Optimization Assignment - Deterministic Algorithm

HUU DANH NGUYEN - 1551023

16 February 2020

Problem:

Rosenbrock function:

$$f(x_1, x_2) = 100(x_2 - x_1^2)^2 + (1 - x_1)^2$$

(a) Solve the problem to minimize $foverR^2$ using the steepest descent and Newton's (unit step, Levenberg–Marquardt, and modified) methods. Start at the point (1.5, 1). Towards which point do the methods converge? How many iterations are required for the different methods?

For this function we have:

$$\Delta f(x_1, x_2) = \frac{-400x_1(x_2 - x_1^2) - 2(1 - x_1)}{200(x_2 - x_1^2)}$$

$$H(x_1, x_2) = \begin{bmatrix} 1200x_1^2 - 400x_2 + 2 & -400x_1 \\ -400x_1 & 200 \end{bmatrix}$$

We have f is always non-negative because f is a sum of squares. It vanishes only at point (1,1). f(x) = 0 with x is only a minimum point for f. Therefore, f must converge to this point.

1. Steepest descent:

   It calculates each step of x_{k+1} written as:

   $$x_{k+1} = x_k + \alpha_k d_k$$

   where α_k is a parameter that indicates the size of the step along the descent direction $-\Delta f(x_k)$ and is updated at each step to ensure that $f(x_k + 1)$ is less than $f(x_k)$ for every k while using the longest possible step along the descent direction. α_k is selected according to a line search along $\Delta f(x_k)$ performed using the backtracking algorithm.

The tolerance is: 10^{-12}
The program is started from point (-1.5,-1) and it converges in 10001 iterations and the converged point is (1.0005093709905537, 1.0010917350733282) and f at that point is 0.0000007884769490

2. Original Newton

   It calculates each step of x_{k+1} written as:

$$x_{k+1} = x_k + \alpha_k d_k$$

   where $d_k = -H^{-1}(x_k)\Delta f(x_k)$ and α_k is equal to 1 all steps
   Thus, it is able to be written as:

$$x_{k+1} = x_k - H^{-1}(x_k)\Delta f(x_k)$$

   The descent direction d_k can be calculated by solving the linear system $H(x_k)d_k = \Delta f(x_k)$ with the Gauss elimination method.

   The tolerance is: 10^{-12}
   The program is started from point (-1.5,-1) and it converges in 6 iterations and the converged point is (1.0, 1.0) and f at that point is 0.0

3. Modified Newton

   It calculates at each step of x_{k+1} written as:

$$x_{k+1} = x_k + \alpha_k d_k$$

   where $d_k = -H^{-1}(x_k)\Delta f(x_k)$ and α_k is selected according to a line search along $\Delta f(x_k)$ performed using the backtracking algorithm.

   The descent direction d_k can be calculated as the original newton.

   The tolerance is: 10^{-12}
   The program is started from point (-1.5,-1) and it converges in 35 iterations and the converged point is (1.0000000168799759, 1.0000000349589215) and f at that point is 0.0000000000000004

4. Levenberg-Marquardt

   This method is written as:

$$f(x_1, x_2) = f_1^2(x_1, x_2) + f_2^2(x_1, x_2) = [10(x_2 - x_1^2)]^2 + [(1 - x_1)]^2$$

   Its Jacobian is:

$$J(x_1, x_2) = \begin{pmatrix} \frac{df_1}{dx_1} & \frac{df_1}{dx_2} \\ \frac{df_2}{dx_1} & \frac{df_2}{dx_2} \end{pmatrix}$$

This method calculates at each step x_{k+1} written as:

$$x_{k+1} = x_k - (J^T J(x_k) + \lambda diag(J^T J(x_k)))^{-1} J^T(x_k) \cdot \begin{bmatrix} f_1(x_k) \\ f_2(x_k) \end{bmatrix}$$

Where λ is the damping parameter.
To update λ during the algorithm execution consists in evaluating if $f(x_{k+1}) < f(x_k)$ at each step: if so then x_{k+1} is accepted as the next iteration point and λ is set to $\frac{\lambda}{\nu}$. Otherwise x_{k+1} is discarded and λ is set to $\lambda\nu$.

We have descent direction d_k:

$$d_k = (J^T J(x_k) + \lambda diag(J^T J(x_k)))^{-1} J^T(x_k) \cdot \begin{bmatrix} f_1(x_k) \\ f_2(x_k) \end{bmatrix}$$

is also calculated with the Gauss elimination method.

$$(J^T J(x_k) + \lambda diag(J^T J(x_k)))d_k = J^T(x_k) \cdot \begin{bmatrix} f_1(x_k) \\ f_2(x_k) \end{bmatrix}$$

The tolerance is: 10^{-12}
The program is started from point (-1.5,-1) and it converges in 47 iterations and the converged point is (0.9999999992010885, 0.9999999984001754) and f at that point is 0.0000000000000000

(b) (Where) Is f convex? Is the point obtained a global minimum?

1. f is not convex on R^2

   f is convex if its Hessian matrix is positive. This is true if its principal minors are non-negative.
   We have its Hessian is:

   $$H(x_1, x_2) = \begin{bmatrix} 1200x_1^2 - 400x_2 + 2 & -400x_1 \\ -400x_1 & 200 \end{bmatrix}$$

   The principal minors are:

   $$a = 1200x_1^2 - 400x_2 + 2, c = 200, ac - b^2$$

   Where $b = -400x_1$.
   a is non-negative if $x_2 \leq 3x_1^2 + \frac{1}{200}$
   $ac - b^2$ is non-negative if $x_2 \leq x_1^2 + \frac{1}{200}$.
   Therefore, f is convex only in $(x_1, x_2) \in R^2 | x_2 \leq 3x_1^2 + \frac{1}{200}, x_2 \leq x_1^2 + \frac{1}{200}$

2. Is the point obtained a global minimum?
   No, it is not. The used methods converge only to a local minimum of the function.