Math 2210 - Section 12.6 Notes

Dylan Zwick

Fall 2008

1 The Chain Rule

1.1 The Calculus I Chain Rule

In calculus I we learned that if we have a composite of two functions, $y(x) = f(g(x))$ then the derivative of the composite was the derivative of the outside function, multiplied by the derivative of the inside function:

$$y'(x) = f'(g(x))g'(x).$$

Example
What is the derivative of $\ln(\sin(x^2 + e^x))$?

1.1.1 The First Version of the Multivariable Chain Rule

If $z = f(x, y)$ is a function of two variables, and both of those variables are in turn functions of a single parameter t, then we can view the function z as a function of the single parameter t.

The idea behind this sentence is much easier to understand than it appears. For example, suppose we have the function $z = \sin(x+y)$, with $x = t^2$ and $y = t^3$, then we could write z as a function of just t, namely $z = \sin(t^2 + t^3)$.

Well, z when expressed like this is just a single variable function, and so if the functions f, x, and y are differentiable, then it makes sense to talk about the derivative of z with respect to t. The relationship between the derivative of z with respect to t, and the other derivatives of f, x, and y are:

Theorem

Let $x = x(t)$ and $y = y(t)$ be differentiable at t, and let $z = f(x, y)$ be differentiable at $(x(t), y(t))$. Then $z = f(x(t), y(t))$ is differentiable at t and

$$\frac{dz}{dt} = \frac{\partial z}{\partial x}\frac{dx}{dt} + \frac{\partial z}{\partial y}\frac{dy}{dt}.$$

This is the first version of the chain rule for multivariable functions. Basically, it's just saying that the amount that z changes when we change t is how much z changes when we change x, multiplied by how much x changes when we change t added to how much z changes when we change y, multiplied by how much y changes when we change t. Again, that's a long sentence, but walk through it and you'll see it's really just logic. The proof is pretty straightforward.

Proof

If we simplify notation and let $\mathbf{p} = (\Delta x, \Delta y)$, and $\Delta z = f(\mathbf{p}+\Delta\mathbf{p}) - f(\mathbf{p})$ then since f is differentiable we have:

$$\begin{aligned}\Delta z = f(\mathbf{p}+\Delta\mathbf{p}) - f(\mathbf{p}) &= \bigtriangledown f(\mathbf{p}) \cdot \Delta\mathbf{p} + \epsilon(\mathbf{p}) \cdot \Delta\mathbf{p} \\ &= f_x(\mathbf{p})\Delta x + f_y(\mathbf{p})\Delta y + \epsilon(\Delta\mathbf{p}) \cdot \Delta\mathbf{p}\end{aligned}$$

where $\epsilon(\mathbf{p}) \to \mathbf{0}$ as $\Delta\mathbf{p} \to \mathbf{0}$.

Now, if we divide both sides by Δt and take the limit as $\Delta t \to 0$ we get:

$$\frac{dz}{dt} = f_x(\mathbf{p})\frac{dx}{dt} + f_y(\mathbf{p})\frac{dy}{dt}.$$

which is what we want to prove.

Example
Find $\dfrac{dw}{dt}$ given $w = x^2y - y^2x$, $x = \cos t$, $y = \sin t$.

1.2 The Second Version of the Multivariable Chain Rule

This is a natural extension of the concepts we just discussed. Suppose that we have a function $z = f(x, y)$ and x and y are themselves functions of two other parameters s and t, say $x = x(s, t)$ and $y = y(s, t)$. Then z itself can be viewed as a function of s and t, and if everything is differentiable we can takes its partial derivative with respect to s or t. The corresponding relations are:

Theorem - Let $x = x(s, t)$ and $y = y(s, t)$ have first partial derivaties at (s, t) and let $z = f(x, y)$ be differentaible at $(x(s, t), y(s, t))$. Then $z = f(x(s, t), y(s, t))$ has first partial derivatives given by:

$$\frac{\partial z}{\partial s} = \frac{\partial f}{\partial x}\frac{\partial x}{\partial s} + \frac{\partial f}{\partial y}\frac{\partial y}{\partial s} \text{ and } \frac{\partial z}{\partial t} = \frac{\partial f}{\partial x}\frac{\partial x}{\partial t} + \frac{\partial f}{\partial y}\frac{\partial y}{\partial t}.$$