MAT-269: 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

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

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
$$g(u) = c_5 \exp(-u^\lambda/2), \qquad \lambda > 0.$$

EX UMBRA IN SOLEM

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 $\mathsf{EC}_p(\boldsymbol{\mu}, \boldsymbol{\Sigma}; g)$.

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:

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

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 $\boldsymbol{x} \sim \text{EC}_N(\boldsymbol{\mu}, \boldsymbol{\Sigma}; g)$ 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

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

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

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

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,

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

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 (4), 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 $h(u) = u^{N/2} g(u)$.

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:

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