More on multivariable chain rules

The "Type I" Chain Rule has some important theoretical consequences.

The Chain Rule—Type I. The derivative of the composition $f(\mathbf{P}(t))$ is

$$\frac{df}{dt}\bigg|_{t=t_0} = \nabla f(\mathbf{P}(t_0)) \cdot \mathbf{P}'(t_0).$$

Theorem.

1. Let $f(x, y)$ be a differentiable function such that $\nabla f(x, y) = \mathbf{0}$ for all (x, y). Then $f(x, y)$ is a constant function.

2. If $g(x, y)$ and $h(x, y)$ are two differentiable functions such that

$$\nabla g(x, y) = \nabla h(x, y)$$

for all (x, y). Then $g(x, y) = h(x, y) + K$ for some constant K.

Chain Rule—Type II

For this situation, consider a function $f(x, y)$ of two variables and suppose that the variables x and y are functions of other variables.

For example, consider x and y as a function of the polar coordinates r and θ. That is,

$$x = r\cos\theta \quad \text{and} \quad y = r\sin\theta.$$

Example. Let $f(x, y) = xy + y^2$. What is the angular rate of change of $f(x, y)$ at the point $(x, y) = (1, 2)$?

Directional derivatives

Partial derivatives only measure rates of change along paths parallel to the axes. Directional derivatives measure the rate of change of a function in any direction.

Example. Consider the function

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

Let's draw its level curves and its graph and calculate a rate of change in a direction other than one parallel to the x- or y-axes at the point $(2, 1)$.

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 as

$$\begin{aligned} x &= a + u_1 h \\ y &= b + u_2 h \end{aligned}$$

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 $\mathbf{P} = (a, b)$, 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. For example, given a function of three variables $f(x, y, z)$ and a position vector $\mathbf{P}$ corresponding to a point (a, b, c), then the definition in the three variable case is

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

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 $f(\mathbf{L}(h))$, where

$$\mathbf{L}(h) = (a + u_1 h)\mathbf{i} + (b + u_2 h)\mathbf{j}$$

is the line that goes through (a, b) at $t = 0$. Note that the vector $\mathbf{u} = u_1\,\mathbf{i} + u_2\,\mathbf{j}$ is used only to indicate a direction, and consequently, it is *always a unit vector*. In other words, $u_1^2 + u_2^2 = 1$.

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