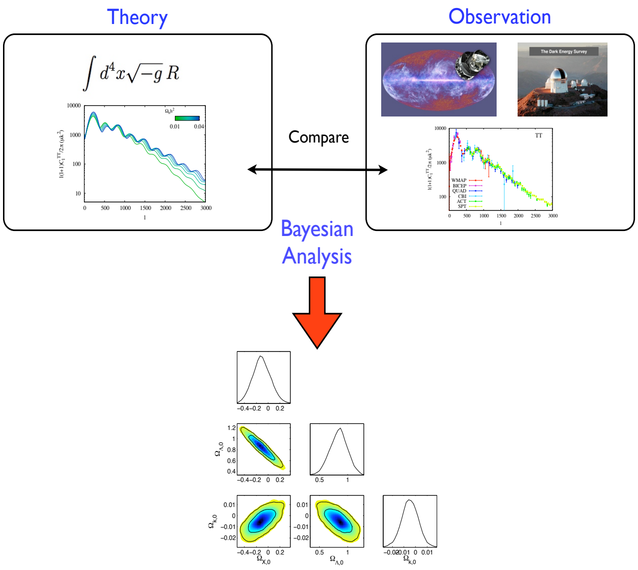
Numerical Methods



José-Alberto Vázquez

ICF-UNAM / Kavli-Cambridge

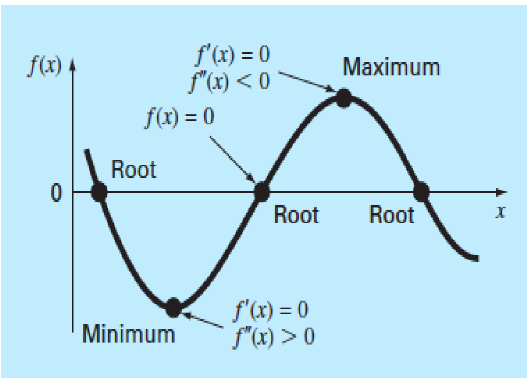
In progress

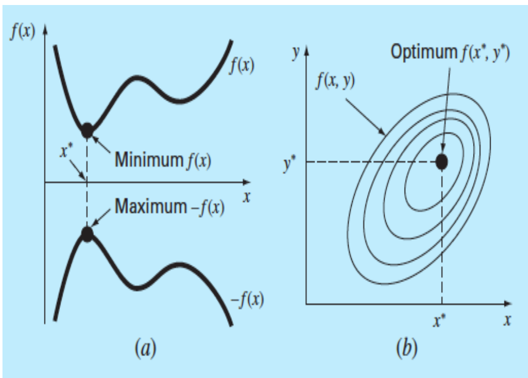
August 12, 2021

1

Optimization

Root location and optimization are related in the sense that both involve guessing and searching for a point on a function. In fact, some optimization methods seek to find an optima by solving the root problem: $f'(x) = 0$. Optimization typically deals with finding the 'best result', or optimum solution, of a problem.





Distinguishing a global from a local extremum can be a very difficult problem for the general case. In some problems (usually the large ones), there may be no practical way to ensure that you have located a global optimum.

Single-variable optimization has the goal of finding the value of x that yields an extremum, either a maximum or minimum of $f(x)$.

1.1 Golden-Section search

As with bisection, we can start by defining an interval that contains a single answer. That is, the interval should contain a single maximum, and hence is called unimodal.

Specifically, if in the neighborhood of the minimum we can find three points $x_0 < x_1 < x_2$ corresponding to $f(x_0) > f(x_1) < f(x_2)$, then there exists a minimum between x_0 and x_2. To search for this minimum, we can choose another point x_3 between x_1 and x_2 as shown in the figure below, with the following two possible outcomes:

If $f(x_3) = f_{3a} > f(x_1)$, the minimum is inside the interval $x_3 - x_0 = a + c$ associated with three new points $x_0 < x_1 < x_3$, i.e., x_2 is replaced by x_3. If $f(x_3) = f_{3b} < f(x_1)$, the minimum is inside the interval $x_2 - x_1 = b$ associated with three new points $x_1 < x_3 < x_2$, i.e., x_0 is replaced by x_1.

In either case, the new search interval $x_3 - x_0 = a + c$ or $x_2 - x_1 = b$ is smaller than the old one $x_2 - x_0 = a + b$, i.e., such an iteration is guaranteed to converge.

We therefore choose x_3 in such a way that the two resulting search intervals will always be the same:

$$a + c = b, \qquad c = b - a$$

The search interval should always be partitioned into two sections with the same ratio:

$$\frac{a+b}{b} = \frac{a+c}{a} = \frac{b}{a} \quad \text{or} \quad \frac{a+b}{b} = \frac{b}{b-c} = \frac{b}{a}$$

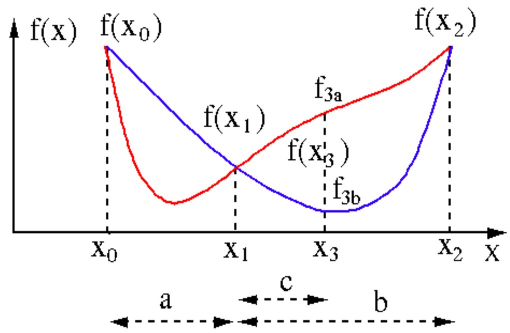
we get

$$a^2 + ab - b^2 = 0$$

solving this quadratic equation for a in terms of b we get

$$a = \frac{-1 \pm \sqrt{5}}{2} b = (0.618, -1.618) b$$

This is the golden ratio between the two sections a and b of the search interval.



1.2 Newton's

A similar open approach can be used to fi nd an optimum of $f(x)$ by defining a new function, $g(x) = f'(x)$. Thus, we can use the following (py: do it)

$$x_{i+1} = x_i - \frac{f'(x_i)}{f''(x_i)}$$

http://scipy-lectures.org/advanced/mathematical_optimization/index.html

1.3 Multidimentional Optimization

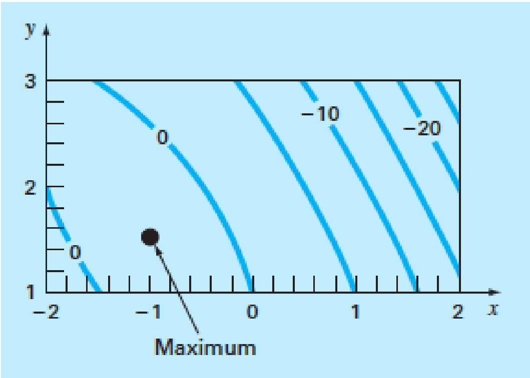
The approaches that do not require derivative evaluation are called nongradient, or direct, methods. Those that require derivatives are called gradient, or descent (or ascent), methods.

1.4 Direct Methods

1.4.1 Random Search

This simple brute force approach works even for discontinuous and nondifferentiable functions. Furthermore, it always finds the global optimum rather than a local optimum. Its major shortcoming is that as the number of independent variables grows, the implementation effort required can become onerous.

More sophisticated search techniques are available. These are heuristic approaches that were developed to handle either nonlinear and/or discontinuous problems that classical optimization cannot usually handle well, if at all. Simulated annealing, tabu search, artificial neural networks, and genetic algorithms are a few. The most widely applied is the genetic algorithm. (here: ver



notas/art geneticos)

(py: do grid, notas MCMC)

1.5 Gradient Methods

Gradient methods explicitly use derivative information to generate efficient algorithms to locate optima. – The Gradient, the Hessian.

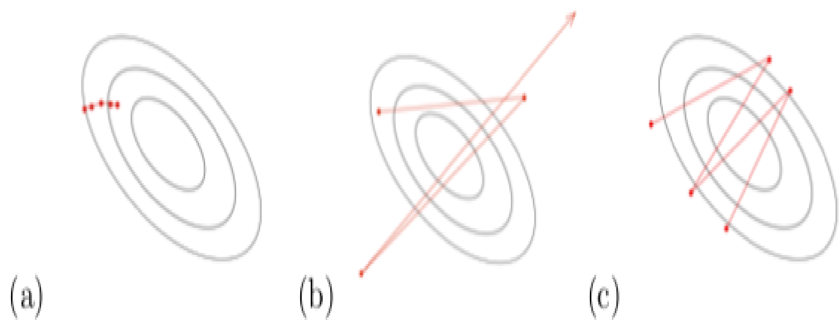
1.5.1 Steepest Ascent Method

You could walk a short distance along the gradient direction. Then you could stop, reevaluate the gradient and walk another short distance. By repeating the process you would eventually get to the top of the hill.

1. determining the 'best' direction to search
2. determining the 'best value' along that search direction

(here: ver notas ANN)

(hw: $f(x, y) = -8x + x^2 + 12y + 4y^2 - 2xy$ using initial guesses $x = 0$ and $y = 0$.)


(a)
(b)
(c)