Lecture 3, 1/24

Definition. Let $f : U \subseteq \mathbf{R}^n \to \mathbf{R}^m$, where U is open, and let $a \in U$. For $j = 1, \ldots, n$ the *jth partial derivative of f at a* is $D_{e_j} f(a)$. Standard abbreviations for this are $D_j f(a)$ and $\frac{\partial f}{\partial x_j}(a)$.

Note that

$$\begin{aligned} D_j f(a) &= \frac{d}{dt} f(a_1, \ldots, a_j + t, \ldots, a_n)\big|_{t=0} \\ &= \frac{d}{dx_j} f(a_1, \ldots, x_j, \ldots, a_n)\big|_{x_j = a_j} \\ &= \frac{\partial}{\partial x_j} f(x_1, \ldots, x_n)\big|_{x=a}, \end{aligned}$$

where in the last line we use "curly" d's to indicate that the other *variables* are held constant during the differentiation with respect to x_j. It is this observation that makes partial differentiation easy (when the function is given by a formula): treat the other variables as constants, and use the usual rules for differentiation from one-variable calculus.

Since all directional derivatives exist for a differentiable function, all partial derivatives also exist. These can be used to compute the matrix of the derivative (relative to the standard bases of $\mathbf{R}^n$ and $\mathbf{R}^m$). Namely,

$$\begin{aligned} f'(a)_{ij} &= \big(f'(a)(e_j)\big)_i \\ &= \big(D_j f(a)\big)_i \\ &= D_j f_i(a) = \frac{\partial f_i}{\partial x_j}(a). \end{aligned}$$

The matrix $\big(D_j f_i(a)\big)_{m \times n}$ is called the *Jacobian matrix* of f at a. Note that the existence of the Jacobian matrix does not ensure that f is differentiable. However, it does indicate what the derivative must be if f is differentiable.

Example. Define $f : \mathbf{R}^2 \to \mathbf{R}$ by

$$f(x) = \begin{cases} \frac{x_1 x_2}{\|x\|}, & \text{if } x \neq 0 \\ 0, & \text{if } x = 0. \end{cases}$$

It is an easy exercise to check that $D_1 f(0) = D_2 f(0) = 0$ (though one must use the definition of directional derivative), and thus the Jacobian matrix of f at the origin is $(0\ 0)$. Therefore if f is differentiable, its derivative must be given by this matrix. We can now check whether this matrix defines a linear map satisfying the definition of the derivative:

$$\lim_{h \to 0} \frac{f(0+h) - f(0) - 0 \cdot h}{\|h\|} = \lim_{h \to 0} \frac{f(h)}{\|h\|} = \lim_{h \to 0} \frac{h_1 h_2}{h_1^2 + h_2^2}$$

does not exist (we saw this in an earlier example). Therefore f is not differentiable at the origin.

Since the existence of the partial derivatives at a point does not guarantee differentiability, it is of interest to find conditions stronger than mere existence that do imply differentiability. The standard result that follows is sufficient but not necessary, but is usually good enough.

Proposition. Let $U \subseteq \mathbf{R}^n$ be open, let $f : U \to \mathbf{R}^m$, and let $a \in U$. If the partial derivatives of f exist in a neighborhood of a, and are continuous at a, then f is differentiable at a.

Proof. Working separately with the component functions of f, we may as well assume that $m = 1$. We will show that the Jacobian matrix of f at a, which exists by hypothesis, does satisfy the definition of derivative. Since U is open there is a ball centered at a contained in U. For h in this ball,

$$\frac{1}{\|h\|}\Big(f(a+h) - f(a) - \sum_{j=1}^{n} D_j f(a) \cdot h_j\Big)$$

$$= \frac{1}{\|h\|}\sum_{j=1}^{n}\Big(f(a + h_1 e_1 + \cdots + h_j e_j) - f(a + h_1 e_1 + \cdots + h_{j-1} e_{j-1}) - D_j f(a) \cdot h_j\Big)$$

$$= \frac{1}{\|h\|}\sum_{j=1}^{n}\Big(h_j \cdot D_j f(a + h_1 e_1 + \cdots + h_{j-1} e_{j-1} + \theta_j h_j e_j) - D_j f(a) \cdot h_j\Big),$$

for some $\theta_j \in (0,1)$, by the mean value theorem,

$$= \sum_{j=1}^{n} \frac{h_j}{\|h\|}\Big(D_j f(a + h_1 e_1 + \cdots + h_{j-1} e_{j-1} + \theta_j h_j e_j) - D_j f(a)\Big).$$

In each term, the factor $h_j/\|h\|$ is bounded by 1 in absolute value, while the second factor tends to 0 as h tends to 0 since $D_j f$ is continuous at a. Therefore the entire sum tends to 0 as h tends to 0. Therefore f is differentiable at a. ∎

The following is a simple use of differentiability at a point, whose proof is nearly the same as the analogous result from one-variable calculus.

Proposition. Let $U \subseteq \mathbf{R}^n$ be open, let $f : U \to \mathbf{R}$, and suppose that f has a local extremum at the point $a \in U$. If f is differentiable at a, then $f'(a) = 0$.

Proof. Suppose for definiteness that f has a local minimum at a. For each $i = 1, \ldots, n$ consider the diffference quotient $\big(f(a + te_i) - f(a)\big)/t$. The numerator is greater than or equal to zero (for all small-enough $t \neq 0$), whereas the denominator may be either positive or negative. The limit as $t \to 0$ must then be zero. But the limit is the partial derivative $D_i f(a)$. Thus the Jacobian matrix of $f'(a)$ vanishes, so that $f'(a) = 0$. ∎

Our next goal is to prove the chain rule for differentiation. To do this, we will use another characterization of differentiability. To motivate it, we observe that if f is differentiable at a, then the quantity $f(a + h) - f(a) - f'(a)(h)$ not only tends to 0 as h tends to 0, but does so more rapidly than h itself. In fact, this quantity tends to 0 even if it is divided by $\|h\|$. Our new characterization involves singling out this quotient for consideration.

Lemma. Let $U \subseteq \mathbf{R}^n$ be open, let $f : U \to \mathbf{R}^m$, and let $a \in U$. Then f is differentiable at a if and only if there is a linear map $T \in L(\mathbf{R}^n, \mathbf{R}^m)$, and a function $\phi : W \to \mathbf{R}^m$, where W is a neighborhood of 0 in $\mathbf{R}^n$ such that $a + W \subseteq U$, satisfying the following three conditions:

(i) $\phi(0) = 0$.
(ii) ϕ is continuous at 0.
(iii) $f(a + h) = f(a) + T(h) + \phi(h)\|h\|$, for $h \in W$.

In this case, $f'(a) = T$.

Proof. (only if): Assuming that f is differentiable at a, define ϕ by

$$\phi(h) = \begin{cases} \frac{f(a+h)-f(a)-f'(a)(h)}{\|h\|}, & \text{if } h \neq 0 \\ 0 & \text{if } h = 0. \end{cases}$$

Then (i) and (iii) (with $T = f'(a)$) follow from the definition of ϕ, while (ii) is true by the definition of the derivative.

(if): Assuming that T and ϕ are given, we have for $h \neq 0$:

$$\frac{f(a + h) - f(a) - T(h)}{\|h\|} = \phi(h),$$

which tends to 0 as h tends to 0 by continuity of ϕ at 0. Therefore f is differentiable at a, and $f'(a) = T$. ■

Theorem. (The *Chain Rule.*) Let $U \subseteq \mathbf{R}^n$ and $V \subseteq \mathbf{R}^m$ be open, let $f : U \to \mathbf{R}^m$ and $g : V \to \mathbf{R}^\ell$ be functions, and let $a \in U$ with $f(a) \in V$. Suppose that f is differentiable at a and g is differentiable at $f(a)$. Then $g \circ f$ is differentiable at a, and $(g \circ f)'(a) = g'\big(f(a)\big) \circ f'(a)$.

Remarks. Notice that the domain of $g \circ f$ is $U \cap f^{-1}(V)$. Since V is a neighborhood of $f(a)$ and f is continuous at a, we know that $f^{-1}(V)$ is a neighborhood of a (i.e. it contains an open set that contains a). Therefore $U \cap f^{-1}(V)$ is a neighborhood of a. Thus it makes sense to try to differentiate $g \circ f$ at a.