Lecture 18: 14.6: Directional Derivatives and the Gradient Vector.
Recall that the partial derivatives

$$f_x(x_0, y_0) = \lim_{h\to 0} \frac{f(x_0+h, y_0) - f(x_0, y_0)}{h} \tag{14.6.1}$$

$$f_y(x_0, y_0) = \lim_{h\to 0} \frac{f(x_0, y_0+h) - f(x_0, y_0)}{h} \tag{14.6.2}$$

represents the rate of change of the function f in the x and y directions.

Suppose that we now want to measure the rate of change in the direction of an arbitrary unit vector $\mathbf{u} = \langle a, b\rangle$. To do this is it is convenient to introduce the position vectors for the point $\mathbf{r}_0 = \langle x_0, y_0\rangle$ and for the point $\mathbf{r}(h) = \langle x(h), y(h)\rangle$ obtained by moving a distance h from $\mathbf{r}_0$ in the direction of the unit vector $\mathbf{u}$:

$$\mathbf{r}(h) = \mathbf{r}_0 + h\mathbf{u}, \tag{14.6.3}$$

The rate of change of f in the direction of the unit vector $\mathbf{u}$ is now given by the derivative of $f(x, y)$ along the curve $\langle x, y\rangle = \mathbf{r}(h)$;

$$\frac{d}{dh} f(\mathbf{r}(h))\Big|_{h=0}, \tag{14.6.4}$$

where $f(\mathbf{r}(h)) = f(\mathbf{r}_0 + h\mathbf{u}) = f(x_0 + ha, y_0 + hb)$. We therefore define the **directional derivative** of f at (x_0, y_0) in the direction of unit vector $\mathbf{u} = \langle a, b\rangle$ by

$$D_{\mathbf{u}} f(x_0, y_0) = \lim_{h\to 0} \frac{f(x_0 + ha, y_0 + hb) - f(x_0, y_0)}{h} \tag{14.6.5}$$

If $\mathbf{u}$ is $\mathbf{i} = \langle 1, 0\rangle$ or $\mathbf{j} = \langle 0, 1\rangle$ it reduces to partial derivatives in the x or y directions:

$$D_{\mathbf{i}} f(x_0, y_0) = f_x(x_0, y_0), \qquad D_{\mathbf{j}} f(x_0, y_0) = f_x(x_0, y_0) \tag{14.6.6}$$

Furthermore, by the chain rule we have since $\mathbf{r}(h) = \langle x(h), y(h)\rangle = \langle x_0 + ha, y_0 + hb\rangle$

$$\frac{d}{dh} f(x(h), y(h)) = \frac{\partial f}{\partial x}\frac{dx}{dh} + \frac{\partial f}{\partial y}\frac{dy}{dh}, \tag{14.6.7}$$

so we conclude that

$$D_{\mathbf{u}} f(x, y) = f_x(x, y)\, a + f_y(x, y)\, b, \qquad \text{if} \qquad \mathbf{u} = \langle a, b\rangle \tag{14.6.8}$$

Note that $\mathbf{u}$ has to be a unit vector in order that $D_{\mathbf{u}} f$ give the rate of change.
Ex. Find the directional derivative of $f(x, y) = x^2 + y^2$ in the direction of $\langle 1, 1\rangle$ at the point $(1, 2)$. **Sol.** A unit vector in the direction is $\mathbf{u} = \langle 1/\sqrt{2}, 1/\sqrt{2}\rangle$ and $D_{\mathbf{u}} f(1, 2) = f_x(1, 2)/\sqrt{2} + f_y(1, 2)/\sqrt{2} = 2/\sqrt{2} + 4\sqrt{2} = 3\sqrt{2}$.

We notice that the expression in (14.6.8) can be written as a dot product:

$$D_{\mathbf{u}} f(x, y) = \langle f_x(x, y), f_y(x, y)\rangle \cdot \langle a, b\rangle \tag{14.6.9}$$

Because of this we introduce a special notation for the so called **gradient** of f:

$$\nabla f(x, y) = \langle f_x(x, y), f_y(x, y)\rangle \tag{14.6.10}$$

also denoted **grad**f. With this notation we hence have

(14.6.11) $$D_{\mathbf{u}}f(x,y) = \nabla f(x,y)\cdot \mathbf{u}$$

This expresses the directional derivative in the direction $\mathbf{u}$ as the scalar projection of the gradient onto $\mathbf{u}$. Using this identity we conclude that the maximum value of $D_{\mathbf{u}}f(x,y)$ is $|\nabla f(x,y)|$ and it occurs when $\mathbf{u}$ is in the direction of $\nabla f(x,y)$. Furthermore $D_{\mathbf{u}}f(x,y)=0$ when $\mathbf{u}$ is perpendicular to $\nabla f(x,y)$. If $\mathbf{r}(t)=\langle x(t),y(t)\rangle$ is a level curve to $f(x,y)$, then $f(\mathbf{r}(t))$ is constant and by the chain rule

$$\frac{d}{dt}f(\mathbf{r}(t)) = \frac{d}{dt}f(x(t),y(t)) = f_x(x,y)x'(t)+f_y(x,y)y'(t) = \nabla f(\mathbf{r}(t)))\cdot\mathbf{r}'(t)=0$$

i.e. the gradient is perpendicular to the tangent of the level curves of f.
You reach the top of a hill fastest by going in the direction of the gradient.

Ex. Find the gradient of $f(x,y)=x^2+y^2$ at the point $(1,2)$ and the directional derivative in the direction $\mathbf{u}=\langle 1/\sqrt{2},1/\sqrt{2}\rangle$, at the point $(1,2)$.

Sol. $\nabla f(x,y) = \langle f_x(x,y), f_y(x,y)\rangle = \langle 2x,2y\rangle$. Hence $\nabla f(1,2)=\langle 2,4\rangle$ and $D_{\mathbf{u}}f(1,2)=\nabla f(1,2)\cdot\mathbf{u} = \langle 2,4\rangle\cdot\langle 1/\sqrt{2},1/\sqrt{2}\rangle = 2/\sqrt{2}+4/\sqrt{2}=3\sqrt{2}$.

Ex. In which direction does $f(x,y)=x^2+y^2$ increase the most at the point $(1,2)$ and what is the maximum rate of increase? In which direction is f constant?

Sol. $\nabla f(1,2)=\langle 2,4\rangle$ so the increase is largest in the direction of the gradient, i.e. in the direction of the unit vector $\mathbf{u}=\langle 2,4\rangle/\sqrt{2^2+4^2}=\langle 1/\sqrt{5},2/\sqrt{5}\rangle$, where $D_{\mathbf{u}}=\langle 2,4\rangle\cdot\langle 1/\sqrt{5},2/\sqrt{5}\rangle = 2\sqrt{5}$. The rate of change vanishes in the direction perpendicular to the gradient $\langle 2,4\rangle$, i.e. in the direction of $\langle 4,-2\rangle$.

Similarly, we now for a function of three variables $f(x,y,z)$ define the directional derivative in the direction of a unit vector $\mathbf{u}=\langle a,b,c\rangle$ at the point (x_0,y_0,z_0) by

(14.6.12) $$D_{\mathbf{u}}f(x_0,y_0,z_0) = \lim_{h\to 0}\frac{f(x_0+ha,y_0+hb,z_0+hc)-f(x_0,y_0,z_0)}{h}.$$

As before

(14.6.13) $$D_{\mathbf{u}}f(x,y,z) = \frac{d}{dh}f(\mathbf{r}(h))\Big|_{h=0}, \qquad \text{if} \quad \mathbf{r}(h)=\mathbf{r}_0+h\mathbf{u}$$

where $\mathbf{r}_0=\langle x_0,y_0,z_0\rangle$. We define the gradient by

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

and as before we have the identity:

(14.6.15) $$D_{\mathbf{u}}f(x,y,y) = \nabla f(x,y,z)\cdot\mathbf{u}.$$

Hence the maximum of $D_{\mathbf{u}}f(x,y,z)$ is $|\nabla f(x,y,z)|$ and it occurs when $\mathbf{u}$ is in the direction of $\nabla f(x,y,z)$ and $D_{\mathbf{u}}f(x,y,z)=0$ when $\mathbf{u}$ is perpendicular to $\nabla f(x,y,z)$.

Let S be the level surface $F(x,y,z)=k$. Suppose that $\mathbf{r}(t)=\langle x(t),y(t),z(t)\rangle$ is a curve on the level surface S. Then since $F(x(t),y(t),z(t))=k$ so

$$\frac{d}{dt}F(\mathbf{r}(t)) = \nabla f(\mathbf{r}(t)))\cdot\mathbf{r}'(t)=0$$

Since this holds for any curve on the surface it follows that the gradient is perpendicular to the tangent vector of any such curve. We therefore define the **tangent plane** to the surface $F(x,y,z)=k$ at the point (x_0,y_0,z_0) to be the plane through (x_0,y_0,z_0) with normal vector $\nabla F(x_0,y_0,z_0)$. The equation of the tangent plane is

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

Ex. Find the tangent plane to $F(x,y,z)=x^2+y^2+z^2=5$ at the point $(1,2,0)$.