In a nutshell: Newton's method for finding extrema

Given a continuous and differentiable real-valued function f of a real variable with one initial approximation of an extremum x_0 where $f^{(2)}(x_0) \neq 0$. If the derivative is already zero, we have already found an extremum. This algorithm uses iteration, Taylor series and solving a trivial linear equation to approximate an extremum.

Parameters:

$\varepsilon_{\text{step}}$ The maximum error in the value of the root cannot exceed this value.
ε_{abs} The value of the function at the approximation of the root cannot exceed this value.
N The maximum number of iterations.

1. Let $k \leftarrow 0$.
2. If $k > N$, we have iterated N times, so stop and return signalling a failure to converge.
3. The next approximation to the extremum will be the root of the 1st-order Taylor series of the derivative at the point x_k, where the 1st-order Taylor series forms a linear polynomial tangent to the function at $(x_k, f^{(1)}(x_k))$.

   Let $x_{k+1} \leftarrow x_k - \frac{f^{(1)}(x_k)}{f^{(2)}(x_k)}$.

   a. If x_{k+1} is not a finite floating-point number, so return signalling a failure to converge.
   b. If $|x_{k+1} - x_k| < \varepsilon_{\text{step}}$ and $|f(x_{k+1}) - f(x_k)| < \varepsilon_{\text{abs}}$, return x_{k+1}.
4. Increment k and return to Step 2.

If this method converges, then if $f^{(2)}(x_{k+1}) > 0$, it is a minimum; if $f^{(2)}(x_{k+1}) < 0$, it is a maximum; but if $f^{(2)}(x_{k+1}) \approx 0$, it could be a maximum, a minimum, or a saddle point.

Convergence

If h is the error, it can be show that the error decreases according to $O(h^2)$. This technique is not guaranteed to converge if there is a root, for the denominator could be arbitrarily small, causing the next approximation to be arbitrarily far from the previous approximation.

Error analysis

The rate of convergence can by found by writing down the 1st-order Taylor series

$$f^{(1)}(r) = f^{(1)}(x_k) + f^{(2)}(x_k)(r - x_k) + \frac{1}{2} f^{(3)}(\xi)(r - x_k)^2$$

and noting that if r is a the x-value at which the extremum occurs, then $f^{(1)}(r) = 0$, and thus we can write this as

$$r - \left(x_k - \frac{f^{(1)}(x_k)}{f^{(2)}(x_k)} \right) = -\frac{1}{2} \frac{f^{(3)}(\xi)}{f^{(2)}(x_k)} (r - x_k)^2$$

and observing that the term in the brackets is x_{k+1}, so

$$r - x_{k+1} = -\frac{1}{2} \frac{f^{(3)}(\xi)}{f^{(2)}(x_k)} (r - x_k)^2$$

Thus, the error of the next approximation is a constant $-\frac{1}{2}\frac{f^{(3)}(\xi)}{f^{(2)}(x_k)}$ multiplied by the error of the previous approximation squared.

Acknowledgement: Jakob Koblinsky noted that the error analysis section continued to refer to *r* as being a root. This has been corrected.