MAT-269: Estimación ML bajo distribuciones de contornos elípticos: caso independiente

Felipe Osorio

fosorios.mat.utfsm.cl

Departamento de Matemática, UTFSM

EX UMBRA IN SOLEM

Un vector aleatorio $\boldsymbol{x} \sim \mathsf{EC}_p(\boldsymbol{\mu}, \boldsymbol{\Sigma}; g)$[1] con $\boldsymbol{\mu} \in \mathbb{R}^p$ y $\boldsymbol{\Sigma} > \boldsymbol{0}$ tiene función de densidad

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

si y sólo si $g : \mathbb{R} \to [0, \infty)$ es tal que

$$\int_0^\infty u^{p/2-1} g(u)\, \mathrm{d}u < \infty,$$

y decimos que $g(\cdot)$ es la generadora de densidad.

[1] Cuando $\boldsymbol{x} \sim \mathsf{EC}_p(\boldsymbol{0}, \boldsymbol{I}; g)$ es usual escribir $\boldsymbol{x} \sim \mathsf{S}_p(g)$.

Resultado 1:

$\boldsymbol{x} \sim \mathsf{EC}_p(\boldsymbol{\mu}, \boldsymbol{\Sigma}; g)$, $\boldsymbol{\Sigma} > \mathbf{0}$, si y sólo si

$$\boldsymbol{x} \stackrel{\mathsf{d}}{=} \boldsymbol{\mu} + R\boldsymbol{B}\boldsymbol{u},$$

donde $R \geq 0$ es independiente de $\boldsymbol{u}$ y $\boldsymbol{B}$ es matriz tal que $\boldsymbol{\Sigma} = \boldsymbol{B}\boldsymbol{B}^\top$.

Resultado 2:

Suponga que $\boldsymbol{x} \sim \mathsf{EC}_p(\boldsymbol{\mu}, \boldsymbol{\Sigma}; g)$ y $\mathsf{E}(R^2) < \infty$. Entonces

$$\mathsf{E}(\boldsymbol{x}) = \boldsymbol{\mu}, \qquad \mathsf{Cov}(\boldsymbol{x}) = \frac{\mathsf{E}(R^2)}{p}\boldsymbol{\Sigma}.$$

Considere $\boldsymbol{x} \stackrel{\mathsf{d}}{=} R\boldsymbol{u} \sim \mathsf{S}_p(g)$, la densidad de R es determinada desde $g(\boldsymbol{x}^\top \boldsymbol{x})$ mediante transformar en coordenadas polares y luego obtener su distribución marginal. La densidad es dada por:

$$h(r) = \frac{2\pi^{p/2-1}}{\Gamma(p/2)}\, r^{p-1} g(r^2).$$

De ahí que $\mathsf{E}(R^k) < \infty$ si y solo si,

$$\int_0^\infty r^{k+p-1} g(r^2)\, \mathsf{d}r < \infty$$

La función característica de $\boldsymbol{x}$ es

$$\phi(\boldsymbol{t}^\top \boldsymbol{t}) = \mathsf{E}\{\exp(iR\boldsymbol{t}^\top \boldsymbol{u})\} = \int_0^\infty \psi(r^2 \boldsymbol{t}^\top \boldsymbol{t}) h(r)\, \mathsf{d}r,$$

con $\psi(\boldsymbol{s}^\top \boldsymbol{s}) = \mathsf{E}(e^{i\boldsymbol{s}^\top \boldsymbol{u}})$ la función característica de $\boldsymbol{u}$.

Sabemos que[2]

$$\mathsf{E}(\boldsymbol{x}) = \boldsymbol{\mu}, \qquad \mathsf{E}\{(\boldsymbol{x}-\boldsymbol{\mu})(\boldsymbol{x}-\boldsymbol{\mu})^\top\} = \frac{\mathsf{E}(R^2)}{p}\boldsymbol{\Sigma},$$

además

$$\mathsf{E}\{(\boldsymbol{x}-\boldsymbol{\mu})\otimes(\boldsymbol{x}-\boldsymbol{\mu})^\top\otimes(\boldsymbol{x}-\boldsymbol{\mu})\} = \mathbf{0}.$$

En efecto, todos los momentos impares de $\boldsymbol{x}-\boldsymbol{\mu}$ son cero. Los primeros momentos de R están relacionados con la función característica $\phi(\cdot)$,

$$\mathsf{E}(R^2) = -2p\phi'(0), \qquad \mathsf{E}(R^4) = 4p(p+2)\phi''(0).$$

El exceso de curtosis, adopta la forma:

$$\begin{aligned}\frac{\mathsf{E}\{(x_i-\mu_i)^4\} - 3[\mathsf{E}\{(x_i-\mu_i)^2\}]^2}{[\mathsf{E}\{(x_i-\mu_i)^2\}]^2} &= \frac{3\,\mathsf{E}(R^4)/(p(p+2)) - 3(\mathsf{E}(R^2)/p)^2}{(\mathsf{E}(R^2)/p)^2}\\ &= 3\Big(\frac{\mathsf{E}(R^4)}{\mathsf{E}^2(R^2)}\,\frac{p}{p+2} - 1\Big) = 3\kappa.\end{aligned}$$

[2] Además, podemos escribir $\mathrm{Cov}(\boldsymbol{x}) = -2\phi'(0)\boldsymbol{\Sigma}$.

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

Estimación ML: caso independiente

Sea $\boldsymbol{x}_1, \dots, \boldsymbol{x}_n$ vectores aleatorios IID cada uno con densidad $\mathsf{EC}_p(\boldsymbol{\mu}, \boldsymbol{\Sigma}; g)$. La densidad conjunta (de la muestra) es:

$$L(\boldsymbol{\theta}) = |\boldsymbol{\Sigma}|^{-n/2} \prod_{i=1}^{n} g\big[(\boldsymbol{x}_i - \boldsymbol{\mu})^\top \boldsymbol{\Sigma}^{-1}(\boldsymbol{x}_i - \boldsymbol{\mu})\big].$$

La media muestral $\overline{\boldsymbol{x}}$ y la matriz de covarianza

$$\boldsymbol{S} = \frac{1}{n-1} \sum_{i=1}^{n} (\boldsymbol{x}_i - \overline{\boldsymbol{x}})(\boldsymbol{x}_i - \overline{\boldsymbol{x}})^\top,$$

son estimadores insesgados de $\boldsymbol{\mu}$ y

$$\boldsymbol{\Psi} = \frac{\mathsf{E}(R^2)}{p} \boldsymbol{\Sigma},$$

respectivamente, donde $R^2 = (\boldsymbol{x} - \boldsymbol{\mu})^\top \boldsymbol{\Sigma}^{-1}(\boldsymbol{x} - \boldsymbol{\mu})$.

Estimación ML: caso independiente

Por la ley de los grandes números, tenemos

$$\overline{\boldsymbol{x}} \xrightarrow{\mathsf{P}} \boldsymbol{\mu}, \qquad \boldsymbol{S} \xrightarrow{\mathsf{P}} \boldsymbol{\Sigma},$$

Además, (asintóticamente) $\mathsf{Cov}(\overline{\boldsymbol{x}}) = \frac{1}{n}\boldsymbol{\Sigma}$, mientras que

$$\begin{aligned} n\,\mathsf{Cov}(\operatorname{vec}\boldsymbol{S}) &= \mathsf{E}\{(\operatorname{vec}\boldsymbol{S} - \operatorname{vec}\boldsymbol{\Sigma})(\operatorname{vec}\boldsymbol{S} - \operatorname{vec}\boldsymbol{\Sigma})^\top\} \\ &\to (\kappa+1)(\boldsymbol{I}_{p^2} + \boldsymbol{K}_p)(\boldsymbol{\Sigma}\otimes\boldsymbol{\Sigma}) + \kappa(\operatorname{vec}\boldsymbol{\Sigma})(\operatorname{vec}\boldsymbol{\Sigma})^\top. \end{aligned}$$

Es decir,

$$\sqrt{n}\begin{pmatrix} \overline{\boldsymbol{x}} - \boldsymbol{\mu} \\ \operatorname{vec}\boldsymbol{S} - \operatorname{vec}\boldsymbol{\Sigma} \end{pmatrix} \to \mathsf{N}\Big(\begin{pmatrix} \boldsymbol{0} \\ \boldsymbol{0} \end{pmatrix}, \begin{pmatrix} \boldsymbol{\Sigma} & \boldsymbol{0} \\ \boldsymbol{0} & \boldsymbol{\Omega} \end{pmatrix}\Big),$$

donde

$$\boldsymbol{\Omega} = 2(\kappa+1)\boldsymbol{N}_p(\boldsymbol{\Sigma}\otimes\boldsymbol{\Sigma}) + \kappa(\operatorname{vec}\boldsymbol{\Sigma})(\operatorname{vec}\boldsymbol{\Sigma})^\top,$$

con $\boldsymbol{N}_p = \frac{1}{2}(\boldsymbol{I}_{p^2} + \boldsymbol{K}_p)$.

Note que

$$\mathsf{E}\{[(\boldsymbol{x}-\boldsymbol{\mu})^\top \boldsymbol{\Sigma}^{-1}(\boldsymbol{x}-\boldsymbol{\mu})]^2\} = \mathsf{E}\{(\boldsymbol{z}^\top \boldsymbol{z})^2\} = \mathsf{E}\{(R^2 \boldsymbol{u}^\top \boldsymbol{u})^2\} = \mathsf{E}(R^4)\,\mathsf{E}\{(\boldsymbol{u}^\top \boldsymbol{u})^2\}$$
$$= \frac{p^2\,\mathsf{E}(R^4)}{\mathsf{E}^2(R^2)} = p(p+2)(\kappa+1),$$

pues $\boldsymbol{z} = \boldsymbol{B}^{-1}(\boldsymbol{x}-\boldsymbol{\mu}) \sim \mathsf{S}_p(g)$ y $\boldsymbol{z} \stackrel{\mathsf{d}}{=} R\boldsymbol{u}$.

Como $\overline{\boldsymbol{x}} \xrightarrow{\mathsf{P}} \boldsymbol{\mu}$ y $\boldsymbol{S} \xrightarrow{\mathsf{P}} \boldsymbol{\Sigma}$,

$$\frac{1}{n}\sum_{i=1}^n [(\boldsymbol{x}_i - \overline{\boldsymbol{x}})^\top \boldsymbol{S}^{-1}(\boldsymbol{x}_i - \overline{\boldsymbol{x}})]^2 \xrightarrow{\mathsf{P}} p(p+2)(\kappa+1).$$

De este modo, un estimador consistente para κ es (Mardia, 1970):[3]

$$\widehat{\kappa} = \frac{1}{p(p+2)}\frac{1}{n}\sum_{i=1}^n [(\boldsymbol{x}_i - \overline{\boldsymbol{x}})^\top \boldsymbol{S}^{-1}(\boldsymbol{x}_i - \overline{\boldsymbol{x}})]^2 - 1.$$

[3] Biometrika **57**, 519-530.

Estimación ML: caso independiente

Para obtener los estimadores máximo verosímiles de $\boldsymbol{\mu}$ y $\boldsymbol{\Sigma}$ asumiremos que $g(\cdot)$ es conocido. La función de log-verosimilitud adopta la forma:

$$\ell(\boldsymbol{\theta}) = -\frac{n}{2}\log|\boldsymbol{\Sigma}| + \sum_{i=1}^{n}\log g\big[(\boldsymbol{x}_i-\boldsymbol{\mu})^\top\boldsymbol{\Sigma}^{-1}(\boldsymbol{x}_i-\boldsymbol{\mu})\big].$$

Diferenciando $\ell(\boldsymbol{\theta})$ con relación a $\boldsymbol{\mu}$, obtenemos[4]

$$\begin{aligned}
\mathsf{d}_\mu\,\ell(\boldsymbol{\theta}) &= \sum_{i=1}^{n}\mathsf{d}_\mu\log g(u_i) = \sum_{i=1}^{n}\frac{g'(u_i)}{g(u_i)}\,\mathsf{d}_\mu u_i\\
&= -\sum_{i=1}^{n}W_g(u_i)\big[(\mathsf{d}\,\boldsymbol{\mu})^\top\boldsymbol{\Sigma}^{-1}(\boldsymbol{x}_i-\boldsymbol{\mu}) + (\boldsymbol{x}_i-\boldsymbol{\mu})^\top\boldsymbol{\Sigma}^{-1}\,\mathsf{d}\,\boldsymbol{\mu}\big]\\
&= (\mathsf{d}\,\boldsymbol{\mu})^\top\boldsymbol{\Sigma}^{-1}\sum_{i=1}^{n}\omega_i(\boldsymbol{\theta})(\boldsymbol{x}_i-\boldsymbol{\mu}),
\end{aligned}$$

donde $\omega_i(\boldsymbol{\theta}) = -2W_g(u_i)$ con $W_g(u_i) = g'(u_i)/g(u_i)$, para $i=1,\dots,n$.

[4] Considere $u_i = (\boldsymbol{x}_i-\boldsymbol{\mu})^\top\boldsymbol{\Sigma}^{-1}(\boldsymbol{x}_i-\boldsymbol{\mu})$ para $i=1,\dots,n$.

Estimación ML: caso independiente

Análogamente, diferenciando $\ell(\boldsymbol{\theta})$ con relación a $\boldsymbol{\Sigma}$, tenemos

$$
\begin{aligned}
\mathsf{d}_\Sigma\, \ell(\boldsymbol{\theta}) &= -\frac{n}{2}\operatorname{tr}\boldsymbol{\Sigma}^{-1}\,\mathsf{d}\,\boldsymbol{\Sigma} + \sum_{i=1}^n W_g(u_i)(\boldsymbol{x}_i-\boldsymbol{\mu})^\top(\mathsf{d}\,\boldsymbol{\Sigma}^{-1})(\boldsymbol{x}_i-\boldsymbol{\mu}) \\
&= -\frac{n}{2}\operatorname{tr}\boldsymbol{\Sigma}^{-1}\,\mathsf{d}\,\boldsymbol{\Sigma} - \sum_{i=1}^n W_g(u_i)(\boldsymbol{x}_i-\boldsymbol{\mu})^\top\boldsymbol{\Sigma}^{-1}(\mathsf{d}\,\boldsymbol{\Sigma})\boldsymbol{\Sigma}^{-1}(\boldsymbol{x}_i-\boldsymbol{\mu}) \\
&= -\frac{n}{2}\operatorname{tr}\boldsymbol{\Sigma}^{-1}\,\mathsf{d}\,\boldsymbol{\Sigma} - \frac{1}{2}\sum_{i=1}^n \omega_i(\boldsymbol{\theta})\operatorname{tr}\boldsymbol{\Sigma}^{-1}(\boldsymbol{x}_i-\boldsymbol{\mu})(\boldsymbol{x}_i-\boldsymbol{\mu})^\top\boldsymbol{\Sigma}^{-1}\,\mathsf{d}\,\boldsymbol{\Sigma} \\
&= \frac{1}{2}\operatorname{tr}\boldsymbol{\Sigma}^{-1}\{\boldsymbol{Q}_\omega(\boldsymbol{\mu}) - n\boldsymbol{\Sigma}\}\boldsymbol{\Sigma}^{-1}\,\mathsf{d}\,\boldsymbol{\Sigma}
\end{aligned}
$$

donde

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

Estimación ML: caso independiente

La condición de primer orden, lleva al siguiente sistema de ecuaciones:

$$\sum_{i=1}^{n} \omega_i(\widehat{\boldsymbol{\theta}})(\boldsymbol{x}_i - \widehat{\boldsymbol{\mu}}) = \boldsymbol{0}$$

$$\boldsymbol{Q}_{\omega}(\widehat{\boldsymbol{\mu}}) - n\widehat{\boldsymbol{\Sigma}} = \boldsymbol{0},$$

que no tiene solución en forma explícita y por tanto métodos iterativos son requeridos.

Por ejemplo, usando una estimación inicial $\boldsymbol{\theta} = \boldsymbol{\theta}^{(k)}$, actualizamos las estimaciones para $\boldsymbol{\mu}$ y $\boldsymbol{\Sigma}$, como:

$$\boldsymbol{\mu}^{(k+1)} = \frac{1}{\sum_{i=1}^{n} \omega_i(\boldsymbol{\theta}^{(k)})} \sum_{i=1}^{n} \omega_i(\boldsymbol{\theta}^{(k)})\boldsymbol{x}_i,$$

$$\boldsymbol{\Sigma}^{(k+1)} = \frac{1}{n} \sum_{i=1}^{n} \omega_i(\boldsymbol{\theta}^{(k)})(\boldsymbol{x}_i - \boldsymbol{\mu}^{(k+1)})(\boldsymbol{x}_i - \boldsymbol{\mu}^{(k+1)})^{\top},$$

a la convergencia del algoritmo[5], hacemos $(\widehat{\boldsymbol{\mu}}, \widehat{\boldsymbol{\Sigma}})$.

[5] Esto es, cuando la secuencia $\{\boldsymbol{\mu}^{(k)}, \boldsymbol{\Sigma}^{(k)}\}$ se 'estabiliza'.

Funciones de pesos $\omega(\theta)$ para algunas distribuciones elípticas

- Normal: $\boldsymbol{x} \sim \mathsf{N}_p(\boldsymbol{\mu}, \boldsymbol{\Sigma})$, tenemos:

$$\omega_i(\boldsymbol{\theta}) = 1, \qquad i = 1, \dots, n.$$

- t-Student: $\boldsymbol{x} \sim t_p(\boldsymbol{\mu}, \boldsymbol{\Sigma}, \nu)$, $\nu > 0$,

$$\omega_i(\boldsymbol{\theta}) = \frac{\nu + p}{\nu + u_i}, \qquad i = 1, \dots, n,$$

la distribución $\mathsf{Cauchy}_p(\boldsymbol{\mu}, \boldsymbol{\Sigma})$, es obtenida para $\nu = 1$.

- Normal contaminada: $\boldsymbol{x} \sim \mathsf{CN}_p(\boldsymbol{\mu}, \boldsymbol{\Sigma}, \epsilon, \gamma)$, con $\epsilon \in [0, 1)$ y $\gamma > 0$,

$$\omega_i(\boldsymbol{\theta}) = \frac{(1-\epsilon)\exp(-u/2) + \epsilon\gamma^{-(p/2+1)}\exp(-u/(2\gamma))}{(1-\epsilon)\exp(-u/2) + \epsilon\gamma^{-p/2}\exp(-u/(2\gamma))}, \qquad i = 1, \dots, n.$$

- Logística: $\boldsymbol{x} \sim \mathsf{L}_p(\boldsymbol{\mu}, \boldsymbol{\Sigma})$, en cuyo caso,

$$\omega_i(\boldsymbol{\theta}) = 2\tanh(u_i/2), \qquad i = 1, \dots, n.$$

- Exponencial Potencia: $\boldsymbol{x} \sim \mathsf{PE}_p(\boldsymbol{\mu}, \boldsymbol{\Sigma}, \lambda)$, $\lambda > 0$,

$$\omega_i(\boldsymbol{\theta}) = \lambda u_i^{\lambda - 1}, \qquad i = 1, \dots, n.$$

EX UMBRA IN SOLEM

Ejemplo: Datos de creatinina (Shih y Weisberg, 1986)[6]

La eliminación de creatinina a través de la orina es una medida importante de la función renal.

Los datos corresponden aun estudio para determinar la eliminación de creatinina en 34 pacientes registrando las siguientes variables:

- peso corporal (WT) en kg.,
- concentración de creatinina sérica (SC) en mg/decilitro,
- Edad (Age) en años, y
- la depuración de creatinina endógena (CR).

Un modelo típico sugerido en farmacocinética es:

$$\mathsf{E}\{\log(\mathsf{CR})\} = \beta_0 + \beta_1 \log(\mathsf{WT}) + \log_2 \log(\mathsf{SC}) + \beta_3 \log(140 - \mathsf{Age}).$$

A continuación asumiremos que $(\log(\mathsf{CR}), \log(\mathsf{WT}), \log(\mathsf{SC}), \log(140 - \mathsf{Age}))^\top$ sigue una distribución normal multivariada.

[6] Technometrics **28**, 231-239.

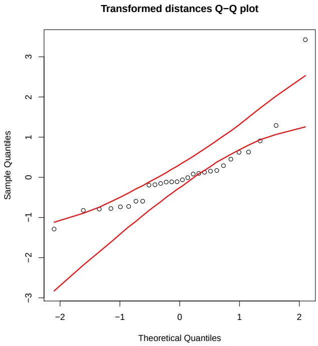
Ejemplo: Datos de creatinina

```
> library(heavy)
> data(creatinine)
> creatinine
      WT      SC Age    CR
1   71.0 0.71253  38 132.0
2   69.0 1.48161  78  53.0
3   85.0 2.20545  69  50.0
4  100.0 1.42505  70  82.0
5   59.0 0.67860  45 110.0
6   73.0 0.75777  65 100.0
7   63.0 1.11969  76  68.0
8   81.0 0.91611  61  92.0

...

26  67.0 1.19886  21  80.0
27  68.0 7.60001  81   4.3
28  72.2 2.10001  43  43.2
29    NA 1.35719  78  75.0
30    NA 1.05183  38  41.0
31 107.0      NA  62 120.0
32  75.0      NA  70  52.0
33  62.0      NA  63  73.0
34  52.0      NA  68  57.0
```

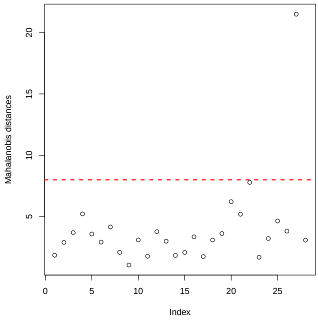
Ejemplo: Datos de creatinina

```
# ajuste con el modelo normal
> fm0 <- heavyFit(~ log(WT) + log(SC) + log(140 - Age) + log(CR),
+          data = creatinine, family = normal())

> fm0 # Salida:
Call:
heavyFit(x = ~log(WT) + log(SC) + log(140 - Age) + log(CR),
          data = creatinine, family = normal())
Converged in 1 iterations

Center:
        log(WT)        log(SC) log(140 - Age)        log(CR)
        4.2763         0.2038         4.4370         4.3069

Scatter matrix estimate:
               log(WT)     log(SC)     log(140 - Age) log(CR)
log(WT)         0.023876
log(SC)        -0.000072    0.248036
log(140 - Age) -0.002317   -0.051800   0.0455740
log(CR)         0.022111   -0.307312   0.0873372      0.457392

Number of Observations: 28
```

Ejemplo: Datos de creatinina



(a)



(b)

Ejemplo: Datos de creatinina

```
# ajuste con el modelo Student-t
> fm1 <- heavyFit(~ log(WT) + log(SC) + log(140 - Age) + log(CR),
+          data = creatinine, family = Student(df = 4))

> fm1
Call:
heavyFit(x = ~log(WT) + log(SC) + log(140 - Age) + log(CR),
         data = creatinine, family = Student(df = 6.67344))
Converged in 38 iterations

Center:
      log(WT)          log(SC) log(140 - Age)          log(CR)
      4.2872           0.1262          4.4446          4.4247

Scatter matrix estimate:
               log(WT)    log(SC)    log(140 - Age) log(CR)
log(WT)         0.021237
log(SC)         0.004514   0.138614
log(140 - Age) -0.004014  -0.030218   0.039962
log(CR)         0.012819  -0.133791   0.051319    0.180992

Number of Observations: 28
```

Debemos destacar que, en general:

$$\mathsf{Cov}(\boldsymbol{x}) = \frac{\mathsf{E}(R^2)}{p}\,\boldsymbol{\Sigma} = \boldsymbol{\Psi}.$$

Desde el procedimiento de estimación se obtuvo,[7] $\widehat{\nu} = 6.673439$. Para la distribución $t_p(\boldsymbol{\mu}, \boldsymbol{\Sigma}, \nu)$, sigue que:

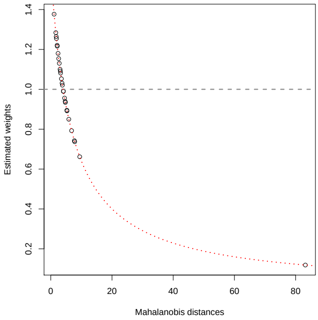
$$\widehat{\boldsymbol{\Psi}}_{\mathsf{Student}} = \frac{\widehat{\nu}}{\widehat{\nu} - 2}\widehat{\boldsymbol{\Sigma}} = \begin{pmatrix} 0.03032 & 0.00645 & -0.00573 & 0.01831 \\ 0.00645 & 0.19793 & -0.04315 & -0.19105 \\ -0.00573 & -0.04315 & 0.05706 & 0.07338 \\ 0.01831 & -0.19105 & 0.07328 & 0.25845 \end{pmatrix},$$

mientras que para la distribución $\mathsf{N}_p(\boldsymbol{\mu}, \boldsymbol{\Sigma})$, tenemos $\mathsf{Cov}(\boldsymbol{x}) = \boldsymbol{\Sigma}$ $(= \boldsymbol{\Psi})$. En efecto,

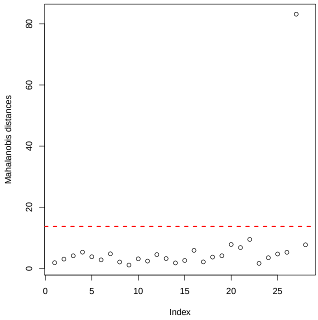
$$\widehat{\boldsymbol{\Psi}}_{\mathsf{Normal}} = \begin{pmatrix} 0.02386 & -0.00007 & -0.00232 & 0.02211 \\ -0.00007 & 0.24804 & -0.05180 & -0.30731 \\ -0.00232 & -0.05180 & 0.04557 & 0.08734 \\ 0.02211 & -0.30731 & 0.08734 & 0.45739 \end{pmatrix}.$$

[7] Para detalles sobre como estimar ν, no se pierda la próxima clase! ☺

Ejemplo: Datos de creatinina



(a)



(b)

Referencias bibliográficas

Mardia, K.V. (1970).
Measures of multivariate skewness and kurtosis with applications.
Biometrika **57**, 519-530.

Maronna, R.A. (1976).
Robust M-estimators of multivariate location and scatter.
The Annals of Statistics **4**, 51-67.

Shih, W.J., Weisberg, S. (1986).
Assessing influence in multiple linear regression with incomplete data.
Technometrics **28**, 231-239.

Xie, F.C., Wei, B.C., Lin, J.G. (207).
Case-deletion influence measures for the data from multivariate t distributions.
Journal of Applied Statistics **34**, 907-921.

EX UMBRA IN SOLEM