2.3 El teorema de Taylor

Dada una función suave de dos variables, $f(x, y)$, junto con un punto de referencia (x_0, y_0), el teorema de Taylor expresa $f(x, y)$ en términos de los valores de f y sus derivadas parciales en el punto (x_0, y_0).

Conviene escribir $x = x_0 + h$, $y = y_0 + k$ y considerar los *incremetos* h, k como las variables preferidas. Por ejemplo, el plano tangente a la superficie $z = f(x, y)$ en el punto (x_0, y_0, z_0) está dado por

$$z - z_0 = m_1(x - x_0) + m_2(y - y_0),$$

donde $m_1 = f_x(x_0, y_0)$, $m_2 = f_y(x_0, y_0)$. Entonces el plano se escribe como

$$z = z_0 + m_1 h + m_2 k, \quad \text{o bien} \quad z = f(x_0, y_0) + h\, f_x(x_0, y_0) + k\, f_y(x_0, y_0).$$

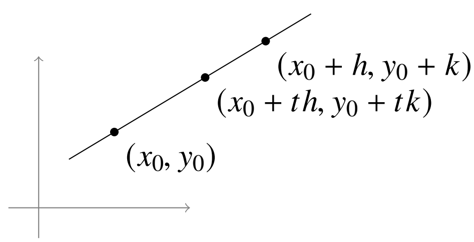
Como se sabe, el lado derecho es el polinomio de primer grado en (h, k) que mejor aproxima la verdadera función $z = f(x, y)$ cerca del punto (x_0, y_0). En síntesis,

$$f(x_0 + h, y_0 + k) = f(x_0, y_0) + h\, f_x(x_0, y_0) + k\, f_y(x_0, y_0) + R_1 f(h, k), \tag{2.1}$$

donde el "error" $R_1 f(h, k)$ es relativamente pequeño[1] en comparación con la distancia $\sqrt{h^2 + k^2}$ entre los puntos (x_0, y_0) y $(x_0 + h, y_0 + k)$.

Esta fórmula (2.1) expresa el teorema de Taylor hasta primer orden de derivadas. Para mejorar la aproximación, hace falta agregar un polinomio de segundo orden, de la forma $ah^2 + 2bhk + ck^2$, donde los coeficientes a, b, c dependen de derivadas parciales de f de segundo orden.

Para obtener estas coeficientes, hay un artificio que reduce el asunto a un problema de Cálculo I. Considérese la recta que pasa por los puntos (x_0, y_0) y $(x_0 + h, y_0 + k)$:



Esta recta se parametriza por una variable t. Los valores de $f(x, y)$ en los puntos de la recta definen una función de t, así:

$$g(t) := f(x_0 + th, y_0 + tk), \quad \text{para} \quad t \in \mathbb{R}.$$

Ahora es cuestión de aplicar el teorema de Taylor (en una variable) a esta función, cerca de $t = 0$:

$$g(t) = g(0) + g'(0)\, t + \tfrac{1}{2} g''(0)\, t^2 + R_2 g(t), \tag{2.2}$$

donde $\lim_{t\to 0} R_2 g(t)/t^2 = 0$. ⟦ Desde luego, se podría seguir hasta tercer orden, etcétera. ⟧

[1] Aquí la frase "relativamente pequeño" significamente que $\lim_{(h,k)\to(0,0)} R_1 f(h, k)/\sqrt{h^2 + k^2} = 0$.

Las derivadas de $g(t)$ se calculan *mediante la regla de la cadena*. Si $x(t) := x_0 + th$, $y(t) := y_0 + tk$, entonces $g(t) = f(x(t), y(t))$, lo cual implica que

$$\begin{aligned} g'(t) &= f_x(x(t), y(t))\, x'(t) + f_y(x(t), y(t))\, y'(t) \\ &= h\, f_x(x(t), y(t)) + k\, f_y(x(t), y(t)) \\ &= \big[h\, f_x + k\, f_y\big](x(t), y(t)). \end{aligned}$$

La derivada de segundo orden usa la regla de la cadena para las funciones f_x y f_y respectivamente:

$$\begin{aligned} g''(t) &= h\big[f_{xx}(x(t), y(t))\, x'(t) + f_{xy}(x(t), y(t))\, y'(t)\big] + k\big[f_{yx}(x(t), y(t))\, x'(t) + f_{yy}(x(t), y(t))\, y'(t)\big] \\ &= h^2\, f_{xx}(x(t), y(t)) + 2hk\, f_{xy}(x(t), y(t)) + k^2\, f_{yy}(x(t), y(t)) \\ &= \big[h^2\, f_{xx} + 2hk\, f_{xy} + k^2\, f_{yy}\big](x(t), y(t)). \end{aligned}$$

(Fíjese que se ha empleado la igualdad $f_{xy} = f_{yx}$ entre las "derivadas parciales mixtas" de 2º orden.)

Al evaluar estas expresiones en $t = 0$, se obtiene el teorema de Taylor de segundo orden para f en (x_0, y_0):

$$\begin{aligned} &f(x_0 + h, y_0 + k) \\ &\qquad = f(x_0, y_0) + \big[h\, f_x + k\, f_y\big](x_0, y_0) + \frac{1}{2}\big[h^2\, f_{xx} + 2hk\, f_{xy} + k^2\, f_{yy}\big](x_0, y_0) + R_2 f(h, k), \end{aligned} \tag{2.3}$$

donde el resto $R_2 f(h, k)$ es relativamente pequeño en comparación con $(h^2 + k^2)$; es decir, $\lim_{(h,k)\to(0,0)} R_2 f(h, k)/(h^2 + k^2) = 0$.

En lo que sigue, la estimación del resto $R_2 f$ no es lo que más nos concierne, sino el cálculo de los términos de segundo orden. Estos se pueden escribir en términos de una matriz:

$$\begin{bmatrix} h & k \end{bmatrix} \begin{bmatrix} f_{xx} & f_{xy} \\ f_{xy} & f_{yy} \end{bmatrix} \begin{bmatrix} h \\ k \end{bmatrix} = h^2\, f_{xx} + 2hk\, f_{xy} + k^2\, f_{yy}\,.$$

Esta matriz merece una notación, $Hf(x, y)$, y un nombre:[2] es el **hessiano** de la función f. Nótese bien que es una *matriz simétrica*, en virtud de la igualdad $f_{xy} = f_{yx}$.

Para funciones de tres variables, el teorema de Taylor funciona de igual manera. Hasta primer orden, se puede escribir:

$$\begin{aligned} &f(x_0 + h, y_0 + k, z_0 + l) \\ &\qquad = f(x_0, y_0, z_0) + h\, f_x(x_0, y_0, z_0) + k\, f_y(x_0, y_0, z_0) + l\, f_z(x_0, y_0, z_0) + R_1 f(h, k, l) \\ &\qquad = f(x_0, y_0, z_0) + \big[h\, f_x + k\, f_y + l\, f_z\big](x_0, y_0, z_0) + R_1 f(h, k, l). \end{aligned}$$

Los términos de segundo orden ahora forman un polinomio de segundo grado en (h, k, l), cuyos coeficientes son las derivadas parciales de segundo orden de f, organizados en una matriz simétrica.

[2] En homenaje al matemático alemán Otto Hesse, quien lo empleó en sus trabajos de mediados del siglo XIX.

La fórmula de Taylor, hasta segundo orden, es

$$
\begin{aligned}
&f(x_0+h, y_0+k, z_0+l) \\
&\qquad = f(x_0,y_0,z_0) + \big[h\,f_x + k\,f_y + l\,f_z\big](x_0,y_0,z_0) \\
&\qquad\qquad + \frac{1}{2}\big[h^2\,f_{xx} + 2hk\,f_{xy} + 2hl\,f_{xz} + k^2\,f_{yy} + 2kl\,f_{yz} + l^2\,f_{zz}\big](x_0,y_0,z_0) + R_2f(h,k,l).
\end{aligned}
$$

El hessiano $Hf(x_0,y_0,z_0)$, una matriz 3×3, aparece en la expresión:

$$
\begin{bmatrix} h & k & l \end{bmatrix}
\begin{bmatrix} f_{xx} & f_{xy} & f_{xz} \\ f_{xy} & f_{yy} & f_{yz} \\ f_{xz} & f_{yz} & f_{zz} \end{bmatrix}
\begin{bmatrix} h \\ k \\ l \end{bmatrix}
= h^2\,f_{xx} + 2hk\,f_{xy} + 2hl\,f_{xz} + k^2\,f_{yy} + 2kl\,f_{yz} + l^2\,f_{zz}\,.
$$

2.4 Máximos y mínimos: puntos críticos

Considérese de nuevo una función de dos variables, $f(x,y)$, con derivadas parciales continuas hasta segundo orden (al menos). Un punto (x_0,y_0) de su dominio es un **punto crítico** de f si sus derivadas parciales de primer orden se anulan en ese punto:

$$
f_x(x_0,y_0) = 0, \qquad f_x(x_0,y_0) = 0.
$$

Esto significa dos cosas. En primer lugar, si $z_0 := f(x_0,y_0)$, la superficie $z = f(x,y)$ posee un *plano tangente horizontal* en el punto (x_0,y_0,z_0).

En segundo lugar, *el término de primer orden* del desarrollo de Taylor de f en (x_0,y_0) *se anula*. Por lo tanto, la forma del grafo en torno a (x_0,y_0,z_0) se rige por el término de segundo orden; la fórmula (2.3) se convierte en:

$$
f(x_0+h, y_0+k) = f(x_0,y_0) + \frac{1}{2}\big[h^2\,f_{xx} + 2hk\,f_{xy} + k^2\,f_{yy}\big](x_0,y_0) + R_2f(h,k).
$$

Ahora todo dependo de el hessiano $Hf(x_0,y_0)$, evaluado en el punto crítico. Si la cantidad $\frac{1}{2}(h^2\,f_{xx} + 2hk\,f_{xy} + k^2\,f_{yy})$ es *positiva* para cualquier $(h,k) \neq (0,0)$, esto significa que al valor $f(x_0,y_0)$ se estará sumando algo positivo y la altura del grafo $f(x_0+h, y_0+k)$ será mayor. Esto a su vez significa que el valor $f(x_0,y_0)$ es un **mínimo local** de la función f.

Análogamente, si $\frac{1}{2}(h^2\,f_{xx} + 2hk\,f_{xy} + k^2\,f_{yy})$ es *negativa* para cualquier $(h,k) \neq (0,0)$, los puntos alrededor de (x_0,y_0,z_0) tendrán menor altura: brevemente, el valor $f(x_0,y_0)$ es un **máximo local** de f.

Hay una tercera posibilidad: para algunos $(h,k) \neq (0,0)$, la función cuadrática de marras es positiva pero para otros (h,k) es negativa. Si esto sucede, el plano tangente horizontal en (x_0,y_0,z_0) *cruza la superficie*, con algunos puntos cercanos de la superficie por encima y otros por debajo del plano tangente. En tal caso, (x_0,y_0,z_0) se llama un **punto de ensilladura** (o un "punto de silla").

El signo de $\frac{1}{2}(h^2\,f_{xx} + 2hk\,f_{xy} + k^2\,f_{yy})$ depende de las propiedades de la matriz simétrica $Hf(x_0,y_0)$. El caso más simple es cuando la matriz es diagonal – cuando $f_{xy}(x_0,y_0) = 0$.

Ejemplo 2.1. Análisis de la función $f(x, y) := 2x - x^2 + 2y^2 - y^4$.

Primero, se calculan las derivadas parciales de primer orden:

$$f_x = 2 - 2x, \qquad f_y = 4y - 4y^3.$$

Para obtener los puntos cr]íticos, se colocan estas iguales a cero:

$$\left\{\begin{aligned} 2 - 2x &= 0 \\ 4y - 4y^3 &= 0 \end{aligned}\right\} \implies \left\{\begin{aligned} &x = 1 \\ &y = 0, \ \pm 1 \end{aligned}\right\} \implies (x, y) = \begin{cases} (1, 0), \\ (1, 1), \\ (1, -1). \end{cases}$$

Entonces hay (exactamente) tres puntos críticos. Se debe analizarlos casos por caso. En primer lugar, se calcula el hessiano:

$$f_{xx} = -2, \quad f_{xy} = 0, \quad f_{yy} = 4 - 12y^2; \qquad Hf(x, y) = \begin{bmatrix} -2 & 0 \\ 0 & 4 - 12y^2 \end{bmatrix}.$$

- En el punto crítico $(1, 1)$, $Hf(1, 1) = \begin{bmatrix} -2 & 0 \\ 0 & -8 \end{bmatrix}$, con lo cual

$$\tfrac{1}{2}(h^2 f_{xx} + 2hk\, f_{xy} + k^2 f_{yy})(1, 1) = -h^2 - 4k^2 < 0 \quad \text{para } (h, k) \neq (0, 0).$$

  Además, $f(1, 1) = 2$. El punto del grafo $(1, 1, 2)$ es un máximo local.

- En el punto crítico $(1, -1)$, sucede lo mismo: se obtiene $-h^2 - 4k^2 < 0$, así que $(1, -1, 2)$ es otro máximo local.

- En el punto crítico $(1, 0)$, se obtiene $Hf(1, 0) = \begin{bmatrix} -2 & 0 \\ 0 & 4 \end{bmatrix}$. Ahora la función $-h^2 + 2k^2$ toma valores positivos y negativos cerca de $(0, 0)$; en efecto, $(h, 0) \mapsto -h^2 < 0$ mientras $(0, k) \mapsto 2k^2 > 0$. Como $f(1, 0) = 1$, se concluye que $(1, 0, 1)$ es un punto de ensilladura.

Lo notable del ejemplo anterior es que en los tres puntos críticos el hessiano es una matriz diagonal. Esto permite determinar los signos del término cuadrática de modo directo.

En casos más generales, habrá que *diagonalizar la matriz* $Hf(x_0, y_0)$. Esto siempre es posible, porque esta es una matriz simétrica real. Además, los **autovalores** (*valores propios*, *eigenwerten*) de $Hf(x_0, y_0)$ son reales: son las entradas diagonales de la matriz diagonalizada.

Se continuará con casos no diagonales (y con tres variables) en la próxima clase.