Numeric Optimization - Assignment #01

103062512 徐丞裕

October 18, 2015

1. Solution:

   (a) Approximation of $f(x)$ at point x_k using second order Taylor series:

   (b) Find optimal point of $\tilde{f}$:

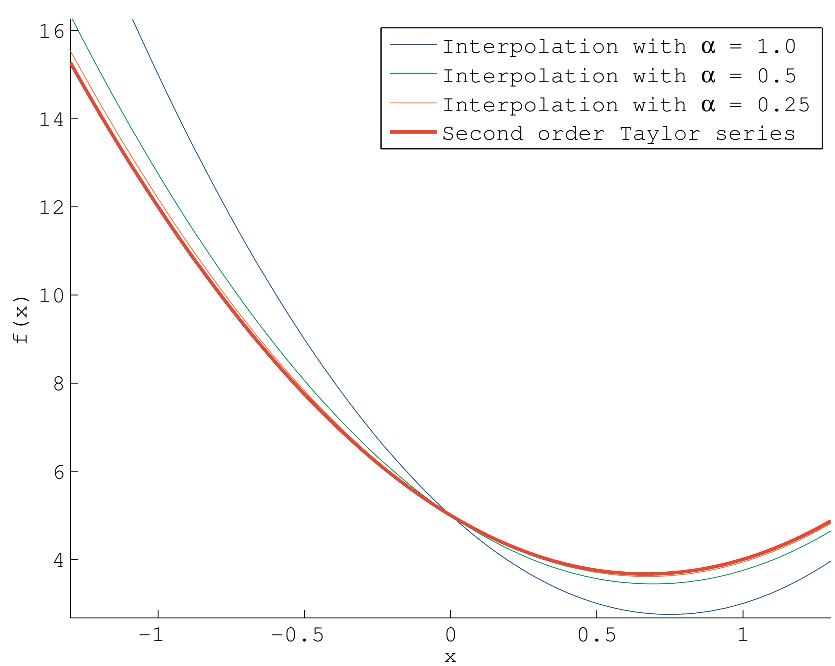
$$\begin{aligned}
\tilde{f}(x) &= \frac{f^{(2)}(x_k)}{2!}((x-x_k)^2 + 2\frac{f^{(1)}(x_k)}{f^{(2)}(x_k)}(x-x_k) + (\frac{f^{(1)}(x_k)}{f^{(2)}(x_k)})^2) + (f(x_k) - \frac{(f^{(1)}(x_k))^2}{2f^{(2)}(x_k)}) \\
\tilde{f}(x) &= \frac{f^{(2)}(x_k)}{2!}((x-x_k) + \frac{f^{(1)}(x_k)}{f^{(2)}(x_k)})^2 + (f(x_k) - \frac{(f^{(1)}(x_k))^2}{2f^{(2)}(x_k)}) \\
x^* &= x_k - \frac{f^{(1)}(x_k)}{f^{(2)}(x_k)}
\end{aligned}$$

   (c) Apply update rule of Newton's method:

$$\begin{aligned}
x_{k+1} &= x_k - H^{-1}g \\
x_{k+1} &= x_k - (f^{(2)}(x_k))^{-1} f^{(1)}(x_k) \\
x_{k+1} &= x_k - \frac{f^{(2)}(x_k)}{f^{(1)}(x_k)}
\end{aligned}$$

2. Solution:

   (a) With small α, the interpolation method is quite similar to Newton's method which also derive a quadratic model near x_k. To show this, please see the graph below:
   Target function: $f(x) = x^4 - 2x^2 + 3x^2 - 4x + 5$



As the behavior of interpolation method is quite similar to the Newton's method with α sufficiently small, one can fail the interpolation method with an example that will make Newton's method to fail.

To fail Newton's method, consider the following function:

$$w(x) = \begin{cases} \frac{2}{3}(-x)^{\frac{3}{2}} & x < 0 \\ 0 & x = 0 \\ \frac{2}{3}x^{\frac{3}{2}} & x > 0 \end{cases}$$

For $w(x)$, the update rule of Newton's method is:

$$x_{k+1} = \begin{cases} x_k - 2\frac{x^{\frac{1}{2}}}{x^{-\frac{1}{2}}} & x > 0 \\ x_k + 2\frac{(-x)^{\frac{1}{2}}}{(-x)^{-\frac{1}{2}}} & x < 0 \end{cases}$$

By choosing $x_0 = 3$, we have:

$$\begin{aligned} x_1 &= 3 - 6 \\ x_2 &= -3 + 6 \\ &\vdots \end{aligned}$$

which will not make Newton's method converge obviously.

(b) Pros & cons:

i. Pros of interpolation method compared to Newton's method:

A. Does not require any information about derivatives.

B. Has behavior of Newton's method when α is sufficiently small.

ii. Cons of interpolation method compared to Newton's method:

A. Solving the linear system may be more and more expensive as the dimension x_k grows.

(c) Let the cubic function used for approximating $f(x)$ be $n(x) = ax^3 + bx^2 + cx + d$, we can obtain a, b, c, and d by solving the following linear system:

$$\begin{bmatrix} x_1^3 & x_1^2 & x_1 & 1 \\ 3x_1^2 & 2x_1 & 1 & 0 \\ x_2^3 & x_2^2 & x_2 & 1 \\ 3x_2^2 & 2x_2 & 1 & 0 \end{bmatrix} \begin{bmatrix} a \\ b \\ c \\ d \end{bmatrix} = \begin{bmatrix} f(x_1) \\ f'(x_1) \\ f(x_2) \\ f'(x_2) \end{bmatrix}$$

3. Solution:

(a) Gradient g and Hessian matrix H of f:

$$g = \begin{pmatrix} \frac{\partial f}{\partial x} \\ \frac{\partial f}{\partial y} \end{pmatrix} = \begin{pmatrix} x \\ 9y \end{pmatrix}$$

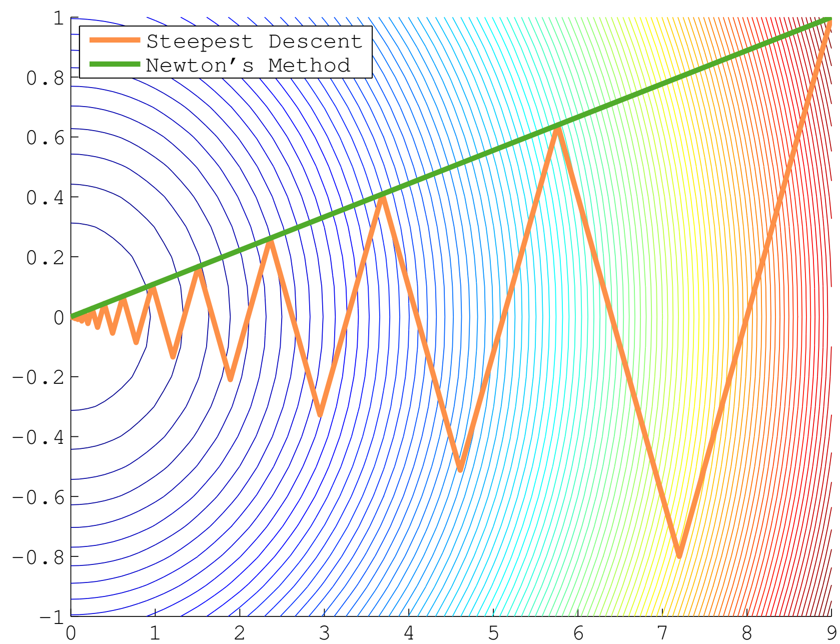
$$H = \begin{bmatrix} \frac{\partial^2 f}{\partial^2 x} & \frac{\partial^2 f}{\partial x \partial y} \\ \frac{\partial^2 f}{\partial y \partial x} & \frac{\partial^2 f}{\partial^2 x} \end{bmatrix} = \begin{bmatrix} 1 & 0 \\ 0 & 9 \end{bmatrix}$$

(b) Convergence criteria: $\|x_{k+1} - x_k\| < 10^{-4}$

i. Steepest descent: 46 iterations.

ii. Newton's method: 1 iteration.

(c) Traces of steepest descent and Newton's method:



Steepest Descent
Newton's Method
1
0.8
0.6
0.4
0.2
0
-0.2
-0.4
-0.6
-0.8
-1
0
1
2
3
4
5
6
7
8
9