Preface

Here are my online notes for my Calculus III course that I teach here at Lamar University. Despite the fact that these are my “class notes”, they should be accessible to anyone wanting to learn Calculus III or needing a refresher in some of the topics from the class.

These notes do assume that the reader has a good working knowledge of Calculus I topics including limits, derivatives and integration. It also assumes that the reader has a good knowledge of several Calculus II topics including some integration techniques, parametric equations, vectors, and knowledge of three dimensional space.

Here are a couple of warnings to my students who may be here to get a copy of what happened on a day that you missed.

1. Because I wanted to make this a fairly complete set of notes for anyone wanting to learn calculus I have included some material that I do not usually have time to cover in class and because this changes from semester to semester it is not noted here. You will need to find one of your fellow class mates to see if there is something in these notes that wasn’t covered in class.

2. In general I try to work problems in class that are different from my notes. However, with Calculus III many of the problems are difficult to make up on the spur of the moment and so in this class my class work will follow these notes fairly close as far as worked problems go. With that being said I will, on occasion, work problems off the top of my head when I can to provide more examples than just those in my notes. Also, I often don’t have time in class to work all of the problems in the notes and so you will find that some sections contain problems that weren’t worked in class due to time restrictions.

3. Sometimes questions in class will lead down paths that are not covered here. I try to anticipate as many of the questions as possible in writing these up, but the reality is that I can’t anticipate all the questions. Sometimes a very good question gets asked in class that leads to insights that I’ve not included here. You should always talk to someone who was in class on the day you missed and compare these notes to their notes and see what the differences are.

4. This is somewhat related to the previous three items, but is important enough to merit its own item. THESE NOTES ARE NOT A SUBSTITUTE FOR ATTENDING CLASS!! Using these notes as a substitute for class is liable to get you in trouble. As already noted not everything in these notes is covered in class and often material or insights not in these notes is covered in class.
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 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 \( \vec{v} = \langle 2, 1 \rangle \).

\[
\vec{v} = \left\langle \frac{1}{5}, \frac{1}{10} \right\rangle \quad \vec{v} = \langle 6, 3 \rangle \quad \vec{v} = \left\langle \frac{2}{\sqrt{5}}, \frac{1}{\sqrt{5}} \right\rangle
\]

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,
Calculus III

\[ \| \mathbf{v} \| = \sqrt{a^2 + b^2 + c^2} \]

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 \( \theta = \frac{\pi}{4} \). The unit vector that points in this direction is given by,

\[ \mathbf{u} = \langle \cos \theta, \sin \theta \rangle \]

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 \( \mathbf{u} = \langle a, b \rangle \) is called the **directional derivative** and is denoted by \( D_{\mathbf{u}} f(x, y) \). The definition of the directional derivative is,

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

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,

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

So, it looks like we have the following relationship.

\[ g'(0) = D_{\mathbf{u}} f(x_0, y_0) \]  \hspace{1cm} (1)
Now, let’s look at this from another perspective. Let’s rewrite \( g(z) \) as follows,

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

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.

\[
g'(z) = f_x(x, y)a + f_y(x, y)b \quad \text{(2)}
\]

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 (2) we get,

\[
g'(0) = f_x(x_0, y_0)a + f_y(x_0, y_0)b \quad \text{(3)}
\]

Now, simply equate (1) and (3) to get that,

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

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

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

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,

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

Let’s work a couple of examples.

**Example 1** Find each of the directional derivatives.

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

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

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

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) \( D_\vec{v} f(x, y, z) \) 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,

\[
D_\vec{u} f(x, y, z) = \left( \frac{1}{\sqrt{10}} \right) (2xz - yz) + (0) (3y^2 z^2 - xz) + \left( \frac{3}{\sqrt{10}} \right) (x^2 + 2y^3 z - xy)
\]

\[
= \frac{1}{\sqrt{10}} \left( 3x^2 + 6y^3 z - 3xy - 2xz + yz \right)
\]
Calculus III

\[ D_u f(x, y, z) = f_x(x, y, z)a + f_y(x, y, z)b + f_z(x, y, z)c \]
\[ = (f_x, f_y, f_z) \cdot (a, b, c) \]

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,

\[ \nabla f = \langle f_x, f_y, f_z \rangle \quad \text{or} \quad \nabla f = \langle f_x, f_y \rangle \]

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

\[ \nabla f = f_x \mathbf{i} + f_y \mathbf{j} + f_z \mathbf{k} \quad \text{or} \quad \nabla f = f_x \mathbf{i} + f_y \mathbf{j} \]

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,

\[ D_u f = \nabla f \cdot \vec{u} \]

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,

\[ D_u f(\vec{x}) = \nabla f \cdot \vec{u} \]

where \( \vec{x} = (x, y, z) \) or \( \vec{x} = (x, y) \) 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 2  Find each of the directional derivative.
(a) \( D_u f(\vec{x}) \) for \( f(x, y) = x \cos(y) \) in the direction of \( \vec{v} = \langle 2, 1 \rangle \). [Solution]
(b) \( D_u f(\vec{x}) \) 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]
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,

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

$$= \frac{1}{\sqrt{5}} \left( 2\cos(y) - x\sin(y) \right)$$

(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,

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

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.

**Theorem**

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

**Proof**

This is a really simple proof. First, if we start with the dot product form \( D_{\vec{u}} f \left( \vec{x} \right) \) 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 \( \theta \) 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 \( D_{\vec{u}} f \left( \vec{x} \right) \) is \( \| \nabla f \left( \vec{x} \right) \| \). 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, \( \nabla f \left( \vec{x} \right) \).

Let’s take a quick look at an example.

**Example 3** 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 \( \left(60,100\right) \) 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.

\[
\nabla f \left( \vec{x} \right) = \langle -0.02x, -0.04y \rangle
\]

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

\[
\nabla f \left(60,100\right) = \langle -1.2, -4 \rangle
\]
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 \((-1.2, -4)\). 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 \( \vec{r}(t) = (x(t), y(t), z(t)) \) be the vector
equation for \( C \) and suppose that \( t_0 \) be the value of \( t \) such that \( \vec{r}(t_0) = (x_0, y_0, z_0) \). 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,

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

\[ (4) \]

Notice that \( \nabla f = (f_x, f_y, f_z) \) and \( \vec{r}'(t) = (x'(t), y'(t), z'(t)) \) so \( (4) \) 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.