Relación 2

David Cabezas Berrido

Ejercicio 1. Una generalización del resultado 5 II (ver debajo del ejercicio) nos dice que las a_{ii} son independientes por ser la matriz I_p diagonal por cajas (de tamaño 1).

Además, el resultado 5 I nos asegura que

$$a_{ii} = \frac{a_{ii}}{1} \sim \chi_n^2, \qquad i = 1, \ldots, p$$

Utilizando que la distribución χ^2 es reproductiva en los grados de libertad y que los elementos de la diagonal son variables aleatorias independientes, obtenemos

$$\operatorname{tr}(A) = \sum_{i=1}^{p} a_{ii} \sim \chi_{np}^2$$

Proposición (Generalizacion del resultado 5 II). *Sea $A \sim W_p(n, \Sigma)$, si Σ es diagonal por cajas, consideramos las cajas del mismo orden en la diagonal de A, es decir:*

$$A = \begin{pmatrix} A_{11} & \cdots & A_{1k} \\ \vdots & \ddots & \vdots \\ A_{k1} & \cdots & A_{kk} \end{pmatrix}, \quad \Sigma = \begin{pmatrix} \Sigma_{11} & \cdots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \cdots & \Sigma_{kk} \end{pmatrix}; \qquad \begin{array}{c} \textit{con } O(A_{jj}) = O(\Sigma_{jj}) \\ \textit{para cada } j = 1, \ldots, k \end{array}$$

notando por $O(C)$ al orden de una matriz C. Entonces, para cada $j = 1, \ldots, k$, $A_{jj} \sim W_{O(A_{jj})}(n, \Sigma_{jj})$, y las A_{jj} son independientes.

Demostración. La prueba es análoga a la del resultado 5 II. El hecho de que $A \sim W_p(n, \Sigma)$ nos asegura la existencia de n vectores aleatorios independientes $Z_1, \ldots, Z_n$ cada uno siguiendo una $N_p(0, \Sigma)$ tal que A se escribe como $A = \sum_{\alpha=1}^{n} Z_\alpha Z'_\alpha$. Podemos dividir cada vector Z_α en subvectores con tantas componentes como las cajas de la diagonal de A: $Z'_\alpha = (Z'_{\alpha 1} \cdots Z'_{\alpha k})$, donde $Z_{\alpha j}$ tiene $O(A_{jj})$ componentes.

Tenemos entonces que, para cada $j = 1, \ldots, k$, $A_{jj} = \sum_{\alpha=1}^{n} Z_{\alpha j} Z'_{\alpha j}$, y cada $Z_{\alpha j} \sim N_{O(A_{jj})}(0, \Sigma_{jj})$, de modo que $A_{jj} \sim W_{O(A_{jj})}(n, \Sigma_{jj})$.

Además, como Σ es diagonal por cajas, no sólo los vectores $Z_1, \ldots, Z_n$ son independientes, sino que los subvectores $Z_{\alpha j}$ con $\alpha = 1, \ldots n$ y $j = 1, \ldots, k$ son independientes. De modo que las cajas A_{jj} con $j = 1, \ldots, k$ son independientes. $\square$

Ejercicio 2. En el caso de que a y b sean linealmente independientes, la matriz $\begin{pmatrix} a & b \end{pmatrix}$ tiene rango máximo, aplicando entonces el resultado 3 con $M = \begin{pmatrix} a' \\ b' \end{pmatrix}$ obtenemos:

$$B = \begin{pmatrix} a' \\ b' \end{pmatrix} A \begin{pmatrix} a & b \end{pmatrix} = \begin{pmatrix} a'Aa & a'Ab \\ b'Aa & b'Ab \end{pmatrix} \sim W_2\left(n, \begin{pmatrix} a' \\ b' \end{pmatrix} \Sigma \begin{pmatrix} a & b \end{pmatrix}\right) \sim W_2\left(n, \begin{pmatrix} a'\Sigma a & a'\Sigma b \\ b'\Sigma a & b'\Sigma b \end{pmatrix}\right)$$

El resultado 5 II nos asegura que si $a'\Sigma b = \Sigma_{B12} = 0$, entonces las variables $a'Aa = B_{11}$ y $b'Ab = B_{22}$ son independientes.

El resultado 3 nos dice que $a'Aa$ y $b'Ab$ siguen Wishart centradas univariantes con n grados de libertad y $a'\Sigma a$, $b'\Sigma b$.

Recíprocamente, si $a'Aa = B_{11}$ y $b'Ab = B_{22}$ son independientes, en particular son incorreladas. Por tanto, $\operatorname{Cov}(B_{11}, B_{22}) = 0$. Utilizando ahora el resultado 1 obtenemos

$$0 = \operatorname{Cov}(B_{11}, B_{22}) = n(\Sigma_{B12}^2 + \Sigma_{B12}^2) = 2n(a'\Sigma b)^2,$$

lo que obliga a $a'\Sigma b = 0$.

En el caso de que a y b sean linealmente dependientes, existen dos formas de razonar: La primera es extender el resultado 3 al caso en el que la matriz M no sea de rango máximo, la demostración es exactamente la misma al caso de rango máximo pero utilizando la propiedad de la DNM con transformaciones lineales de rango no necesariamente básico. Este razonamiento también es válido en el caso $\Sigma \geq 0$.

La segunda es distinguir los casos en los que a y b sean linealmente dependientes:

- Si $a = 0$ ó $b = 0$, una (al menos) de las variables $a'Aa$ y $b'Ab$ es degenerada (constante $= 0$), por lo que claramente son independientes. Mientras que $a'\Sigma b = 0$, de modo que se sigue cumpliendo lo requerido.
- Si ambos son no nulos, existirá una constante $\lambda \neq 0$ tal que $b = \lambda a$. Las variables continuas y no degeneradas $a'Aa$ y $b'Ab = \lambda^2 a'Aa$ claramente no son independientes. Por otra parte, al ser $\Sigma > 0$, se tendrá $a'\Sigma b = \lambda a'\Sigma a \neq 0$. Así que una vez más se cumple lo requerido.

Ejercicio 3. Trabajaremos con la primera expresión de la función de verosimilitud, queremos maximizar en Σ:

$$L(\mu, \Sigma; x_1, \ldots, x_N) = \frac{1}{(2\pi)^{\frac{pN}{2}} |\Sigma|^{\frac{N}{2}}} \exp\left\{-\frac{1}{2}\sum_{\alpha=1}^{N}(x_\alpha - \mu)'\Sigma^{-1}(x_\alpha - \mu)\right\}$$

Usando que el logaritmo es estrictamente creciente, esto equivale a maximizar:

$$\ln L(\mu, \Sigma; x_1, \ldots, x_N) = -\frac{pN}{2}\ln 2\pi - \frac{N}{2}\ln|\Sigma| - \frac{1}{2}\sum_{\alpha=1}^{N}(x_\alpha - \mu)'\Sigma^{-1}(x_\alpha - \mu)$$

Sólo tenemos que maximizar la parte que depende de Σ, también utilizamos que el último sumando coincide con su traza por ser un escalar y eliminamos el $\frac{1}{2}$. Tenemos que maximizar entonces:

$$g(\Sigma) = -N \ln |\Sigma| - \operatorname{tr}\left(\sum_{\alpha=1}^{N}(x_\alpha - \mu)'\Sigma^{-1}(x_\alpha - \mu)\right)$$

Utilizamos la linealidad de la traza para sacar la sumatoria, y que la traza es invariante por permutaciones en el producto de matrices.

$$g(\Sigma) = -N \ln |\Sigma| - \sum_{\alpha=1}^{N} \operatorname{tr}\left(\Sigma^{-1}(x_\alpha - \mu)(x_\alpha - \mu)'\right)$$

Aplicando otra vez la linealidad de la traza y que Σ no depende de α:

$$g(\Sigma) = -N \ln |\Sigma| - \operatorname{tr}\left(\Sigma^{-1}\sum_{\alpha=1}^{N}(x_\alpha - \mu)(x_\alpha - \mu)'\right)$$

Estamos en condiciones de aplicar el Lema de Watson con $D = \sum_{\alpha=1}^{N}(x_\alpha - \mu)(x_\alpha - \mu)'$, que es simétrica y definida positiva (esto último por hipótesis). Obtenemos entonces que $\hat{\Sigma} = \frac{1}{N}D = \frac{1}{N}\sum_{\alpha=1}^{N}(X_\alpha - \mu)(X_\alpha - \mu)'$ maximiza g y por tanto L, luego es el EMV de Σ.

Para comprobar si es insesgado, hacemos uso de la linealidad de la esperanza. Después utilizamos que las variables X_α están identicamente distribuidas por una $N_p(\mu, \Sigma)$ y concluimos que efectivamente $\hat{\Sigma}$ es insesgado en Σ.

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

Ejercicio 4. Con la definición de esperanza con la función de densidad, tenemos que

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

donde $M_S^+(p)$ denota el espacio de matrices simétricas definidas positivas de orden p.

Para cada real $m > p - 1$, sabemos que la función de densidad de la distribución $W_p(m, \Sigma)$ debe integrar 1, es decir

$$\int_{M_S^+(p)} \frac{|A|^{\frac{m-p-1}{2}} \exp\{-\frac{1}{2}\operatorname{tr}(\Sigma^{-1}A)\}}{2^{\frac{mp}{2}}|\Sigma|^{\frac{m}{2}}\Gamma_p(\frac{m}{2})} dA = 1 \tag{2}$$

Multiplicamos y dividimos en (1) por los factores que faltan para utilizar (2) con $m = n + 2r > n \geq p > p - 1$: $\Gamma_p(\frac{n+2r}{2})$, $2^{\frac{2rp}{2}}$ y $|\Sigma|^{\frac{2r}{2}}$. Obtenemos entonces:

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

Reagrupando y usando (2) obtenemos:

$$E[|A|^r] = \frac{\Gamma_p(\frac{n+2r}{2})2^{\frac{2rp}{2}}|\Sigma|^{\frac{2r}{2}}}{\Gamma_p(\frac{n}{2})}\int_{M_S^+(p)} \frac{|A|^{\frac{n+2r-p-1}{2}}\exp\{-\frac{1}{2}\operatorname{tr}(\Sigma^{-1}A)\}}{2^{\frac{(n+2r)p}{2}}|\Sigma|^{\frac{n+2r}{2}}\Gamma_p(\frac{n+2r}{2})}dA = \frac{\Gamma_p(\frac{n+2r}{2})2^{\frac{2rp}{2}}|\Sigma|^{\frac{2r}{2}}}{\Gamma_p(\frac{n}{2})}$$

Por tanto, para cada $r > 0$ tenemos

$$E[|A|^r] = \frac{\Gamma_p(\frac{n+2r}{2})2^{rp}|\Sigma|^r}{\Gamma_p(\frac{n}{2})}$$

Ejercicio 5.

a) Sea $\mathcal{X} = \mathrm{vec}(\mathbf{X}) = \begin{pmatrix} X_1 \\ \vdots \\ X_N \end{pmatrix}$ el vector $pN \times 1$ formado por la concatenación de las observaciones. Buscamos primero una matriz E_α de dimensión $p \times pN$ tal que $X_\alpha = E_\alpha \mathcal{X}$, que claramente debe estar formada por $N-1$ cajas de ceros $p \times p$ y la identidad de orden p en la posición α:

$$E_\alpha = \begin{pmatrix} 0 & \cdots & 0 & I_p^{(\alpha)} & 0 & \cdots & 0 \end{pmatrix}$$

Podemos definir también $\bar{E} = \frac{1}{N}\sum_{\alpha=1}^{N} E_\alpha = \begin{pmatrix} \frac{1}{N} I_p & \cdots & \frac{1}{N} I_p \end{pmatrix}$, de tal forma que $\bar{X} = \bar{E}\mathcal{X}$.

Tenemos entonces

$$\begin{pmatrix} \bar{X} \\ X_\alpha - \bar{X} \end{pmatrix} = \begin{pmatrix} \bar{E} \\ E_\alpha - \bar{E} \end{pmatrix} \mathcal{X} = \begin{pmatrix} \bar{E} \\ E_\alpha - \bar{E} \end{pmatrix} \begin{pmatrix} X_1 \\ \vdots \\ X_N \end{pmatrix}$$

Podemos notar

$$F_\alpha = \begin{pmatrix} \bar{E} \\ E_\alpha - \bar{E} \end{pmatrix} = \begin{pmatrix} \frac{1}{N} I_p & \cdots & \cdots & \cdots & \cdots & \cdots & \frac{1}{N} I_p \\ -\frac{1}{N} I_p & \cdots & -\frac{1}{N} I_p & I_p - \frac{1}{N} I_p \ ^{(\alpha)} & -\frac{1}{N} I_p & \cdots & -\frac{1}{N} I_p \end{pmatrix}$$

Por otra parte, como consecuencia de la independencia de las observaciones, tenemos

$$\mathcal{X} = \begin{pmatrix} X_1 \\ \vdots \\ X_N \end{pmatrix} \sim N_{pN}\left(\begin{pmatrix} \mu \\ \vdots \\ \mu \end{pmatrix}, \begin{pmatrix} \Sigma & 0 & \cdots & 0 \\ 0 & \Sigma & \ddots & \vdots \\ \vdots & \ddots & \ddots & 0 \\ 0 & \cdots & 0 & \Sigma \end{pmatrix} \right)$$

De modo que $\begin{pmatrix} \bar{X} \\ X_\alpha - \bar{X} \end{pmatrix}$ sigue una normal $2p$-variante con vector de medias

$$F_\alpha \begin{pmatrix} \mu \\ \vdots \\ \mu \end{pmatrix} = \begin{pmatrix} \frac{1}{N} I_p & \cdots & \cdots & \cdots & \cdots & \cdots & \frac{1}{N} I_p \\ -\frac{1}{N} I_p & \cdots & -\frac{1}{N} I_p & I_p - \frac{1}{N} I_p \ ^{(\alpha)} & -\frac{1}{N} I_p & \cdots & -\frac{1}{N} I_p \end{pmatrix} \begin{pmatrix} \mu \\ \vdots \\ \mu \end{pmatrix} = \begin{pmatrix} N\frac{1}{N}\mu \\ \mu - N\frac{1}{N}\mu \end{pmatrix} = \begin{pmatrix} \mu \\ 0 \end{pmatrix}$$

y matriz de covarianzas

$$F_\alpha \begin{pmatrix} \Sigma & 0 & \cdots & 0 \\ 0 & \Sigma & \ddots & \vdots \\ \vdots & \ddots & \ddots & 0 \\ 0 & \cdots & 0 & \Sigma \end{pmatrix} F'_\alpha$$

$$= \begin{pmatrix} \frac{1}{N} I_p & \cdots & \cdots & \cdots & \cdots & \cdots & \frac{1}{N} I_p \\ -\frac{1}{N} I_p & \cdots & -\frac{1}{N} I_p & I_p - \frac{1}{N} I_p \ ^{(\alpha)} & -\frac{1}{N} I_p & \cdots & -\frac{1}{N} I_p \end{pmatrix} \begin{pmatrix} \Sigma & 0 & \cdots & 0 \\ 0 & \Sigma & \ddots & \vdots \\ \vdots & \ddots & \ddots & 0 \\ 0 & \cdots & 0 & \Sigma \end{pmatrix} \begin{pmatrix} \frac{1}{N} I_p & -\frac{1}{N} I_p \\ \vdots & \vdots \\ \vdots & -\frac{1}{N} I_p \\ \vdots & I_p - \frac{1}{N} I_p \ ^{(\alpha)} \\ \vdots & -\frac{1}{N} I_p \\ \vdots & \vdots \\ \frac{1}{N} I_p & -\frac{1}{N} I_p \end{pmatrix}$$

$$= \begin{pmatrix} \frac{1}{N} \Sigma & \cdots & \cdots & \cdots & \cdots & \cdots & \frac{1}{N} \Sigma \\ -\frac{1}{N} \Sigma & \cdots & -\frac{1}{N} \Sigma & \Sigma - \frac{1}{N} \Sigma \ ^{(\alpha)} & -\frac{1}{N} \Sigma & \cdots & -\frac{1}{N} \Sigma \end{pmatrix} \begin{pmatrix} \frac{1}{N} I_p & -\frac{1}{N} I_p \\ \vdots & \vdots \\ \vdots & -\frac{1}{N} I_p \\ \vdots & I_p - \frac{1}{N} I_p \ ^{(\alpha)} \\ \vdots & -\frac{1}{N} I_p \\ \vdots & \vdots \\ \frac{1}{N} I_p & -\frac{1}{N} I_p \end{pmatrix} = \begin{pmatrix} \frac{1}{N} \Sigma & 0 \\ 0 & \Sigma - \frac{1}{N} \Sigma \end{pmatrix}$$

Como $\begin{pmatrix} \bar{X} \\ X_\alpha - \bar{X} \end{pmatrix}$ sigue una distribución normal con matriz de covarianzas diagonal por cajas, concluimos que que las variables $\bar{X}$ y $X_\alpha - \bar{X}$ son independientes y por tanto incorreladas: $\mathrm{Cov}(\bar{X}, X_\alpha - \bar{X}) = 0$.

Además, como A es una función medible de las variables $X_\alpha - \bar{X}, \quad \alpha = 1, \ldots, N$; y $\bar{X}$ es independiente de todas ellas, deducimos que $\bar{X}$ y A son independientes.

b) Buscaré primero una matriz C tal que $C\mathbf{X} = C \begin{pmatrix} X'_1 \\ \vdots \\ X'_N \end{pmatrix} = \begin{pmatrix} X'_1 - \bar{X}' \\ \vdots \\ X'_N - \bar{X}' \end{pmatrix}$.

Necesito entonces

$$C\mathbf{X} = I_N \mathbf{X} - \frac{1}{N} \begin{pmatrix} \sum_{\alpha=1}^N X'_\alpha \\ \vdots \\ \sum_{\alpha=1}^N X'_\alpha \end{pmatrix}$$

Desarrollando el segundo sumando:

$$\begin{pmatrix} \sum_{\alpha=1}^{N} X'_\alpha \\ \vdots \\ \sum_{\alpha=1}^{N} X'_\alpha \end{pmatrix} = \begin{pmatrix} \sum_{\alpha=1}^{N} X_{\alpha 1} & \cdots & \sum_{\alpha=1}^{N} X_{\alpha p} \\ \vdots & & \vdots \\ \sum_{\alpha=1}^{N} X_{\alpha 1} & \cdots & \sum_{\alpha=1}^{N} X_{\alpha p} \end{pmatrix} = \begin{pmatrix} 1 & \cdots & 1 \\ \vdots & & \vdots \\ 1 & \cdots & 1 \end{pmatrix} \begin{pmatrix} X_{11} & \cdots & X_{1p} \\ \vdots & & \vdots \\ X_{N1} & \cdots & X_{Np} \end{pmatrix}$$

$$= \begin{pmatrix} 1 & \cdots & 1 \\ \vdots & & \vdots \\ 1 & \cdots & 1 \end{pmatrix} \begin{pmatrix} X'_1 \\ \vdots \\ X'_N \end{pmatrix} = \begin{pmatrix} 1 & \cdots & 1 \\ \vdots & & \vdots \\ 1 & \cdots & 1 \end{pmatrix} \mathbf{X} = \mathbf{1}_N \mathbf{X}$$

Por tanto $C = I_N - \frac{1}{N}\mathbf{1}_N$. Tenemos también

$$\begin{pmatrix} X_1 - \bar{X} & \cdots & X_N - \bar{X} \end{pmatrix} = (C\mathbf{X})' = \mathbf{X}'C' = \mathbf{X}'C$$

Finalmente,

$$A = \sum_{\alpha=1}^{N} (X_\alpha - \bar{X})(X_\alpha - \bar{X})' = \begin{pmatrix} X_1 - \bar{X} & \cdots & X_N - \bar{X} \end{pmatrix} \begin{pmatrix} X'_1 - \bar{X}' \\ \vdots \\ X'_N - \bar{X}' \end{pmatrix} = \mathbf{X}'CC\mathbf{X}$$

Luego obtenemos lo requerido para $B = C^2$.

c) Por ser A la matriz de dispersiones de una muestra aleatoria simple (de tamaño $N > p$) de una distribución $N_p(\mu, \Sigma)$ con $\Sigma > 0$, sabemos que $A \sim W_p(N-1, \Sigma)$.

Siempre que G sea de rango p, tendremos por el resultado 3 $GAG' \sim W_p(N-1, G\Sigma G')$. Como Σ es simétrica y definida positiva, admite una factorización de la forma $\Sigma = CC'$ con C matriz cuadrada no singular de orden p. Tomamos $G = C^{-1}$, que es una matriz no singular de orden p. Tenemos entonces

$$GAG' \sim W_p(N-1, G\Sigma G') = W_p(N-1, C^{-1}CC'C^{-1\prime}) = W_p(N-1, I_p)$$