Key Ideas: Directional Derivatives and the Gradient Vector

- Suppose that the graph of \( f(x, y) = \frac{100}{x^2 + y^4 + 2} \) is a mountain, and I’m standing on it at the point \((1, 1, 25)\). We know that since \( f_x(1, 1) = -12.5 \) this means that I will go down by about 12.5 units for every unit I move in the positive \( x \) direction. Similarly, since \( f_y(1, 1) = -25 \) this means I will go down by 25 units for every unit I go in the positive \( y \) direction. But what if I move in another direction? For example, what if I’m moving in the direction of the vector \( \langle -1/\sqrt{2}, 1/\sqrt{2} \rangle \)? Today we will answer this question, and we will see that at least in the differentiable case, the answer is relatively simple.

- We will begin to equate the notion of “direction” with the idea of “unit vectors”. This makes sense because for any non-zero vector there is a unique unit vector which points in the same direction. But, it is important to make sure that you use a unit vector in the formulas below and not any other vector or the values computed will be incorrect.

- **Definition:** Let \( u = \langle a, b \rangle \) be any unit vector. We define the derivative of \( f \) in the direction of \( u \) to be

  \[
  D_u f(x, y) = \lim_{h \to 0} \frac{f(x + ha, y + hb) - f(x, y)}{h} = \lim_{h \to 0} \frac{f(x + hu) - f(x)}{h}.
  \]

- So, in particular, \( f_x \) is the derivative in the direction \( \langle 1, 0 \rangle \) and \( f_y \) is the derivative in the direction \( \langle 0, 1 \rangle \), but now we can (in theory) take the derivative in any direction. And, of course, the value of this derivative gives a measure of the rate at which \( f \) would increase (or decrease) if the input point were changed from \( (x, y) \) by moving in the direction \( u! \)

- At this point it is convenient to introduce the **gradient vector** of the function \( f(x, y) \):

  \[
  \nabla f(x, y) = \langle f_x(x, y), f_y(x, y) \rangle.
  \]

  This is a vector valued function of two variables. (For each specific \( x \) and \( y \) you get a vector.) It is also sometimes written as \( \text{grad} \ f \).

- Fortunately, we almost never have to resort to using the limit definition of directional derivatives, because for any differentiable function \( f(x, y) \) it is true that

  \[
  D_u f(x, y) = f_x(x, y)a + f_y(x, y)b = \nabla f(x, y) \cdot u.
  \]

  (Why is this true? It is because of the tangent plane. On the tangent plane, \( z \) increases by the same amount for every increase of \( x \) by one and for the same amount for every increase of \( y \) by one...so if you know that \( x \) increases by \( a \) and \( y \) increases by \( b \), the formula above gives the corresponding increase in \( z \).)
• The same basic idea applies for functions of three variables. If \( \mathbf{u} = \langle a, b, c \rangle \) is a unit vector and \( f(x, y, z) \) is a function then we can talk about the directional derivative of \( f \) at this point. Its definition in terms of limits is exactly the same as the definition in the two variable case (when written in vector form!) and it continues to be true that

\[
\nabla f = \langle f_x, f_y, f_z \rangle \quad D_{\mathbf{u}} f = \nabla f(x, y, z) \cdot \mathbf{u}.
\]

• The gradient vector has another use. Not only is it a way to conveniently carry all of your first partial derivatives around together in a single package, it also points in an important direction. It points in the direction that the function increases the fastest. This is a bit of a surprise...in the two variable case we would think of it as just being two slopes put together, but it is easy to check!

**Theorem:** If \( f \) is a differentiable function of two or three variables then the maximum value of the directional derivative \( D_{\mathbf{u}} f \) is \( |\nabla f(x)| \) and it occurs when \( \mathbf{u} \) points in the direction of the vector \( \nabla f \)!

The proof of this fact is just a simple computation and basic trig. Since \( D_{\mathbf{u}} f = \nabla f \cdot \mathbf{u} = |\nabla f| \cos \theta \), this is largest when \( \theta = 0 \) and \( \cos \theta = 1 \).

**Gradient Vectors and Level Curves:** Consider the curve \( f(x, y) = k \) which is the trace of the graph of \( z = f(x, y) \) at height \( k \). You can think of this as a mountain pass that stays at the same height...if you walk along the level curve you neither go up nor down. Then, if you parametrize this path \( \mathbf{r}(t) = \langle x(t), y(t) \rangle \) you have the equation \( f(\mathbf{r}) = k \) whose derivative (according to the chain rule) is

\[
\nabla f \cdot \mathbf{r}' = 0.
\]

In other words, the gradient vector is orthogonal to the level curves. This make sense because going a little bit more to the left or right of this angle brings you closer to the path that stays at the same height, so if you want to follow the curve of steepest ascent you always follow the gradient vector.

• A similar idea applies in the case of level surfaces of functions of three variables. Since we did not cover level surfaces very carefully before, lets first take a moment to think of an explicit example. Consider a function like \( f(x, y, z) = x^2 + y^2 + z \). If you pick a number like 8 and consider all of the points \( (x, y, z) \) in space such that \( f(x, y, z) = 8 \) you get the “level surface” for that function, which looks like a paraboloid starting at height 8 above the origin and coming down from there. These are the set of points for which the function has the value 8.

The gradient vector \( \nabla f \) is normal to the surface \( f(x, y, z) = k \) at each point. So, in particular,

\[
f_x(a, b, c)(x - a) + f_y(a, b, c)(y - b) + f_z(a, b, c)(z - c) = 0
\]

is the equation of the tangent plane to the surface at the point \( (a, b, c) \).

---

**Homework**

**Section 14.6:** Read the section and do these problems: 7–13, 14, 16, 17, 19–21, 40a, 41a, 42a, 52

**Test Next Week!** I don’t believe it, but it is already time for our second test. So, there are no “starred” problems on this assignment after all.