Math 5615 Honors: Newton's Method

Sasha Voronov

University of Minnesota

December 4, 2020

Contraction Mapping

Definition

A mapping $T : X \to X$ of a metric space X with metric d to itself is a *contraction mapping* if there is a number $0 < r < 1$ such that $d(T(x), T(y)) \leq r\, d(x, y)$ for all $x, y \in X$. Such a constant r is called a *contraction constant for T*.

fixed point: $T(x^*) = x^*$

Theorem (Contraction Mapping Theorem)

A contraction mapping $T : X \to X$ of a complete metric space X has a unique fixed point x^*. Moreover, if r is a contraction constant for T, then given any $x_0 \in X$, the iteration $x_{k+1} = T(x_k)$, $k = 0, 1, 2, 3, \ldots$ defines a sequence $\{x_k\}$ that converges to x^*, and for each k, we have

$$d(x_k, x^*) \leq \frac{r^k}{1 - r} d(x_1, x_0).$$

Using the Contraction Mapping Theorem to Solve $f(x) = 0$

Suppose we want to solve an equation $f(x) = 0$ for a continuous function $f : \mathbb{R} \to \mathbb{R}$ on a closed interval $I \subset \mathbb{R}$, such as $x^3 - x - 1 = 0$ on $[1, \infty)$. We do not know a formula for finding a root (well, actually, there is one, but if the equation were $x^5 - x - 1 = 0$, there would be none), so we might try to compute the root approximately.

Idea: create a contraction mapping $g : [a, b] \to [a, b]$, making sure that $f(x) = 0$ has a solution on $[a, b] \subset I$, so that $f(x) = 0 \Leftrightarrow g(x) = x$. Then if we take any $x_0 \in [a, b]$ and do iterations

$$x_{n+1} = g(x_n),$$

x_n will be approaching the true soluition x^* of $g(x) = x$ as $n \to \infty$.

Newton's Method

The only thing left is to find a closed insteval $[a, b] \subset I$ and construct that $g(x)$ so as

1. $g(x) = x \Leftrightarrow f(x) = 0$,
2. $g : [a, b] \to [a, b]$ is a contraction mapping:

$$|g(x) - g(y)| \leq r|x - y| \qquad \text{for some } r, 0 < r < 1.$$

1 is easy: $f(x) = x^3 - x - 1 = 0 \Leftrightarrow x = x^3 - 1 = g_1(x)$, but it is not a contraction anywhere near the expected root of $f(x)$ between 1 and 2 ($f(1) = -1$, $f(2) = 5$): $g_1'(x) = 3x^2 > 1$ and $g_1(x) - g_1(y) = g_1'(c)(x - y)$, $c \in (a, b)$, whence there is no way to have $|g_1(x) - g_1(y)| \leq r|x - y|$ for $r < 1$ anywhere on $[1, 2]$.

Newton's method: Try $g(x) := x - f(x)/f'(x)$, if $f'(x) \neq 0$. Then the iterations will look like: $x_{n+1} := g(x_n)$,

$$x_{n+1} := x_n - \frac{f(x_n)}{f'(x_n)}.$$

Geometric Interpreation

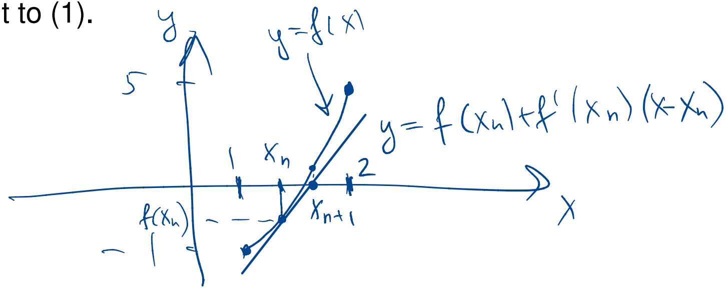
Note that

$$x_{n+1} := x_n - \frac{f(x_n)}{f'(x_n)} \qquad (1)$$

is the x intercept of the tangent line $y = f(x_n) + f'(x_n)(x - x_n)$ to the graph of $y = f(x)$ at $(x_n, f(x_n))$. Indeed,

$$f(x_n) + f'(x_n)(x_{n+1} - x_n) = 0$$

is equivalent to (1).



Newton's Method Justified

If we somehow knew that at the expected solution x^* of $f(x) = 0$, $f'(x^*) \neq 0$, for example, if we knew $f'(x) > 0$ on I, then Condition 1 for $g(x)$ would be satisfied:

$$g(x) = x - \frac{f(x)}{f'(x)} = x \iff \frac{f(x)}{f'(x)} = 0 \iff f(x) = 0$$

In our case, $f'(x) = 3x^2 - 1 > 0$ for $x \geq 1$.

What about Condition 2, contraction mapping, for g? $x^*: f(x^*) = 0$.

Theorem

If f has $f'(x)$ and $f''(x)$ on an open interval containing x^*, $f''(x)$ is continuous there, and $f'(x^*) \neq 0$, then here exists a $\delta > 0$ such that $|g(x) - g(y)| \leq \frac{1}{2}|x - y|$ for $x, y \in [x^* - \delta, x^* + \delta]$.

(works also for any r, $0 < r < 1$, with suitable choice of δ.)

Proof of Newton's Method

If δ is so small that $f'(x) \neq 0$ on $[x^*-\delta, x^*+\delta]$, then $g(x) = x - \frac{f(x)}{f'(x)}$ will be well-defined there.

$|g(x) - g(y)| = |g'(c)| \cdot |x-y|$ for some c between x and y, $x, y \in [x^*-\delta, x^*+\delta]$. Let's compute

$$g'(x) = \left(x - \frac{f(x)}{f'(x)}\right)' = 1 - \frac{f'(x)f'(x) - f''(x)f(x)}{(f'(x))^2}$$

$$= \frac{f''(x)f(x)}{f'(x)^2}.$$

At x^*, $f(x^*)=0$, therefore, $g'(x^*)=0$. By continuity of g' at x^*, for δ suff. small, $|g'(x)| \leq \frac{1}{2}$ for $|x-x^*| < \delta$. Thus, for x,y there

$|g(x)-g(y)| \leq \frac{1}{2}|x-y|$. $\square$

[redacted]

Comment to justify Newton's method:

For small enough δ,

g maps $[x^* - \delta, x^* + \delta]$ to itself, $[x^* - \delta, x^* + \delta]$. Indeed,

$$g(x^*) = x^*, \quad |g(x) - \underbrace{g(x^*)}_{= x^*}| \le \frac{1}{2}|x - x^*| \le \frac{1}{2}\delta$$

$$g(x) \in [x^* - \tfrac{1}{2}\delta, x^* + \tfrac{1}{2}\delta] \subset [x^* - \delta, x^* + \delta]$$