CLARK UNIVERSITY 1887

The chain rule, part 2
Math 131 Multivariate Calculus
D Joyce, Spring 2014

We'll continue the stages to a complete generalization of the chain rule that we started in part 1.

The second step where t becomes a vector $\mathbf{t}$. In general, we'll want $\mathbf{t}$ to be a vector $(t_1, t_2, \ldots, t_n)$, but, for purposes of illustration, let's make $n = 2$, and write $\mathbf{t} = (s, t)$. And, while we're at it, let's have $m = 2$ so that $\mathbf{x} = (x, y)$ where x and y are each functions of both s and t. Then $f \circ \mathbf{x} : \mathbf{R}^2 \to \mathbf{R}$ is defined by

$$(f \circ \mathbf{x})(\mathbf{t}) = f(x(s,t), y(s,t)).$$

We want to find $D(f \circ \mathbf{x})$ which comprises the two partial derivatives $\dfrac{\partial f}{\partial s}$ and $\dfrac{\partial f}{\partial t}$.

But these partials are each derivatives explained in the previous paragraphs. After all, a partial derivative is just an ordinary derivative when the other variables are left constant. So, we have

$$\frac{\partial f}{\partial s} = \frac{\partial f}{\partial x}\frac{\partial x}{\partial s} + \frac{\partial f}{\partial y}\frac{\partial y}{\partial s}$$

and

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

When this is written out in matrix notation, we get the same matrix equation for the chain rule, namely,

$$D(f \circ \mathbf{x}) = Df\ D\mathbf{x},$$

the only difference is, this time the two matrices being multiplied together don't have one row and one column, but two, and in general they'll have n.

Example 1. We need an example to illustrate what's going on. Let $f(x, y) = \sqrt{x^2 + y^2}$, let $x(s, t) = s \ln t$, and let $y(s, t) = \sin x + \cos t$. Then

$$f(x(s,t), y(s,t)) = \sqrt{(s \ln t)^2 + (\sin s + \cos t)^2}.$$

Therefore,

$$\begin{aligned}
\frac{df}{ds} &= \frac{\partial f}{\partial x}\frac{\partial x}{\partial s} + \frac{\partial f}{\partial y}\frac{\partial y}{\partial s} \\
&= \frac{x}{\sqrt{x^2+y^2}} \ln t + \frac{y}{\sqrt{x^2+y^2}} \cos s \\
\frac{df}{dt} &= \frac{\partial f}{\partial x}\frac{\partial x}{\partial t} + \frac{\partial f}{\partial y}\frac{\partial y}{\partial t} \\
&= \frac{x}{\sqrt{x^2+y^2}} \left(\frac{s}{t}\right) + \frac{y}{\sqrt{x^2+y^2}} (-\sin t)
\end{aligned}$$

The derivative of a composition is the product of matrices. We're up to the last step of generalizing the chain rule where f becomes a vector-valued function.

There really isn't much more to do when $\mathbf{f} = (f_1, f_2, \ldots, f_p)$. That's because the derivative $D\mathbf{f}$ of the vector-valued function $\mathbf{f}$ is just a matrix whose rows are the derivatives Df_j of the component functions f_j for $j = 1, 2, \ldots, p$. Since each $D(f_j \circ \mathbf{x}) = Df_j\, D\mathbf{x}$, placing them in rows gives the matrix product

$$\boxed{D(\mathbf{f} \circ \mathbf{x}) = D\mathbf{f}\, D\mathbf{x}}$$

Note that $\mathbf{f}$ is a function $\mathbf{R}^m \to \mathbf{R}^p$ while its derivative $D\mathbf{f}$ is a $p \times m$ matrix; also $\mathbf{x}$ is a function $\mathbf{R}^n \to \mathbf{R}^m$ while its derivative $D\mathbf{x}$ is an $m \times n$ matrix; so the composition $\mathbf{f} \circ \mathbf{x}$ is a function $\mathbf{R}^n \to \mathbf{R}^p$ and its derivative $D(\mathbf{f} \circ \mathbf{x})$ is a $p \times n$ matrix.

Polar/rectangular conversions. We now know that the chain rule is, in general, a product of matrices

$$D(\mathbf{f} \circ \mathbf{x}) = D\mathbf{f}\, D\mathbf{x}.$$

As an important application of this general chain rule, we'll develop the derivatives associated to polar/rectangular conversions.

Suppose $f : \mathbf{R}^2 \to \mathbf{R}$ is a scalar-valued function defined on the plane. Let's use the notation $w = f(x, y)$. Then we have the two partial derivatives $\frac{\partial w}{\partial x}$ and $\frac{\partial w}{\partial y}$ as usual. But sometimes we want to use polar coordinates and find the partials with respect to r and θ. Of course, we could figure them out directly since we know how to express f in terms of r and θ, since $w = f(x, y) = f(x(r, \theta), y(r, \theta))$ where $x(r, \theta) = r \cos\theta$ and $y(r, \theta) = r \sin\theta$. Here the question is: can we figure out what $\frac{\partial w}{\partial r}$ and $\frac{\partial w}{\partial \theta}$ are directly from $\frac{\partial w}{\partial x}$ and $\frac{\partial w}{\partial y}$?

We can treat the vector $\mathbf{x} = (x, y)$ as a function $\mathbf{R}^2 \to \mathbf{R}^2$ as

$$\mathbf{x}(r, \theta) = (x(r, \theta), y(r, \theta)) = (r \cos\theta, r \sin\theta).$$

Then we can apply the general chain rule $D(\mathbf{f} \circ \mathbf{x}) = D\mathbf{f}\, D\mathbf{x}$. We'll get the matrix equation

$$\begin{aligned}
\begin{bmatrix} \frac{\partial w}{\partial r} & \frac{\partial w}{\partial \theta} \end{bmatrix} &= \begin{bmatrix} \frac{\partial w}{\partial x} & \frac{\partial w}{\partial y} \end{bmatrix} \begin{bmatrix} \frac{\partial x}{\partial r} & \frac{\partial x}{\partial \theta} \\ \frac{\partial y}{\partial r} & \frac{\partial y}{\partial \theta} \end{bmatrix} \\
&= \begin{bmatrix} \frac{\partial w}{\partial x} & \frac{\partial w}{\partial y} \end{bmatrix} \begin{bmatrix} \cos\theta & -r\sin\theta \\ \sin\theta & r\cos\theta \end{bmatrix}
\end{aligned}$$

This result is usally summarized in terms of differential operators where the name of the function being differentiated is omitted.

$$\begin{aligned}
\frac{\partial}{\partial r} &= \cos\theta \frac{\partial}{\partial x} + \sin\theta \frac{\partial}{\partial y} \\
\frac{\partial}{\partial \theta} &= -r\sin\theta \frac{\partial}{\partial x} + r\cos\theta \frac{\partial}{\partial y}
\end{aligned}$$

Conversely, you can find the partials with respect to x and y in terms of those for r and θ.

$$\begin{aligned}
\frac{\partial}{\partial x} &= \cos\theta \frac{\partial}{\partial r} - \frac{\sin\theta}{r} \frac{\partial}{\partial \theta} \\
\frac{\partial}{\partial y} &= \sin\theta \frac{\partial}{\partial r} + \frac{\cos\theta}{r} \frac{\partial}{\partial \theta}
\end{aligned}$$

Math 131 Home Page at
`http://math.clarku.edu/~djoyce/ma131/`