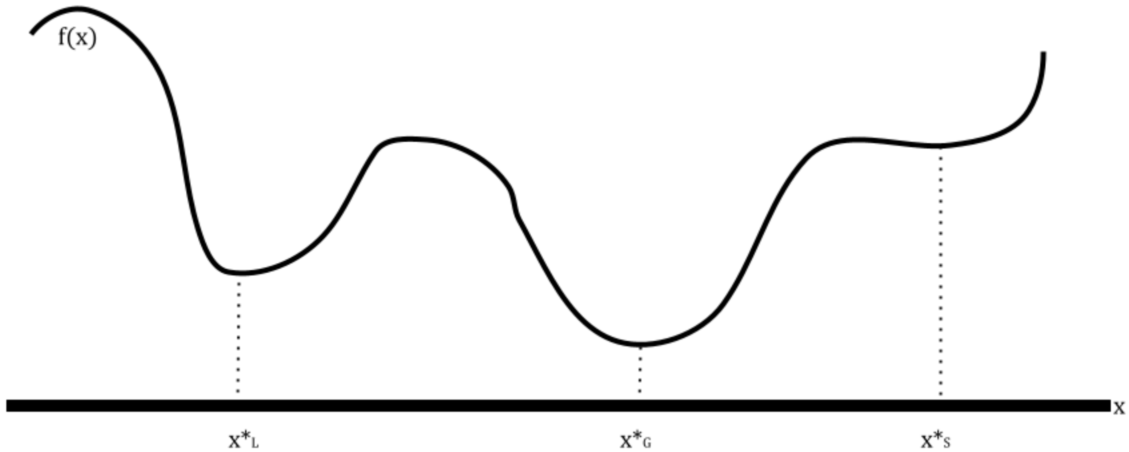
1 One-Variable Optimization (Section 11.1)

Suppose we have a nonlinear function f represented by the graph below,



with the points

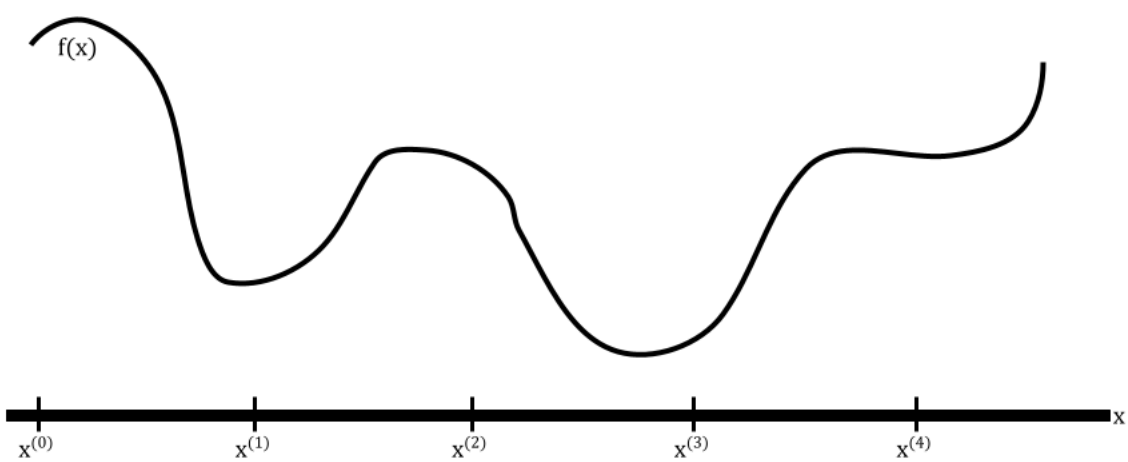
- x_L^*: the local minimum, or the smallest point of $f(x)$ in an open neighborhood around x_L^*,
- x_G^*: the global minimum (and also a local minimum), or the smallest minimum across the entire function, and
- x_S^*: where f increases or decreases on either side of x_S^*.

We're interested in the local minimum. More specifically, the goal is to find the minimum of a nonlinear function $f(x)$ (and we're fine with a local minima).

1D optimization is also relevant for $\mathbb{R}^m$. If $F : \mathbb{R}^m \to \mathbb{R}$, then we can define a line to be $\{u + tv : t \in \mathbb{R}\}$, where $u, v \in \mathbb{R}$. Then, for a fixed $\vec{u}$ and $\vec{v}$, we can find $F(u + tv) = f(t)$. The search to find the minimum depends on what information of f is available. In particular, whether we have access to f' or not.

1.1 Illustrative Strategy (Refining Search)

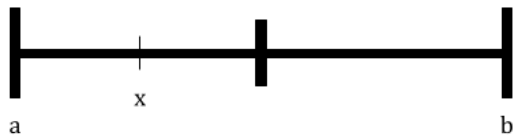
Suppose we have a bound $|f'(x)| \leq M$, where M can be any value greater than or equal to the maximum of $f'(x)$ in the given bounds. The idea is that we have $x^{(k)} = hk$, with h equally spaced intervals (step size) and for $k = 0, 1, 2, \ldots$. This gives us something like



Suppose we consider a particular interval $[a, b]$ with $a < b$. Then,

$$f(x) \geq \min\{f(a), f(b)\} - \frac{1}{2}(b-a)M,$$

where we derived the latter part from the Mean Value Theorem. To see why this works, suppose x is to the left of the midpoint between $[a, b]$. Then,



then, by the Mean Value Theorem,

$$f(x) - f(a) = f'(\xi)(x-a)$$

and

$$f(x) - f(a) \geq -M\frac{1}{2}(b-a) \implies f(x) \geq f(a) - \frac{1}{2}(b-a)M.$$

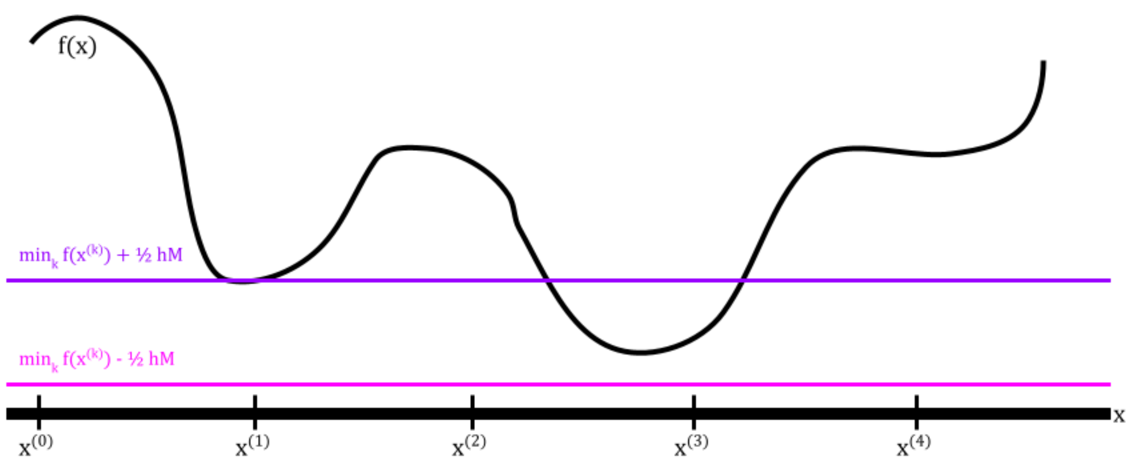
In any case, if we consider all sampled points, then[1]

$$\min_k f(x^{(k)}) \geq \inf_x f(x) \geq \underbrace{\min_k f(x^{(k)}) - \frac{1}{2}hM}_{\text{Lower Bound}}.$$

Suppose we consider the interval $[x^{(j)}, x^{(j+1)}]$ for refinement. We then want to consider the inequality when considering the refined search:

$$\min_k f(x^{(k)}) + \frac{1}{2}hM \geq \min\{f(x^{(j)}), f(x^{(k+1)})\},$$

Visually, these combined ideas would look like

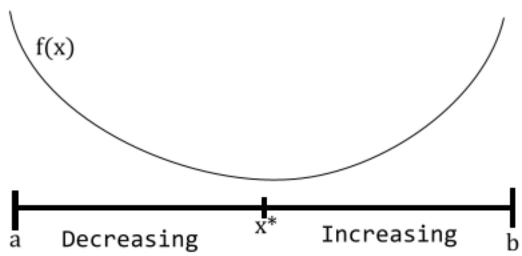


Notice how, for example, $x^{(2)}$ and $x^{(3)}$ has a minimum. So, we would consider this interval in our refined search.

[1]We can roughly think of inf as the "true minimum."

1.2 No Derivative Strategy

Suppose we do not have derivative information. Then, an assumption we can make is that f is unimodal (i.e., one minimum). Such an example is



For something like this, we might consider the **golden section search**. That is,

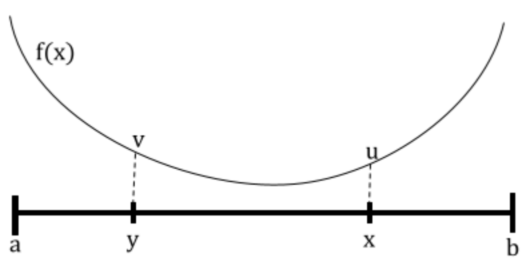
$$r^2 = 1 - r$$

$$r = \frac{1}{2}\left(\sqrt{5} - 1\right) \approx 0.6180\ldots.$$

This method is similar to the bisection method. In particular, we have

$$x = a + r(b - a) \qquad y = a + r^2(b - a).$$

This gives us



The search continues based on the values of u and v. In particular,

- if $v < u$, then we want to update the right bracket.
- if $v \geq u$, then we want to update the left bracket.

This can be modeled into an algorithm which takes the following inputs:

- `a`: the first endpoint.
- `b`: the second endpoint.
- f: the function.
- ϵ: the tolerance.
- M: The maximum number of iterations.

Algorithm 1 Golden Section Search

```
function GoldenSectionSearch(a, b, f, ϵ, M)
    x ← a + r(b − a)
    y ← a + r^2(b − a)
    u ← f(x)
    v ← f(y)
    for k ← 1 to M do
        if v < u then
            b ← x
            x ← y
            u ← v
            y ← a + r^2(b − a)
            v ← f(y)
        else
            a ← y
            y ← x
            v ← u
            x ← a + r(b − a)
            u ← f(x)
        end if
        if |b − a| < ϵ then
            break
        end if
    end for
end function
```