Numerical methods and optimization

Week 3.

The Newton-method

$$x_{k+1} = x_k - \frac{f'(x_k)}{f''(x_k)}$$

and we assume that f is twice differentiable.
If the function being minimized $f(x)$ is not available in closed form or is difficult to differentiate, the derivatives $f'(x)$ and $f''(x)$ in this formula can be approximated by the finite difference formulas as

$$\begin{aligned} f'(x_k) &= \frac{f(x_k + \Delta_x) - f(x_k - \Delta_x)}{2\Delta_x} \\ f''(x_k) &= \frac{f(x_k + \Delta_x) - 2f(x_k) + f(x_k - \Delta_x)}{{\Delta_x}^2} \end{aligned}$$

quasi-Newton method

where Δ_x is a small step size. Substitution of these formulas into

$$x_{k+1} = x_k - \frac{f'(x_k)}{f''(x_k)}$$

leads to

$$x_{k+1} = x_k - \frac{\Delta_x[f(x_k + \Delta_x) - f(x_k - \Delta_x)]}{2[f(x_k + \Delta_x) - 2f(x_k) + f(x_k - \Delta_x)]}$$

The iterative process indicated by this equation is known as the **quasi-Newton method**. To test the convergence of the iterative process, the following criterion can be used:

$$|f'(x_{k+1})| \approx \left| \frac{f(x_k + \Delta_x) - f(x_k - \Delta_x)}{2\Delta_x} \right| \leq \varepsilon$$

Example

Find the minimum of the function

$$f(x) = 0.65 - \frac{0.75}{1 + x^2} - 0.65x \cdot \tan^{-1}\left(\frac{1}{x}\right)$$

using quasi-Newton method with the starting point $x_1 = 0.1$ and the step size $\Delta_x = 0.01$ in central difference formulas. Use $\varepsilon = 0.01$ for checking the convergence.

Solution:

Iteration 1:

$$x_1 = 0.1, \Delta_x = 0.01, \varepsilon = 0.01, f_1 = f(x_1) = -0.188197,$$

$$f_1^+ = f(x_1 + \Delta_x) = -0.195512, f_1^- = f(x_1 - \Delta_x) = -0.180615$$

$$x_2 = x_1 - \frac{\Delta_x(f_1^+ - f_1^-)}{2(f_1^+ - 2f_1 + f_1^-)} = 0.377882$$

Convergence check:

$$f_2^+ = f(x_2 + \Delta_x) = -0.304662, \qquad f_2^- = f(x_2 - \Delta_x) = -0.301916$$

$$|f'(x_2)| = \left| \frac{f_2^+ - f_2^-}{2\Delta_x} \right| = 0.137300 > \varepsilon$$

Iteration 2:

$$f_2 = f(x_2) = -0.303368, f_2^+ = f(x_2 + \Delta_x) = -0.304662,$$

$$f_2^- = f(x_2 - \Delta_x) = -0.301916$$

$$x_3 = x_2 - \frac{\Delta_x(f_2^+ - f_2^-)}{2(f_2^+ - 2f_2 + f_2^-)} = 0.465390$$

Convergence check:

$$f_3^+ = f(x_3 + \Delta_x) = -0.310004, \qquad f_3^- = f(x_3 - \Delta_x) = -0.309650$$

$$|f'(x_3)| = \left| \frac{f_3^+ - f_3^-}{2\Delta_x} \right| = 0.017700 > \varepsilon.$$

Iteration 3:

$$f_3 = f(x_3) = -0.309885, f_3^+ = f(x_3 + \Delta_x) = -0.310004,$$

$$f_3^- = f(x_3 - \Delta_x) = -0.309650$$

$$x_4 = x_3 - \frac{\Delta_x (f_3^+ - f_3^-)}{2(f_3^+ - 2f_3 + f_3^-)} = 0.480600.$$

Convergence check:

$$|f'(x_4)| = \left| \frac{f_4^+ - f_4^-}{2\Delta_x} \right| = 0.000350 < \varepsilon.$$

Since the process has converged, we take the optimum solution as

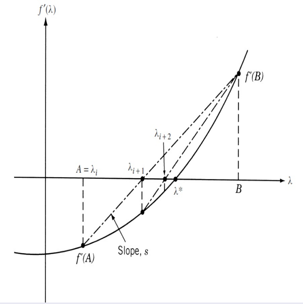
$$x^* \approx x_4 = 0.480600.$$

The secant method uses an equation similar to the previous as

$$f'(x) = f'(x_k) + s(x - x_k) = 0$$

where s is the slope of the line connecting the two points $(A, f'(A))$ and $(B, f'(B))$, where A and B denote two different approximations to the correct solution, x^*.

The slope s can be expressed as



$$s = \frac{f'(B) - f'(A)}{B - A}$$

Equation

$$f'(x) = f'(x_k) + s(x - x_k) = 0$$

approximates the function $f'(x)$ between A and B as a linear equation (secant), and hence the solution of the equation gives the new approximation to the root of $f'(x)$ as

$$x_{k+1} = x_k - \frac{f'(x_k)}{s} = A - \frac{f'(A)(B - A)}{f'(B) - f'(A)}$$

The iterative process given by this equation is known as the secant method. Since the secant approaches the second derivative of $f(x)$ at A as B approaches A, the secant method can also be considered as a quasi-Newton method. It can also be considered as a form of elimination technique since part of the interval, (A, x_{k+1}) in previous figure, is eliminated in every iteration.

The iterative process can be implemented by using the following step-by-step procedure. 1. Set $x_1 = A = 0$ and evaluate $f'(A)$. The value of $f'(A)$ will be negative. Assume an initial trial step length t_0. Set $k = 1$.
2. Evaluate $f'(t_0)$.
3. If $f'(t_0) < 0$, set $A = x_k = t_0$, $f'(A) = f'(t_0)$, new $t_0 = 2t_0$, and go to step 2.
4. If $f'(t_0) \geq 0$, set $B = t_0$, $f'(B) = f'(t_0)$, and go to step 5.
5. Find the new approximate solution of the problem as

$$x_{k+1} = A - \frac{f'(A)(B - A)}{f'(B) - f'(A)}$$

6. Test for convergence:

$$|f'(x_{k+1})| \leq \varepsilon$$

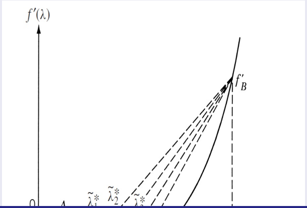
7. If $f'(x_{k+1}) \geq 0$, set new

$$B = x_{k+1}, f'(B) = f'(x_{k+1}), k = k+1,$$

and go to step 5.
8. If $f'(x_{k+1}) < 0$, set new

$$A = x_{k+1}, f'(A) = f'(x_{k+1}), k = k+1,$$

and go to step 7.



Remarks

1. The secant method is identical to assuming a linear equation for $f'(x)$. This implies that the original function $f(x)$, is approximated by a quadratic equation.
2. In some cases we may encounter a situation where the function $f'(x)$ varies very slowly with x, as shown in the previous figure. This situation can be identified by noticing that the point B remains unaltered for several consecutive refits. Once such a situation is suspected, the convergence process can be improved by taking the next value of x_{k+1} as $(A+B)/2$ instead of finding its value from the equation.

Example

Find the minimum of the function

$$f(x) = 0.65 - \frac{0.75}{1 + x^2} - 0.65x \cdot \tan^{-1}\left(\frac{1}{x}\right)$$

using the secant method with an initial step size of $t_0 = 0.1$, $x_1 = 0.0$, and $\varepsilon = 0.01$.

Solution:

$$x_1 = A = 0.0, t_0 = 0.1, f'(A) = -1.02102$$

$$B = A + t_0 = 0.1, f'(B) = -0.744832$$

Since

$$f'(B) < 0,$$

we set new

$$A = 0.1, f'(A) = -0.744832, t_0 = 2 \cdot 0.1 = 0.2, B = x_1 + x_0 = 0.2,$$

and compute $f'(B) = -0.490343$.

Since $f'(B) < 0$, we set new

$$A = 0.2, f'(A) = -0.490343, t_0 = 2 \cdot 0.2 = 0.4, B = x_1 + t_0 = 0.4,$$

and compute $f'(B) = -0.103652$.
Since $f'(B) < 0$, we set new

$$A = 0.4, f'(A) = -0.103652, t_0 = 2 \cdot 0.4 = 0.8, B = x_1 + t_0 = 0.8,$$

and compute $f'(B) = +0.180800$. Since $f'(B) > 0$, we proceed to find x_2.

Iteration 1

Since

$$A = x_1 = 0.4, f'(A) = -0.103652, B = 0.8, f'(B) = +0.180800,$$

we compute

$$x_2 = A - \frac{f'(A)(B-A)}{f'(B) - f'(A)} = 0.545757.$$

Convergence check:

$$|f'(x_2)| = |+0.0105789| > \varepsilon.$$

Iteration 2

Since $f'(x_2) = +0.0105789 > 0$, we set new

$$A = 0.4, f'(A) = -0.103652, B = x_2 = 0.545757,$$

$$f'(B) = f'(x_2) = +0.0105789,$$

and compute

$$x_3 = A - \frac{f'(A)(B - A)}{f'(B) - f'(A)} = 0.490632.$$

Convergence check: $|f'(x_3)| = |+0.00151235| < \varepsilon$. Since the process has converged, the optimum solution is given by $x^* \approx x_3 = 0.490632$.