MAT-269: Sesión 7
Estimación bajo restricciones sobre μ y Σ

Felipe Osorio

fosorios.mat.utfsm.cl

Departamento de Matemática, UTFSM

Estimadores ML restringidos

Suponga $\boldsymbol{x}_1, \dots, \boldsymbol{x}_n$ vectores aleatorios IID desde $\mathsf{N}_p(\boldsymbol{\mu}, \boldsymbol{\Sigma})$ y considere:[1]

1. $\boldsymbol{\mu} = \boldsymbol{\mu}_0$ conocido. Entonces,
$$\widehat{\boldsymbol{\Sigma}} = \frac{1}{n} \sum_{i=1}^{n} (\boldsymbol{x}_i - \boldsymbol{\mu}_0)(\boldsymbol{x}_i - \boldsymbol{\mu}_0)^\top.$$
2. $\boldsymbol{\Sigma} = \boldsymbol{\Sigma}_0$ conocido. De este modo,
$$\widehat{\boldsymbol{\mu}} = \overline{\boldsymbol{x}}.$$
3. $\boldsymbol{\mu} = \gamma \boldsymbol{a}$, $\gamma \in \mathbb{R}$ con $\boldsymbol{a} \in \mathbb{R}^p$ conocido. Luego,
$$\widehat{\gamma}_\Sigma = \frac{\boldsymbol{a}^\top \boldsymbol{\Sigma}^{-1} \overline{\boldsymbol{x}}}{\boldsymbol{a}^\top \boldsymbol{\Sigma}^{-1} \boldsymbol{a}}.$$
Para $\boldsymbol{\Sigma}$ desconocido, tenemos:
$$\widehat{\gamma} = \frac{\boldsymbol{a}^\top \boldsymbol{S}^{-1} \overline{\boldsymbol{x}}}{\boldsymbol{a}^\top \boldsymbol{S}^{-1} \boldsymbol{a}}.$$

[1] Ud. lo deberá resolver como parte de la Tarea 1 (Entrega: 4 Mayo).

4. $\boldsymbol{A\mu} = \boldsymbol{a}$, $\boldsymbol{A} \in \mathbb{R}^{q\times p}$, $\boldsymbol{a} \in \mathbb{R}^q$ matrices conocidas. Luego

$$\widehat{\boldsymbol{\mu}}_{\Sigma} = \overline{\boldsymbol{x}} - \boldsymbol{\Sigma A}^\top(\boldsymbol{A\Sigma A}^\top)^{-1}(\boldsymbol{A}\overline{\boldsymbol{x}} - \boldsymbol{a}),$$

para $\boldsymbol{\Sigma}$ desconocido

$$\widehat{\boldsymbol{\mu}} = \overline{\boldsymbol{x}} - \boldsymbol{SA}^\top(\boldsymbol{ASA}^\top)^{-1}(\boldsymbol{A}\overline{\boldsymbol{x}} - \boldsymbol{a}).$$

5. $\boldsymbol{\Sigma} = \phi\boldsymbol{V}$ con $\boldsymbol{V} > 0$ conocida y $\phi > 0$. Por tanto,

$$\widehat{\boldsymbol{\mu}} = \overline{\boldsymbol{x}}, \qquad \widehat{\phi} = \frac{1}{p}\operatorname{tr}(\boldsymbol{V}^{-1}\boldsymbol{S}).$$

Suponga que

$$\boldsymbol{\Sigma} = \begin{pmatrix} \boldsymbol{\Sigma}_{11} & \boldsymbol{\Sigma}_{12} \\ \boldsymbol{\Sigma}_{21} & \boldsymbol{\Sigma}_{22} \end{pmatrix}, \qquad \text{con} \qquad \boldsymbol{\Sigma}_{12} = \mathbf{0} = \boldsymbol{\Sigma}_{21}^\top,$$

y $\boldsymbol{\mu} = (\boldsymbol{\mu}_1^\top, \boldsymbol{\mu}_2^\top)^\top$. De este modo, tenemos $\boldsymbol{\theta} = (\boldsymbol{\mu}_1, \boldsymbol{\mu}_2, \boldsymbol{\Sigma}_{11}, \boldsymbol{\Sigma}_{22})$, así la función de log-verosimilitud adopta la forma:

$$\ell(\boldsymbol{\theta}) = -\frac{np}{2} \log 2\pi - \frac{n}{2} \log |\boldsymbol{\Sigma}| - \frac{1}{2} \sum_{i=1}^{n} (\boldsymbol{x}_i - \boldsymbol{\mu})^\top \boldsymbol{\Sigma}^{-1} (\boldsymbol{x}_i - \boldsymbol{\mu}).$$

Note que el estimador ML para $\boldsymbol{\mu}$ no depende de $\boldsymbol{\Sigma}$, luego $\widehat{\boldsymbol{\mu}} = \overline{\boldsymbol{x}}$. Ahora, usando que

$$|\boldsymbol{\Sigma}| = \begin{vmatrix} \boldsymbol{\Sigma}_{11} & \mathbf{0} \\ \mathbf{0} & \boldsymbol{\Sigma}_{22} \end{vmatrix} = |\boldsymbol{\Sigma}_{11}||\boldsymbol{\Sigma}_{22}|, \qquad \boldsymbol{\Sigma}^{-1} = \begin{pmatrix} \boldsymbol{\Sigma}_{11}^{-1} & \mathbf{0} \\ \mathbf{0} & \boldsymbol{\Sigma}_{22}^{-1} \end{pmatrix}.$$

Matriz de covarianza diagonal

Sigue que la parte relevante de $\ell(\boldsymbol{\theta})$ es dada por:

$$\begin{aligned}\ell(\boldsymbol{\theta}) = &-\frac{n}{2}\log|\boldsymbol{\Sigma}_{11}| - \frac{1}{2}\sum_{i=1}^{n}(\boldsymbol{x}_i - \boldsymbol{\mu}_1)^\top \boldsymbol{\Sigma}_{11}^{-1}(\boldsymbol{x}_i - \boldsymbol{\mu}_1) \\ &- \frac{n}{2}\log|\boldsymbol{\Sigma}_{22}| - \frac{1}{2}\sum_{i=1}^{n}(\boldsymbol{x}_i - \boldsymbol{\mu}_2)^\top \boldsymbol{\Sigma}_{22}^{-1}(\boldsymbol{x}_i - \boldsymbol{\mu}_2)\end{aligned}$$

que puede ser escrita como:

$$\ell(\boldsymbol{\theta}) = \ell_1(\boldsymbol{\theta}_1) + \ell_2(\boldsymbol{\theta}_2),$$

con $\boldsymbol{\theta}_j = (\boldsymbol{\mu}_j, \boldsymbol{\Sigma}_{jj})$, para $j = 1, 2$, y

$$\ell_j(\boldsymbol{\theta}_j) = -\frac{n}{2}\log|\boldsymbol{\Sigma}_{jj}| - \frac{1}{2}\operatorname{tr}\boldsymbol{\Sigma}_{jj}^{-1}\boldsymbol{Q}_j(\boldsymbol{\mu}_j),$$

donde

$$\boldsymbol{Q}_j(\boldsymbol{\mu}_j) = \sum_{i=1}^{n}(\boldsymbol{x}_i - \boldsymbol{\mu}_j)(\boldsymbol{x}_i - \boldsymbol{\mu}_j)^\top.$$

Matriz de covarianza diagonal

De ahí que, usando el Resultado 2 de la Sesión 6, obtenemos el estimador ML para $\mathbf{\Sigma}$:

$$\widehat{\mathbf{\Sigma}} = \begin{pmatrix} \widehat{\mathbf{\Sigma}}_{11} & \mathbf{0} \\ \mathbf{0} & \widehat{\mathbf{\Sigma}}_{22} \end{pmatrix},$$

con

$$\widehat{\mathbf{\Sigma}}_{jj} = \frac{1}{n}\sum_{i=1}^{n}(\boldsymbol{x}_i - \overline{\boldsymbol{x}}_j)(\boldsymbol{x}_i - \overline{\boldsymbol{x}}_j)^\top,$$

para $j = 1, 2$, donde $\overline{\boldsymbol{x}} = (\overline{\boldsymbol{x}}_1^\top, \overline{\boldsymbol{x}}_2^\top)^\top$.

Suponga que la matriz de datos $\boldsymbol{X}$ es particionada como:

$$\boldsymbol{X} = \begin{pmatrix} \boldsymbol{X}_1 \\ \vdots \\ \boldsymbol{X}_k \end{pmatrix},$$

donde las filas de $\boldsymbol{X}_i \in \mathbb{R}^{n_i \times p}$ son IID $\mathsf{N}_p(\boldsymbol{\mu}_i, \boldsymbol{\Sigma}_{ii})$, para $i = 1, \dots, k$.

Las restricciones más comunes son:

(a) $\boldsymbol{\Sigma}_{11} = \cdots = \boldsymbol{\Sigma}_{kk}$ (digamos, $= \boldsymbol{\Sigma}$).

(b) $\boldsymbol{\Sigma}_{11} = \cdots = \boldsymbol{\Sigma}_{kk}$ y $\boldsymbol{\mu}_1 = \cdots \boldsymbol{\mu}_k$.

Para el caso en (a) note que podemos escribir:

$$\ell(\boldsymbol{\theta}) = \sum_{i=1}^{k} \ell_i(\boldsymbol{\theta}),$$

donde

$$\ell_i(\boldsymbol{\theta}) = -\frac{n_i}{2}\log|\boldsymbol{\Sigma}| - \frac{1}{2}\operatorname{tr}\boldsymbol{\Sigma}^{-1}\{\boldsymbol{Q}_i + n_i(\overline{\boldsymbol{x}}_i - \boldsymbol{\mu}_i)(\overline{\boldsymbol{x}}_i - \boldsymbol{\mu}_i)^\top\},$$

Sea

$$\boldsymbol{S}_i = \frac{1}{n_i}\boldsymbol{Q}_i = \frac{1}{n_i}\sum_{j=1}^{n_i}(\boldsymbol{x}_j - \overline{\boldsymbol{x}}_i)(\boldsymbol{x}_j - \overline{\boldsymbol{x}}_i)^\top,$$

para $i = 1, \dots, k$. Es decir,

$$\ell(\boldsymbol{\theta}) = -\frac{1}{2}\sum_{i=1}^{k}\Big\{n_i \log|\boldsymbol{\Sigma}| + n_i \operatorname{tr}\boldsymbol{\Sigma}^{-1}(\boldsymbol{S}_i + (\overline{\boldsymbol{x}}_i - \boldsymbol{\mu}_i)(\overline{\boldsymbol{x}}_i - \boldsymbol{\mu}_i)^\top)\Big\}.$$

Como no existe restricciones sobre $\boldsymbol{\mu}$ sigue que el MLE de $\boldsymbol{\mu}$ es $\overline{\boldsymbol{x}}_i$ $(i = 1, \dots, k)$. Además, considere

$$\boldsymbol{W} = \sum_{i=1}^{k} n_i \boldsymbol{S}_i, \qquad n = \sum_{i=1}^{k} n_i.$$

De este modo, la log-verosimilitud perfilada, es dada por:

$$\ell_*(\boldsymbol{\Sigma}) = \ell(\widehat{\boldsymbol{\mu}}, \boldsymbol{\Sigma}) = -\frac{n}{2} \log |\boldsymbol{\Sigma}| - \frac{1}{2} \operatorname{tr} \boldsymbol{\Sigma}^{-1} \boldsymbol{W}.$$

De ahí que el MLE para $\boldsymbol{\Sigma}$ adopta la forma:[2]

$$\widehat{\boldsymbol{\Sigma}} = \frac{1}{n} \boldsymbol{W} = \frac{1}{n} \sum_{i=1}^{k} n_i \boldsymbol{S}_i$$

El caso en (b) es análogo (se deja como Ejercicio).

[2] Es decir $\widehat{\boldsymbol{\Sigma}}_{ii} = \widehat{\boldsymbol{\Sigma}}$, para $i = 1, \dots, k$.

MAT-269: Sesión 7
Estimación ML bajo distribuciones de contornos elípticos

Felipe Osorio

fosorios.mat.utfsm.cl

Departamento de Matemática, UTFSM

Sabemos que un vector aleatorio p-variado tiene distribución en la familia de contornos elípticos con parámetros $\boldsymbol{\mu} \in \mathbb{R}^p$ y $\boldsymbol{\Sigma} \geq 0$ si su función característica es de la forma:

$$\varphi_x(\boldsymbol{t}) = \exp(i\boldsymbol{t}^\top \boldsymbol{\mu})\phi(\boldsymbol{t}^\top \boldsymbol{\Sigma}\boldsymbol{t}),$$

mientras que si $\boldsymbol{\Sigma} > 0$ el vector aleatorio $\boldsymbol{x}$ tendrá densidad

$$f(\boldsymbol{x}) = |\boldsymbol{\Sigma}|^{-1/2} g[(\boldsymbol{x} - \boldsymbol{\mu})^\top \boldsymbol{\Sigma}^{-1}(\boldsymbol{x} - \boldsymbol{\mu})], \qquad (1)$$

con $g : \mathbb{R} \to [0, \infty)$ llamada función generadora de densidad,[1] tal que:

$$\int_0^\infty u^{p/2} g(u)\, \mathrm{d}u < \infty.$$

Cuando un vector aleatorio tiene densidad como en (1) anotamos $\boldsymbol{x} \sim \mathsf{EC}_p(\boldsymbol{\mu}, \boldsymbol{\Sigma}; g)$.

[1] Por ejemplo, $g(u) = (2\pi)^{-p/2} \exp(-u/2)$ para el caso normal.

- Normal: $\boldsymbol{x} \sim \mathsf{N}_p(\boldsymbol{\mu}, \boldsymbol{\Sigma})$, con
$$g(u) = c_1 \exp(-u/2).$$
- t-Student: $\boldsymbol{x} \sim t_p(\boldsymbol{\mu}, \boldsymbol{\Sigma}, \nu)$, donde
$$g(u) = c_2(1 + u/\nu)^{-(\nu+p)/2}, \qquad \nu > 0.$$
- Normal contaminada: $\boldsymbol{x} \sim \mathsf{CN}_p(\boldsymbol{\mu}, \boldsymbol{\Sigma}, \epsilon, \gamma)$, con $\epsilon \in [0, 1)$ y $\gamma > 0$,
$$g(u) = c_1\big\{(1-\epsilon)\exp(-u/2) + \epsilon\gamma^{-p/2}\exp(-u/(2\gamma))\big\}.$$
- Cauchy: $\boldsymbol{x} \sim \mathsf{Cauchy}_p(\boldsymbol{\mu}, \boldsymbol{\Sigma})$, con
$$g(u) = c_3(1 + u)^{-(p+1)/2}.$$
- Logística: $\boldsymbol{x} \sim \mathsf{L}_p(\boldsymbol{\mu}, \boldsymbol{\Sigma})$, con
$$g(u) = c_4 \exp(-u)/\{1 + \exp(-u)\}^2.$$
- Exponencial Potencia: $\boldsymbol{x} \sim \mathsf{PE}_p(\boldsymbol{\mu}, \boldsymbol{\Sigma}, \lambda)$, donde
$$g(u) = c_5 \exp(-u^\lambda/2), \qquad \lambda > 0.$$

Observación:

Debemos resaltar que diferentemente al caso de la distribución normal, en el caso general de la familia elíptica, podemos tener los siguientes enfoques:

(a) Modelo dependiente: Supondremos $\boldsymbol{x}_1, \dots, \boldsymbol{x}_n$ tal que su densidad conjunta $\boldsymbol{x} = (\boldsymbol{x}_1^\top, \dots, \boldsymbol{x}_n^\top)^\top$, sigue una distribución de contornos elipticos.

(b) Modelo independiente: Considere $\boldsymbol{x}_1, \dots, \boldsymbol{x}_n$ vectores aleatorios independientes cada uno con distribución $\mathrm{EC}_p(\boldsymbol{\mu}, \boldsymbol{\Sigma}; g)$.

Estimación ML: caso dependiente

Considere la matriz de datos:

$$\boldsymbol{X} = \begin{pmatrix} \boldsymbol{x}_1^\top \\ \vdots \\ \boldsymbol{x}_n^\top \end{pmatrix}$$

distribuído de acuerdo con una distribución en la familia de contornos elípticos con parámetros

$$\boldsymbol{\mu} = (\boldsymbol{\mu}_1^\top, \dots, \boldsymbol{\mu}_n^\top)^\top, \qquad \boldsymbol{\Sigma} = \operatorname{block\,diag}(\boldsymbol{\Sigma}_1, \dots, \boldsymbol{\Sigma}_n),$$

donde $\boldsymbol{\Sigma}_j > \mathbf{0}$, para $j = 1, \dots, n$. De este modo, $\boldsymbol{X}$ tiene densidad de la forma:

$$\prod_{j=1}^n |\boldsymbol{\Sigma}_j|^{-1/2} \, g\Big[\sum_{i=1}^n (\boldsymbol{x}_j - \boldsymbol{\mu}_j)^\top \boldsymbol{\Sigma}_j^{-1} (\boldsymbol{x}_j - \boldsymbol{\mu}_j)\Big]. \qquad (2)$$

Observación:

Evidentemente la distribución matricial en (2) puede ser escrita como una distribución multivariada definiendo:

$$\boldsymbol{x} = \operatorname{vec}(\boldsymbol{X}^\top) = (\boldsymbol{x}_1^\top, \dots, \boldsymbol{x}_n^\top)^\top,$$

de este modo, $\boldsymbol{x} \sim \mathsf{EC}_N(\boldsymbol{\mu}, \boldsymbol{\Sigma}; g)$, con $N = np$.

Resultado 1 (Anderson, Fang y Hsu, 1986)[2]

Sea Ω un conjunto en el espacio paramétrico de $(\boldsymbol{\mu}, \boldsymbol{\Sigma})$, $\boldsymbol{\Sigma} > \mathbf{0}$ tal que si $(\boldsymbol{\mu}, \boldsymbol{\Sigma}) \in \Omega$, entonces $(\boldsymbol{\mu}, c\boldsymbol{\Sigma}) \in \Omega$ para todo $c > 0$. Suponga que g es función tal que $g(\|\boldsymbol{x}\|^2)$ es una densidad en $\mathbb{R}^N$ y $u^{N/2}g(u)$ tiene un máximo finito (positivo) u_g. Suponga que, basado en una (única) observación $\boldsymbol{x}$ desde

$$|\boldsymbol{\Sigma}|^{-1/2} g[(\boldsymbol{x} - \boldsymbol{\mu})^\top \boldsymbol{\Sigma}^{-1} (\boldsymbol{x} - \boldsymbol{\mu})],$$

los MLE bajo normalidad $(\widetilde{\boldsymbol{\mu}}, \widetilde{\boldsymbol{\Sigma}}) \in \Omega$ existen y son únicos y que $\widetilde{\boldsymbol{\Sigma}} > \mathbf{0}$ con prob. 1. Entonces los MLE para $\boldsymbol{x} \sim \mathsf{EC}_N(\boldsymbol{\mu}, \boldsymbol{\Sigma}; g)$ son:

$$\widehat{\boldsymbol{\mu}} = \widetilde{\boldsymbol{\mu}}, \qquad \widehat{\boldsymbol{\Sigma}} = \frac{N}{u_g} \widetilde{\boldsymbol{\Sigma}},$$

y el máximo de la verosimilitud es

$$|\widehat{\boldsymbol{\Sigma}}|^{-1/2} g(u_g).$$

[2] The Canadian Journal of Statistics **14**, 55-59.

Demostración:

Sea $\boldsymbol{B} = |\boldsymbol{\Sigma}|^{-1/N}\boldsymbol{\Sigma}$ y

$$u = (\boldsymbol{x}-\boldsymbol{\mu})^\top \boldsymbol{\Sigma}^{-1}(\boldsymbol{x}-\boldsymbol{\mu}) = \frac{(\boldsymbol{x}-\boldsymbol{\mu})^\top \boldsymbol{B}^{-1}(\boldsymbol{x}-\boldsymbol{\mu})}{|\boldsymbol{\Sigma}|^{1/N}}.$$

Entonces $(\boldsymbol{\mu}, \boldsymbol{B}) \in \Omega$ y $|\boldsymbol{B}| = 1$. Note que

$$\begin{aligned} u^{N/2} &= \Big\{\frac{(\boldsymbol{x}-\boldsymbol{\mu})^\top \boldsymbol{B}^{-1}(\boldsymbol{x}-\boldsymbol{\mu})}{|\boldsymbol{\Sigma}|^{1/N}}\Big\}^{N/2} \\ &= |\boldsymbol{\Sigma}|^{-1/2}\{(\boldsymbol{x}-\boldsymbol{\mu})^\top \boldsymbol{B}^{-1}(\boldsymbol{x}-\boldsymbol{\mu})\}^{N/2}. \qquad (3) \end{aligned}$$

De este modo, la función de verosimilitud es dada por:

$$L(\boldsymbol{\theta}) = \{(\boldsymbol{x}-\boldsymbol{\mu})^\top \boldsymbol{B}^{-1}(\boldsymbol{x}-\boldsymbol{\mu})\}^{N/2} u^{N/2} g(u). \qquad (4)$$

Estimación ML: caso dependiente

Bajo normalidad, tenemos que

$$g(u) = (2\pi)^{-N/2}\exp(-u/2),$$

y el máximo de (4) es alcanzado en

$$\boldsymbol{\mu} = \widetilde{\boldsymbol{\mu}}, \qquad \boldsymbol{B} = \widetilde{\boldsymbol{B}} = |\widetilde{\boldsymbol{\Sigma}}|^{-1/N}\widetilde{\boldsymbol{\Sigma}},$$

y $\widetilde{u} = N$. En general, el máximo de $L(\boldsymbol{\theta})$ es alcanzado en

$$\widehat{\boldsymbol{\mu}} = \widetilde{\boldsymbol{\mu}}, \qquad \widehat{\boldsymbol{B}} = \widetilde{\boldsymbol{B}}, \qquad \widehat{u} = u_g.$$

Entonces,

$$\widehat{\boldsymbol{\Sigma}} = |\widehat{\boldsymbol{\Sigma}}|^{1/N}\widehat{\boldsymbol{B}} = \frac{|\widehat{\boldsymbol{\Sigma}}|^{1/N}}{|\widetilde{\boldsymbol{\Sigma}}|^{1/N}}\,\widetilde{\boldsymbol{\Sigma}}.$$

Estimación ML: caso dependiente

Usando (3), sigue que:

$$\frac{|\widehat{\boldsymbol{\Sigma}}|^{1/N}}{|\widetilde{\boldsymbol{\Sigma}}|^{1/N}} = \frac{(\boldsymbol{x} - \widehat{\boldsymbol{\mu}})^\top \widehat{\boldsymbol{B}}^{-1}(\boldsymbol{x} - \widehat{\boldsymbol{\mu}})/\widehat{u}}{(\boldsymbol{x} - \widetilde{\boldsymbol{\mu}})^\top \widetilde{\boldsymbol{B}}^{-1}(\boldsymbol{x} - \widetilde{\boldsymbol{\mu}})/\widetilde{u}} = \frac{\widetilde{u}}{\widehat{u}},$$

lo que permite obtener:

$$\widehat{\boldsymbol{\Sigma}} = \frac{N}{\widehat{u}}\widetilde{\boldsymbol{\Sigma}}.$$

Ahora, por (3), tenemos que

$$\begin{aligned} L(\widehat{\boldsymbol{\theta}}) &= \{(\boldsymbol{x} - \widehat{\boldsymbol{\mu}})^\top \widehat{\boldsymbol{B}}^{-1}(\boldsymbol{x} - \widehat{\boldsymbol{\mu}})\}^{-N/2}\widehat{u}^{N/2}g(\widehat{u}) \\ &= \{\widehat{u}|\widehat{\boldsymbol{\Sigma}}|^{1/N}\}^{-N/2}\widehat{u}^{N/2}g(\widehat{u}) \\ &= |\widehat{\boldsymbol{\Sigma}}|^{-1/2}g(\widehat{u}), \end{aligned}$$

lo que concluye la prueba.

Estimación ML: caso dependiente

Observación:

Si g es continua y diferenciable, entonces u_g[3] es la solución de:

$$g'(u) + \frac{N}{2u} g(u) = 0,$$

o bien

$$\frac{N}{2u} + W_g(u) = 0,$$

donde $W_g(u) = \mathrm{d} \log g(u) / \mathrm{d}u = g'(u)/g(u)$.

Es fácil notar que para las distribuciones normal y t de Student, $u_g = N$. Para otras distribuciones, u_g debe ser obtenido numéricamente. Por ejemplo, para la distribución logística se debe resolver:

$$\frac{N}{2u} = \tanh\left(\frac{u}{2}\right).$$

[3] u_g maximiza la función $h(u) = u^{N/2} g(u)$.

EX UMBRA IN SOLEM

Ejemplo:

Suponga la densidad conjunta en (2) con $\boldsymbol{\mu}_1 = \cdots = \boldsymbol{\mu}_n = \boldsymbol{\mu}$ y $\boldsymbol{\Sigma}_1 = \cdots = \boldsymbol{\Sigma}_n = \boldsymbol{\Sigma}$ y $n > p$. Bajo normalidad los MLE de $\boldsymbol{\mu}$ y $\boldsymbol{\Sigma}$ son $\widetilde{\boldsymbol{\mu}} = \overline{\boldsymbol{x}}$ y $\widetilde{\boldsymbol{\Sigma}} = \boldsymbol{Q}/n$, donde

$$\overline{\boldsymbol{x}} = \frac{1}{n}\sum_{j=1}^{n} \boldsymbol{x}_j, \qquad \boldsymbol{Q} = \sum_{j=1}^{n} (\boldsymbol{x}_j - \overline{\boldsymbol{x}})(\boldsymbol{x}_j - \overline{\boldsymbol{x}})^\top.$$

De este modo los MLE bajo el modelo elíptico dependiente son:

$$\widehat{\boldsymbol{\mu}} = \overline{\boldsymbol{x}}, \qquad \widehat{\boldsymbol{\Sigma}} = \frac{p}{u_g}\boldsymbol{Q}.$$