Analysis: skeleton notes 5: derivatives

Robin Chapman

11 November 2011

Let I be an interval, and f be a function from I to $\mathbf{R}$. For $a \in I$ we say that f is *differentiable* at a if

$$\lim_{x \to a} \frac{f(x) - f(a)}{x - a}$$

exists. If so we denote this limit by $f'(a)$. Of course, we say that f is *differentiable* on I if it is differentiable at all $a \in I$. Then the function f' is called the *derivative* of f. Another way of writing the definition of differentiablity is

$$f'(a) = \lim_{h \to 0} \frac{f(a+h) - f(a)}{h}$$

provided this limit exists.

It is convenient sometimes in the theory to use an equivalent but more roundabout-seeming definition. We claim that f is differentiable at a if and only if there is a function $\varphi : I \to \mathbf{R}$ such that both

- $f(x) = f(a) + (x - a)\varphi(x)$ for all x, and
- φ is continuous at a.

Also $\varphi(a) = f'(a)$. Note that $\varphi(x)$ must equal $(f(x) - f(a))/(x - a)$ when $x \neq a$. To define $\varphi(a)$ so that φ is continuous at a then we must have

$$\varphi(a) = \lim_{x \to a} \varphi(x) = \lim_{x \to a} \frac{f(x) - f(a)}{x - a}.$$

As a consequence, $f(x) \to f(a) + (a - a)\varphi(a) = f(a)$ as $x \to a$ and so if f is differentiable at a it is continuous there.

Combining differentiable functions via the standard arithmetic operations gives differentiable functions. That is if f and g are differentiable at a then so are $f + g$, $f - g$, fg and f/g; moreover the derivatives are what you think they are. (Of course for f/g we need $g(a) \neq 0$; then it follows that $g(x) \neq 0$ when x is close enough to a). The sum and difference rules are very easy to prove, so I'll look at the product and quotient rules in detail.

As f and g are differentiable at a then there are functions φ and ψ, continuous at a with $f(x) = f(a)+(x-a)\varphi(x)$ and $g(x) = g(a)+(x-a)\psi(x)$; moreover $\varphi(a) = f'(a)$ and $\psi(a) = g'(a)$. To prove differentiability of fg at a, consider

$$\begin{aligned} & \frac{f(x)g(x) - f(a)g(a)}{x-a} \\ = & \frac{(f(a) + (x-a)\varphi(x))(g(a) + (x-a)\psi(x)) - f(a)g(a)}{x-a} \\ = & \varphi(x)g(a) + f(a)\psi(x) + (x-a)\varphi(x)\psi(x). \end{aligned}$$

As $\varphi(x) \to \varphi(a) = f'(a)$ as $x \to a$ (since φ is continuous at a) and similarly $\psi(x) \to g'(a)$ as $x \to a$ then

$$\frac{f(x)g(x) - f(a)g(a)}{x-a} \to f'(a)g(a) + f(a)g'(a)$$

as $x \to a$. This means that fg is differentiable at a with derivative $f'(a)g(a)+f(a)g'(a)$.

We can write f/g as $f(1/g)$ and so using the product rule we can reduce the quotient rule to the case of $1/g$. This time we consider

$$\begin{aligned} \frac{g(x)^{-1} - g(a)^{-1}}{x-a} &= -\frac{g(x) - g(a)}{(x-a)g(a)g(x)} \\ &= -\frac{(x-a)\psi(x)}{(x-a)g(a)g(x)} = -\frac{\psi(x)}{g(a)g(x)}. \end{aligned}$$

As $x \to a$ then $\psi(x) \to g'(a)$ and $g(x) \to g(a)$ (since g is continuous at a). Hence

$$\frac{g(x)^{-1} - g(a)^{-1}}{x-a} \to -\frac{g'(a)}{g(a)^2}$$

so that $1/g$ is differentiable at a with derivative $-g'(a)/g(a)^2$.

A similar argument proves the "chain rule" for the differentiability of the composite of two functions. Let $f : I \to \mathbf{R}$ and $g : J \to \mathbf{R}$ be functions such that $g \circ f$ makes sense, and suppose that f is differentiable at a and g at $f(a)$. Then $f(x) = f(a) + (x-a)\varphi(x)$ and $g(y) = g(f(a)) + (y - f(a))\psi(y)$ where φ is continuous at a, ψ is continuous at $f(a)$; moreover $\varphi(a) = f'(a)$ and $\psi(f(a)) = g'(f(a))$. We now consider

$$\frac{g(f(x)) - g(f(a))}{x-a} = \frac{g(f(a)) + (f(x) - f(a))\psi(f(x)) - g(f(a))}{x-a}$$

$$= \frac{(x-a)\varphi(x)\psi(f(x))}{x-a} = \varphi(x)\psi(f(x)).$$

As $x \to a$ then $f(x) \to f(a)$, $\varphi(x) \to \varphi(a) = f'(a)$ and $\psi(f(x)) \to \varphi(f(a)) = g'(f(a))$. Therefore

$$\frac{g(f(x)) - g(f(a))}{x-a} \to f'(a)g'(f(a))$$

which is the familiar "chain rule".

There is also a rule for calculating the derivative of an inverse function. Suppose that $f : I \to \mathbf{R}$ is a continuous increasing (or decreasing) function, and that f is differentiable at some $a \in I$ with $f'(a) \neq 0$. Let F be the inverse function of f. Write $b = f(a)$ and $f(I) = J$. Let (y_n) be a sequence of points in J converging to b, and with each $y_n \neq b$. Write $x_n = F(y_n)$. Then $a = F(b)$, $x_n \neq a$ and by the continuity of F, $x_n \to a$. Then

$$\frac{F(y_n) - F(b)}{y_n - b} = \frac{x_n - a}{f(x_n) - f(a)} \to \frac{1}{f'(a)}.$$

Hence $F'(b)$ exists and equals $1/f'(a)$.

Derivatives are often used to detect local maxima or minima. Suppose that f is defined (a, b), $c \in (a, b)$, $f'(c)$ exists and f is maximized over (a, b) at c, that is $f(x) \leq f(c)$ for all $x \in (a, b)$. Then there is a sequence (c_n) of points in (a, b) converging to c with each $c < c_n < b$. Then

$$\frac{f(c_n) - f(c)}{c_n - c} \leq 0$$

and so

$$f'(c) = \lim_{n\to\infty} \frac{f(c_n) - f(c)}{c_n - c} \leq 0.$$

Similarly there is a sequence (c'_n) of points in (a, b) converging to c with each $c > c'_n > a$, which leads to $f'(c) \geq 0$. Hence $f'(c) = 0$ and the argument adapts to the case of a local minimum too.

It's essential that we look at "interior" points of intervals here. On a closed interval a function may be maximized or minimized at the endpoints where the derivative may be nonzero.

Applying the Boundedness theorem and this maximum/minimum corollary leads to Rolle's theorem:

If $a < b$, f is continuous on $[a, b]$ and differentiable on (a, b) and moreover $f(a) = f(b)$ then there is $c \in (a, b)$ with $f'(c) = 0$.

For a proof see my last year's "proofs of major theorems". An almost immediate corollary is the Mean Value Theorem"

> If $a < b$, f is continuous on $[a, b]$ and differentiable on (a, b) then there is $c \in (a, b)$ with $f(b) - f(a) = f'(c)(b - a)$.

This is most usefully applied when we know numbers m and M such that $m \leq f'(x) \leq M$ for all $x \in (a, b)$. Then $m(b - a) \leq f(b) - f(a) \leq M(b - a)$. A corollary is that if $f'(x) > 0$ on an interval then f is strictly increasing there.

Another way of writing the Mean Value theorem is that

$$f(a + h) = f(a) + hf'(a + th)$$

for some $0 < t < 1$, where f is continuous and differentiable on the interval between a and $a + h$. There is a more precise "second mean value theorem" stating that

$$f(a + h) = f(a) + hf'(a) + \frac{h^2}{2}f''(a + t_1 h)$$

where $0 < t_1 < 1$ under the hypothesis that f is twice differentiable between x and $x + h$. There is an "n-th mean value theorem" also known as Taylor's theorem (with remainder term).