Estadística Multivariante Derivación matricial

Trabajo B

Antonio R. Moya Martín-Castaño
Elena Romero Contreras
Nuria Rodríguez Barroso
Universidad de Granada
anmomar85@correo.ugr.es
elenaromeroc@correo.ugr.es
rbnuria6@gmail.com

Índice

1. Introducción

En este apartado se va a tratar la derivación respecto a vectores y matrices, que es muy necesaria en estadística multivariante sobre todo desde el punto de vista de la optimización. Así, permite calcular datos tales como estimador máximo verosímil, matrices de información de Fisher, o cotas tipo Crámer-Rao. Más importancia tiene todavía este tema si tenemos en cuenta que, si ya la derivación vectorial puede dar lugar a cálculos costosos, en el caso de la matricial se pueden generar un enorme número de derivadas que pueden resultar difícil de ordenar con sentido en una matriz.

Muchas son las aproximaciones que se han dado y en este caso lo que se hará es seguir una línea en la que las matrices se conviertan en vectores, mucho más fáciles de manejar. **********

2. Diferencial primera y jacobianos

2.1. Diferencial de una función vectorial

Definición 2.2.1: Consideramos una función vectorial $f : S \to \mathbb{R}^m$ con $S \subset \mathbb{R}^n$. Sea **c** un punto interior de S y consideremos una bola cerrada con centro en **c** y radio **r**, $B(c, r)$. Sea u un punto de $\mathbb{R}^n$ tal que $||u|| \leq r$ es decir, $c + r \in B(c, r)$.
Diremos que f es diferenciable en **c** si existe una matriz real de orden m x n que depende de c y no de u y que cumple que $f(c + u) - f(c) = A(c)u + r_c(u)$ con $\lim_{u\to 0} \frac{r_c(u)}{||u||} = 0$. Además, se define la primera diferencial de f en el punto c con incremento u como: $df(c; u) = A(c)u$.

Definición 2.2.2: Sea $f : S \subset \mathbb{R}^n \to \mathbb{R}^m$ y sea $f_i : S \to \mathbb{R}$ su i-ésima componente. Sea c un punto interior de S y e_j el j-ésimo vector de la base canónica de $\mathbb{R}^n$. Se define la derivada parcial de f respecto a la j-ésima coordenada como:

$$D_j f_i(c) = \lim_{t\to 0} \frac{f_i(c + te_j) - f_i(c)}{t}, t \in \mathbb{R}$$

Teorema 2.1. *(Primer teorema de identificación para funciones vectoriales)*
*Sea $f : S \subset \mathbb{R}^n \to \mathbb{R}^m$ diferenciable en un punto c interior de S y $u \in \mathbb{R}^n$. Entonces $df(c; u) = (Df(c))u$ donde $Df(c)$ es una matriz mxn cuyos elementos $D_j f_i(c)$ son las derivadas parciales de **f** evaluadas en c y que recibe el nombre de matriz jacobiana. Recíprocamente, si $A(c)$ es una matriz que verifica que $df(c; u) = A(c)u \forall u \in \mathbb{R}^n$, entonces $A(c) = Df(c)$.*

Teorema 2.2. *Sea $f : S \subset \mathbb{R}^n \to \mathbb{R}^m$ diferenciable en un punto c interior de S. Sea T un subconjunto de $\mathbb{R}^m$ tal que $f(x) \in T \forall x \in S$ y supongamos que $g : T \to \mathbb{R}^p$ es diferenciable en un punto **b** (b = f(c)) de T. Entonces la función compuesta $h : S \to \mathbb{R}^p$ definida por $h(x) = g(f(x))$, es diferenciable en c y $Dh(c) = Dg(b)Df(c)$.*

Teorema 2.3. *(Regla de invarianza de Cauchy)*
*En el ambiente del teorema anterior, si **f** es diferenciable en **c** y **g** lo es en **b=f(c)**, entonces la diferenciable $h = gof$ es $dh(c; u) = dg(b; df(c; u))$.*

Ejercicio 2.1: Sea $h : \mathbb{R}^k \to \mathbb{R}$ definida por $h(\beta) = (y - X\beta)^t(y - X\beta)$ donde $y \in \mathbb{R}^n$ y $X \in \mathbb{M}_{n\times k}$. Haciendo uso de la regla de invarianza de Cauchy demostrar que

$$dh(c; u) = dg(y - Xc; df(c; u)) = dg(y - Xc; -Xu) = -2(y - Xc)^t Xu$$

y con ello $Dh(c) = -2(y - Xc)^t X$.

Solución: Para comenzar, tengamos en cuenta que h es composición de $f : \mathbb{R}^k \to \mathbb{R}^n$ y $g : \mathbb{R}^n \to \mathbb{R}$, con $f(c) = (y - Xc)$ con X una matriz nxk y con $y \in \mathbb{R}^n$; y $g(X) = X^t X$, de modo que $h(x) = g(f(x)) \forall x \in \mathbb{R}^k$. Así, es claro que $dh(c; u) = dg(y - Xc; df(c; u))$. ************************ Parece evidente pero no se que escribir xd***************

Ejercicio 2.2: Sea $F(X) = AG(X)B$, donde $A_{m\times r}$ y $B_{s\times p}$ son matrices constantes y $G(X)_{r\times s}$ es una función diferenciable. Calcular $DF(C)$ a partir de la definición de diferencial matricial.

Solución:

Ejercicio 2.3: Si $X_{n\times n}$ es una matriz simétrica y $F : \mathbb{M}_{n\times q} \to \mathbb{M}_{m\times p}$ es diferenciable, demostrar que $d\text{Vec}(F(X)) = D_n DF(X) d\text{Vech}(X)$, mientras que $d\text{Vec}(F(X)) = N_n DF(X) d\text{Vec}(X)$ donde $N_n = \frac{1}{2}[I_{n^2} + K_{nn}]$.

Solución:

3. Matrices jacobianas y derivadas matriciales

Ejercicio 3.1: A partir de las relacioens existentes entre la derivada matricial y la matriz jacobiana, verificar las siguientes expresiones:

a) Sea $X_{n\times n}$ y $F(X) = \text{tr}[X]$. Entonces $DF(X) = \text{Vec}^t(I_n)$.

b) Sea ahora $X_{n\times q}$ y $F(X) = X$. Entonces $DF(X) = I_q \otimes I_n = I_{nq}$.

Ejercicio 3.2: Sea $X_{n\times q}$. Demostra las siguientes igualdades:

a) $\dfrac{\partial X^t}{\partial X} = K_{qn}$.

b) $\dfrac{\partial X}{\partial X^t} = K_{nq}$.

c) $\dfrac{\partial X^t}{\partial X^t} = \text{Vec}(I_q)\text{Vec}^t(I_n)$.

Ejercicio 3.3: Demostrar que si $X_{n\times n}$ es no singular entonces $\dfrac{\partial X^{-1}}{\partial X} = -\text{Vec}((X^{-1})^t)\text{Vec}^t(X^{-1})$.

Solución:

4. Diferencial segunda y hessianos