En (1, 0, 2), estas derivadas parciales son 2 y 1, respectivamente. Por tanto, de acuerdo con la Ecuación (1), el plano tangente es

$$z = 2(x-1) + 1(y-0) + 2, \qquad \text{es decir,} \qquad z = 2x + y. \qquad \blacktriangle$$

Escribimos $\mathbf{D}f(x_0, y_0)$ para denotar la matriz fila

$$\left[\frac{\partial f}{\partial x}(x_0, y_0) \quad \frac{\partial f}{\partial y}(x_0, y_0)\right],$$

de modo que la definición de diferenciabilidad afirma que

$$\begin{aligned} f(x_0, y_0) &+ \mathbf{D}f(x_0, y_0)\begin{bmatrix} x - x_0 \\ y - y_0 \end{bmatrix} \\ &= f(x_0, y_0) + \left[\frac{\partial f}{\partial x}(x_0, y_0)\right](x - x_0) + \left[\frac{\partial f}{\partial y}(x_0, y_0)\right](y - y_0) \end{aligned} \qquad (3)$$

es una buena aproximación de f cerca de (x_0, y_0). Como anteriormente, "buena" se toma en el sentido de que la expresión (3) difiere de $f(x, y)$ en algo pequeño multiplicado por $\sqrt{(x-x_0)^2 + (y-y_0)^2}$. Decimos que la expresión (3) es la ***mejor aproximación lineal*** de f cerca de (x_0, y_0).

Diferenciabilidad: caso general

Ahora ya estamos preparados para dar una definición de diferenciabilidad de funciones f de $\mathbb{R}^n$ en $\mathbb{R}^m$, basándonos en lo expuesto anteriormente. La derivada $\mathbf{D}f(\mathbf{x}_0)$ de $f = (f_1, \ldots, f_m)$ en un punto $\mathbf{x}_0$ es una matriz $\mathbf{T}$ cuyos elementos son $t_{ij} = \partial f_i / \partial x_j$ evaluados en $\mathbf{x}_0$.[2]

Definición Diferenciable, *n* variables, *m* funciones Sea U un conjunto abierto en $\mathbb{R}^n$ y sea $f\colon U \subset \mathbb{R}^n \to \mathbb{R}^m$ una función dada. Decimos que f es ***diferenciable*** en $\mathbf{x}_0 \in U$ si las derivadas parciales de f existen en $\mathbf{x}_0$ y si

$$\lim_{\mathbf{x} \to \mathbf{x}_0} \frac{\|f(\mathbf{x}) - f(\mathbf{x}_0) - \mathbf{T}(\mathbf{x} - \mathbf{x}_0)\|}{\|\mathbf{x} - \mathbf{x}_0\|} = 0, \qquad (4)$$

donde $\mathbf{T} = \mathbf{D}f(\mathbf{x}_0)$ es la matriz $m \times n$ con elementos $\partial f_i / \partial x_j$ evaluadas en $\mathbf{x}_0$ y $\mathbf{T}(\mathbf{x} - \mathbf{x}_0)$ es el producto de $\mathbf{T}$ por $\mathbf{x} - \mathbf{x}_0$ (visto como una matriz columna). Decimos que $\mathbf{T}$ es la ***derivada*** de f en $\mathbf{x}_0$.

[2] Resulta que solamente necesitamos postular la existencia de *alguna* matriz que proporcione la mejor aproximación lineal cerca de $\mathbf{x}_0 \in \mathbb{R}^n$, ya que de hecho esta matriz es *necesariamente* la matriz cuyo elemento ij-ésimo es $\partial f_i / \partial x_j$.