Computational Physics and Astrophysics

Solving Nonlinear Equations

Kostas Kokkotas
University of Tübingen, Germany

and

Pablo Laguna
Georgia Institute of Technology, USA

Spring 2012

Solving nonlinear equations

The classical root-finding problem consists of, given a function $f(x)$ with $x \in (a, b)$, finding the value(s) r such that

$$f(r) = 0$$

The most common approach involves a recurrence relation:

$$x_n = g(x_{n-1})$$

such that

$$\lim_{n \to \infty} x_n = r$$

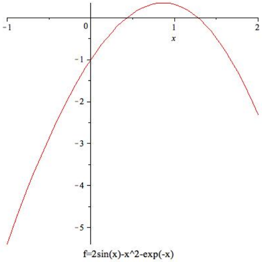
Independently of the method under consideration, one needs to answer the following key points:

- Best choice for the initial guess x_0.
- Bracketing the root.
- Under which conditions the method converges.
- Speed of convergence.

Bisection method

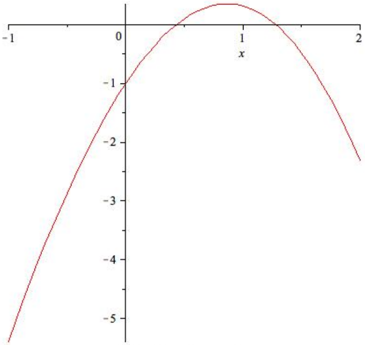
Consider the following function

$$f(x) = 2\sin(x) - x^2 - e^{-x}$$



- We are interested finding the root between $x_0 = 0$ and $x_1 = 1$.
- Notice that $f(0) = -1$ and $f(1) = 0.31506$
- Therefore, there must be at least one root since the function changes sign

We use then the following recurrence procedure



- $x_2 = (x_0 + x_1)/2 = 0.5$
  $\rightarrow f(0.5) = 0.1023$
- $x_3 = (x_0 + x_2)/2 = 0.25$
  $\rightarrow f(0.25) = -0.1732$
- $x_4 = (x_3 + x_2)/2 = 0.375$
  $\rightarrow f(0.375) = -0.0954$
- $x_5 = (x_4 + x_2)/2 = 0.4375$
  $\rightarrow f(0.4375) = 0.0103$
- $x_6 = (x_4 + x_5)/2 = 0.40625$
  $\rightarrow f(0.40625) = -0.0408$
- ...
- $x_n = r_1 \approx 0.4310378790$

The other root is $r_2 = 1.279762546$.

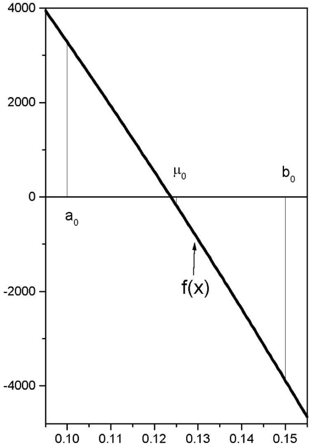
Bisection method algorithm

Consider the interval $[a_0, b_0]$. If $f(a_0) \cdot f(b_0) < 0$, then there is at least one root within this interval. Next define, $\mu_0 = (a_0 + b_0)/2$. Then, either:

1. $f(\mu_0) \cdot f(a_0) < 0$
2. $f(\mu_0) \cdot f(b_0) < 0$
3. $f(\mu_0) = 0$

If (3), the root has been found, else we set a new interval

$$[a_1, b_1] = \begin{cases} [\mu_0, b_0] & \text{if} \quad (2) \\ [a_0, \mu_0] & \text{if} \quad (1) \end{cases}$$



Bisection

```
REPEAT
    SET x3 = (x1 + x2)/2
    IF f(x3) · f(x1) < 0
        SET x2 = x3
    ELSE
        SET x1 = x3
    ENDIF
UNTIL |f(x3)| < E
```

Bisection method convergence

Define the error at a given iteration as

$$\varepsilon_n = |r - x_n|$$

For this method

$$\varepsilon_n \leq |a_n - b_n|$$

At every step the error is half of the previous step

$$\varepsilon_n = \frac{\varepsilon_{n-1}}{2} = \frac{\varepsilon_{n-2}}{2^2} = \cdots = \frac{\varepsilon_0}{2^n}$$

Thus, if we demand an error E, the number of steps n to reach this accuracy is

$$n = \log_2 \frac{\varepsilon_0}{E}$$

Linear interpolation

- Assume that the function $f(x)$ in the interval (x_1, x_2) is such that $f(x_1)$ and $f(x_2)$ have opposite signs.
- Introduce the straight line function $y(x)$ connecting the points $(x_1, f(x_1))$ and $(x_2, f(x_2))$:

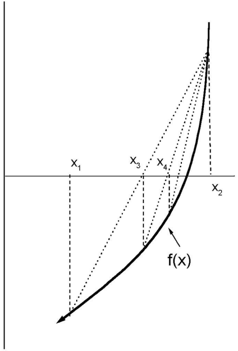
$$y(x) = f(x_1) + \frac{f(x_1) - f(x_2)}{x_1 - x_2}(x - x_1)$$

- Denote by x_3 the position at which $y(x)$ crosses the axis Ox. This point is given by

$$x_3 = \frac{x_2 f(x_1) - x_1 f(x_2)}{f(x_1) - f(x_2)} = x_2 - \frac{f(x_2)}{f(x_2) - f(x_1)}(x_2 - x_1)$$

- If in the (x_1, x_2) interval, $f(x) \approx y(x)$, then x_3 is a good approximation to the root r of $f(x)$.

Linear interpolation: Recurrence relation



With x_3, there are three choices:

1. If $f(x_1) \cdot f(x_3) < 0$ set $x_2 = x_3$
2. If $f(x_2) \cdot f(x_3) < 0$ set $x_1 = x_3$
3. If $f(x_3) = 0$ root has been found

Recurence relation:

$$x_{n+2} = x_{n+1} - \frac{f(x_{n+1})}{f(x_{n+1}) - f(x_n)}(x_{n+1} - x_n)$$

Linear interpolation

```
REPEAT
      SET x3 = x2 - f(x2) · (x2 - x1) / (f(x2) - f(x1))
      IF f(x3) · f(x1) < 0
            SET x2 = x3
      ELSE
            SET x1 = x3
      ENDIF
UNTIL |f(x3)| < E
```

Linear interpolation : Convergence

THEOREM: If ξ is the root of the equation $f(x) = 0$, and $\varepsilon_n = x_n - \xi$ is the error associated with the root estimate x_n, the convergence rate of the linear interpolation method is

$$\varepsilon_{n+1} = k \cdot \varepsilon_n^{1.618}$$

PROOF: From $x_n = \xi + \varepsilon_n$, we have that

$$\begin{aligned} f(x_n) &= f(\xi + \varepsilon_n) = f(\xi) + \varepsilon_n f'(\xi) + \frac{\varepsilon_n^2}{2} f''(\xi) \\ &= \varepsilon_n f'(\xi) + \frac{\varepsilon_n^2}{2} f''(\xi) \end{aligned}$$

then from the recurrence relation

$$x_{n+2} = x_{n+1} - \frac{f(x_{n+1})}{f(x_{n+1}) - f(x_n)} (x_{n+1} - x_n)$$

one has that

$$\varepsilon_{n+2} = \varepsilon_{n+1} - \frac{\varepsilon_{n+1} f'(\xi) + \frac{1}{2}\varepsilon_{n+1}^2 f''(\xi)}{f'(\xi)(\varepsilon_{n+1} - \varepsilon_n) + \frac{1}{2} f''(\xi)(\varepsilon_{n+1}^2 - \varepsilon_n^2)} \cdot (\varepsilon_{n+1} - \varepsilon_n)$$

$$\varepsilon_{n+2} = \varepsilon_{n+1} - \frac{\varepsilon_{n+1} f'(\xi) + \frac{1}{2}\varepsilon_{n+1}^2 f''(\xi)}{f'(\xi) + \frac{1}{2} f''(\xi)(\varepsilon_{n+1} + \varepsilon_n)}$$

$$\varepsilon_{n+2} = \varepsilon_{n+1}\left[1 - \frac{f'(\xi) + \frac{1}{2}\varepsilon_{n+1} f''(\xi)}{f'(\xi) + \frac{1}{2} f''(\xi)(\varepsilon_{n+1} + \varepsilon_n)}\right]$$

$$\varepsilon_{n+2} = \varepsilon_{n+1}\left[\frac{\frac{1}{2}\varepsilon_n f''(\xi)}{f'(\xi) + \frac{1}{2} f''(\xi)(\varepsilon_{n+1} + \varepsilon_n)}\right]$$

$$\varepsilon_{n+2} \approx \varepsilon_{n+1}\varepsilon_n \frac{1}{2}\frac{f''(\xi)}{f'(\xi)}$$

Let $\varepsilon_{n+1} = k\,\varepsilon_n^m$, then

$$\begin{aligned}\varepsilon_{n+2} &= k\,\varepsilon_{n+1}^m = k^{m+1}\,\varepsilon_n^{m^2}\\ \varepsilon_{n+1}\varepsilon_n &= k\,\varepsilon_n^{m+1}\end{aligned}$$

thus

$$\varepsilon_{n+2} \approx \varepsilon_{n+1}\varepsilon_n \frac{1}{2}\frac{f''(\xi)}{f'(\xi)}$$

implies

$$k^{m+1}\,\varepsilon_n^{m^2} = k\,\varepsilon_n^{m+1}\frac{f''(\xi)}{2f'(\xi)}$$

Therefore

$$\begin{aligned}k^m &= \frac{f''(\xi)}{2f'(\xi)}\\ m^2 &= m+1 \Rightarrow m = 1.618\end{aligned}$$

$$\varepsilon_{n+1} = k\cdot\varepsilon_n^{1.618}$$

Newton Raphson method

If in the neighborhood of the root r of the equation $f(x) = 0$ the 1st and 2nd derivatives of $f(x)$ are continuous, it is possible to develop a root finding method which is faster than the bisection and linear interpolation method.

Let x_{n+1} and x_n be respectively the $n+1$ and n root iterations of the equation $f(x) = 0$ such that $x_{n+1} = x_n + \delta_n$ with $\delta_n \ll 1$.

Then:

$$\begin{aligned} f\left(x_{n+1}\right) &= f\left(x_n + \delta_n\right) \\ &= f\left(x_n\right) + \delta_n f'\left(x_n\right) + \frac{\delta_n^2}{2} f''\left(x_n\right) + \cdots \end{aligned}$$

In the limit $n \to \infty$, we have that $f(x_{n+1}) = 0$, thus

$$0 = f(x_n) + \delta_n f'(x_n) \quad \Rightarrow \quad \delta_n = -\frac{f(x_n)}{f'(x_n)}$$

Newton Raphson method: Recurrence relation

Then

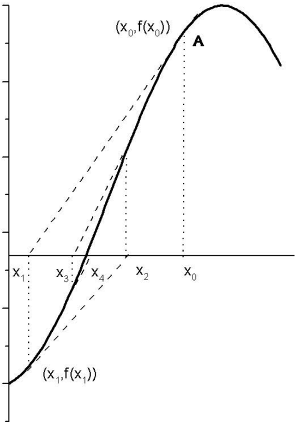
$$x_{n+1} = x_n - \frac{f(x_n)}{f'(x_n)}$$

Notice problems when

- $f'(x_n) = 0$
- $x_{n+2} = x_n$ non-convergent cylce

Recall linear interpolation recurrence relation

$$x_{n+1} = x_n - f(x_n)\left[\frac{(x_n - x_{n-1})}{f(x_n) - f(x_{n-1})}\right]$$



Newton Raphson method : Convergence

If $f(r) = 0$ such that $x_n = r + \varepsilon_n$, from the recurrence relation

$$
\begin{aligned}
r + \varepsilon_{n+1} &= r + \varepsilon_n - \frac{f(r+\varepsilon_n)}{f'(r+\varepsilon_n)} \\
\varepsilon_{n+1} &= \varepsilon_n - \frac{f(r) + \varepsilon_n f'(r) + \frac{1}{2}\varepsilon_n^2 f''(r)}{f'(r) + \varepsilon_n f''(r)} \\
\varepsilon_{n+1} &= \varepsilon_n - \frac{\varepsilon_n f'(r) + \frac{1}{2}\varepsilon_n^2 f''(r)}{f'(r) + \varepsilon_n f''(r)} \\
\frac{\varepsilon_{n+1}}{\varepsilon_n} &= 1 - \frac{f'(r) + \frac{1}{2}\varepsilon_n f''(r)}{f'(r) + \varepsilon_n f''(r)} \\
\frac{\varepsilon_{n+1}}{\varepsilon_n} &= \frac{\frac{1}{2}\varepsilon_n f''(r)}{f'(r) + \varepsilon_n f''(r)} \approx \varepsilon_n \frac{1}{2}\frac{f''(r)}{f'(r)}
\end{aligned}
$$

we get that

$$\varepsilon_{n+1} = -\frac{f''(r)}{2f'(r)} \cdot \varepsilon_n^2$$

That is, quadratic convergence

Newton-Raphson

```
IF f'(x) ≠ 0
    REPEAT
        SET x = x − f(x)/f'(x)
    UNTIL |f(x)| < E
ENDIF
```

Newton Halley method

Recall

$$f\left(x_{n+1}\right) = f\left(x_n + \delta_n\right) = f\left(x_n\right) + \delta_n f'\left(x_n\right) + \frac{\delta_n^2}{2} f''\left(x_n\right) + \ldots$$

$$f\left(x_n\right) + \delta_n \left[f'\left(x_n\right) + \frac{\delta_n}{2} f''\left(x_n\right) \right] = 0$$

$$\delta_n = -\frac{f(x_n)}{f'(x_n) + \frac{\delta_n}{2} f''(x_n)}$$

Substitute the previous Newton Raphson result

$$\delta_n \approx -\frac{f(x_n)}{f'(x_n)}$$

in the denominator

Then

$$\delta_n = -\frac{f(x_n)}{f'(x_n) - \frac{f(x_n)f''(x_n)}{2\,f'(x_n)}}$$

so

$$\begin{aligned} x_{n+1} &= x_n - \frac{f(x_n)}{f'(x_n) - \frac{f(x_n)f''(x_n)}{2\,f'(x_n)}} \\ x_{n+1} &= x_n - \frac{2\,f(x_n)f'(x_n)}{2\,f'(x_n)f'(x_n) - f(x_n)f''(x_n)} \end{aligned}$$

Notice that for $f''(x_n) = 0$ we recover that traditional Newton-Raphson method.

Newton Halley's methods achieves cubic convergence:

$$\varepsilon_{n+1} = -\left[\frac{1}{6}\frac{f'''(\xi)}{f'(\xi)} - \frac{1}{4}\left(\frac{f''(\xi)}{f'(\xi)}\right)^2\right]\cdot \varepsilon_n^3$$

Newton-Halley

```
IF f'(x) ≠ 0 AND f''(x) ≠ 0
    REPEAT
        SET x = x − 2 f(x) f'(x)/(2 f'(x) f'(x) − f(x) f''(x))
    UNTIL |f(x)| < E
ENDIF
```

EXAMPLE: Use the Newton-Raphson and Newton-Halley method to calculate square root of a number Q (here $Q = 9$).
Set $f(x) = x^2 - Q$, then

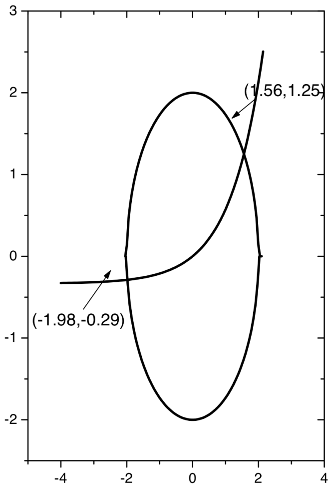
$$x_{n+1} = \frac{1}{2}\left(x_n + \frac{Q}{x_n}\right) \quad \text{(Raphson)} \quad x_{n+1} = \frac{x_n^3 + 3\,x_n\,Q}{3\,x_n^2 + Q} \quad \text{(Halley)}$$

Newton	Error	Halley	Error
x_0=15	$\varepsilon_0 = 12$	x_0=15	$\varepsilon_0 = 12$
x_1=7.8	$\varepsilon_1 = 4.8$	x_1=5.526	$\varepsilon_1 = 2.5$
x_2=4.477	$\varepsilon_2 = 1.477$	x_2=3.16024	$\varepsilon_2 = 0.16$
x_3=3.2436	$\varepsilon_3 = 0.243$	x_3=3.00011	$\varepsilon_3 = 1.05 \times 10^{-4}$
x_4=3.0092	$\varepsilon_4 = 9.15 \times 10^{-3}$	x_4=3.0000...	$\varepsilon_4 = 3.24 \times 10^{-14}$

An example of two non-linear equations is the following:

$$
\begin{aligned}
f(x,y) &= e^x - 3y - 1 \\
g(x,y) &= x^2 + y^2 - 4
\end{aligned}
$$

$f(x,y) = 0$ and $g(x,y) = 0$ are two curves on the xy plane.



Let's assume that after $n+1$ iterations the method converged to the solution (x_{n+1}, y_{n+1}) i.e. $f(x_{n+1}, y_{n+1}) \approx 0$ and $g(x_{n+1}, y_{n+1}) \approx 0$. Then if $x_{n+1} = x_n + \varepsilon_n$ and $y_{n+1} = y_n + \delta_n$, then

$$0 \approx f(x_{n+1}, y_{n+1}) = f(x_n+\varepsilon_n, y_n+\delta_n) \approx f(x_n, y_n)+\varepsilon_n \left(\frac{\partial f}{\partial x}\right)_n +\delta_n \left(\frac{\partial f}{\partial y}\right)_n$$

$$0 \approx g(x_{n+1}, y_{n+1}) = g(x_n+\varepsilon_n, y_n+\delta_n) \approx g(x_n, y_n)+\varepsilon_n \left(\frac{\partial g}{\partial x}\right)_n +\delta_n \left(\frac{\partial g}{\partial y}\right)_n$$

Solving for ε_n and δ_n

$$\varepsilon_n = \frac{-f\frac{\partial g}{\partial y} + g\frac{\partial f}{\partial y}}{\frac{\partial f}{\partial x}\frac{\partial g}{\partial y} - \frac{\partial g}{\partial x}\frac{\partial f}{\partial y}} \quad \text{and} \quad \delta_n = \frac{-g\frac{\partial f}{\partial x} + f\frac{\partial g}{\partial x}}{\frac{\partial f}{\partial x}\frac{\partial g}{\partial y} - \frac{\partial g}{\partial x}\frac{\partial f}{\partial y}}$$

Since $x_{n+1} = x_n + \varepsilon_n$ and $y_{n+1} = y_n + \delta_n$, the recurrence relations are:

$$
\begin{aligned}
x_{n+1} &= x_n - \left(\frac{f \cdot g_y - g \cdot f_y}{f_x \cdot g_y - g_x \cdot f_y} \right)_n \\
y_{n+1} &= y_n - \left(\frac{g \cdot f_x - f \cdot g_x}{f_x \cdot g_y - g_x \cdot f_y} \right)_n
\end{aligned}
$$

where $f_x = \frac{\partial f}{\partial x}$