Lecture 13: Inverse functions

Today, we’ll begin with a classical application of the calculus: obtaining numerical solutions to equations.

Situation: We would like to solve an equation

\[ f(x) = 0. \]

Here \( f \) is a function and we should imagine that its first and second derivatives are continuous in some interval. We approach this problem by what is usually called Newton’s method.

First we make an initial guess \( x_0 \). Probably, we are not too lucky and

\[ f(x_0) \neq 0. \]

Thus what we do is that we calculate \( f'(x_0) \). We obtain the linear approximation

\[ f(x) \approx f(x_0) + f'(x_0)(x - x_0). \]

We solve for the \( x_1 \) which makes the linear approximation equal to zero. That is

\[ x_1 = x_0 - \frac{f(x_0)}{f'(x_0)}. \]

Usually, we are still not lucky and

\[ f(x_1) \neq 0. \]

Thus we obtain \( x_2 \) from \( x_1 \) in the same way, and so on. For general \( j \), we get

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

A question we should ask, which is extremely practical, is how fast does the sequence \( \{f(x_j)\} \) converge to 0. If we knew that, we should really know how many steps of Newton’s method we should have to apply to get a good approximation to a zero.

This is a job for the Mean Value Theorem. In the interval \([x_0, x_1]\) (or \([x_1, x_0]\) depending on which of \( x_0 \) or \( x_1 \) is greater) there is a point \( c \) so that

\[ f'(c) = \frac{f(x_1) - f(x_0)}{x_1 - x_0}. \]

Thus

\[ f(x_1) - f(x_0) = -\frac{f(x_0)}{f'(x_0)} f'(c). \]
Suppose that in our interval where all the action is taking place we have an upper bound $M$ for $|f''(x)|$. Then

$$
|f(x_1)| \leq |f(x_0)||1 - \frac{f'(c)}{f'(x_0)}|
$$

$$
= \frac{|f(x_0)|}{|f'(x_0)|}|f'(x_0) - f'(c)|
$$

$$
\leq \frac{M|f(x_0)|}{|f'(x_0)|}|x_0 - c|
$$

$$
\leq \frac{M|f(x_0)|^2}{|f'(x_0)|^2}.
$$

Suppose further that in our interval, we have a lower bound on the absolute value of the derivative,

$$
|f'(x)| > \frac{1}{K}.
$$

Then we conclude

$$
|f(x_1)| \leq K^2M|f(x_0)|^2.
$$

Moreover, we have the same thing for every $j$.

$$
|f(x_j)| \leq K^2M|f(x_{j-1})|^2.
$$

To get any benefit from these inequalities, we must have $|f(x_0)| < \frac{1}{K^2M}$. In other words, our initial guess should be pretty good. But, if this is true, and $|f(x_0)| = \frac{r}{K^2M}$ with $r < 1$, we get from these inequalities:

$$
|f(x_1)| \leq \frac{r^2}{K^2M},
$$

$$
|f(x_2)| \leq \frac{r^4}{K^2M},
$$

and in general

$$
|f(x_j)| \leq \frac{r^{2^j}}{K^2M}.
$$

This is a pretty fast rate of convergence. It is double exponential.

We encapsulate all of this as a theorem:
Theorem Let $I$ be an interval and $f$ a function which is twice continuously differentiable on $I$. Suppose that for every $x \in I$, we have $|f''(x)| < M$ and $|f'(x)| > \frac{1}{K}$. Then if we pick $x_0 \in I$, and we define the sequence $\{x_j\}$ by

$$x_j = x_{j-1} - \frac{f(x_{j-1})}{f'(x_{j-1})}.$$ 

Then if we assume that each $x_j$ is in $I$ and that $|f(x_0)| < \frac{rJ}{K^2M}$, then we obtain the estimate

$$|f(x_j)| \leq \frac{r^{2^j}}{K^2M}.$$ 

Just to know an equation has a solution, we often need a lot less. We say that a function $f$ is strictly increasing on an interval $[a, b]$ if for every $x, y \in [a, b]$ with $x < y$, we have $f(x) < f(y)$.

Theorem Let $f$ be continuous and increasing on $[a, b]$. The $f$ has an inverse uniquely defined from $[f(a), f(b)]$ to $[a, b]$.

Proof For $c \in [f(a), f(b)]$, we want $x$ with $f(x) = c$. Since

$$f(a) \leq c \leq f(b),$$

and $f$ is continuous, we have that there exists such an $x$ by the Intermediate Value theorem. Because $f$ is strictly increasing, this $c$ is unique.

With $f$ as above, if $f$ is differentiable at a point $x$ with nonzero derivative, we will show that its inverse is differentiable at $f(x)$.

Theorem Let $f$ be a strictly increasing continuous function on $[a, b]$. Let $g$ be its inverse. Suppose $f'(x)$ is defined and nonzero for some $x \in (a, b)$. Then $g$ is differentiable at $f(x)$ and

$$g'(f(x)) = \frac{1}{f'(x)}.$$ 

Proof By the differentiability of $f$ at $x$, we get

$$f(y) = f(x) + f'(x)(y - x) + o(y - x).$$

Now we solve for $y - x$.

$$(y - x) = \frac{f(y) - f(x)}{f'(x)} + o(y - x).$$
We rewrite this as
\[ g(f(y)) - g(f(x)) = \frac{f(y) - f(x)}{f'(x)} + o(y - x). \]

Finally we simply observe that anything that is \( o(y - x) \) is also \( o(f(y) - f(x)) \) since
\[ \lim_{y \to x} \frac{f(y) - f(x)}{y - x} = f'(x). \]

Thus we have obtained our desired result.

**Example** The function \( e^x \) is strictly increasing on the whole real line. Thus it has an inverse from the positive reals to the real line. We call this inverse function \( \log \). We have
\[ \frac{d}{dx}(\log x) = \frac{1}{x}. \]

**Application: logarithmic differentiation**

An important fact about \( \log \) is that
\[ \log(ab) = \log a + \log b. \]

This gives rise to a nice way of thinking of the product and quotient rules (and generalizations.)

Instead of calculating \( \frac{d}{dx}(fg) \), we calculate \( \frac{d}{dx}(\log fg) \).
We get
\[ \frac{d}{dx}(\log fg) = \frac{\frac{d}{dx}(fg)}{fg}, \]
but on the other hand
\[ \frac{d}{dx}(\log fg) = \frac{d}{dx}(\log f + \log g) = \frac{f'}{f} + \frac{g'}{g}. \]

Solving we get
\[ \frac{d}{dx}(fg) = (\frac{f'}{f} + \frac{g'}{g})fg = f'g + g'f. \]

The same idea works for arbitrarily long products and quotients.