One of the most powerful tools for differentiating functions is the chain rule. We are all familiar with the one dimensional version, $\frac{d}{dx}(f \circ g)(x) = \frac{df}{dx}(g(x))\frac{dg}{dx}(x)$ and this generalizes to the multivariable case. Suppose we have a function $\mathbf{F} : \mathbb{R}^n \to \mathbb{R}^m$ and another function $\mathbf{G} : \mathbb{R}^m \to \mathbb{R}^l$. Then the composition, $\mathbf{G} \circ \mathbf{F}$ is defined. Now if $\mathbf{F}$ is differentiable on an open subset U of $\mathbb{R}^n$ containing a point $\mathbf{a}$ and $\mathbf{G}$ is differentiable on an open subset of $\mathbb{R}^m$ containing $\mathbf{F}(U)$ then $\mathbf{G} \circ \mathbf{F}$ is differentiable at $\mathbf{a}$, and:

$$D(\mathbf{G} \circ \mathbf{F})(\mathbf{a}) = D\mathbf{G}(\mathbf{F}(\mathbf{a})) \cdot D\mathbf{F}(\mathbf{a})$$

Here $\cdot$ denotes matrix multiplication. If we apply this to two single variable functions we get the traditional chain rule.

This rule can be useful in many situations, and can even be used to derive many other differentiation rules. For example, if you let $\mathbf{F} : \mathbb{R} \to \mathbb{R}^2$ such that $\mathbf{F}(x) = (f(x), g(x))$ and $G : \mathbb{R}^2 \to \mathbb{R}$ such that $G(u, v) = uv$ then applying the chain rule to $G \circ \mathbf{F}$ yields the well known product rule from single variable calculus.

Many times a function can be decomposed into the composition of functions. When computing the derivatives of such functions, we can either try to attack it as is, likely using the single variable chain rule when computing the partials, or we can decompose it and use the multivariable chain rule. Either approach will give the same results.

Tutorial Problems

1. Let $\mathbf{F}(x, y) = f(g(x), h(y), k(x, y))$, where $f : \mathbb{R}^3 \to \mathbb{R}$. Assume all functions are differentiable. Find $\mathbf{F}_x$ and $\mathbf{F}_y$.

Let $u(x, y) = (g(x), h(y), k(x, y))$. Then $\mathbf{F} = f \circ u$, so $D\mathbf{F}(x, y) = Df(u(x, y)) \cdot Du(x, y)$. Now $Df(u(x, y)) = [\, D_1 f \;\; D_2 f \;\; D_3 f \,]$, where $D_i f$ is the ith partial of f evaluated at $u(x, y)$.

$$Du(x, y) = \begin{bmatrix} g'(x) & 0 \\ 0 & h'(y) \\ \partial k/\partial x(x, y) & \partial k/\partial y(x, y) \end{bmatrix}$$

Multiplying gives $\mathbf{F}_x = D_1 f g'(x) + D_3 f \partial f/\partial x(x, y)$ and $\mathbf{F}_y = D_2 f h'(y) + D_3 f \partial f/\partial y(x, y)$.

2. Let $F(x, y) = x^3 y$, where $x^3 + tx = 8$ and $ye^y = t$. Find $\frac{dF}{dt}(0)$.

Considering F as a function of t, we have $F(t) = u \circ v(t)$ where $u(x, y) = x^3 y$ and $v(t) = (x(t), y(t))$, such that $x(t)$ satisfies $x^3 + tx = 8$ and $y(t)$ satisfies $ye^y = t$. Now $DF(0) = Du(v(0)) \cdot Dv(0)$. Now we find $v(0)$ by solving the equations $x^3 + 0x = 8$ and

$ye^y = 0$. We get $v(0) = (2, 0)$. Now $Du = [\,3x^2y \;\; x^3\,]$, so $Du(v(0)) = [\,0 \;\; 8\,]$. We find $Dv(0)$ by implicitly differentiating the constaining equations for x and y. For x, we get $3x^2x' + x + tx' = 0$. Subbing in $t = 0$, $x = 2$, we get $12x'(0) + 2 + 0x'(0) = 0$, $x'(0) = -1/6$. For y, we get $y'e^y + y'ye^y = 1$. Subbing in $y = 0$, $t = 0$, we get $y'(0) + 0y'(0) = 1$ and so $y'(0) = 1$. So $Dv(0) = [\,-1/6 \;\; 1\,]^T$. Multiplying the matrices we get $dF/dx(0) = 8$.

3. For $\mathbf{F}, \mathbf{G} : \mathbb{R} \to \mathbb{R}^2$, both differentiable, find $D(\mathbf{F} \cdot \mathbf{G})$ where $\cdot$ denotes the dot product in $\mathbb{R}^n$.

$\mathbf{F} \bullet \mathbf{G} = \mathbf{F}_1\mathbf{G}_1 + \mathbf{F}_2\mathbf{G}_2$. We define two functions u and v such that $u \circ v(x) = \mathbf{F}(x) \bullet \mathbf{G}(x)$. Let $v(x) = (\mathbf{F}_1(x), \mathbf{F}_2(x), \mathbf{G}_1(x), \mathbf{G}_2(x))$. Let $u(x_1, x_2, y_1, y_2) = x_1y_1 + x_2y_2$. Now $Du \circ v(x) = Du(v(x)) \cdot Dv(x)$. We have $Du(v(x)) = [\,\mathbf{G}_1 \;\; \mathbf{G}_2 \;\; \mathbf{F}_1 \;\; \mathbf{F}_2\,]$ and $Dv(x) = [\,\mathbf{F}_1'(x) \;\; \mathbf{F}_2'(x) \;\; \mathbf{G}_1' \;\; \mathbf{G}_2'\,]^T$. Multiplying gives $D\mathbf{F} \bullet \mathbf{G} = \mathbf{F}_1'\mathbf{G}_1 + \mathbf{F}_2'\mathbf{G}_2 + \mathbf{F}_1\mathbf{G}_1' + \mathbf{F}_2\mathbf{G}_2' = D\mathbf{F} \bullet \mathbf{G} + \mathbf{F} \bullet D\mathbf{G}$.

4. I have a differentiable function $F : \mathbb{R}^2 \to \mathbb{R}$. I switched into polar coordinates and found that at $(r, \theta) = (1, \pi/4)$ we have $\partial F/\partial r = 2$ and $\partial F/\partial\theta = -1$. What are the partials with respect to x and y at the point $(1/\sqrt{2}, 1/\sqrt{2})$?

Considering F as a function of r and θ, we put $F = u \circ v$, where $v(r, \theta) = (r\cos\theta, r\sin\theta)$ and u is a real valued function in x and y. We are given that $DF = [\,2 \;\; -1\,]$, but we can also calculate it in terms of Du and Dv using the chain rule. (Note that $v(1, \pi/4) = (1/\sqrt{2}, 1/\sqrt{2}.)$ Now $DF(1, \pi/4) = Du(1/\sqrt{2}, 1/\sqrt{2}) \cdot Dv(1, \pi/4)$. $Du(1/\sqrt{2}, 1/\sqrt{2} = [\,u_x(1/\sqrt{2}, 1/\sqrt{2}) \;\; u_y(1/\sqrt{2}, 1/\sqrt{2})\,]$ which are the values we are trying to find. Next we have:

$$Dv(1, \pi/4) = \begin{bmatrix} \cos\theta & -r\sin\theta \\ \sin\theta & r\cos\theta \end{bmatrix}_{(1,\pi/4)} = \begin{bmatrix} 1/\sqrt{2} & -1/\sqrt{2} \\ 1/\sqrt{2} & 1/\sqrt{2} \end{bmatrix}$$

Note that $Dv(1, \pi/4)^{-1} = Dv(1, \pi/4)^T$. Now we have $Du(1/\sqrt{2}, 1/\sqrt{2}) \cdot Dv(1, \pi/4) = [\,2 \;\; -1\,]$. Multiplying both sides on the right by $Dv(1, \pi/4)^T$ gives $Du(1/\sqrt{2}, 1/\sqrt{2}) = [\,3/\sqrt{2} \;\; 1/\sqrt{2}\,]$.

5. I have a function that measures temperature and pressure at a point in $\mathbb{R}^3$, given by $F(x,y,z) = (-z^{1/2}+\sin(x+y), 1/z^3+e^{-xy})$. I want to know what the rates of change are as I move along the surface $z = e^{-xy}$. Find the partial derivatives with respect to x and y.

We define $G : \mathbb{R}^2 \to \mathbb{R}^2$ to be the value of F on the surface. That is $G = F \circ u$ where $u(x,y) = (x, y, e^{-xy})$. Now $DG(a,b) = DF(u(a,b)) \cdot Du(a,b)$. We find that:

$$\begin{aligned} DF(u(a,b)) &= \begin{bmatrix} \cos(x+y) & \cos(x+y) & -1/2z^{-1/2} \\ -ye^{-xy} & -xe^{-xy} & -3/z^4 \end{bmatrix}_{u(a,b)} \\ &= \begin{bmatrix} \cos(a+b) & \cos(a+b) & -1/2e^{1/2ab} \\ -be^{-ab} & -ae^{-ab} & -3e^{4ab} \end{bmatrix} \end{aligned}$$

Next we have:

$$Du(a,b) = \begin{bmatrix} 1 & 0 \\ 0 & 1 \\ -be^{-ab} & -ae^{-ab} \end{bmatrix}$$

Multiplying gives:

$$DG(a,b) = \begin{bmatrix} \cos(a+b) + b/2e^{-1/2ab} & \cos(a+b) + a/2e^{-1/2ab} \\ -be^{-ab} + 3be^{3ab} & -ae^{-ab} + 3ae^{3ab} \end{bmatrix}$$

6. Let $f(x,y) = (e^{e^{xy}+\sin y}, e^{xy}\sin y)$. Calculate Df first directly from the definition, then by writing f as the composition of two functions.

First we calculate Df directly:

$$Df = \begin{bmatrix} ye^{xy}e^{e^{xy}+\sin y} & (xe^{xy} + \cos y)e^{e^{xy}+\sin y} \\ ye^{xy}\sin y & xe^{xy}\sin y + e^{xy}\cos y \end{bmatrix}$$

Next we write $f = u \circ v$ where $v(x,y) = (e^{xy}, \sin y)$ and $u(\hat{x},\hat{y}) = (e^{\hat{x}+\hat{y}}, \hat{x}\hat{y})$. Now we compute $Df(x,y) = Du(v(x,y)) \cdot Dv(x,y)$. First, we have:

$$Du(v(x,y)) = \begin{bmatrix} e^{\hat{x}+\hat{y}} & e^{\hat{x}+\hat{y}} \\ \hat{y} & \hat{x} \end{bmatrix}_{v(x,y)} = \begin{bmatrix} e^{e^{xy}+\sin y} & e^{e^{xy}+\sin y} \\ \sin y & e^{xy} \end{bmatrix}$$

Secondly, we have:

$$Dv(x,y) = \begin{bmatrix} ye^{xy} & xe^{xy} \\ 0 & \cos y \end{bmatrix}$$

Multiplying we see that we get the same result.