MA 105 : Calculus

D1 - Lecture 17

Sandip Singh

Department of Mathematics

Autumn 2023, IIT Bombay, Mumbai

Rules for the total derivative

Rule 1: Just like in the one variable case, if f and g are differentiable

$$D(f+g)(x) = Df(x) + Dg(x)$$

and

$$D(cf)(x) = cDf(x),\ \forall c \in \mathbb{R}.$$

Rule 2: Somewhat harder, but only because the notation gets more cumbersome, is the Chain rule: Let $U \subseteq \mathbb{R}^m$, $V \subseteq \mathbb{R}^n$, $f : U \to \mathbb{R}^n$ be a function which differentiable at $x_0 \in U$ and $f(U) \subseteq V$. Let $g : V \to \mathbb{R}^\ell$ be a function which is differentiable at $f(x_0)$. Then $g \circ f : U \to \mathbb{R}^\ell$ is differentiable at x_0 and

$$D(g \circ f)(x_0)_{\ell \times m} = Dg(f(x_0))_{\ell \times n} \circ Df(x_0)_{n \times m},$$

where $\circ$ on the right hand side denotes the matrix multiplication.

The Chain Rule: Applications

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 2: With notation as above

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

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

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

Remark: We omit the proof of the above theorem here but if you are interested in seeing that I can upload it on Moodle later.

An application to tangents of curves

A simple example to verify the chain rule: Let $z = f(x, y) = xy$, $x(t) = t^3$ and $y(t) = t^2$. Then $z(t) = t^5$, so $z'(t) = 5t^4$.

On the other hand, using the chain rule we get

$$z'(t) = \frac{\partial f}{\partial x}\frac{dx}{dt} + \frac{\partial f}{\partial y}\frac{dy}{dt} = y \cdot 3t^2 + x \cdot 2t = 3t^4 + 2t^4 = 5t^4.$$

Example: A continuous mapping $c : I \to \mathbb{R}^n$ on an interval $I \subseteq \mathbb{R}$ 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

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

exists and is nonzero, it represents the tangent vector to the curve $c(t)$ at the point $c(t_0)$.

For an example, consider the curve $c(t) = (t, \sqrt{1-t^2})$ in $\mathbb{R}^2$ and defined on the interval $[-1, 1]$. Observe that the curve $c(t)$ represents the **upper unit semicircle** centered at the origin.

You can verify easily that whenever $c'(t_0)$ exists and is nonzero, the tangent line to the circle $c(t)$ at the point $c(t_0)$ is the line that passes through the point $c(t_0)$ and is parallel to the tangent vector $c'(t_0)$. $\square$

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 surface $z = f(x, y)$.

Let us compute the tangent vector to the curve at $c(t_0) = (x_0, y_0, z_0)$. 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

$$k'(t_0) = \frac{\partial f}{\partial x}(x_0, y_0)g'(t_0) + \frac{\partial f}{\partial y}(x_0, y_0)h'(t_0).$$

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 to the surface $z = f(x, y)$ at the point (x_0, y_0, z_0) on the surface 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

$$c'(t_0) = \left(g'(t_0), h'(t_0), \frac{\partial f}{\partial x}(x_0, y_0)g'(t_0) + \frac{\partial f}{\partial y}(x_0, y_0)h'(t_0)\right)$$

at the point $c(t_0) = (x_0, y_0, z_0)$ on the curve $c(t)$ lies on the plane, we need only check that its dot product with the normal vector is 0. But this is now clear. $\square$

The Gradient of scalar fields

When viewed as a row vector rather than as a matrix, the derivative matrix of $f : U \to \mathbb{R}$ at a point $(a_1, a_2, \dots, a_m) \in U \subseteq \mathbb{R}^m$ is called the gradient of f at $(a_1, a_2, \dots, a_m)$ and is denoted as $\nabla f(a_1, a_2, \dots, a_m)$. Thus

$$\nabla f(a_1, a_2, \dots, a_m) = \left(\frac{\partial f}{\partial x_1}(a_1, a_2, \dots, a_m), \dots, \frac{\partial f}{\partial x_m}(a_1, a_2, \dots, a_m) \right).$$

For the case $m = 2$, in terms of the coordinate vectors $\mathbf{i}$ and $\mathbf{j}$ the gradient can be written as

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

As we will see in the next slide, the gradient is related to the directional derivative in the direction v:

$$\nabla_v f = \nabla f \cdot v.$$

Another application of the chain rule: **Directional derivatives**

Let $U \subset \mathbb{R}^3$ and let $f : U \to \mathbb{R}$ be differentiable. We want to relate the directional derivative to the gradient.

We consider the (differentiable) curve $c(t) = (x_0, y_0, z_0) + tv$, where $v = (v_1, v_2, v_3)$ is a **unit vector**. We can rewrite $c(t)$ as $c(t) = (x_0 + tv_1, y_0 + tv_2, z_0 + tv_3)$. We apply the chain rule to compute the derivative of the function $f(c(t))$ (observe that $\frac{d}{dt} f(c(t))$ at $t = 0$ is same as the directional derivative of f at (x_0, y_0, z_0) in the direction of the vector $v = (v_1, v_2, v_3)$):

$$\frac{d}{dt} f(c(0)) = \frac{\partial f}{\partial x}(x_0, y_0, z_0)v_1 + \frac{\partial f}{\partial y}(x_0, y_0, z_0)v_2 + \frac{\partial f}{\partial z}(x_0, y_0, z_0)v_3$$

and this can be rewritten as

$$\nabla_v f(x_0, y_0, z_0) = \frac{d}{dt} f(c(0)) = \nabla f(x_0, y_0, z_0) \cdot v.$$

Of course, the same argument works when $U \subset \mathbb{R}^2$ and f is a function of two variables.

The Chain Rule and Gradients

The preceding argument is a special case of a more general fact.

Let $c(t)$ be a (differentiable) curve in $\mathbb{R}^3$. Then, by writing $c(t) = (x(t), y(t), z(t))$ and then using the chain rule for the derivative of $f(c(t))$ we obtain that

$$\frac{d}{dt} f(c(t)) = \nabla f(c(t)) \cdot c'(t).$$

Going back to the directional derivative, we can ask ourselves the following question. In what direction is f changing fastest at a given point (x_0, y_0, z_0)? In other words, in which direction does the directional derivative attain its largest value?

Using what we have just learnt, we are looking for a unit vector $v = (v_1, v_2, v_3)$ such that

$$\nabla f(x_0, y_0, z_0) \cdot v$$

is as large as possible (we assume that $\nabla f(x_0, y_0, z_0) \neq (0, 0, 0)$).

We rewrite the preceding dot product as

$$\nabla f(x_0, y_0, z_0) \cdot v = \|\nabla f(x_0, y_0, z_0)\| \|v\| \cos\theta$$

where θ is the angle between v and $\nabla f(x_0, y_0, z_0)$. Since v is a unit vector, this gives

$$\nabla f(x_0, y_0, z_0) \cdot v = \|\nabla f(x_0, y_0, z_0)\| \cos\theta.$$

The maximum value on the right hand side is obviously attained when $\theta = 0$, that is, when v points in the direction of ∇f. In other words the function is increasing fastest in the direction v given by ∇f. Thus the unit vector that we seek is

$$v = \frac{\nabla f(x_0, y_0, z_0)}{\|\nabla f(x_0, y_0, z_0)\|}.$$

Surfaces defined implicitly

So far we have only been considering surfaces of the form $z = f(x, y)$, where f was a function on a subset of $\mathbb{R}^2$. We now consider a more general type of surface S defined implicitly:

$$S = \{(x, y, z) \,|\, f(x, y, z) = b\}$$

where b is a constant. Most surfaces we have come across are usually described in this form: for instance, the sphere which is given by $x^2 + y^2 + z^2 = r^2$ or the right circular cone $x^2 + y^2 - z^2 = 0$. Let us try to understand what a tangent plane is more precisely.

If S is a surface, a tangent plane to S at a point $s_0 \in S$ (if it exists) is a plane that contains the tangent lines at s_0 to all curves passing through s_0 and lying on S.

If $c(t)$ is a curve on the surface S given by $f(x,y,z)=b$, we see that $f(c(t))=b$, and hence

$$\frac{d}{dt}f(c(t))=0.$$

On the other hand, by the chain rule,

$$0=\frac{d}{dt}f(c(t))=\nabla f(c(t))\cdot c'(t).$$

Thus, if $s_0=c(t_0)=(x_0,y_0,z_0)$ is a point on the surface, we see that

$$\nabla f(c(t_0))\cdot c'(t_0)=0,$$

for every curve $c(t)$ on the surface S passing through $s_0=(x_0,y_0,z_0)$.

Hence, if $\nabla f(x_0,y_0,z_0)\neq 0$, then $\nabla f(x_0,y_0,z_0)$ is perpendicular to the tangent plane of S at (x_0,y_0,z_0). How to determine a vector perpendicular to the tangent plane at the point (x_0,y_0,z_0) on the surface S given by $z=f(x,y)$? It is determined by $\nabla(z-f(x,y))$ at the point (x_0,y_0,z_0).

The equation of the tangent plane

Since we know that the gradient of f is normal to the level surface $S = \{(x, y, z) \in \mathbb{R}^3 \mid f(x, y, z) = c\}$ (provided the gradient is nonzero), it allows us to write down the equation of the tangent plane of S at the point $s_0 = (x_0, y_0, z_0)$. The equation of this plane is (here f_x, f_y, f_z denote the respective partial derivatives of f)

$$f_x(x_0, y_0, z_0)(x-x_0)+f_y(x_0, y_0, z_0)(y-y_0)+f_z(x_0, y_0, z_0)(z-z_0) = 0.$$

For the curve $f(x, y) = c$, by considering y as a function of x (implicitly) we obtain (by differentiating $f(x, y)$ with respect to x and using the chain rule) that

$$f_x(x_0, y_0) + f_y(x_0, y_0)\frac{dy}{dx} = 0 \Rightarrow \frac{dy}{dx} = -\frac{f_x(x_0, y_0)}{f_y(x_0, y_0)}$$

and hence the equation of the tangent line to the curve $f(x, y) = c$ passing through (x_0, y_0) is

$$y - y_0 = -\frac{f_x(x_0, y_0)}{f_y(x_0, y_0)}(x - x_0),$$

that is, $f_x(x_0, y_0)(x - x_0) + f_y(x_0, y_0)(y - y_0) = 0.$

Gravitational force as gradient of the potential energy

Let $\mathbf{r}$ denote the position vector

$$x\mathbf{i} + y\mathbf{j} + z\mathbf{k},$$

of a point $P = (x, y, z)$ in $\mathbb{R}^3$. Instead of writing $\|\mathbf{r}\|$, it is customary to write r for $\sqrt{x^2 + y^2 + z^2}$. This notation is very useful.

For instance, the Newton's Law of Gravitation can be expressed as

$$\mathbf{F} = -\frac{GMm}{r^3} \cdot \mathbf{r}$$

where the mass M is assumed to be at the origin, $\mathbf{r}$ denotes the position vector of the mass m, G is a constant and $\mathbf{F}$ denotes the gravitational force between the two (point) masses.

A simple computation shows that

$$\nabla\left(\frac{1}{r}\right) = -\frac{\mathbf{r}}{r^3}.$$

Thus the gravitational force at any point can be expressed as the gradient of a function called the potential function/energy.

Keeping our previous discussion in mind, we know that if

$$V = -\frac{GMm}{r}, \quad \mathbf{F} = -\frac{GMm}{r^3} \cdot \mathbf{r} = -\nabla V.$$

What are the level surfaces of V? Clearly, r must be a constant on these level sets, so the level surfaces are spheres.

Since $\mathbf{F} = -\nabla V$, we see that the gravitational force $\mathbf{F}$ is orthogonal to the sphere and points towards the origin. $\square$

Review - problems involving the gradient

Exercise 1: Find the points on the hyperboloid $x^2 - y^2 + 2z^2 = 1$ where the normal line is parallel to the line that joins the points $(3, -1, 0)$ and $(5, 3, 6)$.

Solution: The hyperboloid is an implicitly defined surface. A normal vector at a point (x_0, y_0, z_0) on the hyperboloid is given by the gradient of the function $x^2 - y^2 + 2z^2$ at (x_0, y_0, z_0):

$$\nabla f(x_0, y_0, z_0) = (2x_0, -2y_0, 4z_0).$$

We require this vector to be parallel to the line joining the points $(3, -1, 0)$ and $(5, 3, 6)$. This line lies in the same direction as the vector $(5 - 3, 3 + 1, 6 - 0) = (2, 4, 6)$. Thus we need only solve the equations

$$(2x_0, -2y_0, 4z_0) = \lambda(2, 4, 6),$$

for some $\lambda \in \mathbb{R}$ such that the point (x_0, y_0, z_0) lies on the hyperboloid. By solving the above equations, we find that $x_0 = \lambda$, $y_0 = -2\lambda$ and $z_0 = (3/2)\lambda$. Substituting x_0, y_0, z_0 in the equation of the hyperboloid yields $\lambda = \pm\sqrt{2/3}$. $\square$

Problems involving the gradient, continued

Exercise 2: Find the directions in which the directional derivative of $f(x, y) = x^2 + \sin xy$ at the point $(1, 0)$ has the value 1.

Solution: We compute ∇f first:

$$\nabla f(x, y) = (2x + y \cos xy, x \cos xy),$$

so at $(1, 0)$ we get, $\nabla f(1, 0) = (2, 1)$.

To find the directional derivative in the direction $v = (v_1, v_2)$ (where v is a unit vector), we simply take the dot product with the gradient:

$$\nabla_v f(1, 0) = \nabla f(1, 0) \cdot v = (2, 1) \cdot (v_1, v_2) = 2v_1 + v_2.$$

This will have value "1" when $2v_1 + v_2 = 1$, subject to $v_1^2 + v_2^2 = 1$, which yields $v_1 = 0, v_2 = 1$ or $v_1 = 4/5, v_2 = -3/5$. $\square$

Review of the gradient

Exercise 3: Find $D_u F(2,2,1)$ where D_u denotes the directional derivative of the function $F(x,y,z) = 3x - 5y + 2z$ and u is the unit vector in the direction of the outward normal to the sphere $x^2 + y^2 + z^2 = 9$ at the point $(2,2,1)$.

Solution: The unit outward normal to the sphere $g(x,y,z) = x^2 + y^2 + z^2 = 9$ at $(2,2,1)$ is given by

$$u = \frac{\nabla g(2,2,1)}{\|\nabla g(2,2,1)\|}.$$

We see that $\nabla g(2,2,1) = (4,4,2)$ so the corresponding unit vector u is $\frac{1}{3}(2,2,1)$.

To get the directional derivative we simply take the dot product of $\nabla F(2,2,1) = (3,-5,2)$ with $u = \frac{1}{3}(2,2,1)$:

$$D_u F(2,2,1) = (3,-5,2) \cdot \frac{1}{3}(2,2,1) = -\frac{2}{3}.$$

□