que en general $\sum_{\substack{i=1 \\ i \neq j}}^{n} m_i X_i + b$ aproxima a X_j con un error menor a δ lo cual implica que existe una relación casi lineal entre las variables.

Ejercicio 5

Primero carguemos los datos

```
data <- read.delim("comida_francesa.txt")
```

Para elegir el número de factores, haremos un breve análisis de componentes principales y el número de factores que elegiremos serán el número de componentes principales que acumulen más del 85% de variabilidad (elección arbitraria).

Como nota, usaremos la función `prcomp()` en lugar de `princomp()` ya que la segunda utiliza el factor $1/n$ en lugar de $1/(n-1)$ para calcular el estimador de la matriz de correlaciones, lo que lo hace menos compatible con el resto de funciones de R[1].

```
pc <- prcomp(data[, 2:ncol(data)], center = T, scale = T)
summary(pc)
```

```
## Importance of components:
##                          PC1    PC2     PC3     PC4     PC5     PC6     PC7
## Standard deviation     2.082 1.3524 0.79560 0.35768 0.23903 0.13631 0.03151
## Proportion of Variance 0.619 0.2613 0.09043 0.01828 0.00816 0.00265 0.00014
## Cumulative Proportion  0.619 0.8803 0.97077 0.98904 0.99720 0.99986 1.00000
```

Como con tres componentes principales acumulamos el 88% de la varianza, usaremos dos factores para nuestro análisis.

```
fa_2 <- factanal(data[, 2:ncol(data)], factors = 2)
```

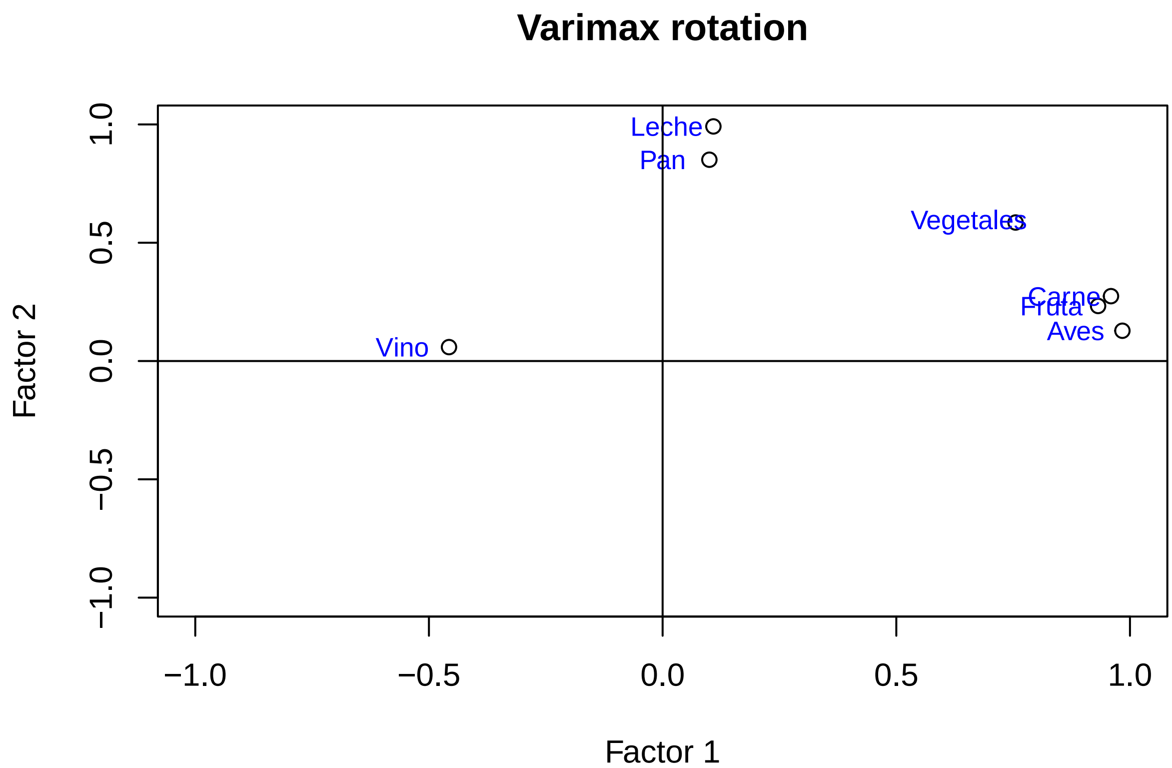
Interpretación

Grafiquemos los factores e interpretemos.

```
plot(fa_2$loadings[,1],
     fa_2$loadings[,2],
     xlab = "Factor 1",
     ylab = "Factor 2",
     ylim = c(-1,1),
     xlim = c(-1,1),
     main = "Varimax rotation")

text(fa_2$loadings[,1]-0.1,
     fa_2$loadings[,2],
      colnames(data[, 2:length(data)]),
      col="blue",
     cex = 0.8)
abline(h = 0, v = 0)
```

[1] 1




```
fa_2
```

```
##
## Call:
## factanal(x = data[, 2:ncol(data)], factors = 2)
##
## Uniquenesses:
##       Pan Vegetales     Fruta     Carne      Aves     Leche      Vino
##     0.266     0.087     0.078     0.005     0.016     0.005     0.788
##
## Loadings:
##           Factor1 Factor2
## Pan        0.100   0.851
## Vegetales  0.755   0.585
## Fruta      0.932   0.233
## Carne      0.959   0.274
## Aves       0.984   0.128
## Leche      0.109   0.992
## Vino      -0.457
##
##                Factor1 Factor2
## SS loadings      3.557   2.199
## Proportion Var   0.508   0.314
## Cumulative Var   0.508   0.822
##
## Test of the hypothesis that 2 factors are sufficient.
```

```
## The chi square statistic is 23.47 on 8 degrees of freedom.
## The p-value is 0.00282
```

Como podemos observar, el ruido (Uniqueness) de las variables es relativamente bajo para todas, salvo por el vino que es casi del 0.8. Lo que en general nos indica que no es un mal modelo de entrada ya que nuestros factores pueden representar las muestras sin la necesidad de corregir con un error grande (la matriz diagonal Ψ en la fórmula $\Sigma = \Lambda\Lambda^T + \Psi$).

Otra señal de que no tenemos un mal modelo es que podemos observar en la consola de R que la varianza acumulada (Cumulative Var) de nuestro modelo es del 82.2%, por lo que a ojo es una buena representación o resumen del total de los datos.

Finalmente, analicemos las cargas para nuestros factores habiendo usado la rotación varimax (que es la estándar en la función `factanal()`):

- Para el primer factor observamos que:
  1. Los gastos en vegetales, fruta, carne y aves componen principalmente este factor. Por lo que en este factor podríamos decir que se compacta la información al respecto de los gastos en alimentos ricos en proteína ó fibra.
  2. El pan y la leche son casi ignorados por el factor.
  3. El vino juega un papel casi importante, pero en comparación de los alimentos ricos en proteína o fibra es bastante menor, por lo que tampoco lo tomaríamos en cuenta.
- Para el segundo factor tenemos:
  1. El pan y la leche son las principales variables que componen a este factor. Entonces este factor nos compacta información al respecto de cuánto dinero gastan las familias en pan y en leche.
  2. Salvo los vegetales, todas las demás variables son bastante ignoradas en este factor.
  3. Los vegetales, al igual que el vino para el factor anterior, según criterios podrían jugar un papel relevante en la información dada por este factor. Nosotros la ignoraremos y concentraremos la interpretación en gastos en pan y leche.

Rotación

Para nuestro análisis anterior, usamos la rotación *varimax* que es la que viene por defecto en la función `factana()`, pero hagamos el contraste con otros dos análisis que usen otras rotaciones. Usemos en un análisis la rotación *promax* y en otro no hagamos ninguna rotación y veamos si hay alguna mejoría de la interpretación.

Esperaríamos encontrar un modelo con cargas o muy cercanas a 0, o muy cercanas a 1. Usando *varimax* podríamos decir que es un buen modelo ya que sólo teníamos una carga por factor que estaba cerca del 0.5 (el vino para el factor 1 y los vegetales para el factor 2), pero veamos si alguna de estas rotaciones mejora la interpretación de la información.

```r
fa_2.none <- factanal(data[, 2:ncol(data)], factors = 2, rotation = "none")
fa_2.promax <- factanal(data[, 2:ncol(data)], factors = 2, rotation = "promax")
fa_2
```

```
##
## Call:
## factanal(x = data[, 2:ncol(data)], factors = 2)
##
## Uniquenesses:
##       Pan Vegetales     Fruta     Carne      Aves     Leche      Vino
##     0.266     0.087     0.078     0.005     0.016     0.005     0.788
##
## Loadings:
##           Factor1 Factor2
## Pan        0.100   0.851
## Vegetales  0.755   0.585
```

```
## Fruta      0.932   0.233
## Carne      0.959   0.274
## Aves       0.984   0.128
## Leche      0.109   0.992
## Vino      -0.457
##
##                Factor1 Factor2
## SS loadings      3.557   2.199
## Proportion Var   0.508   0.314
## Cumulative Var   0.508   0.822
##
## Test of the hypothesis that 2 factors are sufficient.
## The chi square statistic is 23.47 on 8 degrees of freedom.
## The p-value is 0.00282
```

```
fa_2.promax
```

```
##
## Call:
## factanal(x = data[, 2:ncol(data)], factors = 2, rotation = "promax")
##
## Uniquenesses:
##       Pan Vegetales     Fruta     Carne      Aves     Leche      Vino
##     0.266     0.087     0.078     0.005     0.016     0.005     0.788
##
## Loadings:
##           Factor1 Factor2
## Pan                0.866
## Vegetales  0.717   0.416
## Fruta      0.960
## Carne      0.983
## Aves       1.032  -0.123
## Leche              1.011
## Vino      -0.498   0.181
##
##                Factor1 Factor2
## SS loadings      3.718   1.994
## Proportion Var   0.531   0.285
## Cumulative Var   0.531   0.816
##
## Factor Correlations:
##         Factor1 Factor2
## Factor1   1.000  -0.378
## Factor2  -0.378   1.000
##
## Test of the hypothesis that 2 factors are sufficient.
## The chi square statistic is 23.47 on 8 degrees of freedom.
## The p-value is 0.00282
```

- Para el caso en el que no hacemos rotación tenemos:

```
fa_2.none$loadings[,1] # Cargas del primer factor
```

```
##        Pan  Vegetales      Fruta      Carne       Aves      Leche       Vino
##  0.6017038  0.9555287  0.8781203  0.9253912  0.8548299  0.6949917 -0.3244171
```

```r
fa_2.none$loadings[,2] # Cargas del segundo factor
```

```
##          Pan    Vegetales        Fruta        Carne         Aves        Leche 
##  0.609806864 -0.002118175 -0.388920828 -0.373125523 -0.503448531  0.715579085 
##         Vino 
##  0.327289904
```

1. En el primer factor tenemos bastantes atributos cercanos a 1, que son vegetales, fruta, carne y aves. Pero desafortunadamente tenemos muchas más variables con carga cercana a 0.5 (tres atributos, a saber pan, leche y vino).

2. Mientras que en el segundo factor ocurre algo peor, ya que ni siquiera tenemos cargas cernanas a 1, sólo la de vegetales que es cercana a 0. Todas las demás están cerca del 0.5 y no nos ayudan a una buena y condensada interpretación de los datos.

3. Podemos concluir que esta rotación no es mejor que la rotación *varimax*, por mucho.

- Para la rotación promax tenemos:

```r
fa_2.promax$loadings[,1] # Cargas del primer factor
```

```
##         Pan   Vegetales       Fruta       Carne        Aves       Leche 
## -0.02466404  0.71684006  0.96021163  0.98328858  1.03199363 -0.03723432 
##        Vino 
## -0.49776922
```

```r
fa_2.promax$loadings[,2] # Cargas del segundo factor
```

```
##           Pan     Vegetales         Fruta         Carne          Aves 
##  0.8657128472  0.4163619554  0.0004790853  0.0367788886 -0.1228212636 
##         Leche          Vino 
##  1.0110207218  0.1811405475
```

1. Curiosamente, para ambos factores sucedió lo mismo, y es que son bastante análogos a los factores de cuando usamos la rotación *varimax*.

Por buscar diferencias, para el primer factor la carga a los vegetales es menor que en la varimax, pero a cambio de eso, la carga a las variables que no aportaban al factor son aún menores (las de pan y leche). Mientras que para el segundo factor podríamos decir que este sí es un poco mejor (muy poco a decir verdad) que su análogo en la rotación varimax; comparemos las cargas

```r
fa_2$loadings[,2]        # Cargas del segundo factor con rotación varimax
```

```
##        Pan  Vegetales      Fruta      Carne       Aves      Leche       Vino 
## 0.85082428 0.58542890 0.23268939 0.27419588 0.12801962 0.99159433 0.05889403
```

```r
fa_2.promax$loadings[,2] # Cargas del segundo factor con rotación promax
```

```
##           Pan     Vegetales         Fruta         Carne          Aves 
##  0.8657128472  0.4163619554  0.0004790853  0.0367788886 -0.1228212636 
##         Leche          Vino 
##  1.0110207218  0.1811405475
```

Podemos ver que en general las cargas que aportaban interpretación se hicieron un poco más pequeñas o un poco más grandes (según si eran más cercanas a 0 o a 1 respectivamente), y la que no aportaba interpretación (vegetales), bajó su carga considerablemente haciendo que afecte un poco menos en la interpretación.

2. Podemos concluir que no hay grandes diferencias entre estas dos rotaciones, y es bastante subjetivo quedarse con una o con la otra. Si nos preguntan a nosotros, quizás optaríamos por esta, la rotada con *promax* porque en general hace un poco más extremos los valores a 0 y algunos cercanos a 1 los hace más grandes.

Podemos concluir que a grandes rasgos no, las rotaciones que usamos en este documento no encontraron una diferencia significativa para la interpretación de los datos. Sin embargo, por preferencias subjetivas podríamos elegir que la rotación promax mejoró un poco la interpretación de los datos, pero siguen teniendo exactamente la misma relevancia e interpretación todos los atributos en las cargas de ambos factores para las dos rotaciones.

[1] https://stats.stackexchange.com/questions/242864/princomp-outputs-seemingly-wrong-pca-scores-with-cor-true-input-argument

[2] Para la realización de este ejercicio fue de mucha ayuda el siguiente artículo

https://www.geo.fu-berlin.de/en/v/soga/Geodata-analysis/factor-analysis/A-simple-example-of-FA/index.html