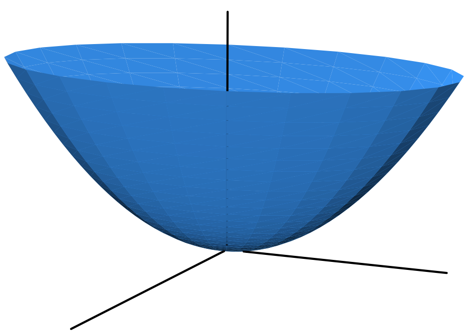
Directional derivatives and the gradient

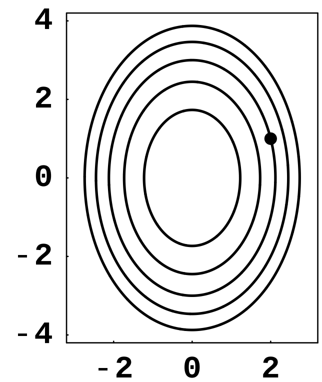
Today we will discuss directional derivatives and related theoretical consequences.

Example. Last class we started to discuss the directional derivative for the function

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

at the point $(2, 1)$ in the direction $\mathbf{u} = \frac{1}{\sqrt{2}}(\mathbf{i} + \mathbf{j})$.





We use the vector $\mathbf{u}$ as a direction vector for a line through the point $(2, 1)$. In parametric form, the line is

$$\begin{cases} x(h) = 2 + \dfrac{h}{\sqrt{2}} \\ y(h) = 1 + \dfrac{h}{\sqrt{2}}. \end{cases}$$

Then we compute the directional derivative of $f(x, y)$ at $(2, 1)$ in the $\mathbf{u}$-direction by calculating the limit

$$\begin{aligned} \lim_{h \to 0} \frac{f(x(h), y(h)) - f(2, 1)}{h} &= \lim_{h \to 0} \frac{\left[2(2 + \frac{h}{\sqrt{2}})^2 + (1 + \frac{h}{\sqrt{2}})^2\right] - 9}{h} \\ &= \lim_{h \to 0} \frac{2(4 + 2\sqrt{2}\,h + \frac{h^2}{2}) + (1 + \sqrt{2}\,h + \frac{h^2}{2}) - 9}{h} \\ &= \lim_{h \to 0} \frac{5\sqrt{2}\,h + \frac{3}{2}h^2}{h} = 5\sqrt{2} \approx 7.1. \end{aligned}$$

On the web site there are links to two animations that illustrate the concept of a directional derivative.

Definition of a Directional Derivative. We start with the two-variable case. Define the "directional derivative of $f(x, y)$ at the point (a, b) in the $\mathbf{u}$-direction" by parametrizing the line through (a, b) using the direction vector $\mathbf{u}$. In other words, if $\mathbf{u} = u_1\mathbf{i} + u_2\mathbf{j}$, then the line is written in vector form as

$$\mathbf{L}(h) = (a\mathbf{i} + b\mathbf{j}) + h\mathbf{u}$$

or in parametric form as $(x, y) = (a + u_1 h, b + u_2 h)$.

Then we compute

$$D_{\mathbf{u}} f(a, b) = \lim_{h \to 0} \frac{f(x, y) - f(a, b)}{h}.$$

Using vector notation with $\mathbf{P} = a\mathbf{i} + b\mathbf{j}$, the same limit is written as

$$D_{\mathbf{u}} f(\mathbf{P}) = \lim_{h \to 0} \frac{f(\mathbf{P} + h\mathbf{u}) - f(\mathbf{P})}{h}.$$

This vector notation generalizes nicely to functions of three variables or, in fact, to any number of variables.

Computing directional derivatives. A directional derivative for $f(x, y)$ at the point (a, b) can be computed by applying the Chain Rule to the composition $f(\mathbf{L}(h))$. Note that the vector $\mathbf{u} = u_1\,\mathbf{i} + u_2\,\mathbf{j}$ is used only to indicate the direction, and consequently, it is *always a unit vector*. In other words, $u_1^2 + u_2^2 = 1$.

Theorem. $D_{\mathbf{u}} f(a, b) = [\nabla f(a, b)] \cdot \mathbf{u}$.

Example. Calculate the directional derivative of $f(x, y) = e^x \sin y$ at the point $(\ln 2, \pi/6)$ in the direction of $2\mathbf{i} + \mathbf{j}$.

This theorem tells us know how a function changes in any given direction, and in particular, it indicates directions of most rapid increase or decrease for the function. Since $\mathbf{u}$ is a unit vector,

$$\begin{aligned} D_{\mathbf{u}} f(a,b) &= (\nabla f(a,b)) \cdot \mathbf{u} \\ &= |\nabla f(a,b)|\,|\mathbf{u}|\,\cos\theta \\ &= |\nabla f(a,b)|\,\cos\theta, \end{aligned}$$

where θ is the angle between $\mathbf{u}$ and the gradient vector $\nabla f(a,b)$.

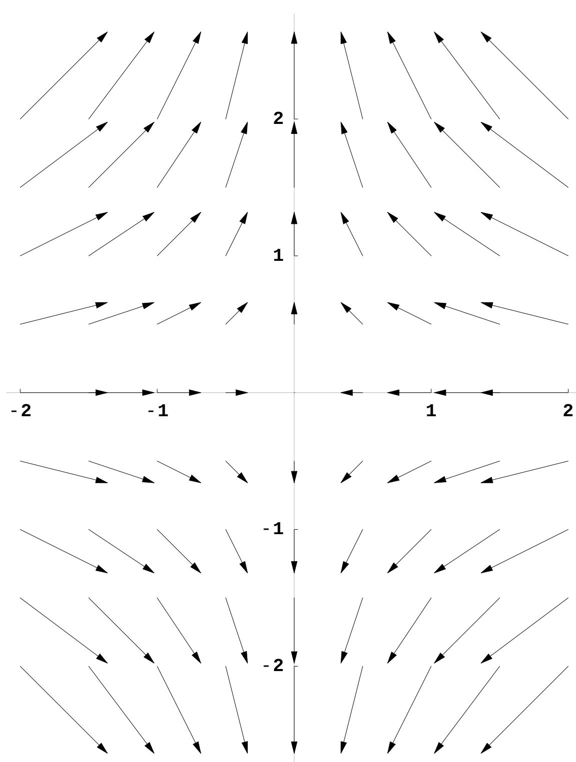
For what values of θ is this number largest? smallest? zero?

Theorem. The function $f(x,y)$ increases most rapidly in the direction of the gradient. The function is "constant" in directions perpendicular to the gradient.

Example. The gradient of $f(x,y) = \frac{1}{4}(y^2 - x^2)$ is

$$\nabla f(x,y) = \tfrac{1}{2}(-x\mathbf{i} + y\mathbf{j}).$$

Here is the gradient *vector field* that it generates.



Another important theoretical application of the Chain Rule is the fact that the gradient vector is always perpendicular to its corresponding level set.

Theorem. The gradient vector $\nabla f(a,b)$ is perpendicular to the level set of level $f(a,b)$.

Example. Once again consider $f(x,y) = \frac{1}{4}(y^2 - x^2)$. Its level sets are hyperbolas that are perpendicular to the gradient vector field.