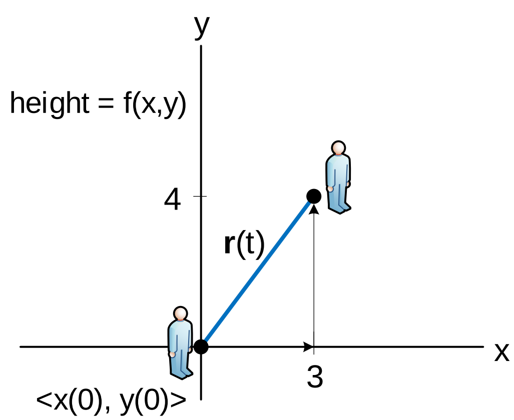
Multivariable Differentiation – The Directional Derivative

So far, we are only able to analyze the rate of change, i.e. the derivative, of multivariable functions with respect to one variable at a time. In other words, we can determine how a multivariable function changes when we change one variable of the other, but not if multiple variables are changing simultaneously. To accomplish this, we used *partial derivatives*. In this lesson we introduce what are referred to as *directional derivatives*, allowing us to determine how a function changes when multiple variables change simultaneously. We'll call on our knowledge of vectors for this task. In the process we will introduce one of the most important vectors in multivariable calculus called the *gradient vector*.

Directional Derivative

Rather than formally derive an expression for the directional derivative, we'll instead develop the expression heuristically. Suppose a person is walking on a terrain and their horizontal position is described with the vector function, $\boldsymbol{r}(t) = \langle x(t), y(t)\rangle$. Furthermore, assume that the height of the terrain is described by the function, $f(x, y)$. As we already know, the rate of change in height with respect to distance a person experiences when walking in the x or y direction ***only*** is given by $\partial f/\partial x$ and $\partial f/\partial y$, respectively.

The question we would like to answer is: "What is the rate of change in height the person experiences when they do not walk in a straight line?". One way to try and answer this question is the notion of a *weighted average.* For example, let's say the person walks, in some unit of time, three paces in the x direction followed by four paces in the y direction, as shown in the figure below.



The weighted average of the rates of change for each direction is given by

$$\begin{aligned} weighted\ average &= \frac{\frac{\partial f}{\partial x}(x\ distance) + \frac{\partial f}{\partial y}(y\ distance)}{total\ distance} \\ &= \frac{\frac{\partial f}{\partial x}(3) + \frac{\partial f}{\partial y}(4)}{\sqrt{3^2 + 4^2}} \\ &= \frac{\partial f}{\partial x}\left(\frac{3}{5}\right) + \frac{\partial f}{\partial y}\left(\frac{4}{5}\right) \end{aligned}$$

Notice that if we organize the weight in a vector, $\langle\frac{3}{5},\frac{4}{5}\rangle$, the result is a unit vector that points in the direction of motion. Organizing the partial derivatives in a similar fashion the weighted average can be represented using a dot product as follows.

$$weighted\ average = \langle\frac{\partial f}{\partial x},\frac{\partial f}{\partial y}\rangle \cdot \langle\frac{3}{5},\frac{4}{5}\rangle$$

As it turns out this is exactly how the directional derivative is defined! We use the notation, $D_{\boldsymbol{u}}f$, to represent the rate of change of a multivariable function, f, in the direction of the vector, $\boldsymbol{u}$. Therefore, for a two variable function, the directional derivative is given as

$$D_{\boldsymbol{u}}f = \langle\frac{\partial f}{\partial x},\frac{\partial f}{\partial y}\rangle \cdot \langle u_x, u_y\rangle$$

Where, $\boldsymbol{u} = \langle u_x, u_y\rangle$, is a unit vector that points in the direction of motion. However, it is the first vector that turns out to be one of the most important vectors in multivariable calculus. We refer to it as the *gradient vector*. The vector can be represented with an even more fundamental operator called the del operator, which is given the symbol of an upside-down Greek delta, ∇.

$$\nabla = \langle\frac{\partial}{\partial x_1},\dots,\frac{\partial}{\partial x_n}\rangle$$

The gradient vector is then written as

$$\nabla f = \langle\frac{\partial}{\partial x_1},\dots,\frac{\partial}{\partial x_n}\rangle f = \langle\frac{\partial f}{\partial x_1},\dots,\frac{\partial f}{\partial x_n}\rangle$$

Using this notation, the directional derivative can be expressed as

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

In this lesson we'll spend a significant amount of time looking at various properties of the gradient vector. However, before we do let's work out some specific examples using what we have learned so far with regard to the directional derivative.

The Directional Derivative
Let $f(x,y)$ be a function of two variables and let $\boldsymbol{u}$ denote a unit vector. Then the derivative of $f(x,y)$ in the direction of $\boldsymbol{u}$ is called the *directional derivative, $D_{\boldsymbol{u}}f$*. $$D_{\boldsymbol{u}}f = \nabla f \cdot \boldsymbol{u}$$ Where, $$\nabla f = \langle\frac{\partial f}{\partial x},\frac{\partial f}{\partial y}\rangle \quad and \quad \boldsymbol{u} = \langle u_x, u_y\rangle$$ The definition can be extended to three or more dimensions as follows where $$\nabla f = \langle\frac{\partial f}{\partial x_1},\dots,\frac{\partial f}{\partial x_n}\rangle \quad and \quad \boldsymbol{u} = \langle u_{x_1},\dots,u_{x_n}\rangle$$

Example 1: Calculate the directional derivative in the direction of $\boldsymbol{v}$ at the given point.

a. $f(x,y) = ln(x^2 + y^2)$ $\quad \boldsymbol{v} = \langle 3, -2 \rangle \quad P = (1,0)$

b. $f(x,y,z) = xe^{-yz}$ $\quad \boldsymbol{v} = \langle 1, 1, 1 \rangle \quad P = (1,2,0)$

Solution:

a. Since the direction vector needs to be of unit length, we write the directional derivative as

$$\begin{aligned} D_{\boldsymbol{u}}f &= \nabla f \cdot \frac{\boldsymbol{v}}{\|\boldsymbol{v}\|} \\ &= \frac{1}{\sqrt{3^2 + 2^2}} \langle \frac{\partial f}{\partial x}, \frac{\partial f}{\partial y} \rangle \cdot \langle 3, -2 \rangle \\ &= \frac{1}{\sqrt{13}} \langle \frac{2x}{x^2 + y^2}, \frac{2y}{x^2 + y^2} \rangle \cdot \langle 3, -2 \rangle \\ &= \frac{6x - 4y}{\sqrt{13}(x^2 + y^2)} \end{aligned}$$

Evaluating at the given point we have

$$D_{\boldsymbol{u}}f|_{(1,0)} = \frac{6 \cdot 1 - 4 \cdot 0}{\sqrt{13}(1^2 + 0^2)} = \frac{6}{\sqrt{13}}$$

b.

$$\begin{aligned} D_{\boldsymbol{u}}f &= \nabla f \cdot \frac{\boldsymbol{v}}{\|\boldsymbol{v}\|} \\ &= \frac{1}{\sqrt{1^2 + 1^2 + 1^2}} \langle \frac{\partial f}{\partial x}, \frac{\partial f}{\partial y}, \frac{\partial f}{\partial z} \rangle \cdot \langle 1, 1, 1 \rangle \\ &= \frac{1}{\sqrt{3}} \langle e^{-yz}, -zxe^{-yz}, yxe^{-yz} \rangle \cdot \langle 1, 1, 1 \rangle \\ &= \frac{e^{-yz}(1 - zx + yx)}{\sqrt{3}} \end{aligned}$$

Evaluating at the given point we have

$$D_{\boldsymbol{u}}f|_{(1,2,0)} = \frac{e^{-2\cdot 0}\big(1 - (0 \cdot 1) + (2 \cdot 1)\big)}{\sqrt{3}} = -\frac{1}{\sqrt{3}}$$

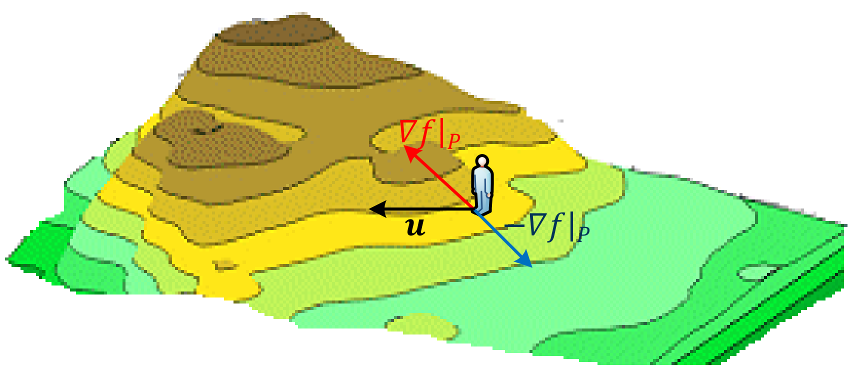
The Gradient Vector

As mentioned, the gradient vector is one of the most important vectors in multivariable calculus. As such the rest of this lesson focusing on two of its key properties. The first is the fact it points in the direction of greatest increase of a function. The second, and somewhat related property, is the fact that it is normal to the level curves of a surface. Before we start, we list some useful algebraic properties of the gradient vector.

Algebraic Properties of the Gradient Vector
If $f(x,y,z)$ and $g(x,y,z)$ are differentiable functions and c is a constant, then *i.* $\nabla(f+g) = \nabla f + \nabla g$ *ii.* $\nabla(cf) = c\nabla f$ *iii.* **Product Rule for Gradients***:* $\nabla(fg) = g\nabla f + f\nabla g$ iv. **Chain Rule for Gradients***:* If $F(t)$ is a differentiable function of one variable, then $$\nabla\Big(F\big(f(x,y,z)\big)\Big) = F'\big(f(x,y,z)\big)\nabla f$$

Direction of Greatest Increase

Imagine standing at a certain location on a mountain where the height can be described by the function, $h = f(x,y)$. You would like to get to the top of the mountain by walking the shortest distance. In other words, you need to determine in which direction to begin walking from your current location such that your height will increase the most.



The directional derivate gives you the rate of change of height based on the direction of travel.

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

Recall the dot product can be written as follows:

$$D_{\boldsymbol{u}}f = \nabla f \cdot \boldsymbol{u} = \|\nabla f\|\|\boldsymbol{u}\|\,cos(\theta) = \|\nabla f\|\,cos(\theta)$$

Where, we used the fact that the unit vector, $\boldsymbol{u}$, has a length of one.

The maximum value of the cosine function is 1.0 and it occurs $\theta = 0^\circ$. Therefore, we experience the greatest increase in height when $\boldsymbol{u}$ points in the same direction as ∇f. In other words, the gradient vector, ∇f, points in the direction of maximum increase of f!

Note also that the cosine is -1.0 for $\theta = 180^{\circ}$. Therefore, $-\nabla f$ points in the direction of maximum decrease of f. In both cases the magnitude of the maximum rate of increase/decrease is equal to $\|\nabla f\|$. Finally, even though the proof was based on a two dimensional function it applies to any number of dimensions.

Gradient Vector as the Direction of Maximum Increase
Let f be a differentiable function at a fixed point, P, with $\nabla f\|_P \neq 0$.   • ∇f points in the direction of the maximum rate of ***increase*** of f at P, and the maximum rate of ***increase*** is $\|\nabla f\|$.   • $-\nabla f$ points in the direction of the maximum rate of ***decrease*** of f at P, and the maximum rate of ***decrease*** is $\|\nabla f\|$.

Example 2: Find the unit vector that points in the maximum direction of increase at P. Then find the magnitude of the rate of increase.

a. $f(x, y) = cos(2x - y^2) \qquad P = (\pi/4, 0)$

b. $f(x, y, z) = 4xyz - y^2z^3 + 4z^3y \quad P = (2,3,1)$

Solution:

a. We know that the gradient vector points in the direction of maximum increase. Therefore,

$$\begin{aligned} \nabla f &= \langle \frac{\partial f}{\partial x}, \frac{\partial f}{\partial y} \rangle \\ &= \langle -2\, sin(2x - y^2), 2y\, sin(2x - y^2) \rangle \end{aligned}$$

This vector points in the direction of maximum increase for any point, (x, y), on the surface. At the point the direction is

$$\nabla f|_{(\pi/4,0)} = \langle -2\, sin(\pi/2 - 0^2), 2 \cdot 0\, sin(\pi/2 - 0^2) \rangle = \langle -2,0 \rangle$$

The unit vector that points in the same direction is given by

$$\widehat{\nabla f}\Big|_P = \frac{\nabla f|_P}{\|\nabla f|_P\|} = \frac{\langle -2,0 \rangle}{\sqrt{2^2}} = \langle -1,0 \rangle$$

Finally, the magnitude of the maximum rate of increase is

$$\|\nabla f|_P\| = 2$$

b. We perform the same steps from *a.*, but with three dimensions.

The gradient vector function is

$$\nabla f = \langle \frac{\partial f}{\partial x}, \frac{\partial f}{\partial y}, \frac{\partial f}{\partial z} \rangle$$
$$= \langle 4yz, 4xz - 2yz^3 + 4z^3, 4xy - 3y^2z^2 + 12z^2y \rangle$$

The gradient evaluated at the given point is

$$\nabla f|_{(2,3,1)} = \langle 12, 8 - 6 + 4, 24 - 27 + 36 \rangle = \langle 12, 6, 33 \rangle$$

The unit vector is

$$\widehat{\nabla f}\Big|_P = \frac{\nabla f|_P}{\|\nabla f|_P\|} = \frac{3\langle 4, 2, 11\rangle}{\sqrt{12^2 + 6^2 + 33^2}} = \frac{1}{\sqrt{141}}\langle 4, 2, 11\rangle$$

Finally, the maximum rate of increase is

$$\|\nabla f|_P\| = 3\sqrt{141}$$

Example 3: The temperature at a point (x, y, z) in a room is $T(x, y, z) = \frac{xz}{x^2+y^2}$. Find the direction in which the temperature decreases most rapidly at the point $(-3,4,1)$. Then find the rate of decrease of the temperature in that direction.

Solution: The vector that points in the direction of the maximum rate of decrease in temperature is the negative gradient vector, $-\nabla f$.

$$-\nabla f = -\langle \frac{\partial f}{\partial x}, \frac{\partial f}{\partial y}, \frac{\partial f}{\partial z} \rangle$$
$$= -\langle \frac{z(x^2 + y^2) - 2x^2z}{(x^2 + y^2)^2}, \frac{-2xyz}{(x^2 + y^2)^2}, \frac{x}{x^2 + y^2} \rangle$$

The direction vector evaluated at $(-3,4,1)$ is

$$-\nabla f|_{(-3,4,1)} = -\langle \frac{7}{625}, \frac{24}{625}, \frac{-3}{25} \rangle = \frac{1}{25}\langle -\frac{7}{25}, -\frac{24}{25}, 3 \rangle$$

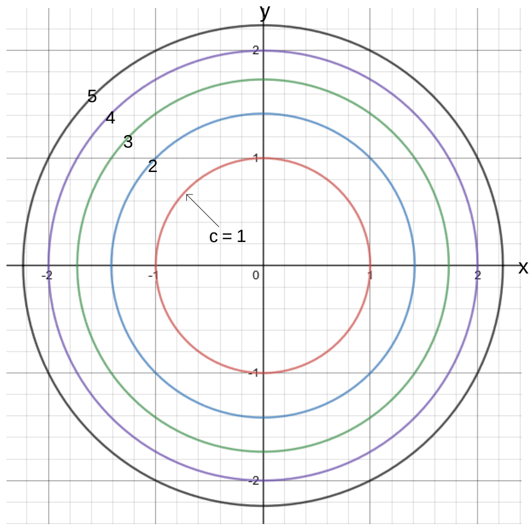
And the rate of decrease is

$$\|\nabla f|_P\| = \frac{1}{25}\sqrt{\left(\frac{7}{25}\right)^2 + \left(\frac{24}{25}\right)^2 + 9} = \frac{\sqrt{10}}{25} \cong 0.1265$$

Gradient as the Normal to Level Curves

In an earlier lesson we learned that level curves are projections of horizontal traces in x-y plane. In other words, level curves represent coordinates where the function, $f(x,y)$, is constant, i.e. $f(x,y) = c$. An example is shown below.

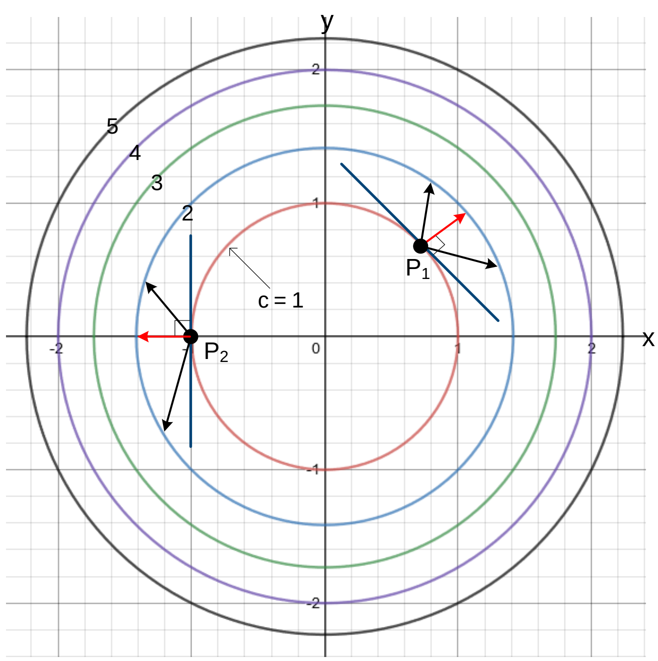
$$f(x,y) = x^2 + y^2 = c$$



Level surfaces are completely analogous for three variable functions. For example, the function from the example 3 represents the temperature at various locations, (x,y,z), in space. Therefore, the following represents *surfaces* where the temperature is constant.

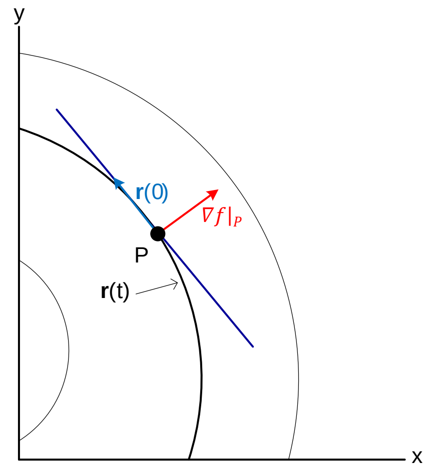
$$T(x,y,z) = \frac{xz}{x^2+y^2} = c$$

The property we wish to show in this section is that the gradient is normal to level curves/surfaces. Before we formally prove this let's argue it heuristically using level curves as shown in the figure below.



Starting at an arbitrary point, P_i, on one of the level curves we can move in any arbitrary direction. However, by definition, moving in the direction of gradient vector allows us to move from one level curve to the next in the shortest distance. The figure makes it clear that we achieve this if we travel in the direction of the red vector, indicating it is the gradient vector. Furthermore, the figure shows this vector is perpendicular to a tangent line to the curve at P_i. From these it follows that the gradient vector is perpendicular to the level curve. Although a heuristic argument, it is indeed quite convincing. Nonetheless, a formal proof is shown below.

We start by parameterizing a level curve with the vector function, $\boldsymbol{r}(t) = \langle x(t), y(t)\rangle$, such that $\boldsymbol{r}(0) = P$, and $\boldsymbol{r}'(t) \neq 0$.



The level curve can then be represented as

$$f(\boldsymbol{r}(t)) = c$$

Next, using the multivariable chain rule, we differentiate with respect to t,

$$\frac{\partial}{\partial t}\Big(f\big(\boldsymbol{r}(t)\big)\Big) = \frac{\partial}{\partial t}(c)$$

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

and rewrite the left-hand side as a dot product.

$$\langle\frac{\partial f}{\partial x}, \frac{\partial f}{\partial y}\rangle \cdot \langle\frac{dx}{dt}, \frac{dy}{dt}\rangle = 0$$

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

Finally, since $\boldsymbol{r}'(t)$ is tangent to the level curve for all t, and the dot product of the gradient vector and this tangent vector is zero, they must be orthogonal! The proof easily extends to level surfaces as well and is left as an exercise.

We can also easily derive the equation for the tangent line for level curves at a point, (x_0, y_0). Starting with the point slope formula of a line, we have

$$(y - y_0) = \left(\left.\frac{dy}{dx}\right|_{(x_0,y_0)}\right)(x - x_0)$$

The level curve that corresponds to the tangent line above can be written as

$$f(x, y) = c$$

Now, we use the implicit differentiation formula we derived in a previous lesson, to find the slope of the tangent line at (x_0, y_0).

$$\left.\frac{dy}{dx}\right|_{(x_0,y_0)} = -\frac{f_x(x_0, y_0)}{f_y(x_0, y_0)}$$

Substituting into the point-slope formula above, we can write a general form for the tangent line to the level curve at the point (x_0, y_0).

$$(y - y_0) = \left(-\frac{f_x(x_0, y_0)}{f_y(x_0, y_0)}\right)(x - x_0)$$

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

This technique can be repeated to find the *tangent plane* to level surfaces, which is given as

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

Gradient Vector as a Normal Vector		
Let P be a point on a level curve, $f(x, y) = c$, or on a level surface, $f(x, y, z) = c$, and assume that $\nabla f\|_P \neq 0$. Then $\nabla f\|_P$ is a vector that is normal to the tangent line/plane to the curve/surface at the point P. Moreover, the tangent line/plane to the curve/surface at the point P has the equation.		
Tangent Line	:	$f_x(x_0, y_0)(x - x_0) + f_y(x_0, y_0)(y - y_0) = 0$
Tangent Plane	:	$f_x(x_0, y_0, z_0)(x - x_0) + f_y(x_0, y_0, z_0)(y - y_0) + f_z(x_0, y_0, z_0)(z - z_0) = 0$

Example 4: Find a normal vector and tangent line to the level curve of $f(x, y)$, at $P = \left(3, \frac{1}{3}\right)$.

$$f(x, y) = ln(xy)$$

Solution: To find the level curve we need to find the 'level' at P.

$$f\left(3, \frac{1}{3}\right) = ln(1) = 0$$

Therefore, the level curve is given as

$$ln(xy) = 0$$

The normal vector is given by the gradient vector.

$$\nabla f|_{\left(3,\frac{1}{3}\right)} = \left\langle \frac{1}{x}, \frac{1}{y} \right\rangle\Big|_{\left(3,\frac{1}{3}\right)} = \left\langle \frac{1}{3}, 3 \right\rangle$$

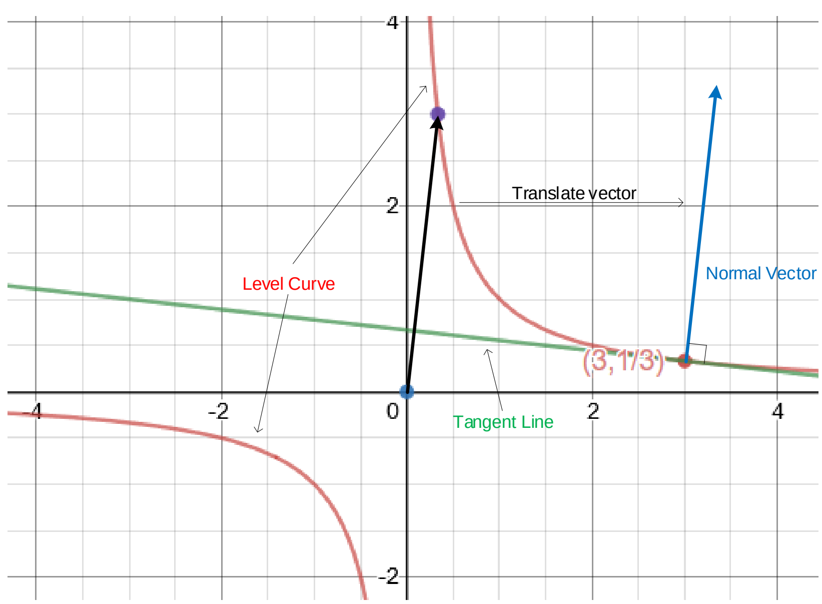
The tangent line is then found as

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

$$\frac{1}{3}(x - 3) + 3\left(y - \frac{1}{3}\right) = 0$$

$$y = \frac{2}{3} - \frac{1}{9}x$$

The figure below shows the level curve, tangent line, and normal vector.



Example 4: Find a normal vector and tangent plane for the level surface of the given function at $P = (2, 2, 2)$.

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

Solution: To find the level surface we again need to find the 'level' at the point indicated.

$$f(2, 2, 2) = 4 + 4 + 4 = 12$$

Therefore, the level surface, *a sphere*, is given as

$$x^2 + y^2 + z^2 = 12$$

The normal vector is given by the gradient vector.

$$\nabla f|_{(2,2,2)} = \langle 2x, 2y, 2z\rangle|_{(2,2,2)} = \langle 4, 4, 4\rangle$$

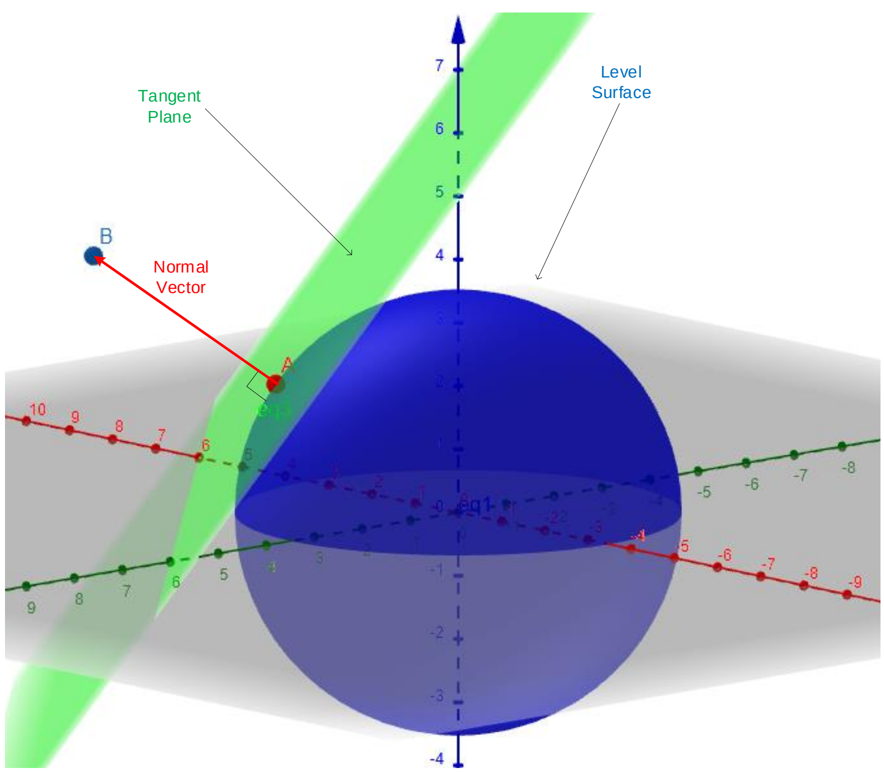
Lastly, the tangent plane is then found as

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

$$4(x - 2) + 4(y - 2) + 4(z - 2) = 0$$

$$4x + 4y + 4z = 24$$

The figure below again shows the level surface, tangent plane, and the normal vector.



Final Summary for Multivariable Differentiation – The Directional Derivative

The Directional Derivative
Let $f(x,y)$ be a function of two variables and let $\boldsymbol{u}$ denote a unit vector. Then the derivative of $f(x,y)$ in the direction of $\boldsymbol{u}$ is called the *directional derivative,* $D_{\boldsymbol{u}}f$. $$D_{\boldsymbol{u}}f = \nabla f \cdot \boldsymbol{u}$$ Where, $$\nabla f = \langle \frac{\partial f}{\partial x}, \frac{\partial f}{\partial y} \rangle \quad \textit{and} \quad \boldsymbol{u} = \langle u_x, u_y \rangle$$ The definition can be extended to three or more dimensions as follows where $$\nabla f = \langle \frac{\partial f}{\partial x_1}, \dots, \frac{\partial f}{\partial x_n} \rangle \quad \textit{and} \quad \boldsymbol{u} = \langle u_{x_1}, \dots, u_{x_n} \rangle$$
Algebraic Properties of the Gradient Vector
If $f(x,y,z)$ and $g(x,y,z)$ are differentiable functions and c is a constant, then   *i.* $\nabla(f+g) = \nabla f + \nabla g$ *ii.* $\nabla(cf) = c\nabla f$ *iii.* **Product Rule for Gradients***:* $\nabla(fg) = g\nabla f + f\nabla g$ iv. **Chain Rule for Gradients***:* If $F(t)$ is a differentiable function of one variable, then $$\nabla\Big(F\big(f(x,y,z)\big)\Big) = F'\big(f(x,y,z)\big)\nabla f$$
Gradient Vector as the Direction of Maximum Increase
Let f be a differentiable function at a fixed point, P, with $\nabla f\|_P \neq 0$.   • ∇f points in the direction of the maximum rate of ***increase*** of f at P, and the maximum rate of ***increase*** is $\lVert \nabla f \rVert$.   • $-\nabla f$ points in the direction of the maximum rate of ***decrease*** of f at P, and the maximum rate of ***decrease*** is $\lVert \nabla f \rVert$.
Gradient Vector as a Normal Vector
Let P be a point on a level curve, $f(x,y) = c$, or on a level surface, $f(x,y,z) = c$, and assume that $\nabla f\|_P \neq 0$. Then $\nabla f\|_P$ is a vector that is normal to the tangent line/plane to the curve/surface at the point P. Moreover, the tangent line/plane to the curve/surface at the point P has the equation   *Tangent Line* : $f_x(x_0,y_0)(x-x_0) + f_y(x_0,y_0)(y-y_0) = 0$ *Tangent Plane* : $f_x(x_0,y_0,z_0)(x-x_0) + f_y(x_0,y_0,z_0)(y-y_0) + f_z(x_0,y_0,z_0)(z-z_0) = 0$

By: ferrantetutoring