Infinitesimal Calculus 3

Lecture 13, Sunday December 4, 2022
Ari Feiglin

Definition 13.0.1:

If $f\colon \mathbb{R}^2 \longrightarrow \mathbb{R}$ is differentiable at (x_0, y_0) then the **tangent plane** to f at (x_0, y_0) is the set of points (x, y, z) such that

$$(z - z_0) = A(x - x_0) + B(y - y_0)$$

This is equivalent to

$$A(x - x_0) + B(y - y_0) - (z - z_0) = 0$$

And such a normal vector to this plane is $(A, B, -1)$ (in general the normal to $a(x - x_0) + b(y - y_0) + c(z - z_0) = 0$ is (a, b, c)). And such by our previous theorem, the normal to the tangent plane is

$$\vec{n} = (\partial_x f(x_0, y_0), \partial f_y(x_0, y_0), -1)$$

Notice that if we define $v_1 = (1, 0, \partial_x f(x_0, y_0)$ and $v_2 = (0, 1, \partial_y f(x_0, y_0))$ which are tangent to the x and y partial functions of f ($(x, f'(x))$ is tangent to $f(x)$), then we see that $v_1 \cdot n = v_2 \cdot n = 0$ which should be intuitive. So the tangent plane is spanned by $\big\{(1, 0, \partial_x f), (0, 1, \partial_y f)\big\}$.

Definition 13.0.2:

If f is a real valued function in $\mathbb{R}^n$ and $u \in \mathbb{R}^n$ is a unit vector, the **directional derivative** in the direction of u is defined by:

$$D_u f(\vec{x}) = \lim_{t\to 0} \frac{f(\vec{x} + tu) - f(\vec{x})}{t}$$

Notice that if $u = e_k = (0, \dots, 1, \dots, 0)$ then $D_u f = \partial_k f$ (the kth partial derivative).

Proposition 13.0.3:

If $f(x, y)$ is defined in a neighborhood of (x_0, y_0) and differentiable there, suppose $u = (a, b)$ is a unit vector then $D_u f(x_0, y_0)$ exists and is equal to

$$D_u f(x_0, y_0) = a\partial_x f(x_0, y_0) + b\partial_y f(x_0, y_0)$$

Proof:

We know that by definition of differentiability

$$f(x, y) = f(x_0, y_0) + A\Delta x + B\Delta y + \varepsilon(\Delta x, \Delta y)$$

Where $A = \partial_x f(x_0, y_0)$ and $B = \partial_y f(x_0, y_0)$. Then

$$D_u f(x_0, y_0) = \lim_{t\to 0} \frac{f(x_0 + ta, y_0 + tb) - f(x_0, y_0)}{t} = \lim_{t\to 0} \frac{Ata + Btb + \varepsilon(\Delta x, \Delta y)}{t} = Aa + Bb + \lim_{t\to 0} \frac{\varepsilon(\Delta x, \Delta y)}{t} = Aa + Bb$$

The last limit being equal to 0 is true since $\Delta x^2 + \Delta y^2 = t^2(a^2 + b^2)$ and thus it is a constant times the limit of $\frac{\varepsilon(\Delta\vec{x})}{\|\vec{x}\|}$ which is 0. And recalling what A and B are equal to, this proves our proposition.

■

Definition 13.0.4:

Suppose $f\colon \mathbb{R}^n \longrightarrow \mathbb{R}$ is defined in a neighborhood of $\vec{x}$ and has partial derivatives at $\vec{x}$. Then we define the **gradient** of f at $\vec{x}$ to be:

$$\nabla f(\vec{x}) = \begin{pmatrix} \partial_{e_1} f(\vec{x}) \\ \cdots \\ \partial_{e_n} f(\vec{x}) \end{pmatrix}$$

Thus by our above proposition, if f is differentiable then

$$D_u f = u \cdot \nabla f$$

Notice that if θ is the angle between u and the gradient of f, then

$$D_u f = \|u\| \cdot \|\nabla f\| \cdot \cos\theta = \|\nabla f\| \cdot \cos\theta$$

Thus the directional derivative is maximal when u is parallel with ∇f ($\theta = 0$), and minimal when u is pointed in the opposite direction of ∇f ($\theta = \pi$). Thus the gradient of a function informs us of its maximal rate of change.

Definition 13.0.5:

If $E \subseteq \mathbb{R}^n$ and $f\colon E \longrightarrow \mathbb{R}$, if f is defined in a neighborhood of $\vec{v}_0 \in E$ then we define the kth partial derivative of f at $\vec{v}_0$ to be:

$$\partial_{x_k} f(\vec{v}_0) = \lim_{\Delta x \to 0} \frac{f(\vec{v}_0 + \Delta x \cdot e_k) - f(\vec{v}_0)}{\Delta x}$$

And f is differentiable at $(x_1, \ldots, x_n)$ if there exist numbers $A_1, \ldots, A_n$ and an ε function such that

$$f(x_1 + \Delta x_1, \ldots, x_n + \Delta x_n) = f(x_1, \ldots, x_n) + \sum_{k=1}^{n} A_k \Delta x_k + \varepsilon(\Delta x_1, \ldots, \Delta x_n)$$

Alternatively if we define $\vec{x} = (x_1, \ldots, x_n)$ and $h = (\Delta x_1, \ldots, \Delta x_n)$, then f is differentiable at $\vec{x}$ if and only if there exists a linear transformation $L\colon \mathbb{R}^n \longrightarrow \mathbb{R}$, and an ε function such that

$$f(\vec{x} + h) = f(x) + L(h) + \varepsilon(h)$$

This is true since the linear transformation L corresponds to mapping e_k to A_k.