

Minimization with derivatives

Suppose we have not only $f(x)$ but $f'(x)$ available to us and we want to minimize $f(x)$ on $[a, b]$.

Remark. If you are approximating $f'(x)$ using $f(x)$ evaluations using one of the methods that we've covered before, you DON'T have $f'(x)$ available (roundoff error will kill you, and you'll be better off with a ~~straight~~ straight Brent's method code - see Mathematica demo).

②

So suppose we have a triple

a, b, c such that $f(a) > f(b) > f(c)$. Then b is a local minimum. We want to show that b is a global minimum. To do this we will use the Intermediate Value Theorem. This theorem states that if f is continuous on $[a, c]$ and $f(a) < f(c)$, then there exists some $d \in (a, c)$ such that $f(d) = f(a)$.



and we know there's a min in (a, c) because $f(b) < f(a), f(c)$.

Observe that the sign of $f'(b)$ can be used to predict whether (a, b) or (b, c) contains the min.

Further if we know (say)

$$f(b), f'(b) < 0$$

$$f(c), f'(c) > 0$$



(3)

we can guess the root of f' between b and c using the secant line method.

(The secant method converges superlinearly, so we get superlinear convergence near the min.)

If our prediction for the root of f' is outside (b, c) , or too close to the endpoints to evaluate stably, we will just bisect (b, c) .

This gives you the derivative form of Brent's method (due to Brent).

(mathematica demo)

You can read about this in
the excerpt from "Numerical
Recipes" (Section 10.3) posted
on the course page.