To understand single-variable functions better, higher-ordered derivatives are used (as you have seen before). The same is true for multi-variable functions. It is particularly important in dealing with optimization problems.

What are higher-order derivatives?

Let $f(x,y) = x^2 + x^2 y + y^3$. Then

$$f_x(x,y) = 2x + 2xy \quad \text{and} \quad f_y(x,y) = x^2 + 3y^2.$$

Now we can talk about partial derivatives of $f_x$ and $f_y$ in terms of $x$ and $y$:

- Partial derivative of $f_x$ with respect to $x = f_{xx}(x,y) = 2 + 2y$
- Partial derivative of $f_x$ with respect to $y = f_{xy}(x,y) = 2x$
- Partial derivative of $f_y$ with respect to $x = f_{yx}(x,y) = 2 - x$
- Partial derivative of $f_y$ with respect to $y = f_{yy}(x,y) = 6y$

The 2nd partial derivatives, (also called iterated partial derivatives) are written:

$$f_{xx} = \frac{\partial}{\partial x} \left( \frac{\partial f}{\partial x} \right) = \frac{\partial^2 f}{\partial x^2}, \quad f_{xy} = \frac{\partial}{\partial y} \left( \frac{\partial f}{\partial x} \right) = \frac{\partial^2 f}{\partial y \partial x},$$

$$f_{yx} = \frac{\partial}{\partial x} \left( \frac{\partial f}{\partial y} \right) = \frac{\partial^2 f}{\partial x \partial y}, \quad f_{yy} = \frac{\partial}{\partial y} \left( \frac{\partial f}{\partial y} \right) = \frac{\partial^2 f}{\partial y^2}.$$
Lecture 20: Higher order partial derivatives

Monday, November 7, 2016 8:27 AM

Ex. Find 2nd-order partial derivatives of \( f(x,y) = \log(x^2+y) \).

**Solution.** \( f_x = \frac{2x}{x^2+y} \), \( f_y = \frac{1}{x^2+y} \).

So

\[
 f_{xx} = \frac{2(x^2+y) - (2x)(2x)}{(x^2+y)^2} = \frac{2y - 2x^2}{(x^2+y)^2}.
\]

\[
 f_{xy} = \frac{-2x}{(x^2+y)^2}, \quad f_{yx} = \frac{-2x}{(x^2+y)^2}, \quad f_{yy} = \frac{-1}{(x^2+y)^2}.
\]

In the above example, you see \( f_{xy} = f_{yx} \). This is true in a fairly general setting: if all the 2nd-order derivatives are continuous.

So for nice function it is better to choose a good order.

Ex. Let \( g(x,y) = e^{(2x+3y)} + xy \). Find \( g_{xy} \).

**Solution.** As you can see, a part of \( g \) is a complicated function of \( x \). So it takes time to compute \( g_x \), and then \( g_{xy} \). But since it is a nice function, \( g_{xy} = g_{yx} \). And \( g_y = x \). So \( g_{xy} = g_{yx} = 1 \).

We can talk about iterated partial derivatives of more than
Lecture 20: Higher order partial derivatives

Monday, November 7, 2016   8:45 AM

2 variable functions and the same technique can be used:

Ex. Let \( g(x, y, z) = \ln(\cos(x^2 + y^2)) + xy^2z^3 \). Find \( g_{xyz} \).

Solution. Since \( g \) is a nice function, we can choose any order we like to find \( g_{xyz} \).

\( g \) is easiest as a function of \( z \). So let’s compute \( g_z \) first:

\[
 g_z = 3xy^2z^2.
\]

Now it is easy enough to work with any order of \( x \) and \( y \):

\[
 g_{zx} = 3y^2z^2 \quad \text{and} \quad g_{zxy} = 6yz^2.
\]

Hence \( g_{xyz} = g_{zxy} = 6yz^2 \).

Optimization problems: finding maximum and minimum of a multi-variable function.

As we have seen before, if directional derivative of \( f \) in the direction of \( \mathbf{u} \) is positive, then \( f \) increases in the direction of \( \mathbf{u} \). And we have seen the directional derivative of \( f \) in the direction of \( \nabla f(p_0) \) is \( \|\nabla f(p_0)\| \). So if \( f \) has a local maximum at \( p_0 \), then

\[
 \nabla f(p_0) = \mathbf{0}.
\]
Similarly we have $D_\alpha f(p_0) < 0$ implies $f$ decreases in the direction of $\alpha$. And directional derivative in the direction of $\nabla f(p_0)$ is $-||\nabla f(p_0)||$. So we get

If $f$ is differentiable at $p_0$ and $f$ has a local maximum or a local minimum at $p_0$, then $\nabla f(p_0) = 0$.

For instance consider the following combination of “mountains” and “lake”.

Because of the above important box we define a critical point of $f$ as follows:

**Definition.** We say $p_0$ is a critical point of $f$ if either $f$ is not differentiable at $p_0$, or $\nabla f(p_0) = 0$.

So we have

If $f$ has a local max or a local min at $p_0$, then $p_0$ is a critical point of $f$. 
Not all the critical points are going to give us local extreme values.

If $f$ does NOT have a local max or a local min at a critical point $p_0$, then $p_0$ is called a saddle point.