2.7 Directional Derivatives

To this point we've only looked at the two partial derivatives $f_x(x, y)$ and $f_y(x, y)$. Recall that these derivatives represent the rate of change of f as we vary x (holding y fixed) and as we vary y (holding x fixed) respectively. We now need to discuss how to find the rate of change of f if we allow both x and y to change simultaneously. The problem here is that there are many ways to allow both x and y to change. For instance, one could be changing faster than the other and then there is also the issue of whether or not each is increasing or decreasing. So, before we get into finding the rate of change we need to get a couple of preliminary ideas taken care of first. The main idea that we need to look at is just how are we going to define the changing of x and/or y.

Let's start off by supposing that we wanted the rate of change of f at a particular point, say (x_0, y_0). Let's also suppose that both x and y are increasing and that, in this case, x is increasing twice as fast as y is increasing. So, as y increases one unit of measure x will increase two units of measure.

To help us see how we're going to define this change let's suppose that a particle is sitting at (x_0, y_0) and the particle will move in the direction given by the changing x and y. Therefore, the particle will move off in a direction of increasing x and y and the x coordinate of the point will increase twice as fast as the y coordinate. Now that we're thinking of this changing x and y as a direction of movement we can get a way of defining the change. We know from Calculus II that vectors can be used to define a direction and so the particle, at this point, can be said to be moving in the direction,

$$\vec{v} = \langle 2, 1 \rangle$$

Since this vector can be used to define how a particle at a point is changing we can also use it to describe how x and/or y is changing at a point. For our example we will say that we want the rate of change of f in the direction of $\vec{v} = \langle 2, 1 \rangle$. In this way we will know that x is increasing twice as fast as y is. There is still a small problem with this however. There are many vectors that point in the same direction. For instance, all of the following vectors point in the same direction as $\vec{v} = \langle 2, 1 \rangle$.

$$\vec{v} = \left\langle \frac{1}{5}, \frac{1}{10} \right\rangle \qquad \vec{v} = \langle 6, 3 \rangle \qquad \vec{v} = \left\langle \frac{2}{\sqrt{5}}, \frac{1}{\sqrt{5}} \right\rangle$$

We need a way to consistently find the rate of change of a function in a given direction. We will do this by insisting that the vector that defines the direction of change be a unit vector. Recall that a unit vector is a vector with length, or magnitude, of 1. This means that for the example that we started off thinking about we would want to use

$$\vec{v} = \left\langle \frac{2}{\sqrt{5}}, \frac{1}{\sqrt{5}} \right\rangle$$

since this is the unit vector that points in the direction of change.

For reference purposes recall that the magnitude or length of the vector $\vec{v} = \langle a, b, c \rangle$ is given by,

$$\|\vec{v}\| = \sqrt{a^2 + b^2 + c^2}$$

For two dimensional vectors we drop the c from the formula.

Sometimes we will give the direction of changing x and y as an angle. For instance, we may say that we want the rate of change of f in the direction of $\theta = \frac{\pi}{3}$. The unit vector that points in this direction is given by,

$$\vec{u} = \langle \cos(\theta), \sin(\theta) \rangle$$

Okay, now that we know how to define the direction of changing x and y its time to start talking about finding the rate of change of f in this direction. Let's start off with the official definition.

Definition

The rate of change of $f(x, y)$ in the direction of the unit vector $\vec{u} = \langle a, b \rangle$ is called the **directional derivative** and is denoted by $D_{\vec{u}} f(x, y)$. The definition of the directional derivative is,

$$D_{\vec{u}} f(x, y) = \lim_{h \to 0} \frac{f(x + ah, y + bh) - f(x, y)}{h}$$

So, the definition of the directional derivative is very similar to the definition of partial derivatives. However, in practice this can be a very difficult limit to compute so we need an easier way of taking directional derivatives. It's actually fairly simple to derive an equivalent formula for taking directional derivatives.

To see how we can do this let's define a new function of a single variable,

$$g(z) = f(x_0 + az, y_0 + bz)$$

where x_0, y_0, a, and b are some fixed numbers. Note that this really is a function of a single variable now since z is the only letter that is not representing a fixed number.

Then by the definition of the derivative for functions of a single variable we have,

$$g'(z) = \lim_{h \to 0} \frac{g(z + h) - g(z)}{h}$$

and the derivative at $z = 0$ is given by,

$$g'(0) = \lim_{h \to 0} \frac{g(h) - g(0)}{h}$$

If we now substitute in for $g(z)$ we get,

$$g'(0) = \lim_{h \to 0} \frac{g(h) - g(0)}{h} = \lim_{h \to 0} \frac{f(x_0 + ah, y_0 + bh) - f(x_0, y_0)}{h} = D_{\vec{u}} f(x_0, y_0)$$

So, it looks like we have the following relationship.

$$g'(0) = D_{\vec{u}} f(x_0, y_0) \tag{13.2}$$

Now, let's look at this from another perspective. Let's rewrite $g(z)$ as follows,

$$g(z) = f(x, y) \quad \text{where } x = x_0 + az \text{ and } y = y_0 + bz$$

We can now use the chain rule from the previous section to compute,

$$g'(z) = \frac{dg}{dz} = \frac{\partial f}{\partial x}\frac{dx}{dz} + \frac{\partial f}{\partial y}\frac{dy}{dz} = f_x(x, y)\,a + f_y(x, y)\,b$$

So, from the chain rule we get the following relationship.

$$g'(z) = f_x(x, y)\,a + f_y(x, y)\,b \tag{13.3}$$

If we now take $z = 0$ we will get that $x = x_0$ and $y = y_0$ (from how we defined x and y above) and plug these into Equation 13.3 we get,

$$g'(0) = f_x(x_0, y_0)\,a + f_y(x_0, y_0)\,b \tag{13.4}$$

Now, simply equate Equation 13.2 and Equation 13.4 to get that,

$$D_{\vec{u}} f(x_0, y_0) = g'(0) = f_x(x_0, y_0)\,a + f_y(x_0, y_0)\,b$$

If we now go back to allowing x and y to be any number we get the following formula for computing directional derivatives.

2D Directional Derivative Formula

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

This is much simpler than the limit definition. Also note that this definition assumed that we were working with functions of two variables. There are similar formulas that can be derived by the same type of argument for functions with more than two variables. For instance, the directional derivative of $f(x, y, z)$ in the direction of the unit vector $\vec{u} = \langle a, b, c\rangle$ is given by,

3D Directional Derivative Formula

$$D_{\vec{u}} f(x, y, z) = f_x(x, y, z)\,a + f_y(x, y, z)\,b + f_z(x, y, z)\,c$$

Let's work a couple of examples.

Example 74

Find each of the directional derivatives.

(a) $D_{\vec{u}} f(2, 0)$ where $f(x, y) = x\mathbf{e}^{xy} + y$ and $\vec{u}$ is the unit vector in the direction of $\theta = \frac{2\pi}{3}$.

(b) $D_{\vec{u}} f(x, y, z)$ where $f(x, y, z) = x^2 z + y^3 z^2 - xyz$ in the direction of $\vec{v} = \langle -1, 0, 3\rangle$.

Solution

(a) $D_{\vec{u}} f(2, 0)$ where $f(x, y) = x\mathbf{e}^{xy} + y$ and $\vec{u}$ is the unit vector in the direction of $\theta = \frac{2\pi}{3}$.

We'll first find $D_{\vec{u}} f(x, y)$ and then use this a formula for finding $D_{\vec{u}} f(2, 0)$. The unit vector giving the direction is,

$$\vec{u} = \left\langle \cos\left(\frac{2\pi}{3}\right), \sin\left(\frac{2\pi}{3}\right)\right\rangle = \left\langle -\frac{1}{2}, \frac{\sqrt{3}}{2}\right\rangle$$

So, the directional derivative is,

$$D_{\vec{u}} f(x, y) = \left(-\frac{1}{2}\right)\left(\mathbf{e}^{xy} + xy\mathbf{e}^{xy}\right) + \left(\frac{\sqrt{3}}{2}\right)\left(x^2\mathbf{e}^{xy} + 1\right)$$

Now, plugging in the point in question gives,

$$D_{\vec{u}} f(2, 0) = \left(-\frac{1}{2}\right)(1) + \left(\frac{\sqrt{3}}{2}\right)(5) = \frac{5\sqrt{3} - 1}{2}$$

(b) $D_{\vec{u}} f(x, y, z)$ where $f(x, y, z) = x^2 z + y^3 z^2 - xyz$ in the direction of $\vec{v} = \langle -1, 0, 3\rangle$.

In this case let's first check to see if the direction vector is a unit vector or not and if it isn't convert it into one. To do this all we need to do is compute its magnitude.

$$\|\vec{v}\| = \sqrt{1 + 0 + 9} = \sqrt{10} \neq 1$$

So, it's not a unit vector. Recall that we can convert any vector into a unit vector that points in the same direction by dividing the vector by its magnitude. So, the unit vector that we need is,

$$\vec{u} = \frac{1}{\sqrt{10}}\langle -1, 0, 3\rangle = \left\langle -\frac{1}{\sqrt{10}}, 0, \frac{3}{\sqrt{10}}\right\rangle$$

The directional derivative is then,

$$\begin{aligned} D_{\vec{u}} f(x, y, z) &= \left(-\frac{1}{\sqrt{10}}\right)(2xz - yz) + (0)\left(3y^2 z^2 - xz\right) + \left(\frac{3}{\sqrt{10}}\right)\left(x^2 + 2y^3 z - xy\right) \\ &= \frac{1}{\sqrt{10}}\left(3x^2 + 6y^3 z - 3xy - 2xz + yz\right) \end{aligned}$$

There is another form of the formula that we used to get the directional derivative that is a little nicer and somewhat more compact. It is also a much more general formula that will encompass both of the formulas above.

Let's start with the second one and notice that we can write it as follows,

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

In other words, we can write the directional derivative as a dot product and notice that the second vector is nothing more than the unit vector $\vec{u}$ that gives the direction of change. Also, if we had used the version for functions of two variables the third component wouldn't be there, but other than that the formula would be the same.

Now let's give a name and notation to the first vector in the dot product since this vector will show up fairly regularly throughout this course (and in other courses). The **gradient of** f or **gradient vector of** f is defined to be,

$$\nabla f = \langle f_x, f_y, f_z \rangle \qquad \text{or} \qquad \nabla f = \langle f_x, f_y \rangle$$

Or, if we want to use the standard basis vectors the gradient is,

$$\nabla f = f_x \vec{i} + f_y \vec{j} + f_z \vec{k} \qquad \text{or} \qquad \nabla f = f_x \vec{i} + f_y \vec{j}$$

The definition is only shown for functions of two or three variables, however there is a natural extension to functions of any number of variables that we'd like.

With the definition of the gradient we can now say that the directional derivative is given by,

Directional Derivative Gradient Formula

$$D_{\vec{u}} f = \nabla f \cdot \vec{u}$$

where we will no longer show the variable and use this formula for any number of variables. Note as well that we will sometimes use the following notation,

$$D_{\vec{u}} f(\vec{x}) = \nabla f \cdot \vec{u}$$

where $\vec{x} = \langle x, y, z \rangle$ or $\vec{x} = \langle x, y \rangle$ as needed. This notation will be used when we want to note the variables in some way, but don't really want to restrict ourselves to a particular number of variables. In other words, $\vec{x}$ will be used to represent as many variables as we need in the formula and we will most often use this notation when we are already using vectors or vector notation in the problem/formula.

Let's work a couple of examples using this formula of the directional derivative.

Example 75

Find each of the directional derivatives.

(a) $D_{\vec{u}} f(\vec{x})$ for $f(x, y) = x \cos(y)$ in the direction of $\vec{v} = \langle 2, 1 \rangle$.

(b) $D_{\vec{u}} f(\vec{x})$ for $f(x, y, z) = \sin(yz) + \ln(x^2)$ at $(1, 1, \pi)$ in the direction of $\vec{v} = \langle 1, 1, -1 \rangle$.

Solution

(a) $D_{\vec{u}} f(\vec{x})$ for $f(x, y) = x \cos(y)$ in the direction of $\vec{v} = \langle 2, 1 \rangle$.

Let's first compute the gradient for this function.

$$\nabla f = \langle \cos(y), -x \sin(y) \rangle$$

Also, as we saw earlier in this section the unit vector for this direction is,

$$\vec{u} = \left\langle \frac{2}{\sqrt{5}}, \frac{1}{\sqrt{5}} \right\rangle$$

The directional derivative is then,

$$\begin{aligned} D_{\vec{u}} f(\vec{x}) &= \langle \cos(y), -x \sin(y) \rangle \cdot \left\langle \frac{2}{\sqrt{5}}, \frac{1}{\sqrt{5}} \right\rangle \\ &= \frac{1}{\sqrt{5}} (2 \cos(y) - x \sin(y)) \end{aligned}$$

(b) $D_{\vec{u}} f(\vec{x})$ for $f(x, y, z) = \sin(yz) + \ln(x^2)$ at $(1, 1, \pi)$ in the direction of $\vec{v} = \langle 1, 1, -1 \rangle$.

In this case are asking for the directional derivative at a particular point. To do this we will first compute the gradient, evaluate it at the point in question and then do the dot product. So, let's get the gradient.

$$\nabla f(x, y, z) = \left\langle \frac{2}{x}, z \cos(yz), y \cos(yz) \right\rangle$$

$$\nabla f(1, 1, \pi) = \left\langle \frac{2}{1}, \pi \cos(\pi), \cos(\pi) \right\rangle = \langle 2, -\pi, -1 \rangle$$

Next, we need the unit vector for the direction,

$$\|\vec{v}\| = \sqrt{3} \qquad \vec{u} = \left\langle \frac{1}{\sqrt{3}}, \frac{1}{\sqrt{3}}, -\frac{1}{\sqrt{3}} \right\rangle$$

Finally, the directional derivative at the point in question is,

$$D_{\vec{u}} f(1,1,\pi) = \langle 2, -\pi, -1\rangle \cdot \left\langle \frac{1}{\sqrt{3}}, \frac{1}{\sqrt{3}}, -\frac{1}{\sqrt{3}} \right\rangle$$
$$= \frac{1}{\sqrt{3}}(2 - \pi + 1)$$
$$= \frac{3 - \pi}{\sqrt{3}}$$

Before proceeding let's note that the first order partial derivatives that we were looking at in the majority of the section can be thought of as special cases of the directional derivatives. For instance, f_x can be thought of as the directional derivative of f in the direction of $\vec{u} = \langle 1, 0\rangle$ or $\vec{u} = \langle 1, 0, 0\rangle$, depending on the number of variables that we're working with. The same can be done for f_y and f_z

We will close out this section with a couple of nice facts about the gradient vector. The first tells us how to determine the maximum rate of change of a function at a point and the direction that we need to move in order to achieve that maximum rate of change.

Maximum Rate of Change

The maximum value of $D_{\vec{u}} f(\vec{x})$ (and hence then the maximum rate of change of the function $f(\vec{x})$) is given by $\|\nabla f(\vec{x})\|$ and will occur in the direction given by $\nabla f(\vec{x})$.

Proof

This is a really simple proof. First, if we start with the dot product form $D_{\vec{u}} f(\vec{x})$ and use a nice fact about dot products as well as the fact that $\vec{u}$ is a unit vector we get,

$$D_{\vec{u}} f = \nabla f \cdot \vec{u} = \|\nabla f\| \; \|\vec{u}\| \cos(\theta) = \|\nabla f\| \cos(\theta)$$

where θ is the angle between the gradient and $\vec{u}$.

Now the largest possible value of $\cos(\theta)$ is 1 which occurs at $\theta = 0$. Therefore the maximum value of $D_{\vec{u}} f(\vec{x})$ is $\|\nabla f(\vec{x})\|$ Also, the maximum value occurs when the angle between the gradient and $\vec{u}$ is zero, or in other words when $\vec{u}$ is pointing in the same direction as the gradient, $\nabla f(\vec{x})$.

Let's take a quick look at an example.

Example 76

Suppose that the height of a hill above sea level is given by $z = 1000 - 0.01x^2 - 0.02y^2$. If you are at the point $(60, 100)$ in what direction is the elevation changing fastest? What is the maximum rate of change of the elevation at this point?

Solution

First, you will hopefully recall from the Quadric Surfaces section that this is an elliptic paraboloid that opens downward. So even though most hills aren't this symmetrical it will at least be vaguely hill shaped and so the question makes at least a little sense.

Now on to the problem. There are a couple of questions to answer here, but using the theorem makes answering them very simple. We'll first need the gradient vector.

$$\nabla f(\vec{x}) = \langle -0.02x, -0.04y \rangle$$

The maximum rate of change of the elevation will then occur in the direction of

$$\nabla f(60, 100) = \langle -1.2, -4 \rangle$$

The maximum rate of change of the elevation at this point is,

$$\|\nabla f(60, 100)\| = \sqrt{(-1.2)^2 + (-4)^2} = \sqrt{17.44} = 4.176$$

Before leaving this example let's note that we're at the point $(60, 100)$ and the direction of greatest rate of change of the elevation at this point is given by the vector $\langle -1.2, -4 \rangle$. Since both of the components are negative it looks like the direction of maximum rate of change points up the hill towards the center rather than away from the hill.

The second fact about the gradient vector that we need to give in this section will be very convenient in some later sections.

Fact

The gradient vector $\nabla f(x_0, y_0)$ is orthogonal (or perpendicular) to the level curve $f(x, y) = k$ at the point (x_0, y_0). Likewise, the gradient vector $\nabla f(x_0, y_0, z_0)$ is orthogonal to the level surface $f(x, y, z) = k$ at the point (x_0, y_0, z_0).

Proof

We're going to do the proof for the $\mathbb{R}^3$ case. The proof for the $\mathbb{R}^2$ case is identical. We'll also need some notation out of the way to make life easier for us let's let S be the level surface given by $f(x, y, z) = k$ and let $P = (x_0, y_0, z_0)$. Note as well that P will be on S.

Now, let C be any curve on S that contains P. Let $\vec{r}(t) = \langle x(t), y(t), z(t) \rangle$ be the vector equation for C and suppose that t_0 be the value of t such that $\vec{r}(t_0) = \langle x_0, y_0, z_0 \rangle$. In other words, t_0 be the value of t that gives P.

Because C lies on S we know that points on C must satisfy the equation for S. Or,

$$f(x(t), y(t), z(t)) = k$$

Next, let's use the Chain Rule on this to get,

$$\frac{\partial f}{\partial x}\frac{dx}{dt} + \frac{\partial f}{\partial y}\frac{dy}{dt} + \frac{\partial f}{\partial z}\frac{dz}{dt} = 0$$

Notice that $\nabla f = \langle f_x, f_y, f_z \rangle$ and $\vec{r}'(t) = \langle x'(t), y'(t), z'(t) \rangle$ so this becomes,

$$\nabla f \cdot \vec{r}'(t) = 0$$

At, $t = t_0$ this is,

$$\nabla f(x_0, y_0, z_0) \cdot \vec{r}'(t_0) = 0$$

This then tells us that the gradient vector at P , $\nabla f(x_0, y_0, z_0)$, is orthogonal to the tangent vector, $\vec{r}'(t_0)$, to any curve C that passes through P and on the surface S and so must also be orthogonal to the surface S.

As we will be seeing in later sections we are often going to be needing vectors that are orthogonal to a surface or curve and using this fact we will know that all we need to do is compute a gradient vector and we will get the orthogonal vector that we need. We will see the first application of this in the next chapter.

2.8 Applications of Partial Derivatives

In this section we'll take a look at a couple of applications of partial derivatives. The applications here are either very similar to applications we saw for derivatives of single variable functions or extensions of those applications.

For example we will be looking at the tangent plane to a surface rather than tangent lines to curves as we did with single variable functions.

In addition we be finding relative and absolute extrema of multi-variable functions. The difference in this chapter compared to the last time we saw these applications is that they will often involve a lot more work. Because of the increased difficulty of the problems we'll be restricting ourselves to finding the relative and absolute extrema of functions of two variables only.

We will also be looking at Lagrange Multipliers. This is a method that will allow us to optimize a function that is subject to some constraint. That is to say optimizing a function of two or three variables where the variables must also satisfy some constraint (usually in the form on an equation involving the variables).

2.8 Tangent Planes and Linear Approximations

Earlier we saw how the two partial derivatives f_x and f_y can be thought of as the slopes of traces. We want to extend this idea out a little in this section. The graph of a function $z = f(x, y)$ is a surface in $\mathbb{R}^3$ (three dimensional space) and so we can now start thinking of the plane that is "tangent" to the surface as a point.

Let's start out with a point (x_0, y_0) and let's let C_1 represent the trace to $f(x, y)$ for the plane $y = y_0$ (*i.e.* allowing x to vary with y held fixed) and we'll let C_2 represent the trace to $f(x, y)$ for the plane $x = x_0$ (*i.e.* allowing y to vary with x held fixed). Now, we know that $f_x(x_0, y_0)$ is the slope of the tangent line to the trace C_1 and $f_y(x_0, y_0)$ is the slope of the tangent line to the trace C_2. So, let L_1 be the tangent line to the trace C_1 and let L_2 be the tangent line to the trace C_2.

The tangent plane will then be the plane that contains the two lines L_1 and L_2. Geometrically this plane will serve the same purpose that a tangent line did in Calculus I. A tangent line to a curve was a line that just touched the curve at that point and was "parallel" to the curve at the point in question. Well tangent planes to a surface are planes that just touch the surface at the point and are "parallel" to the surface at the point. Note that this gives us a point that is on the plane. Since the tangent plane and the surface touch at (x_0, y_0) the following point will be on both the surface and the plane.

$$(x_0, y_0, z_0) = (x_0, y_0, f(x_0, y_0))$$

What we need to do now is determine the equation of the tangent plane. We know that the general equation of a plane is given by,

$$a(x - x_0) + b(y - y_0) + c(z - z_0) = 0$$

where (x_0, y_0, z_0) is a point that is on the plane, which we have. Let's rewrite this a little. We'll move the x terms and y terms to the other side and divide both sides by c. Doing this gives,

$$z - z_0 = -\frac{a}{c}(x - x_0) - \frac{b}{c}(y - y_0)$$

Now, let's rename the constants to simplify up the notation a little. Let's rename them as follows,

$$A = -\frac{a}{c} \qquad B = -\frac{b}{c}$$

With this renaming the equation of the tangent plane becomes,

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

and we need to determine values for A and B.

Let's first think about what happens if we hold y fixed, *i.e.* if we assume that $y = y_0$. In this case the equation of the tangent plane becomes,

$$z - z_0 = A(x - x_0)$$

This is the equation of a line and this line must be tangent to the surface at (x_0, y_0) (since it's part of the tangent plane). In addition, this line assumes that $y = y_0$ (*i.e.* fixed) and A is the slope of this line. But if we think about it this is exactly what the tangent to C_1 is, a line tangent to the surface at (x_0, y_0) assuming that $y = y_0$. In other words,

$$z - z_0 = A(x - x_0)$$

is the equation for L_1 and we know that the slope of L_1 is given by $f_x(x_0, y_0)$. Therefore, we have the following,

$$A = f_x(x_0, y_0)$$

If we hold x fixed at $x = x_0$ the equation of the tangent plane becomes,

$$z - z_0 = B(y - y_0)$$

However, by a similar argument to the one above we can see that this is nothing more than the equation for L_2 and that it's slope is B or $f_y(x_0, y_0)$. So,

$$B = f_y(x_0, y_0)$$

The equation of the tangent plane to the surface given by $z = f(x, y)$ at (x_0, y_0) is then,

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

Also, if we use the fact that $z_0 = f(x_0, y_0)$ we can rewrite the equation of the tangent plane as,

$$\begin{aligned} z - f(x_0, y_0) &= f_x(x_0, y_0)(x - x_0) + f_y(x_0, y_0)(y - y_0) \\ z &= f(x_0, y_0) + f_x(x_0, y_0)(x - x_0) + f_y(x_0, y_0)(y - y_0) \end{aligned}$$

We will see an easier derivation of this formula (actually a more general formula) in the next section so if you didn't quite follow this argument hold off until then to see a better derivation.

Example 77

Find the equation of the tangent plane to $z = \ln(2x + y)$ at $(-1, 3)$.

Solution

There really isn't too much to do here other than taking a couple of derivatives and doing some quick evaluations.

$$f(x,y) = \ln(2x+y) \qquad z_0 = f(-1,3) = \ln(1) = 0$$
$$f_x(x,y) = \frac{2}{2x+y} \qquad f_x(-1,3) = 2$$
$$f_y(x,y) = \frac{1}{2x+y} \qquad f_y(-1,3) = 1$$

The equation of the plane is then,

$$z - 0 = 2(x+1) + (1)(y-3)$$
$$z = 2x + y - 1$$

One nice use of tangent planes is they give us a way to approximate a surface near a point. As long as we are near to the point (x_0, y_0) then the tangent plane should nearly approximate the function at that point. Because of this we define the **linear approximation** to be,

$$L(x,y) = f(x_0,y_0) + f_x(x_0,y_0)(x-x_0) + f_y(x_0,y_0)(y-y_0)$$

and as long as we are "near" (x_0, y_0) then we should have that,

$$f(x,y) \approx L(x,y) = f(x_0,y_0) + f_x(x_0,y_0)(x-x_0) + f_y(x_0,y_0)(y-y_0)$$

Example 78

Find the linear approximation to $z = 3 + \frac{x^2}{16} + \frac{y^2}{9}$ at $(-4, 3)$.

Solution

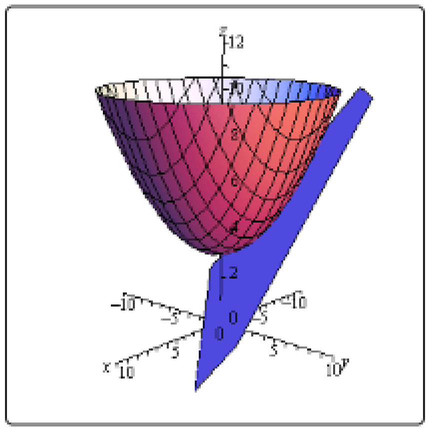
So, we're really asking for the tangent plane so let's find that.

$$f(x,y) = 3 + \frac{x^2}{16} + \frac{y^2}{9} \qquad f(-4,3) = 3 + 1 + 1 = 5$$
$$f_x(x,y) = \frac{x}{8} \qquad f_x(-4,3) = -\frac{1}{2}$$
$$f_y(x,y) = \frac{2y}{9} \qquad f_y(-4,3) = \frac{2}{3}$$

The tangent plane, or linear approximation, is then,

$$L(x,y) = 5 - \frac{1}{2}(x+4) + \frac{2}{3}(y-3)$$

For reference purposes here is a sketch of the surface and the tangent plane/linear approximation.


z
12
10
8
6
4
2
0
−10
−5
5
x 10
0
−10
−5
5
10 y

2.9 Gradient Vector, Tangent Planes and Normal Lines

In this section we want to revisit tangent planes only this time we'll look at them in light of the gradient vector. In the process we will also take a look at a normal line to a surface.

Let's first recall the equation of a plane that contains the point (x_0, y_0, z_0) with normal vector $\vec{n} = \langle a, b, c \rangle$ is given by,

$$a(x - x_0) + b(y - y_0) + c(z - z_0) = 0$$

When we introduced the gradient vector in the section on directional derivatives we gave the following fact.

Fact

The gradient vector $\nabla f(x_0, y_0)$ is orthogonal (or perpendicular) to the level curve $f(x, y) = k$ at the point (x_0, y_0). Likewise, the gradient vector $\nabla f(x_0, y_0, z_0)$ is orthogonal to the level surface $f(x, y, z) = k$ at the point (x_0, y_0, z_0).

Actually, all we need here is the last part of this fact. This says that the gradient vector is always orthogonal, or *normal*, to the surface at a point.

Also recall that the gradient vector is,

$$\nabla f = \langle f_x, f_y, f_z \rangle$$

So, the tangent plane to the surface given by $f(x, y, z) = k$ at (x_0, y_0, z_0) has the equation,

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

This is a much more general form of the equation of a tangent plane than the one that we derived in the previous section.

Note however, that we can also get the equation from the previous section using this more general formula. To see this let's start with the equation $z = f(x, y)$ and we want to find the tangent plane to the surface given by $z = f(x, y)$ at the point (x_0, y_0, z_0) where $z_0 = f(x_0, y_0)$. In order to use the formula above we need to have all the variables on one side. This is easy enough to do. All we need to do is subtract a z from both sides to get,

$$f(x, y) - z = 0$$

Now, if we define a new function

$$F(x, y, z) = f(x, y) - z$$

we can see that the surface given by $z = f(x, y)$ is identical to the surface given by $F(x, y, z) = 0$ and this new equivalent equation is in the correct form for the equation of the tangent plane that we derived in this section.

So, the first thing that we need to do is find the gradient vector for F.

$$\nabla F = \langle F_x, F_y, F_z \rangle = \langle f_x, f_y, -1 \rangle$$

Notice that

$$F_x = \frac{\partial}{\partial x}\Big(f(x,y) - z\Big) = f_x \qquad F_y = \frac{\partial}{\partial y}\Big(f(x,y) - z\Big) = f_y$$
$$F_z = \frac{\partial}{\partial z}\Big(f(x,y) - z\Big) = -1$$

The equation of the tangent plane is then,

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

Or, upon solving for z, we get,

$$z = f(x_0, y_0) + f_x(x_0, y_0)(x - x_0) + f_y(x_0, y_0)(y - y_0)$$

which is identical to the equation that we derived in the previous section.

We can get another nice piece of information out of the gradient vector as well. We might on occasion want a line that is orthogonal to a surface at a point, sometimes called the **normal line**. This is easy enough to get if we recall that the equation of a line only requires that we have a point and a parallel vector. Since we want a line that is at the point (x_0, y_0, z_0) we know that this point must also be on the line and we know that $\nabla f(x_0, y_0, z_0)$ is a vector that is normal to the surface and hence will be parallel to the line. Therefore, the equation of the normal line is,

$$\vec{r}(t) = \langle x_0, y_0, z_0 \rangle + t\,\nabla f(x_0, y_0, z_0)$$

Example 79

Find the tangent plane and normal line to $x^2 + y^2 + z^2 = 30$ at the point $(1, -2, 5)$.

Solution

For this case the function that we're going to be working with is,

$$F(x, y, z) = x^2 + y^2 + z^2$$

and note that we don't have to have a zero on one side of the equal sign. All that we need is a constant. To finish this problem out we simply need the gradient evaluated at the point.

$$\nabla F = \langle 2x, 2y, 2z \rangle$$
$$\nabla F(1, -2, 5) = \langle 2, -4, 10 \rangle$$

The tangent plane is then,

$$2\,(x-1)-4\,(y+2)+10\,(z-5)=0$$

The normal line is,

$$\vec{r}\,(t)=\langle 1,-2,5\rangle+t\,\langle 2,-4,10\rangle=\langle 1+2t,-2-4t,5+10t\rangle$$