Various Topics Related to the Gradient Vector

Finding the Maximum and Miniumum Values of the Directional Derivative:

Say we have a function $z = f(x,y)$, which is differentiable and has a nonzero gradient vector at the point (x_0,y_0).

For any unit vector $\mathbf{u}$, the directional derivative of f at the point (x_0,y_0) in the direction of $\mathbf{u}$ is $D_{\mathbf{u}} f(x_0,y_0) = \mathbf{u} \cdot \nabla f(x_0,y_0)$.

Let us ask the question, what is the maximum possible value that $D_{\mathbf{u}} f(x_0,y_0)$ can have? And for which unit vector does it achieve this maximum value?

Let θ denote the angle between $\mathbf{u}$ and $\nabla f(x_0,y_0)$. 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 $\nabla f(x_0,y_0)$. If $\theta = \pi$, then $\mathbf{u}$ points in the opposite direction from $\nabla f(x_0,y_0)$. If $\theta = \frac{\pi}{2}$, then $\mathbf{u}$ is orthogonal (i.e., perpendicular) to $\nabla f(x_0,y_0)$.

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 $\theta \in (\frac{\pi}{2},\pi)$.

Thus, the directional derivative has its maximum value when $\theta = 0$, i.e.,when $\mathbf{u}$ points in the direction of $\nabla f(x_0,y_0)$, and this maximum value is $|\nabla f(x_0,y_0)|$. Furthermore, the directional derivative has its minimum value when $\theta = \pi$, i.e.,when $\mathbf{u}$ points in the direction of $-\nabla f(x_0,y_0)$, and this maximum value is $-|\nabla f(x_0,y_0)|$. The directional derivative has a value of zero when $\theta = \frac{\pi}{2}$, i.e., when $\mathbf{u}$ is orthogonal to $\nabla f(x_0,y_0)$. (If the unit vector in the direction of $\nabla f(x_0,y_0)$ is $< a,b >$, then there are two unit vectors orthogonal to $\nabla f(x_0,y_0)$, namely, $< -b,a >$ and $< b,-a >$.)

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.

For example, consider $z = f(x,y) = x^2 + y^2$. $\nabla f(x,y) = < 2x, 2y >$. At the point $(1,8)$, the gradient is $\nabla f(1,8) = < 2, 16 >$, whose magnitude is $\sqrt{260}$ or $2\sqrt{65}$. The unit vector in the direction of $\nabla f(1,8)$ is $\frac{1}{2\sqrt{65}} < 2, 16 > = < \frac{1}{\sqrt{65}}, \frac{8}{\sqrt{65}} >$, the unit vector in the opposite direction is $< \frac{-1}{\sqrt{65}}, \frac{-8}{\sqrt{65}} >$, and the orthogonal unit vectors 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 four unit vectors. The first result is $\frac{130}{\sqrt{65}}$, but if we rationalize the denominator, we get $\frac{130\sqrt{65}}{65} = 2\sqrt{65}$, which equals the magnitude of $\nabla f(1,8)$. The second result is $\frac{-130}{\sqrt{65}}$, and the third and fourth results are 0.

The Relationship Between Gradient Vectors and Level Curves:

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 $12x - 4y + 5z = 33$.