

1 Reminders from 1-dimensional calculus

Consider a function of one variable $f: D \rightarrow \mathbb{R}$ with domain $D \subset \mathbb{R}$. Around any point $x_0 \in D$, f can be approximated by a degree 1 polynomial

$$f(x) \approx f(x_0) + f'(x_0)(x - x_0)$$

called the **linearization** or **linear approximation** of f at x_0 , sometimes denoted $L(x)$. Here we are assuming that the derivative $f'(x_0)$ exists.

Example 1. Consider $f(x) = \sin(3x)$. Let us find the linear approximation of f at the point $x = 2$. The derivative is

$$f'(x) = 3 \cos(3x)$$

$$f'(2) = 3 \cos(6)$$

so that the linear approximation is

$$\begin{aligned} f(x) &\approx f(x_0) + f'(x_0)(x - x_0) \\ &= f(2) + f'(2)(x - 2) \\ &= \sin(6) + 3 \cos(6)(x - 2). \quad \square \end{aligned}$$

More generally, assuming f has enough derivatives, Taylor's theorem says that around x_0 , f can be approximated by a degree n polynomial

$$\begin{aligned} f(x) &\approx f(x_0) + f'(x_0)(x - x_0) + \frac{f''(x_0)}{2}(x - x_0)^2 + \dots + \frac{f^{(n)}(x_0)}{n!}(x - x_0)^n \\ &= \sum_{k=0}^n \frac{f^{(k)}(x_0)}{k!}(x - x_0)^k =: P_n(x) \end{aligned}$$

called the Taylor polynomial (of degree n) of f at x_0 .

Note that Taylor polynomials P_n are a good approximation of the function f only when x is close to x_0 . See for example the fun graphics here:

http://en.wikipedia.org/wiki/Taylor%27s_theorem

To emphasize that x should be close to x_0 , let us denote the difference $h := x - x_0$. We can thus rewrite the Taylor polynomials as

$$\begin{aligned} f(x_0 + h) &\approx f(x_0) + f'(x_0)h + \frac{f''(x_0)}{2}h^2 + \dots + \frac{f^{(n)}(x_0)}{n!}h^n \\ &= \sum_{k=0}^n \frac{f^{(k)}(x_0)}{k!}h^k \end{aligned}$$

and in particular the linear approximation of f around x_0 is

$$f(x_0 + h) \approx f(x_0) + f'(x_0)h.$$

Geometrically, the graph of the linear approximation L is the line that is tangent to the graph of f at x_0 .

2 Linear approximation and tangent planes

The notion discussed above has an analogue in multivariate calculus.

Consider a function of *two* variables $f: D \rightarrow \mathbb{R}$ with domain $D \subset \mathbb{R}^2$. Around any point $(x_0, y_0) \in D$, f can be approximated by a degree 1 polynomial

$$f(x, y) \approx f(x_0, y_0) + f_x(x_0, y_0)(x - x_0) + f_y(x_0, y_0)(y - y_0) =: L(x, y)$$

called the **linearization** or **linear approximation** of f at (x_0, y_0) . Here we are assuming that the partial derivatives $f_x(x_0, y_0), f_y(x_0, y_0)$ exist.

Example 2. Consider $f(x, y) = \sin(3x + y)$. Let us find the linear approximation of f at the point $(2, 1)$. The partial derivatives are

$$f_x(x, y) = 3 \cos(3x + y)$$

$$f_x(2, 1) = 3 \cos(7)$$

$$f_y(x, y) = \cos(3x + y)$$

$$f_y(2, 1) = \cos(7)$$

so that the linear approximation is

$$\begin{aligned} f(x, y) &\approx f(x_0, y_0) + f_x(x_0, y_0)(x - x_0) + f_y(x_0, y_0)(y - y_0) \\ &= f(2, 1) + f_x(2, 1)(x - 2) + f_y(2, 1)(y - 1) \\ &= \sin(7) + 3 \cos(7)(x - 2) + \cos(7)(y - 1). \quad \square \end{aligned}$$

There are Taylor polynomials in several variables, but they are slightly trickier and we will not discuss them here. The linear approximation $L(x, y)$ described above is the degree 1 Taylor polynomial of f .

Again, $L(x, y)$ is a good approximation of the function f only when (x, y) is close to (x_0, y_0) . To emphasize that, let us denote the differences $h := x - x_0, k := y - y_0$. We can thus rewrite the linear approximation as

$$f(x_0 + h, y_0 + k) \approx f(x_0, y_0) + f_x(x_0, y_0)h + f_y(x_0, y_0)k.$$

Geometrically, the graph of the linear approximation L is the plane that is tangent to the graph of f at (x_0, y_0) .

Upshot: Finding the tangent plane to the graph of a function is the same as finding the linear approximation of the function.

Example 3. Consider $f(x, y) = x^2y^3$. Let us find the equation of the plane tangent to the graph of f at the point $(2, 1, f(2, 1)) = (2, 1, 4)$. The partial derivatives are

$$\begin{aligned}f_x(x, y) &= 2xy^3 \\f_x(2, 1) &= 4 \\f_y(x, y) &= 3x^2y^2 \\f_y(2, 1) &= 12\end{aligned}$$

so that the linear approximation at $(2, 1)$ is

$$\begin{aligned}f(x, y) &\approx f(x_0, y_0) + f_x(x_0, y_0)(x - x_0) + f_y(x_0, y_0)(y - y_0) \\&= f(2, 1) + f_x(2, 1)(x - 2) + f_y(2, 1)(y - 1) \\&= 4 + 4(x - 2) + 12(y - 1).\end{aligned}$$

Therefore the tangent plane has equation

$$z = 4 + 4(x - 2) + 12(y - 1)$$

which can be rewritten as

$$z = 4x + 12y - 16. \square$$

3 Differentiability

Now we know how to approximate a function by a degree 1 polynomial, assuming the function is nice enough. How do we know that the linear approximation is reasonably good? Not all functions can be approximated by a degree 1 polynomial. That is what differentiability measures.

The function $f(x, y)$, close to (x_0, y_0) , can always be written as

$$f(x_0 + h, y_0 + k) = f(x_0, y_0) + f_x(x_0, y_0)h + f_y(x_0, y_0)k + E(h, k)$$

where $E(h, k)$ can be thought as the error term, the failure of f to be equal to its linear approximation at (x_0, y_0) .

Definition: The function f is **differentiable** at (x_0, y_0) if the error term $E(h, k)$ satisfies

$$\lim_{(h,k) \rightarrow (0,0)} \frac{|E(h, k)|}{\sqrt{h^2 + k^2}} = 0.$$

Analytically, this means that the linear approximation is “reasonably good”, in the sense that the error term has order higher than 1, and thus becomes negligible when (h, k) is very small.

Geometrically, it means that the graph of f has a tangent plane at the point $(x_0, y_0, f(x_0, y_0))$. By analogy, consider a function of one variable which is not differentiable, for example the absolute value function $f(x) = |x|$ at $x_0 = 0$. Then the graph of f at x_0 does not have a tangent line.

Facts about differentiability

Proposition: If f is differentiable at (x_0, y_0) , then f is continuous at (x_0, y_0) .

Proof: The condition

$$\lim_{(h,k) \rightarrow (0,0)} \frac{|E(h,k)|}{\sqrt{h^2 + k^2}} = 0$$

guarantees a fortiori

$$\lim_{(h,k) \rightarrow (0,0)} E(h,k) = 0.$$

Therefore, as (x, y) gets very close to (x_0, y_0) , we have

$$\begin{aligned} \lim_{(h,k) \rightarrow (0,0)} f(x_0 + h, y_0 + k) &= \lim_{(h,k) \rightarrow (0,0)} [f(x_0, y_0) + f_x(x_0, y_0)h + f_y(x_0, y_0)k + E(h, k)] \\ &= f(x_0, y_0) + 0 + 0 + 0 \\ &= f(x_0, y_0). \blacksquare \end{aligned}$$

Example 4: The function

$$f(x, y) = \begin{cases} \frac{xy}{x^2 + y^2} & \text{if } (x, y) \neq (0, 0) \\ 0 & \text{if } (x, y) = (0, 0) \end{cases}$$

is **not** differentiable at $(0, 0)$. In fact, it is not even continuous at $(0, 0)$ because the limit

$$\lim_{(x,y) \rightarrow (0,0)} \frac{xy}{x^2 + y^2}$$

does not exist.

However, note that f **does** have both partial derivatives at $(0, 0)$, since f is identically zero along the x -axis and y -axis. More precisely, we have

$$\begin{aligned} f_x(0, 0) &= \lim_{h \rightarrow 0} \frac{f(0 + h, 0) - f(0, 0)}{h} = \lim_{h \rightarrow 0} \frac{0}{h} = 0 \\ f_y(0, 0) &= \lim_{k \rightarrow 0} \frac{f(0, 0 + k) - f(0, 0)}{k} = \lim_{k \rightarrow 0} \frac{0}{k} = 0. \square \end{aligned}$$

If a function is differentiable at a point (x_0, y_0) , then it has both partial derivatives $f_x(x_0, y_0)$ and $f_y(x_0, y_0)$. However, example 4 shows that the converse does not hold. In other words, the existence of both partial derivatives **does not guarantee** differentiability. Morally, that is because limits in several variables are more complicated than limits in one variable.

Here is an even worse example.

Example 5: The function

$$f(x, y) = \begin{cases} \frac{y^3}{x^2 + y^2} & \text{if } (x, y) \neq (0, 0) \\ 0 & \text{if } (x, y) = (0, 0) \end{cases}$$

is **not** differentiable at $(0, 0)$. It **is** continuous at $(0, 0)$. Moreover, it has both partial derivatives

$$f_x(0, 0) = \lim_{h \rightarrow 0} \frac{f(0 + h, 0) - f(0, 0)}{h} = \lim_{h \rightarrow 0} \frac{0}{h} = 0$$

$$f_y(0, 0) = \lim_{k \rightarrow 0} \frac{f(0, 0 + k) - f(0, 0)}{k} = \lim_{k \rightarrow 0} \frac{k}{k} = 1.$$

In fact, it has all directional derivatives (which will be discussed in section 14.6). However, the error term

$$\begin{aligned} E(h, k) &= f(x_0 + h, y_0 + k) - f(x_0, y_0) - f_x(x_0, y_0)h - f_y(x_0, y_0)k \\ &= f(h, k) - f(0, 0) - f_x(0, 0)h - f_y(0, 0)k \\ &= \frac{k^3}{h^2 + k^2} - k \\ &= \frac{-h^2k}{h^2 + k^2} \end{aligned}$$

does **not** satisfy

$$\lim_{(h,k) \rightarrow (0,0)} \frac{|E(h, k)|}{\sqrt{h^2 + k^2}} = 0$$

since this limit does not exist. \square

If partial derivatives are not enough, how to tell if a function is differentiable? The following theorem provides sufficient conditions.

Theorem: If both partial derivatives f_x and f_y exist near a point (x_0, y_0) and are continuous at (x_0, y_0) , then f is differentiable at (x_0, y_0) .

Example 6: The function

$$f(x, y) = x^3y + e^xy^5 + \cos(xy)$$

is differentiable on all of \mathbb{R}^2 .