16 Optimization

This lecture describes methods for the optimization of a real-valued function $f(x)$ on a bounded real interval $[a, b]$. We will describe methods for determining the maximum of $f(x)$ on $[a, b]$, i.e., we solve

$$\max_{a \leq x \leq b} f(x). \tag{1}$$

The same methods applied to $-f(x)$ determine the minimum of $f(x)$ on $[a, b]$.

We first note that if the derivative $f'(x)$ exists, is continuous on $[a, b]$, and is not too difficult to evaluate, then it can be attractive to solve (1) by first computing all distinct zeros of $f'(x)$, say $z_1 < z_2 < \cdots < z_\ell$, in the interior of the interval $[a, b]$, and then evaluate $f(x)$ at these zeros and at the endpoints a and b. The z_j can be minima, maxima, or inflection points of $f(x)$. The function $f(x)$ achieves its maximum at one these zeros or at the endpoints. The zeros of $f'(x)$ can be computed by one of the methods of Lectures 6-7.

The remainder of this lecture describes methods that do not require evaluation of the derivative. These methods are attractive to use when $f'(x)$ is either not available or very complicated to compute. The first method *Golden Section Search* (GSS) is analogous to bisection. The second method applies interpolation by a quadratic polynomial.

Let $\mathbf{N}(x)$ denote an open real interval that contains x. The function $f(x)$ is said to have a *local maximum* at x^* if there is an open interval $\mathbf{N}(x^*)$, such that

$$f(x^*) \geq f(x), \qquad x \in \mathbf{N}(x^*) \cap [a, b].$$

A solution of (1) is referred to as a *global maximum* of $f(x)$ on $[a, b]$. There may be several global maxima. Every global maximum is a local maximum. The converse is, of course, not true.

The optimization methods to be described determine a local maximum of $f(x)$ in $[a, b]$. Sometimes it is known from the background of the problem that there is at most one local maximum, which then necessarily also is the global maximum. In general, however, a function can have several local maxima, which may not be global maxima. A function, of course, may have several global maxima.

Example 16.1: The function $f(x) = \sin(x)$ has global maxima at $\pi/2$ and $5\pi/2$ in the interval $[0, 3\pi]$; see Figure 1. □

Example 16.2: The function $f(x) = e^{-x}\sin(x)$ has local maxima at $\pi/4$ and $9\pi/4$ in the interval $[0, 3\pi]$; see Figure 2. The local maximum at $\pi/4$ also is the global maximum. □

Many optimization methods are myopic; they only know the function in a small neighborhood of the current approximation of the maximum. There are engineering techniques for steering myopic optimization methods towards a global maximum. These techniques are referred to as *evolutionary methods*, *genetic algorithms*, or *simulated annealing*.

16.1 Golden section search

This method is analogous to bisection in the sense that the original interval $[a_1, b_1] = [a, b]$ is replaced by a sequence of intervals $[a_j, b_j]$, $j = 2, 3, \ldots$, of decreasing lengths, so that each interval $[a_j, b_j]$ contains at least one local maximum of $f(x)$. GSS divides each one of the intervals $[a_j, b_j]$, $j = 1, 2, \ldots$, into three subintervals $[a_j, x_j^{\text{left}}]$, $[x_j^{\text{left}}, x_j^{\text{right}}]$, and $[x_j^{\text{right}}, b_j]$, where x_j^{left} and x_j^{right} are cleverly chosen points, with $a_j < x_j^{\text{left}} < x_j^{\text{right}} < b_j$. The method then discards the right-most or left-most subinterval, so that the remaining interval is guaranteed to contain a local maximum of $f(x)$.

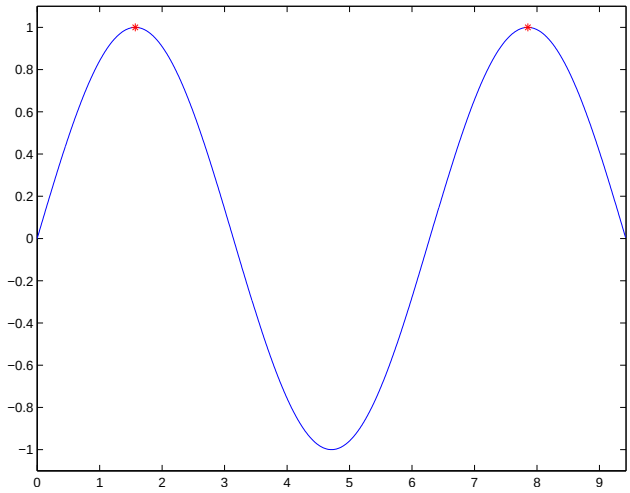

Figure 1: The function $f(x) = \sin(x)$ for $0 \leq x \leq 3\pi$. The function has two local maxima (at $x = \pi/2$ and $x = 9\pi/2$). The corresponding function values are marked by $*$ (in red).

Thus, let $a_1 < x_1^{\text{left}} < x_1^{\text{right}} < b_1$ and evaluate $f(x)$ at these points. If $f(x_1^{\text{left}}) \geq f(x_1^{\text{right}})$, then $f(x)$ has a local maximum in the interval $[a_1, x_1^{\text{right}}]$ and we set $[a_2, b_2] := [a_1, x_1^{\text{right}}]$; otherwise $f(x)$ has a local maximum in $[x_1^{\text{left}}, b_1]$ and we set $[a_2, b_2] := [x_1^{\text{left}}, b_1]$. We now repeat the process for the interval $[a_2, b_2]$, i.e., we select interior points x_2^{left} and x_2^{right}, such that $a_2 < x_2^{\text{left}} < x_2^{\text{right}} < b_2$ and evaluate $f(x)$ at these points. Similarly, as above we obtain the new interval $[a_3, b_3]$, which is guaranteed to contain a local maximum of $f(x)$. The computations proceed in this manner until a sufficiently small interval, say $[a_k, b_k]$, which contains a local maximum of $f(x)$, has been determined.

We would like to choose the interior points x_j^{left} and x_j^{right}, so that the interval $[a_{j+1}, b_{j+1}]$ is substantially shorter than the interval $[a_j, b_j]$.

Example 16.3: The choices

$$x_j^{\text{left}} = a_j + \frac{1}{3}(b_j - a_j), \qquad x_j^{\text{right}} = a_j + \frac{2}{3}(b_j - a_j), \tag{2}$$

secure that the length of the new interval $[a_{j+1}, b_{j+1}]$ is $2/3$ of the length of the previous interval $[a_j, b_j]$. This subdivision requires that $f(x)$ be evaluated at x_j^{left} and x_j^{right} for every j. □

There is a better choice of interior points than in Example 16.3. The following selection reduces the length of the intervals $[a_j, b_j]$ faster than in Example 16.3 and requires only one function evaluation for every $j \geq 2$. The trick is to choose the interior points x_j^{left} and x_j^{right} in $[a_j, b_j]$ so that one of them becomes an endpoint of the next interval $[a_{j+1}, b_{j+1}]$ and the other one becomes one of the interior points x_{j+1}^{left} or x_{j+1}^{right} of this interval. We then only have to evaluate $f(x)$ at one interior point.

Let the interior points be given by

$$x_j^{\text{left}} = a_j + \rho(b_j - a_j), \qquad x_j^{\text{right}} = a_j + (1 - \rho)(b_j - a_j) \tag{3}$$

for some constant $0 < \rho < 1$ to be determined. The choice $\rho = 1/3$ yields (2).

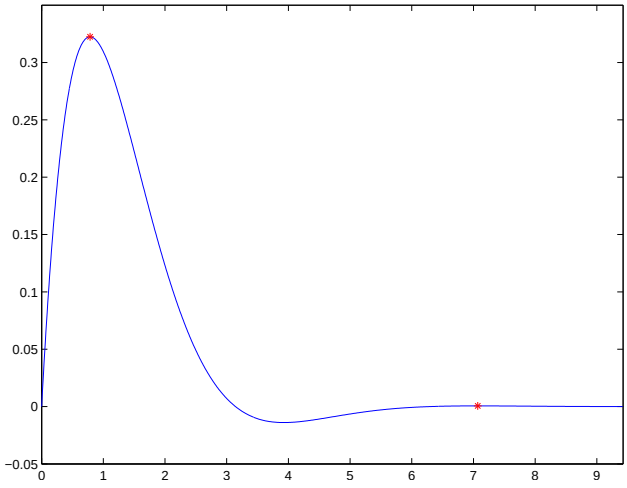

Figure 2: The function $f(x) = e^{-x}\sin(x)$ for $0 \le x \le 3\pi$. The function has two local maxima (at $x = \pi/4$ and $x = 9\pi/4$). The corresponding function values are marked by $*$ (in red).

Assume that $f(x_j^{\text{left}}) \ge f(x_j^{\text{right}})$. Then $f(x)$ has a local maximum in $[a_j, x_j^{\text{right}}]$. We set

$$a_{j+1} := a_j, \qquad b_{j+1} := x_j^{\text{right}} \tag{4}$$

and would like

$$x_{j+1}^{\text{right}} = x_j^{\text{left}}. \tag{5}$$

The latter requirement determines the parameter ρ.

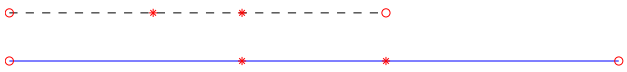

Figure 3: The bottom line depicts the interval $[a_j, b_j]$ with the endpoints a_j and b_j marked by o (in red), and the interior points x_j^{left} and x_j^{right} marked by $*$ (in red). The top line shows the interval $[a_{j+1}, b_{j+1}] = [a_j, x_j^{\text{right}}]$ that may be determined in step j of GSS. The endpoints are marked by o (in red) and the interior points x_{j+1}^{left} and x_{j+1}^{right} by $*$ (in red).

There are many ways to determine ρ. One of the simplest is based on the observation that

$$\frac{b_{j+1} - x_{j+1}^{\text{right}}}{b_{j+1} - a_{j+1}} = \frac{b_j - x_j^{\text{right}}}{b_j - a_j}, \tag{6}$$

cf. Figure 3. Using first (4) and (5), and subsequently (3), the left-hand side simplifies to

$$\frac{b_{j+1} - x_{j+1}^{\text{right}}}{b_{j+1} - a_{j+1}} = \frac{x_j^{\text{right}} - x_j^{\text{left}}}{x_j^{\text{right}} - a_j} = \frac{(1-2\rho)(b_j - a_j)}{(1-\rho)(b_j - a_j)} = \frac{1-2\rho}{1-\rho}.$$

The right-hand side of (6) can be simplified similarly. It follows from the right-hand side equation of (3) that $x_j^{\text{right}} = b_j - \rho(b_j - a_j)$ and, therefore,

$$\frac{b_j - x_j^{\text{right}}}{b_j - a_j} = \frac{\rho(b_j - a_j)}{b_j - a_j} = \rho.$$

Thus, equation (6) is equivalent to

$$\frac{1 - 2\rho}{1 - \rho} = \rho,$$

which can be expressed as

$$\rho^2 - 3\rho + 1 = 0.$$

The only root smaller than unity is given by

$$\rho = \frac{1}{2}(3 - \sqrt{5}) \approx 0.382.$$

It follows that the lengths of the intervals $[a_j, b_j]$ decreases as j increases according to

$$b_{j+1} - a_{j+1} = x_j^{\text{right}} - a_j = (1 - \rho)(b_j - a_j) \approx 0.618\,(b_j - a_j).$$

In particular, the length decreases less in each step than for the bisection method of Lectures 6-7. We conclude that GSS is useful for determining a local maximum to low accuracy. If high accuracy is desired, then faster methods should be employed, such as the method of the following section.

Example 16.4: We apply GSS to the function of Example 16.1 with $[a_1, b_1] = [0, 3\pi]$. We obtain $x_1^{\text{left}} = 3.60$, $x_1^{\text{right}} = 5.82$, and $f(x_1^{\text{left}}) = f(x_1^{\text{right}}) = -0.44$. We therefore choose the next interval to be $[a_2, b_2] = [0, x_1^{\text{right}}]$ and obtain $x_2^{\text{right}} = 2.22$ and $x_2^{\text{left}} = x_1^{\text{right}} = 5.82$. We have $f(x_2^{\text{right}}) = 0.79$ and it is clear how the method will determine the local maximum at $x = \pi/2$. The progress of GSS is illustrated by Figure 3.

The reason that this local maximum is determined depends on that we discard the subinterval $[x_1^{\text{right}}, b_1]$ when $f(x_1^{\text{left}}) \geq f(x_1^{\text{right}})$. If this inequality would have been $f(x_1^{\text{left}}) > f(x_1^{\text{right}})$, then the other local maximum would have been found. This example illustrates that it generally is a good idea to graph the function first, and then choose $[a_1, b_1]$ so that it contains the desired maximum. □

Example 16.5: We apply GSS to the function of Example 16.2 with $[a_1, b_1] = [0, 3\pi]$. Similarly as in Example 16.4, we obtain $x_1^{\text{left}} = 3.60$, $x_1^{\text{right}} = 5.82$, $f(x_1^{\text{left}}) = -1.2 \cdot 10^{-2}$, and $f(x_1^{\text{right}}) = -1.2 \cdot 10^{-3}$. We therefore choose the next interval to be $[a_2, b_2] = [x_1^{\text{left}}, b_1]$, and obtain $x_2^{\text{left}} = 5.82$ and $x_2^{\text{right}} = 7.20$. We evaluate $f(x_2^{\text{right}}) = 5.9 \cdot 10^{-4}$ and it is clear the method will determine the local maximum at $x = 5\pi/2$. If we would have liked to determine the global maximum, then a smaller initial interval $[a_1, b_1]$ containing $x = \pi/4$ should be used. □

16.2 Quadratic interpolation

We found in Lectures 15 that the secant method gives faster convergence than bisection. Analogously, we will see that interpolation by a sequence of quadratic polynomials yields faster convergence towards a local maximum than GSS. The rationale is that GSS only uses the function values to determines whether the function at a new node is larger or smaller than at an adjacent node. We describe in this section a method based on interpolating the function by a sequence of quadratic polynomial and determine the maxima of

the latter. We proceed as follows: Let the interval $[a_j, b_j]$ be determined in step $j-1$ of GSS, and let $\hat{x}_j$ denote the point inside this interval at which $f(\hat{x}_j)$ is known, i.e., $\hat{x}_j$ is either x_j^{left} or x_j^{right}. Instead of evaluating $f(x)$ by a new point determined by GSS, we fit a quadratic polynomial $q_j(x)$ to the data $(a_j, f(a_j))$, $(b_j, f(b_j))$, and $(\hat{x}_j, f(\hat{x}_j))$. If this polynomial has a maximum in the interior of $[a_j, b_j]$, then this will be our new interior point. Denote this point by x_j^{polmax} and evaluate $f(x_j^{\text{polmax}})$.

We remark that it is easy to determine x_j^{polmax}. If the leading coefficient of $q_j(x)$ is negative, then $q_j(x)$ is concave ($-q_j(x)$ is convex) and has a maximum. The point x_j^{polmax} is the zero of the linear polynomial $q_j'(x)$.

Thus, assume that $a_j < x_j^{\text{polmax}} < b_j$ and that $q_j(x)$ achieves its maximum at x_j^{polmax}. We then discard the point in the set $\{a_j, \tilde{x}_j, b_j\}$ furthest away from x_j^{polmax}. This gives us three nodes at which the function $f(x)$ is known, and we can compute a new quadratic polynomial, whose maximum we determine, and so on.

Interpolation by a sequence of quadratic polynomials usually gives much faster convergence than GSS, however, it is not guaranteed to converge. Thus, quadratic interpolation has to be safeguarded. For instance, one may have to apply GSS when quadratic interpolation fails to converge to a maximum in the last interval determined by GSS.

Example 16.6. The quadratic polynomial determined by interpolating $f(x)$ at the nodes $\{a_1, x_1^{\text{left}}, b_1\}$ or at the nodes $\{a_1, x_1^{\text{right}}, b_1\}$ in Example 16.5 is not convex. Hence, we have to reduce the interval by GSS before applying polynomial interpolation. □

We finally comment on the accuracy of the computed local maximum by any optimization method. Assume that we have determined an approximation $\hat{x}$ of the local maximum x^*. Let $f(x)$ have a continuous second derivative $f''(x)$ in a neighborhood of x^*. Then, in view of that $f'(x^*)$ vanishes, Taylor expansion of $f(x)$ at x^* yields

$$f(\hat{x}) - f(x^*) = \frac{1}{2}(x_j - x^*)^2 f''(\tilde{x}),$$

where $\tilde{x}$ lives in the interval with endpoints x^* and $\hat{x}$. This formula implies that the error in the computed maximum $\hat{x}$ of the function $f(x)$ can be expected to be proportional to square root of the error in the function value $f(\hat{x})$. In particular, we cannot expect to be able to determine a local maximum with a relative error smaller than the square root of machine epsilon.

Exercise 16.1

Exercise 16.1. Apply GSS with initial interval $[0, \pi]$ to determine the global maximum of the function in Example 16.2. Carry out 4 steps by GSS. □

Exercise 16.2

Apply quadratic interpolation with initial interval $[0, \pi]$ to determine the global maximum of the function in Example 16.2. Carry out 4 steps. How does the accuracy of the computed approximate maximum compare with the accuracy of the computed approximate maximum determined in Exercise 16.1? Estimate empirically the rate of convergence of the quadratic interpolation method. For instance, study the quotients

$$s_j = \left| \frac{x_{j+1}^{\text{polmax}} - x^*}{x_j^{\text{polmax}} - x^*} \right|, \qquad j = 1, 2, 3, \dots ,$$

or their logarithms. □

Exercise 16.3

Compute the minimum of the function in Example 16.2 by carrying out 4 GSS steps. □

Exercise 16.4

Exercise 16.4. GSS is analogous to bisection, and the quadratic interpolation method is analogous to the secant method. The former applies local quadratic interpolation to determine a maximum, while the latter applies local linear interpolation to determine a zero. Describe an analogue of regula falsi for optimization. □