

CS 450: Numerical Analysis¹

Nonlinear Equations

University of Illinois at Urbana-Champaign

¹*These slides have been drafted by Edgar Solomonik as lecture templates and supplementary material for the book “Scientific Computing: An Introductory Survey” by Michael T. Heath ([slides](#)).*

Solving Nonlinear Equations

- ▶ Solving (systems of) nonlinear equations corresponds to root finding:
 - ▶ $f(x^*) = 0$
 - ▶ $f(\mathbf{x}^*) = 0$
 - ▶ $\mathbf{f}(\mathbf{x}^*) = \mathbf{0}$
 - ▶ Algorithms for root-finding make it possible to solve systems of nonlinear equations and employ a similar methodology to finding minima in optimization.
- ▶ Main algorithmic approach: find successive roots of local linear approximations of f :

Conditions for Existence of Solution

▶ *Intermediate value theorem* for univariate problems:

▶ A function has a unique *fixed point* $g(x^*) = x^*$ in a given closed domain if it is *contractive* and contained in that domain,

$$\|g(x) - g(z)\| \leq \gamma \|x - z\|$$

Multiple Roots and Degeneracy

- ▶ If x^* is a root of f with *multiplicity* m , its $m - 1$ derivatives are also zero at x^* ,

$$f(x^*) = f'(x^*) = f''(x^*) = \dots = f^{(m-1)}(x^*) = 0.$$

- ▶ Increased multiplicity affects conditioning and convergence:

Bisection Algorithm

- ▶ Assume we know the desired root exists in a bracket $[a, b]$ and $\text{sign}(f(a)) \neq \text{sign}(f(b))$:

- ▶ Bisection subdivides the interval by a factor of two at each step by considering $f(c_k)$ at $c_k = (a_k + b_k)/2$:

Rates of Convergence

- ▶ Let \mathbf{x}_k be the k th iterate and $\mathbf{e}_k = \mathbf{x}_k - \mathbf{x}^*$ be the error, bisection obtains *linear convergence*, $\lim_{k \rightarrow \infty} \|\mathbf{e}_k\| / \|\mathbf{e}_{k-1}\| \leq C$:

- ▶ r th order convergence implies that $\|\mathbf{e}_k\| / \|\mathbf{e}_{k-1}\|^r \leq C$

Convergence of Fixed Point Iteration

- ▶ Fixed point iteration: $x_{k+1} = g(x_k)$ is locally linearly convergent if for $x^* = g(x^*)$, we have $|g'(x^*)| < 1$:

- ▶ It is quadratically convergent if $g'(x^*) = 0$:

Newton's Method

Demo: Newton's Method
Demo: Convergence of Newton's Method

- ▶ Newton's method is derived from a *Taylor series* expansion of f at x_k :

- ▶ Newton's method is *quadratically convergent* if started sufficiently close to x^* so long as $f'(x^*) \neq 0$:

Secant Method

Demo: Secant Method

Demo: Convergence of the Secant Method

▶ The *Secant method* approximates $f'(x_k) \approx \frac{f(x_k) - f(x_{k-1})}{x_k - x_{k-1}}$:

▶ The convergence of the Secant method is *superlinear* but not quadratic:

Nonlinear Tangential Interpolants

- ▶ Secant method uses a linear interpolant based on points $f(x_k), f(x_{k-1})$, could use more points and higher-order interpolant:
- ▶ Quadratic interpolation (Muller's method) achieves convergence rate $r \approx 1.84$:
- ▶ Inverse quadratic interpolation resolves the problem of nonexistence/nonuniqueness of roots of polynomial interpolants:

Achieving Global Convergence

- ▶ Hybrid bisection/Newton methods:

- ▶ Bounded (damped) step-size:

Systems of Nonlinear Equations

- ▶ Given $\mathbf{f}(\mathbf{x}) = [f_1(\mathbf{x}) \ \cdots \ f_m(\mathbf{x})]^T$ for $\mathbf{x} \in \mathbb{R}^n$, seek \mathbf{x}^* so that $\mathbf{f}(\mathbf{x}^*) = \mathbf{0}$

- ▶ At a particular point \mathbf{x} , the *Jacobian* of \mathbf{f} , describes how \mathbf{f} changes in a given direction of change in \mathbf{x} ,

$$\mathbf{J}_{\mathbf{f}}(\mathbf{x}) = \begin{bmatrix} \frac{df_1}{dx_1}(\mathbf{x}) & \cdots & \frac{df_1}{dx_n}(\mathbf{x}) \\ \vdots & & \vdots \\ \frac{df_m}{dx_1}(\mathbf{x}) & \cdots & \frac{df_m}{dx_n}(\mathbf{x}) \end{bmatrix}$$

Multivariate Newton Iteration

- ▶ Fixed-point iteration $\mathbf{x}_{k+1} = \mathbf{g}(\mathbf{x}_k)$ achieves local convergence so long as $|\lambda_{\max}(\mathbf{J}_g(\mathbf{x}^*))| < 1$ and quadratic convergence if $\mathbf{J}_g = \mathbf{O}$:

Multidimensional Newton's Method

- ▶ Newton's method corresponds to the fixed-point iteration

$$g(\mathbf{x}) = \mathbf{x} - \mathbf{J}_f^{-1}(\mathbf{x})\mathbf{f}(\mathbf{x})$$

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

Estimating the Jacobian using Finite Differences

- ▶ To obtain $\mathbf{J}_f(\mathbf{x}_k)$ at iteration k , can use finite differences:

- ▶ $n + 1$ function evaluations are needed: $f(\mathbf{x})$ and $f(\mathbf{x} + h\mathbf{e}_i), \forall i \in \{1, \dots, n\}$, which correspond to $m(n + 1)$ scalar function evaluations if $\mathbf{J}_f(\mathbf{x}_k) \in \mathbb{R}^{m \times n}$.

Quasi-Newton Methods

In solving a nonlinear equation, seek approximate Jacobian $\mathbf{J}_f(\mathbf{x}_k)$ for each \mathbf{x}_k

- ▶ Find $\mathbf{B}_{k+1} = \mathbf{B}_k + \delta\mathbf{B}_k \approx \mathbf{J}_f(\mathbf{x}_{k+1})$, so as to approximate *secant equation*

$$\mathbf{B}_{k+1} \underbrace{(\mathbf{x}_{k+1} - \mathbf{x}_k)}_{\delta\mathbf{x}} = \underbrace{\mathbf{f}(\mathbf{x}_{k+1}) - \mathbf{f}(\mathbf{x}_k)}_{\delta\mathbf{f}}$$

- ▶ *Broyden's method* solves the secant equation and minimizes $\|\delta\mathbf{B}_k\|_F$:

$$\delta\mathbf{B}_k = \frac{\delta\mathbf{f} - \mathbf{B}_k\delta\mathbf{x}}{\|\delta\mathbf{x}\|^2} \delta\mathbf{x}^T$$

Safeguarding Methods

- ▶ Can dampen step-size to improve reliability of Newton or Broyden iteration:

- ▶ *Trust region methods* provide general step-size control: