CS 450: Numerical Analysis
Lecture 14
Chapter 5 – Nonlinear Equations
Newton’s Method for Systems of Nonlinear Equations

Edgar Solomonik

Department of Computer Science
University of Illinois at Urbana-Champaign
Newton’s and secant method provide basic approaches for solving a univariate nonlinear equation:

\[
x_{k+1}^{\text{Newton}} = x_k - \frac{f(x_k)}{f'(x_k)}
\]

\[
x_{k+1}^{\text{Secant}} = x_k - \frac{f(x_k) \cdot \frac{x_k - x_{k-1}}{f(x_k) - f(x_{k-1})}}
\]

Inverse (quadratic) interpolation can provide better convergence:

Interpolate quadratic polynomial \( q \) so that \( q(f(x_i)) = x_i \) for \( i \in \{k, k-1, k-2\} \), then pick approximate root as \( x_{k+1} = q(0) \).
Systems of Nonlinear Equations

Given \( f(x) = \begin{bmatrix} f_1(x) \\ \vdots \\ f_m(x) \end{bmatrix} \) for \( x \in \mathbb{R}^n \), seek \( x^* \in \mathbb{R}^n \) so that \( f(x^*) = 0 \).

\( x^* \) must simultaneously set to zero all components of \( f \):
\[ f_1(x^*) = \cdots = f_m(x^*) = 0. \]

At a particular point \( x \), the Jacobian of \( f \), describes how \( f \) changes in a given direction of change in \( x \),
\[
J_f(x) = \begin{bmatrix}
\frac{df_1}{dx_1}(x) & \cdots & \frac{df_1}{dx_n}(x) \\
\vdots & & \vdots \\
\frac{df_m}{dx_1}(x) & \cdots & \frac{df_m}{dx_n}(x)
\end{bmatrix}
\]

Our local approximation is given by
\[
f(x + \delta x) \approx f(x) + J_f(x)\delta x,
\]

note that when \( m = 1 \) the Jacobian corresponds to the gradient of \( f \).
Multivariate Fixed-Point and Newton Iteration

- Fixed-point iteration $x_{k+1} = g(x_k)$ achieves local convergence so long as $|\lambda_{\text{max}}(J_g(x^*))| < 1$:
  
  \[ \text{Given starting point } x_0 \text{ close enough to } x^*, \text{ we will have } |\lambda_{\text{max}}(J_g(x_i))|, \forall i. \]

- Newton’s method corresponds to the fixed-point iteration
  
  \[ g(x) = x - J_f^{-1}(x)f(x) \]

  Note that generally Newton’s method iteration has a fixed-point $\bar{x}$ whenever $f(\bar{x}) = 0$, i.e. we have found a root of $f$, namely $x^* = \bar{x}$.
  
  A necessary assumption is that $J_f(x^*)$ is nonsingular, otherwise we can find nonzero solutions $y$ to $J_f(x^*)y = f(x^*) = 0$. 

Convergence of Newton Iteration

- Newton’s method achieves quadratic local convergence if \( ||J^{-1}_f(x^*)|| \) is bounded:

\[
e_k = x_k - x^* = g(x_{k-1}) - x^* \\
= x_{k-1} - J^{-1}_f(x_{k-1})f(x_{k-1}) - x^* \\
= J^{-1}_f(x_{k-1})(f(x_{k-1}) - J_f(x_{k-1})(x^* - x_{k-1})) \\
||e_k|| \leq ||J^{-1}_f(x_{k-1})||O(||x^* - x_{k-1}||^2) \\
= ||J^{-1}_f(x_{k-1})||O(||e_{k-1}||^2)
\]
Convergence of Newton Iteration (II)

- Quadratic convergence is achieved when the Jacobian of a fixed-point iteration is zero at the solution, which is true for Newton’s method:

\[
g(x) = x - J_f^{-1}(x)f(x)
\]

\[
J_g(x) = I - J_f^{-1}(x)J_f(x) - \sum_i f_i(x)H_f^{(i)}(x)
\]

\[
= I - I - O = O
\]

where \(H_f^{(i)}\) is the \(i\)th component of the derivative of \(J_f^{-1}(x)\) of \(f\).
To obtain $J_f(x_k)$ at iteration $k$, can use finite differences:

For $n = 1$, we have $J_f \approx (1/h)(f(x_k + h) - f(x_k))$.

More generally, the $i$th column of $j_i$ of the Jacobian is given by

$$j_i \approx (1/h)(f(x_k + he_i) - f(x_k)).$$

$n + 1$ function evaluations are needed: $f(x), f(x + he_i) \forall i \in \{1, \ldots, n\}$, which correspond to $m(n + 1)$ scalar function evaluations.
Cost of Multivariate Newton Iteration

- What is the cost of solving \( J_f(x_k)s_k = f(x_k) \)?
  \( O(n^3) \)

- What is the cost of Newton’s iteration overall?
  For \( k \) steps, \( O(n^3k + kn^2\gamma_{func-eval}) \).