MAT-266: Distribución de formas cuadráticas

Felipe Osorio

fosorios.mat.utfsm.cl

Departamento de Matemática, UTFSM

Resultado 1:

Si $\boldsymbol{X} \sim \mathsf{N}_p(\boldsymbol{\mu}, \boldsymbol{I})$ y $\boldsymbol{A} \in \mathbb{R}^{p\times p}$ es matriz simétrica. Entonces $\boldsymbol{X}^\top \boldsymbol{A}\boldsymbol{X} \sim \chi^2(k;\theta)$ sólo si $\boldsymbol{A}$ es idempotente, en cuyo caso los grados de libertad y el parámetro de no centralidad están dados por

$$k = \operatorname{rg}(\boldsymbol{A}) = \operatorname{tr}(\boldsymbol{A}), \qquad \text{y} \qquad \theta = \frac{1}{2}\boldsymbol{\mu}^\top \boldsymbol{A}\boldsymbol{\mu},$$

respectivamente.

Demostración:

Suponga que $\boldsymbol{A}$ es idempotente de rango k. Entonces existe una matriz ortogonal $\boldsymbol{P}$ tal que

$$\boldsymbol{P}^\top \boldsymbol{A}\boldsymbol{P} = \begin{pmatrix} \boldsymbol{I}_k & \boldsymbol{0} \\ \boldsymbol{0} & \boldsymbol{0} \end{pmatrix}.$$

Sea $\boldsymbol{Y} = \boldsymbol{P}^\top \boldsymbol{X}$, entonces $\boldsymbol{Y} \sim \mathsf{N}_p(\boldsymbol{P}^\top \boldsymbol{\mu}, \boldsymbol{I})$, y

$$\boldsymbol{X}^\top \boldsymbol{A}\boldsymbol{X} = \boldsymbol{Y}^\top \begin{pmatrix} \boldsymbol{I}_k & \boldsymbol{0} \\ \boldsymbol{0} & \boldsymbol{0} \end{pmatrix} \boldsymbol{Y} = \sum_{i=1}^k Y_i^2,$$

que sigue una distribución chi-cuadrado con k grados de libertad.

Para el parámetro de no centralidad θ, note que

$$\begin{aligned}\mathsf{E}\{\chi^2(k;\theta)\} = k + 2\theta = \mathsf{E}(\boldsymbol{X}^\top \boldsymbol{A}\boldsymbol{X}) &= \operatorname{tr}(\mathsf{E}(\boldsymbol{X}\boldsymbol{X}^\top)\boldsymbol{A}) \\ &= \operatorname{tr}((\boldsymbol{I} + \boldsymbol{\mu}\boldsymbol{\mu}^\top)\boldsymbol{A}) = k + \boldsymbol{\mu}^\top \boldsymbol{A}\boldsymbol{\mu},\end{aligned}$$

y de ahí que $\theta = \frac{1}{2}\boldsymbol{\mu}^\top \boldsymbol{A}\boldsymbol{\mu}$.

Ahora, suponga que $\boldsymbol{X}^\top \boldsymbol{A}\boldsymbol{X} \sim \chi^2(k;\theta)$. Si $\boldsymbol{A}$ tiene rango r, entonces para $\boldsymbol{P}$ matriz ortogonal $p \times p$,

$$\boldsymbol{P}^\top \boldsymbol{A}\boldsymbol{P} = \begin{pmatrix} \boldsymbol{\Lambda}_1 & \boldsymbol{0} \\ \boldsymbol{0} & \boldsymbol{0} \end{pmatrix},$$

con $\boldsymbol{\Lambda}_1 = \operatorname{diag}(\lambda_1, \dots, \lambda_r)$, donde $\lambda_1, \dots, \lambda_r$ son los valores propios no nulos de $\boldsymbol{A}$. Sea $\boldsymbol{Y} = \boldsymbol{P}^\top \boldsymbol{X}$, entonces

$$\boldsymbol{X}^\top \boldsymbol{A}\boldsymbol{X} = \boldsymbol{Y}^\top \boldsymbol{P}^\top \boldsymbol{A}\boldsymbol{P}\boldsymbol{Y} = \sum_{j=1}^{r} \lambda_j Y_j^2 = U.$$

Tenemos que $\boldsymbol{Y} \sim \mathsf{N}_p(\boldsymbol{\delta}, \boldsymbol{I})$ con $\boldsymbol{\delta} = \boldsymbol{P}^\top \boldsymbol{\mu}$, de modo que $Y_j^2 \sim \chi^2(1, \delta_j^2/2)$ con función característica

$$\varphi_{Y_j^2}(t) = (1 - 2it)^{-1/2} \exp\Big(\frac{it\delta_j^2}{1 - 2it}\Big),$$

por la independencia de $Y_1, \dots, Y_r$ sigue que

$$\begin{aligned}
\varphi_U(t) &= \prod_{j=1}^r (1 - 2it\lambda_j)^{-1/2} \exp\Big(\frac{it\lambda_j\delta_j^2}{1 - 2it\lambda_j}\Big) \\
&= \exp\Big(it \sum_{j=1}^r \frac{\lambda_j \delta_j^2}{1 - 2it\lambda_j}\Big) \prod_{j=1}^r (1 - 2it\lambda_j)^{-1/2}.
\end{aligned}$$

Como $\boldsymbol{X}^\top \boldsymbol{A}\boldsymbol{X} \sim \chi^2(k;\theta)$ tiene función característica

$$\varphi_{X^\top AX}(t) = (1-2it)^{-k/2}\exp\Big(\frac{2it\theta}{1-2it}\Big),$$

entonces desde las dos expresiones anteriores debemos tener $r = k$, $\lambda_j = 1$, $\forall j$ y $\theta = \sum_j \delta_j^2/2$. Consecuentemente $\boldsymbol{P}^\top \boldsymbol{A}\boldsymbol{P}$ tiene la forma

$$\boldsymbol{P}^\top \boldsymbol{A}\boldsymbol{P} = \begin{pmatrix} \boldsymbol{I}_k & \boldsymbol{0} \\ \boldsymbol{0} & \boldsymbol{0} \end{pmatrix},$$

que es idempotente. Luego

$$\boldsymbol{P}^\top \boldsymbol{A}\boldsymbol{P} = (\boldsymbol{P}^\top \boldsymbol{A}\boldsymbol{P})(\boldsymbol{P}^\top \boldsymbol{A}\boldsymbol{P}) = \boldsymbol{P}^\top \boldsymbol{A}^2\boldsymbol{P} \quad\Longrightarrow\quad \boldsymbol{A}^2 = \boldsymbol{A}.$$

Resultado 2:

Si $\boldsymbol{X} \sim \mathsf{N}_p(\boldsymbol{\mu}, \boldsymbol{\Sigma})$ donde $\boldsymbol{\Sigma}$ es no singular y $\boldsymbol{X}$, $\boldsymbol{\mu}$ y $\boldsymbol{\Sigma}$ son particionados como

$$\boldsymbol{X} = \begin{pmatrix} \boldsymbol{X}_1 \\ \boldsymbol{X}_2 \end{pmatrix}, \qquad \boldsymbol{\mu} = \begin{pmatrix} \boldsymbol{\mu}_1 \\ \boldsymbol{\mu}_2 \end{pmatrix}, \qquad \boldsymbol{\Sigma} = \begin{pmatrix} \boldsymbol{\Sigma}_{11} & \boldsymbol{\Sigma}_{12} \\ \boldsymbol{\Sigma}_{21} & \boldsymbol{\Sigma}_{22} \end{pmatrix},$$

donde $\boldsymbol{X}_1$, $\boldsymbol{\mu}_1$ son $k \times 1$ y $\boldsymbol{\Sigma}_{11}$ es $k \times k$. Entonces

$$U = (\boldsymbol{X} - \boldsymbol{\mu})^\top \boldsymbol{\Sigma}^{-1}(\boldsymbol{X} - \boldsymbol{\mu}) - (\boldsymbol{X}_1 - \boldsymbol{\mu}_1)^\top \boldsymbol{\Sigma}_{11}^{-1}(\boldsymbol{X}_1 - \boldsymbol{\mu}_1) \sim \chi^2(p - k).$$

Demostración:

Considere $\boldsymbol{\Sigma} = \boldsymbol{B}\boldsymbol{B}^\top$, donde $\boldsymbol{B}$ es no singular y particione $\boldsymbol{B}$ como

$$\boldsymbol{B} = \begin{pmatrix} \boldsymbol{B}_1 \\ \boldsymbol{B}_2 \end{pmatrix}, \qquad \boldsymbol{B}_1 \in \mathbb{R}^{k \times p}.$$

Luego,

$$\boldsymbol{\Sigma} = \boldsymbol{B}\boldsymbol{B}^\top = \begin{pmatrix} \boldsymbol{B}_1\boldsymbol{B}_1^\top & \boldsymbol{B}_1\boldsymbol{B}_2^\top \\ \boldsymbol{B}_2\boldsymbol{B}_1^\top & \boldsymbol{B}_2\boldsymbol{B}_2^\top \end{pmatrix},$$

de donde sigue que $\boldsymbol{\Sigma}_{11} = \boldsymbol{B}_1\boldsymbol{B}_1^\top$. Ahora, sea $\boldsymbol{Z} = \boldsymbol{B}^{-1}(\boldsymbol{X} - \boldsymbol{\mu}) \sim \mathsf{N}_p(\boldsymbol{0}, \boldsymbol{I})$. De este modo,

$$\begin{pmatrix} \boldsymbol{B}_1 \\ \boldsymbol{B}_2 \end{pmatrix} \boldsymbol{Z} = \begin{pmatrix} \boldsymbol{X}_1 - \boldsymbol{\mu}_1 \\ \boldsymbol{X}_2 - \boldsymbol{\mu}_2 \end{pmatrix}.$$

Entonces

$$\begin{aligned} U &= \boldsymbol{Z}^\top \boldsymbol{Z} - \boldsymbol{Z}^\top \boldsymbol{B}_1^\top (\boldsymbol{B}_1 \boldsymbol{B}_1^\top)^{-1} \boldsymbol{B}_1 \boldsymbol{Z} \\ &= \boldsymbol{Z}^\top (\boldsymbol{I} - \boldsymbol{B}_1^\top (\boldsymbol{B}_1 \boldsymbol{B}_1^\top)^{-1} \boldsymbol{B}_1) \boldsymbol{Z} \\ &= \boldsymbol{Z}^\top (\boldsymbol{I} - \boldsymbol{H}_1) \boldsymbol{Z}, \end{aligned}$$

con $\boldsymbol{H}_1 = \boldsymbol{B}_1^\top (\boldsymbol{B}_1 \boldsymbol{B}_1^\top)^{-1} \boldsymbol{B}_1$.

Note que $\boldsymbol{H}_1$ es simétrica e idempotente y por tanto también lo es $\boldsymbol{C} = \boldsymbol{I} - \boldsymbol{H}_1$. De donde sigue que $U \sim \chi^2(\nu)$, con $\nu = \mathrm{rg}(\boldsymbol{C}) = p - k$.

Suponga que $\boldsymbol{X} \sim \mathsf{N}_p(\mathbf{0}, \boldsymbol{\Sigma})$. Una condición para que $\boldsymbol{X}^\top \boldsymbol{A}\boldsymbol{X}$ tenga una distribución chi-cuadrado es:[1]

$$\boldsymbol{\Sigma}\boldsymbol{A}\boldsymbol{\Sigma}\boldsymbol{A} = \boldsymbol{\Sigma}\boldsymbol{A},$$

en cuyo caso los grados de libertad son $k = \operatorname{rg}(\boldsymbol{A}\boldsymbol{\Sigma})$. Si $\boldsymbol{\Sigma}$ es no singular, la condición resulta $\boldsymbol{A}\boldsymbol{\Sigma}\boldsymbol{A} = \boldsymbol{A}$.

Resultado 3:

Si $\boldsymbol{X} \sim \mathsf{N}_p(\mathbf{0}, \boldsymbol{\Sigma})$ donde $\boldsymbol{\Sigma}$ tiene rango k $(\leq p)$ y si $\boldsymbol{A}$ es una inversa generalizada de $\boldsymbol{\Sigma}$ $(\boldsymbol{\Sigma}\boldsymbol{A}\boldsymbol{\Sigma} = \boldsymbol{\Sigma})$, entonces $\boldsymbol{X}^\top \boldsymbol{A}\boldsymbol{X} \sim \chi^2(k)$.

[1] Esto representa una generalización del Resultado 1.

Demostración:

Considere $\boldsymbol{Y} = \boldsymbol{B}\boldsymbol{X}$ donde $\boldsymbol{B}$ es una matriz no singular $p \times p$ tal que

$$\boldsymbol{B}\boldsymbol{\Sigma}\boldsymbol{B}^\top = \begin{pmatrix} \boldsymbol{I}_k & \boldsymbol{0} \\ \boldsymbol{0} & \boldsymbol{0} \end{pmatrix}.$$

Particionando $\boldsymbol{Y} = (\boldsymbol{Y}_1^\top, \boldsymbol{Y}_2^\top)^\top$ donde $\boldsymbol{Y}_1$ es un vector $k \times 1$ sigue que $\boldsymbol{Y}_1 \sim \mathsf{N}_k(\boldsymbol{0}, \boldsymbol{I})$ y $\boldsymbol{Y}_2 = \boldsymbol{0}$ con probabilidad 1.

Ahora, note que

$$\begin{aligned} \begin{pmatrix} \boldsymbol{I}_k & \boldsymbol{0} \\ \boldsymbol{0} & \boldsymbol{0} \end{pmatrix} &= \boldsymbol{B}\boldsymbol{\Sigma}\boldsymbol{B}^\top = \boldsymbol{B}\boldsymbol{\Sigma}\boldsymbol{A}\boldsymbol{\Sigma}\boldsymbol{B}^\top \\ &= \boldsymbol{B}\boldsymbol{\Sigma}\boldsymbol{B}^\top\boldsymbol{B}^{-\top}\boldsymbol{A}\boldsymbol{B}^{-1}\boldsymbol{B}\boldsymbol{\Sigma}\boldsymbol{B}^\top \\ &= \begin{pmatrix} \boldsymbol{I}_k & \boldsymbol{0} \\ \boldsymbol{0} & \boldsymbol{0} \end{pmatrix} \boldsymbol{B}^{-\top}\boldsymbol{A}\boldsymbol{B}^{-1} \begin{pmatrix} \boldsymbol{I}_k & \boldsymbol{0} \\ \boldsymbol{0} & \boldsymbol{0} \end{pmatrix}. \end{aligned}$$

Luego, con probabilidad uno,

$$\begin{aligned}
\boldsymbol{X}^\top \boldsymbol{A}\boldsymbol{X} &= \boldsymbol{Y}^\top \boldsymbol{B}^{-\top}\boldsymbol{A}\boldsymbol{B}^{-1}\boldsymbol{Y} = (\boldsymbol{Y}_1^\top, \boldsymbol{0})\boldsymbol{B}^{-\top}\boldsymbol{A}\boldsymbol{B}^{-1}\begin{pmatrix}\boldsymbol{Y}_1\\ \boldsymbol{0}\end{pmatrix}\\
&= (\boldsymbol{Y}_1^\top, \boldsymbol{0})\begin{pmatrix}\boldsymbol{I}_k & \boldsymbol{0}\\ \boldsymbol{0} & \boldsymbol{0}\end{pmatrix}\boldsymbol{B}^{-\top}\boldsymbol{A}\boldsymbol{B}^{-1}\begin{pmatrix}\boldsymbol{I}_k & \boldsymbol{0}\\ \boldsymbol{0} & \boldsymbol{0}\end{pmatrix}\begin{pmatrix}\boldsymbol{Y}_1\\ \boldsymbol{0}\end{pmatrix}\\
&= (\boldsymbol{Y}_1^\top, \boldsymbol{0})\begin{pmatrix}\boldsymbol{I}_k & \boldsymbol{0}\\ \boldsymbol{0} & \boldsymbol{0}\end{pmatrix}\begin{pmatrix}\boldsymbol{Y}_1\\ \boldsymbol{0}\end{pmatrix} = \boldsymbol{Y}_1^\top \boldsymbol{Y}_1 \sim \chi^2(k).
\end{aligned}$$

Resultado 4:

Si $\boldsymbol{X} \sim \mathsf{N}_p(\boldsymbol{\mu}, \boldsymbol{\Sigma})$, donde $\boldsymbol{\Sigma}$ es no singular, y $\boldsymbol{A}$ es una matriz simétrica $p \times p$. Entonces $\boldsymbol{X}^\top \boldsymbol{A}\boldsymbol{X} \sim \chi^2(k; \lambda)$, donde $k = \operatorname{rg}(\boldsymbol{A})$, $\lambda = \boldsymbol{\mu}^\top \boldsymbol{A}\boldsymbol{\mu}/2$ si y sólo si $\boldsymbol{A}\boldsymbol{\Sigma}$ es matriz idempotente.

Demostración:

Considere $\boldsymbol{Y} = \boldsymbol{B}\boldsymbol{X}$, donde $\boldsymbol{B}$ es una matriz no singular $p \times p$ tal que $\boldsymbol{B}\boldsymbol{\Sigma}\boldsymbol{B}^\top = \boldsymbol{I}_p$. Entonces

$$\boldsymbol{X}^\top \boldsymbol{A}\boldsymbol{X} = \boldsymbol{Y}^\top \boldsymbol{B}^{-\top} \boldsymbol{A}\boldsymbol{B}^{-1}\boldsymbol{Y},$$

donde $\boldsymbol{Y} \sim \mathsf{N}_p(\boldsymbol{B}\boldsymbol{\mu}, \boldsymbol{I})$. Desde el Resultado 1 sigue que $\boldsymbol{X}^\top \boldsymbol{A}\boldsymbol{X}$ tiene distribución chi-cuadrado sólo si $\boldsymbol{B}^{-\top}\boldsymbol{A}\boldsymbol{B}^{-1}$ es idempotente. Esto es equivalente a mostrar que $\boldsymbol{A}\boldsymbol{\Sigma}$ es idempotente.

Si $\boldsymbol{A\Sigma}$ es idempotente, tenemos

$$\boldsymbol{A} = \boldsymbol{A\Sigma A} = \boldsymbol{A}\boldsymbol{B}^{-1}\boldsymbol{B}^{-\top}\boldsymbol{A}, \qquad (\boldsymbol{\Sigma} = \boldsymbol{B}^{-1}\boldsymbol{B}^{-\top})$$

así, pre- y post-multiplicando por $\boldsymbol{B}^{-\top}$ y $\boldsymbol{B}^{-1}$, obtenemos

$$\boldsymbol{B}^{-\top}\boldsymbol{A}\boldsymbol{B}^{-1} = (\boldsymbol{B}^{-\top}\boldsymbol{A}\boldsymbol{B}^{-1})(\boldsymbol{B}^{-\top}\boldsymbol{A}\boldsymbol{B}^{-1}),$$

y por tanto es idempotente.

Por otro lado, si $\boldsymbol{B}^{-\top}\boldsymbol{A}\boldsymbol{B}^{-1}$ es idempotente, entonces

$$\boldsymbol{B}^{-\top}\boldsymbol{A}\boldsymbol{B}^{-1} = (\boldsymbol{B}^{-\top}\boldsymbol{A}\boldsymbol{B}^{-1})(\boldsymbol{B}^{-\top}\boldsymbol{A}\boldsymbol{B}^{-1}) = \boldsymbol{B}^{-\top}\boldsymbol{A\Sigma A}\boldsymbol{B}^{-1},$$

es decir $\boldsymbol{A} = \boldsymbol{A\Sigma A}$ y de ahí que $\boldsymbol{A\Sigma}$ es idempotente.

Ejemplo:

Sea $X_1, \dots, X_n$ variables aleatorias IID $\mathsf{N}(\theta, \sigma^2)$, en este caso podemos definir $\boldsymbol{X} = (X_1, \dots, X_n)^\top$ tal que $\boldsymbol{X} \sim \mathsf{N}_n(\theta \mathbf{1}_n, \sigma^2 \boldsymbol{I})$. Considere la forma cuadrática

$$Q = \frac{1}{\sigma^2} \sum_{i=1}^{n} (X_i - \overline{X})^2 = \frac{1}{\sigma^2} \boldsymbol{X}^\top \boldsymbol{C} \boldsymbol{X} = \boldsymbol{X}^\top \boldsymbol{A} \boldsymbol{X},$$

con $\boldsymbol{C} = \boldsymbol{I}_n - \frac{1}{n} \mathbf{1}\mathbf{1}^\top$ y $\boldsymbol{A} = \boldsymbol{C}/\sigma^2$. De esta manera

$$\boldsymbol{A}\boldsymbol{\Sigma} = \boldsymbol{I}_n - \frac{1}{n} \mathbf{1}\mathbf{1}^\top,$$

que es idempotente. En efecto,

$$\boldsymbol{C}^2 = \Big(\boldsymbol{I}_n - \frac{1}{n} \mathbf{1}\mathbf{1}^\top\Big)\Big(\boldsymbol{I}_n - \frac{1}{n} \mathbf{1}\mathbf{1}^\top\Big) = \boldsymbol{I}_n - \frac{1}{n} \mathbf{1}\mathbf{1}^\top - \frac{1}{n} \mathbf{1}\mathbf{1}^\top + \frac{1}{n^2} \mathbf{1}\mathbf{1}^\top \mathbf{1}\mathbf{1}^\top = \boldsymbol{C}.$$

Además

$$\mathrm{rg}(\boldsymbol{A}) = \mathrm{rg}(\boldsymbol{C}) = \mathrm{tr}\left(\boldsymbol{I}_n - \frac{1}{n}\mathbf{1}\mathbf{1}^\top\right) = n - 1,$$

y

$$\lambda = \frac{\theta^2}{2}\mathbf{1}^\top\boldsymbol{A}\mathbf{1} = \frac{\theta^2}{2\sigma^2}\mathbf{1}^\top\Big(\boldsymbol{I}_n - \frac{1}{n}\mathbf{1}\mathbf{1}^\top\Big)\mathbf{1} = 0.$$

Finalmente,

$$Q = \frac{1}{\sigma^2}\sum_{i=1}^n (X_i - \overline{X})^2 \sim \chi^2(n-1).$$