1 Nonlinear equations

A system of nonlinear equations is a problem of finding a n variables $[x_1, \ldots, x_n]$ which satisfy n equations,

$$f_i(x_1, ..., x_n) = 0, i = 1 \ldots n \ , \qquad (1)$$

where the functions f_i are generally non-linear.

1.1 Modified Newton method

Suppose that the vector $\mathbf{x} = [x_1, \ldots, x_n]$ is close to the solution. Let us try to find the step $\Delta\mathbf{x}$ which would brings us to the solution, $f_i(\mathbf{x} + \Delta\mathbf{x}) = 0$. The first order Taylor expansion gives a system of linear equations

$$f_i(\mathbf{x}) + \sum_k \frac{\partial f_i}{\partial x_k} \Delta x_k = 0 \qquad (2)$$

or in the matrix form

$$J\Delta\mathbf{x} = -\mathbf{f}(\mathbf{x}), \qquad (3)$$

where

$$J_{ik} \equiv \frac{\partial f_i}{\partial x_k}, \ \mathbf{f}(\mathbf{x}) \equiv [f_1(\mathbf{x}), \ldots, f_n(\mathbf{x})] \ . \qquad (4)$$

The solution $\Delta\mathbf{x}$ to the linear system (3) gives the approximate direction and the step-size towards the solution. The *modified* Newton's method instead of the full step uses a more conservative approach:

---

repeat

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

  $\lambda = 1$

  repeat

    $\mathbf{z} = \mathbf{x} + \lambda\Delta\mathbf{x}$

    $\lambda = \lambda/2$

  until $||f(\mathbf{z})|| \leq (1 - \lambda)||f(\mathbf{x})||$ or $\lambda \leq 0.01$

  $\mathbf{x} = \mathbf{z}$

until converged

---

2 Optimization

Optimization is a problem of finding minimum (or maximum) of a given real (non-linear) function $f(\mathbf{p})$ of an n-dimensional argument $\mathbf{p} = [x_1, \ldots, x_n]$.

2.1 Downhill simplex method

The *downhill simplex method* (also called Nelder-Mead method) is a commonly used nonlinear optimization algorithm in an n-dimensional space. The minimum is found by transforming a simplex (a polytope of $n+1$ vertexes) according to the function values at the vertexes, moving it downhill until it converges towards the minimum (maximum).

To discuss the algorithm we need the following definitions:

Simplex: a figure (polytope) represented by $n+1$ points, called vertexes, $[\mathbf{p}_1, \ldots, \mathbf{p}_{n+1}]$, where each point $\mathbf{p}_k$ is an n-dimensional vector.

Highest point: the vertex, $\mathbf{p}_{\mathrm{hi}}$, with the largest value of the function: $f(\mathbf{p}_{\mathrm{hi}}) = \max_{(k)} f(\mathbf{p}_k)$.

Lowest point: the vertex, $\mathbf{p}_{\mathrm{lo}}$, with the smallest value of the function: $f(\mathbf{p}_{\mathrm{lo}}) = \min_{(k)} f(\mathbf{p}_k)$.

Centroid: the center of gravity of all points, except for the highest: $\mathbf{p}_{\mathrm{ce}} = \frac{1}{n} \sum_{(k \neq \mathrm{hi})} \mathbf{p}_k$

The simplex is moved downhill by a combination of the following elementary operations:

Reflection: The highest point is reflected against the centroid, $\mathbf{p}_{\mathrm{hi}} \to \mathbf{p}_{\mathrm{re}} = \mathbf{p}_{\mathrm{ce}} + (\mathbf{p}_{\mathrm{ce}} - \mathbf{p}_{\mathrm{hi}})$.

Expansion: The lowest point doubles its distance from the centroid, $\mathbf{p}_{\mathrm{lo}} \to \mathbf{p}_{\mathrm{ex}} = \mathbf{p}_{\mathrm{ce}} + 2(\mathbf{p}_{\mathrm{lo}} - \mathbf{p}_{\mathrm{ce}})$.

Contraction: The highest point halves its distance from the centroid, $\mathbf{p}_{\mathrm{hi}} \to \mathbf{p}_{\mathrm{co}} = \mathbf{p}_{\mathrm{ce}} + \frac{1}{2}(\mathbf{p}_{\mathrm{hi}} - \mathbf{p}_{\mathrm{ce}})$.

Reduction: All points, except for the lowest, move towards the lowest points halving the distance. $\mathbf{p}_{k \neq \mathrm{lo}} = \frac{1}{2}(\mathbf{p}_k + \mathbf{p}_{\mathrm{lo}})$.

Finally here is a possible algorithm for a downhill simplex method:

---

repeat

  try Reflection

  if $f(\mathbf{p}_{\mathrm{re}}) < f(\mathbf{p}_{\mathrm{hi}})$

    accept Reflection

  elseif $f(\mathbf{p}_{\mathrm{re}}) < f(\mathbf{p}_{\mathrm{lo}})$

    do Expansion

  else

    try Contraction

    if $f(\mathbf{p}_{\mathrm{co}}) < f(\mathbf{p}_{\mathrm{hi}})$

      accept Contraction

    else

      do Reduction

until converged
