Lecture 9
Newton's Method

Jinn-Liang Liu
2017/6/1

Newton's method is a method for finding successively approximations to a root (*unknown solution*) x^* of a nonlinear equation

$$g(x) = 0. \tag{9.1}$$

We first choose an initial guess $x^{(0)}$ that is sufficiently close to x^* and then determine the tangent line to the curve of $g(x)$ at the point $\left(x^{(0)}, g(x^{(0)})\right)$ with the slope of $g'(x^{(0)})$. This line intercepts the x-axis at $x^{(1)}$ which is expected to be closer to x^* than $x^{(0)}$. Then we have another tangent line through $\left(x^{(1)}, g(x^{(1)})\right)$ and obtain $x^{(2)}$ and so on. In other words, we are iteratively solving the following linearized equation

$$\begin{aligned} g'(x^{(0)})w &= g(x^{(0)}), \;\; w = x^{(0)} - x^{(1)}, && (9.2\text{a}) \\ g'(x^{(0)})w &= \lim_{t \to 0} \frac{g(x^{(0)} + tw) - g(x^{(0)})}{t}, \text{ or} \\ -g'(x^{(0)})x^{(1)} &= -g'(x^{(0)})x^{(0)} + g(x^{(0)}) && (9.2\text{b}) \end{aligned}$$

where $x^{(1)}$ is next iterate (unknown) to be solved then $x^{(2)}$ (with $x^{(0)}$ replaced by $x^{(1)}$) and so on. Newton's method thus generates a sequence of approximate solutions $\left\{x^{(n)}\right\}_{n=0}^{\infty}$ to the exact solution x^*. Three main questions concerning with this method are:

(1) How do we guarantee the convergence, i.e., $\lim_{n\to\infty} x^{(n)} = x^*$ with what initial guess $x^{(0)}$?

(2) How fast is the convergence? (Answer: The convergence order is two (quadratic convergence) if the sequence converges.)

(3) How do we generalize this idea to a nonlinear PDE? This lecture is concerned with this question.

Question 9.1. What is linearity? For $f_1(x) = x$, $f_2(x) = x^2$, $f_3(x) = e^x$, $f_4(X) = AX$ with $A = \begin{bmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{bmatrix}$, $X = \begin{bmatrix} x_1 \\ x_2 \end{bmatrix}$, $f_5(u) = \frac{du(x)}{dx}$, $f_6(u) = u\frac{du}{dx}$, $f_7(u) = -\frac{d^2u}{dx^2}$, $f_8(u) = \int u(x)dx$, which one is linear?

HW 9.1. Read Newton's method from, for example, Wikipedia and prove (1) and (2).

Example 9.1. Consider first the nonlinear equation

$$ax = e^x =: f(x) \tag{9.3}$$

where $a \neq 0$ is a scalar. Here, we can think a as a linear operator acting on x. Then

$$\begin{aligned} g(x) &= ax - f(x) = 0, \;\; g'(x) = a - f'(x) = a - e^x & (9.4) \\ g'(x^{(0)}) &= \frac{g(x^{(0)}) - 0}{x^{(0)} - x^{(1)}} & (9.5) \end{aligned}$$

$$\begin{aligned} g'(x^{(0)})w &= g(x^{(0)}), \;\; w = x^{(0)} - x^{(1)}, \text{ or} & (9.6a) \\ \left[a - f'(x^{(0)})\right] x^{(1)} &= -f'(x^{(0)})x^{(0)} + f(x^{(0)}), & (9.6b) \end{aligned}$$

where the last equation is a linearized equation.

Example 9.2. We next consider the following coupled nonlinear system with two *unknown solutions* (independent variables) (x_1, x_2), the linear operator A (a matrix), and two nonlinear functions $f_1(x_1, x_2)$ and $f_2(x_1, x_2)$,

$$\begin{cases} a_{11}x_1 + a_{12}x_2 = f_1(x_1, x_2) \\ a_{21}x_1 + a_{22}x_2 = f_2(x_1, x_2) \end{cases} \tag{9.7}$$

$$\begin{bmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = \begin{bmatrix} f_1(x_1, x_2) \\ f_2(x_1, x_2) \end{bmatrix} \tag{9.8}$$

$$AX = F(X), \;\; G(X) = AX - F(X) = 0 \tag{9.9}$$

$$A = \begin{bmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{bmatrix}, \quad X = \begin{bmatrix} x_1 \\ x_2 \end{bmatrix}, \quad F(X) = \begin{bmatrix} f_1(x_1, x_2) \\ f_2(x_1, x_2) \end{bmatrix}$$

$$F'(X) := J(X) = J(x_1, x_2) = \begin{bmatrix} \frac{\partial f_1}{\partial x_1} & \frac{\partial f_1}{\partial x_2} \\ \frac{\partial f_2}{\partial x_1} & \frac{\partial f_2}{\partial x_2} \end{bmatrix} \;\; \text{(Jacobian operator)} \tag{9.10}$$

$$G'(X) = A - F'(X) \tag{9.11}$$

$$\begin{aligned} G'(X^{(0)})W &= G(X^{(0)}), \;\; W = X^{(0)} - X^{(1)}, \text{ or} & (9.12a) \\ \left[A - F'(X^{(0)})\right] X^{(1)} &= -F'(X^{(0)})X^{(0)} + F(X^{(0)}), & (9.12b) \end{aligned}$$

HW 9.2. Prove (9.12a) using (9.2a). Write (9.12b) explicitly in a system of two linear equations.

Example 9.3. This last example is an elliptic semilinear (nonlinear) PDE with now an *unknown solution* $u(x)$, the positive linear operator $-\frac{d^2}{dx^2}$ (acting on ?), and the nonlinear functional (a function of a function) $f(u)$,

$$-u''(x) = f(u) = e^u \tag{9.13}$$

$$G(u) = -u''(x) - f(u) = 0 \tag{9.14}$$

$$\begin{aligned} G'(u)w \ :&= \lim_{t\longrightarrow 0} \frac{G(u+tw)-G(u)}{t} \quad (G'(u) \text{ is an operator}) \qquad (9.15\text{a})\\ &= \lim_{t\longrightarrow 0} \frac{[-u''(x)-tw''-f(u+tw)]-[-u''(x)-f(u)]}{t}\\ &= -w'' - \lim_{t\longrightarrow 0} \frac{f(u+tw)-f(u)}{t}\\ &= -w'' - \lim_{t\longrightarrow 0} \frac{f(u+tw)-f(u)}{tw} w\\ &= -w'' - f'(u)w \qquad (9.15\text{b}) \end{aligned}$$

Note that the differentiations in w'' and $f'(u)$ are different, i.e., $w' = \frac{dw(x)}{dx}$ and $f'(u) = \frac{df(u)}{du}$ etc. The linearized problem of (9.13) is thus

$$G'(u^{(0)})w = -w'' - f'(u^{(0)})w = G(u^{(0)}), \ \ w = u^{(0)} - u^{(1)} \tag{9.16}$$

If (9.16) is discretized by, for example, the finite difference method (FDM), we obtain a linear system like (9.12) as

$$\begin{aligned} G'(U^{(0)})W &= G(U^{(0)}), \ \ W = U^{(0)} - U^{(1)} &\quad (9.17\text{a})\\ \left[A - F'(U^{(0)})\right] U^{(1)} &= -F'(U^{(0)})U^{(0)} + F(U^{(0)}) &\quad (9.17\text{b}) \end{aligned}$$

where $G'(U^{(0)}) = A - F'(U^{(0)})$, A is a coefficient matrix corresponding to the discretization of $-w''$, $W = [W_1, \cdots, W_N]^T$, $F'(U^{(0)})$ is a diagonal matrix with entries $d_i = f'(U^{(0)}(x_i))$ with $U^{(0)}(x_i) =: U_i^{(0)} \approx u(x_i)$, $U^{(0)} = \left[U_1^{(0)}, \cdots, U_N^{(0)}\right]^T$, $G(U^{(0)}) = AU^{(0)} - F(U^{(0)})$, and $F(U^{(0)}) = \left[f(U_1^{(0)}), \cdots, f(U_N^{(0)})\right]^T$.

The monotone iterative method with FDM for (9.13) is to replace (9.18) by a more general form

$$[A - D]\, U^{(1)} = -DU^{(0)} + F(U^{(0)}) \tag{9.18}$$

where the matrix D can be a constant diagonal matrix or a variable diagonal matrix. Of course if $D = F'(U^{(0)})$, we have Newton's method.

Project 9.1. Consider the following 1D nonlinear Poisson problem

$$\begin{aligned} -u''(x) &= \lambda e^u \ , \quad \forall\, x \in (0,1) && (9.19a) \\ u(0) &= u(1) = 0 && (9.19b) \end{aligned}$$

where the exact solution $u(x) = \ln\left[\cosh^2\left(\frac{\mu}{2}\right) \cdot \cosh^{-2}\left(\mu\left(x - \frac{1}{2}\right)\right)\right]$, $\mu = 2$, $\lambda = 2\mu^2 \cosh^{-2}\left(\frac{\mu}{2}\right)$, and $\cosh(\theta) = \frac{e^\theta + e^{-\theta}}{2}$. Implement the central finite difference method, the conjugate gradient method, and Newton's method (NM) to find approximate solutions $U(x)$ to $u(x)$. Given a total number of nodes N, the mesh size $\Delta x = h = \frac{1}{N-1}$. The maximum error of an approximate solution $U(x)$ is defined as

$$\begin{aligned} E^u &= ||e(x)||_\infty = ||u(x) - U(x)||_\infty \\ &= \max_{1 \le i \le N} |e_i| = \max_{1 \le i \le N} |u_i - U_i| \,. && (9.20) \end{aligned}$$

Example: For $N = 5$, we have $h = \frac{1}{4}$ and

$$A = \begin{bmatrix} 1 & 0 & 0 & 0 & 0 \\ \frac{-1}{h^2} & \frac{2}{h^2} & \frac{-1}{h^2} & 0 & 0 \\ 0 & \frac{-1}{h^2} & \frac{2}{h^2} & \frac{-1}{h^2} & 0 \\ 0 & 0 & \frac{-1}{h^2} & \frac{2}{h^2} & \frac{-1}{h^2} \\ 0 & 0 & 0 & 0 & 1 \end{bmatrix} \text{(See (1.17))}, \qquad (9.21)$$

$$\begin{aligned} U^{(0)} &= \begin{bmatrix} U_1^{(0)} \\ U_2^{(0)} \\ U_3^{(0)} \\ U_4^{(0)} \\ U_5^{(0)} \end{bmatrix} = \begin{bmatrix} 0 \\ \frac{1}{4} \\ \frac{1}{2} \\ \frac{1}{4} \\ 0 \end{bmatrix} \text{ or } = \begin{bmatrix} 0 \\ 0 \\ 0 \\ 0 \\ 0 \end{bmatrix}, \quad \overrightarrow{x} = \begin{bmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \\ x_5 \end{bmatrix} = \begin{bmatrix} 0 \\ \frac{1}{4} \\ \frac{1}{2} \\ \frac{3}{4} \\ 1 \end{bmatrix}, \\ \overrightarrow{u} &= \begin{bmatrix} u_1 = u(x_1) \\ u_2 = u(x_2) \\ u_3 = u(x_3) \\ u_4 = u(x_4) \\ u_5 = u(x_5) \end{bmatrix} && (9.22) \end{aligned}$$

$$
\begin{aligned}
D &= \begin{bmatrix} 0 & 0 & 0 & 0 & 0 \\ 0 & d_2 & 0 & 0 & 0 \\ 0 & 0 & d_3 & 0 & 0 \\ 0 & 0 & 0 & d_4 & 0 \\ 0 & 0 & 0 & 0 & 0 \end{bmatrix}, \\
d_2 &= f'(U^{(0)}(x_2)) = \lambda e^{U^{(0)}(x_2)} = \lambda e^{U_2^{(0)}} = \lambda e^{\frac{1}{4}} \qquad (9.23)
\end{aligned}
$$

$$
\begin{aligned}
F(U^{(0)}) &= \begin{bmatrix} 0 \\ F_2 \\ F_3 \\ F_4 \\ 0 \end{bmatrix}, \\
F_2 &= f(U^{(0)}(x_2)) = \lambda e^{U^{(0)}(x_2)} = \lambda e^{U_2^{(0)}} = \lambda e^{\frac{1}{4}} = d_2 \qquad (9.24)
\end{aligned}
$$

$$
[A - D]\, U^{(1)} = -DU^{(0)} + F(U^{(0)}) \qquad \Longrightarrow \qquad \widehat{A}\, U^{(1)} = \overrightarrow{b} \qquad (9.25)
$$

$$
\widehat{A} = A - D,\ \overrightarrow{b} = \begin{bmatrix} 0 \\ b_2 \\ b_3 \\ b_4 \\ 0 \end{bmatrix},\ \ b_2 = -d_2 U_2^{(0)} + d_2 \qquad (9.26)
$$

Algorithm NM: Newton's Method Find approximate solution $U(x)$ of (9.19).

Input: N, $l\max$, $EUTol$.

Step 1. Set A, $U^{(0)}$, $\overrightarrow{x}$ in (9.21) and (9.22).

Step 2. Set the exact solution $\overrightarrow{u}$ in (9.22).

Step 3. For $l = 1, \cdots, l\max$, do Step 4-8. (Newton's iteration.)

Step 4. Set the diagonal matrix (vector) D in (9.23).

Step 5. Set $\widehat{A}$ and $\overrightarrow{b}$ in (9.25).

Step 6. Call CG or SOR or JM to solve $\widehat{A}\,U^{(1)} = \overrightarrow{b}$.

Step 7. Compute $EU = \max_{1 \le i \le N} \left| U_i^{(1)} - U_i^{(0)} \right|$.

Step 8. If $EU < EUTol$ then compute $E^u = \max_{1 \le i \le N} \left| u_i - U_i^{(1)} \right|$ and stop, else set $U^{(0)} = U^{(1)}$ and go to Step 3.

Output:

N	l	E^u
5		
9		
17		
33		
65		
129		

Question 9.2. (1) Verify the exact solution satisfy (9.19). (2) Write (9.17b) explicitly in a system of five linear equations for the problem in Project 9.1, i.e., $N = 5$. (3) How do you choose an initial guess for Newton's method in your program?

HW 9.3. Can you show numerically the quadratic convergence by your program? Answer this question by writing the mathematical formulas before programming.

HW 9.4. Use (9.15a) as the definition of the operator $G'(u)w$ to prove that (a) $g'(x^{(0)})w = \left[a - f'(x^{(0)})\right] w$ in (9.2) with $u = x^{(0)}$ and $G'(u) = g'(x^{(0)})$, (b) $G'(X^{(0)})W = \left[A - F'(X^{(0)})\right] W$ in (9.12) with $u = X^{(0)}$ and $G'(u) = G'(X^{(0)})$.