MA 105 D1 Lecture 10

Ravi Raghunathan

Department of Mathematics

Autumn 2014, IIT Bombay, Mumbai

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 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 $x = (x_1, x_2)$ is defined as

$$\nabla_v = \lim_{t \to 0} \frac{f(x + tv) - f(x)}{t} = \lim_{t \to 0} \frac{f(x_1 + tv_1, x_2 + tv_2) - f((x_, x_2))}{t}.$$

It measures the rate of change of the function f along the path $x + tv$

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

Consider the function

$$f(x_1, x_2) = \begin{cases} 1 & \text{if } x_1 = 0 \text{ or if } x_2 = 0 \\ 0 & \text{otherwise.} \end{cases}$$

It should be clear to you that since this function is constant along the two axes,

$$\frac{\partial f}{\partial x_1}(0,0) = 0 \quad \text{and} \quad \frac{\partial f}{\partial x_2}(0,0) = 0$$

On the other hand, $f(x_1, x_2)$ is not continuous at the origin! Thus, a function may have both partial derivatives (and, in fact, any directional derivative - see the next slide) but still not be continuous. This suggests that for a function of two variables, just requiring that both partial derivatives exist is not a good or useful definition of "differentiability".

Last class we studied the following function from Exercise 5.5:

$$\frac{x^2y^2}{x^2y^2+(x-y)^2} \quad \text{for} \quad (x,y)\neq(0,0).$$

Let us further set $f(0,0)=0$. You can check that every directional derivative exists and is equal to 0, except along $y=x$ when the directional derivative is not defined - I had earlier said it was 1. However, we have already seen that the function is not continuous at the origin since we have shown that $\lim_{(x,y)\to 0} f(x,y)$ does not exist. For an example with directional derivatives in all directions see Exercise 5.3(i).

Conclusion: All directional derivatives may exist at a point even if the function is discontinuous.

Let us go back and examine the notion of differentiability for a function of $f(x)$ of one variable. Suppose f is differentiable at the point x_0, What is the equation of the tangent line through $(x_0, f(x_0))$?

$$y = f(x_0) + f'(x_0)(x - x_0)$$

as the equation for the tangent line. If we consider the difference $f(x) - f(x_0) - f'(x_0)(x - x_0)$ we get the distance of a point on the tangent line from the curve $y = f(x)$. Writing $h = (x - x_0)$, we see that the difference can be rewritten

$$f(x_0 + h) - f(x_0) - f'(x_0)h$$

The tangent line is close to the function f - how close?- so close that even after dividing by h the distance goes to 0. A few lectures ago we wrote this as

$$|f(x_0 + h) - f(x_0) - f'(x_0)h| = o(h)|h|$$

where $o(h)$ is a function that goes to 0 as h goes to 0.

The preceding idea generalises to two (or more) dimensions. 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 curve*. 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 "$o(h)$" 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

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

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

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

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

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

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

for some function $o(h, k)$ 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 $\mathbb{R}^2$.

The matrix $Df(x_0, y_0)$ is called the Derivative matrix 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 $\mathbf{i}$ and $\mathbf{j}$ 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 $\mathcal{C}^1$ function is differentiable.

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.

The Chain Rule

We now study the situation where we have composition of functions. We assume that $x, y : I \to \mathbb{R}$ are differentiable functions from some interval (open or closed) to $\mathbb{R}$. Thus the pair $(x(t), y(t))$ defines a function from I to $\mathbb{R}^2$. Suppose we have a function $f : \mathbb{R}^2 \to \mathbb{R}$ which is differentiable. We would like to study the derivative of the composite function $z(t) = f(x(t), y(t))$ from I to $\mathbb{R}$.

Theorem 27: With notation as above

$$\frac{dz}{dt} = \frac{\partial z}{\partial x}\frac{dx}{dt} + \frac{\partial z}{\partial y}\frac{dy}{dt}.$$

For a function $w = f(x, y, z)$ in three variables the chain rule takes the form

$$\frac{dw}{dt} = \frac{\partial w}{\partial x}\frac{dx}{dt} + \frac{\partial w}{\partial y}\frac{dy}{dt} + \frac{\partial w}{\partial z}\frac{dz}{dt}$$

An application to tangents of curves

Example: Let us verify this rule in a simple case. Let $z = xy$, $x = t^3$ and $y = t^2$.
Then $z = t^5$ so $z'(t) = 5t^4$. On the other hand, using the chain rule we get

$$z'(t) = y \cdot 3t^2 + x \cdot 2t = 3t^4 + 2t^4 = 5t^4.$$

Example: A mapping $c : I \to \mathbb{R}^n$ of an interval I to $\mathbb{R}$ is called a path or curve in $\mathbb{R}^n$, $(n = 2, 3)$. The function $c(t)$ will be given by a tuple of functions form
Let us consider a curve $c(t)$ in $\mathbb{R}^3$. Each point on the curve will be given by a triple of coordinates which will depend on t. That is, the curve can be described by a triple of functions $(g(t), h(t), k(t))$. We can write

$$c(t) = g(t)\mathbf{i} + h(t)\mathbf{j} + k(t)\mathbf{k}, \quad \text{and} \quad c'(t_0) = g'(t_0)\mathbf{i} + h'(t_0)\mathbf{j} + k'(t_0)\mathbf{k},$$

represents its tangent or velocity vector at the point $c(t_0)$.

So far our example has nothing to do with the chain rule. Suppose $z = f(x, y)$ is a surface, and our curve given by $c(t) = (g(t), h(t), f(g(t), h(t))$ lies on the $z = f(x, y)$. Let us compute the tangent vector to the curve at $c(t_0)$. It is given by

$$c'(t_0) = g'(t_0)\mathbf{i} + h'(t_0)\mathbf{j} + k'(t_0)\mathbf{k},$$

where $k(t) = (f(g(t), h(t))$. Using the chain rule we see that

$$k'(t_0) = \frac{\partial f}{\partial x} g'(t_0) + \frac{\partial f}{\partial y} h'(t_0).$$

We can further show that this tangent vector lies on the tangent plane to the surface $z = f(x, y)$. Indeed we have already seen that the tangent plane has the equation

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

A normal vector to this plane is given by

$$\left(-\frac{\partial f}{\partial x}(x_0, y_0), -\frac{\partial f}{\partial y}(x_0, y_0), 1\right).$$

Thus, to verify that the tangent vector lies on the plane, we need only check that its dot product with normal vector is 0. But this is now clear.

Just to give a concrete example of what we are talking about, take a curve $(g(t), h(t))$ in the unit disc $x^2 + y^2 \leq 1$ in the xy plane. Then

$$\left(g(t), h(t), \sqrt{1 - g(t)^2 - h(t)^2}\right)$$

lies on the upper hemisphere

$$z = \sqrt{1 - x^2 - y^2}.$$

For concreteness, we can take $I = \left[0, \frac{1}{\sqrt{2}}\right]$, $g(t) = t$ and $h(t) = t^2$