The partial derivatives of a multi-variable function, $\frac{\partial f}{\partial x}$ and $\frac{\partial f}{\partial y}$, tell us the rate of change of the function along $x$-direction and $y$-direction.

But what about other directions?

On the $x$-$y$ plane, each direction can be represented by a unit vector $\vec{u}$. We are going to define the directional derivative of a function $z=f(x,y)$ at $(x_0, y_0)$ in the direction $\vec{u}$:

On the $x$-$y$ plane, consider the line $l$ passing through $(x_0, y_0)$ and parallel to the unit vector $\vec{u}$. Passing through the line $l$, there is a unique vertical plane $\alpha$, and $\alpha$ intersects the graph of $z=f(x,y)$ along a curve $C$, so $C$ projects to $l$ on the $x$-$y$ plane.

If we start at $(x_0, y_0)$ and travel along $\vec{u}$ direction for a distance $h$, and arrive at $(x, y)$, then the vector with initial point $(x_0, y_0)$ and terminal point $(x, y)$ is $<x-x_0, y-y_0> = h\vec{u}$.

If we know $\vec{u} = <a, b>$ ($a^2 + b^2 = 1$), then $<x-x_0, y-y_0> = h\vec{u} = <ha, hb>$.

$\Rightarrow \begin{cases} x = x_0 + ha \\ y = y_0 + hb \end{cases} \Rightarrow f(x, y) = f(x_0 + ha, y_0 + hb)$.

So the rate of change of the function along $\vec{u}$ direction at $(x_0, y_0)$ is

$$Daf(x_0, y_0) = \lim_{h \to 0} \frac{f(x_0 + ha, y_0 + hb) - f(x_0, y_0)}{h}$$

We call it the directional derivative of $f$ at $(x_0, y_0)$ in the direction $\vec{u}$. 
There is a faster way to compute the directional derivative, if we know the following theorem:

**Theorem.** If \( f \) is a differentiable function of \( x \) and \( y \), then \( f \) has a directional derivative in the direction of any unit vector \( \vec{u} = \langle a, b \rangle \) and

\[
D_{\vec{u}} f(x, y) = \frac{\partial f}{\partial x}(x, y) \cdot a + \frac{\partial f}{\partial y}(x, y) \cdot b
\]

**Proof.** For a fixed \((x_0, y_0)\) in the domain of \( f \), we define the function

\[
g(h) = f(x_0 + ha, y_0 + hb)
\]

Then we get

\[
g'(0) = \lim_{h \to 0} \frac{g(h) - g(0)}{h} = \lim_{h \to 0} \frac{f(x_0 + ha, y_0 + hb) - f(x_0, y_0)}{h}
\]

\[
= D_{\vec{u}} f(x_0, y_0)
\]

But we can also write

\[
g(h) = f(x, y), \quad x = x_0 + ha, \quad y = y_0 + hb
\]

so the Chain Rule implies

\[
g'(h) = \frac{\partial f}{\partial x}(x, y) \cdot \frac{dx}{dh} + \frac{\partial f}{\partial y}(x, y) \cdot \frac{dy}{dh}
\]

\[
= a \frac{\partial f}{\partial x}(x, y) + b \frac{\partial f}{\partial y}(x, y)
\]

When \( h = 0 \), \( x = x_0 \), \( y = y_0 \). So

\[
g'(0) = a \frac{\partial f}{\partial x}(x_0, y_0) + b \frac{\partial f}{\partial y}(x_0, y_0)
\]

By the unit vector \( \vec{u} \) forms an angle \( \theta \) with the positive \( x \)-axis, then

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

If the unit vector \( \vec{u} \) is the directional derivative of \( f \) by

\[
D_{\vec{u}} f(x_0, y_0) = \frac{\partial f}{\partial x}(x_0, y_0) \cdot \cos \theta + \frac{\partial f}{\partial y}(x_0, y_0) \cdot \sin \theta
\]

**Example.** Find the directional derivative \( D_{\vec{u}} f(x, y) \) if \( f(x, y) = x^2 - 3xy + 4y^2 \)

and \( \vec{u} \) is the unit vector given by angle \( \theta = \frac{\pi}{6} \). What is \( D_{\vec{u}} f(1, 2) \)?

\[
D_{\vec{u}} f(x, y) = \frac{\partial f}{\partial x}(x, y) \cos \frac{\pi}{6} + \frac{\partial f}{\partial y}(x, y) \sin \frac{\pi}{6} = (2x - 3y) \cdot \frac{\sqrt{3}}{2} + (-3x + 8y) \cdot \frac{1}{2}
\]

\[
D_{\vec{u}} f(1, 2) = (3 \cdot 1^2 - 3 \cdot 2) \cdot \frac{\sqrt{3}}{2} + (-3 \cdot 1 + 8 \cdot 2) \cdot \frac{1}{2} = \frac{13 - 3\sqrt{3}}{2}
\]
We have seen that \( D_{\mathbf{u}} f(x, y) = \frac{\partial f}{\partial x}(x, y), \mathbf{u} + \frac{\partial f}{\partial y}(x, y), \mathbf{u} \) if \( \mathbf{u} = \langle a, b \rangle \) is a unit vector. We can rewrite the above equation as
\[
D_{\mathbf{u}} f(x, y) = \langle \frac{\partial f}{\partial x}(x, y), \frac{\partial f}{\partial y}(x, y) \rangle \cdot \langle a, b \rangle
\]
We now define the gradient of \( f \) to be the vector
\[
\nabla f(x, y) = \langle \frac{\partial f}{\partial x}(x, y), \frac{\partial f}{\partial y}(x, y) \rangle
\]
then \( D_{\mathbf{u}} f(x, y) = \nabla f(x, y) \cdot \mathbf{u} \)

Example. Find the gradient of the function \( f(x, y) = x^2y^3 - 4y \) at \((2, -1)\), and find the directional derivative in the direction of the vector \( \mathbf{v} = \langle 2, 5 \rangle \).

\[
\nabla f(x, y) = \langle 2xy^3, 3x^2y^2 - 4 \rangle
\]
so \( \nabla f(2, -1) = \langle -4, 8 \rangle \)

Note \( \mathbf{v} \) is not a unit vector, so we first compute the unit vector in the direction of \( \mathbf{v} \), which is
\[
\mathbf{v} = \frac{2}{\sqrt{29}}, \frac{5}{\sqrt{29}}
\]

So \( D_{\mathbf{v}} f(2, -1) = \nabla f(2, -1) \cdot \mathbf{v} = \langle -4, 8 \rangle \cdot \langle \frac{2}{\sqrt{29}}, \frac{5}{\sqrt{29}} \rangle = \frac{32}{\sqrt{29}} \)

We can extend the above constructions of directional derivatives and gradient vectors to functions of \( n \) variables (in particular, \( 3 \) variables)

The gradient gives a solution to the maximization of directional derivatives:

**Theorem.** If \( f \) is a differentiable function of \( n \) variables, then the maximum value of \( D_{\mathbf{u}} f(x_1, \ldots, x_n) \) is \( |\nabla f(x_1, \ldots, x_n)| \), and it is achieved when \( \mathbf{u} = \frac{\nabla f(x_1, \ldots, x_n)}{|\nabla f(x_1, \ldots, x_n)|} \)
Proof. \[ \nabla f \cdot \mathbf{u} = |\nabla f| \cdot |\mathbf{u}| \cdot \cos \theta = |\nabla f| \cdot \cos \theta \]
so \( \nabla f \cdot \mathbf{u} \) obtains maximum when \( \theta = 0 \).

Example. \( f(x,y) = xe^y \). In what direction does \( f \) have the maximum rate of change at \((2,0)\)? What's the maximum rate of change?

\[ \nabla f(x,y) = \langle e^y, xe^y \rangle, \text{ so } \nabla f(2,0) = \langle 1, 2 \rangle \]

The maximum rate of change is along \( \nabla f(2,0) = \langle 1, 2 \rangle \) direction.
The maximum rate of change is \( |\nabla f(2,0)| = \sqrt{5} \).
The gradient has another important application, which is related to level sets:

Theorem. If \( F(x_1, \ldots, x_n) = C \) is a level set, then \( \nabla F \) is a normal vector for this level set in the space \( \mathbb{R}^n \). In particular, when \( n = 2 \), \( \nabla f(x,y) \) is perpendicular to the tangent line of \( f(x,y) = C \) at \((x,y)\); when \( n = 3 \), \( \nabla f(x,y,z) \) is perpendicular to the tangent plane of \( f(x,y,z) = C \) at \((x,y,z)\).

Example. Find the equation of the tangent plane at \((-2,1,-3)\) to the ellipsoid \( \frac{x^2}{4} + y^2 + \frac{z^2}{9} = 3 \).

Let \( F(x,y,z) = \frac{x^2}{4} + y^2 + \frac{z^2}{9} \), then \( \nabla F(x,y,z) = \langle \frac{x}{2}, 2y, \frac{2z}{9} \rangle \)

so \( \nabla F(-2,1,-3) = \langle -1, 2, -\frac{2}{3} \rangle \)

so the tangent plane has equation

\[-1 \cdot (x+2) + 2 \cdot (y-1) - \frac{2}{3} \cdot (z+3) = 0.\]

\[\Rightarrow 3x - 6y + 2z + 18 = 0\]