Nonlinear Equations

- Often (most of the time?) the relevant system of equations is not linear in the unknowns.
- Then, cannot decompose as $Ax = b$. Oh well.
- Instead write as:
  
  \begin{align*}
  (1) \quad f(x) &= 0 \quad \text{function of one variable (1-D)} \\
  (2) \quad f(x) &= 0 \quad x = (x_1, x_2, \ldots, x_n), \quad f = (f_1, f_2, \ldots, f_n) \quad (n-D) 
  \end{align*}

- Not guaranteed to have any real solutions, but generally do for astrophysical problems.
Solutions in 1-D

- Generally, solving multi-D equations is much harder, so we'll start with the 1-D case first...
- By writing $f(x) = 0$ we have reduced the problem to solving for the roots of $f$.
- In 1-D it is usually possible to trap or bracket the desired roots and hunt them down.
- Typically all root-finding methods proceed by iteration, improving a trial solution until some convergence criterion is satisfied.
Function Pathologies

- Before blindly applying a root-finding algorithm to a problem, it is essential that the form of the equation in question be understood: graph it!
- For smooth functions good algorithms will always converge, provided the initial guess is good enough.
- Pathologies include discontinuities, singularities, multiple or very close roots, or no roots at all!
Numerical Root Finding

• Suppose $f(a)$ and $f(b)$ have opposite sign.

• Then, if $f$ is continuous on the interval $(a,b)$, there must be at least one root between $a$ and $b$ (this is the Intermediate Value Theorem).

• Such roots are said to be bracketed.

Bracketed root  Many roots  No roots
Example Application

• Use root finding to calculate the equilibrium temperature of the ISM.

• The ISM is a very diffuse plasma.
  – Heated by nearby stars and cosmic rays.
  – Cooled by a variety of processes:
    • Bremsstrahlung: collisions between electrons and ions
    • Atom-electron collisions followed by radiative decay
    • Thermal radiation from dust grains
Example, Cont'd

- Equilibrium temperature given when:

  Rate of Heating \( H = \) Rate of Cooling \( C \)

  - Often \( H \) is not a function of temperature \( T \).
  - Usually \( C \) is a complex, nonlinear function of \( T \).

- To solve, find \( T \) such that \( H - C(T) = 0 \).
Bracketing and Bisection

• NRiC 9.1 lists some simple bracketing routines.
• Once bracketed, root is easy to find by bisection:
  – Evaluate $f$ at interval midpoint $(a + b) / 2$.
  – Root must be bracketed by midpoint and whichever $a$ or $b$ gives $f$ of opposite sign.
  – Bracketing interval decreases by 2 each iteration:
    $$\varepsilon_{n+1} = \varepsilon_n / 2.$$ 
  – Hence to achieve error tolerance of $\varepsilon$ starting from interval of size $\varepsilon_0$ requires $n = \log_2(\varepsilon_0/\varepsilon)$ steps.
Convergence

• Bisection converges linearly (first power of $\varepsilon$).

• Methods in which
  
  \[ \varepsilon_{n+1} = (\text{constant}) \times (\varepsilon_n)^m \quad m > 1 \]

  are said to converge superlinearly.

• Note error actually decreases exponentially for bisection. It converges "linearly" because successive figures are won linearly with computational effort.
Tolerance

- What is a practical tolerance $\varepsilon$ for convergence?
- Cannot be less than round-off error.
- For single-precision (float) accuracy, typically take $\varepsilon = 10^{-6}$ in fractional error.
  
  i.e. $\frac{f(x) - f(x_r)}{f(x_r)} \sim 10^{-6}$

  where $x =$ numerical solution, $x_r =$ actual root.
- When $f(x_r) = 0$ this fails, so use $\varepsilon = 10^{-6}$ as absolute error (or perhaps use $\varepsilon(|a| + |b|)/2$).
Newton-Raphson Method

• Can one do better than linear convergence? **Duh!**
• Expand $f(x)$ in a Taylor series:

$$f(x + \delta) = f(x) + f'(x) \delta + f''(x) \frac{\delta^2}{2} + ...$$

• For $\delta << x$, drop higher order terms, so:

$$f(x + \delta) = 0 \Rightarrow \delta = -\frac{f(x)}{f'(x)}$$

• $\delta$ is correction added to current guess of root:

i.e. $x_{i+1} = x_i + \delta$
• Graphically, Newton-Raphson (NR) uses tangent line at $x_i$ to find zero crossing, then uses $x$ at zero crossing as next guess:

• Note: only works near root ($\delta << x$)
Newton-Raphson, Cont'd

- When higher order terms important, NR fails spectacularly. Other pathologies exist too:

  - Shoots to infinity
  - Never converges
Newton-Raphson, Cont'd

- Why use NR if it fails so badly?
- Rate of convergence:
  \[ \varepsilon_{i+1} = \varepsilon_i - \frac{f(x_i)}{f'(x_i)} \]

- Taylor expand \( f(x_i) \) & \( f'(x_i) \) to get:
  \[ \varepsilon_{i+1} = - \varepsilon_i^2 \frac{f''(x_i)}{f'(x_i)} \quad \text{[quadratic!]} \]

- Note both \( f(x) \) and \( f'(x) \) must be evaluated each iteration, plus both must be continuous near root.
- Best use of NR is to "polish-up" bisection root.
Nonlinear Systems of Equations

- Consider the system \( f(x,y) = 0, \ g(x,y) = 0 \). Plot zero contours of \( f \) & \( g \):

- No information in \( f \) about \( g \), and vice versa.
  - In general, no good method for finding roots.
Nonlinear Systems, Cont'd

- If you are near root, best bet is NR.
  
e.g. For $\mathbf{F}(\mathbf{x}) = \mathbf{0}$, choose $x_{i+1} = x_i + \delta$, where
  
  $$\mathbf{F}'(\mathbf{x}) \, \delta = - \mathbf{F}(\mathbf{x})$$

- This is a matrix equation: $\mathbf{F}'(\mathbf{x})$ is a matrix with elements $\frac{\partial F_i}{\partial x_j}$. The matrix is called the Jacobian.

- Written out:

\[
\frac{\partial f}{\partial x} \delta_x + \frac{\partial f}{\partial y} \delta_y = - f(x, y) \\
\frac{\partial g}{\partial x} \delta_x + \frac{\partial g}{\partial y} \delta_y = - g(x, y)
\]
Nonlinear Systems, Cont'd

- Given initial guess, must evaluate matrix elements and RHS, solve system for $\delta$, and compute next iteration $x_{i+1}$. Then repeat (must solve $2 \times 2$ linear system each time).

- Essentially the non-linear system has been linearized to make it easier to work with.

- NriC 9.7 discusses a global convergence strategy that combines multi-D NR with "backtracking" to improve chances of finding solutions.
Example: Interstellar Chemistry

- ISM is multiphase plasma consisting of electrons, ions, atoms, and molecules.
- Originally, the ISM was thought to be too hostile for molecules.
- But in 1968-69, radio observations discovered absorption/emission lines of \( \text{NH}_3, \text{H}_2\text{CO}, \text{H}_2\text{O}, \ldots \)
- Lots of organic molecules, e.g. \( \text{CH}_3\text{CH}_2\text{OH} \) (ethanol).
Example, Cont'd

• In some places, all atoms have been incorporated into molecules.

• For example, molecular clouds: dense, cold clouds of gas composed primarily of molecules.
  
  \( (T \sim 30 \text{ K}, n \sim 10^6 \text{ cm}^{-3}, M \sim 10^{5-6} M_\odot, R \sim 10 - 100 \text{ pc}) \).

• How do you predict what abundances of different molecules should be, given \( n \) and \( T \)?

• Need to solve a chemical reaction network.
Example, Cont'd

• Consider reaction between two species A and B:

  \[ A + B \rightarrow AB \quad \text{reaction rate} = n_A n_B R_{AB} \]

• Reverse also possible:

  \[ AB \rightarrow A + B \quad \text{reaction rate} = n_{AB} R'_{AB} \]

• In equilibrium:

  (1) \[ n_A n_B R_{AB} = n_{AB} R'_{AB} \]

  (2) \[ n_A + n_{AB} = n_A^0 \quad \text{\textbackslash Normalizations: #} \]

  (3) \[ n_B + n_{AB} = n_B^0 \quad \text{\textbackslash A \& B conserved} \]
Example, Cont'd

- Substitute normalization equations into reaction equations to get quadratic in $n_{AB}$, easily solved.

- However, many more possible reactions:
  - $AC + B \leftrightarrow AB + C$ (exchange reaction)
  - $ABC \leftrightarrow AB + C$ (dissociation reaction)

- Wind up with large nonlinear system describing all forward/reverse reactions, involving known reaction rates $R$. Must solve given fixed $n^0$ & $T$. 
Numerical Derivatives

• For NR and function minimization, often need derivatives of functions. It's always better to use an analytical derivative if it's available.

• If you're stuck, could try:

\[ f'(x) \approx \frac{f(x+h) - f(x)}{h} \]

• However, this is very susceptible to RE. Better:

\[ f'(x) \approx \frac{f(x+h) - f(x-h)}{2h} \]

• Read NriC 5.7 before trying this!