Chong, Edwin Kah Pin., and Stanislaw H. Żak. *An Introduction to Optimization*. Fourth edition., John Wiley & Sons, Inc., 2013.

2 ONE-DIMENSIONAL SEARCH METHODS

2.1 INTRODUCTION

- One-dimensional search methods play an important role in multidimensional optimization problems. Iterative algorithms for solving such optimization problems typically involve a line search at every iteration.

- Let $f: \mathbb{R}^n \to \mathbb{R}$ be a function that we wish to minimize. Iterative algorithms for finding a minimizer of f are of the form

$$\boldsymbol{x}^{(k+1)} = \boldsymbol{x}^{(k+1)} + \alpha_k \boldsymbol{d}^{(k)},$$

  where $\boldsymbol{x}^{(0)}$ is a given initial point and $\alpha_k \geq 0$ is chosen to minimize $\phi_k(\alpha) = f\left(\boldsymbol{x}^{(k)} + \alpha \boldsymbol{d}^{(k)}\right)$. The vector $\boldsymbol{d}^{(k)}$ is called the **search direction**.

- Note that choice of α_k involves a one-dimensional minimization.

2.2 GOLDEN SECTION SEARCH

- The search methods allow us to determine the **minimizer** of a function $f: \mathbb{R} \to \mathbb{R}$ over a closed interval, say $[a_0, b_0]$.

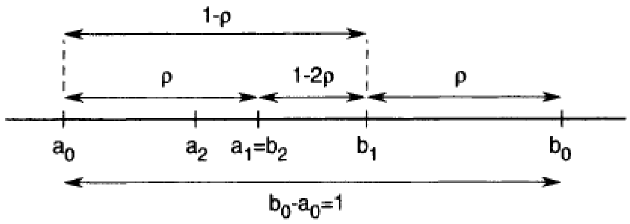
- We assume that the objective function f is **unimodal**, which means that f has only one local minimizer.

- We choose the intermediate points in such a way that the reduction in the range is **symmetric**, in the sense that

$$a_1 - a_0 = b_0 - b_1 = \rho(b_0 - a_0),$$

  where $\rho < 1/2$.

- If $f(a_1) < f(b_1)$, then the minimizer must lie in the range $[a_0, b_1]$. If, on the other hand, $f(a_1) \geq f(b_1)$, then the minimizer is located in the range $[a_1, b_0]$.



- Without loss of generality, imagine that the original range $[a_0, b_0]$ is of unit length. Then, to have **only one new evaluation** of f, it is enough to choose ρ so that $\rho(b_1 - a_0) = b_1 - b_2$.

- Because $b_1 - a_0 = 1 - \rho$ and $b_1 - b_2 = 1 - 2\rho$, we have

$$\rho(b_1 - a_0) = b_1 - b_2$$
$$\rho(1-\rho) = 1 - 2\rho$$
$$\rho^2 - 3\rho + 1 = 0 \Rightarrow \rho = \frac{3 \pm \sqrt{5}}{2}$$

- Because we require that $\rho < 1/2$, we take $\rho = (3-\sqrt{5})/2 \approx 0.3820$.
- Using the golden section rule means that at every stage (except the first stage) of the uncertainty range reduction, the objective function need only be evaluated at **one new point**.
- The uncertainty range is reduced by the ratio $1-\rho \approx 0.6180$ at every stage. Hence, N steps of reduction using the golden section method reduces the range by the factor $(1-\rho)^N \approx (0.6180)^N$.

Example 2.2.1. Use the golden section search to find the value of x that minimizes

$$f(x) = x^4 - 14x^3 + 60x^2 - 70x$$

in the range $[0, 2]$. Locate this value of x to within a range of 0.3.

After N stages the range $[0, 2]$ is reduced by $(0.6180)^N$. So, we choose N so that

$$2(0.6180)^N \leq 0.3.$$

Four stages of reduction will do; that is, $N = 4$.

Iteration 1.

$$f(a_1) = f(0.7640) = -24.3608 \qquad f(b_1) = f(1.2360) = -18.9596,$$

where $a_1 = a_0 + \rho(b_0 - a_0) = 0.7640$ and $b_1 = b_0 - \rho(b_0 - a_0) = 1.2360$. Thus, $f(a_1) < f(b_1)$, so the uncertainty interval is reduced to $[a_0, b_1] = [0, 1.2360]$ with the range of 1.2360.

Iteration 2.

$$f(a_2) = f(0.4722) = -21.0989 \qquad f(b_2) = f(a_1) = f(0.7640) = -24.3608,$$

where $a_2 = a_0 + \rho(b_1 - a_0) = 0.4722$. Thus, $f(a_2) > f(b_2)$, so the uncertainty interval is reduced to $[a_2, b_1] = [0.4722, 1.2360]$ with the range of 0.7638.

Iteration 3.

$$f(a_3) = f(b_2) = f(0.7640) = -24.3086 \qquad f(b_3) = f(0.9442) = -23.5930,$$

where $b_3 = b_2 - \rho(b_1 - a_2) = 0.9442$. Thus, $f(a_3) < f(b_3)$, so the uncertainty interval is reduced to $[a_2, b_3] = [0.4722, 0.9442]$ with the range of 0.4721.

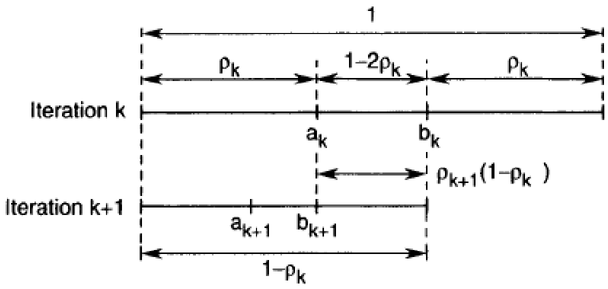
Iteration 4.

$$f(a_4) = f(0.6525) = -23.8375 \qquad f(b_4) = f(a_3) = f(0.7640) = -24.3608,$$

where $a_4 = a_2 + \rho(b_3 - a_2) = 0.6525$. Thus, $f(a_4) > f(b_4)$, so the uncertainty interval is reduced to $[a_4, b_3] = [0.6525, 0.9442]$ with the range of $0.2917 < 0.3$.

2.3 FIBONACCI SEARCH

- Suppose now that we are allowed to vary the value ρ from stage to stage, so that at the kth stage in the reduction process we use a value ρ_k, at the next stage we use a value ρ_{k+1}, and so on.



- Without loss of generality, imagine that the range $[a_{k-1}, b_{k-1}]$ is of unit length. Then, to have only one new evaluation of f, it is enough to choose ρ_{k+1} so that

$$\rho_{k+1}(b_k - a_{k-1}) = b_k - b_{k+1}.$$

- Because $b_k - a_{k-1} = 1 - \rho_k$ and $b_k - b_{k+1} = 1 - 2\rho_k$, we have

$$\rho_{k+1}(b_k - a_{k-1}) = b_k - b_{k+1}$$
$$\rho_{k+1}(1 - \rho_k) = 1 - 2\rho_k$$
$$\rho_{k+1} = \frac{1 - \rho_k - \rho_k}{1 - \rho_k} \Rightarrow \rho_{k+1} = 1 - \frac{\rho_k}{1 - \rho_k}$$

- Then, after N iterations of the algorithm, the uncertainty range is reduced by a factor of $(1 - \rho_1)(1 - \rho_2) \cdots (1 - \rho_N)$.

- The sequence $\rho_1, \rho_2, \ldots$ that minimizes the factor above is

$$\rho_1 = 1 - \frac{F_N}{F_{N+1}}, \rho_2 = 1 - \frac{F_{N-1}}{F_N}, \ldots, \rho_k = 1 - \frac{F_{N-k+1}}{F_{N-k+2}}, \ldots, \rho_N = 1 - \frac{F_1}{F_2},$$

  where the F_k are the elements of the Fibonacci sequence.

- The **Fibonacci sequence** is defined as follows. First, let $F_{-1} = 0$ and $F_0 = 1$ by convention. Then, for $k \geq 0$, $F_{k+1} = F_k + F_{k-1}$.

- The first eight elements in the Fibonacci sequence are as follows: $F_1 = 1, F_2 = 2, F_3 = 3, F_4 = 5, F_5 = 8, F_6 = 13, F_7 = 21, F_8 = 34$.

- In the Fibonacci search method, the uncertainty range is reduced by the factor:

$$(1-\rho_1)(1-\rho_2)\cdots(1-\rho_N) = \frac{F_N}{F_{N+1}}\frac{F_{N-1}}{F_N}\cdots\frac{F_1}{F_2} = \frac{1}{F_{N+1}}.$$

- Fibonacci method is better than the golden section method in that it gives a **smaller** final uncertainty range.
- In the final iteration of the Fibonacci search method, because $\rho_N = 1 - F_1/F_2 = 1/2$, the two intermediate points coincide in the middle of the uncertainty interval, and therefore we cannot further reduce the uncertainty range.
- To get around this problem, we perform the new evaluation for the last iteration using $\rho_N = 1/2 - \varepsilon$, where ε is a small number (e.g., $\varepsilon = 0.1$). Therefore, in the worst case, the reduction factor in the uncertainty range for the Fibonacci method is

$$(1-\rho_1)(1-\rho_2)\cdots(1-\rho_N) = \frac{F_N}{F_{N+1}}\frac{F_{N-1}}{F_N}\cdots\frac{F_2}{F_3}\left(\frac{F_1}{F_2}+\varepsilon\right) = \frac{F_2}{F_{N+1}}\left(\frac{F_1}{F_2}+\varepsilon\right) = \frac{1+2\varepsilon}{F_{N+1}}.$$

Example 2.3.1. Use the Fibonacci search to find the value of x that minimizes

$$f(x) = x^4 - 14x^3 + 60x^2 - 70x$$

in the range $[0,2]$. Locate this value of x to within a range of 0.3.

After N stages the range $[0,2]$ is reduced by $(1+2\varepsilon)/F_{N+1}$. So, we choose N so that

$$2\left(\frac{1+2\varepsilon}{F_{N+1}}\right) \le 0.3.$$

Let $\varepsilon = 0.1$. Four stages of reduction will do; that is, $N = 4$.

Iteration 1.

$$f(a_1) = f(0.7500) = -24.3398 \qquad f(b_1) = f(1.2500) = -18.6523,$$

where $a_1 = a_0 + \rho(b_0 - a_0) = 0.7500$ and $b_1 = b_0 - \rho(b_0 - a_0) = 1.2500$. Thus, $f(a_1) < f(b_1)$, so the uncertainty interval is reduced to $[a_0, b_1] = [0, 1.250]$ with the range of 1.2500.

Iteration 2.

$$f(a_2) = f(0.5000) = -21.6875 \qquad f(b_2) = f(a_1) = f(0.7500) = -24.3398,$$

where $a_2 = a_0 + \rho(b_1 - a_0) = 0.5000$. Thus, $f(a_2) > f(b_2)$, so the uncertainty interval is reduced to $[a_2, b_1] = [0.5000, 1.2500]$ with the range of 1.2500.

Iteration 3.

$$f(a_3) = f(b_2) = f(0.7500) = -24.3398 \qquad f(b_3) = f(1.0000) = -23.000,$$

where $b_3 = b_2 - \rho(b_1 - a_2) = 1.0000$. Thus, $f(a_3) < f(b_3)$, so the uncertainty interval is reduced to $[a_2, b_3] = [0.5000, 1.0000]$ with the range of 0.5000.

Iteration 4.

$$f(a_4) = f(0.7000) = -24.1619 \qquad f(b_4) = f(a_3) = f(0.7500) = -24.3398,$$

where $a_4 = a_2 + \rho(b_3 - a_2) = 0.7000$. Thus, $f(a_4) > f(b_4)$, so the uncertainty interval is reduced to $[a_4, b_3] = [0.7000, 1.0000]$ with the range of 0.3.

2.4 NEWTON'S METHOD

- Assume that at each measurement point $x^{(k)}$ we can calculate $f(x^{(k)})$, $f'(x^{(k)})$ and $f''(x^{(k)})$. We can fit a quadratic function through $x^{(k)}$ that matches its first and second derivatives with that of the function f. This quadratic has the form

$$q(x) = f(x^{(k)}) + f'(x^{(k)})(x - x^{(k)}) + \frac{1}{2} f''(x^{(k)})(x - x^{(k)})^2.$$

- Then, instead of minimizing f, we minimize its approximation q. The first-order necessary condition for a minimizer of q yields

$$0 = q'(x) = f'(x^{(k)}) + f''(x^{(k)})(x - x^{(k)}).$$

- Setting $x = x^{(k+1)}$, we obtain

$$x^{(k+1)} = x^{(k)} - \frac{f'(x^{(k)})}{f''(x^{(k)})}.$$

- Newton's method works well if $f''(x) > 0$ everywhere. However, if $f''(x) < 0$ for some x, Newton's method may fail to converge to the minimizer.

Example 2.4.1. Use Newton's method, find the minimizer of

$$f(x) = \frac{1}{2}x^2 - \sin x.$$

The initial value is $x^{(0)} = 0.5$. The required accuracy is $\varepsilon = 10^{-5}$, in the sense that we stop when $|x^{(k+1)} - x^{(k)}| < \varepsilon$.

We compute

$$f'(x) = x - \cos x\,, \quad f''(x) = 1 + \sin x.$$

Hence,

$$x^{(1)} = x^{(0)} - \frac{f'(x^{(0)})}{f''(x^{(0)})} = 0.5 - \frac{0.5 - \cos 0.5}{1 + \sin 0.5} = 0.7552 \quad \text{and} \quad |x^{(1)} - x^{(0)}| = 0.2552$$

$$x^{(2)} = x^{(1)} - \frac{f'(x^{(1)})}{f''(x^{(1)})} = x^{(1)} - \frac{0.02710}{1.685} = 0.7391 \quad \text{and} \quad |x^{(2)} - x^{(1)}| = 0.1608$$

$$x^{(3)} = x^{(2)} - \frac{f'(x^{(2)})}{f''(x^{(2)})} = x^{(2)} - \frac{9.461 \times 10^{-5}}{1.673} = 0.7390 \quad \text{and} \quad |x^{(3)} - x^{(2)}| = 5.65 \times 10^{-5}$$

$$x^{(4)} = x^{(3)} - \frac{f'(x^{(3)})}{f''(x^{(3)})} = x^{(3)} - \frac{1.17 \times 10^{-9}}{1.673} = 0.7390 \quad \text{and} \quad |x^{(4)} - x^{(3)}| = 7.06 \times 10^{-10}$$

Note that $f'(x^{(4)}) = -8.6 \times 10^{-6} \approx 0$ and $f''(x^{(4)}) = 1.673 > 0$, so $x^* \approx x^{(4)}$ is a strict local minimizer.

2.5 SECANT METHOD

- Newton's method for minimizing f uses **second derivatives** of f.
- If the second derivative is not available, we may attempt to approximate it using first derivative information. In particular, we may approximate $f''(x^{(k)})$ with

$$\frac{f'(x^{(k)}) - f'(x^{(k-1)})}{x^{(k)} - x^{(k-1)}}.$$

- Using the foregoing approximation of the second derivative, we obtain the algorithm

$$x^{(k+1)} = x^{(k)} - \frac{x^{(k)} - x^{(k-1)}}{f'(x^{(k)}) - f'(x^{(k-1)})} f'(x^{(k)}) = \frac{f'(x^{(k)})x^{(k-1)} - f'(x^{(k-1)})x^{(k)}}{f'(x^{(k)}) - f'(x^{(k-1)})},$$

  called the **secant method**.
- Note that the algorithm requires two initial points $x^{(-1)}$ and $x^{(0)}$.

Example 2.5.1. Use Newton's method, find the minimizer of

$$f(x) = \frac{1}{2}x^2 - \sin x.$$

The starting points are $x^{(-1)} = 0.4$ and $x^{(0)} = 0.5$. The required accuracy is $\varepsilon = 10^{-5}$, in the sense that we stop when $|x^{(k+1)} - x^{(k)}| < \varepsilon$.

We compute $f'(x) = x - \cos x$. Hence,

$$x^{(1)} = x^{(0)} - \frac{x^{(0)} - x^{(-1)}}{f'(x^{(0)}) - f'(x^{(-1)})} f'(x^{(0)}) = 0.5 - \frac{0.5 - 0.4}{-0.3776 - (-0.5211)}(-0.3776)$$
$$= 0.7632 \quad \text{and} \quad |x^{(1)} - x^{(0)}| = 0.2632$$

$$x^{(2)} = x^{(1)} - \frac{x^{(1)} - x^{(0)}}{f'(x^{(1)}) - f'(x^{(0)})} f'\left(x^{(1)}\right) = 0.7632 - \frac{0.7632 - 0.5}{0.0405 - (-0.3776)}(0.0405)$$
$$= 0.7377 \quad \text{and} \quad \left|x^{(2)} - x^{(1)}\right| = 0.0255$$

$$x^{(3)} = x^{(2)} - \frac{x^{(2)} - x^{(1)}}{f'(x^{(2)}) - f'(x^{(1)})} f'\left(x^{(2)}\right) = 0.7377 - \frac{0.7377 - 0.7632}{-0.0024 - (0.0405)}(-0.0024)$$
$$= 0.7391 \quad \text{and} \quad \left|x^{(3)} - x^{(2)}\right| = 0.0014$$

$$x^{(4)} = x^{(3)} - \frac{x^{(3)} - x^{(2)}}{f'(x^{(3)}) - f'(x^{(2)})} f'\left(x^{(3)}\right) = 0.7391 - \frac{0.7391 - 0.7377}{0.0000 - (-0.0024)}(0.0000)$$
$$= 0.7391 \quad \text{and} \quad \left|x^{(4)} - x^{(3)}\right| = 7.46 \times 10^{-6}$$

Note that $f'\left(x^{(4)}\right) \approx 0$ and $f''\left(x^{(4)}\right) = 1.674 > 0$, so $x^* \approx x^{(4)}$ is a strict local minimizer.