Multivariable chain rules

Today we start our discussion of multivariable chain rules. There are two basic types, and both generalize the single variable chain rule.

Chain Rule—Type I

Consider a vector-valued function $\mathbf{P}(t) = x(t)\mathbf{i} + y(t)\mathbf{j}$ which parameterizes a curve in the xy-plane and a function $f(x, y)$. How is the derivative df/dt related to the partial derivatives of $f(x, y)$ and the derivative $\mathbf{P}'(t)$?

Example. Consider $\mathbf{P}(t) = t\mathbf{i} + t^2\mathbf{j}$ and $f(x, y) = 2x^2 + y^2$. Let's calculate the derivative df/dt at $t = 1$.

The problem with this approach is that it ignores the fact that the function in question is really a composition of two functions. We can illustrate this composition symbolically by making a dependency chart.

Chain Rule. The derivative of the composition $f(\mathbf{P}(t))$ is given by

$$\frac{df}{dt} = \frac{\partial f}{\partial x}\frac{dx}{dt} + \frac{\partial f}{\partial y}\frac{dy}{dt}.$$

Example. Back to the composition of $\mathbf{P}(t) = t\mathbf{i} + t^2\mathbf{j}$ and $f(x, y) = 2x^2 + y^2$. Let's calculate the derivative df/dt at $t = 1$ using the Chain Rule.

This version of the Chain Rule has an important formulation in terms of the gradient of f.

Definition. Given a function $f(x, y)$ that is differentiable at the point (a, b). Then the *gradient vector* of f at (a, b) is the vector

$$\nabla f(a,b) = \frac{\partial f}{\partial x}(a,b)\,\mathbf{i} + \frac{\partial f}{\partial y}(a,b)\,\mathbf{j}.$$

Sometimes the gradient vector of f is denoted $\mathbf{grad} f(a, b)$.

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

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

Example. Once again we return to $\mathbf{P}(t) = t\mathbf{i} + t^2\mathbf{j}$ and $f(x, y) = 2x^2 + y^2$.

Animation of this chain rule

Example. Use the polar curve $r = \cos 2\theta$ to parameterize a curve $\mathbf{P}(t)$ in the xy-plane and consider the composition $f(\mathbf{P}(t))$ where

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

This chain rule has some important theoretical implications as well.

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.