Análisis Matemtico. Curso 2020-21.

Resumen de las semanas 3 y 4

Oes de Landau. Se utilizan para expresar una comparación entre dos funciones $f(x)$ y $\varphi(x)$. Habitualmente $\varphi(x)$ tiene un punto especial x_0, en el que vale cero o infinito, y para $x \neq x_0$ tenemos $0 < \varphi(x) < \infty$. Entonces la fórmula "$f(x) = \mathrm{O}\big(\varphi(x)\big)$" (o grande de Landau) significa lo siguiente cuando f es una función escalar:

$$|f(x)| \;\leq\; \text{constante} \cdot \varphi(x) \quad , \quad \text{para } x \quad \text{cercano a } x_0 \; ,$$

y significa esto otro cuando f es una función vectorial:

$$\|f(x)\| \;\leq\; \text{constante} \cdot \varphi(x) \quad , \quad \text{para } x \quad \text{cercano a } x_0 \; .$$

La fórmula "$f(x) = \mathrm{O}(1)$" significa que $f(x)$ es acotada en una región previamente indicada.

Por su parte, la fórmula "$f(x) = \mathrm{o}\big(\varphi(x)\big)$" (o pequeña de Landau) significa lo siguiente cuando f es una función escalar:

$$\lim_{x \to x_0} \frac{|f(x)|}{\varphi(x)} \;=\; 0 \; ,$$

y significa esto otro cuando f es una función vectorial:

$$\lim_{x \to x_0} \frac{\|f(x)\|}{\varphi(x)} \;=\; 0 \; .$$

Diferenciabilidad y desarrollo de Taylor de orden 1.

Una función $f(x)$, escalar o vectorial, se dice **diferenciable en $\boldsymbol{x_0}$** si está definida en un entorno de x_0 y admite un **desarrollo de Taylor de orden 1 centrado en $\boldsymbol{x_0}$,** es decir que se tiene:

$$f(x_0 + h) \;=\; P_1(x_0 + h) + \mathrm{o}\big(\|h\|\big) \; , \tag{1}$$

Siendo P_1 un polinomio de grado no mayor que 1 (escalar o vectorial, según sea f). Esto obliga a que haya una función lineal L tal que

$$P_1(x_0 + h) \;=\; f(x_0) + L(h) \, , \quad \text{es decir} \quad P(x) \;=\; f(x_0) + L(x - x_0) \; .$$

Una tal f es continua en x_0 (el recíproco no es cierto). El polinomio P_1 está determinado, de manera única, por una cualquiera de las condiciones siguientes (que son equivalentes):

1. Es un polinomio de grado ≤ 1 y además $f(x) - P_1(x) = \mathrm{o}\big(\|x - x_0\|^1\big)$.
2. Es un polinomio de grado ≤ 1, satisface $P_1(x_0) = f(x_0)$ y también $(P_1)_{x_i}(x_0) = f_{x_i}(x_0)$ para $i = 1, \ldots, n$. Es decir, las derivadas de órdenes 0 y 1 de P_1 en x_0 coninciden con las correspondientes derivadas de f en este mismo punto.

Es consecuencia del desarrollo (1) que f es derivable en $t = 0$ a lo largo de *cualquier* camino diferenciable $\alpha(t)$ con $\alpha(0) = x_0$, y, sobre todo, *la derivada sólo depende de la velocidad con la que el camino pasa por* x_0, concretamente:

$$\alpha(0) = x_0 \quad \Longrightarrow \quad \left.\frac{d}{dt}\right|_{t=0} f\big(\alpha(t)\big) \;=\; L\big(\alpha'(0)\big) \; . \tag{2}$$

Esto es un caso particular de la **regla de la cadena,** que repasamos más abajo. En particular, podemos calcular la función lineal L así:

$$\text{para todo } h \in \mathbb{R}^n \quad , \quad L(h) \;=\; D_h f(x_0) \;\stackrel{\text{def}}{=}\; \left.\frac{d}{dt}\right|_{t=0} f\big(\, x_0 + th \,\big) \; . \tag{3}$$

La función lineal L se llama **diferencial de f en $\boldsymbol{x_0}$,** se denota $\boldsymbol{(df)_{x_0}}$ y puede calcularse por la siguiente fórmula (tomando h como una columna):

$$(df)_{x_0}(h) \;=\; h_1\, f_{x_1}(x_0) + h_2\, f_{x_2}(x_0) + \cdots + h_n\, f_{x_n}(x_0) \;=\; \big[\; f_{x_1}(x_0) \mid f_{x_2}(x_0) \mid \;\cdots\; \mid f_{x_n}(x_0) \;\big]\, h\ ,$$

porque si $\{\mathbf{e}_1, \ldots, \mathbf{e}_n\}$ es la base estándar de $\mathbb{R}^n$ entonces la fórmula (3) nos da:

$$L(\mathbf{e}_i) \;=\; D_{\mathbf{e}_i} f(x_0) \;=\; f_{x_i}(x_0) \quad , \quad i = 1, 2, \ldots n\ .$$

La matriz $(Df)_{x_0} \;=\; \big[\; f_{x_1}(x_0) \mid \;\cdots\; \mid f_{x_n}(x_0) \;\big]$ se llama **matriz jacobiana de f en el punto $\boldsymbol{x_0}$**. Esta matriz es una **fila $\boldsymbol{1 \times n}$,** cuando f es escalar, y es $\boldsymbol{k \times n}$ cuando f es función vectorial con valores en $\mathbb{R}^k$.

Las derivadas $f_{x_i}(x_0)$ son las **columnas,** no las filas, de la jacobiana de f en x_0.

Atención. Para $n \geq 2$ existen funciones f derivables a lo largo de las rectas que pasan por x_0 y que cumplen la fórmula (3) para alguna función lineal L, y sin embargo o no son derivables a lo largo de algunos caminos no rectilíneos al pasar por x_0, o no cumplen (2) para algunos de esos caminos. Algunas de estas funciones pueden, incluso ¡ser discontinuas en x_0! Tales funciones no admiten un desarrollo del tipo (1), con lo cual la función lineal L tiene muy poca utilidad para ellas. No son diferenciables en x_0.

Afortunadamente, tenemos una condición que se cumple frecuentemente y es suficiente (no necesaria) para que una función de dos o más variables sea diferenciable:

Si para x cercano a x_0 existen las derivadas $f_{x_i}(x)$, $i = 1, \ldots, n$, y son funciones continuas de x, entonces $f(x)$ es diferenciable en x_0.

Para una función de una sola variable $\alpha(t)$, escalar o vectorial, la existencia de la derivada $\alpha'(t_0)$ es suficiente para que la función sea diferenciable en t_0: $\alpha(t_0 + h) = h\,\alpha'(t_0) + \mathrm{o}(|h|)$.
A una función continua $\alpha(t)$: intervalo $\to E$, con $E \subseteq \mathbb{R}^k$, la llamamos **camino en E.**

Regla de la cadena para diferenciales. Si f es diferenciable en x_0 y g es diferenciable en $y_0 = f(x_0)$, entonces la función compuesta $g \circ f$ es diferenciable en x_0 y su diferencial es la compuesta de las diferenciales:

$$\big(\, d(g \circ f) \,\big)_{x_0} \;=\; (dg)_{f(x_0)} \circ (df)_{x_0}\ .$$

Regla de la cadena para jacobianas. En las misma condiciones del párrafo anterior, se tiene:

$$D(g \circ f)_{x_0} \;=\; (Dg)_{f(x_0)}\, Df_{x_0}\ . \tag{4}$$

Es decir, la matriz jacobiana de la compuesta es el resultado de multiplicar las matrices jacobianas de las dos funciones componentes, escribiendo dichas matrices en el mismo orden en que se escribe la composición $g \circ f$.

En particular, si $f(t) \equiv \big(y_1(t), \ldots, y_m(t)\big)$: intervalo $\to \mathbb{R}^m$ es un **camino diferenciable** y $g(y_1, \ldots, y_m)$ es una función diferenciable en cada punto $f(t)$ del camino, entonces

$$\boxed{\frac{d}{dt}\, g\big(f(t)\big) \;=\; \frac{d}{dt}\, g\big(\, y_1(t)\,, \ldots\,, y_m(t)\,\big) \;=\; \frac{dy_1}{dt}\, g_{y_1}\big(f(t)\big) + \cdots + \frac{dy_m}{dt}\, g_{y_m}\big(f(t)\big) \;\in \mathbb{R}^k} \tag{5}$$

La regla general (4), con $f(x) \equiv \big(y_1(x), \ldots, y_m(x)\big)$ función de n variables, consiste en n igualdades cada una de las cuales es el resultado de elegir un $i \in \{1, 2, \ldots, n\}$ y sustituir por doquier en (5) la letra t por la x y el operador $\dfrac{d}{dt}$ por el $\dfrac{\partial}{\partial x_i}$.

Funciones de clase $\mathcal{C}^k$.

Dado un abierto $U \subseteq \mathbb{R}^n$. Se dice que $f : U \to \mathbb{R}^m$ es **diferenciable de clase $\mathcal{C}^k$ en U** si existen y son continuas en todo U las derivadas parciales de f de órdenes desde cero hasta k (entendiendo que la derivada de orden cero es la propia f).
Decimos que **f es $\mathcal{C}^\infty$ en U** o que **es suave en U,** si es $\mathcal{C}^k$ en U para todo k.

El **teorema de Schwarz** dice que, para una tal función, el orden de derivación no importa (en derivadas de orden no mayor que k). Como primer ejemplo, si f es $\mathcal{C}^2$ entonces $f_{x_1 x_2} \equiv f_{x_2 x_1}$. Como segundo ejemplo, si $f(x, y, z)$ es $\mathcal{C}^3$ entonces

$$f_{xxy} \equiv f_{xyx} \equiv f_{yxx} \quad \text{y también} \quad f_{xyz} \equiv f_{xzy} \equiv f_{yxz} \equiv f_{yzx} \equiv f_{zxy} \equiv f_{zyx} \ .$$

Pero sí importa, y mucho, cuántas veces se deriva respecto de cada variable. Por ejemplo, para la mayoría de las $f(x, y)$ de clase $\mathcal{C}^3$ la función f_{xyy} es distinta de la función f_{xxy}.

La suma de funciones $\mathcal{C}^k$ es $\mathcal{C}^k$. Al multiplicar dos funciones de clase $\mathcal{C}^k$ (escalares o matriciales) resulta una función de clase $\mathcal{C}^k$. La compuesta de funciones $\mathcal{C}^k$ es $\mathcal{C}^k$. Etc, etc.

Desarrollo de Taylor de orden 2.

Si f es de clase $\mathcal{C}^2$ en un entorno del punto $x_0 \in \mathbb{R}^n$, definimos la **matriz hessiana de f en x_0** como la tabla de las derivadas segundas:

$$\mathrm{Hess}\,(f)_{x_0} \ = \ \left[\, f_{x_i x_j} \,\right]_{1 \le i,j \le n} ,$$

que es una matriz simétrica $n \times n$. Definimos la **forma hessiana de f en x_0** como la forma cuadrática en $\mathbb{R}^n$ cuya matriz en la base estándar es la matriz hessiana:

$$\mathrm{Hess}\,(f)_{x_0}(v) \ = \ v^t \left(\mathrm{Hess}(f)_{x_0}\right) v \quad \text{para todo } \ v \in \mathbb{R}^n \ .$$

Podemos determinar esta forma cuadrática, de manera única, por la siguiente identidad:

$$\left.\frac{d^2}{dt^2}\right|_{t=0} f\big(\,x_0 + tv\,\big) \ = \ \mathrm{Hess}\,(f)_{x_0}(v) \quad \text{para todo } \ v \in \mathbb{R}^n \ . \tag{6}$$

El **desarrollo de Taylor de orden 2 centrado en x_0** de una $f \in \mathcal{C}^2$ se define por la fórmula:

$$f(x_0 + h) \ = \ P_2(x_0 + h) + \mathrm{o}\big(\|h\|^2\big) \ ,$$

donde P_2 es un polinomio de grado no mayor que 2. Este polinomio está determinado, de manera única, por una cualquiera de las condiciones siguientes (que son equivalentes):

1. Es un polinomio de grado ≤ 2 y $f(x) - P_2(x) = \mathrm{o}\big(\|x - x_0\|^2\big)$.

2. Es un polinomio de grado ≤ 2 y la derivadas parciales de órdenes 0, 1 y 2 de P_2 en x_0 coinciden con las correspondientes derivadas parciales de f en este mismo punto.

Tenemos la siguiente fórmula:

$$P_2(x_0 + h) \ = \ f(x_0) + (df)_{x_0}(h) + \frac{1}{2!}\,\mathrm{Hess}\,(f)_{x_0}(h) \ , \tag{7}$$

equivalente a esta otra:

$$P_2(x) \ = \ f(x_0) + (Df)_{x_0}\,(x - x_0) + \frac{1}{2}\,(x - x_0)^t \left(\mathrm{Hess}\,(f)_{x_0}\right)(x - x_0) \ .$$

Tenemos, pues, los desarrollos equivalentes:

$$f(x_0 + h) \ = \ f(x_0) + (df)_{x_0}(h) + \frac{1}{2}\,\mathrm{Hess}\,(f)_{x_0}(h) + \mathrm{o}\big(\|h\|^2\big) \ , \tag{8}$$

$$f(x) \ = \ f(x_0) + (Df_{x_0})\,(x - x_0) + \frac{1}{2}\,(x - x_0)^t \left(\mathrm{Hess}\,(f)_{x_0}\right)(x - x_0) + \mathrm{o}\big(\|x - x_0\|^2\big) \ .$$

Máximos y mínimos locales.

Punto crítico de f. Es cualquier punto en el que se anulan todas las derivadas parciales primeras de f.

Los siguientes resultados son consecuencia de las fórmulas (3) y (6):

1. Para que x_0 sea un **máximo local** de f es necesario (no suficiente) que x_0 sea un punto crítico de f y que la hessiana de f en x_0 sea semidefinida negativa.

2. Para que x_0 sea un **mínimo local** de f es necesario (no suficiente) que x_0 sea un punto crítico de f y que la hessiana de f en x_0 sea semidefinida positiva.

Los siguientes resultados se obtienen a partir de (8), viendo que la parte $\mathrm{o}\big(\|h\|^2\big)$ se puede acotar por un múltiplo arbitrariamente pequeño de $|\mathrm{Hess}\,(f)_{x_0}(h)|$

1. Para que x_0, sea un **máximo local estricto** de f es suficiente (no necesario) que x_0 sea un punto crítico de f y que la hessiana de f en x_0 sea definida negativa.

2. Para que x_0, sea un **mínimo local estricto** de f es suficiente (no necesario) que x_0 sea un punto crítico de f y que la hessiana de f en x_0 sea definida positiva.

Desarrollo de Taylor de orden k.

Para f de clase $\mathcal{C}^k$ en un entorno de x_0, existe un polinomio $P_k(x)$ determinado de manera única por una cualquiera de las condiciones siguientes (que son equivalentes):

1. Es un polinomio de grado $\leq k$ y $f(x) - P_k(x) = \mathrm{o}\big(\|x - x_0\|^k\big)$.

2. Es un polinomio de grado $\leq k$ y la derivadas parciales de órdenes $0, 1, 2, \ldots, k$ de P_k en x_0 coinciden con las correspondientes derivadas parciales de f en este mismo punto.

Para una tal función, el **desarrollo de Taylor de orden k centrado en $\boldsymbol{x_0}$** es la fórmula

$$f(x) \;=\; P_k(x) + \mathrm{o}\big(\|x - x_0\|^k\big) \;,$$

y llamamos **resto de Taylor de orden k** a la parte $\mathrm{o}\big(\|x - x_0\|^k\big)$ de dicha fórmula.

La condición 1. arriba indicada permite muchas veces determinar P_k. También podemos calcularlo por la condición 2., que equivale a la siguiente fórmula:

$$P_k(x_0 + h) \;=\; f(x_0) + \varphi_1(h) + \frac{1}{2!}\,\varphi_2(h) + \frac{1}{3!}\,\varphi_3(h) + \cdots + \frac{1}{k!}\,\varphi_k(h) \;,$$

siendo cada $\varphi_j(h)$ el polinomio homogéneo de grado j definido por la siguiente identidad:

$$\left.\frac{d^j}{dt^j}\right|_{t=0} f\big(\,x_0 + th\,\big) \;=\; \varphi_j(h) \quad \text{para todo } \; h \in \mathbb{R}^n \;.$$