LESSON 11

Directional Derivatives and the Gradient Vector

Contents

In this lesson we consider a surface defined by a function in two variables, $z = f(x, y)$ and want to explore the rate of change of z as we move in a direction other than the x-direction or y direction. We want to know how to find this rate of change (this **directional derivative**), and how to determine the direction of maximum rate of change. In this investigation we will come across what is called the Gradient Vector, which is actually a vector function in two variables. We will define this formally and come to understand its significance.

11.1 Directional Derivatives

Suppose we have a function $z = f(x, y)$ and we want to find the slope of the surface in the direction $\mathbf{u} = \langle a, b\rangle$, where $\mathbf{u}$ is a unit vector ($|\mathbf{u}| = 1$ and need *not* be pointing in the x or y directions).

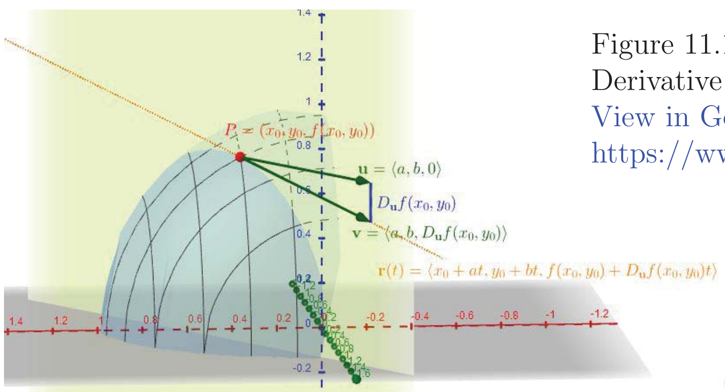

Figure 11.1.1: The Directional Derivative
View in Geogebra:
https://www.geogebra.org/3d/hmm3tynt

We are going to define the notation $D_{\mathbf{u}}f(x_0, y_0)$ to be the **directional derivative** of f in the direction of $\mathbf{u}$. That is, this is the rate of change of z as we move in the $\mathbf{u}$ direction. To come up with a formula for $D_{\mathbf{u}}f(x_0, y_0)$ we note that if we use h as a scalar for $\mathbf{u}$, then as $h \to 0$, our direction vector shrinks to a length of 0. Moreover, if we move away from our point $P = (x_0, y_0, f(x_0, y_0))$ in the $\mathbf{u}$ direction scaled by h we get $\Delta x = ha$ and $\Delta y = hb$. This idea can be seen from the top view of the diagram in the figure 11.1.2 below.

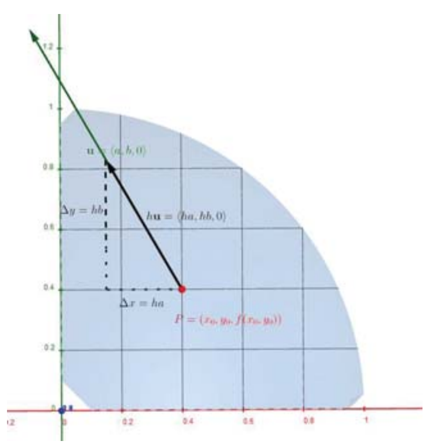

Figure 11.1.2: The Directional Derivative
View in Geogebra:
https://www.geogebra.org/3d/hmm3tynt

So our change in z is $\Delta z = f(x_0 + \Delta x, y_0 + \Delta y) - f(x_0, y_0) = f(x_0 + ha, y_0 + hb) - f(x_0, y_0)$. We can now define our directional derivative to be

Definition 11.1.1

The **directional derivative** of f at (x_0, y_0) in the direction of $\mathbf{u} = \langle a, b \rangle$ is

$$D_{\mathbf{u}} f(x_0, y_0) = \lim_{h \to 0} \frac{f(x_0 + ha, y_0 + hb) - f(x_0, y_0)}{h}$$

if this limit exists.

While this is a working definition, it is not easy to use for quick computations. Instead we use the following theorem:

Theorem 11.1.1

If f is a differentiable function in x and y, then f has a directional derivative in the direction of any unit vector $\mathbf{u} = \langle a, b \rangle$ and

$$D_{\mathbf{u}} f(x, y) = f_x(x, y)a + f_y(x, y)b.$$

Proof : Let $g(h) = f(x(h), y(h))$ where $x = x_0 + ha$ and $y = y_0 + hb$. Then, by the chain rule, we get

$$g'(h) = \frac{\partial f}{\partial x}\frac{dx}{dh} + \frac{\partial f}{\partial y}\frac{dy}{dh} = f_x(x, y)a + f_y(x, y)b.$$

Note, from the definitions of $x(h)$ and $y(h)$, that if $h = 0$ then $x = x_0$ and $y = y_0$. Therefore, by the above equation, we have $g'(0) = f_x(x_0, y_0)a + f_y(x_0, y_0)b$.

On the other hand, by the definition of single-variable derivatives, we have

$$g'(0) = \lim_{h \to 0} \frac{g(h) - g(0)}{h} = \lim_{h \to 0} \frac{f(x_0 + ha, y_0 + hb) - f(x_0, y_0)}{h} = D_{\mathbf{u}} f(x_0, y_0).$$

Setting these two computations for $g'(0)$ equal we get

$$D_{\mathbf{u}} f(x_0, y_0) = f_x(x_0, y_0)a + f_y(x_0, y_0)b$$

and generalizing this for any x_0 and y_0 we get our desired result.

■

Example 11.1.1 Given $f(x, y) = \sqrt{1 - x^2 - y^2}$, find the slope of the surface at the point $\left(\frac{2}{5}, \frac{2}{5}\right)$ in the direction of the unit vector $\mathbf{u}$ which makes an angle of $\frac{2\pi}{3}$ with the positive x-axis. Note this is exactly the example shown in figures 11.1.1 and 11.1.2.

Exercise 11.1.1 Find $D_{\mathbf{u}}f(x, y)$ if $f(x, y) = x^3 - 3xy^2$ and $\mathbf{u}$ is in the direction of the plane $x - y = 0$. Note this gives two possible directions. What is $D_{\mathbf{u}}f(-2, -2)$? What does this represent?

11.2 The Gradient Vector

Looking back at our formula

$$D_{\mathbf{u}}f(x,y) = f_x(x,y)a + f_y(x,y)b$$

let's note that this can be rewritten as a dot product in the following way

$$D_{\mathbf{u}}f(x,y) = \langle f_x(x,y), f_y(x,y)\rangle \cdot \langle a,b\rangle.$$

The vector on the left of the dot product is our **gradient vector** and is actually a vector function in two variables.

Definition 11.2.1

Given a function $f(x,y)$ which is differentiable in x and y, then the **gradient vector** is defined to be

$$\bigtriangledown f(x,y) = \langle f_x(x,y), f_y(x,y)\rangle.$$

Exercise 11.2.1 Given $f(x,y) = xy^3 + 2x^2y$, find the gradient vector. Then find the directional derivative to f at the point $(2,-1)$ in the direction of $\mathbf{v} = \langle 2,3\rangle$.

11.3 Functions in Three Variables

For a function $w = f(x, y, z)$ in three variables, we can define the directional derivative $D_{\mathbf{u}}f(x, y, z)$ in the direction of $\mathbf{u} = \langle a, b, c\rangle$ similarly to how we did when working in two variables, simply noting that this new directional derivative tells us the rate that we move between the level surfaces of f (the rate that w changes) from a given point in a given direction in three dimensions. Everything we've done simply scales up:

Definition 11.3.1

The **directional derivative** of f at (x_0, y_0, z_0) in the direction of $\mathbf{u} = \langle a, b, c\rangle$ is

$$D_{\mathbf{u}}f(x_0, y_0) = \lim_{h\to 0} \frac{f(x_0 + ha, y_0 + hb, z_0 + hc) - f(x_0, y_0, z_0)}{h}$$

if this limit exists.

Definition 11.3.2

Given a function $f(x, y, z)$ which is differentiable in x, y, and z, then the **gradient vector** is defined to be

$$\bigtriangledown f(x, y, z) = \langle f_x(x, y, z), f_y(x, y, z), f_z(x, y, z)\rangle.$$

Theorem 11.3.1

If f is a differentiable function in x, y and z, then f has a directional derivative in the direction of any unit vector $\mathbf{u} = \langle a, b, c\rangle$ and

$$\begin{aligned} D_{\mathbf{u}}f(x, y, z) &= f_x(x, y, z)a + f_y(x, y, z)b + f_z(x, y, z)c \\ &= \bigtriangledown f(x, y, z) \cdot \mathbf{u} \end{aligned}$$

Example 11.3.1 If $f(x, y, z) = x \sin(yz)$, find the gradient of f and use it to find the directional derivative at $(1, 3, 0)$ in the direction of $\langle 1, 2, -1\rangle$.

11.4 Maximizing the Directional Derivative

Suppose we want to know the direction in which our outputs are increasing most rapidly. In two dimensions this would be the direction in which our surface is most steeply uphill and in three dimensions this would be the direction in which we are moving between level surfaces most quickly. It turns out that the answer is quite simple:

Theorem 11.4.1

Given a differentiable function $z = f(x, y)$, and a point (x_0, y_0), then $\bigtriangledown f(x_0, y_0)$ *is* the direction in which z is increasing most rapidly from (x_0, y_0) and $|\bigtriangledown f(x_0, y_0)|$ is the maximum rate of change of z from (x_0, y_0).
Similarly, given a differentiable function $w = f(x, y, z)$ and a point (x_0, y_0, z_0), then $\bigtriangledown f(x_0, y_0, z_0)$ is the direction in which w is increasing most rapidly from (x_0, y_0, z_0) and $|\bigtriangledown f(x_0, y_0, z_0)|$ is the maximum rate of change of w from (x_0, y_0, z_0).

In other words, if we want to find the direction of steepest ascent given a function in two variables from a certain point, all we need to do is plug that point into the gradient vector. To find the slope in that direction all we need to do is take the magnitude of that direction.

Proof : By the definition of the dot product we have

$$D_{\mathbf{u}} f = \bigtriangledown f \cdot \mathbf{u} = |\bigtriangledown f||\mathbf{u}|\cos(\theta) = |\bigtriangledown f|\cos(\theta)$$

where θ is the angle between $\mathbf{u}$ and $\bigtriangledown f$. Note that $|\mathbf{u}| = 1$, which is how it disapeared in the above equation. Thus, $D_{\mathbf{u}} f$ is maximized when $\cos(\theta) = 1$ which occurs when $\bigtriangledown f$ and $\mathbf{u}$ are pointing in the same direction. This is exactly the case when $D_{\mathbf{u}} f = |\bigtriangledown f|$.

■

Exercise 11.4.1 Given $f(x, y) = 4x^4 + 4xy + 3y^2$, find the direction of maximal increase from $(-1, -1)$ and the value of the slope in that direction.

One last note about the direction of maximum increase is that it is always at right angle to the level curves (or level surfaces) of our function. Given a function $z = f(x, y)$, if we set z equal to some constant, we will get a level curve. Traveling along this level curve means we have no change in z so it becomes intuitive that the direction of greatest increase in z is when we travel orthogonaly to these level curves.

11.5 Tangent Planes to Level Surfaces

Suppose we have a level surface given by $f(x, y, z) = k$ for some real number k. As was noted at the end of the previous page, the gradient vector to a level surface will be orthogonal to the surface.

Proof : Let $C = \mathbf{r}(t) = \langle x(t), y(t), z(t)\rangle$ be a curve on the surface given by $f(x, y, z) = k$. Then we can write $f(x(t), y(t), z(t)) = k$ since the curve is defined to be on the level surface. Thus

$$\begin{aligned} \frac{\partial}{\partial t} f(x(t), y(t), z(t)) &= \frac{\partial}{\partial t}(k) \\ \frac{\partial f}{\partial x}\frac{dx}{dt} + \frac{\partial f}{\partial y}\frac{dy}{dt} + \frac{\partial f}{\partial z}\frac{dz}{dt} &= 0 \qquad \text{by the chain rule} \\ \langle f_x, f_y, f_z\rangle \cdot \langle x'(t), y'(t), z'(t)\rangle &= 0 \\ \triangledown f \cdot \mathbf{r}'(t) &= 0 \end{aligned}$$

Since $\mathbf{r}(t)$ is an arbitrary curve on the surface, then $\mathbf{r}'(t)$ is a vector tangent to the surface in any direction. Thus $\triangledown f$ is orthogonal to any tangent on the surface and hence orthogonal to the surface.

Definition 11.5.1

Given a level surface $f(x, y, z) = k$ and a point P on this surface, the **normal line** to the surface through this point is the line through the point that is orthogonal to the surface.

Example 11.5.1 Find the tangent plane and normal line to the surface $x^2 + \frac{1}{4}y^2 + \frac{1}{9}z^2 = 3$ at the point $(1, -2, -3)$.