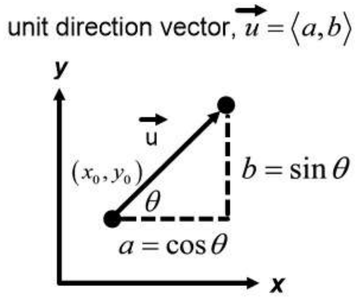
Derivations/Proofs for Directional Derivative and Gradient Vector

Definition of Directional Derivative

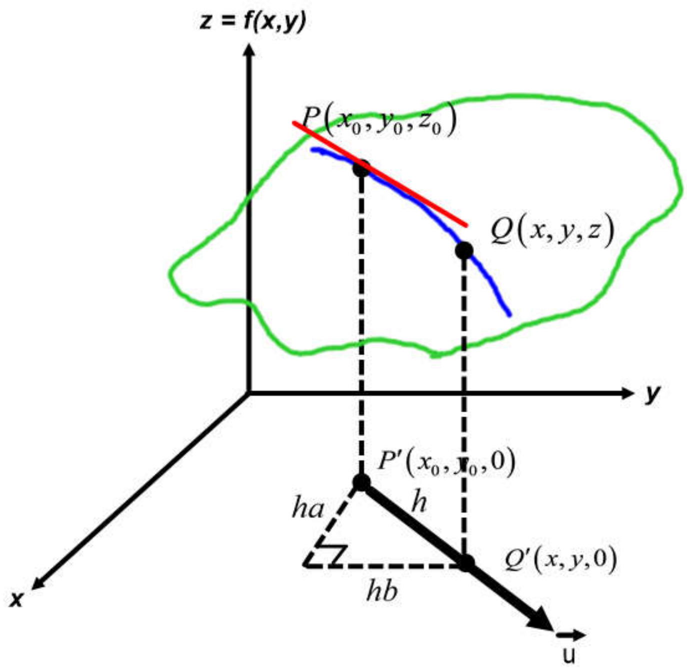
Partial derivatives in the direction of the x and y axes are defined by the following limits:

$$f_x(x_0, y_0) = \lim_{h\to 0} \frac{f(x_0+h, y_0) - f(x_0, y_0)}{h}$$
$$f_y(x_0, y_0) = \lim_{h\to 0} \frac{f(x_0, y_0+h) - f(x_0, y_0)}{h}$$

If we want to find the derivative in any direction, we can first specify that direction by using a unit direction vector in the domain:



Then, considering a small motion in this direction in 3D…



…we can imagine starting at an initial point, $P(x_0, y_0, z_0)$ and moving in the direction of the unit direction vector a small distance *h* in the domain. On the surface, this brings us to point *Q(x,y,z)*. The derivative in this direction is then the change in z (height) on the surface divided by this distance *h*.

Vector $\overrightarrow{P'Q'} = h\vec{u} = \langle ha, hb \rangle$ is parallel to $\vec{u}$ so there is some scalar *h* that scales *u* back to this small change in the direction. This means that:

$$\begin{aligned} x - x_0 &= ha \\ y - y_0 &= hb \end{aligned} \quad \text{so} \quad \begin{aligned} x &= x_0 + ha \\ y &= y_0 + hb \end{aligned}$$

The change in *z* over *h* is then…

$$\frac{\Delta z}{h}=\frac{z-z_0}{h}=\frac{f\left(x_0+ha,\, y_0+hb\right)-f\left(x_0,\, y_0\right)}{h}$$

…and we can define the directional derivative as:

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

If the unit vector is in the direction of the x or y axis, this expression becomes the derivatives f_x and f_y so the regular partial derivatives are just special cases of this more general directional derivative.

Directional Derivative can be computed from the partial derivatives

Define a function *g* of a variable representing the small change in the direction (in the domain), *h*:

$$g\left(h\right)=f\left(x_0+ha,\, y_0+hb\right)$$

Then by the usual definition of a derivative near *h=0*:

$$\begin{aligned}g'\left(0\right)&=\lim_{h\to 0}\frac{g\left(h\right)-h\left(0\right)}{h}=\lim_{h\to 0}\frac{f\left(x_0+ha,\, y_0+hb\right)-f\left(x_0,\, y_0\right)}{h}\\&=D_u f\left(x_0,y_0\right)\end{aligned}$$

But at the same time, we could also write the function *g(h)* as a function of *x* and *y*: $g\left(h\right)=f\left(x,y\right)$
and by the Chain Rule:

$$\begin{aligned}g'\left(h\right)&=\frac{\partial f}{\partial x}\frac{dx}{dh}+\frac{\partial f}{\partial y}\frac{dy}{dh}\\&=f_x\left(x,y\right)a+f_y\left(x,y\right)b\end{aligned}$$

When $h=0,\ x=x_0\ and\ y=y_0$, therefore: $g'\left(0\right)=f_x\left(x_0,y_0\right)a+f_y\left(x_0,y_0\right)b$
But also $g'\left(0\right)=D_u f\left(x_0,y_0\right)$ so by the Transitive Property:

$$D_u f\left(x_0,y_0\right)=f_x\left(x_0,y_0\right)a+f_y\left(x_0,y_0\right)b$$

$$or$$

$$D_u f\left(x_0,y_0\right)=f_x\left(x_0,y_0\right)\cos\theta+f_y\left(x_0,y_0\right)\sin\theta$$

Definition of the Gradient Vector

The result from the last section can be written as the dot product of two vectors:

$$D_u f(x,y) = f_x(x,y)a + f_y(x,y)b$$
$$D_u f(x,y) = \langle f_x(x,y),\ f_y(x,y)\rangle \cdot \langle a,\ b\rangle$$
$$D_u f(x,y) = \langle f_x(x,y),\ f_y(x,y)\rangle \cdot \vec{u}$$

The first vector is very useful, and is called the **Gradient Vector:**

$$\nabla f(x,y) = \langle f_x(x,y),\ f_y(x,y)\rangle$$

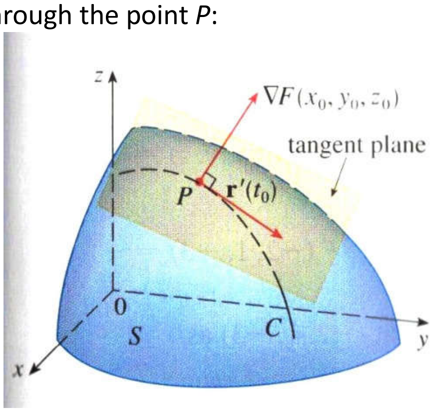
Directional derivative defined using the gradient vector

So because of the way the gradient is defined:

$$D_u f(x,y) = \langle f_x(x,y),\ f_y(x,y)\rangle \cdot \vec{u}$$
$$D_u f(x,y) = \nabla f(x,y) \cdot \vec{u}$$

Why gradient is perpendicular to tangent plane (tangent to a level surface)

Suppose *S* is a surface with equation *F(x,y,z)=k* and let $P_0(x_0, y_0, z_0)$ be a point on *S*. Let *C* be any curve that lies on surface *S* and passes through the point *P*:



Curve *C* can be described using a vector function, $\vec{r}(t) = \langle x(t), y(t), z(t)\rangle$. Define t_0 to be the parameter value along the curve at the point *P*, so $\vec{r}_0(t) = \langle x_0, y_0, z_0\rangle$. This point must satisfy the equation for the level surface,

$$F(x(t), y(t), z(t)) = k$$

If we differentiate this equation with respect to the parameter and use the Chain Rule:

$$\frac{d}{dt}\left[F(x(t), y(t), z(t))\right] = \frac{d}{dt}[k]$$
$$\frac{\partial F}{\partial x}\frac{\partial x}{\partial t} + \frac{\partial F}{\partial y}\frac{\partial y}{\partial t} + \frac{\partial F}{\partial z}\frac{\partial z}{\partial t} = 0$$

But $\nabla F = \langle F_x, F_y, F_z\rangle$ and $\vec{r}'(t) = \langle x'(t), y'(t), z'(t)\rangle$ so we can rewrite the left side:

$$\nabla F \cdot \vec{r}'(t) = 0$$

At point *P*: $\nabla F(x_0, y_0, z_0) \cdot \vec{r}'(t_0) = 0$ which means that the gradient vector at this point is perpendicular to the tangent vector $\vec{r}'(t_0)$ along any curve *C* on *S* that passes through *P*, and means the gradient is perpendicular to what we can call the **tangent plane** at this point to the surface.

Equation of the tangent plane to a surface

This also means that the gradient is a normal vector to the tangent plane, and because planes are defined using a normal vector and a point…

$$\vec{n} = \langle a, b, c \rangle = \langle F_x, F_y, F_z \rangle \qquad \vec{r_0} = \langle x_0, y_0, z_0 \rangle$$

…the equation for the tangent plane is given by…

$$ax + by + cz = \vec{n} \bullet \vec{r_0}$$
$$F_x x + F_y y + F_z z = \langle F_x, F_y, F_z \rangle \bullet \langle x_0, y_0, z_0 \rangle$$
$$F_x x + F_y y + F_z z = F_x x_0 + F_y y_0 + F_z z_0$$

$$F_x (x - x_0) + F_y (y - y_0) + F_z (z - z_0) = 0$$

…where the partial derivatives are evaluated at the point.