UNIVERSITAT DE BARCELONA

Maximizar funciones

MULTIVARIATE ANALYSIS
MESIO (15-16)

PROF. SERGI CIVIT
PROF. MIQUEL SALICRÚ

Función de verosimilitud

Sea $x_1, x_2, \ldots, x_n$ una muestra aleatoria representativa de la variable $X \approx N_p(\mu, \Sigma)$. Ya que $x_1, x_2, \ldots, x_n$ son independientes, la función de densidad conjunta es el producto de marginales:

$$L(x_1, \ldots, x_n) = \prod_{i=1}^{n} \left\{ \frac{1}{|\Sigma|^{1/2} (2\pi)^{p/2}} e^{-\frac{1}{2}(x_i-\mu)'\Sigma^{-1}(x_i-\mu)} \right\}$$

$$= (2\pi)^{-np/2} |\Sigma|^{-n/2} \exp\left\{ -\tfrac{1}{2} \sum_{i=1}^{n} (x_i - \mu)'\Sigma^{-1}(x_i - \mu) \right\}$$

Interés práctico

Estimación de parámetros. $\hat{\mu}$ y $\hat{\Sigma}$ maximizan la función de verosimilitud: son solución del sistema de ecuaciones

$$\frac{\partial L(x_1,\ldots,x_n)}{\partial \mu} = 0 \quad , \quad \frac{\partial L(x_1,\ldots,x_n)}{\partial \Sigma} = 0$$

Test de hipótesis. $H_0 : \theta \in \Theta_0$ se rechaza en favor de $H_1 : \theta \in \Theta$ si

$$\Lambda = \frac{\max_{\theta \in \Theta_0} L(x_1,\ldots,x_n,\theta)}{\max_{\theta \in \Theta} L(x_1,\ldots,x_n,\theta} < c$$

Asintoticamente: $-2\ln\Lambda \approx \chi^2_{\nu-\nu_0}$

Optimización de funciones

Funciones de una variable $y = f(x) : R \longrightarrow R$

f ’(x)=0 y f ”(x)≠0

Funciones vectoriales y matriciales

Definición de la derivada: escalar/vector y escalar/matriz

Formulario de derivación: funciones más relevantes

Derivación vectorial y matricial 1/3

Escalar/escalar $y = f(x) : R \longrightarrow R$

$$y' = \frac{\partial y}{\partial x} = f'(x)$$

Escalar/vector $y = f(x_1, \ldots, x_p) : R^p \longrightarrow R$

$$y' = \frac{\partial y}{\partial x} = \left(\frac{\partial y}{\partial x_1}, \ldots, \frac{\partial y}{\partial x_p}\right)$$

Escalar/matriz $y = f(x_{11}, \ldots, x_{pq}) : R^{pxq} \longrightarrow R$

$$y' = \frac{\partial y}{\partial A} = \begin{pmatrix} \frac{\partial y}{\partial x_{11}} & \cdots & \frac{\partial y}{\partial x_{1q}} \\ \vdots & & \vdots \\ \frac{\partial y}{\partial x_{p1}} & \cdots & \frac{\partial y}{\partial x_{pq}} \end{pmatrix}$$

Derivación vectorial y matricial 2/3

Matriz/matriz (caso general): $X \approx M_{np}, Y(X) \approx M_{mq}, \frac{\partial Y(X)}{\partial X} \approx M_{pm\,x\,nq}$

$$Y(X) = \begin{pmatrix} y_{11}(X) & \cdots & y_{1q}(X) \\ \vdots & & \vdots \\ y_{m1}(X) & \cdots & y_{mq}(X) \end{pmatrix} \Rightarrow \frac{\partial Y(X)}{\partial X} = \begin{pmatrix} \frac{\partial y_{11}}{\partial X} & \cdots & \frac{\partial y_{1q}}{\partial X} \\ \vdots & & \vdots \\ \frac{\partial y_{m1}}{\partial X} & \cdots & \frac{\partial y_{mq}}{\partial X} \end{pmatrix}_{\text{MATRIZ POR CAJAS}}$$

$$\downarrow$$

$$\frac{\partial y_{ij}(X)}{\partial X} = \left(\frac{\partial y_{ij}(X)}{\partial x_{kr}}\right)_{\substack{k=1,\ldots,p \\ r=1,\ldots,n}}$$

Derivación vectorial y matricial 3/3

Vector/escalar, vector/vector, vector/matriz

Matriz/escalar, matriz/vector, matriz/matriz

VECTOR/VECTOR

$$\frac{\partial Y(X)}{\partial X} = \begin{pmatrix} \frac{\partial y_1(x)}{\partial x_1} & .... & \frac{\partial y_1(x)}{\partial x_p} \\ \vdots & & \vdots \\ \frac{\partial y_q(x)}{\partial x_1} & .... & \frac{\partial y_q(x)}{\partial x_p} \end{pmatrix}$$

Qué derivar? 1/2

$$\ln L(x_1,\dots,x_n) = \frac{-np}{2}\ln(2\pi) \boxed{-\frac{n}{2}\ln|\Sigma|} \boxed{-\tfrac{1}{2}\sum_{i=1}^{n}(x_i-\mu)'\Sigma^{-1}(x_i-\mu)}$$

Igualdad 1.

$$-\frac{1}{2}\sum_{i=1}^{n}(x_i-\mu)'\Sigma^{-1}(x_i-\mu) = -\frac{1}{2}\mathrm{tr}\left[\sum_{i=1}^{n}(x_i-\mu)'\Sigma^{-1}(x_i-\mu)\right]$$

$$= -\frac{1}{2}\sum_{i=1}^{n}\mathrm{tr}[(x_i-\mu)'\Sigma^{-1}(x_i-\mu)] = -\frac{1}{2}\sum_{i=1}^{n}\mathrm{tr}[\Sigma^{-1}(x_i-\mu)(x_i-\mu)']$$

$$= -\frac{1}{2}\mathrm{tr}\left[\Sigma^{-1}\cdot\sum_{i=1}^{n}(x_i-\mu)(x_i-\mu)'\right] = \boxed{-\frac{n}{2}\mathrm{tr}\left[\Sigma^{-1}\cdot S_0\right]}$$

Qué derivar? 2/2

Igualdad 2.

$$-\frac{1}{2}\sum_{i=1}^{n}(x_i-\mu)'\Sigma^{-1}(x_i-\mu)=\boxed{-\frac{n}{2}\mathrm{tr}\left[\Sigma^{-1}\cdot S\right]-\frac{n}{2}(\bar{X}-\mu)'\Sigma^{-1}(\bar{X}-\mu)}$$

$$nS_0=\sum_{i=1}^{n}(x_i-\mu)(x_i-\mu)'$$

$$=\sum_{i=1}^{n}(x_i-\bar{X}+\bar{X}-\mu)(x_i-\bar{X}+\bar{X}-\mu)'=nS+n(\bar{X}-\mu)(\bar{X}-\mu)'$$

Formulario: escalar/vector

$$\frac{\partial\, a'x}{\partial x}=a'$$

$$\frac{\partial\,(x-a)'A(x-a)}{\partial x}=2A(x-a)\quad (A=A')$$

$$\frac{\partial\, x'Ay}{\partial x}=Ay\quad (A=A')$$

Formulario: escalar/matriz

$$\frac{\partial \log|X|}{\partial X} = 2X^{-1} - \mathrm{diag}(X^{-1})$$

$$\frac{\partial |X|}{\partial X} = |X|(2X^{-1} - \mathrm{diag}(X^{-1}))$$

$$\frac{\partial\, \mathrm{tr}(XY)}{\partial X} = 2Y - \mathrm{diag}(Y)$$

$$\frac{\partial\, \mathrm{tr}(X^{-1}Y)}{\partial X} = -2X^{-1}YX^{-1} + \mathrm{diag}(X^{-1}YX^{-1})$$

$$\frac{\partial (Y-\alpha)'X^{-1}(Y-\alpha)}{\partial X} = -2X^{-1}((Y-\alpha)(Y-\alpha)')X^{-1} + \mathrm{diag}(X^{-1}((Y-\alpha)(Y-\alpha)')X^{-1})$$

X, Y matrices simétricas nxn

Traza de matrices

$\mathrm{tr}(A \pm B) = \mathrm{tr}(A) \pm \mathrm{tr}(B)$

$\mathrm{tr}(A \cdot B) = \mathrm{tr}(B \cdot A)$ (si A, B compatibles)

$\mathrm{tr}(B^{-1} \cdot A \cdot B) = \mathrm{tr}(A)$ (si B no singular)

$x'\Sigma x = \mathrm{tr}(x'\Sigma x) = \mathrm{tr}(\Sigma x x')$

$\mathrm{tr}(A \cdot A') = \sum_{ij} a_{ij}^2$

Determinante de matrices

$$|A \cdot B| = |A| \cdot |B|$$

$$|A^{-1}| = 1/|A|$$

$$|B \cdot A \cdot B^{-1}| = |A|$$

$$|\lambda A| = \lambda^p |A|$$

$$|I_n + Abb'| = 1 + b'Ab \quad \begin{array}{l} \text{A no singular nxn} \\ \text{b nector nx1} \end{array}$$

Referencias

Graybill, F. A. (1983). Matrices with applications in statistics.

Magnus, J. R., & Neudecker, H. (1995). Matrix differential calculus with applications in statistics and econometrics.