UNIVERSIDAD EAFIT®

DEPARTAMENTO DE CIENCIAS MATEMÁTICAS
Estadística 2 Ing Mat CM0423
parcial 1 (25%)

Nombre: David Plazas Escudero Código: 201710005101 Nota: 4.8

Profesor: Francisco Zuluaga Grupo: 001 Febrero 21 de 2019

1) Sea $y \sim (\mu, \Sigma)$ y sea A Una matriz de constantes de dimensión PxP.

Demuestre que $E(y'Ay) = tr(A\Sigma) + \mu' A\mu$. tenga presente que y es de dimesion px1

el valor esperado de la traza es la traza del valor esperado
Tr(a) = a siendo a escalar
Tr(X+Y) = tr(X) + Tr(Y)
Tr(aX)= a Tr(X)

1.25

$$y = \begin{bmatrix} y_1 \\ \vdots \\ y_p \end{bmatrix}_{p\times 1} \quad \mu = \begin{bmatrix} \mu_1 \\ \vdots \\ \mu_p \end{bmatrix}_{p\times 1}$$

(*)tr(XYZ)=tr(ZXY)=tr(YZX) siempre y cuando el producto exista

$$A = \begin{bmatrix} a_{11} & \cdots & a_{1p} \\ \vdots & \ddots & \vdots \\ a_{p1} & \cdots & a_{pp} \end{bmatrix}$$

Tenga presente que $\Sigma = E(yy') - \mu\mu'$ (1)

$E(y'Ay) = tr(A\Sigma) + \mu' A \mu$ (2) (1) en (2)

$$\underset{1\times p}{y'}\ \underset{p\times p}{A}\ \underset{p\times 1}{y} = [\ ]_{1\times 1}$$

$$\underset{p\times p}{A}\ \underset{(p\times 1)}{y}\ \underset{(1\times p)}{y'} \quad (p\times p)$$

$= tr[A(E(yy') - \mu\mu')] + \mu' A \mu$

$$\underset{1\times p}{\mu'}\ \underset{p\times p}{A}\ \underset{p\times 1}{\mu} = [\ ]_{1\times 1} \text{ Escalar}$$

$= tr[AE(yy') - A\mu\mu'] + tr(\mu' A \mu)$

$$\underset{1\times p}{y'}\ \underset{p\times p}{A}\ \underset{p\times 1}{y} = 1\times 1$$

Porque $\mu' A \mu$ es escalar y traza de escalar es el mismo.

$$y_{p\times 1}\, y'_{1\times p} = [\ ]_{p\times p}$$

$= tr[AE(yy')] - tr(\underset{(p\times p)(p\times 1)(1\times p)}{A\mu\mu'}) + tr(\underset{(1\times p)(p\times p)(p\times 1)}{\mu' A \mu})$ $= 1\times 1$

pxp

Luego, $tr(A\mu\mu') = tr(\mu' A\mu)$ por (*) y el producto existe.

$= tr[AE(yy')] - tr(\mu' A \mu) + tr(\mu' A \mu)$

$= tr[AE(yy')]$

$= tr[E(Ayy')] = E(tr(Ayy')) = E(tr(y'Ay)) = E(y'Ay)$

ambos productos existen

Porque $y'Ay$ es escalar.

Scanned with CamScanner

UNIVERSIDAD EAFIT
Abierta al mundo

(4) : $D_s = \begin{bmatrix} s_1 & 0 & \cdots & 0 \\ 0 & s_2 & \cdots & 0 \\ \vdots & & \ddots & \\ 0 & 0 & \cdots & s_p \end{bmatrix} \Longrightarrow D_s^{-1} = \begin{bmatrix} s_1^{-1} & 0 & \cdots & 0 \\ 0 & s_2^{-1} & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & s_p^{-1} \end{bmatrix}$

2) Sea $D_s = diag(s_1, s_2, \ldots, s_p)$, $Y = \begin{bmatrix} y_{11} & \cdots & y_{1p} \\ \vdots & \ddots & \vdots \\ y_{n1} & \cdots & y_{np} \end{bmatrix}_{(n \times p)}$,J una matrix de unos de (nxn)

1.25

a) Muestre que la siguiente operación produce una matrix de datos estandarizados

$Z = \left(I_n - \frac{J}{n}\right) Y D_s^{-1}$

Analicemos primero

$\left(I_n - \frac{J}{n}\right)Y$

$Y - \frac{J}{n}Y$

$= Y - \frac{1}{n}\underbrace{\begin{bmatrix} 1 & \cdots & 1 \\ \vdots & & \vdots \\ 1 & \cdots & 1 \end{bmatrix}}_{n \times n}\underbrace{\begin{bmatrix} y_{11} & \cdots & y_{1p} \\ \vdots & & \vdots \\ y_{n1} & \cdots & y_{np} \end{bmatrix}}_{n \times p}$

$= Y - \frac{1}{n}\begin{bmatrix} y_{11}+y_{21}+\ldots+y_{n1} & \cdots & y_{1p}+y_{2p}+\ldots+y_{np} \\ y_{11}+y_{21}+\ldots+y_{n1} & \ddots & y_{1p}+y_{2p}+\ldots+y_{np} \\ \vdots & & \vdots \\ y_{11}+y_{21}+\ldots+y_{n1} & \cdots & y_{1p}+y_{2p}+\ldots+y_{np} \end{bmatrix}$

$= Y - \begin{bmatrix} \frac{y_{11}+\ldots+y_{n1}}{n} & \cdots & \frac{y_{1p}+y_{2p}+\ldots+y_{np}}{n} \\ \vdots & \ddots & \vdots \\ \frac{y_{11}+\ldots+y_{n1}}{n} & \cdots & \frac{y_{1p}+\ldots+y_{np}}{n} \end{bmatrix}$

$= Y - \begin{bmatrix} \bar{y}_1 & \cdots & \bar{y}_p \\ \vdots & \ddots & \\ \bar{y}_1 & \cdots & \bar{y}_p \end{bmatrix}$

$= \begin{bmatrix} y_{11}-\bar{y}_1 & \cdots & y_{1p}-\bar{y}_p \\ \vdots & \ddots & \\ y_{n1}-\bar{y}_1 & \cdots & y_{np}-\bar{y}_p \end{bmatrix} = Y_c$

Luego, $\underset{(4)}{Y_c D_s^{-1}} = \begin{bmatrix} y_{11}-\bar{y}_1 & \cdots & y_{1p}-\bar{y}_p \\ \vdots & \ddots & \vdots \\ y_{n1}-\bar{y}_1 & \cdots & y_{np}-\bar{y}_p \end{bmatrix}\begin{bmatrix} s_1^{-1} & \cdots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \cdots & s_p^{-1} \end{bmatrix}$

$= \begin{bmatrix} \frac{y_{11}-\bar{y}_1}{s_1} & \cdots & \frac{y_{1p}-\bar{y}_p}{s_p} \\ \vdots & & \vdots \\ \frac{y_{n1}-\bar{y}_1}{s_1} & \cdots & \frac{y_{np}-\bar{y}_p}{s_p} \end{bmatrix}_{(n \times p)} = Y_s$

(Datos estandarizados)

donde

$Z_{ij} = \frac{y_{ij}-\bar{y}_j}{s_j}$ $\quad i = 1, \ldots, n \quad j = 1, \ldots, p$

b) Demuestre que $S_Z = \frac{Z'Z}{n-1} = R$, Es decir demuestre que la matriz de covarianza muestral de los datos estadarizados es igual a la matriz de correlacion de los datos originales.

$S_{xy} = \sum_{i=1}^{n} \frac{(x_i-\bar{x})(y_i-\bar{y})}{n-1}$

$S_Z = \frac{Z'Z}{n-1}$

$= \frac{1}{n-1}\begin{bmatrix} \frac{y_{11}-\bar{y}_1}{s_1} & \frac{y_{21}-\bar{y}_1}{s_1} & \cdots & \frac{y_{n1}-\bar{y}_1}{s_1} \\ \vdots & \vdots & \ddots & \vdots \\ \frac{y_{1p}-\bar{y}_p}{s_p} & \frac{y_{2p}-\bar{y}_p}{s_p} & \cdots & \frac{y_{np}-\bar{y}_p}{s_p} \end{bmatrix}\begin{bmatrix} \frac{y_{11}-\bar{y}_1}{s_1} & \cdots & \frac{y_{1p}-\bar{y}_p}{s_p} \\ \frac{y_{21}-\bar{y}_1}{s_1} & \cdots & \frac{y_{2p}-\bar{y}_p}{s_p} \\ \vdots & \ddots & \vdots \\ \frac{y_{n1}-\bar{y}_1}{s_1} & \cdots & \frac{y_{np}-\bar{y}_p}{s_p} \end{bmatrix} = \begin{bmatrix} \frac{\sum(y_{i1}-\bar{y}_1)^2}{(n-1)s_1^2} & \frac{\sum(y_{i1}-\bar{y}_1)(y_{i2}-\bar{y}_2)}{(n-1)s_1 s_2} & \cdots & \boxed{\cdot} \\ \vdots & & & \\ \frac{\sum(y_{ip}-\bar{y}_p)(y_{i1}-\bar{y}_1)}{(n-1)s_p s_1} & \cdots & & \frac{\sum(y_{ip}-\bar{y}_p)^2}{(n-1)s_p^2} \end{bmatrix}$

$= \begin{bmatrix} \frac{s_{11}}{s_1^2} & \frac{s_{12}}{s_1 s_2} & \cdots & \frac{s_{1p}}{s_1 s_p} \\ \vdots & & & \\ \frac{s_{p1}}{s_p s_1} & \frac{s_{p2}}{s_p s_2} & \cdots & \frac{s_{pp}}{s_p^2} \end{bmatrix} = \begin{bmatrix} r_{11} & r_{12} & \cdots & r_{1p} \\ \vdots & & \ddots & \vdots \\ r_{p1} & r_{p2} & \cdots & r_{pp} \end{bmatrix} = R$

$\boxed{\cdot} = \frac{\sum(y_{i1}-\bar{y}_1)(y_{ip}-\bar{y}_p)}{(n-1)s_1 s_p}$

Scanned with CamScanner

$$xy - x\bar{y} - \bar{x}y + \bar{x}\bar{y}$$

UNIVERSIDAD EAFIT
Abierta al mundo

$$\sigma_{xy} = E((X-\mu_x)(Y-\mu_y))$$
$$= E(XY) - E(X)\mu_y - \mu_x E(Y) + E(\;\;)$$
$$\sigma_{xy} = E(XY) - \mu_x \mu_y$$
$$E(S_{xy}) = \sigma_{xy} \longrightarrow E(XY) = \sigma_{xy} + \mu_x \mu_y$$

3) Demuestre que $s_{xy} = \frac{1}{n-1}(\sum_{i=1}^{n} x_i y_i \;-\; n\bar{x}\bar{y})$ es un estimador insesgado de σ_{xy}

Pista: tenga presente $\text{cov}(\bar{x},\bar{y}) = \frac{cov(x,y)}{n} = \frac{\sigma_{xy}}{n}$

$$\sigma_{xy} = E(XY) - \mu_x \mu_y$$
$$\Longrightarrow \sigma_{\bar{x}\bar{y}} = E(\bar{x}\bar{y}) - \mu_{\bar{x}}\mu_{\bar{y}}$$
$$\frac{\sigma_{xy}}{n} = E(\bar{x}\bar{y}) - \mu_x \mu_y$$
$$E(\bar{x}\bar{y}) = \frac{\sigma_{xy}}{n} + \mu_x \mu_y$$

$$S_{xy} = \frac{1}{n-1}\sum_{i=1}^{n} x_i y_i - n\bar{x}\bar{y}$$

$$E(S_{xy}) = \frac{1}{n-1}\left(\sum_{i=1}^{n} E(x_i y_i) - nE(\bar{x}\bar{y})\right)$$

$$= \frac{1}{n-1}\left[\sum_{i=1}^{n}(\sigma_{xy} + \mu_x\mu_y) - n\left(\frac{\sigma_{xy}}{n} + \mu_x\mu_y\right)\right]$$

$$= \frac{1}{n-1}\left[(n\sigma_{xy} + \cancel{n\mu_x\mu_y}) - \frac{\cancel{n}}{\cancel{n}}\sigma_{xy} - \cancel{n\mu_x\mu_y}\right]$$

1.25

$$= \frac{n\sigma_{xy} - \sigma_{xy}}{n-1} = \sigma_{xy}$$

Como $E(S_{xy}) = \sigma_{xy} \Rightarrow S_{xy}$ es insesgado de σ_{xy}

UNIVERSIDAD EAFIT
Abierta al mundo

4. En el archivo café.txt se encuentra información sobre diferentes caracteriticas asociadas diferentes variedades de café. La valoración sensorial es un instrumento que se ha venido aplicando para el control de calidad, aceptabilidad y comercialización del café

```
Intar : Intensidad del aroma
Aroma : Aroma de la bebida
Cuerpo: Cuerpo
Acidez: Acidez
Astr : Astringencia
Amargo : Amargo
```

a) Encuentre el vector de medias, las matriz de covarianza muestral y de correlaciones muestral.

b) Encuentre det(S) y tr(S)

Defina la siguientes combinaciones lineales para las variables.

z_1 = 0.5827Intar+0.6409Aroma+0.2746Cuerpo+0.3107Acidez+0.2210Amargo+0.1693Astr

z_2 = 0.5335Intar-0.6936Aroma+-0.1116Cuerpo+0.4469Acidez-.0020Amargo+0.1478Astr

c) Encuentre $\bar{z}$ y S_z → Falta

d) Encuentre R_z *a partir de* S_z → Falta

e) Implemente En R La siguiente formula y validela con los resultados del numeral a

Matriz de covarianzas muestrales $S = \frac{1}{n-1}\left(\sum_{i=1}^{n} y_i y_i' - n\bar{y}\bar{y}'\right)$

10