Nombre: Pedro Ramos Suárez.

1. **Sea A una matriz aleatoria con distribución $\mathbf{W_p(n, I_p)}$. Probar que $\mathrm{tr}(\mathbf{A})$. tiene distribución $\chi^2_{\mathbf{np}}$.**

El resultado 1 sobre marginalización asegura que:

$$\frac{a_{ii}}{\sigma_i^2} \sim \chi^2_n, \quad i = 1, \dots, p$$

y como $\sigma_i^2 = 1, \forall i = 1, \dots, p$ (por ser $\Sigma = I_p$), se tiene que

$$a_{ii} \sim \chi^2_n, \forall i = 1, \dots, p$$

.

Por otro lado, como $A \sim W_p(n, I_p)$ será, por definición, $A \overset{d}{=} \sum_{\alpha=1}^n Z_\alpha Z'_\alpha$, con $Z_\alpha \sim N_p(0, I_p)$ independientes, $\forall \alpha = 1, \dots, n \quad (n \geq p)$. Para cada α, podemos escribir $Z_\alpha = \begin{pmatrix} Z_{\alpha_1} \\ \vdots \\ Z_{\alpha_p} \end{pmatrix}$. Como I_p es una matriz diagonal, las variables aleatorias $Z_{\alpha_1}, \dots, Z_{\alpha_p}$ componentes del vector Z_α son independientes, $\forall \alpha = 1, \dots, n$ (pues en el caso de una DNM incorrelación implica independencia), con $Z_{\alpha_i} \sim N(0,1), i = 1, \dots, p$. Puesto que

$$a_{ii} \overset{d}{=} \sum_{\alpha=1}^n Z_{\alpha_i} Z'_{\alpha_i}, \quad i = 1, \dots, p$$

es claro que $a_{11}, \dots, a_{pp}$ son independientes. Aplicando finalmente la propiedad de reproductividad de la distribución χ^2 llegamos al resultado buscado:

$$\mathrm{tr}(A) = \sum_{i=1}^p a_{ii} \sim \chi^2_{np}$$

donde hemos tenido en cuenta que $\sum_{i=1}^p n = np$.

2. Sea A una matriz aleatoria con distribución $\mathbf{W_p(n, \Sigma)}$ $(\mathbf{\Sigma > 0})$. Probar que, para cualesquiera vectores $\mathbf{a, b \in \mathbb{R}^p}$, las variables aleatorias $\mathbf{a'Aa}$ y $\mathbf{b'Ab}$ son independientes si y sólo si $\mathbf{a'\Sigma b = 0}$.

- a y b linealmente independientes: Consideramos la matriz $M = \begin{pmatrix} a & b \end{pmatrix}_{p\times 2}$, de rango 2. Tenemos:

$$M'AM = \begin{pmatrix} a' \\ b' \end{pmatrix} A \begin{pmatrix} a & b \end{pmatrix} = \begin{pmatrix} a'Aa & a'Ab \\ b'Aa & b'Ab \end{pmatrix}_{4\times 4}$$

  y, por la propiedad de reproductividad bajo transformaciones lineales rectangulares de la distribución de Wishart, se cumple que

$$M'AM \sim W_2(n, M'\Sigma M) \equiv W_2(n, \begin{pmatrix} a'\Sigma a & a'\Sigma b \\ b'\Sigma a & b'\Sigma b \end{pmatrix})$$

  $\Leftarrow$ Si $a'\Sigma b = 0$, el resultado sobre marginalización por bloques en una distribución de Wishart nos asegura entonces que $a'Aa$ y $b'Ab$ son independientes.

  $\Rightarrow$ Si $a'Aa$ y $b'Ab$ son independientes, se tiene que:

$$0 = Cov(a'Aa, b'Ab) = n((a'\Sigma b)^2 + (a'\Sigma b)^2) = 2n(a'\Sigma b)^2$$

  Luego ha de ser $a'\Sigma b = 0$.

- a y b linealmente dependiente: Existirá entonces $\alpha \in \mathbb{R}$ tal que $a = \alpha b$, luego

$$a'Aa = (\alpha b')A\alpha b = \alpha^2(b'Ab)$$

  esto es, las variables $a'Aa$ y $b'Ab$ no son independientes. Por otro lado, será $a'\Sigma b = \alpha(b'\Sigma b) \neq 0$ ya que Σ es definida positva. Así pues, el contrarrecíproco se cumple trivialmente.

- $a = 0$ ó $b = 0$: Entonces $a'Aa = 0$ ó $b'Ab = 0$, es decir, una de las variables es degenerada, con $P[a'Aa = 0] = 1$ ó $P[b'Ab = 0] = 1$, y una variable aleatoria degenerada es independiente de cualquier otra. Además, será $a'\Sigma b = 0$, de manera que el resultado se cumple trivialmente.

3. Probar que si $\mathbf{X_1, \ldots, X_N}$ constituyen una muestra aleatoria simple de una distribución $\mathbf{N_p}(\mu, \mathbf{\Sigma})$ $(\mathbf{\Sigma > 0})$, y el parámetro vector de medias μ es conocido, entonces el estimador máximo-verosímil de $\mathbf{\Sigma}$ es:

$$\hat{\mathbf{\Sigma}} := \frac{\mathbf{1}}{\mathbf{N}} \sum_{\alpha=\mathbf{1}}^{\mathbf{N}} (\mathbf{X}_\alpha - \mu)(\mathbf{X}_\alpha - \mu)'$$

(bajo la condición de ser esta matriz definida positiva). Comprobar si este estimador es o no insesgado.

Vamos a buscar el máximo en Σ de la función $\ln(L(\mu, \Sigma))$, que será el mismo que el de $L(\mu, \Sigma)$ por ser el logaritmo estrictamente creciente.

$$\ln(L(\mu, \Sigma)) = -\frac{pN}{2} \ln(2\pi) - \frac{N}{2} \ln(|\Sigma|) - \frac{1}{2} \sum_{\alpha=1}^{N} (X_\alpha - \mu)' \Sigma^{-1} (X_\alpha - \mu)$$

donde $\sum_{\alpha=1}^{N} (X_\alpha - \mu)' \Sigma^{-1} (X_\alpha - \mu)$ es un escalar, de modo que coincide con su traza:

$$\sum_{\alpha=1}^{N} (X_\alpha - \mu)' \Sigma^{-1} (X_\alpha - \mu) = \operatorname{tr}(\sum_{\alpha=1}^{N} (X_\alpha - \mu)' \Sigma^{-1} (X_\alpha - \mu)) = \sum_{\alpha=1}^{N} \operatorname{tr}(X_\alpha - \mu)' \Sigma^{-1} (X_\alpha - \mu)) =$$

$$= \sum_{\alpha=1}^{N} \operatorname{tr}(\Sigma^{-1} (X_\alpha - \mu)(X_\alpha - \mu)') = \operatorname{tr}(\sum_{\alpha=1}^{N} (\Sigma^{-1} (X_\alpha - \mu)(X_\alpha - \mu)') =$$

$$= \operatorname{tr}(\Sigma^{-1} \sum_{\alpha=1}^{N} (X_\alpha - \mu)(X_\alpha - \mu)')$$

Así, llegamos a

$$\ln(L(\mu, \Sigma)) = -\frac{pN}{2} \ln(2\pi) - \frac{N}{2} \ln(|\Sigma|) - \frac{1}{2} \operatorname{tr}(\Sigma^{-1} \sum_{\alpha=1}^{N} (X_\alpha - \mu)(X_\alpha - \mu)')$$

Consideramos ahora la función

$$f(\Sigma) = 2[\ln(L(\mu, \Sigma)) + \frac{pN}{2} \ln(2\pi)] = -N \ln(|\Sigma|) - \operatorname{tr}(\Sigma^{-1} \sum_{\alpha=1}^{N} (X_\alpha - \mu)(X_\alpha - \mu)')$$

y aplicamos el lema de Watson con $G = \Sigma$ y $D = \sum_{\alpha=1}^{N} (X_\alpha - \mu)(X_\alpha - \mu)'$, ambas matrices simétricas y definidas positivas, para obtener que f alcanza el único máximo respecto a Σ en $\Sigma = \frac{1}{N} \sum_{\alpha=1}^{N} (X_\alpha - \mu)(X_\alpha - \mu)'$. Es claro que f y $\ln(L(\mu, \Sigma))$ alcanzan el máximo para Σ en el mismo punto, de donde deducimos entonces que $\hat{\Sigma} = \frac{1}{N} \sum_{\alpha=1}^{N} (X_\alpha - \mu)(X_\alpha - \mu)'$ es el estimador máximo versímil de Σ.

Comprobamos que es insesgado:

$$E[\hat{\Sigma}] = E[\frac{1}{N} \sum_{\alpha=1}^{N} (X_\alpha - \mu)(X_\alpha - \mu)'] = \frac{1}{N} \sum_{\alpha=1}^{N} E[(X_\alpha - \mu)(X_\alpha - \mu)'] = \frac{1}{N} \sum_{\alpha=1}^{N} \Sigma = \Sigma$$

luego $\hat{\Sigma}$ es un estimador insesgado.

4. Sea A una matriz aleatoria con distribución $\mathbf{W_p(n, \Sigma)}$ $(\mathbf{\Sigma > 0})$. Probar usando la función de densidad de Wishart, que

$$\mathbf{E[|A|^r] = |\Sigma|^r 2^{pr} \frac{\Gamma_p(\frac{1}{2}n + r)}{\Gamma_p(\frac{1}{2}n)}}, \quad \forall \mathbf{r > 0}$$

[OBSERVACIÓN: La definición de la 'densidad de Wishart', y de la correspondiente 'función característica de Wishart', sigue siendo válida, por extensión, tomando como parámetro 'grados de libertad' cualquier número real n tal que $\mathbf{n > p - 1}$ (aunque la definición implícita en términos de vectores $\mathbf{Z}_\alpha \sim \mathbf{N_p(0, \Sigma)}$ $(\mathbf{\Sigma > 0})$ independientes sólo se aplicaría para el caso en que n es entero, con $\mathbf{n \geq p}$)]

Sea $\mathcal{M}_p$ el espacio de las matrices reales, simétricas y definidas positivas de orden p. Entonces:

$$E[|A|^r] = \int_{\mathcal{M}_p} |A|^r \frac{|A|^{\frac{n-p-1}{2}} \exp\{-\frac{1}{2}\operatorname{tr}(\Sigma^{-1}A)\}}{2^{\frac{np}{2}} |\Sigma|^{\frac{n}{2}} \Gamma_p(\frac{n}{2})} dA =$$

$$= \int_{\mathcal{M}_p} \frac{|A|^{\frac{n+2r-p-1}{2}} \exp\{-\frac{1}{2}\operatorname{tr}(\Sigma^{-1}A)\}}{2^{\frac{np}{2}} |\Sigma|^{\frac{n}{2}} \Gamma_p(\frac{n}{2})} \cdot \frac{2^{pr}\Gamma_p(\frac{n+2r}{2})|\Sigma|^r}{2^{pr}\Gamma_p(\frac{n+2r}{2})|\Sigma|^r} dA =$$

$$= \int_{\mathcal{M}_p} \frac{|A|^{\frac{n+2r-p-1}{2}} \exp\{-\frac{1}{2}\operatorname{tr}(\Sigma^{-1}A)\}}{2^{\frac{p(n+2r)}{2}} |\Sigma|^{\frac{n+2r}{2}} \Gamma_p(\frac{n+2r}{2})} \cdot \frac{2^{pr}\Gamma_p(\frac{n+2r}{2})|\Sigma|^r}{\Gamma_p(\frac{n}{2})} dA =$$

$$= \frac{2^{pr}\Gamma_p(\frac{n+2r}{2})|\Sigma|^r}{\Gamma_p(\frac{n}{2})} \cdot \int_{\mathcal{M}_p} \frac{|A|^{\frac{n+2r-p-1}{1}} \exp\{-\frac{1}{2}\operatorname{tr}(\Sigma^{-1}A)\}}{2^{\frac{p(n+2r)}{2}} |\Sigma|^{\frac{n+2r}{2}} \Gamma_p(\frac{n+2r}{2})} dA$$

Usando que:

$$f(A) = \frac{|A|^{\frac{n+2r-p-1}{2}} \exp\{-\frac{1}{2}\operatorname{tr}(\Sigma^{-1}A)\}}{2^{\frac{p(n+2r)}{2}} |\Sigma|^{\frac{n+2r}{2}} \Gamma_p(\frac{n+2r}{2})}$$

es la función de distribución de Wishart con $n + 2r$ grados de libertad ($W_p(n + 2r, \Sigma)$), tenemos:

$$\int_{\mathcal{M}_p} \frac{|A|^{\frac{n+2r-p-1}{1}} \exp\{-\frac{1}{2}\operatorname{tr}(\Sigma^{-1}A)\}}{2^{\frac{p(n+2r)}{2}} |\Sigma|^{\frac{n+2r}{2}} \Gamma_p(\frac{n+2r}{2})} dA = 1$$

y por lo tanto:

$$E[|A|^r] = \frac{|\Sigma|^r 2^{pr}\Gamma_p(\frac{n+2r}{2})}{\Gamma_p(\frac{n}{2})} = |\Sigma|^r 2^{pr} \frac{\Gamma_p(\frac{1}{2}n + r)}{\Gamma_p(\frac{1}{2}n)}, \quad \forall r > 0$$

5. Sea $\mathbf{X_1,\ldots,X_N}$ una muestra aleatoria simple de una distribución $\mathbf{N_p}(\mu,\boldsymbol{\Sigma})$ $(\boldsymbol{\Sigma}>\mathbf{0})$. Sea $\mathbf{A}=\sum_{\alpha=1}^{\mathbf{N}}(\mathbf{X}_\alpha-\bar{\mathbf{X}})(\mathbf{X}_\alpha-\bar{\mathbf{X}})'$ la matriz de dispersiones muestral, con $\bar{\mathbf{X}}=\frac{\mathbf{1}}{\mathbf{N}}\sum_{\alpha=1}^{\mathbf{N}}\mathbf{X}_\alpha$ el vector de medias muestral.

(a) Probar que la matriz $\mathrm{Cov}(\bar{\mathbf{X}},\mathbf{X}_\alpha-\bar{\mathbf{X}})$ es nula, para $\alpha=\mathbf{1},\ldots,\mathbf{N}$. Deducir, entonces, que $\bar{\mathbf{X}}$ y $\mathbf{A}$ son independientes.

Sea $\begin{pmatrix}\bar{X}\\ X_\alpha-\bar{X}\end{pmatrix}$, $\alpha=1,\ldots,N$. Entonces:

$$\begin{pmatrix}\bar{X}\\ X_\alpha-\bar{X}\end{pmatrix}=\begin{pmatrix}\frac{1}{N}I_p & \ldots & \frac{1}{N}I_p & \ldots & \frac{1}{N}I_p\\ -\frac{1}{N}I_p & \ldots & \frac{N-1}{N}I_p & \ldots & -\frac{1}{N}I_p\end{pmatrix}\begin{pmatrix}X_1\\ \vdots\\ X_\alpha\\ \vdots\\ X_N\end{pmatrix}=M\begin{pmatrix}X_1\\ \vdots\\ X_\alpha\\ \vdots\\ X_N\end{pmatrix}$$

Notemos que $\begin{pmatrix}X_1\\ \vdots\\ X_N\end{pmatrix}\sim N_{pN}(\bar{\mu},\bar{\Sigma})$, con

$$\bar{\mu}=\begin{pmatrix}\mu\\ \vdots\\ \mu\end{pmatrix}\quad \text{y}\quad \bar{\Sigma}=\begin{pmatrix}\Sigma & 0 & \ldots & 0\\ 0 & \Sigma & \ldots & 0\\ \vdots & \vdots & \ddots & \vdots\\ 0 & 0 & \ldots & \Sigma\end{pmatrix}=\mathrm{diag}(\Sigma,\ldots,\Sigma)$$

por se los vectores $X_\alpha\sim N_p(\mu,\Sigma)$ independientes $\forall\alpha=1,\ldots,N$.

Por el resultado sobre transformaciones lineales de rango máximo de una DNM tenemos que $\begin{pmatrix}\bar{X}\\ X_\alpha-\bar{X}\end{pmatrix}\sim N_{2p}(M\bar{\mu},M\bar{\Sigma}M'),\forall\alpha=1,\ldots,N$, con:

$$M\bar{\Sigma}M'=\begin{pmatrix}\frac{1}{N}I_p & \ldots & \frac{1}{N}I_p & \ldots & \frac{1}{N}I_p\\ -\frac{1}{N}I_p & \ldots & \frac{N-1}{N}I_p & \ldots & -\frac{1}{N}I_p\end{pmatrix}\begin{pmatrix}\Sigma & 0 & \ldots & 0\\ 0 & \Sigma & \ldots & 0\\ \vdots & \vdots & \ddots & \vdots\\ 0 & 0 & \ddots & \Sigma\end{pmatrix}\begin{pmatrix}\frac{1}{N}I_p & -\frac{1}{N}I_p\\ \vdots & \vdots\\ \frac{1}{N}I_p & \frac{N-1}{N}I_p\\ \vdots & \vdots\\ \frac{1}{N}I_p & -\frac{1}{N}I_p\end{pmatrix}=$$

$$=\begin{pmatrix}\frac{1}{N}\Sigma & \ldots & \frac{1}{N}\Sigma & \ldots & \frac{1}{N}\Sigma\\ -\frac{1}{N}\Sigma & \ldots & \frac{N-1}{N}\Sigma & \ldots & -\frac{1}{N}\Sigma\end{pmatrix}\begin{pmatrix}\frac{1}{N}I_p & -\frac{1}{N}I_p\\ \vdots & \vdots\\ \frac{1}{N}I_p & \frac{N-1}{N}I_p\\ \vdots & \vdots\\ \frac{1}{N}I_p & -\frac{1}{N}I_p\end{pmatrix}=\begin{pmatrix}\frac{1}{N}\Sigma & 0_{p\times p}\\ 0_{p\times p0} & \frac{N-1}{N}\Sigma\end{pmatrix}$$

Tenemos entonces que $Cov(\bar{X},X_\alpha-\bar{X})=0_{p\times p}$ y, además, como en el caso de una DNM la incorrelación implica independencia, los vectores aleatorios $\bar{X}$ y $X_\alpha-\bar{X}$ son independientes $\forall\alpha=1,\ldots,N$.

Finalmente, al ser A una función medible de los vectores $X_\alpha-\bar{X}$, tenemos que A y $\bar{X}$ son independientes.

(b) Encontrar una matriz B tal que $\mathbf{A} = \mathbf{X}'\mathbf{B}\mathbf{X}$, con $\mathbf{X} = (\mathbf{X_1}, \ldots, \mathbf{X_N})'$ (matriz $(\mathbf{N} \times \mathbf{p})-$dimensional).

$$A = \sum_{\alpha=1}^{N} (X_\alpha - \bar{X})(X_\alpha - \bar{X})' = \begin{pmatrix} X_1 - \bar{X} & \ldots & X_N - \bar{X} \end{pmatrix} \begin{pmatrix} X_1' - \bar{X}' \\ \vdots \\ X_N' - \bar{X}' \end{pmatrix} =$$

$$= [\begin{pmatrix} X_1 & \ldots & X_N \end{pmatrix} - \begin{pmatrix} \bar{X} & \ldots & \bar{X} \end{pmatrix}][\begin{pmatrix} X_1' \\ \vdots \\ X_N' \end{pmatrix} - \begin{pmatrix} \bar{X}' \\ \vdots \\ \bar{X}' \end{pmatrix}] =$$

$$= [\begin{pmatrix} X_1 & \ldots & X_N \end{pmatrix} - \frac{1}{N} \begin{pmatrix} \sum_{\alpha=1}^{N} X_\alpha & \ldots & \sum_{\alpha=1}^{N} X_\alpha \end{pmatrix}][\begin{pmatrix} X_1' \\ \vdots \\ X_N' \end{pmatrix} - \frac{1}{N} \begin{pmatrix} \sum_{\alpha=1}^{N} X_\alpha' \\ \vdots \\ \sum_{\alpha=1}^{N} X_\alpha' \end{pmatrix}] =$$

$$= [\begin{pmatrix} X_1 & \ldots & X_N \end{pmatrix} (I_N - \frac{1}{N} \begin{pmatrix} 1 & \ldots & 1 \\ \vdots & \ddots & \vdots \\ 1 & \ldots & 1 \end{pmatrix})][(I_N - \frac{1}{N} \begin{pmatrix} 1 & \ldots & 1 \\ \vdots & \ddots & \vdots \\ 1 & \ldots & 1 \end{pmatrix}) \begin{pmatrix} X_1' \\ \vdots \\ X_N' \end{pmatrix}] =$$

$$= X'BX$$

donde

$$B = (I_N - \frac{1}{N} \begin{pmatrix} 1 & \ldots & 1 \\ \vdots & \ddots & \vdots \\ 1 & \ldots & 1 \end{pmatrix})(I_N - \frac{1}{N} \begin{pmatrix} 1 & \ldots & 1 \\ \vdots & \ddots & \vdots \\ 1 & \ldots & 1 \end{pmatrix}) =$$

$$= I_N - \frac{1}{N} \begin{pmatrix} 1 & \ldots & 1 \\ \vdots & \ddots & \vdots \\ 1 & \ldots & 1 \end{pmatrix} = \begin{pmatrix} \frac{N-1}{N} & -\frac{1}{N} & \ldots & -\frac{1}{N} \\ -\frac{1}{N} & \frac{N-1}{N} & \ldots & -\frac{1}{N} \\ \vdots & \vdots & \ddots & \vdots \\ -\frac{1}{N} & -\frac{1}{N} & \ldots & \frac{N-1}{N} \end{pmatrix}$$

(c) Suponiendo que $\mathbf{N} > \mathbf{p}$, encontrar alguna matriz $(\mathbf{p} \times \mathbf{p})-$dimensional G tal que $\mathbf{GAG}' \sim \mathbf{W_p}(\mathbf{N} - \mathbf{1}, \mathbf{I_p})$.

Puesto que $A \sim W_p(N-1, \Sigma)$, tenemos por el resultado de reproductividad bajo transformaciones lineales rectangulares de una distribución de Wishart que $GAG' \sim W_p(N-1, G\Sigma G')$. Así pues, basta encontrar una matriz G de orden p tal que $G\Sigma G' = I_p$.

Tomamos una matriz C de dimensión $p \times p$ no singular tal que $\Sigma = CC'$ (que existe por ser $\Sigma > 0$) y, llamando $G = C^{-1}$ llegamos a que

$$G\Sigma G' = GCC'G' = C^{-1}CC'(C^{-1})' = C^{-1}CC'(C')^{-1} = I_p$$

como buscábamos.