EAD 115

Numerical Solution of Engineering and Scientific Problems

David M. Rocke
Department of Applied Science
One-Dimensional Unconstrained Optimization

• Given a function $f(x)$, find its maximum value (or its minimum value).
• We may not know how many maxima $f()$ has.
• We need methods of finding local maxima.
• We need methods of attempting to find the global maximum, though this can be difficult.
A graph of a function $f(x)$ against $x$ shows:

- **Roots** where $f(x) = 0$.
- **Minimum** where $f'(x) = 0$ and $f''(x) > 0$.
- **Maximum** where $f'(x) = 0$ and $f''(x) < 0$.

The graph illustrates the behavior of the function at critical points based on the first and second derivatives.
Global and Local Optimization

• We will mainly discuss ways of finding local optima
• One method of attempting to find the global optimum is to find locate local optima repeatedly from diverse, often random starting points
• This can be surprisingly effective considering how simple a method it is
• Other methods for global optimization include simulated annealing, Markov chain Monte Carlo, genetic and other evolutionary algorithms, and tabu search

• These methods are more complicated than repeated restarted optimum finding, but not necessarily more effective

• Biological or physical analogy does not guarantee good performance

• Hard work and careful tuning are required; there are no magic black boxes!
Bracket Methods

• Suppose we have an interval that is thought to contain a single maximum of a function $f(x)$, so that $f(x)$ is increasing from the lower end to the maximum, and decreasing from the maximum to the upper end

• We want a method similar to bisection for solving this problem
Adding new points

• In bisection, we add one additional new point in the middle, and pick either the left or the right interval based on which one has the function changing sign

• This does not provide enough information for finding a maximum—we will need at least two additional points for this
(a) Eliminate $x_1$ and $x_2$ from consideration.

(b) New iterations.

Old $x_1$ becomes new $x_2$.

Old $x_2$ becomes new $x_1$. 

Extremum (maximum) at $x_1$. 

$f(x)$ is the function.
No maximum here
No maximum here
No maximum here
No maximum here
Even Interval Division Wastes a Function Evaluation
Golden Section Search uses Function Evaluations Efficiently
\[(x - rx) = r^2 x\]

\[1 - r = r^2\]

\[r^2 - r + 1 = 0\]

\[r = \frac{\sqrt{5} - 1}{2}\]

\[r = 0.6180\]
FUNCTION Gold (xlow, xhigh, maxit, es, fx)
R = (5^{0.5} - 1)/2
xt = xlow; xu = xhigh
iter = 1
d = R * (xu - xt)
x1 = xt + d; x2 = xu - d
f1 = f(x1)
f2 = f(x2)

IF f1 > f2 THEN
  xo = x1
  fx = f1
ELSE
  xo = x2
  fx = f2
END IF
DO
  d = R*d
  IF f1 > f2 THEN
    xt = x2
    x2 = x1
    x1 = x1 + d
    f2 = f1
    f1 = f(x1)
  ELSE
    xu = x1
    x1 = x2
    x2 = xu - d
    f1 = f2
    f2 = f(x2)
  END IF
  iter = iter + 1
  IF f1 > f2 THEN
    xo = x1
    fx = f1
  ELSE
    xo = x2
    fx = f2
  END IF
  IF xo = 0. THEN
    ea = (1. - R) * ABS((xu - xt)/xo) * 100.
  END IF
  IF ea <= es OR iter >= maxit EXIT
END DO
Gold = xo
END Gold

(a) Maximization
(b) Minimization
Importance of Number of Function Evaluations

- In small problems this does not matter
- Reduced function evaluations means faster performance
- This matters if a large number of optimizations needs to be performed
- This matters if one function evaluation is expensive in computation time
Error Analysis for Golden Section Search

- At the end of each iteration, we have an interval that is “known” to contain the optimum
- We analyze the case where the left-hand interval is discarded; the other case is symmetric
- Old points are $x_l, x_2, x_1,$ and $x_u$
- New points are $x_2, x_1,$ and $x_u, x_1$ is guess
\[ x_1 - x_2 = x_l + r(x_u - x_l) - [x_u - r(x_u - x_l)] \]
\[ = -(x_u - x_l) + 2r(x_u - x_l) \]
\[ = (2r - 1)(x_u - x_l) \]
\[ = .236(x_u - x_l) \]

\[ x_u - x_1 = x_u - [x_l + r(x_u - x_l)] \]
\[ = (1 - r)(x_u - x_l) \]
\[ = .382(x_u - x_l) \]
Example

• $f(x) = 2 \sin x - \frac{x^2}{10}$
• Initial interval is $[0, 4]$
• Graph function
• Use golden section search to find a maximum
<table>
<thead>
<tr>
<th>i</th>
<th>$x_i$</th>
<th>$f_i$</th>
<th>$x_2$</th>
<th>$f_2$</th>
<th>$x_1$</th>
<th>$f_1$</th>
<th>$x_u$</th>
<th>$f_u$</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>0</td>
<td>0</td>
<td>1.5279</td>
<td>1.7647</td>
<td>2.4721</td>
<td>0.6300</td>
<td>4.0000</td>
<td>-3.114</td>
</tr>
<tr>
<td>2</td>
<td>0</td>
<td>0</td>
<td>0.9443</td>
<td>1.5310</td>
<td>1.5279</td>
<td>1.7647</td>
<td>2.4721</td>
<td>0.6300</td>
</tr>
<tr>
<td>3</td>
<td>0.9443</td>
<td>1.5310</td>
<td>1.5279</td>
<td>1.7647</td>
<td>1.8885</td>
<td>1.5432</td>
<td>2.4721</td>
<td>0.6300</td>
</tr>
<tr>
<td>4</td>
<td>0.9443</td>
<td>1.5310</td>
<td>1.3050</td>
<td>1.7595</td>
<td>1.5279</td>
<td>1.7647</td>
<td>1.8885</td>
<td>1.5432</td>
</tr>
<tr>
<td>5</td>
<td>1.3050</td>
<td>1.7595</td>
<td>1.5279</td>
<td>1.7647</td>
<td>1.6656</td>
<td>1.7136</td>
<td>1.8885</td>
<td>1.5432</td>
</tr>
<tr>
<td>6</td>
<td>1.3050</td>
<td>1.7595</td>
<td>1.4427</td>
<td>1.7755</td>
<td>1.5279</td>
<td>1.7647</td>
<td>1.6656</td>
<td>1.7136</td>
</tr>
<tr>
<td>7</td>
<td>1.3050</td>
<td>1.7595</td>
<td>1.3901</td>
<td>1.7742</td>
<td>1.4427</td>
<td>1.7755</td>
<td>1.5279</td>
<td>1.7647</td>
</tr>
<tr>
<td>8</td>
<td>1.3901</td>
<td>1.7742</td>
<td>1.4427</td>
<td>1.7755</td>
<td>1.4752</td>
<td>1.7732</td>
<td>1.5279</td>
<td>1.7647</td>
</tr>
</tbody>
</table>
Quadratic Interpolation

• If golden section search is analogous to bisection, then the equivalent of linear interpolation (false position) is quadratic interpolation

• We approximate the function over the interval by a quadratic (parabola), and solve the quadratic

• Requires three points instead of two
Fig 13.6

- True maximum
- True function
- Quadratic approximation of maximum
- Quadratic function

The diagram shows a graph with the function $f(x)$ plotted against $x$. The true function is indicated by the solid line, and the quadratic approximation is shown in dashed line. The maximum point is marked with a circle.
Given three points find the quadratic joining them

\((x_0, f(x_0))\)
\((x_1, f(x_1))\)
\((x_2, f(x_2))\)

\(f(x) = ax^2 + bx + c\)

\(f(x_0) = ax_0^2 + bx_0 + c\)

\(f(x_1) = ax_1^2 + bx_1 + c\)

\(f(x_2) = ax_2^2 + bx_2 + c\)
\[
\begin{align*}
    f(x_0) &= ax_0^2 + bx_0 + c \\
    f(x_1) &= ax_1^2 + bx_1 + c \\
    f(x_2) &= ax_2^2 + bx_2 + c
\end{align*}
\]

\[
\begin{bmatrix}
    x_0^2 & x_0 & 1 \\
    x_1^2 & x_1 & 1 \\
    x_2^2 & x_2 & 1
\end{bmatrix}
\begin{bmatrix}
    a \\
    b \\
    c
\end{bmatrix}
= 
\begin{bmatrix}
    f(x_0) \\
    f(x_1) \\
    f(x_2)
\end{bmatrix}
\]
\[
x_3 = \frac{f(x_0)(x_1^2 - x_2^2) + f(x_1)(x_2^2 - x_0^2) + f(x_2)(x_0^2 - x_1^2)}{2f(x_0)(x_1 - x_2) + 2f(x_1)(x_2 - x_0) + 2f(x_2)(x_0 - x_1)}
\]

- Initial endpoints are \(x_0\) and \(x_2\)
- Initial middle point is \(x_1\)
- New middle point as guess for the optimum is \(x_3\)
- Discard one of \(x_0\) or \(x_3\) using the same rule as golden section search
- Error estimate usually by change in estimate
<table>
<thead>
<tr>
<th>i</th>
<th>$x_0$</th>
<th>$f_0$</th>
<th>$x_1$</th>
<th>$f_1$</th>
<th>$x_2$</th>
<th>$f_2$</th>
<th>$x_3$</th>
<th>$f_3$</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>0.0000</td>
<td>0.0000</td>
<td>1.0000</td>
<td>1.5829</td>
<td>4.0000</td>
<td>-3.114</td>
<td>1.5055</td>
<td>1.7691</td>
</tr>
<tr>
<td>2</td>
<td>1.0000</td>
<td>1.5829</td>
<td>1.5055</td>
<td>1.7691</td>
<td>4.0000</td>
<td>-3.114</td>
<td>1.4903</td>
<td>1.7714</td>
</tr>
<tr>
<td>3</td>
<td>1.0000</td>
<td>1.5829</td>
<td>1.4903</td>
<td>1.7714</td>
<td>1.5055</td>
<td>1.7691</td>
<td>1.4256</td>
<td>1.7757</td>
</tr>
<tr>
<td>4</td>
<td>1.0000</td>
<td>1.5829</td>
<td>1.4256</td>
<td>1.7757</td>
<td>1.4903</td>
<td>1.7714</td>
<td>1.4266</td>
<td>1.7757</td>
</tr>
<tr>
<td>5</td>
<td>1.4256</td>
<td>1.7757</td>
<td>1.4266</td>
<td>1.7757</td>
<td>1.4903</td>
<td>1.7714</td>
<td>1.4275</td>
<td>1.7757</td>
</tr>
</tbody>
</table>
Newton’s Method

• Open rather than bracketing method, analogous to Newton-Raphson
• Also uses a quadratic model of the function
• Quadratic model is at a point not over an interval
• Optimum is when the derivative is 0
• Use Newton-Raphson on the derivative
\[ f(x) \doteq f(x_i) + f'(x_i)(x - x_i) + 0.5f''(x_i)(x - x_i)^2 \]
\[ f'(x) \doteq f'(x_i) + 0.5f''(x_i)2(x - x_i) \]
\[ 0 \doteq f'(x_i) + f''(x_i)(x - x_i) \]
\[ f''(x_i)(x - x_i) \doteq -f'(x_i) \]
\[ (x - x_i) \doteq -f'(x_i) / f''(x_i) \]
\[ x_{i+1} \doteq x_i - f'(x_i) / f''(x_i) \]
<table>
<thead>
<tr>
<th>i</th>
<th>x</th>
<th>f(x)</th>
<th>f'(x)</th>
<th>f''(x)</th>
</tr>
</thead>
<tbody>
<tr>
<td>0</td>
<td>2.50000</td>
<td>0.57194</td>
<td>-2.10229</td>
<td>-1.39694</td>
</tr>
<tr>
<td>1</td>
<td>0.99508</td>
<td>1.57859</td>
<td>0.88985</td>
<td>-1.87761</td>
</tr>
<tr>
<td>2</td>
<td>1.46901</td>
<td>1.77385</td>
<td>-0.09058</td>
<td>-2.18965</td>
</tr>
<tr>
<td>3</td>
<td>1.42764</td>
<td>1.77573</td>
<td>-0.00020</td>
<td>-2.17954</td>
</tr>
<tr>
<td>4</td>
<td>1.42755</td>
<td>1.77573</td>
<td>0.00000</td>
<td>-2.17952</td>
</tr>
</tbody>
</table>
Error behavior of optimization methods in dimension 1

- Golden section search has linear convergence with ratio $\phi$ and an error bound
- Quadratic interpolation has linear convergence and an error estimate from change in the estimate
- Newton’s method has quadratic convergence and an error estimate from change in the estimate
## Golden Section Search

<table>
<thead>
<tr>
<th>i</th>
<th>$x$</th>
<th>$E_t$</th>
<th>$E_a$</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>1.5279</td>
<td>0.1003</td>
<td>0.9442</td>
</tr>
<tr>
<td>2</td>
<td>1.5279</td>
<td>0.1003</td>
<td>0.5836</td>
</tr>
<tr>
<td>3</td>
<td>1.5279</td>
<td>0.1003</td>
<td>0.3607</td>
</tr>
<tr>
<td>4</td>
<td>1.5279</td>
<td>0.1003</td>
<td>0.2229</td>
</tr>
<tr>
<td>5</td>
<td>1.5279</td>
<td>0.1003</td>
<td>0.1378</td>
</tr>
<tr>
<td>6</td>
<td>1.4427</td>
<td>0.0151</td>
<td>0.0852</td>
</tr>
<tr>
<td>7</td>
<td>1.4427</td>
<td>0.0151</td>
<td>0.0526</td>
</tr>
<tr>
<td>8</td>
<td>1.4427</td>
<td>0.0151</td>
<td>0.0325</td>
</tr>
</tbody>
</table>
## Quadratic Interpolation

<table>
<thead>
<tr>
<th>I</th>
<th>x</th>
<th>$E_t$</th>
<th>$E_a$</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>1</td>
<td>0.4276</td>
<td>-</td>
</tr>
<tr>
<td>2</td>
<td>1.5055</td>
<td>0.0779</td>
<td>0.5055</td>
</tr>
<tr>
<td>3</td>
<td>1.4903</td>
<td>0.0627</td>
<td>0.0152</td>
</tr>
<tr>
<td>4</td>
<td>1.4256</td>
<td>0.0020</td>
<td>0.0647</td>
</tr>
<tr>
<td>5</td>
<td>1.4266</td>
<td>0.0010</td>
<td>0.0010</td>
</tr>
</tbody>
</table>
Newton’s Method

<table>
<thead>
<tr>
<th></th>
<th>$x$</th>
<th>$E_t$</th>
<th>$E_a$</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>2.50000</td>
<td>1.07245</td>
<td></td>
</tr>
<tr>
<td>2</td>
<td>0.99508</td>
<td>0.43247</td>
<td>1.5049</td>
</tr>
<tr>
<td>3</td>
<td>1.46901</td>
<td>0.04146</td>
<td>0.4739</td>
</tr>
<tr>
<td>4</td>
<td>1.42764</td>
<td>$9 \times 10^{-5}$</td>
<td>0.0414</td>
</tr>
<tr>
<td>5</td>
<td>1.42755</td>
<td>$5 \times 10^{-10}$</td>
<td>$9 \times 10^{-5}$</td>
</tr>
</tbody>
</table>
Pitfalls of Newton’s Method

• Different starting points may lead to different solutions
• Iterations may diverge
• The former requires repeated search for the optimal optimum = global optimum
• The latter may require limiting step size, or requiring an increase in function value at each iteration
<table>
<thead>
<tr>
<th>Starting Point</th>
<th>Solution [x/f(x)]</th>
</tr>
</thead>
<tbody>
<tr>
<td>0</td>
<td>Oscillate</td>
</tr>
<tr>
<td>1</td>
<td>1.4276 / 1.7757</td>
</tr>
<tr>
<td>2</td>
<td>1.4276/ 1.7757</td>
</tr>
<tr>
<td>3</td>
<td>-1.7463/ -2.2742</td>
</tr>
<tr>
<td>4</td>
<td>5.2671/ -4.4723</td>
</tr>
<tr>
<td>5</td>
<td>5.2671/ -4.4723</td>
</tr>
<tr>
<td>6</td>
<td>5.2671/ -4.4723</td>
</tr>
<tr>
<td>7</td>
<td>7.0689/ -3.5823</td>
</tr>
<tr>
<td>8</td>
<td>7.0689/ -3.5823</td>
</tr>
</tbody>
</table>