Lecture 10 Applications of the Mean Value theorem

Last time, we proved the mean value theorem:

**Theorem** Let $f$ be a function continuous on the interval $[a, b]$ and differentiable at every point of the interior $(a, b)$. Then there is $c \in (a, b)$ so that

$$f'(c) = \frac{f(b) - f(a)}{b - a}.$$

On first glance, this seems like not a very quantitative statement. There is a point $c$ in the interval $(a, b)$ where the equation holds, but we can’t use the theorem to guess exactly where that point $c$ is, and so it is hard for us to use the mean value theorem to obtain information about large scale changes in the function $f$ from the value of its first derivative.

But in fact, this objection is somewhat misleading. The mean value theorem is really the central result in Calculus, a result which permits a number of rigorous quantitative estimates? How does that work? The trick is to apply the mean value theorem, primarily on intervals where the derivative of the function $f$ is not changing too much. As it turns out, understanding second derivatives is key to effectively applying the mean value theorem. We will spend this lecture giving some examples.

I sloppily assigned a homework problem in which you were to prove that if $f$ was a function and $c$ a point so that the first derivative $f'(x)$ was defined everywhere in an interval having $c$ in its interior, and so that this first derivative function $f'$ was differentiable at $c$, then

$$f''(c) = \lim_{h \to 0} \frac{f(c + h) + f(c - h) - 2f(c)}{h^2}.$$ 

It is tempting to try to prove this just by comparing the first derivative of $f$ to the difference quotients

$$\frac{f(c + h) - f(c)}{h},$$

and

$$\frac{f(c) - f(c - h)}{h},$$

subtracting them and dividing by $h$. This certainly gives us the expression whose limit we’d like to take. But to relate it to the second derivative, we need a little more than the differential approximation for $f$, which only estimates $f(c + h) - f(c)$ to within $o(h)$. We need an estimate that is within $o(h^2)$ because of the $h^2$ in the denominator of the expression under the limit. We will get such an estimate by using the second derivative to get a differential approximation for the first derivative and then using the mean value theorem in quite small intervals. We proceed.
Theorem (Taylor approximation, version 1) Let \( f \) be a function which is continuous on an interval \( I \) having \( c \) on its interior and suppose that \( f'(x) \) is defined everywhere in \( I \). Suppose further that \( f''(c) \) is defined. Then for \( h \) sufficiently small that \([c, c+h] \subset I\), we have

\[
f(c+h) = f(c) + hf'(c) + \frac{h^2}{2} f''(c) + o(h^2).
\]

Proof We see from the differential approximation for \( f'(x) \) that

\[
f'(c+t) = f'(c) + tf''(c) + o(t),
\]

for \( t < h \). Since we restrict to \( t < h \), we can replace \( o(t) \) by \( o(h) \). (The equation above depends on both \( t \) and \( h \).) So we record:

\[
f'(c+t) = f'(c) + tf''(c) + o(h).
\]

Now our plan is to use the above expression for \( f' \) together with the mean value theorem on small interval to obtain a good estimate for \( f(c+h) - f(c) \). We should specify what these intervals are going to be. We will pick a natural number \( n \), which will be the number of equal pieces into which we divide the interval \([c, c+h]\). We define points \( x_j \) where \( j \) will run from 0 to \( n \) as follows:

\[
x_j = c + \frac{jh}{n}.
\]

We observe that we can calculate \( f(c+h) - f(c) \), which we are interested in, by understanding \( f(x_j) - f(x_{j-1}) \) for each \( j \) from 1 to \( n \). Precisely, we have

\[
f(c+h) - f(c) = \sum_{j=1}^{n} f(x_j) - f(x_{j-1}),
\]

since the sum telescopes to \( f(x_n) - f(x_0) = f(c+h) - f(c) \). Now we will understand each term in the sum using the mean value theorem on \([x_{j-1}, x_j]\). There is \( y_j \in (x_{j-1}, x_j) \) with

\[
\frac{h}{n} f'(y_j) = f(x_j) - f(x_{j-1}).
\]

Thus we rewrite the sum

\[
f(c+h) - f(c) = \sum_{j=1}^{n} \frac{h}{n} f'(y_j).
\]

Now we estimate \( f'(y_j) \) using the differential approximation, equation (1). We conclude

\[
f'(y_j) = f'(c) + (y_j - c) f''(c) + o(h).
\]

Now we use the fact that \( y_j \in (x_{j-1}, x_j) \) to estimate \( y_j - c = \frac{jh}{n} + O\left(\frac{h}{n}\right) \). Now we combine everything.

\[
f(c+h) - f(c) = \sum_{j=1}^{n} \left[ \frac{h}{n} f'(c) + \frac{jh^2}{n^2} f''(c) + O\left(\frac{h^2}{n^2}\right) + o\left(\frac{h^2}{n}\right) \right].
\]
Now we sum all the terms in the square bracket separately. The terms constant in \( j \) get multiplied by \( n \). Just the second term depends on \( j \) (linearly!) and we recall our old notation

\[
S_1(n) = \sum_{j=1}^{n} j,
\]

to write

\[
f(c + h) - f(c) = hf'(c) + h^2 \frac{S_1(n)}{n^2} f''(c) + O(h^2) + o(h^2).
\]

Observe that this equality holds for every choice of \( n \). We take the limit as \( n \) goes to infinity and obtain the desired result, remembering that

\[
\lim_{n \to \infty} \frac{S_1(n)}{n^2} = \frac{1}{2}.
\]

Let’s take a deep breath. A great deal happened in that argument. It might take a moment to digest. But at least you have the power to use the above Theorem to solve your homework problem.

A bit of further reflection shows that if we have enough derivatives we adapt the above argument to give us an estimate for \( f(c + h) - f(c) \) up to \( o(h^m) \) for any \( m \). Let’s do this. We adopt the notation that where defined, \( f^{(j)} \) denotes the \( j \)th derivative of \( f \).

**Theorem (Taylor approximation, version 2)** Let \( f \) be a function which is continuous on an interval \( I \) having \( c \) on its interior and suppose that \( f'(x), \ldots, f^{(m-2)}(x) \) are defined and continuous everywhere in \( I \). Suppose that \( f^{(m+1)} \) is defined everywhere on \( I \) and that \( f^{(m)}(c) \) is defined. Then for \( h \) sufficiently small that \([c, c + h] \subset I\), we have

\[
f(c + h) = f(c) + \sum_{k=1}^{m} \frac{h^k}{k!} f^{(k)}(c) + o(h^m).
\]

**Proof** We will prove this by induction on \( m \). We observe that the base case \( m = 2 \) is just the previous theorem. Thus we need only perform the induction step. We use the induction hypothesis to get an appropriate estimate for the first derivative anywhere in the interval \([c, c + h] \].

\[
f'(c + t) = f'(c) + \sum_{k=2}^{m} \frac{f^{(k)}(c)}{(k - 1)!} t^{k-1} + o(h^{m-1}).
\]

Now we proceed as in the proof of the previous theorem. We choose \( n \) and let \( x_j = c + \frac{h}{n} \). We observe that

\[
f(c + h) - f(c) = \sum_{j=1}^{n} f(x_j) - f(x_{j-1}),
\]
and we use the mean value theorem to find \( y_j \in (x_{j-1}, x_j) \) with

\[
f(x_j) - f(x_{j-1}) = f'(y_j) \frac{h}{n}.
\]

Now we use equation (2) to estimate \( f'(y_j) \) as before and we obtain

\[
f(c + h) - f(c) = \sum_{j=1}^{n} \left[ \sum_{k=1}^{m} \frac{h^k}{(k-1)!} f^{(k)}(c) j^{k-1} + O(\frac{h^k}{n^k}) \right] = 0 + o(\frac{h^n}{n}).
\]

Now summing in \( j \), we obtain

\[
f(c + h) - f(c) = \sum_{k=1}^{m} \frac{h^k}{(k-1)!} \frac{S_{k-1}(n)}{n^k} + O(\frac{h^k}{n}) + o(\frac{h^n}{n}).
\]

Letting \( n \) tend to infinity and using the fact that

\[
\lim_{n \to \infty} \frac{S_{k-1}(n)}{n^k} = \frac{1}{k},
\]

we obtain the desired result.

**Remarks**

1. You should be tempted to ask, does this mean that every function which is infinitely differentiable everywhere can be given as the sum of a power series. We have shown that if the function is \( n \) times differentiable near \( c \), then at \( c + t \) near \( c \), we have

\[
f(c + t) = T_{n,f,c}(t) + o(t^n),
\]

where \( T_{n,f,c} \) is the degree \( n \) Taylor approximation to \( f \) near \( c \). It is tempting to try to take the limit as \( n \to \infty \). This doesn’t work because of the definition of \( o(t^n) \). We just know that

\[
\lim_{t \to 0} \frac{f(c + t) - T_{n,f,c}(t)}{t^n} = 0,
\]

but for different \( t \), the rate at which the limit goes to 0 can differ substantially.

2. Am I not then pulling a fast one in light of 1. Didn’t I say in the proofs that

\[
f(c + t) = f(c) + tf'(c) + o(h),
\]

when in fact, it should be \( o(t) \) but with a different rate of convergence for each \( t \). No. But it’s because these estimates are all coming from the same limit. Suppose it weren’t true that

\[
f(c + t) = f(c) + tf'(c) + o(h).
\]
Then there is no $\delta > 0$ so that $|h| < \delta$ implies

$$|f(c + t) - f(c) - tf'(c)| \leq \epsilon h,$$

for every $t < h$. Then we could pick a sequence $t_j$ going to zero, for which the absolute value is greater than $\epsilon h$ and hence $\epsilon t_j$. This would violate the definition of the derivative at $c$.

3. What’s going on in these proofs? How come there are all these complicated sums showing up. Calculus isn’t really about sums is it? There must be some way we can make them all go away. In fact, we can by subsuming them into a definition. It’s a definition that’s coming up soon: the definition of the integral. What we really did in the proof of the first theorem is to calculate:

$$f(c + h) - f(c) = \int_c^{c+h} f'(x)dx = \int_c^{c+h} [f'(c) + xf''(c) + o(h)]dx.$$