Ch 2.3, part 1: Newton's Method

Thursday, September 18, 2025 4:49 PM

Newton's method (a.k.a. Newton-Raphson method)

Fundamental, ubiquitous algorithm (partly because it extends beyond 1D root-finding to multi-dimensional root-finding and optimization)

Derivation

In equations, recall we want to solve $f(p)=0$

General technique in STEM: replace problem with a simpler approximation

... in particular, linearization

Do 1st order Taylor Series, $f(p) \approx f(p_0) + f'(p_0)\cdot(p-p_0)$, good approximation when $|p-p_0|$ small

so, solve $0 = f(p_0) + f'(p)\cdot(p-p_0)$ for p,

i.e., $p = p_0 - \dfrac{f(p_0)}{f'(p_0)}$

but this was only approximate, so do this repeatedly:

$$\boxed{p_k = p_{k-1} - \frac{f(p_{k-1})}{f'(p_{k-1})} \qquad \text{Newton's Method}}$$

[Aside: next week you'll see the multivariate version,

$$\vec{p}_k = \vec{p}_{k-1} - J(\vec{p}_{k-1})^{-1}\cdot F(\vec{p}_{k-1})$$

for solving $F(\vec{p}) = \vec{0}$, J is Jacobian of F]

See https://www.geogebra.org/m/DGFGBJyU
for a great interactive demo of Newton's method

In pictures,

Approximate the function $f(x)$ using its tangent line, and find a zero of the tangent line (which is easy since it's a line)

When the tangent line is a good approximation of the function, Newton's method will converge rapidly

Ch 2.3: Newton, p. 2

Thursday, September 18, 2025 4:50 PM

Convergence, part 1

⚠ In general, Newton's method is <u>not</u> "globally convergent". That is, you can't start at <u>any</u> starting p_0. As we'll see later, this can be partially remedied, eg. with safeguarding and hybrid methods.

<u>Thm 2.6</u> Local convergence of Newton's Method

Let $f \in C^2([a,b])$ and p is a root of f $(f(p)=0)$ inside (a,b), and p is <u>simple root</u> $(f'(p) \neq 0)$, then if Newton's method is initialized close enough to p, then the sequence (p_k) generated by Newton's method will converge to p

(i.e., $\exists\, \delta > 0$ s.t. $(\forall\, p_0$ with $|p_0 - p| < \delta)$, $p_k \to p$).

<u>proof</u>

We can rewrite $p_{k+1} = \underbrace{p_k - \frac{f(p_k)}{f'(p_k)}}_{g(p_k)}$ as $p_{k+1} = g(p_k)$

i.e., we've converted root-finding $(f(p)=0)$ into a fixed-point problem,

$f(p)=0$ iff $-\frac{f(p)}{f'(p)} = 0$ (since we assume $f'(p) \neq 0$ at the root) $\quad p = g(p)$, $g(p) := p - \frac{f(p)}{f'(p)}$

iff $p - \frac{f(p)}{f'(p)} = p$.

So, use <u>contraction mapping theorem</u>. We'll find an interval $x \in (p-\delta, p+\delta)$ such that

1. g maps this interval into this interval
2. g is a contraction, ie., $|g'(x)| \leq L < 1$ on this interval.

First, is g <u>well-defined</u>? $\frac{1}{f'(x)}$ is a problem if $f'(x)=0$. We assume $f'(p) \neq 0$ at the true root p, and by continuity, there's also some region $(p-\delta_1, p+\delta_2)$ where $f'(x) \neq 0$. So g is well-defined and continuous on this region. In fact, since $f \in C^2$, $g \in C^1$ on this region.

Now, show g is a contraction, i.e., want $|g'(x)|$ small.

Well. $g'(p) = 0$ in fact. To see this, use the quotient rule:

$$g'(x) = \frac{f(x) f''(x)}{(f'(x))^2} \quad \text{so} \quad g'(p) = \frac{\overbrace{\cancel{f(p)}}^{=0 \text{ since a root}} f''(p)}{\underbrace{(f'(p))^2}_{\neq 0 \text{ since a simple root}}} = 0$$

Quotient Rule: $(f/g)' = \frac{f'g - fg'}{g^2}$

so... $g(x) = x - \frac{f(x)}{f'(x)}$, $g'(x) = 1 - \left[\frac{f'(x)f'(x)}{f'(x)^2} - \frac{f(x)f''(x)}{f'(x)^2}\right]$

cancel

Ch 2.3: Newton, p. 3

Thursday, September 18, 2025 4:50 PM

thus by continuity of g', there's an interval around p where $g'(x)$ is almost 0, i.e., $\forall L > 0$ (in particular, choose some $L < 1$) $\exists\, \delta_2$ s.t. $\forall x \in [p - \delta_2, p + \delta_2]$, $|g'(x)| \leq L$. ✓

Let $\delta = \min(\delta_1, \delta_2)$,

$I = [p + \delta, p - \delta]$. It remains to show g maps I into I.

We'll use the fact that I is symmetric about p and that g is a contraction on I.

Take $x \in I$, then since $p \in I$ also, and g is a contraction on I,

$$|g(x) - p| \underset{\text{since } p = g(p)}{=} |g(x) - g(p)| \leq L\,|x - p| \quad \Big\} \text{ since } g \text{ is a contraction on } I$$
$$< |x - p|$$
$$< \delta \quad \text{since } x \in I = [p - \delta, p + \delta]$$

and $|g(x) - p| < \delta$ means $g(x) \in I$

So, we can apply fixed-pt thm ("contraction mapping" / "Banach fixed pt.") to get convergence. $\square$

Convergence, part 2 (rate, ie., local quadratic convergence)

Helper Theorem for generic fixed-pt. iteration to solve $p = g(p)$

Thm 2.9 Let p be a solution to $x = g(x)$, and suppose $\underbrace{g'(p) = 0}_{\text{as in Newton's method}}$ and g'' is continuous and bounded $|g''(x)| < M$ on some open interval $(p - \delta_1, p + \delta_2)$, $\delta_1, \delta_2 > 0$. Then $\exists\, \delta > 0$ s.t. if $|p_0 - p| \leq \delta$, $p_k := g(p_{k-1})$, then p_n converges quadratically to p and $\exists\, K$ st. $(\forall k \geq K)$ $|p_{k+1} - p| < \frac{M}{2} |p_k - p|^2$. *

proof As in previous theorem, close enough to p, we have $k < 1$ with $|g'(x)| \leq k$ near p, and g maps $[p - \delta, p + \delta]$ into $[p - \delta, p + \delta]$. (all due to continuity properties)

Now, Taylor expand around p:

$$g(x) = \underbrace{g(p)}_{p \text{ fixed pt.}} + \underbrace{\cancel{g'(p)(x - p)}}_{g'(p) = 0 \text{ by assumption}} + g''(\xi)/_2\,(x - p)^2, \quad \xi \text{ between } x \text{ and } p$$

Choosing $x = p_k$, $\quad p_{k+1} = g(p_k)$

$$\text{So} \quad p_{k+1} - p = g''(\xi)/_2\,(p_k - p)^2 \quad \text{or } \xi_k$$

In previous theorem, under these conditions, $p_k \to p$.

What about ξ? ξ_k between p and p_k, so by squeeze thm., $\xi_k \to p$ also.

Thus

$$\lim_{k \to \infty} \frac{|p_{k+1} - p|}{|p_k - p|^2} = \lim_{k \to \infty} \left| \frac{g''(\xi_k)}{2} \right| = \left| \frac{g''(p)}{2} \right| < M/_2$$

which means $p_k \to p$ quadratically. $\square$

Ch 2.3: Newton, p. 4

Thursday, September 18, 2025 4:51 PM

Recall our earlier discussion of simple roots

(Thm 2.11) A root p of $f \in C^1([a,b])$, $p \in (a,b)$, is called simple if $f'(p) \neq 0$.

(Careful with notation:

root finding $f(p) = 0$, then $f'(p) \neq 0$ is good since it means a simple root.

fixed-pt. $g(p) = p$, then $g'(p) = 0$ is good since it means fast convergence.)

So, putting it altogether

Thm If p is a simple root of f, then if Newton's method is initialized sufficiently close to p, it will converge, and at a quadratic rate.

Note:

For methods that can be cast as fixed-pt. iterations $p_{n+1} = g(p_n)$,

where $g(x) = x - \phi(x) f(x)$, need $\phi(p) \neq 0$,

and for quadratic convergence, need $g'(p) = 0$, which is true iff $\underbrace{\phi(p) = 1/f'(p)}$
(or any superlinear convergence)

Newton's method defines $\phi(x) = 1/f'(x)$ $\forall x$ to ensure this.