ECE 4960

Spring 2017

Lecture 14

Nonlinear Equations and Optimization:

Newton-Like Methods

Edwin C. Kan

School of Electrical and Computer Engineering

Cornell University

Observation on the Jacobian Matrix in the Newton's Method

1. If x is transformed to $ax + b$ with a and b being constant, or in the vector case: where *[A]* is a constant nondegenerate matrix and is a constant vector, the Jacobian evaluation and the search steps will NOT change at all. The Newton method has "**affine invariance**".
2. *[J]* can be sparse if there are many f_i's having a zero coefficient (i.e., no dependence or correlation) on x_j.
3. $[J]^{-1}$ will be best calculated directly or symbolically. The more accurate we can estimate J_{ij}, the better search we get (i.e., if the function is linear to that variable, we will get the correct solution in one step).

The Quasi –Newton Method

- The **Quasi-Newton** method: If the analytical expression of J_{ij} is not available, we can evaluate J_{ij} numerically, similar to our local analysis to whatever accuracy with a small Δx_i.

$$J_{ij} \equiv \frac{\partial f_j}{\partial x_i} \cong \frac{f_j\left(x_1, x_2, \ldots, x_i + \Delta x_i, \ldots, x_n\right) - f_j\left(x_1, x_2, \ldots, x_i, \ldots, x_n\right)}{\Delta x_i}$$

- Notice that each J_{ij} needs an evaluation of f, which means the formulation of the Jacobian is quite expensive, even if it is sparse.
- The quasi-Newton method will show quadratic convergence in the basin of attraction. However, due to possible errors in the local analysis, the basin of attraction may not be the same.

The Secant Method

- The **Secant** method: We can also estimate the “gradient” from two previous guesses.

$$J_{ij} \equiv \frac{\partial f_j}{\partial x_i} \cong \frac{f_j\left(x_1^{(k)}, x_2^{(k)}, \ldots, x_n^{(k)}\right) - f_j\left(x_1^{(k-1)}, x_2^{(k-1)}, \ldots, x_n^{(k-1)}\right)}{x_i^{(k)} - x_i^{(k-1)}}$$

- Notice that the numerator in the right-hand side is the same for the same *j*, which makes the computation very cheap.
- The secant method is notoriously slow due to the inaccuracy introduced in the Jacobian evaluation, and often will not show quadratic convergence.
- The line-search method is usually necessary to stabilize the search at the expense of more evaluation of *f*.
- Need two initial guesses to start iteration.

Hacker Practice

Use the quasi-Newton method with line search to solve the same nonlinear equation by making $x^{(0)} = 1$ and the local analysis of the Jacobian matrix by 10^{-4} perturbation.

$$f(x) = e^{100x} - 1 = 0$$

$$\Delta x^{(k)} = -\left[\frac{f\left(1.0001x^{(k)}\right) - f\left(x^{(k)}\right)}{0.0001x^{(k)}}\right]^{-1} f\left(x^{(k)}\right)$$

Report $x^{(k)}$, $\Delta x^{(k)}$, $f(x^{(k)})$.

What is the convergence behavior in comparison with the Newton method with the Jacobian matrix evaluated symbolically?