Lecture 4 : Monday April 6th

jacques@ucsd.edu

KEY CONCEPTS : TANGENT HYPERPLANE, GRADIENT, DIRECTIONAL DERIVATIVE, LEVEL CURVE

KNOW HOW TO FIND EQUATION OF TANGENT HYPERPLANE, GRADIENT, DIRECTIONAL DERIVATIVES, DIRECTION OF FASTEST INCREASE.

4.1 Differentiability

Recall that $f : \mathbb{R}^n \to \mathbb{R}$ is differentiable at a point $a \in \mathbb{R}^n$ if all partial derivatives $f_j(a)$ exist and

$$\lim_{x \to a} \frac{f(x) - f(a) - \sum_{j=1}^{n} f_j(a)(x_j - a_j)}{d(x,a)} = 0.$$

It is not enough that all the partial derivatives $f_j(a)$ exist for f to be differentiable at a. Remember here that $x \to a$ means $(x_1, x_2, \dots, x_n) \to (a_1, a_2, \dots, a_n)$. Let's do some examples. In all the examples $p = (x, y)$.

Example 1. The function $f(x,y) = xy$ is differentiable at $a = (0,0)$, since $f_x(a) = 0 = f_y(a)$ as we checked in a previous example, and

$$\lim_{p \to a} \frac{xy}{\sqrt{x^2 + y^2}} = 0.$$

Let's use the ϵ-δ definition of limits to see that this limit is zero. Let $g(x,y)$ be the function in the limit. For every $\epsilon > 0$ we must find a $\delta > 0$ such that $d(p,a) < \delta$ ensures $|g(x,y)| < \epsilon$. We claim that $|xy| \le x^2 + y^2$. To see this, note that it is equivalent to $|x|^2 + |y|^2 - |xy| \ge 0$. But this is true since

$$0 \le (|x| - |y|)^2 = |x|^2 + |y|^2 - 2|xy| \le |x|^2 + |y|^2 - |xy|.$$

Now we use this to show that f is differentiable:

$$\begin{aligned} |g(x,y)| < \epsilon \quad &\leftarrow \quad \left| \frac{xy}{\sqrt{x^2+y^2}} \right| < \epsilon \\ &\leftarrow \quad \frac{x^2+y^2}{\sqrt{x^2+y^2}} < \epsilon \\ &\leftarrow \quad \sqrt{x^2+y^2} < \epsilon \\ &\leftarrow \quad d(p,a) < \epsilon. \end{aligned}$$

So putting $\epsilon = \delta$ in the definition of the limit we get $\lim g(x, y) = 0$ as required for f to be differentiable at $(0, 0)$. There's an easier way to check the limit is zero, using squeezing: note that

$$0 \leq \frac{|xy|}{\sqrt{x^2 + y^2}} \leq \frac{|xy|}{\sqrt{y^2}} = |x|$$

so the function in the limit is between 0 and $|x|$. Since both of these have a limit of zero as $(x, y) \to (0, 0)$, so must $xy/\sqrt{x^2 + y^2}$.

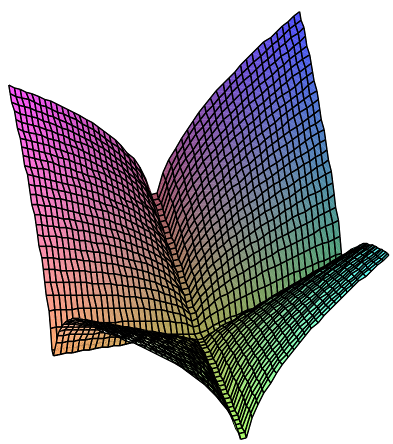
Example 2. This example is almost a follow up to the last one, where we now put $f(x, y) = \sqrt{|xy|}$. This function is not differentiable at $a = (0, 0)$, since $f_x(0, 0) = f_y(0, 0) = 0$ as we saw last lecture, but

$$\lim_{(x,y) \to a} \frac{\sqrt{|xy|}}{\sqrt{x^2 + y^2}}$$

does not exist. To see that the limit fails, along the line $y = mx$ for $m \neq 0$ the limit is

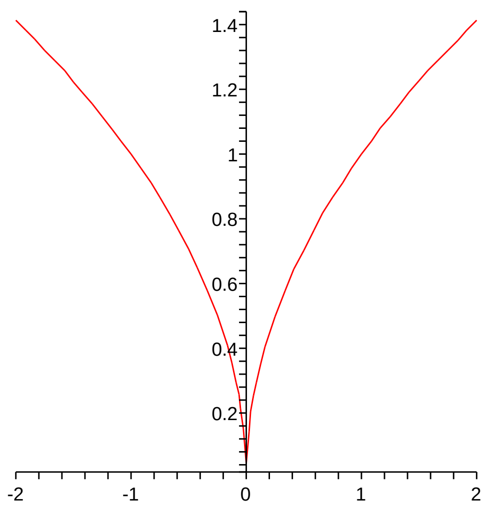
$$\lim_{x \to 0} \frac{\sqrt{m}|x|}{\sqrt{x^2 + m^2 x^2}} = \frac{1}{\sqrt{m}}$$

and this depends on m. So f is not differentiable at $(0, 0)$.



$f(x, y) = \sqrt{|xy|}$

Not differentiable at $(0, 0)$.



$f(x, y) = \sqrt{|x|}$

Not differentiable at 0.

4.2 Tangent plane

In the case of a function of two variables, we represent a function $f(x, y)$ graphically as a surface $z = f(x, y)$. Provided the surface is smooth enough, we can speak of the plane tangent to $z = f(x, y)$ at a point $a \in \mathbb{R}^2$. For example, the surface $z = x^2 + y^2$ represents a paraboloid. At the point $(0, 0)$, the tangent plane is horizontal and clearly given by the equation $z = 0$ – i.e. the xy-plane. For our purposes, differentiability of an n-variable function $f : \mathbb{R}^n \to \mathbb{R}$ at a point is enough to guarantee the existence of a tangent hyperplane at that point. We generally imagine that $f : \mathbb{R}^n \to \mathbb{R}$ is a surface in $n+1$ dimensions defined by the equation $z = f(x_1, x_2, \ldots, x_n)$. For $n = 1$, the tangent hyperplane is just a tangent line whose slope is given by derivatives. Here is the definition of the tangent hyperplane:

Definition of tangent hyperplane.

Let $f : \mathbb{R}^n \to \mathbb{R}$ be a function which is differentiable at $a \in \mathbb{R}^n$. Then the equation of the tangent hyperplane to f at a is given by

$$z = f(a) + \sum_{j=1}^{n} f_j(a)(x_j - a_j).$$

Example 1. Let's check this makes sense for the paraboloid $z = x^2 + y^2$ at $(x, y) = (0, 0)$. First note that if $f(x, y) = x^2 + y^2$ then $f_x(0, 0) = 0 = f_y(0, 0)$. Next, to check differentiability, we see $d((x, y), (0, 0)) = \sqrt{x^2 + y^2}$ and

$$\lim_{(x,y)\to(0,0)} \frac{f(x, y) - f(0, 0) - x f_x(0, 0) - y f_y(0, 0)}{\sqrt{x^2 + y^2}} = \lim_{(x,y)\to(0,0)} \frac{x^2 + y^2}{\sqrt{x^2 + y^2}} = 0.$$

Therefore f is differentiable at $(0, 0)$. The equation of the tangent plane is

$$z = f(0, 0) + x f_x(0, 0) + y f_y(0, 0) = 0$$

as we expected.

Example 2. As another example, consider $f(x, y) = x + y + xy$. Then $z = x + y + xy$ has a tangent plane at $(x, y) = (0, 0)$ since f is differentiable at $(0, 0)$:

$$\lim_{(x,y)\to(0,0)} \frac{f(x, y) - f(0, 0) - x f_x(0, 0) - y f_y(0, 0)}{\sqrt{x^2 + y^2}} = \lim_{(x,y)\to(0,0)} \frac{x + y + xy - x - y}{\sqrt{x^2 + y^2}} = 0$$

where the last step is left as an exercise. So the tangent plane is

$$z = f(0, 0) + x f_x(0, 0) + y f_y(0, 0) = x + y.$$

4.2 Gradient and directional derivatives

The derivatives $f_j(a)$ of a function $f : \mathbb{R}^n \to \mathbb{R}$ denote the slope of the surface $z = f(x_1, x_2, \ldots, x_n)$ in the direction of the x_j-axis at the point a. The following definition holds all this information in the form of a single vector.

Definition of gradient.

The gradient of a function $f : \mathbb{R}^n \to \mathbb{R}$ is the vector of partial derivatives

$$\nabla f = (f_1, f_2, \ldots, f_n).$$

Note that ∇f is sometimes written grad(f). So the vector $\nabla f(a)$ gives the slope or rate of change of f at a in the direction of the co-ordinate axes – in the direction of the jth co-ordinate for each j. This also gives us a convenient way to write the sum in the definition of differentiability, and in the definition of the tangent plane:

$$\sum_{j=1}^{n} f_j(a)(x_j - a_j) = \nabla f(a) \cdot (x - a)$$

where the $\cdot$ is the dot or scalar product of the vectors $\nabla f(a)$ and $x - a$. More generally, we can talk about the rate of change of f in the direction of any vector u.

Definition of directional derivative.

Given a unit vector (we sometimes call this a direction) u, the rate of change of f at a in the direction of u is defined by

$$\nabla f(a) \cdot u = \frac{\partial f}{\partial x_1}(a)u_1 + \frac{\partial f}{\partial x_2}(a)u_2 + \cdots + \frac{\partial f}{\partial x_n}(a)u_n = \sum_{j=1}^{n} \frac{\partial f}{\partial x_j}(a)u_j.$$

Here $\nabla f(a) \cdot u$ is the dot product of $\nabla f(a)$ and u. This is otherwise known as the directional derivative of f at a in the direction of u and gives the slope of f in all directions.

The gradient of a function is extremely important when considering optimization problems for functions of several variables. The basic geometric significance of the gradient is as follows:

Proposition.

Let $f : \mathbb{R}^n \to \mathbb{R}$ and suppose ∇f exists. Then ∇f is the direction along which f is increasing the fastest.

This is true because the rate of change of f at a point a in the direction of a vector u is given by the directional derivative $\nabla f(a) \cdot u = \|\nabla f(a)\| \|u\| \cos\phi$ where ϕ is the angle between $\nabla f(a)$ and u. If u is a unit vector, then this is a maximum when $\phi = 0$. This means u and $\nabla f(a)$ are parallel vectors, as required.

Example. The surface $z = x + y + xy = f(x,y)$ increases fastest in the direction $\nabla f(0,0) = (1,1)$ at $(x,y) = (0,0)$. In general, $\nabla f = (1+y, 1+x)$. Now to work out the directional derivative of f at $(0,0)$ in the direction of the unit vector $(1/\sqrt{2}, 1/\sqrt{2})$, we compute

$$\nabla f(0,0) \cdot (1/\sqrt{2}, 1/\sqrt{2}) = \sqrt{2}.$$

This is the slope of f from $(0,0)$ in the direction of $(1/\sqrt{2}, 1/\sqrt{2})$.

4.3 Level curves

The gradient ∇f for $f : \mathbb{R}^n \to \mathbb{R}$ can be considered as a function $\nabla f : \mathbb{R}^n \to \mathbb{R}^n$ since to each $a \in \mathbb{R}^n$ it assigns the vector $\nabla f(a) \in \mathbb{R}^n$. We may represent ∇f as a vector field, where from each point a we draw the vector $\nabla f(a)$. This vector field in one picture represents the rate of change of f at all points in all directions. If f represents a surface and we pour water on a point on that surface under the influence of gravity, then the water will flow along the lines of steepest descent from that point. This path is represented by the direction of steepest gradient at each point. If we were to draw contours on the surface it is intuitively clear that the gradient should be perpendicular to these contours. This leads us to the definition of level curves. For a function $f : \mathbb{R}^2 \to \mathbb{R}$, the intersection of f with a given plane is a set of curves. For planes parallel to the xy-plane, defined by $z = c$, the level curves of $z = f(x,y)$ are exactly the curves $f(x,y) = c$ for various values of c. These are exactly the same as contour lines on a topographic map. What we said above is that the gradient is perpendicular to the level curves.

Proposition.

> Let $f : \mathbb{R}^2 \to \mathbb{R}$ and suppose ∇f exists. Then $\nabla f(a)$ is perpendicular to the level curve passing through a.

Example. Let $f(x,y) = \sqrt{1 - x^2 - y^2}$ for $x^2 + y^2 \leq 1$. Then $z = f(x,y)$ denotes a hemisphere and is differentiable at all points in the disc $x^2 + y^2 < 1$. The level curves of this surface are circles $f(x,y) = c$ for $c < 1$: in other words they are

$$\sqrt{1 - x^2 - y^2} = c \to x^2 + y^2 = 1 - c^2.$$

The radius of this circle is $\sqrt{1 - c^2}$. For example at height $c = 1/\sqrt{2}$ the level curve is $x^2 + y^2 = 1/2$. Now the gradient is given by

$$\nabla f = (-x/\sqrt{1 - x^2 - y^2}, -y/\sqrt{1 - x^2 - y^2}).$$

At the point $(x, y) = (1/\sqrt{2}, 0)$ on the level curve of height $1/2$, the tangent vector is clearly parallel to the y-axis – the direction of the contour at that point is $(0, 1)$. Since $\nabla(1/\sqrt{2}, 0) = (-1, 0)$, we see that

$$\nabla f(1/\sqrt{2}, 0) \cdot (0, 1) = 0$$

and so the gradient is indeed perpendicular to the level curve at the point $(1/\sqrt{2}, 0)$.