Various Topics Related to the Gradient Vector

Finding the Maximum Value of the Directional Derivative:

For the function $z = f(x,y) = x^2 + y^2$, we have the gradient $\nabla f = < 2x, 2y >$. At the point $(x_0, y_0) = (1,8)$, $\nabla f(1,8) = < 2, 16 >$.

For any unit vector $\mathbf{u} = < a, b >$, the directional derivative of f at the point $(1,8)$ in the direction of $\mathbf{u}$ is $D_{\mathbf{u}} f(1,8) = \mathbf{u} \cdot \nabla f(1,8) = < a, b > \cdot < 2, 16 > = 2a + 16b$.

Let us ask the question, what is the maximum possible value that $D_{\mathbf{u}} f(1,8)$ can have? And for which unit vector does it achieve this maximum value?

Since $\mathbf{u}$ is a unit vector, $\sqrt{a^2 + b^2} = 1$, so $b = \pm\sqrt{1 - a^2}$, where $a \in [-1, 1]$. If we substitute this into the formula $2a + 16b$, we get $2a \pm 16\sqrt{1 - a^2}$. The maximum value this can have is obtained if we choose the "plus-or-minus" to be "plus." Using a graphing calculator, we find the maximum value of $2a + 16\sqrt{1 - a^2}$ to be approximately 16.125, which occurs at the approximate value $a = 0.124$. Using the formula $b = \sqrt{1 - a^2}$ and substituting 0.124 for a, we compute 0.992 as the approximate value of b. Thus, at the point $(1,8)$, the maximum value of the directional derivative will be about 16.125, and this will occur in the direction of the unit vector that is approximately $< 0.124, 0.992 >$.

For any unit vector $\mathbf{u}$, let θ denote the angle between $\mathbf{u}$ and $\nabla f(1,8)$. As discussed in Chapter 11, θ must lie in the interval $[0, \pi]$. If $\theta = 0$, then $\mathbf{u}$ points in the same direction as $\nabla f(1,8)$. If $\theta = \pi$, then $\mathbf{u}$ points in the opposite direction from $\nabla f(1,8)$. If $\theta = \frac{\pi}{2}$, then $\mathbf{u}$ is orthogonal (i.e., perpendicular) to $\nabla f(1,8)$.

In the case where $\theta = 0$, $\mathbf{u}$ would be the unit vector in the direction of $< 2, 16 >$, which we can easily compute: Since the magnitude of $< 2, 16 >$ is $\sqrt{260}$ or $2\sqrt{65}$, we would have $\frac{1}{2\sqrt{65}} < 2, 16 > = < \frac{1}{\sqrt{65}}, \frac{8}{\sqrt{65}} >$. For this $\mathbf{u}$, we obtain $D_{\mathbf{u}} f(1,8) = < \frac{1}{\sqrt{65}}, \frac{8}{\sqrt{65}} > \cdot < 2, 16 >$ $= \frac{2}{\sqrt{65}} + \frac{128}{\sqrt{65}} = \frac{130}{\sqrt{65}}$.

Here, the first component of $\mathbf{u}$ is $\frac{1}{\sqrt{65}}$, whose decimal value is approximately 0.124, and the second component is $\frac{8}{\sqrt{65}}$, whose decimal value is approximately 0.992. The exact value of $D_{\mathbf{u}} f(1,8)$ is $\frac{130}{\sqrt{65}}$, whose decimal value is approximately 16.125. But this is the value we obtained earlier when we approximated the maximum value of $D_{\mathbf{u}} f(1,8)$ via a graphing calculator. Furthermore, $\mathbf{u}$ is the vector which we approximated earlier as $< 0.124, 0.992 >$. Thus, the unit vector in the same direction as $\nabla f(1,8)$, i.e., the unit vector where $\theta = 0$, is the unit vector that produces the maximum value of the directional derivative.

If we rationalize the denominator for the maximum directional derivative, $\frac{130}{\sqrt{65}} = \frac{130\sqrt{65}}{65} = 2\sqrt{65}$, which happens to be the length of $\nabla f(1,8)$. Thus, the maximum value of the directional derivative at the point $(1,8)$ is equal to the magnitude of $\nabla f(1,8)$.

In summary, $D_{\mathbf{u}} f(1,8)$ achieves its maximum value when $\mathbf{u}$ points in the direction of $\nabla f(1,8)$ (i.e., when $\theta = 0$), and the maximum value of $D_{\mathbf{u}} f(1,8)$ is equal to the magnitude of $\nabla f(1,8)$. Is this a coincidence? No! In general, so long as $\nabla f(x_0,y_0) \neq \mathbf{0}$, then $D_{\mathbf{u}} f(x_0,y_0)$ will achieve its maximum value when $\mathbf{u}$ has the same direction as $\nabla f(x_0,y_0)$ (i.e., when the angle between $\mathbf{u}$ and $\nabla f(x_0,y_0)$ is 0), and the maximum value of $D_{\mathbf{u}} f(x_0,y_0)$ will equal the magnitude of $\nabla f(x_0,y_0)$. Let us explore *why* this is so.

Recall that for any nonzero vectors $\mathbf{a}$ and $\mathbf{b}$, if θ is the angle between them, then $\mathbf{a} \cdot \mathbf{b} = |\mathbf{a}||\mathbf{b}|\cos\theta$.

Hence, $D_{\mathbf{u}} f(x_0,y_0) = \mathbf{u} \cdot \nabla f(x_0,y_0) = |\mathbf{u}||\nabla f(x_0,y_0)|\cos\theta = |\nabla f(x_0,y_0)|\cos\theta$. From this we can see the following:

- Since $\cos\theta$ is positive for $\theta \in [0, \frac{\pi}{2})$, the directional derivative is positive when $\theta \in [0, \frac{\pi}{2})$.
- Since $\cos\theta$ is negative for $\theta \in (\frac{\pi}{2}, \pi]$, the directional derivative is negative when $\theta \in (\frac{\pi}{2}, \pi]$.
- Since $\cos\frac{\pi}{2} = 0$, the directional derivative is zero when $\theta = \frac{\pi}{2}$.
- Since $\cos 0 = 1$, the directional derivative has the value $|\nabla f(x_0,y_0)|$ when $\theta = 0$. This is the *maximum* value of the directional derivative.
- Since $\cos\theta \in (0,1)$ for $\theta \in (0, \frac{\pi}{2})$, the value of the directional derivative must be *between* 0 and $|\nabla f(x_0,y_0)|$ when $\theta \in (0, \frac{\pi}{2})$.
- Since $\cos\pi = -1$, the directional derivative has the value $-|\nabla f(x_0,y_0)|$ when $\theta = \pi$. This is the *minimum* value of the directional derivative.
- Since $\cos\theta \in (-1,0)$ for $\theta \in (\frac{\pi}{2}, \pi)$, the value of the directional derivative must be *between* $-|\nabla f(x_0,y_0)|$ and 0 when $\theta \in (\frac{\pi}{2}, \pi)$.

Picture this: We start with a unit vector positioned with its tail at the point (x_0,y_0), pointing in the same direction as $\nabla f(x_0,y_0)$, so $\theta = 0$. For this vector, the directional derivative has its maximum value, equal to $|\nabla f(x_0,y_0)|$, which is a positive number. Now keep the tail fixed at (x_0,y_0) and rotate the unit vector away from $\nabla f(x_0,y_0)$. As we rotate it away, θ increases and the value of the directional derivative steadily declines toward 0. The directional derivative reaches 0 when $\theta = \frac{\pi}{2}$, i.e., when our unit vector is orthogonal (perpendicular) to $\nabla f(x_0,y_0)$. As we rotate our unit vector even further away from $\nabla f(x_0,y_0)$, θ further increases, and the directional derivative becomes negative. As we continue our rotation, the value of the directional derivative continues to decrease. But bear in mind that we are now dealing with a negative value, so saying that it “decreases” means its absolute value is increasing. The minimum value of the directional derivative (i.e., the negative value with the largest absolute value) is obtained when $\theta = \pi$, i.e., when our unit vector is the exact opposite of $\nabla f(x_0,y_0)$, and this minimum value is equal to $-|\nabla f(x_0,y_0)|$, which is a negative number.

In the case of the example we have been considering, since the unit vector in the direction of $\nabla f(1,8)$ is $< \frac{1}{\sqrt{65}}, \frac{8}{\sqrt{65}} >$, the two unit vectors orthogonal to $\nabla f(1,8)$ are $< \frac{8}{\sqrt{65}}, \frac{-1}{\sqrt{65}} >$ and $< \frac{-8}{\sqrt{65}}, \frac{1}{\sqrt{65}} >$. Go ahead and compute the directional derivative at the point $(1,8)$ for each of these; you will find it is zero. Furthermore, the unit vector in the opposite direction from $\nabla f(1,8)$ is $< \frac{-1}{\sqrt{65}}, \frac{-8}{\sqrt{65}} >$. Go ahead and compute the directional derivative for this unit vector; you will find it is $\frac{-130}{\sqrt{65}}$, which is $-|\nabla f(1,8)|$.

The Relationship Between Gradient Vectors and Level Curves:

For the function $z = f(x,y) = x^2 + y^2$, the circle in the x,y plane centered at the origin with radius 6, $x^2 + y^2 = 36$, is a level curve for the function, i.e., it is the level curve $z = 36$. Suppose we parameterize this circle in the usual way: Let $x = 6\cos t$, $y = 6\sin t$. At any point P_t, the velocity vector is $\mathbf{v}(t) = < \frac{dx}{dt}, \frac{dy}{dt} > = < -6\sin t, 6\cos t >$ and the speed is constant, namely, $v(t) = 6$. Thus, the unit tangent vector at P_t is $\mathbf{T}(t) = \frac{1}{v(t)}\mathbf{v}(t) = \frac{1}{6} < -6\sin t, 6\cos t > = < -\sin t, \cos t >$.

Consider any point on this level curve, such as the point generated when $t = \frac{\pi}{3}$, which is $(3, 3\sqrt{3})$. At this point, the unit tangent vector is $\mathbf{T}(\frac{\pi}{3}) = < -\frac{\sqrt{3}}{2}, \frac{1}{2} >$. We can compute the directional derivative of $f(x,y)$ in the direction of $< -\frac{\sqrt{3}}{2}, \frac{1}{2} >$, but we must first find $\nabla f(3, 3\sqrt{3})$, which turns out to be $< 6, 6\sqrt{3} >$. Thus, the directional derivative is $< -\frac{\sqrt{3}}{2}, \frac{1}{2} > \bullet < 6, 6\sqrt{3} > = 0$.

This should ring a bell. When is the directional derivative zero? When it is computed in a direction perpendicular to the gradient vector. Is that the case in this example? Well, $\mathbf{T}(\frac{\pi}{3}) = < -\frac{\sqrt{3}}{2}, \frac{1}{2} >$ and $\nabla f(3, 3\sqrt{3}) = < 6, 6\sqrt{3} >$. The dot product of these two vectors is zero, so yes, indeed: The gradient vector *is* orthogonal (perpendicular) to the unit tangent vector. But the unit tangent vector has the same direction as the velocity vector and the tangent line, so we can also say the gradient vector is orthogonal to the level curve's velocity vector and tangent line.

Here is what we have observed so far: At the point generated when $t = \frac{\pi}{3}$, the gradient vector is orthogonal to the level curve's unit tangent vector, velocity vector, and tangent line, and the directional derivative as we move along the curve at this instant is 0. In fact, this is true not only for $t = \frac{\pi}{3}$, but for *all* values of t. In other words, as we move along the level curve, at *any* instant, the gradient vector will be orthogonal to the level curve's unit tangent vector, velocity vector, and tangent line, and the directional derivative as we move along the curve at this instant is 0.

Indeed, we can generalize further. This insight applies not just to this particular function, $f(x,y) = x^2 + y^2$, and to this particular level curve, $z = 36$. It applies just as well to *any* function and to *any* level curve.

In general, if we are given a level curve $z = c$ for a function $z = f(x,y)$, then:

- At any point on the level curve, the gradient vector ∇f will be orthogonal to the level curve's tangent line, to its velocity vector, and to its unit tangent vector. For brevity, we shall simply say ∇f is "orthogonal to the level curve."
- As we move along the curve, the directional derivative at any instant must be zero. In other words, at any point on the curve, if we move in the direction indicated by the velocity vector, then the directional derivative will be zero.

It makes sense that the directional derivative should be zero as we move along a level curve. After all, as we move through the x,y plane along the level curve $z = c$, the value of the function $f(x,y)$ does not change–it is fixed at the constant c. Since the directional derivative is the rate of change of the function as we move in a particular direction, and since a function with a constant value has a rate of change equal to zero, it's only natural to expect that the directional derivative would be zero as we move along the level curve.

Picture a series of level curves for $z = f(x,y)$, such as $z = 10$, $z = 20$, $z = 30$, $z = 40$. The graph of the function is a surface in x,y,z space, which we can picture as consisting of hills and valleys. The series of level curves in the x,y plane is like a topographic map, where each level curve indicates a specified "altitude." If we are on the surface and walking along a path corresponding to a given level curve on the map, then our altitude does not change. If we turn at a right angle to the path, we will be moving either *uphill* as steeply as possible or *downhill* as steeply as possible (depending on whether we turn in the direction of the gradient vector or in the opposite direction).

Based on this analogy, we say that ∇f points in the *direction of steepest ascent*, whereas $-\nabla f$ points in the *direction of steepest descent*.

Hyper-Surfaces, **Three-Dimensional Gradient Vectors**, **Level Surfaces**, **and Tangent Planes**:

All the concepts developed so far can be carried over to higher dimensions. Say we have a function $w = F(x,y,z)$, which has a three-dimensional domain and a one-dimensional range. The domain is x,y,z space or some subset thereof, and the range is the w axis or some subset thereof. We cannot actually draw the graph, since it would exist in "four-dimensional" space (i.e., x,y,z,w space), which does not physically exist. However, we can refer to the graph theoretically as a **hyper-surface**.

The set of all points in x,y,z space such that the function has a fixed value, c, is known as a **level surface**, which can be denoted as $F(x,y,z) = c$ or simply $w = c$.

Since the function F has three independent variables, it has three partial derivatives: $\frac{\partial F}{\partial x}$, $\frac{\partial F}{\partial y}$, and $\frac{\partial F}{\partial z}$. These could also be denoted $\frac{\partial w}{\partial x}$, $\frac{\partial w}{\partial y}$, and $\frac{\partial w}{\partial z}$, or as F_x, F_y, and F_z. The gradient vector is $\nabla F = < \frac{\partial F}{\partial x}, \frac{\partial F}{\partial y}, \frac{\partial F}{\partial z} >$.

As mentioned previously, at any given point, a surface may or may not have a tangent plane. Later we will explore the conditions under which we are guaranteed the existence of a tangent plane. For now, just assume the tangent plane exists at a given point. Then every line which is tangential to the surface at this point must lie in this plane (indeed, the tangent plane can be thought of as the union of all possible tangent lines at the given point).

If we are given a level surface $w = c$ for a function $w = F(x,y,z)$, then:

- At any point on the level surface, the gradient vector ∇F will be orthogonal to the level surface's tangent plane. For brevity, we shall simply say ∇F is "orthogonal to the level surface."
- As we move across the level surface along any path, when we pass through a given point, our tangent vector will lie in the surface's tangent plane at that point, and the directional derivative of $F(x,y,z)$ at that point, in the direction of the velocity vector, must be zero.

To find an equation for the tangent plane of the level surface at a specified point, (x_0,y_0,z_0), all we need is a normal vector for the plane. $\nabla F(x_0,y_0,z_0)$ can serve that role. Since $\nabla F(x_0,y_0,z_0) = < F_x(x_0,y_0,z_0), F_y(x_0,y_0,z_0), F_z(x_0,y_0,z_0) >$, the equation of the tangent plane is $F_x(x_0,y_0,z_0)(x - x_0) + F_y(x_0,y_0,z_0)(y - y_0) + F_z(x_0,y_0,z_0)(z - z_0) = 0$. This could also be written as $\nabla F(x_0,y_0,z_0) \cdot < x - x_0, y - y_0, z - z_0 > = 0$.

For example, suppose we have the function $w = F(x,y,z) = x^3 + y^4 + z^5$. At the point $(2,-1,1)$ in x,y,z space, the value of the function is $w = F(2,-1,1) = 10$. So the point $(2,-1,1,10)$ lies on the graph of F in x,y,z,w space (which is a hyper-surface).

Consider the level surface $w = 10$. The point $(2,-1,1)$ lies on this level surface. The equation of the level surface is $x^3 + y^4 + z^5 = 10$.

$F_x = 3x^2$, $F_y = 4y^3$, and $F_z = 5z^4$, so $\nabla F(x,y,z) = < 3x^2, 4y^3, 5z^4 >$. At the point $(2,-1,1)$, we get $\nabla F(2,-1,1) = < 12, -4, 5 >$.

The tangent plane to the surface at the point $(2,-1,1)$ is thus $12(x - 2) - 4(y + 1) + 5(z - 1) = 0$, or $12x - 4y + 5z = 33$.

Before we go on, let's touch on one other point...

Recall that any surface in x,y,z space passes the **Vertical Line Test** if any vertical line intersects the surface at no more than one point. Another way to phrase the test is this: Any specified values for x and y can produce only one corresponding value of z.

A level surface for a function $w = F(x,y,z)$ may or may not pass the Vertical Line Test.

In the above example, our level surface, $x^3 + y^4 + z^5 = 10$, did pass the Vertical Line Test: If we specify $x = x_1$ and $y = y_1$, then we get a unique value of z, namely, $z = \sqrt[5]{10 - (x_1)^3 - (y_1)^4}$.

On the other hand, the level surface $x^3 + y^4 + z^2 = 10$ does not pass the Vertical Line Test: If we specify $x = x_1$ and $y = y_1$, then we get two values of z, namely, $z = \pm\sqrt{10 - (x_1)^3 - (y_1)^4}$.

Tangent Planes for a Function with a Two-Dimensional Domain:

Let's go back to the more familiar, and easier to visualize, situation, where our function has a two-dimensional domain, and its graph is an ordinary surface in x,y,z space (a surface which passes the Vertical Line Test). Say our function is $z = f(x,y)$. Say we have a particular point in the domain, (x_0,y_0). Let $z_0 = f(x_0,y_0)$. So (x_0,y_0,z_0) is a point on the surface. How can we write an equation for its tangent plane at this point?

We capitalize on what we already know. Specifically, we invent a function with a three-dimensional domain, and then do what we learned to do in the previous section.

Rewrite the equation $z = f(x,y)$ as $f(x,y) - z = 0$. Now let $F(x,y,z) = f(x,y) - z$. This is a function with a three-dimensional domain (which is x,y,z space or some subset thereof), a one-dimensional range (the w axis or some subset thereof), and a four-dimensional graph, which is a hyper-surface in x,y,z,w space. We could refer to F as the "hyper-function" for f.

The point (x_0,y_0,z_0) lies in the domain of F, and $F(x_0,y_0,z_0) = f(x_0,y_0) - z_0 = z_0 - z_0 = 0$. So the point (x_0,y_0,z_0) lies on the level surface $F(x,y,z) = 0$. Likewise, every point on the graph of the function $f(x,y)$ lies on this same level surface. In fact, the graph of the function $z = f(x,y)$ *is* the level surface!!

$F_x = \frac{\partial}{\partial x}F(x,y,z) = \frac{\partial}{\partial x}(f(x,y) - z) = \frac{\partial}{\partial x}f(x,y) - \frac{\partial}{\partial x}z = f_x(x,y) - 0 = f_x(x,y)$
$F_y = \frac{\partial}{\partial y}F(x,y,z) = \frac{\partial}{\partial y}(f(x,y) - z) = \frac{\partial}{\partial y}f(x,y) - \frac{\partial}{\partial y}z = f_y(x,y) - 0 = f_y(x,y)$
$F_z = \frac{\partial}{\partial z}F(x,y,z) = \frac{\partial}{\partial z}(f(x,y) - z) = \frac{\partial}{\partial z}f(x,y) - \frac{\partial}{\partial z}z = 0 - 1 = -1.$
So $\nabla F = < F_x,F_y,F_z > = < f_x(x,y),f_y(x,y),-1 >$.

At the point (x_0,y_0,z_0), we get $\nabla F(x_0,y_0,z_0) = < f_x(x_0,y_0),f_y(x_0,y_0),-1 >$.

At the point (x_0,y_0,z_0), the tangent plane has normal vector $< f_x(x_0,y_0),f_y(x_0,y_0),-1 >$, so the equation of the tangent plane is $f_x(x_0,y_0)(x - x_0) + f_y(x_0,y_0)(y - y_0) - 1(z - z_0) = 0$. So $f_x(x_0,y_0)(x - x_0) + f_y(x_0,y_0)(y - y_0) - z + z_0 = 0$, or $z = f_x(x_0,y_0)(x - x_0) + f_y(x_0,y_0)(y - y_0) + z_0$. This could also be written as $z = \nabla f(x_0,y_0) \cdot < x - x_0, y - y_0 > + z_0$.

For instance, consider the function $z = f(x,y) = x^2 + y^2$, whose graph is a circular paraboloid. At the point $(x_0,y_0) = (1,8)$, $z_0 = f(1,8) = 65$, so the point $(1,8,65)$ lies on the circular paraboloid. This function has the gradient $\nabla f = < 2x, 2y >$, so $\nabla f(1,8) = < 2,16 >$. Thus, the equation of the tangent plane at the point $(1,8,65)$ is $z = 2(x-1) + 16(y-8) + 65$.