Nonlinear equations and systems

An equation is **non-linear** if it involves non-linear components of the unknown, such as powers of the unknowns (e.g. $x^3, x^{1/4}, x^{-2}$, etc.) or other non-linear functions of the unknown (e.g. $\sin x, \log x, e^x$, etc.). In general, a non-linear equation takes the form

$$f(x) = 0 \qquad (1)$$

where $f : \mathbf{R} \to \mathbf{R}$ is a non-linear function of $x \in \mathbf{R}$.
A number x such that (1) is satisfied is called a **root** (or **zero** or **solution**) of f. A non-linear equation may have one or more roots, or no roots.

Example:
Let $f(x) = 4x^2 - 4x - 1 - e^x$. The equation $f(x) = 0$ has 3 roots. The roots are the intersection of functions $4x^2 - 4x - 1$ and e^x.
Let $f(x) = \ln x - e^x$. The equation $f(x) = 0$ has no roots.

Similarly, a non-linear system of equations is a set of equations of which at least one is non-linear, and which involve an unknown vector. In general, a non-linear equation takes the form

$$f(x) = 0 \qquad (2)$$

where $f : \mathbf{R}^n \to \mathbf{R}^n$ is a non-linear vector function of $x \in \mathbf{R}^n$, i.e.

$$f = \begin{bmatrix} f_1(x_1, \dots, x_n) \\ \cdots \\ f_n(x_1, \dots, x_n) \end{bmatrix}$$

and 0 is the zero vector in $\mathbf{R}^n$. A vector $x \in \mathbf{R}^n$ such that (2) is satisfied is called a **root** (or **zero** or **solution**) of the system. A non-linear system may have one or more roots, or no roots.

Example: The 2×2 non-linear system

$$5x_1^2 - x_2^2 = 0$$
$$4x_2 - \sin x_1 - \cos x_2 = 0$$

has 2 solutions.

Most non-linear equations and systems cannot be solved exactly by standard mathematical techniques, that is, we cannot arrive to an analytical formula for their roots. Instead, numerical techniques are used. They are often iterative methods and are referred to as **non-linear solvers**. That is, we start with some initial guess $x^{(0)}$ and compute successive approximations $x^{(k)}$ to the root x, for $k = 1, \dots$ until some **stopping criterion** is satisfied.

The most common initial guess is $x^{(0)} = 0$.
The most common stopping criteria are

$$\frac{||f(x^{(k)})||}{||f(x^{(0)})||} \le \varepsilon \text{ or } ||f(x^{(k)})|| \le \varepsilon,$$

where $f(x^{(k)})$ is called the **residual** and ε is the desired precision of the approximation, or

$$\frac{||x^{(k+1)} - x^{(k)}||}{||x^{(k+1)}||} \le \varepsilon \text{ or } ||x^{(k+1)} - x^{(k)}|| \le \varepsilon.$$

Iterative methods for non-linear equations do not always **converge** (i.e. a criterion may never be reached).

Fixed point iteration methods

A **fixed point** of a function g is an x such that $x = g(x)$. Let $f(x) = 0$ be the non-linear equation to be solved, and let x be a root. We construct g so that

$$f(x) = 0 \Leftrightarrow x = g(x).$$

Then a **fixed point iteration** method for approximating the root of f is:

guess $x^{(0)}$
for $k = 1$ to maxit
  $x^{(k)} = g(x^{(k-1)})$
  if stopping criterion is satisfied, $x = x^{(k)}$, exit
endfor

The above is a **one-step** iteration scheme.
Many convergence properties of the fixed point iteration scheme are based on the properties of the function g.

Example:
Let $f(x) = x + \ln x$. Then: $g(x) = -\ln x, g(x) = e^{-x}, g(x) = \frac{x+e^{-x}}{2}$. However, not all 3 fixed point iteration methods arising from the above gs have the same convergence properties.

Newton's method

Consider linearizing a non-linear function, by replacing $f(x)$ by its tangent line at the current point of approximation $x^{(k)}$. We have

$$y = f(x^{(k)}) + f'(x^{(k)})(x - x^{(k)}) \qquad (3)$$

We are looking for x that makes $y = 0$ in (3). This will be the new approximation $x^{(k+1)}$ to the root of f.

$$\begin{aligned} 0 &= f(x^{(k)}) + f'(x^{(k)})(x^{(k+1)} - x^{(k)}) \\ &\Rightarrow x^{(k+1)} = x^{(k)} - \frac{f(x^{(k)})}{f'(x^{(k)})} \end{aligned} \qquad (4)$$

Relation (4) gives **Newton's iteration**. If we define

$$g(x) = x - \frac{f(x)}{f'(x)} \qquad (5)$$

then Newton's iteration becomes

$$x^{(k+1)} = g(x^{(k)})$$

i.e. it takes the form of a fixed point iteration.
Note that f must be differentiable and
$f'(x^{(k)}) \neq 0$.

Newton's method

guess $x^{(0)}$
for $k = 1$ to maxit
  $x^{(k)} = x^{(k-1)} - \frac{f(x^{(k-1)})}{f'(x^{(k-1)})}$
  if $|f(x^{(k)})| \leq \varepsilon$, $x = x^{(k)}$, exit
endfor

Newton's method requires one function and one derivative evaluation per iteration.
Newton's method may converge when started at a certain initial guess, but may diverge if started at another initial guess. However, when it converges, it does so rapidly.

The secant method

The secant method can be derived by considering Newton's method and approximating $f'(x^{(k)})$ with

$$\frac{f(x^{(k)}) - f(x^{(k-1)})}{x^{(k)} - x^{(k-1)}}.$$

From (4), we get

$$x^{(k+1)} = x^{(k)} - f(x^{(k)}) \frac{x^{(k)} - x^{(k-1)}}{f(x^{(k)}) - f(x^{(k-1)})} \qquad (6)$$

Relation (6) gives the **secant iteration**. Note that $f(x^{(k)}) \neq f(x^{(k-1)})$.

The secant iteration is a **two-step** scheme. The secant method requires two initial guesses. we usually make a guess $x^{(0)}$ and pick a $x^{(-1)}$ close to it, such that $f(x^{(0)}) \neq f(x^{(-1)})$.

The secant method

guess $x^{(0)}$ and $x^{(-1)}$
for $k = 1$ to maxit
  $x^{(k)} = x^{(k-1)} - f(x^{(k-1)}) \frac{x^{(k-1)} - x^{(k-2)}}{f(x^{(k-1)}) - f(x^{(k-2)})}$
  if $|f(x^{(k)})| \leq \varepsilon$, $x = x^{(k)}$, exit
endfor

The secant method requires one function evaluation per iteration.
It does not always converge. When it converges, it does so slower that Newton's.

The bisection method

The bisection method is based on the fact that a continuous function f that changes sign in an interval $[L, R]$ must have at least one zero in that interval.

The bisection method is applicable if two points L and R can be found, such that $f(L)f(R) < 0$, and assuming f is continuous. It approximates one of the zeros of f in $[L, R]$ by halving the interval at each iteration.

The bisection method

let$L$ and R be such that $f(L)f(R) < 0$
for $k = 1$ to maxit
    $M = (L + R)/2$
    if $|R - L|/2 \leq \varepsilon, x = M$, exit
    if $f(L)f(M) < 0$ then $R = M$
    elseif $f(L)f(M) > 0$ then $L = M$
    else $x = M$, exit
endfor

The bisection method requires only one function evaluation per iteration.
At the kth iteration bisection reaches a root of f within precision $\frac{R-L}{2^k}$, where $[L, R]$ is the initial interval.
It always converges (whenever it is applicable).
It converges very slowly (slower than Newton's and secant). However, it is useful for computing initial guesses for either Newton's or the secant method.

If f has several roots in $[L, R]$, there is no guarantee as to which roots bisection converges to.

Existence and uniqueness of roots and convergence of fixed point iteration

Intermediate Value Theorem:
Assumption
- f is continuous in $[a, b]$

Result
- for all $\gamma \in (f(a), f(b))$, there exists $\alpha \in (a, b)$ such that $f(\alpha) = \gamma$.

Mean Value Theorem:
Assumption
- f is continuous in $[a, b]$ and differentiable in (a, b)

Result
- there exists $\xi \in (a, b)$ such that $f'(\xi) = \frac{f(b)-f(a)}{b-a}$.

Assume that $g(x) = x \Leftrightarrow f(x) = 0$. The following theorems give sufficient (but not necessary) conditions for the respective results.

Theorem: **Existence** of a root of $f(x)$
Assumptions
- $f(x)$ is continuous in $[a, b]$
- $f(a)f(b) < 0$

Result
- there exists at least one root of $f(x)$ in (a, b).

Theorem: **Existence** of a fixed point of $g(x)$
- $g(x)$ is continuous in I =$[a, b]$
- $g(x) \in I$ for all $x \in I$

Result
- there exists at least one fixed point of $g(x)$ in I.

Theorem: **Uniqueness** of a root of $f(x)$
Assumptions
- $f(x)$ is differentiable in $I = (a, b)$
- $f'(x) \neq 0$ $\forall x \in I$

Result
-the root is unique in I.

Theorem 1: **Convergence** of fixed-point iteration $x^{(k+1)} = g(x^{(k)})$.
Assumptions
- $g'(x)$ exists in (a, b)
- $g(x)$ maps $I = [a, b]$ to itself
- $|g'(x)| \le \lambda < 1 \ \forall x \in (a, b)$
- $x^{(0)} \in I$

Results
- there **exists a unique** fixed point α of g in I
- the iteration scheme $x^{(k+1)} = g(x^{(k)})$ cvgs to α
- $|x^{(k)} - \alpha| \le \lambda^k |x^{(0)} - \alpha| \le \lambda^k \max\{x^{(0)} - a, b - x^{(0)}\}$
- $|x^{(k)} - \alpha| \le \frac{\lambda^k}{1-\lambda} |x^{(0)} - x^{(1)}|$

Convergence speed of iterative methods

A sequence $x^{(0)}, \ldots, x^{(n)}, \ldots$ cvgs to α, if $\forall \varepsilon > 0, \exists n_0$, such that $\forall n \ge n_0$, we have $||x^{(n)} - \alpha|| \le \varepsilon$.

We are not only interested in whether a sequence cvgs but also how fast it cvgs. This is important when assessing the efficiency of several iterative methods.

A sequence $x^{(0)}, \ldots, x^{(n)}, \ldots$ cvgs to α with **order** $\beta \ge 1$, if there exists a constant $\kappa > 0$ such that

$$\lim_{k \to \infty} \frac{||x^{(k+1)} - \alpha||}{||x^{(k)} - \alpha||^\beta} = \kappa$$

The constant κ is called the **asymptotic error constant.**

If $\beta = 1, \kappa < 1$, linear convergence
If $\beta = 2, \kappa > 0$, quadratic convergence

Determining the order of convergence

Theorem:
assumptions
- $x^{(k+1)} = g(x^{(k)})$ cvgs to a fixed pt α
- $g \in \mathbf{C}^\beta$ near α
- $g'(\alpha) = \ldots = g^{(\beta-1)}(\alpha) = 0$ and $g^{(\beta)}(\alpha) \ne 0$

Results
- the order of convergence of $x^{(k+1)} = g(x^{(k)})$ is β
- the asymptotic error constant is $\frac{1}{\beta!} |g^{(\beta)}(\alpha)|$.

Newton's method for nonlinear systems

We extend Newton's method to a system of n nonlinear equations in n variables. Consider, as an example, the case $n = 2$.

$$f(x) \equiv \begin{pmatrix} f_1(x_1, x_2) \\ f_2(x_1, x_2) \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix} \equiv 0.$$

Linearize each component of f by replacing it by its tangent plane at the current point of approximation $x^{(k)} = (x_1^{(k)}, x_2^{(k)})$. The equations of the tangent planes of f_1, f_2 at $x^{(k)}$ are

$$\begin{aligned} y_1 &= f_1(x^{(k)}) + \frac{\partial f_1}{\partial x_1}(x_1 - x_1^{(k)}) + \frac{\partial f_1}{\partial x_2}(x_2 - x_2^{(k)}) \\ y_2 &= f_2(x^{(k)}) + \frac{\partial f_2}{\partial x_1}(x_1 - x_1^{(k)}) + \frac{\partial f_2}{\partial x_2}(x_2 - x_2^{(k)}). \end{aligned}$$

In matrix form, we have

$$\begin{pmatrix} y_1 \\ y_2 \end{pmatrix} = \begin{pmatrix} f_1(x^{(k)}) \\ f_2(x^{(k)}) \end{pmatrix} + \begin{bmatrix} \frac{\partial f_1}{\partial x_1} & \frac{\partial f_1}{\partial x_2} \\ \frac{\partial f_2}{\partial x_1} & \frac{\partial f_2}{\partial x_2} \end{bmatrix} \begin{pmatrix} x_1 - x_1^{(k)} \\ x_2 - x_2^{(k)} \end{pmatrix}$$

Hence, $y = f(x^{(k)}) + J(x^{(k)})(x - x^{(k)})$, where $J(x)$ is a 2×2 matrix that depends on x and is called the **Jacobian**.

To get the next approximation $x^{(k+1)}$, set $y = 0$. We have

$$x^{(k+1)} = x^{(k)} - [J(x^{(k)})]^{-1} f(x^{(k)}).$$

In the general case for any $n \geq 1$ the Jacobian is an $n \times n$ matrix with entries $J_{ij} = \frac{\partial f_i}{\partial x_j}$.

Newton's method for systems

guess $x^{(0)}$
for $k = 1$ to maxit
  evaluate $J(x)$ at $x^{(k-1)}$
  solve $J(x^{(k-1)}) s_{k-1} = -f(x^{(k-1)})$
  $x^{(k)} = x^{(k-1)} + s_{k-1}$
  if $\|f(x^{(k)})\| \leq \varepsilon$, $x = x^{(k)}$, exit
endfor

Newton's method requires:
- the evaluation of the n^2 entries of J at $x^{(k)}$
- the solution of an $n \times n$ linear system of equations
- addition of vectors
- norm calculations