CLARK UNIVERSITY 1887

Total derivatives
Math 131 Multivariate Calculus
D Joyce, Spring 2014

Last time. We found that the total derivative of a scalar-valued function, also called a scalar field, $\mathbf{R}^n \to \mathbf{R}$, is the gradient

$$\nabla f = (f_{x_1}, f_{x_2}, \ldots, f_{x_n}) = \left(\frac{\partial f}{\partial x_1}, \frac{\partial f}{\partial x_2}, \ldots, \frac{\partial f}{\partial x_n}\right).$$

When $n = 2$ the gradient, $\nabla f = (f_x, f_y)$, gives the slopes of the tangent plane in the x-direction and the y-direction.

Total derivatives to vector-valued functions. Let $\mathbf{f} : \mathbf{R}^n \to \mathbf{R}^m$ be a vector-valued function. As always, a vector valued function is determined by its m scalar-valued component functions:

$$\mathbf{f}(\mathbf{x}) = (f_1(\mathbf{x}), f_1(\mathbf{x}), \ldots, f_m(\mathbf{x})).$$

We'll say $\mathbf{f}$ is *differentiable* if all its component functions are differentiable, and in that case, we'll take the *derivative* of $\mathbf{f}$, denoted $D\mathbf{f}$, to be the $m \times n$ matrix of partial derivatives of the component functions:

$$D\mathbf{f}(x_1, x_2, \ldots, x_n) = \begin{bmatrix} \frac{\partial f_1}{\partial x_1} & \frac{\partial f_1}{\partial x_2} & \cdots & \frac{\partial f_1}{\partial x_n} \\ \frac{\partial f_2}{\partial x_1} & \frac{\partial f_2}{\partial x_2} & \cdots & \frac{\partial f_2}{\partial x_n} \\ \vdots & \vdots & \ddots & \vdots \\ \frac{\partial f_m}{\partial x_1} & \frac{\partial f_m}{\partial x_2} & \cdots & \frac{\partial f_m}{\partial x_n} \end{bmatrix}.$$

Thus, the ij-th entry is $\dfrac{\partial f_i}{\partial x_j}$, the j-th partial derivative of the i-component function f_i.

In terms of limits, a logically equivalent definition says $\mathbf{f}$ is differentiable at $\mathbf{a}$ if (1) all the mn partial derivatives exist, and (2) the linear function

$$\mathbf{h}(\mathbf{x}) = \mathbf{f}(\mathbf{a}) + D\mathbf{f}(\mathbf{a})\,(\mathbf{x} - \mathbf{a})$$

is a good approximation of $\mathbf{f}$ near $\mathbf{a}$ in the sense that

$$\lim_{\mathbf{x}\to\mathbf{a}} \frac{\mathbf{f}(\mathbf{x}) - \mathbf{h}(\mathbf{x})}{\|\mathbf{x} - \mathbf{a}\|} = \mathbf{0}.$$

Note that when $m = 1$, this definition says the derivative Df of the scalar-valued function f is the $1 \times n$ row-matrix

$$Df(x_1, x_2, \ldots, x_n) = \begin{bmatrix} \frac{\partial f}{\partial x_1} & \frac{\partial f}{\partial x_2} & \cdots & \frac{\partial f}{\partial x_n} \end{bmatrix}$$

which is the same thing as the gradient ∇f, but written out as a row-matrix rather than an n-tuple.

Example 1. This derivative $D\mathbf{f}$ looks complicated, but it isn't, really. For an example, let $\mathbf{f} : \mathbf{R}^3 \to \mathbf{R}^4$ be defined by

$$\mathbf{f}(x, y, z) = (x + 2y + 3z, xyz, \cos x, \sin x).$$

Then $D\mathbf{f}$ is the 4×3 matrix $D\mathbf{f}(x_1, x_2, \ldots, x_n)$

$$= \begin{bmatrix} \frac{\partial f_1}{\partial x} & \frac{\partial f_1}{\partial y} & \frac{\partial f_1}{\partial z} \\ \frac{\partial f_2}{\partial x} & \frac{\partial f_2}{\partial y} & \frac{\partial f_2}{\partial z} \\ \frac{\partial f_3}{\partial x} & \frac{\partial f_3}{\partial y} & \frac{\partial f_3}{\partial z} \\ \frac{\partial f_4}{\partial x} & \frac{\partial f_4}{\partial y} & \frac{\partial f_4}{\partial z} \end{bmatrix} = \begin{bmatrix} 1 & 2 & 3 \\ yz & xz & xy \\ -\sin x & 0 & 0 \\ \cos x & 0 & 0 \end{bmatrix}$$

Rules of differentiation. Typically, when derivatives of multivariant functions are actually computed, they're computed one partial derivative at a time. Partial derivatives are just ordinary derivatives when only one variable actually varies, so no new rules of differentiation are needed for them. But there are rules for gradients and total derivatives.

The usual properties of derivatives for functions of one variable $f : \mathbf{R} \to \mathbf{R}$ are the sum rule, difference rule, product rule, quotient rule, and chain rule. For the most part, these properties are enjoyed by functions $\mathbf{f} : \mathbf{R}^n \to \mathbf{R}^m$, but the product and quotient rules will need to be modified and restricted.

For instance, derivatives are linear in the sense that they preserve addition and scalar multiplication, and therefore they preserve subtraction and, in general, all linear combinations. So, for vector-valued differentiable functions $\mathbf{R}^n \to \mathbf{R}^m$, we have

$$\boxed{\begin{array}{rcl} D(\mathbf{f}+\mathbf{g}) & = & D\mathbf{f}+D\mathbf{g} \\ D(c\mathbf{f}) & = & cD\mathbf{f} \\ D(\mathbf{f}-\mathbf{g}) & = & D\mathbf{f}-D\mathbf{g} \\ D(c_1\mathbf{f}_1+\cdots+c_r\mathbf{f}_r) & = & c_1D\mathbf{f}_1+\cdots+c_rD\mathbf{f}_r \end{array}}$$

When the functions are scalar-valued functions $\mathbf{R}^n \to \mathbf{R}$, i.e., when $m = 1$, we can also write these rules in terms of gradients as

$$\boxed{\begin{array}{rcl} \nabla(f+g) & = & \nabla f+Dg \\ \nabla(cf) & = & c\,\nabla f \\ \nabla(f-g) & = & \nabla f-\nabla g \\ \nabla(c_1f_1+\cdots+c_rf_r) & = & c_1\nabla f_1+\cdots+c_r\nabla f_r \end{array}}$$

The product and quotient rules apply to scalar-valued functions f and g, both $\mathbf{R}^n \to \mathbf{R}$:

$$\boxed{\begin{array}{rcl} \nabla(fg) & = & (\nabla f)g+f\nabla g \\ \\ \nabla\left(\dfrac{f}{g}\right) & = & \dfrac{(\nabla f)g-f\nabla g}{g^2} \end{array}}$$

The product and quotient rules don't easily generalize to vector-valued functions.

Later, we'll see what happens to the chain rule.

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