Desarrollo de Taylor y extremos en varias variables

El polinomio de Taylor en varias variables

Recordemos que para una función f de una variable, el *polinomio de Taylor de orden n* en a viene dado por

$$f(a)+\frac{f'(a)}{1!}(x-a)+\frac{f''(a)}{2!}(x-a)^2+\cdots+\frac{f^{(n)}(a)}{n!}(x-a)^n.$$

En varias variables es similar, salvo que ahora hay que acompañar cada variable con su derivada parcial. Por ejemplo, en dos variables, el polinomio de Taylor de orden 1 en el punto $\vec{a}=(a_1,a_2)$ es (notemos que $1!=1$)

$$f(\vec{a})+\frac{\partial f}{\partial x}(\vec{a})\ (x-a_1)+\frac{\partial f}{\partial y}(\vec{a})\ (y-a_2).$$

En general, en m variables, si $\vec{a}=(a_1,a_2,\dots,a_m)$ el polinomio de orden 1 es

$$f(\vec{a})+\frac{\partial f}{\partial x_1}(\vec{a})\ (x_1-a_1)+\frac{\partial f}{\partial x_2}(\vec{a})\ (x_2-a_2)+\cdots+\frac{\partial f}{\partial x_m}(\vec{a})\ (x_m-a_m)$$

que se puede abreviar como

$$\boxed{f(\vec{a})+\sum_{i=1}^{m}\frac{\partial f}{\partial x_i}(\vec{a})\ (x_i-a_i)} \tag{1}$$

o todavía más como $f(\vec{a})+\nabla f(\vec{a})\cdot(\vec{x}-\vec{a})$, donde ∇ es el *operador gradiente*, que seguro que ha aparecido en cursos de Física, definido como

$$\nabla f=\Big(\frac{\partial f}{\partial x_1},\frac{\partial f}{\partial x_2},\dots,\frac{\partial f}{\partial x_m}\Big).$$

En dimensión $m=2$, con $x_1=x$, $x_2=y$; al igualar (1) a z se obtiene la ecuación del *plano tangente*. Esto es lógico porque el plano tangente es la mejor aproximación lineal. En más dimensiones lo que resultan son *hiperplanos* tangentes (planos de dimensión m).

Para el polinomio de orden dos, hay que tener en cuenta que el producto de dos cosas de grado 1, como xy, se considera de grado 2. Entonces lo que hay que hacer es sumar a (1) los términos (notemos que $2!=2$)

$$\frac{1}{2}\frac{\partial^2 f}{\partial x_1^2}(\vec{a})\ (x_1-a_1)^2+\frac{1}{2}\frac{\partial^2 f}{\partial x_1\partial x_2}(\vec{a})\ (x_1-a_1)(x_2-a_2)+\cdots+\frac{1}{2}\frac{\partial^2 f}{\partial x_m^2}(\vec{a})\ (x_m-a_m)^2.$$

El caso de dos variables, que es bastante común, correspondería a

$$\frac{1}{2}\frac{\partial^2 f}{\partial x^2}(\vec{a})\ (x-a_1)^2+2\cdot\frac{1}{2}\frac{\partial^2 f}{\partial x\partial y}(\vec{a})\ (x-a_1)(y-a_2)+\frac{1}{2}\frac{\partial^2 f}{\partial y^2}(\vec{a})\ (y-a_2)^2$$

donde se ha supuesto $\frac{\partial^2 f}{\partial x \partial y} = \frac{\partial^2 f}{\partial x \partial y}$, que se cumple siempre que f tenga unas ligeras propiedades de regularidad (todas las funciones del curso).

Escrito con sumatorios, entonces el polinomio de Taylor de orden 2 en $\vec{a}$ es

$$\boxed{f(\vec{a}) + \sum_{i=1}^{m} \frac{\partial f}{\partial x_i}(\vec{a})\ (x_i - a_i) + \frac{1}{2}\sum_{i=1}^{m}\sum_{j=1}^{m} \frac{\partial^2 f}{\partial x_i \partial x_j}(\vec{a})\ (x_i - a_i)(x_j - a_j).} \tag{2}$$

En notación ultracompacta esto es $f(\vec{a}) + \nabla f(\vec{a}) \cdot (\vec{x} - \vec{a}) + (\vec{x} - \vec{a}) \cdot \mathcal{H}f(\vec{a})\ (\vec{x} - \vec{a})$, donde $\mathcal{H}f$ es la *matriz hessiana*, la que tiene como elemento ij la derivada parcial segunda $\partial^2 f/\partial x_i \partial x_j$, y $\vec{x}-\vec{a}$ se entiende como un vector columna para que tenga sentido multiplicarlo por la matriz.

En este curso, el caso más destacado es el de dos variables ($m = 2$). En este caso

$$\mathcal{H}f = \begin{pmatrix} \frac{\partial^2 f}{\partial x^2} & \frac{\partial^2 f}{\partial x \partial y} \\ \frac{\partial^2 f}{\partial y \partial x} & \frac{\partial^2 f}{\partial y^2} \end{pmatrix}$$

y si $\vec{a} = \vec{0} = (0,0)$, el polinomio de Taylor de orden 2 es

$$f(\vec{0}) + x\frac{\partial f}{\partial x}(\vec{0}) + y\frac{\partial f}{\partial y}(\vec{0}) + \frac{x^2}{2}\frac{\partial^2 f}{\partial x^2}(\vec{0}) + xy\frac{\partial^2 f}{\partial x \partial y}(\vec{0}) + \frac{y^2}{2}\frac{\partial^2 f}{\partial y^2}(\vec{0}), \tag{3}$$

donde se ha usado la igualdad de las derivadas parciales cruzadas.

Los polinomios de Taylor de orden mayor son análogos pero se usan mucho menos en la práctica.

Ya hemos mencionado que los polinomios de Taylor de orden 1 están relacionados con el plano tangente, los de orden dos tienen que ver con los extremos (máximos y mínimos) y, en general, para cualquier orden sirven para aproximar una función.

Ejemplo. Sea $f(x,y) = e^x \cos(x+y)$. Calculemos los polinomios de Taylor de órdenes 1 y 2 en el origen $(0,0)$. Unos cálculos prueban

$$\frac{\partial f}{\partial x} = e^x \cos(x+y) - e^x \operatorname{sen}(x+y), \qquad \frac{\partial f}{\partial y} = -e^x \operatorname{sen}(x+y),$$

$$\frac{\partial^2 f}{\partial x^2} = -2e^x \operatorname{sen}(x+y), \qquad \frac{\partial^2 f}{\partial y^2} = -e^x \cos(x+y),$$

$$\frac{\partial^2 f}{\partial x \partial y} = \frac{\partial^2 f}{\partial x \partial y} = -e^x \operatorname{sen}(x+y) - e^x \cos(x+y).$$

Según (1), el polinomio de Taylor de orden 1 es

$$1 + 1 \cdot x + 0 \cdot y = 1 + x,$$

y, según (2), el de orden 2 es

$$1 + \big(1 \cdot x + 0 \cdot y\big) + \frac{1}{2}\big(0 \cdot x^2 - 1 \cdot xy - 1 \cdot yx - 1 \cdot y^2\big) = 1 + x - xy - \frac{1}{2}y^2.$$

Ejemplo. Sea $f(x,y)=x^2-y^2+2x$. Calculemos el plano tangente en $(1,1)$. Esto se reduce a calcular el polinomio de Taylor de orden 1 y éste requiere el gradiente.

$$\nabla f = \big(\frac{\partial f}{\partial x}, \frac{\partial f}{\partial y}\big) = (2x+2,-2y).$$

En el punto $(1,1)$ resulta $(4,-2)$. La ecuación del plano tangente es igual el polinomio de Taylor de orden 1 a z, esto es:

$$z = 2+4(x-1)-2(y-1) \qquad \text{o equivalentemente} \qquad z = 4x-2y.$$

Ejemplo. Con los resultados del primer ejemplo, veamos cuánto aproximan los polinomios de Taylor de órdenes 1 y 2 en el origen a la función $f(x,y)=e^x\cos(x+y)$ si $x=0.1$, $y=0.2$.

Si llamamos $T_!$ y T_2 a los polinomios de Taylor de órdenes 1 y 2 respectivamente, se tiene

$$\begin{cases} T_1(0.1,0.2) = 1+0.1 = 1.1 \\ T_2(0.1,0.2) = 1+0.1-0.1\cdot 0.2-\frac{1}{2}0.2^2 = 1.06. \end{cases}$$

Teniendo en cuenta que $f(0.1,0.2)=1.0558\dots$, la primera aproximación tiene un error relativo menor que un $5\,\%$ y la segunda menor que un $0.5\,\%$.

Referencias. Para recordar la fórmula de Taylor en una variable, los alumnos más avanzados agradecerán leer [Spi84]. Una lectura más ligera es [LE99a, §8.10]. El caso de varias variables de orden general y con término de error está en [MT98, §3.2].

Máximos y mínimos en varias variables.

El gradiente ∇f y la matriz hessiana $\mathcal{H}f$ son los análogos de f' y f'' en lo que respecta a los extremos relativos. Esencialmente, hay un máximo relativo o un mínimo relativo si $\nabla f = \vec{0}$ y $\mathcal{H}f$ es “negativa” o “positiva”, respectivamente. Por supuesto, no está nada claro qué sentido tiene el signo de una matriz. Aquí nos centraremos en el caso de dos variables, aunque trataremos de conservar cierta generalidad.

El *criterio de la derivada primera* en varias variables es:

Si una función diferenciable $f:\mathbb{R}^n \longrightarrow \mathbb{R}$ *alcanza un extremo (máximo o mínimo) relativo en* $\vec{a}$*, entonces* $\nabla f(\vec{a}) = \vec{0}$.

Habitualmente se dice que f tiene un *punto crítico* en $\vec{a}$ para referirse a la condición $\nabla f(\vec{a}) = \vec{0}$.

Geométricamente en el caso de dos variables se adivina la importancia de los puntos críticos, porque son aquellos en los cuales el plano tangente es horizontal. Si la función quedase por debajo de este plano se alcanzaría un máximo relativo y si quedase por encima un mínimo. Sin embargo decidir entre estas posibilidades no es obvio.

El *criterio de la derivada segunda* en varias variables para $f : \mathbb{R}^n \longrightarrow \mathbb{R}$ diferenciable es:

a) *Si* $\nabla f(\vec{a}) = \vec{0}$ *y* $\mathcal{H}f(\vec{a})$ *es* definida negativa *entonces* f *alcanza un máximo relativo en* $\vec{a}$.
b) *Si* $\nabla f(\vec{a}) = \vec{0}$ *y* $\mathcal{H}f(\vec{a})$ *es* definida positiva *entonces* f *alcanza un mínimo relativo en* $\vec{a}$.

Obviamente la pregunta es qué significa ser definida negativa y positiva. No veremos la definición general pero sí el siguiente criterio para el caso de dos variables.

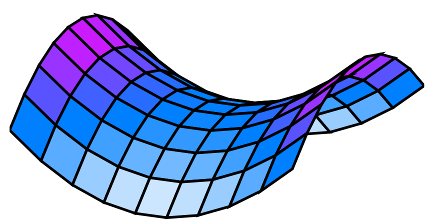
$\mathcal{H}f(\vec{a})$ *es definida negativa si* $\frac{\partial^2 f}{\partial x^2}(\vec{a}) < 0$ *y* $\det\big(\mathcal{H}f(\vec{a})\big) > 0$

$\mathcal{H}f(\vec{a})$ *es definida positiva si* $\frac{\partial^2 f}{\partial x^2}(\vec{a}) > 0$ *y* $\det\big(\mathcal{H}f(\vec{a})\big) > 0$.

¿Qué ocurre si no se aplica el criterio de la derivada segunda? ¿Quiere decir que no hay un extremo relativo? Esencialmente sí, siempre que la derivada parcial segunda y el determinante no se anulen.

Un *punto de silla* es un punto en el que caminando sobre la gráfica en una dirección observamos un mínimo pero en otra observamos un máximo. Corresponde a un dibujo del tipo:

Punto de silla
(silla de montar)



Entonces el criterio de la derivada segunda se completa a:

c) *Si* $\nabla f(\vec{a}) = \vec{0}$ *y* $\det\big(\mathcal{H}f(\vec{a})\big) < 0$ *entonces* f *alcanza un punto de silla en* $\vec{a}$.

No es sorprendente que en este caso se diga que $\mathcal{H}f(\vec{a})$ es *indefinida*.

En resumidas cuentas, el caso de dos variables se comporta como el de una cuando $\det \mathcal{H}f > 0$ *pensando que* f'' *es* $\frac{\partial^2 f}{\partial x^2}$, *mientras que si* $\det \mathcal{H}f < 0$ *hay un punto de silla.*

El caso de más variables es similar con criterios un poco más complicados para decidir si $\mathcal{H}f(\vec{a})$ es definida positiva, definida negativa o indefinida.

Ejemplo. Vamos a hallar los extremos relativos de $f(x,y) = 6x + 6y + 2xy - 5x^2 - 2y^2$. Para ello comenzamos calculando los puntos críticos planteando la ecuación $\nabla f = \vec{0}$:

$$\nabla f = \vec{0} \longleftrightarrow \begin{cases} \frac{\partial f}{\partial x} = 0 \\ \frac{\partial f}{\partial y} = 0 \end{cases} \longleftrightarrow \begin{cases} -10x + 2y + 6 = 0 \\ -4y + 2x + 6 = 0. \end{cases}$$

Al resolver el sistema se obtiene que el único punto crítico corresponde a $x = 1$, $y = 2$. Además

$$\frac{\partial^2 f}{\partial x^2}(1,2) = -10 < 0, \qquad \det\big(\mathcal{H}f(\vec{a})\big) = \begin{vmatrix} -10 & 2 \\ 2 & -4 \end{vmatrix} = 36 > 0.$$

Entonces en $(1,2)$ se alcanza un máximo relativo que es $f(1,2) = 9$.

Ejemplo. Estudiemos ahora el carácter de los puntos críticos de $f(x,y) = 3xy - x^3 - y^3$.Los puntos críticos son las soluciones de:

$$\nabla f = \vec{0} \longleftrightarrow \begin{cases} 3y - 3x^2 = 0 \\ 3x - 3y^2 = 0. \end{cases}$$

Para resolver el sistema basta despejar y sustituir y se obtienen los puntos $\vec{a} = (0,0)$ y $\vec{b} = (1,1)$. La matriz hessiana es:

$$\mathcal{H}f = \begin{pmatrix} -6x & 3 \\ 3 & -6y \end{pmatrix}, \quad \text{entonces} \quad \mathcal{H}f(\vec{a}) = \begin{pmatrix} 0 & 3 \\ 3 & 0 \end{pmatrix} \quad \text{y} \quad \mathcal{H}f(\vec{b}) = \begin{pmatrix} -6 & 3 \\ 3 & -6 \end{pmatrix}.$$

Los determinantes son -9 y 27, respectivamente. Según los criterios c) y a) en $\vec{a}$ se alcanza un punto de silla y en $\vec{b}$ un máximo relativo.

Referencias. El método para calcular extremos en dos variables, está sin muchas explicaciones en [LE99b, §12.8]. En [MT98, §3.3] se trata el tema desde una perspectiva más general y se muestra la relación con el desarrollo de Taylor aunque, de nuevo, se centra en el caso de dos variables. Los criterios generales para decidir si una matriz simétrica es definida positiva, definida negativa o indefinida, están por ejemplo en [HZ12].

Referencias

[HZ12] M.J. Hernández, E.; Vázquez and M.A. Zurro. *Álgebra lineal y Geometría.* Pearson/Addison Wesley, tercera edition, 2012.

[LE99a] R.P. Larson, R.E; Hostetler and B.H. Edwards. *Cálculo y Geometría Analítica.* (Vol.I). McGraw Hill, 1999.

[LE99b] R.P. Larson, R.E.; Hostetler and B.H. Edwards. *Cálculo y Geometría Analítica.* (Vol.II). McGraw Hill, sexta edition, 1999.

[MT98] J. Marsden and A. Tromba. *Cálculo Vectorial.* Pearson/Addison Wesley, cuarta edition, 1998.

[Spi84] M. Spivak. *Calculus Vol. I, II.* Editorial Reverté, Barcelona, 1984.