MA 105 D1 &D2 Lecture 16

Ravi Raghunathan

Department of Mathematics

December 21, 2020

Recap: Directional derivatives

The total derivative

More variables

Partial Derivatives

As before, U will denote a subset of $\mathbb{R}^2$. Given a function $f : U \to \mathbb{R}$, we can fix one of the variables and view the function f as a function of the other variable alone. We can then take the derivative of this one variable function.

To make things precise, fix x_2.
Definition: The partial derivative of $f : U \to \mathbb{R}$ with respect to x_1 at the point (a, b) is defined by

$$\frac{\partial f}{\partial x_1}(a, b) := \lim_{x_1 \to a} \frac{f((x_1, b)) - f((a, b))}{x_1 - a}.$$

Similarly, one can define the partial derivative with respect to x_2. In this case the variable x_1 is fixed and f is regarded only as a function of x_2:

$$\frac{\partial f}{\partial x_2}(a, b) := \lim_{x_2 \to b} \frac{f((a, x_2)) - f((a, b))}{x_2 - b}.$$

Directional Derivatives

The partial derivatives are special cases of the directional derivative. Let $v = (v_1, v_2)$ be a unit vector. Then v specifies a direction in $\mathbb{R}^2$.

Definition: The directional derivative of f in the direction v at a point $p = (p_1, p_2)$ is denoted by $\nabla_v f(x)$ and is defined as

$$\lim_{t \to 0} \frac{f(p + tv) - f(x)}{t} = \lim_{t \to 0} \frac{f((p_1 + tv_1, p_2 + tv_2)) - f((p_1, p_2))}{t}.$$

$\nabla_v f(x)$ measures the rate of change of the function f at x along the path $p + tv$.

If we take $v = (1, 0)$ in the above definition, we obtain $\frac{\partial f}{\partial x_1}(p)$, while $v = (0, 1)$ yields $\frac{\partial f}{\partial x_2}(p)$.

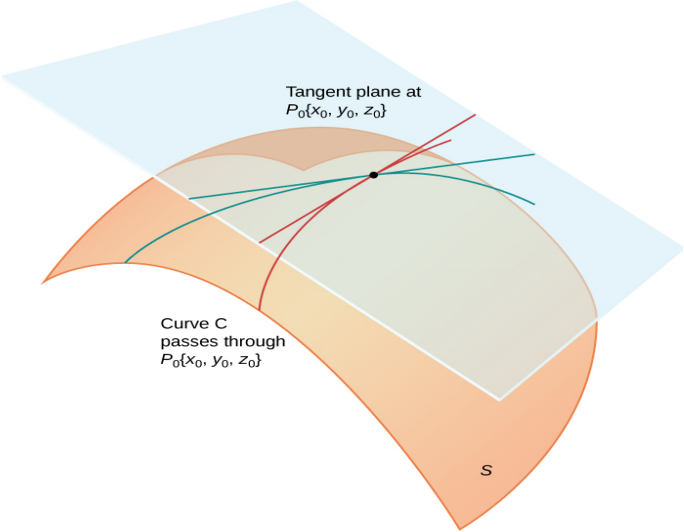
A function may have both partial derivatives, and even all directional derivatives, at a point but it need not be continuous. Exercise 5.5 gives an example of the former situation. Here is an example of the latter situation.

$$f(x, y) = \begin{cases} \frac{xy^2}{x^2+y^4} & \text{if } x \neq 0 \\ 0 & \text{if } x = 0 \end{cases}$$

Exercise 5.8 gives an example of a function in two variables for which neither partial derivative exists at $(0, 0)$ but such that the function is continuous at $(0, 0)$.

Exercise 5.11 gives an example of a function $f(x, y)$ such that f_x and f_y exist at $(0, 0)$ but none of the other directional derivatives exist.

The tangent plane in a picture



https://openstax.org/books/calculus-volume-3/pages/4-4-tangent-planes-and-linear-approximations

The tangent plane

Let $f(x, y)$ be a function which has both partial derivatives. In the two variable case we need to look at the distance between the surface $z = f(x, y)$ and its tangent plane.

Let us first recall how to find the equation of a plane passing through the point $P = (x_0, y_0, z_0)$. It is the graph of the function

$$z = g(x, y) = z_0 + a(x - x_0) + b(y - y_0).$$

Let us determine the tangent plane to $z = f(x, y)$ passing through a point $P = (x_0, y_0, z_0)$ *on the surface*. In other words, we have to determine the constants a and b.

If we fix the y variable and treat $f(x, y)$ only as a function of x, we get a curve. Similarly, if we treat $g(x, y)$ as function only of x, we obtain a line. The tangent to the curve must be the same as the line passing through (x_0, y_0, z_0), and, in any event, their slopes must be the same. Thus, we must have

$$\frac{\partial f}{\partial x}(x_0, y_0) = \frac{\partial g}{\partial x}(x_0, y_0) = a.$$

Arguing in exactly the same way, but fixing the x variable and varying the y variable we obtain

$$\frac{\partial f}{\partial y}(x_0, y_0) = \frac{\partial g}{\partial y}(x_0, y_0) = b.$$

Hence, the equation of the tangent plane to $z = f(x, y)$ at the point (x_0, y_0) is

$$z = f(x_0, y_0) + \frac{\partial f}{\partial x}(x_0, y_0)(x - x_0) + \frac{\partial f}{\partial y}(x_0, y_0)(y - y_0)$$

Differentiability for functions of two variables

We now define differentiability for functions of two variables by imitating the one variable definition, but using the "$\varepsilon_1(h)$" version.
We let $(x, y) = (x_0, y_0) + (h, k) = (x_0 + h, y_0 + k)$
Definition A function $f : U \to \mathbb{R}$ is said to be differentiable at a point (x_0, y_0) if $\frac{\partial f}{\partial x}(x_0, y_0)$, and $\frac{\partial f}{\partial y}(x_0, y_0)$ exist and

$$\lim_{(h,k)\to 0} \frac{\left| f(x_0 + h, y_0 + k) - f(x_0, y_0) - \frac{\partial f}{\partial x}(x_0, y_0)h - \frac{\partial f}{\partial y}(x_0, y_0)k \right|}{\|(h, k)\|} = 0,$$

This is saying that the distance between the tangent plane and the surface is going to zero even after dividing by $\|(h, k)\|$. We could rewrite this as

$$\left| f((x_0, y_0) + (h, k)) - f(x_0, y_0) - \frac{\partial f}{\partial x}(x_0, y_0)h - \frac{\partial f}{\partial y}(x_0, y_0)k \right| = \varepsilon_1(h, k)\|(h, k)\|$$

where $\varepsilon_1(h, k)$ is a function that goes to 0 as $\|(h, k)\| \to 0$. This form of differentiability now looks exactly like the one variable version.

The derivative as a linear map

We can rewrite the differentiability criterion once more as follows. We define the 1×2 matrix

$$Df(x_0, y_0) = \begin{pmatrix} \frac{\partial f}{\partial x}(x_0, y_0) & \frac{\partial f}{\partial y}(x_0, y_0) \end{pmatrix}.$$

A 1×2 matrix can be multiplied by a column vector (which is 2×1 matrix) to give a real number. In particular:

$$\begin{pmatrix} \frac{\partial f}{\partial x}(x_0, y_0) & \frac{\partial f}{\partial y}(x_0, y_0) \end{pmatrix} \begin{pmatrix} h \\ k \end{pmatrix} = \frac{\partial f}{\partial x}(x_0, y_0)h + \frac{\partial f}{\partial y}(x_0, y_0)k$$

that is,

$$Df(x_0, y_0) \begin{pmatrix} h \\ k \end{pmatrix} = \frac{\partial f}{\partial x}(x_0, y_0)h + \frac{\partial f}{\partial y}(x_0, y_0)k$$

The definition of differentiability can thus be reformulated using matrix notation.

Definition: The function $f(x, y)$ is said be differentiable at a point (x_0, y_0) if there exists a matrix denoted $Df((x_0, y_0))$ with the property that

$$f((x_0, y_0) + (h, k)) - f(x_0, y_0) - Df(x_0, y_0)\begin{pmatrix} h \\ k \end{pmatrix} = \varepsilon_1(h, k)\|(h, k)\|,$$

for some function $\varepsilon_1(h, k)$ which goes to zero as (h, k) goes to zero. Viewing the derivative as a matrix allows us to view it as a linear map from $\mathbb{R}^2 \to \mathbb{R}$. Given a 1×2 matrix A and two column vectors v and w, we see that

$$A \cdot (v + w) = A \cdot v + A \cdot w \quad \text{and} \quad A \cdot (\lambda v) = \lambda(A \cdot v),$$

for any real number λ. As we have seen before, functions satisfying the above two properties are called linear functions or linear maps. Thus, the map $v \to A \cdot v$ gives a linear map from $\mathbb{R}^2$ to $\mathbb{R}$.

The matrix $Df(x_0, y_0)$ is called the total derivative of the function $f(x, y)$ at the point (x_0, y_0).

The Gradient

When viewed as a row vector rather than as a matrix, the Derivative matrix is called the gradient and is denoted $\nabla f(x_0, y_0)$. Thus

$$\nabla f(x_0, y_0) = \left(\frac{\partial f}{\partial x}(x_0, y_0), \frac{\partial f}{\partial y}(x_0, y_0) \right).$$

In terms of the coordinate vectors **i** and **j** the gradient can be written as

$$\nabla f(x_0, y_0) = \frac{\partial f}{\partial x}(x_0, y_0)\mathbf{i} + \frac{\partial f}{\partial y}(x_0, y_0)\mathbf{j}.$$

A criterion for differentiability

Before we state the criterion, we note that with our definition of differentiability, every differentiable function is continuous.

Theorem 26: Let $f : U \to \mathbb{R}$. If the partial derivatives $\frac{\partial f}{\partial x}(x, y)$ and $\frac{\partial f}{\partial y}(x, y)$ exist and are continuous in a neighbourhood of a point (x_0, y_0) (that is in a region of the plane of the form $\{(x, y) \mid \|(x, y) - (x_0, y_0)\| < r\}$ for some $r > 0$. Then f is differentiable at (x_0, y_0).

We omit the proof of this theorem. However, we note that a function whose partial derivatives exist and are continuous is said to be continuously differentiable or of class $\mathcal{C}^1$. The theorem says that every function that is $\mathcal{C}^1$ in a small disc around a point is differentiable at that point.

Three variables

For the next few slides, we will assume that $f : U \to \mathbb{R}$ is a function of three variables, that is, U is a subset of $\mathbb{R}^3$. In this case, if we denote the variables by x, y and z, we get three partial derivatives as follows: we hold two of the variables constant and vary the third. For instance if y and z are kept fixed while x is varied, we get the partial derivative with respect to x at the point (a, b, c):

$$\frac{\partial f}{\partial x}(a, b, c) = \lim_{x \to a} \frac{f(x, b, c) - f(a, b, c)}{x - a}.$$

In a similar way we can define the partial derivatives

$$\frac{\partial f}{\partial y}(a, b, c) \quad \text{and} \quad \frac{\partial f}{\partial z}(a, b, c).$$

Once we have the three partial derivatives we can once again define the gradient of f:

$$\nabla f(a, b, c) = \left(\frac{\partial f}{\partial x}(a, b, c), \frac{\partial f}{\partial y}(a, b, c), \frac{\partial f}{\partial z}(a, b, c) \right).$$

Differentiability in three variables

Exercise 1: Formulate a definition of differentiability for a function of three variables.

Exercise 2: Formulate the analogue of Theorem 26 for a function of three variables.

We can also define differentiability for functions from $\mathbb{R}^m$ to $\mathbb{R}^n$ where m and n are any positive integers. We will do this in detail in this course when m and n have the values 1 and 2 and 3.

Finally, the rules for the partial derivatives of sums, differences, products and quotients of functions $f, g : U \to \mathbb{R}$, ($U \subset \mathbb{R}^m$, $m = 2, 3$) are exactly analogous to those for the derivative of functions of one variable.