IECD-325: Formas cuadráticas

Felipe Osorio

felipe.osorio@uv.cl

Resultado 1:

Sea $\boldsymbol{Z} \sim \mathsf{N}_p(\boldsymbol{0}, \boldsymbol{I})$ y sea $U = \boldsymbol{Z}^\top \boldsymbol{Z}$. Entonces $U \sim \chi^2(p)$, con función de densidad

$$f(u) = \frac{1}{2^{p/2}\Gamma(p/2)}\, u^{p/2-1} \exp(-u/2), \qquad u > 0.$$

Demostración:

Como U es una función de variables aleatorias normales, entonces su función característica asume la forma

$$\begin{aligned}\varphi_U(t) = \mathsf{E}\{\exp(itU)\} &= \int_{\mathbb{R}^p} \exp(itu)(2\pi)^{-p/2} \exp(-\tfrac{1}{2}\boldsymbol{z}^\top \boldsymbol{z})\, \mathrm{d}\boldsymbol{z} \\ &= (2\pi)^{-p/2} \int_{\mathbb{R}^p} \exp(-\tfrac{1}{2}(1-2it)\boldsymbol{z}^\top \boldsymbol{z})\, \mathrm{d}\boldsymbol{z} = (1-2it)^{-p/2},\end{aligned}$$

que corresponde a la función característica de una variable aleatoria chi-cuadrado con p grados de libertad.

Definición 1 (Distribución chi-cuadrado no central):

Si $\boldsymbol{Y} \sim \mathsf{N}_p(\boldsymbol{\mu}, \boldsymbol{I})$, entonces $U = \boldsymbol{Y}^\top \boldsymbol{Y}$ tiene distribución chi-cuadrado no central con p grados de libertad y parámetro de no centralidad $\lambda = \boldsymbol{\mu}^\top \boldsymbol{\mu}/2$, en cuyo caso anotamos $U \sim \chi^2(p; \lambda)$.

Resultado 2:

Sea $\boldsymbol{Y} \sim \mathsf{N}_p(\boldsymbol{\mu}, \boldsymbol{I})$ donde $\boldsymbol{\mu} = (\mu_1, \dots, \mu_p) \neq \boldsymbol{0}$ y sea $U = \boldsymbol{Y}^\top \boldsymbol{Y}$. Entonces la función característica de U es dada por

$$\varphi_U(t) = (1 - 2it)^{-p/2} \exp\Big(\frac{2it\lambda}{1 - 2it}\Big),$$

con $\lambda = \boldsymbol{\mu}^\top \boldsymbol{\mu}/2$.

Chi-cuadrado no central

Demostración:

Como $Y_1, \dots, Y_p$ son variables aleatorias independientes, tenemos

$$\varphi_U(t) = \prod_{j=1}^{p} \varphi_{Y_j^2}(t).$$

Ahora, la función característica asociada a la variable aleatoria Y_j^2 es dada por

$$\begin{aligned}\varphi_{Y_j^2}(t) &= \int_{-\infty}^{\infty} \exp(ity_j^2)(2\pi)^{-1/2} \exp\{-\tfrac{1}{2}(y_j - \mu_j)^2\} \, \mathrm{d}y_j \\ &= \exp\Big\{\frac{\mu_j^2}{2}\Big(\frac{1}{1-2it}\Big) - \frac{\mu_j^2}{2}\Big\} \int_{-\infty}^{\infty} (2\pi)^{-1/2} \exp\Big\{-\frac{(1-2it)}{2}\Big(y_j - \frac{\mu_j}{1-2it}\Big)^2\Big\} \, \mathrm{d}y_j\end{aligned}$$

de este modo,

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

y por tanto la función característica de la variable $U = \sum_{j=1}^{p} Y_j^2$, asume la forma

$$\varphi_U(t) = (1 - 2it)^{-p/2} \exp\Big(\frac{2it\lambda}{1-2it}\Big), \qquad \lambda = \boldsymbol{\mu}^\top \boldsymbol{\mu}/2.$$

Observación:

La función característica de la variable $U = \boldsymbol{Y}^\top \boldsymbol{Y}$, puede ser escrita como

$$\begin{aligned}
\varphi_U(t) &= (1-2it)^{-p/2} \exp\Big(\frac{\lambda}{1-2it} - \lambda\Big) \\
&= (1-2it)^{-p/2} e^{-\lambda} \sum_{k=0}^{\infty} \frac{\{\lambda/(1-2it)\}^k}{k!} \\
&= \sum_{k=0}^{\infty} \frac{e^{-\lambda}\lambda^k}{k!} (1-2it)^{-(p+2k)/2}.
\end{aligned}$$

Es decir, la función característica de U es un promedio ponderado con pesos Poisson de funciones características de variables aleatorias chi-cuadrado con $p+2k$ grados de libertad.

Usando la relación entre funciones características y sus correspondientes funciones de densidad, sigue que la chi-cuadrado no central tiene la siguiente representación:

$$U|Z \sim \chi^2(p+2z), \qquad Z \sim \text{Poisson}(\lambda), \tag{1}$$

con densidad

$$f(u) = \sum_{k=0}^{\infty} \frac{e^{-\lambda}\lambda^k}{k!} \, \frac{1}{2^{p/2+k}\Gamma(\frac{p}{2}+k)} \, u^{p/2+k-1} \exp(-u/2), \quad u > 0.$$

La representación en (1) es muy útil para obtener los momentos de una variable aleatoria con distribución chi-cuadrado no central.

Chi-cuadrado no central

El valor esperado de $U \sim \chi^2(p; \lambda)$ es dado por

$$\begin{aligned} \mathsf{E}(U) &= \mathsf{E}\{\mathsf{E}(U|Z)\} = \mathsf{E}(p + 2Z) \\ &= p + 2\,\mathsf{E}(Z) = p + 2\lambda, \end{aligned}$$

mientras que la varianza de U puede ser calculada como

$$\begin{aligned} \mathsf{var}(U) &= \mathsf{E}\{\mathsf{var}(U|Z)\} + \mathsf{var}\{\mathsf{E}(U|Z)\} \\ &= \mathsf{E}\{2(p + 2Z)\} + \mathsf{var}(p + 2Z) \\ &= 2p + 4\lambda + 4\lambda = 2p + 8\lambda. \end{aligned}$$

Resultado 3:

Si $\boldsymbol{X} \sim \mathsf{N}_p(\boldsymbol{\mu}, \boldsymbol{\Sigma})$ donde $\boldsymbol{\Sigma}$ es matriz no singular. Entonces

(a) $(\boldsymbol{X} - \boldsymbol{\mu})^\top \boldsymbol{\Sigma}^{-1} (\boldsymbol{X} - \boldsymbol{\mu}) \sim \chi^2(p)$.

(b) $\boldsymbol{X}^\top \boldsymbol{\Sigma}^{-1} \boldsymbol{X} \sim \chi^2(p; \lambda)$, donde $\lambda = \frac{1}{2} \boldsymbol{\mu}^\top \boldsymbol{\Sigma}^{-1} \boldsymbol{\mu}$.

Demostración:

Considere $\boldsymbol{\Sigma} = \boldsymbol{B}\boldsymbol{B}^\top$ con $\boldsymbol{B}$ no singular. Para probar (a), tome

$$\boldsymbol{Z} = \boldsymbol{B}^{-1}(\boldsymbol{X} - \boldsymbol{\mu}),$$

luego $\boldsymbol{Z} \sim \mathsf{N}_p(\boldsymbol{0}, \boldsymbol{I})$ y de este modo

$$(\boldsymbol{X} - \boldsymbol{\mu})^\top \boldsymbol{\Sigma}^{-1} (\boldsymbol{X} - \boldsymbol{\mu}) = \boldsymbol{Z}^\top \boldsymbol{Z} \sim \chi^2(p; 0).$$

Chi-cuadrado no central

Para probar (b), sea $\boldsymbol{Y} = \boldsymbol{B}^{-1}\boldsymbol{X}$, luego

$$\boldsymbol{Y} \sim \mathsf{N}_p(\boldsymbol{B}^{-1}\boldsymbol{\mu}, \boldsymbol{I}),$$

y

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

que por definición tiene una distribución chi-cuadrado no central, con parámetro de no centralidad

$$\lambda = \frac{1}{2}(\boldsymbol{B}^{-1}\boldsymbol{\mu})^\top (\boldsymbol{B}^{-1}\boldsymbol{\mu}) = \frac{1}{2}\boldsymbol{\mu}^\top \boldsymbol{\Sigma}^{-1}\boldsymbol{\mu}.$$

Definición 2 (Distribución F no central):

Sea $X_1 \sim \chi^2(\nu_1; \lambda)$ y $X_2 \sim \chi^2(\nu_2)$ variables aleatorias independientes. Entonces,

$$F = \frac{X_1/\nu_1}{X_2/\nu_2} \sim \mathsf{F}(\nu_1, \nu_2, \lambda),$$

es decir F sigue una distribución F no central con ν_1 y ν_2 grados de libertad y parámetro de no centralidad λ.

Definición 3 (Distribución Beta no central):

Considere $U_1 \sim \chi^2(\nu_1; \lambda)$, $U_2 \sim \chi^2(\nu_2)$ tal que U_1 y U_2 son variables aleatorias independientes. Entonces,

$$G = \frac{U_1}{U_1 + U_2} \sim \mathsf{Beta}(\nu_1, \nu_2, \lambda),$$

esto es, G sigue una distribución Beta no central con parámetros de forma y escala ν_1 y ν_2, respectivamente y parámetro de no centralidad λ.

Definición 4 (Distribución t de Student no central):

Si $Y \sim \mathsf{N}(\mu, \sigma^2)$ y $U/\sigma^2 \sim \chi^2(\nu)$ son independientes, entonces

$$T = \frac{Y}{\sqrt{U/\nu}} \sim t(\nu; \lambda), \quad \lambda = \mu/\sigma,$$

es llamada una variable aleatoria con distribución t de Student no central con ν grados de libertad y parámetro de no centralidad λ.

Observación:

Si $Z \sim \mathsf{N}(0, 1)$, $U \sim \chi^2(\nu)$, δ es una constante, y Z es independiente de U, entonces

$$T = \frac{Z + \delta}{\sqrt{U/\nu}} \sim t(\nu; \delta).$$

Además,

$$t^2(\nu; \lambda) \stackrel{\mathsf{d}}{=} \mathsf{F}(1, \nu, \lambda^2/2).$$

Resultado 4:

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 5:

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 $\nu = \operatorname{rg}(\boldsymbol{C}) = p - k$.

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 6:

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 4.

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 7:

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 $\boldsymbol{X} \sim \mathsf{N}_n(\theta \boldsymbol{1}_n, \sigma^2 \boldsymbol{I})$. 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 $\boldsymbol{C} = \boldsymbol{I}_n - \frac{1}{n} \boldsymbol{1}\boldsymbol{1}^\top$ y $\boldsymbol{A} = \boldsymbol{C}/\sigma^2$. De esta manera

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

que es idempotente. En efecto,

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

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,

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

Resultado 8:

Sea $\boldsymbol{X} \sim \mathsf{N}_p(\boldsymbol{\mu}, \boldsymbol{\Sigma})$, $Q_1 = \boldsymbol{X}^\top \boldsymbol{A}\boldsymbol{X}$ y $Q_2 = \boldsymbol{X}^\top \boldsymbol{B}\boldsymbol{X}$. Entonces Q_1 y Q_2 son independientes si y sólo si $\boldsymbol{A}\boldsymbol{\Sigma}\boldsymbol{B} = \boldsymbol{0}$.

Demostración:

Tenemos $\boldsymbol{\Sigma} = \boldsymbol{T}\boldsymbol{T}^\top$, y defina $\boldsymbol{G}_1 = \boldsymbol{T}^\top \boldsymbol{A}\boldsymbol{T}$, $\boldsymbol{G}_2 = \boldsymbol{T}^\top \boldsymbol{B}\boldsymbol{T}$. Note que si $\boldsymbol{A}\boldsymbol{\Sigma}\boldsymbol{B} = \boldsymbol{0}$, entonces

$$\boldsymbol{G}_1\boldsymbol{G}_2 = (\boldsymbol{T}^\top \boldsymbol{A}\boldsymbol{T})(\boldsymbol{T}^\top \boldsymbol{B}\boldsymbol{T}) = \boldsymbol{T}^\top \boldsymbol{A}\boldsymbol{\Sigma}\boldsymbol{B}\boldsymbol{T} = \boldsymbol{0}.$$

Debido a la simetría de $\boldsymbol{G}_1$ y $\boldsymbol{G}_2$, sigue que

$$\boldsymbol{0} = (\boldsymbol{G}_1\boldsymbol{G}_2)^\top = \boldsymbol{G}_2^\top \boldsymbol{G}_1^\top = \boldsymbol{G}_2\boldsymbol{G}_1.$$

Como $\boldsymbol{G}_1\boldsymbol{G}_2 = \boldsymbol{G}_2\boldsymbol{G}_1$ existe una matriz ortogonal $\boldsymbol{P}$ que simultáneamente diagonaliza $\boldsymbol{G}_1$ y $\boldsymbol{G}_2$, esto es:

$$\boldsymbol{P}^\top \boldsymbol{G}_1 \boldsymbol{P} = \boldsymbol{P}^\top \boldsymbol{T}^\top \boldsymbol{A}\boldsymbol{T}\boldsymbol{P} = \boldsymbol{D}_1,$$
$$\boldsymbol{P}^\top \boldsymbol{G}_2 \boldsymbol{P} = \boldsymbol{P}^\top \boldsymbol{T}^\top \boldsymbol{B}\boldsymbol{T}\boldsymbol{P} = \boldsymbol{D}_2.$$

De este modo,

$$\boldsymbol{0} = \boldsymbol{G}_1\boldsymbol{G}_2 = \boldsymbol{P}\boldsymbol{D}_1\boldsymbol{P}^\top \boldsymbol{P}\boldsymbol{D}_2\boldsymbol{P}^\top = \boldsymbol{P}\boldsymbol{D}_1\boldsymbol{D}_2\boldsymbol{P}^\top$$

lo que es verdad si $\boldsymbol{D}_1\boldsymbol{D}_2 = \boldsymbol{0}$. Como $\boldsymbol{D}_1$ y $\boldsymbol{D}_2$ son diagonales, sus elementos diagonales deben ocurrir en posiciones diferentes. Es decir,

$$\boldsymbol{D}_1 = \begin{pmatrix} \boldsymbol{M}_1 & \boldsymbol{0} \\ \boldsymbol{0} & \boldsymbol{0} \end{pmatrix}, \qquad \boldsymbol{D}_2 = \begin{pmatrix} \boldsymbol{0} & \boldsymbol{0} \\ \boldsymbol{0} & \boldsymbol{M}_2 \end{pmatrix}.$$

Sea $\boldsymbol{Y} = \boldsymbol{P}^\top \boldsymbol{T}^{-1} \boldsymbol{X}$, entonces

$$\begin{aligned} Q_1 &= \boldsymbol{X}^\top \boldsymbol{A} \boldsymbol{X} = \boldsymbol{X}^\top \boldsymbol{T}^{-\top} \boldsymbol{P} \boldsymbol{P}^\top \boldsymbol{T}^\top \boldsymbol{A} \boldsymbol{T} \boldsymbol{P} \boldsymbol{P}^\top \boldsymbol{T}^{-1} \boldsymbol{X} = \boldsymbol{Y}^\top \boldsymbol{D}_1 \boldsymbol{Y}, \\ Q_2 &= \boldsymbol{X}^\top \boldsymbol{A} \boldsymbol{X} = \boldsymbol{X}^\top \boldsymbol{T}^{-\top} \boldsymbol{P} \boldsymbol{P}^\top \boldsymbol{T}^\top \boldsymbol{B} \boldsymbol{T} \boldsymbol{P} \boldsymbol{P}^\top \boldsymbol{T}^{-1} \boldsymbol{X} = \boldsymbol{Y}^\top \boldsymbol{D}_2 \boldsymbol{Y}. \end{aligned}$$

Además,

$$\mathsf{Cov}(\boldsymbol{Y}) = \mathsf{Cov}(\boldsymbol{P}^\top \boldsymbol{T}^{-1} \boldsymbol{X}) = \boldsymbol{P}^\top \boldsymbol{T}^{-1} \mathsf{Cov}(\boldsymbol{X}) \boldsymbol{T}^{-\top} \boldsymbol{P} = \boldsymbol{I}.$$

En efecto, $\boldsymbol{Y} \sim \mathsf{N}_p(\boldsymbol{P}^\top \boldsymbol{T}^{-1} \boldsymbol{\mu}, \boldsymbol{I})$.

Ahora, particionando adecuadamente $\boldsymbol{Y}$, sigue que

$$\boldsymbol{Y}^\top \boldsymbol{D}_1 \boldsymbol{Y} = (\boldsymbol{Y}_1^\top, \boldsymbol{Y}_2^\top) \begin{pmatrix} \boldsymbol{M}_1 & \boldsymbol{0} \\ \boldsymbol{0} & \boldsymbol{0} \end{pmatrix} \begin{pmatrix} \boldsymbol{Y}_1 \\ \boldsymbol{Y}_2 \end{pmatrix} = \boldsymbol{Y}_1^\top \boldsymbol{M}_1 \boldsymbol{Y}_1,$$
$$\boldsymbol{Y}^\top \boldsymbol{D}_2 \boldsymbol{Y} = (\boldsymbol{Y}_1^\top, \boldsymbol{Y}_2^\top) \begin{pmatrix} \boldsymbol{0} & \boldsymbol{0} \\ \boldsymbol{0} & \boldsymbol{M}_2 \end{pmatrix} \begin{pmatrix} \boldsymbol{Y}_1 \\ \boldsymbol{Y}_2 \end{pmatrix} = \boldsymbol{Y}_2^\top \boldsymbol{M}_2 \boldsymbol{Y}_2,$$

y la independencia entre Q_1 y Q_2 sigue desde la independencia entre $\boldsymbol{Y}_1$ y $\boldsymbol{Y}_2$.

Resultado 9:

Sea $\boldsymbol{X} \sim \mathsf{N}_p(\boldsymbol{\mu}, \boldsymbol{\Sigma})$, $Q = \boldsymbol{X}^\top \boldsymbol{A}\boldsymbol{X}$ y $\boldsymbol{U} = \boldsymbol{B}\boldsymbol{X}$. Entonces Q y $\boldsymbol{U}$ son independientes si y sólo si $\boldsymbol{B}\boldsymbol{\Sigma}\boldsymbol{A} = \boldsymbol{0}$.

Ejemplo:

Considere $X_1, \dots, X_n$ muestra aleatoria desde $\mathsf{N}(\theta, \sigma^2)$, así

$$\boldsymbol{X} = (X_1, \dots, X_n)^\top \sim \mathsf{N}_n(\theta\boldsymbol{1}, \sigma^2\boldsymbol{I}_n).$$

Tenemos

$$\overline{X} = \frac{1}{n}\sum_{i=1}^{n} X_i = \frac{1}{n}\boldsymbol{1}^\top\boldsymbol{X}, \qquad S^2 = \frac{1}{n-1}\sum_{i=1}^{n}(X_i - \overline{X})^2 = \frac{1}{n-1}\boldsymbol{X}^\top\boldsymbol{C}\boldsymbol{X}.$$

Como $\boldsymbol{C}\boldsymbol{1} = \boldsymbol{0}$ sigue la independencia entre $\overline{X}$ y S^2.

Recordatorio 1:

Suponga $\boldsymbol{A}$ matriz $m \times m$, simétrica e idempotente. Entonces,

(a) $a_{ii} \geq 0$, $i = 1, \dots, n$.

(b) $a_{ii} \leq 1$, $i = 1, \dots, n$.

(c) $a_{ij} = a_{ji} = 0$, para todo $j \neq i$, si $a_{ii} = 0$ o $a_{ii} = 1$.

Demostración:

Como $\boldsymbol{A}$ es simétrica e idempotente, tenemos

$$\boldsymbol{A} = \boldsymbol{A}^2 = \boldsymbol{A}^\top \boldsymbol{A},$$

de ahí que

$$a_{ii} = \sum_{j=1}^{m} a_{ji}^2,$$

que claramente es no negativo.

Además, podemos escribir

$$a_{ii} = a_{ii}^2 + \sum_{j \neq i} a_{ji}^2.$$

Por tanto, $a_{ii} \geq a_{ii}^2$ y de este modo (b) es satisfecha.

Si $a_{ii} = 0$ o $a_{ii} = 1$, entonces $a_{ii} = a_{ii}^2$ y debemos tener

$$\sum_{j \neq i} a_{ji}^2 = 0,$$

lo que junto con la simetría de $\boldsymbol{A}$, establece (c).

Lema 1:

Sean $\boldsymbol{A}_1, \dots, \boldsymbol{A}_k$ matrices $m \times m$ simétricas e idempotentes y suponga que

$$\boldsymbol{A}_1 + \cdots + \boldsymbol{A}_k = \boldsymbol{I}_k.$$

Entonces $\boldsymbol{A}_i \boldsymbol{A}_j = \boldsymbol{0}$ para todo $i \neq j$.

Demostración:

Considere cualquiera de esas matrices, digamos $\boldsymbol{A}_h$ y denote su rango por r. Como $\boldsymbol{A}_h$ es simétrica e idempotente, existe una matriz ortogonal $\boldsymbol{P}$ tal que

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

Para $j \neq h$, defina $\boldsymbol{B}_j = \boldsymbol{P}^\top \boldsymbol{A}_j \boldsymbol{P}$, y note que

$$\boldsymbol{I}_m = \boldsymbol{P}^\top \boldsymbol{P} = \boldsymbol{P}^\top \Big(\sum_{j=1}^k \boldsymbol{A}_j \Big) \boldsymbol{P} = \sum_{j=1}^k \boldsymbol{P}^\top \boldsymbol{A}_j \boldsymbol{P} = \begin{pmatrix} \boldsymbol{I}_r & \boldsymbol{0} \\ \boldsymbol{0} & \boldsymbol{0} \end{pmatrix} + \sum_{j \neq h} \boldsymbol{B}_j.$$

O equivalentemente,

$$\sum_{j\neq h} \boldsymbol{B}_j = \begin{pmatrix} \mathbf{0} & \mathbf{0} \\ \mathbf{0} & \boldsymbol{I}_{m-r}. \end{pmatrix}$$

Claramente, dado que $\boldsymbol{A}_j$ es simétrica e idempotente, sigue que $\boldsymbol{B}_j$ también lo es. Por Recordatorio 1, sus elementos diagonales son no negativos. Además, $(\boldsymbol{B}_j)_{ll} = 0$ para $l = 1, \dots, r$. Así, por la parte (c) del Recordatorio 1 sigue que $\boldsymbol{B}_j$ debe ser de la forma

$$\boldsymbol{B}_j = \begin{pmatrix} \mathbf{0} & \mathbf{0} \\ \mathbf{0} & \boldsymbol{C}_j \end{pmatrix},$$

donde $\boldsymbol{C}_j$ es matriz $(m-r) \times (m-r)$, simétrica e idempotente. Ahora, para cualquier $j \neq h$.

$$\boldsymbol{P}^\top \boldsymbol{A}_h \boldsymbol{A}_j \boldsymbol{P} = (\boldsymbol{P}^\top \boldsymbol{A}_h \boldsymbol{P})(\boldsymbol{P}^\top \boldsymbol{A}_j \boldsymbol{P}) = \begin{pmatrix} \boldsymbol{I}_r & \mathbf{0} \\ \mathbf{0} & \mathbf{0} \end{pmatrix} \begin{pmatrix} \mathbf{0} & \mathbf{0} \\ \mathbf{0} & \boldsymbol{C}_j \end{pmatrix} = \mathbf{0},$$

lo que es verdad, sólo si $\boldsymbol{A}_h \boldsymbol{A}_j = \mathbf{0}$, pues $\boldsymbol{P}$ es no singular. Notando que h es arbitrareo, la prueba es completa.

Lema 2:

Sean $\boldsymbol{A}_1, \dots, \boldsymbol{A}_k$ matrices simétricas de orden $m \times m$ y defina

$$\boldsymbol{A} = \boldsymbol{A}_1 + \boldsymbol{A}_2 + \cdots + \boldsymbol{A}_k.$$

Considere las siguientes afirmaciones,

(a) $\boldsymbol{A}_i$ es idempotente, para $i = 1, \dots, k$.

(b) $\boldsymbol{A}$ es idempotente.

(c) $\boldsymbol{A}_i \boldsymbol{A}_j = \boldsymbol{0}$, para todo $i \neq j$.

Entonces si dos condiciones son satisfechas, la tercera condición debe ser verdadera.

Demostración:

Primero mostraremos que (a) y (b) implica (c). Como $\boldsymbol{A}$ es simétrica e idempotente, existe una matriz ortogonal $\boldsymbol{P}$ tal que

$$\boldsymbol{P}^\top \boldsymbol{A} \boldsymbol{P} = \boldsymbol{P}^\top (\boldsymbol{A}_1 + \cdots + \boldsymbol{A}_k) \boldsymbol{P} = \begin{pmatrix} \boldsymbol{I}_r & \boldsymbol{0} \\ \boldsymbol{0} & \boldsymbol{0} \end{pmatrix}, \qquad (1)$$

donde $r = \operatorname{rg}(\boldsymbol{A})$.

Sea $\boldsymbol{B}_i = \boldsymbol{P}^\top \boldsymbol{A}_i \boldsymbol{P}$, para $i = 1, \dots, k$. y note que $\boldsymbol{B}_i$ es simétrica e idempotente. Por el Recordatorio 1, tenemos que $\boldsymbol{B}_i$ debe ser de la forma

$$\boldsymbol{B}_i = \begin{pmatrix} \boldsymbol{C}_i & \boldsymbol{0} \\ \boldsymbol{0} & \boldsymbol{0} \end{pmatrix},$$

donde la matriz $r \times r$, $\boldsymbol{C}_i$ debe ser simétrica e idempotente. Por (1), tenemos

$$\boldsymbol{C}_1 + \cdots + \boldsymbol{C}_k = \boldsymbol{I}_r$$

Por el Lema 1, sigue que $\boldsymbol{C}_i \boldsymbol{C}_j = \boldsymbol{0}$ para $i \neq j$, de donde obtenemos $\boldsymbol{B}_i \boldsymbol{B}_j = \boldsymbol{0}$ y de ahí que $\boldsymbol{A}_i \boldsymbol{A}_j = \boldsymbol{0}$, para $i \neq j$.

(a) y (c) implican (c) sigue de notar

$$
\begin{aligned}
\boldsymbol{A}^2 &= \Big(\sum_{i=1}^{k} \boldsymbol{A}_i\Big)^2 = \sum_{i=1}^{k}\sum_{j=1}^{k} \boldsymbol{A}_i\boldsymbol{A}_j = \sum_{i=1}^{k} \boldsymbol{A}_i^2 + \sum\sum_{i\neq j} \boldsymbol{A}_i\boldsymbol{A}_j \\
&= \sum_{i=1}^{k} \boldsymbol{A}_i = \boldsymbol{A}
\end{aligned}
$$

Finalmente, para probar que (b) y (c) implican (a). Suponga que (c) es verdad, entonces $\boldsymbol{A}_i\boldsymbol{A}_j = \boldsymbol{A}_j\boldsymbol{A}_i$ para todo $i \neq j$ y las matrices $\boldsymbol{A}_1, \dots, \boldsymbol{A}_k$ pueden ser diagonalizadas simultáneamente. Esto es, existe una matriz ortogonal $\boldsymbol{Q}$ tal que

$$
\boldsymbol{Q}^\top \boldsymbol{A}_i \boldsymbol{Q} = \boldsymbol{D}_i, \qquad i = 1, \dots, k,
$$

donde cada una de las matrices $\boldsymbol{D}_1, \dots, \boldsymbol{D}_k$ es diagonal.

Además,

$$\boldsymbol{D}_i\boldsymbol{D}_j = \boldsymbol{Q}^\top\boldsymbol{A}_i\boldsymbol{Q}\boldsymbol{Q}^\top\boldsymbol{A}_j\boldsymbol{Q} = \boldsymbol{Q}^\top\boldsymbol{A}_i\boldsymbol{A}_j\boldsymbol{Q} = \boldsymbol{0}, \qquad i \neq j. \tag{2}$$

Como $\boldsymbol{A}$ es simétrica e idempotente, también lo es la matriz diagonal

$$\boldsymbol{Q}^\top\boldsymbol{A}\boldsymbol{Q} = \boldsymbol{D}_1 + \cdots + \boldsymbol{D}_k,$$

y cada elemento diagonal de $\boldsymbol{Q}^\top\boldsymbol{A}\boldsymbol{Q}$ debe ser 0 o 1. y por (2), lo mismo es válido para para los elementos diagonales de $\boldsymbol{D}_1,\dots,\boldsymbol{D}_k$.

De este modo, $\boldsymbol{D}_i$ es simétrica e idempotente y de ahí que tambiél lo es

$$\boldsymbol{A}_i = \boldsymbol{Q}\boldsymbol{D}_i\boldsymbol{Q}^\top, \qquad i = 1,\dots,k,$$

lo que termina la prueba.

Observación:

Suponga que las condiciones del Lema 2 son satisfechas. Entonces (a), implica que $\operatorname{rg}(\boldsymbol{A}_i) = \operatorname{tr}(\boldsymbol{A}_i)$, y desde (b), sigue que

$$\operatorname{rg}(\boldsymbol{A}) = \operatorname{tr}(\boldsymbol{A}) = \operatorname{tr}\Big(\sum_{i=1}^{k} \boldsymbol{A}_i\Big) = \sum_{i=1}^{k} \operatorname{tr}(\boldsymbol{A}_i) = \sum_{i=1}^{k} \operatorname{rg}(\boldsymbol{A}_i).$$

Resultado 3 (Teorema de Cochran):

Sea $\boldsymbol{X} \sim \mathsf{N}_p(\boldsymbol{\mu}, \boldsymbol{\Sigma})$, con $\boldsymbol{\Sigma} > \mathbf{0}$. Suponga que $\boldsymbol{A}_i$ es una matriz simétrica de orden $p \times p$ con rango r_i, para $i = 1, \dots, k$, y

$$\boldsymbol{A} = \boldsymbol{A}_1 + \cdots + \boldsymbol{A}_k,$$

es de rango r. Considere las condiciones

(a) $\boldsymbol{A}_i\boldsymbol{\Sigma}$ es idempotente para $i = 1, \dots, k$.

(b) $\boldsymbol{A}\boldsymbol{\Sigma}$ es idempotente.

(c) $\boldsymbol{A}_i\boldsymbol{\Sigma}\boldsymbol{A}_j = \mathbf{0}$ para $i \neq j$.

(d) $r = \sum_{i=1}^k r_i$.

si dos de (a), (b) y (c) se satisfacen, o si (b) (d) son satisfechas, entonces

(i) $\boldsymbol{X}^\top\boldsymbol{A}_i\boldsymbol{X} \sim \chi^2(r_i, \lambda_i)$, con $\lambda_i = \boldsymbol{\mu}^\top\boldsymbol{A}\boldsymbol{\mu}/2$, $i = 1, \dots, k$.

(ii) $\boldsymbol{X}^\top\boldsymbol{A}\boldsymbol{X} \sim \chi^2(r, \lambda)$, con $\lambda = \boldsymbol{\mu}^\top\boldsymbol{A}\boldsymbol{\mu}/2$.

(iii) $\boldsymbol{X}^\top\boldsymbol{A}_1\boldsymbol{X}, \boldsymbol{X}^\top\boldsymbol{A}_2\boldsymbol{X}, \dots, \boldsymbol{X}^\top\boldsymbol{A}_k\boldsymbol{X}$ son mutuamente independientes.

Demostración:

Tenemos que $\boldsymbol{\Sigma} = \boldsymbol{T}\boldsymbol{T}^\top$ y las condiciones (a)-(d), pueden ser expresadas como:

(a) $\boldsymbol{T}^\top \boldsymbol{A}_i \boldsymbol{T}$ es idempotente para $i = 1, \dots, k$.

(b) $\boldsymbol{T}^\top \boldsymbol{A} \boldsymbol{T}$ es idempotente.

(c) $(\boldsymbol{T}^\top \boldsymbol{A}_i \boldsymbol{T})(\boldsymbol{T}^\top \boldsymbol{A}_j \boldsymbol{T}) = \boldsymbol{0}$ para $i \neq j$.

(d) $\operatorname{rg}(\boldsymbol{T}^\top \boldsymbol{A} \boldsymbol{T}) = \sum_{i=1}^k \operatorname{rg}(\boldsymbol{T}^\top \boldsymbol{A}_i \boldsymbol{T})$.

Como $\boldsymbol{T}^\top \boldsymbol{A}_1 \boldsymbol{T}, \boldsymbol{T}^\top \boldsymbol{A}_2 \boldsymbol{T}, \dots, \boldsymbol{T}^\top \boldsymbol{A}_k \boldsymbol{T}$ y $\boldsymbol{T}^\top \boldsymbol{A} \boldsymbol{T}$ satisfacen las condiciones del Lema 2.[1] Entonces, las condiciones (a)-(d) se satisfacen.

Sabemos que (a) implica (i) y (b) implica (ii). Mientras que, Resultado 1 con (c), garantiza (iii), lo que completa la prueba.

[1] y Observación en slide 10.