Notas de Cálculo Diferencial e Integal II (2002) - Clase 11

Teorema 1 (Diferencial de la función compuesta (Regla de la Cadena)). *Consideremos las funciones* $f = (f_1, \dots, f_m) \colon \mathbf{U} \to \mathbb{R}^m$ *donde* $\mathbf{U} \subset \mathbb{R}^n$ *es abierto;* $g \colon \mathbf{V} \to \mathbb{R}$*; donde* $\mathbf{V} \subset f(\mathbf{U}) \subset \mathbb{R}^m$ *también es abierto. Supongamos que cada función coordenada* f_j $(j = 1, \dots, m)$ *es diferenciable en un punto* $a \in \mathbf{U}$*, y que* g *es diferenciable en* $b = f(a)$*. Entonces, la función compuesta* $h = g \circ f \colon \mathbf{U} \to \mathbb{R}$ *es diferenciable en el punto* a*, y sus derivadas parciales verifican:*

$$\partial_i \big(g \circ f\big)(a) = \partial g_1(b) \partial_i f_1(a) + \cdots + \partial g_m(b) \partial_i f_m(a) \quad \textit{para cada } i = 1, \dots, n. \tag{1}$$

Observación. Consideremos el caso $m = n = 1$. Hemos visto que para $f \colon (a, b) \to \mathbb{R}$, las nociones de función derivable y función diferenciable coinciden. Es claro entonces, que se trata de una generalización de la regla de la cadena de funciones reales. En particular, en la ecuación (1) queda un único sumando con el producto de las derivadas.

Observación. Veamos como se escribe la fórmula (1) en la notación de Leibnitz del cálculo diferencial. Designemos mediante $x = (x_1, \dots, x_n)$ e $y = (y_1, \dots, y_m)$ los puntos de $\mathbf{U}$ y $\mathbf{V}$ respectivamente, $z = g(y_1, \dots, y_m)$, y la composición consiste en considerar que cada variable y_j $(j = 1, \dots, m)$ depende (es una función) de las variables $(x_1, \dots, x_n)$, y tenemos $y_j = f_j(x_1, \dots, x_n)$ $(j = 1, \dots, m)$. Decimos entonces que las y_j $(j = 1, \dots, m)$ son las *variables dependientes*, mientras que las x_i $(i = 1, \dots, n)$ son las *variables independientes*; y escribimos

$$\partial_i h = \frac{\partial z}{\partial x_i}, \qquad \partial_j g = \frac{\partial z}{\partial y_j}, \qquad \partial_i f_j = \frac{\partial y_j}{x_i},$$

con lo que la regla de la cadena (1) se lee

$$\frac{\partial z}{\partial x_i} = \frac{\partial z}{y_1} \frac{\partial y_1}{\partial x_i} + \cdots + \frac{\partial z}{y_m} \frac{\partial y_m}{\partial x_i}.$$

Esta notación es especialmente útil para recordar.

Demostración. Veremos que se verifica la definición de función diferenciable para h, y al calcular las constantes en la definición, obtendremos (1). Como $\mathbf{U}$ es abierto, existe una bola $B(a, \varepsilon) \subset \mathbf{U}$. Para cada $v = (v_1, \dots, v_n)$ con

$a + v \in B(a, \varepsilon)$, aplicando la defineción de diferenciabilidad de cada función f_j $(j = 1, \dots, m)$ tenemos

$$\begin{aligned} f_j(a+v) &= f_j(a) + \sum_{i=1}^{n} \partial_i f_j(a) v_i + \|v\| p_j(v) \qquad (2) \\ \lim_{v \to 0} p_j(v) &= p(0) = 0. \end{aligned}$$

Sea $w = (w_1, \dots, w_m) = f(a+v) - f(a) = \big(f_1(a+v) - f_1(a), \dots, f_m(a+v) - f_m(a)\big)$. Como $w + b = w + f(a) = f(a+v) \in \mathbf{V}$, aplicando la defineción de diferenciabilidad, ahora a la función g, tenemos

$$\begin{aligned} g(b+w) &= g(b) + \sum_{j=1}^{m} \partial_j g(a) w_j + \|w\| q(w) \\ \lim_{w \to 0} q(w) &= q(0) = 0. \end{aligned}$$

Sustituyendo en la expresión anterior el factor $w_j = f_j(a+v) - f_j(a)$ $(j = 1, \dots, m)$ según lo obtenido en (2), tenemos

$$\begin{aligned} h(a+v) &= g\big(f(a+v)\big) = g(w+b) \\ &= g(b) + \sum_{j=1}^{m} \partial_j g(a) \Big[\sum_{i=1}^{n} \partial_i f_j(a) v_i + \|v\| p_j(v) \Big] + \|w\| q(w) \end{aligned}$$

Reordenando, como $g(b) = g\big(f(a)\big) = h(a)$, cambiando el orden en la suma doble, e introduciendo la función auxiliar $R(v)$, podemos escribir

$$h(a+v) = h(a) + \sum_{i=1}^{n} \Big[\sum_{j=1}^{m} \partial_j g(a) \partial_i f_j(a) \Big] v_i + \|v\| R(v) \qquad (3)$$

$$R(v) = \sum_{j=1}^{m} \partial_j g(b) p_j(v) + \frac{\|w\|}{\|v\|} q(w) \qquad (4)$$

Resta entonces, para obtener la defineción de diferenciabilidad para la función h en el punto a, ver que $\lim_{v \to 0} R(v) = 0$, y observar que las constantes en (3), que serían las derivadas parciales, cumplen la fórmula (1) (que es inmediato). Veamos entonces que $\|w\|/\|v\|$ está acotado. Tenemos,

$$|w_j| = \Big| \sum_{i=1}^{n} \partial_i f_j(a) v_i + \|v\| p_j(v) \Big| \le \sum_{i=1}^{n} \big|\partial_i f_j(a)\big| |v_i| + \|v\| |p_j(v)|$$

Como $|v_i|/\|v\| \leq 1$ y $p_j(v) \to 0$, resulta que $\|w\|/\|v\|$ está acotado. Finalmente, si $v \to 0$, por continuidad, $w \to 0$, y de la defnición (4) obtenemos que $R(v) \to 0$. Esto concluye la demostración. □

Veamos ahora los dos casos particulares, en los que $m = 1$ o $n = 1$.

Corolario 1. *Sean* $f\colon \mathbf{U} \to \mathbb{R}$, *con* $\mathbf{U} \subset \mathbb{R}^n$ *abierto;* $g\colon \mathbf{I} \to \mathbb{R}$, *con* $\mathbf{I} \subset \mathbb{R}$ *abierto, con* $f(\mathbf{U}) \subset \mathbf{I}$. *Si* f *es diferenciable en un punto* $a \in \mathbf{U}$, *y* g *es diferenciable (derivable) en* $b = f(a)$, *entonces* $h = g \circ f$ *es diferenciable en el punto* a, *y*

$$\frac{\partial h}{\partial x_i}(a) = g'(b)\frac{\partial f}{\partial x_i}(a) \qquad (i = 1, \dots, n)$$

Decimos que una función $\lambda\colon \mathbf{I} \to \mathbb{R}^n$, con $\mathbf{I}$ un intervalo real, es una *curva.*

Corolario 2. *Sean* $\lambda = (\lambda_1, \dots, \lambda_n)\colon \mathbf{I} \to \mathbb{R}^n$, *donde* $\mathbf{I}$ *es un intervalo abierto;* $g\colon \mathbf{V} \to \mathbb{R}$, *con* $\mathbf{V} \subset \mathbb{R}^n$ *abierto, con* $f(\mathbf{I}) \subset \mathbf{V}$. *Si* λ *es diferenciable en un punto* $a \in \mathbf{I}$, *y* g *es diferenciable en* $b = \lambda(a)$, *entonces* $\phi = g \circ f\colon \mathbf{I} \to \mathbb{R}$ *es diferenciable (derivable) en el punto* a, *y*

$$\phi'(a) = \partial_1 g(b)\lambda_1'(a) + \cdots + \partial_n g(b)\lambda_n'(a) = \langle \nabla g(b), \lambda'(a)\rangle,$$

donde introdujimos las notaciones $\nabla g(b) = \big(\partial_1 g(b), \dots, \partial_n g(b)\big)$ *que llamamos* vector gradiente *de* g *en el punto* b, *y* $\lambda'(a) = \big(\lambda_1'(a), \dots, \lambda_n'(a)\big)$, *que llamamos* vector tangente *a la curva* λ *en el punto* a.

Observación. En el caso particular, en el que la curva es una recta (o un segmento abierto de recta) $\lambda(t) = b + tv$, como el vector tangente $\lambda'(t) = v$ (es constante), el corolario anterior nos da

$$\phi'(a) = \langle \nabla g(b), v\rangle = \frac{\partial g}{\partial v}(b)$$

donde en la última igualdad utilizamos la definición de derivada direccional, y el corolario se reduce a un resultado conocido.

Además, tomando $\lambda\colon (-\varepsilon, \varepsilon) \to \mathbb{R}^n$ una curva arbitraria, diferenciable en $a = 0$ que verifique $\lambda(0) = b$, $\lambda'(0) = v$, tenemos, con $\phi_\lambda = g \circ \lambda$

$$\phi_\lambda'(a) = \langle \nabla g(b), v\rangle = \frac{\partial g}{\partial v}(b)$$

que es el mismo resultado. Concluímos, que para calcular la derivada direccional de g en el punto b con respecto de v, podemos componer g con *cualquier* curva λ que verifique $\lambda(0) = b$, $\lambda'(0) = v$.