SI 211: Numerical Analysis Homework 7

Prof. Boris Houska

Due on Dec 2, 2020, 23:59

Name: Chengrui Zhang, Student Number: 2019233183

1. P1
   Ans:

   - a): The main idea of Newtons method is to use iteration method to solve the nonlinear equation $f(x) = 0$. The iteration method start with an initial guess x_0 and then solve the linear equation sustems: $f(x_k) + M(x_k)(x_{k+1} - x_k) = 0$, for $k \in \{0, 1, 2, \cdots\}$. Here, the matrix $M(x_k) \in R^{n\times n}$ is chosen by $f(x_k) + M(x_k)(x_{k+1} - x_k) \approx f(x)$.
   - b): The main idea of Newtons method is also to use iteration method to solve the nonlinear equation $F'(x) = 0, (F(x) = \frac{1}{2}||f(x)||_2^2)$. The difference between Gauss Newtons method and Newtons method is the step direction. In the Newtons method, $\Delta x_k = -M(x_k)^{-1}F'(x_k)^T = -M(x_k)^{-1}f'(x_k)^T f(x_k), M = Hessian$, while in the Gauss Newtons method, it uses $M(x_k) = f'(x_k)^T f'(x_k)$ instead of computing Hessian matrix.
   - c): Since $||x_{k+1} - x^*|| \leq \kappa||x_k - x^*|| + \frac{w}{2}||x_k - x^*||_2^2$, we know that 1) if $\kappa \neq 0$, the convergence rate is in general linear; 2) if we choose $M(x_k) = J(x_k)$, we have $\kappa = 0$ and $||x_{k+1} - x^*|| \leq \frac{w}{2}||x_k - x^*||_2^2$, and in this case, the convergence rate is called quadratic.
   - d): For the problem $min_x F(x)$ with $F(x) = \frac{1}{2}||f(x)||_2^2$, if we have $f''(x) = 0$, then the algorithm will converge in one step.
   - e): The Armijo line search condition is a condition that limit the search region of the line search optimization problem $min_{\alpha_k \in [0,1]} F(x_k + \alpha_k \Delta x_k)$, the condition is:
     $F(x_k + \alpha_k \Delta_k) \leq F(x_k) + c\alpha_k F'(x_k)\Delta x_k$.
     This condition can ensure that the line search parameter is not excessively large and if M is positive definite, this condition ensures that we get a strict descent of the objective function.

2. P2
Ans:

2.

$f(x) = x^3 - 2x + 2 \quad \therefore f'(x) = 3x^2 - 2$

$$\therefore x_{k+1} = x_k - \frac{1}{f'(x_k)} \cdot f(x_k) = x_k - \frac{x_k^3 - 2x_k + 2}{3x_k^2 - 2}$$

$$= \frac{1}{3x_k^2 - 2} \cdot \left[2x_k^3 - 2\right]$$

$$\because x_0 = 0 \Downarrow x_1 = 1 \Downarrow x_2 = 0 \Downarrow x_3 = 1 \Downarrow \cdots$$

$\therefore$ if we chose x_0 be the initial point, the newton's method will not converge

3. P3

Ans:

- a):

$$f(x,y) = \frac{1}{2}(x-1)^2 + \frac{1}{2}(10(y-x^2))^2 + \frac{1}{2}y^2$$

$$J = \nabla f(x,y) = \left[\frac{\partial f}{\partial x}, \frac{\partial f}{\partial y}\right]^T$$

$$= \begin{bmatrix} x-1 + 10(y-x^2)\cdot(-20x) \\ 10(y-x^2)\cdot 10 + y \end{bmatrix}^T$$

$$= \begin{bmatrix} 200x^3 - 200xy + x - 1 \\ -100x^2 + 101y \end{bmatrix}$$

$$H = \nabla^2 f(x,y) = \begin{bmatrix} \frac{\partial^2 f}{\partial x^2} & \frac{\partial^2 f}{\partial x \partial y} \\ \frac{\partial^2 f}{\partial y \partial x} & \frac{\partial^2 f}{\partial y^2} \end{bmatrix}$$

$$= \begin{bmatrix} 600x^2 - 200y + 1 & -200x \\ -200x & 101 \end{bmatrix}$$

$$f(x,y) = \frac{1}{2}\|[x-1, 10(y-x^2), y]\|_2^2 \qquad R(x,y) = [x-1, 10(y-x^2), y]$$

$$J_R = \begin{bmatrix} 1 & 0 \\ -20x & 10 \\ 0 & 1 \end{bmatrix} \qquad M = J_R^T \cdot J_R = \begin{bmatrix} 400x^2+1 & -200x \\ -200x & 101 \end{bmatrix}$$

$$\therefore M = H \text{ if and only if } 200x^2 - 200y = 0$$

$$\iff y = x^2$$

- b): The code is shown in **HW7_3.m**. The track of this method is shown in Table. 1, and the figure is shown in Fig. 1

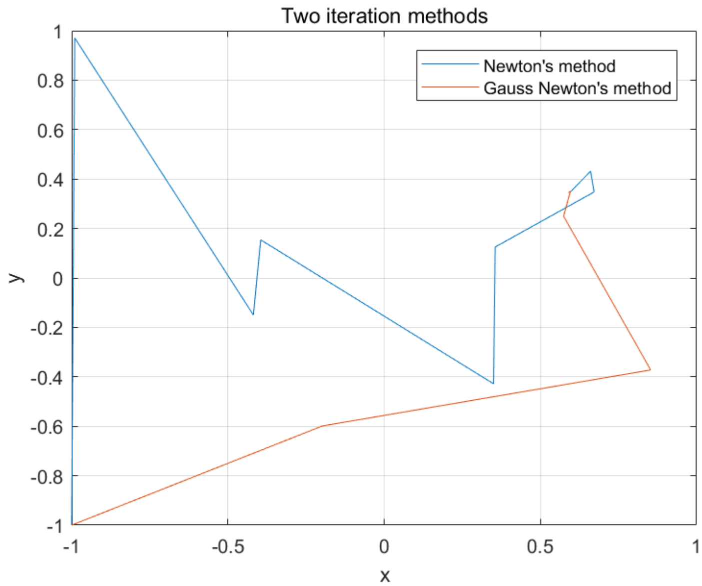

Figure 1: Two iteration methods

iteration	1	2	3	4	5	6	7	8	9	10	11
x_1	-1	-0.9902	-0.4185	-0.3950	0.3502	0.3553	0.6721	0.6605	0.5966	0.5922	0.5911
y_1	-1	0.9706	-0.1503	0.1540	-0.4285	0.1250	0.3478	0.4318	0.3484	0.3472	0.3459

Table 1: Track of Newtons method

- c): The code is shown in **HW7_3.m**. The track of this method is shown in Table. 2, and the difference between exact and approximate Hessian is shown in Fig. 2

iteration	1	2	3	4	5	6	7	8
x_2	-1	-0.1976	0.8528	0.5742	0.5963	0.5896	0.5915	0.5910
y_2	-1	-0.5988	-0.3724	0.2496	0.3515	0.3441	0.3464	0.3458

Table 2: Track of Gauss Newtons method

- d): The above two methods are all have good performance, the final results are nearly the same. The iteration paths is shown in Fig. 1, we can find the Gauss Newtons method is more directly to the result. Newtons method needs 11 iterations while the Gauss Newtons method only needs 8 iterations. The runtime of these two methods are 18.1ms and 3.8ms respectively. Gauss Newtons method is fast than the Newtons method.

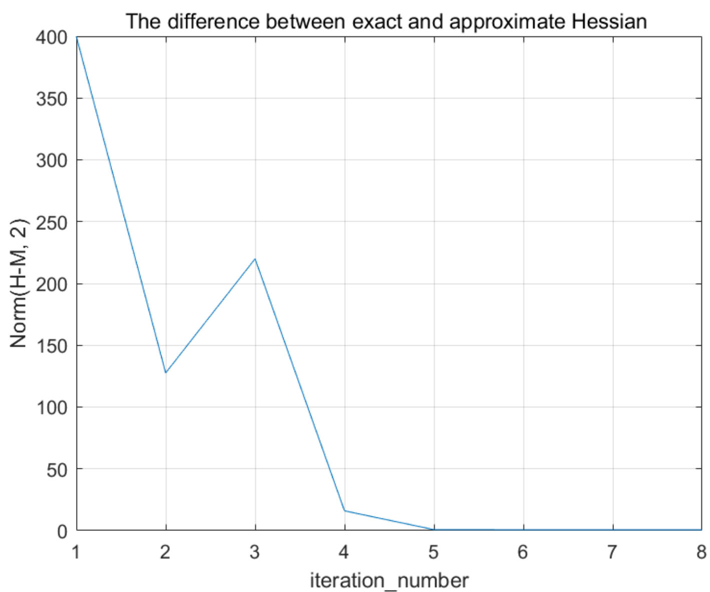

Figure 2: The difference between exact and approximate Hessian