Newtons Method for Solving System of Nonlinear Equations

Inge Söderkvist, 041119

These notes complement the lecture about this topic in the course MAM208 at LTU. At the lecture the theory is explained in more details.

Assume that we have a set of \( n \) nonlinear equations \( f_i(x) = 0, \ i = 1 \ldots n \), depending on the vector \( x = [x_1, \ldots, x_n]^T \). Using the vector notation \( f(x) = [f_1(x), \ldots, f_n(x)]^T \), we can write the equation as a system

\[
f(x) = 0,
\]

where \( f : \mathbb{R}^n \to \mathbb{R}^n \). The system can be solved by an iteration similar to the famous Newton-Raphson method for a scalar nonlinear equation. The method described below is a generalisation of Newton-Raphson method and it is usually called Newtions method for solving a system of nonlinear equations.

In a neighborhood of a point \( x^{(k)} \in \mathbb{R}^n \) we can approximate \( f \) by a first order Taylor expansion as

\[
f(x^{(k)} + p) = f(x^{(k)}) + J(x^{(k)}) p + O(\|p\|^2),
\]

(1)

where \( J(x^{(k)}) \in \mathbb{R}^{n \times n} \) is the Jacobian matrix to \( f(x) \) evaluated at \( x^{(k)} \). The Jacobian matrix is defined as

\[
J = \begin{bmatrix}
\frac{\partial f_1}{\partial x_1} & \cdots & \frac{\partial f_1}{\partial x_n} \\
\vdots & \ddots & \vdots \\
\frac{\partial f_n}{\partial x_1} & \cdots & \frac{\partial f_n}{\partial x_n}
\end{bmatrix}.
\]

By ignoring the term \( O(\|p\|^2) \) in (1) and setting the right hand side to zero, we end up with the following algorithm: (Cf. the derivation of Newton-Raphson for the case when \( n = 1 \)).

---

\(^1\)Minor changes. Ove Edlund, 051124
Algorithm Newton
Guess $x^{(0)}$ as an approximation of the solution.

$k = 0$
Repeat until convergence

Compute $J(x^{(k)})$ and $f(x^{(k)})$.
Solve the linear system $J(x^{(k)}) p = -f(x^{(k)})$
$x^{(k+1)} = x^{(k)} + p$

$k = k + 1$

The index $(k)$ is an iteration index and $x^{(k)}$ is the vector $x$ after $k$ iterations.
The algorithm above converges fast (quadratically) close to the solution.
As with Newton-Raphson's method it is necessary to have good starting values.

Example: Assume that we have the system

\[
\begin{align*}
    x_1^2 + x_2^2 - 1 & = 0, \\
    (x_1 - 2)^2 + (x_2 - 2)^2 - 4 & = 0.
\end{align*}
\]

(The system describes the intersection of two circles $\in \mathbb{R}^2$.)

Then we have $n = 2$,

\[
    f(x) = \begin{bmatrix}
    x_1^2 + x_2^2 - 1 \\
    (x_1 - 2)^2 + (x_2 - 2)^2 - 4
\end{bmatrix},
    J(x) = \begin{bmatrix}
    2x_1 & 2x_2 \\
    2(x_1 - 2) & 2(x_2 - 2)
\end{bmatrix}.
\]

If we set $x^{(0)} = [0, 1]^T$ we get, at the first iteration, the linear system

\[
    J p = -f \iff \begin{bmatrix}
    0 & 2 \\
    -4 & -2
\end{bmatrix} p = -\begin{bmatrix}
    0 \\
    1
\end{bmatrix},
\]

Solving this system we get $p = [0.25, 0]^T$ and $x^{(1)} = [0.25, 1]^T$. Continuing the iteration we get, after 3 iterations, $x^{(3)} = [0.294, 0.956]^T$ and $f(x^{(3)}) = 10^{-4}[0.136, 0.136]^T$. 