Queens College, CUNY, Department of Computer Science

Numerical Methods

CSCI 361 / 761

Fall 2017

Instructor: Dr. Sateesh Mane

October 4, 2017

4 Lecture 4

4.1 Root-finding in multi-dimensions

- This section is about finding a root of a nonlinear equation $f(\boldsymbol{x}) = 0$ where $\boldsymbol{x} = \{x_1, x_2, \dots, x_n\}$ is a multi-dimensional vector (array). An important detail to bear in mind: the function "f" does not have to be a scalar. It could also be multi-dimensional:

$$\boldsymbol{f}(\boldsymbol{x}) = \mathbf{0} \,. \tag{4.1.1}$$

  Then we have to solve n simultaneous equations in n unknowns (for example). All of this can be very difficult.

- Root-finding in multi-dimensions is **much** harder than in one dimension.

- Why? Consider that bisection will not work, even in two dimensions (x, y). How do we bracket a root $(x_{\rm root}, y_{\rm root})$? At the least, we would have to find a "box" (not just two points) to enclose a root. But how would we formulate conditions for the function f to "change sign" from one side of the box to another?

- Hence bisection is out. What about Newton-Raphson?

- Newton-Raphson will work. Or at least, Newton-Raphson *can* work. Let us study Newton-Raphson in multi-dimensions.

4.2 Newton-Raphson in multi-dimensions

- Let us recall the Newton-Raphson iteration formula in one dimension. The iteration requires a division by $f'(x)$. This will not work in multi-dimensions. How do we divide by f' in multi-dimensions?

- We return to the Taylor series expression. Let us use the variable $i = 1, 2, \dots, n$ to index the equations and variables. We have n equations, of the form

$$f_i(\boldsymbol{x} + \Delta\boldsymbol{x}) = f_i(\boldsymbol{x}) + \sum_{j=1}^{n} \frac{\partial f_i}{\partial x_j}\,\Delta x_j + O(|\Delta\boldsymbol{x}|^2) \qquad (i = 1, 2, \dots, n)\,. \tag{4.2.1}$$

- The **Jacobian** matrix $\boldsymbol{J}$ is defined via the set of partial derivatives

$$J_{ij} = \frac{\partial f_i}{\partial x_j} \qquad\qquad (1 \le i, j \le n)\,. \tag{4.2.2}$$

- Hence the above equation can be expressed in matrix/vector notation as

$$\boldsymbol{f}(\boldsymbol{x} + \Delta\boldsymbol{x}) = \boldsymbol{f}(\boldsymbol{x}) + \boldsymbol{J}\cdot\Delta\boldsymbol{x} + O(|\Delta\boldsymbol{x}|^2)\,. \tag{4.2.3}$$

- We now follow the prescription for the one-dimensional Newton-Raphson case. We set the left hand side to zero and we say the remainder term (on the right) is negligible. This yields the multi-dimensional iteration formula

$$\boldsymbol{f}(\boldsymbol{x}) + \boldsymbol{J}\cdot\Delta\boldsymbol{x} = 0\,. \tag{4.2.4}$$

  Rearrange terms to obtain

$$\boldsymbol{J}\cdot\Delta\boldsymbol{x} = -\boldsymbol{f}(\boldsymbol{x})\,. \tag{4.2.5}$$

  This equation must be solved for $\Delta\boldsymbol{x}$ using matrix techniques.

- Formally we can write

$$\Delta\boldsymbol{x} = -\boldsymbol{J}^{-1}\cdot\boldsymbol{f}(\boldsymbol{x})\,. \tag{4.2.6}$$

  However this is a terrible way to solve matrix equations, in general. There are much better techniques, which should be employed.

- The "avoid division by zero" problem $f'(x) \ne 0$ is generalized to the condition $\det(\boldsymbol{J}) \ne 0$.

- Hence we iterate

$$\boldsymbol{x}_{\text{new}} = \boldsymbol{x}_{\text{old}} + \Delta\boldsymbol{x}\,. \tag{4.2.7}$$

- We must specify a convergence criterion. It may not necessarily be the same in all dimensions.

- In general, this is all much more complicated than the one-dimensional case.