**Int. Journal of Math. Analysis, Vol. 1, 2007, no. 6, 279 - 284**

# Convergence of an Adaptive Newton Algorithm

**Sanjay Kumar Khattri**

Stord/Haugesund University College
Norway
sanjay.khattri@hsh.no

**Abstract**

The Newton - Krylov iteration is the most prominent iterative method for solving non-linear system of equations ($\mathbf{F}(\mathbf{x})$). Roughly speaking, the Newton - Krylov iteration consists of solving a series of linear systems (Jacobian systems) of the form $\boldsymbol{J}\,\mathbf{x} = \mathbf{b}$. Solving non-linear system of equations is very costly due to time involved in solving the large Jacobian systems. We adaptively define the tolerance of linear systems $\boldsymbol{J}\,\mathbf{x} = \mathbf{b}$ based on the accuracy of the global system ($\mathbf{F}(\mathbf{x})$). We prove the convergence of the method. Reported numerical work shows that the new approach is computationally very efficient.

**Mathematics Subject Classification:** 90C53, 65B99, 34A34

**Keywords:** Newton; Krylov; Nonlinear Iteration; Symmetric Jacobian

## 1 Introduction

This reseach is concerned with efficient solution of non-linear system of equations with symmetric Jacobian. Let us consider the nonlinear system $\mathbf{F}(\mathbf{x}) = 0$. Here, $\mathbf{F}$ is vector function. That is $\mathbf{F} = [F_1, F_2, \dots, F_n]^{\mathrm{T}}$, and $\mathbf{x}$ is the unknown vector. Let the vector be $\mathbf{x} = (x_1, x_2, \dots, x_n)^{\mathrm{T}}$. A Newton iteration for solving $\mathbf{F}(\mathbf{x}) = 0$ is given as

$$\boldsymbol{J}(\mathbf{x}_k)\,\Delta\mathbf{x}_k = -\,\mathbf{F}(\mathbf{x}_k), \tag{1}$$

$$\mathbf{x}_{k+1} = \mathbf{x}_k + \Delta\mathbf{x}_k \quad k = 0, 1, 2, \dots, m. \tag{2}$$

Here, the equation (1) is referred to as the Newton correction step, and $\boldsymbol{J}$ is the Jacobian ($\boldsymbol{J} = [\partial F_i/\partial x_j]$) [6, 8]. We assume that the Jacobian is symmetric in nature. For starting the above Newton iteration, we need to assume an initial value, $\mathbf{x}_0$, of the solution vector, $\mathbf{x}$. It is known that if the initial guess ($\mathbf{x}_0$) is close to the exact solution, and the Jacobian is invertible then the above Newton iteration will converge quadratically. That is $\|\mathbf{F}(\mathbf{x}_{k+1})\| \leq \mathrm{C}\,\|\mathbf{F}(\mathbf{x}_k)\|^2$. The most costly part of a Newton iteration is solving the Newton correction step equation (1). Roughly speaking, the Newton method consists of solving a series of Newton correction steps [7]. Solving equations (1) to a fixed tolerance can be computationally very expensive.

Let us define the tolerance of the Newton correction steps adaptively. Newton iteration where the tolerance of the correction step is defined adaptively is called Adaptive Newton method. Let the tolerance of the $k^{th}$ Newton correction step be $r_k$. Thus, at the $k^{th}$ step, we solve the equation

$$\boldsymbol{J}(\mathbf{x}_k)\,\Delta\mathbf{x}_k = -\,\mathbf{F}(\mathbf{x}_k) + r_k. \tag{3}$$

Let us further assume that after $k$ Newton iterations the tolerance $r_k$, and the norm of the vectors $\mathbf{F}(\mathbf{x}_k)$ (residual vector) and $\Delta\mathbf{x}_k$ (difference vector) are related as

$$\|\mathbf{r}_k\| \le \mathrm{C}_1\,\|\mathbf{F}(\mathbf{x}_k)\|^2 \quad \text{and} \quad \|\mathbf{r}_k\| \le \mathrm{C}_2\,\|\Delta\mathbf{x}_k{}^2\|. \tag{4}$$

Then, we prove the following quadratic convergence results for the Newton iteration

$$\|\mathbf{F}(\mathbf{x}_{k+1})\| \le \mathrm{C}\,\|\mathbf{F}(\mathbf{x}_k)\|^2 \quad \text{and} \quad \|\mathbf{x}_{k+1} - \mathbf{x}^\star\| \le \mathrm{C}\,\|\mathbf{x}_k - \mathbf{x}^\star\|^2.$$

Here, $\mathbf{x}^\star$ is the exact solution of the nonlinear system $\mathbf{F}$. Let us first show that if the Jacobian matrix is symmetric and Lipschitz continuous then its inverse is bounded. For a symmetric matrix $\boldsymbol{A}$, there exists a number $\Bbbk > 0$ such that the following two inequalities are equivalent

$$\|\boldsymbol{A}^{-1}\| \le \frac{1}{\Bbbk} \quad \text{and} \quad \|\boldsymbol{A}\,\mathbf{v}\| \ge \Bbbk\,\|\mathbf{v}\|, \tag{5}$$

see [9]. Here, $\boldsymbol{A}^{-1}$ is the inverse of the matrix $\boldsymbol{A}$. For a Lipschitz continuous matrix $\boldsymbol{B}$, there exists a number $L > 0$ such that

$$\|\boldsymbol{B}(\mathbf{y}) - \boldsymbol{B}(\mathbf{x})\| \le \; L\,\|\mathbf{y} - \mathbf{x}\|. \tag{6}$$

Now let us bound the inverse of the Jacobian matrix. For a vector $\mathbf{v}$, we can write

$$\|\boldsymbol{J}(\mathbf{x}_k)\,\mathbf{v}\| = \|\boldsymbol{J}(\mathbf{x}_{k+1})\mathbf{v} + (\boldsymbol{J}(\mathbf{x}_k) - \boldsymbol{J}(\mathbf{x}_{k+1}))\mathbf{v}\|. \tag{7}$$

Using the following inequality $\|a + b\| \ge \|a\| - \|b\|$. We get

$$\|\boldsymbol{J}(\mathbf{x}_k)\,\mathbf{v}\| \ge \|\boldsymbol{J}(\mathbf{x}_{k+1})\mathbf{v}\| - \|(\boldsymbol{J}(\mathbf{x}_k) - \boldsymbol{J}(\mathbf{x}_{k+1}))\mathbf{v}\|, \tag{8}$$

using the inequality (5), and also the matrix norm inequality $\|\boldsymbol{A}\mathbf{x}\| \le \|\boldsymbol{A}\|\,\|x\|$

$$\|\boldsymbol{J}(\mathbf{x}_k)\,\mathbf{v}\| \ge \Bbbk\,\|\mathbf{v}\| - \|\boldsymbol{J}(\mathbf{x}_k) - \boldsymbol{J}(\mathbf{x}_{k+1})\|\,\|\mathbf{v}\|. \tag{9}$$

Using the Lipschitz continuity of the Jacobian given by the equation (6), we get

$$\|\boldsymbol{J}(\mathbf{x}_k)\,\mathbf{v}\| \ge \Bbbk\,\|\mathbf{v}\| - L\,\|\mathbf{x}_k - \mathbf{x}_{k+1}\|\,\|\mathbf{v}\|, \tag{10}$$

$$\ge (\Bbbk - L\,\|\mathbf{x}_k - \mathbf{x}_{k+1}\|)\;\|\mathbf{v}\|. \tag{11}$$

Since the Jacobian is symmetric. Thus, using the inequality (5), the inverse of the Jacobian is bounded as

$$\|\boldsymbol{J}(\mathbf{x}_k)^{-1}\| \le \left(\frac{1}{\Bbbk - L\,\|\mathbf{x}_k - \mathbf{x}_{k+1}\|}\right) \tag{12}$$

## 2 Convergence of the Adaptive Newton Method

From the multi dimensional mean value lemma

$$\|\mathbf{F}(\mathbf{x}) - \mathbf{F}(\mathbf{y}) - \boldsymbol{J}(\mathbf{y})\,(\mathbf{x}-\mathbf{y})\| \le \frac{l}{2}\,\|\mathbf{x}-\mathbf{y}\|^2 \ . \tag{13}$$

By the equations (2) and (3)

$$\mathbf{x}_{k+1} = \mathbf{x}_k - \boldsymbol{J}(\mathbf{x}_k)^{-1}\ [\mathbf{F}(\mathbf{x}_k) + r_k] \ \ .$$

Combining the mean value lemma (13) and the above equation

$$\left\|\mathbf{F}(\mathbf{x}_{k+1}) - \mathbf{F}(\mathbf{x}_k) + \boldsymbol{J}(\mathbf{x}_k)\ \left[\boldsymbol{J}(\mathbf{x}_k)^{-1}\ (\mathbf{F}(\mathbf{x}_k) + \mathbf{r}_k)\right]\right\| \le \frac{l}{2}\ \left\|\boldsymbol{J}(x_k)^{-1}\,(\mathbf{F}(\mathbf{x}_k) + \mathbf{r})\right\|^2,$$

since $\boldsymbol{J}\,\boldsymbol{J}^{-1} = \mathbf{I}$

$$\|\mathbf{F}(\mathbf{x}_{k+1}) - \mathbf{F}(\mathbf{x}_k) + \mathbf{F}(\mathbf{x_k}) + \mathbf{r}\| \le \frac{l}{2}\ \left\|\boldsymbol{J}(x_k)^{-1}\ (\mathbf{F}(\mathbf{x}_k) + \mathbf{r}_k)\right\|^2,$$

using $\|x+y\| \le \|x\| - \|y\|$

$$\begin{aligned}
\|\mathbf{F}(\mathbf{x}_{k+1})\| &\le \frac{l}{2}\left\|\boldsymbol{J}(x_k)^{-1}\,(\mathbf{F}(\mathbf{x}_k)+\mathbf{r}_k)\right\|^2 + \|\mathbf{r}_k\|,\\
&\le \frac{l}{2}\left[\|\boldsymbol{J}(\mathbf{x}_k)^{-1}\|^2\,\|\mathbf{F}(\mathbf{x}_k)+\mathbf{r}_k\|^2\right] + \|\mathbf{r}_k\|,\\
&\le \frac{l}{2}\left[\|\boldsymbol{J}(\mathbf{x}_k)^{-1}\|^2\ (\|\mathbf{F}(\mathbf{x}_k)\| + \|\mathbf{r}_k\|)^2\right] + \|\mathbf{r}_k\|,\\
&\le \frac{l}{2}\left[\|\boldsymbol{J}(\mathbf{x}_k)^{-1}\|^2\ \left(\|\mathbf{F}(\mathbf{x}_k)\|^2 + \|\mathbf{r}_k\|^2 + 2\,\|\mathbf{F}(\mathbf{x}_k)\|\,\mathbf{r}_k\right)\right] + \|\mathbf{r}_k\|,\\
&\le \frac{l}{2}\left[\|\boldsymbol{J}(\mathbf{x}_k)^{-1}\|^2\ \left(\|\mathbf{F}(\mathbf{x}_k)\|^2 + \mathrm{C}_1{}^2\|\mathbf{F}(\mathbf{x}_k)\|^4 + 2\,\mathrm{C}_1\,\|\mathbf{F}(\mathbf{x}_k)\|^3\right)\right] + \mathrm{C}_1\|\mathbf{F}(\mathbf{x}_k)\|^2,\\
&\le \frac{l}{2}\|\mathbf{F}(\mathbf{x}_k)\|^2\ \left[\|\boldsymbol{J}(\mathbf{x}_k)^{-1}\|^2\ \left(1.0 + \mathrm{C}_1{}^2 + 2\,\mathrm{C}_1\,\|\mathbf{F}(\mathbf{x}_k)\|\right) + \mathrm{C}_1\right].
\end{aligned}$$

Thus,

$$\boxed{\|\mathbf{F}(\mathbf{x}_{k+1})\| \le \mathrm{C}\,\|\mathbf{F}(\mathbf{x}_k)\|^2}$$

This is our first main result. The fundamental theorem of calculus asserts that there is $t \in [0,1]$ such that

$$\mathbf{F}(\mathbf{z}) - \mathbf{F}(\mathbf{x}) = \int_0^1 \boldsymbol{J}[\mathbf{x} + t\,(\mathbf{z}-\mathbf{x})]\,(\mathbf{z}-\mathbf{x})\,dt. \tag{14}$$

By the equations (2) and (3)

$$\mathbf{x}_{k+1} - \mathbf{x}^\star = \mathbf{x}_k - \boldsymbol{J}(\mathbf{x}_k)^{-1}\ [\mathbf{F}(\mathbf{x}_k) + \mathbf{r}_k] - \mathbf{x}^\star,$$

where $\mathbf{x}^\star$ is the exact solution of the system $\mathbf{F}(\mathbf{x}) = 0$,

$$\mathbf{x}_{k+1} - \mathbf{x}^\star = (\mathbf{x}_k - \mathbf{x}^\star) + \boldsymbol{J}(\mathbf{x}_k)^{-1}\,[\mathbf{F}(\mathbf{x}^\star) - \mathbf{F}(\mathbf{x}_k)] - \boldsymbol{J}(\mathbf{x}_k)^{-1}\,\mathbf{r}_k.$$

Using equation (14)

$$\begin{aligned}
\mathbf{x}_{k+1} - \mathbf{x}^\star &= (\mathbf{x}_k - \mathbf{x}^\star) + \boldsymbol{J}(\mathbf{x}_k)^{-1}\left(\int_0^1 [\boldsymbol{J}(\mathbf{x}_k + t\,(\mathbf{x}_k - \mathbf{x}^\star))]\,(\mathbf{x}^\star - \mathbf{x}_k)\,dt\right) - \boldsymbol{J}(\mathbf{x}_k)^{-1}\,\mathbf{r}_k,\\
&= (\mathbf{x}_k - \mathbf{x}^\star) + \boldsymbol{J}(\mathbf{x}_k)^{-1}\left(\int_0^1 [\boldsymbol{J}(\mathbf{x}_k + t\,(\mathbf{x}_k - \mathbf{x}^\star))]\,(\mathbf{x}^\star - \mathbf{x}_k)\,dt\right) - \boldsymbol{J}(\mathbf{x}_k)^{-1}\,\mathbf{r}_k,\\
&= \boldsymbol{J}(\mathbf{x}_k)^{-1}\int_0^1 [\boldsymbol{J}(\mathbf{x} + t\,(\mathbf{x}^\star - \mathbf{x}_k)) - \boldsymbol{J}(\mathbf{x}_k)]\,(\mathbf{x}^\star - \mathbf{x}_k)\,dt - \boldsymbol{J}(\mathbf{x}_k)^{-1}\,\mathbf{r}_k.
\end{aligned}$$

Taking norm of both the sides of the above equation and using $\|x - y\| \le \|x\| + \|y\|$, $\|x\,y\| \le \|x\|\,\|y\|$, $\|\int x\| \le \int \|x\|$

$$\|\mathbf{x}_{k+1} - \mathbf{x}^\star\| \le \|\boldsymbol{J}(\mathbf{x}_k)^{-1}\| \int_0^1 \|\boldsymbol{J}(\mathbf{x}_k + t\,(\mathbf{x}^\star - \mathbf{x}_k)) - \boldsymbol{J}(\mathbf{x}_k)\|\,\|\mathbf{x}^\star - \mathbf{x}_k\|\,dt + \|\boldsymbol{J}(\mathbf{x}_k)^{-1}\|\,\|\mathbf{r}_k\|,$$

by the Lipschitz continuity of the Jacobian. That is $\|\boldsymbol{J}(\mathbf{x}_k + t\,(\mathbf{x}^\star - \mathbf{x}_k)) - \boldsymbol{J}(\mathbf{x}_k)\| \le L\,t\,\|\mathbf{x}^\star - \mathbf{x}_k\|$. We get

$$\begin{aligned}
\|\mathbf{x}_{k+1} - \mathbf{x}^\star\| &\le \|\mathbf{x}^\star - \mathbf{x}_k\|\ \|\boldsymbol{J}(\mathbf{x}_k)^{-1}\| \int_0^1 L\,t\,\|\mathbf{x}^\star - \mathbf{x}_k\|\,dt + \|\boldsymbol{J}(\mathbf{x}_k)^{-1}\|\,\|\mathbf{r}_k\|,\\
&\le \|\mathbf{x}^\star - \mathbf{x}_k\|^2\,\|\boldsymbol{J}(\mathbf{x}_k)^{-1}\|\,\frac{L}{2} + \|\boldsymbol{J}(\mathbf{x}_k)^{-1}\|\,\|\mathbf{r}_k\|, \qquad (15)
\end{aligned}$$

since $\|\mathbf{r}_k\| \le \mathrm{C}_2\,\|\mathbf{x}^\star - \mathbf{x}_k\|^2$

$$\|\mathbf{x}_{k+1} - \mathbf{x}^\star\| \le \|\mathbf{x}^\star - \mathbf{x}_k\|^2\,\|\boldsymbol{J}(\mathbf{x}_k)^{-1}\|\,\frac{L}{2} + \mathrm{C}_2\,\|\boldsymbol{J}(\mathbf{x}_k)^{-1}\|\,\|\mathbf{x}^\star - \mathbf{x}_k\|^2.$$

Thus,

$$\boxed{\|\mathbf{x}_{k+1} - \mathbf{x}^\star\| \le \|\mathbf{x}_k - \mathbf{x}^\star\|^2\left[\|\boldsymbol{J}(\mathbf{x}_k)^{-1}\|\tfrac{L}{2} + \|\boldsymbol{J}(\mathbf{x}_k)^{-1}\|\right]}$$

## 3 Numerical Work

We are solving the simplified Poisson Boltzmann equation (16) on $\Omega = [-1, 1] \times [-1, 1]$ with $k = 1.0$ [3, 4, 5]. Problems with discontinuity in $\epsilon$ are of practical applications [4]. The domain $\Omega$ is divided into four equal sub-domains as shown in the Figure 1 based on the medium properties $\epsilon$. It should be noted that elliptic problems with discontinuous coefficients can produce very ill conditioned linear systems.

$$-\operatorname{div}(\epsilon\,\operatorname{grad} p) + k\,\sinh(p) = f \quad \text{in} \quad \Omega \quad \text{and} \quad p(x, y) = x^3 + y^3 \quad \text{on} \quad \partial\Omega_{\mathrm{D}}. \qquad (16)$$

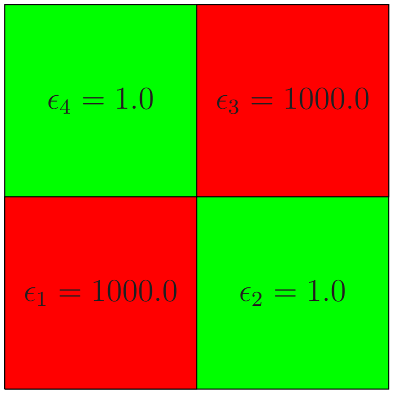


Figure 1: In the sub-domain $\Omega_i$, $\epsilon = \epsilon_i$, $i = 1, \ldots, 4$.

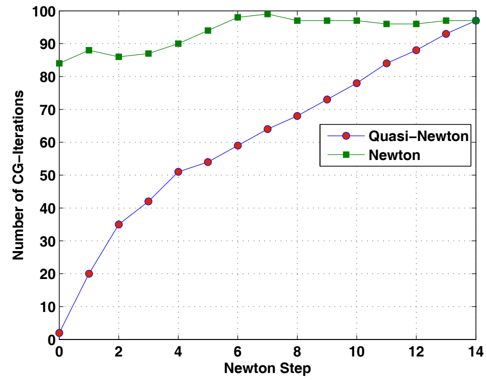


Figure 2: Computational efficiency of Quasi-Newton and Newton methods.

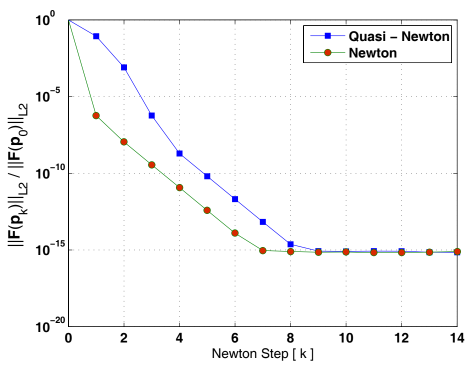


Figure 3: Convergence of residual vector $\mathbf{A}(\mathbf{p})$.

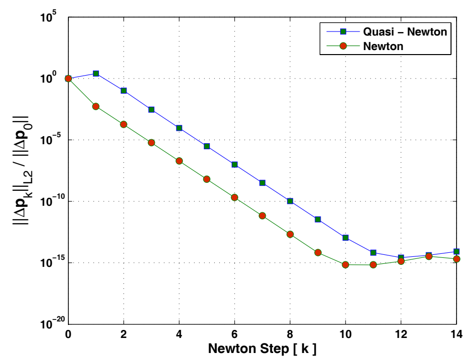


Figure 4: Convergence of the difference vector $\Delta\mathbf{p}$.

Here, the source function $f$ is

$$f = 2\,y\,(y-1) + 2\,x\,(x-1) - 100\,(x-1)\,y\,(y-1)\,\exp\left[x\,(x-1)\,y\,(y-1)\right].$$

For solving the linear systems, formed by the method of finite volumes [1, 2], we are using the ILU-preconditioned Conjugate-Gradient (CG) solver. For the Newton algorithm the tolerance of the CG method is $1.0 \times 10^{-15}$, while for the quasi-Newton method the tolerance of the CG varies with the Newton iteration $k$ as follows: $1.0 \times 10^{-(k+1)}$, $k = 0, 2, \ldots, 14$. The distribution of $\epsilon$ is given by the Figure 1. Thus, in the first and third quadrants of the domain $\epsilon = 100.0$, and in the second and fourth quadrants of the domain $\epsilon = 1.0$.

Figures 2, 3 and 4 report the outcome of our numerical work. The Figures 3 and 4 compare convergence of the Quasi-Newton and Newton methods. While the Figure 2 is comparing the computational efficiency of the Quasi-Newton and the Newton methods. In the Figures 3 and 4, it is interesting to note that the convergence rate of both the methods is same. The Figure 2 presents the computational work required by the Quasi-Newton and Newton methods. We observe here that our approach require approximately half the work needed by the Newton method. Thus, even if initial iterations of the Newton-Krylov algorithm are solved approximately, the convergence rate of the algorithm remains unaffected, and such an approximation saves a substantial amount of computational effort.

## References

[1] S. K. Khattri, Nonlinear elliptic problems with the method of finite volumes, Differential Equations and Nonlinear Mechanics, (2006).

[2] S. K. Khattri, Analyzing Finite Volume for Single Phase Flow in Porous Media, Journal of Porous Media, (2006).

[3] B. Aksoyw, Adaptive Multilevel Numerical Methods with Applications in Diffusive Biomolecular Reactions," PhD Thesis, The University of California, San Diego, 2001.

[4] F. Fogolari, A. Brigo and H. Molinari, The Poisson Boltzmann equation for Biomolecular electrostatics: A Tool for Structural Biology, Journal of Molecular Recognition, John Wiley & Sons Ltd., **15** (2002), 377 - 392.

[5] M. Host, R. E. Kozack, F. Saied and S. Subramaniam, Treatment of Electrostatic Effects in Proteins: Multigrid-based Newton Iterative Method for Solution of the Full Nonlinear Poisson-Boltzmann Equation, Proteins: Structure, Function, and Genetics, **18** (1994), 231–245.

[6] C. G. Broyden, J. E. Dennis and J. J. Morè, On the local and superlinear convergence of quasi-Newton methods, J. Inst. Math. Appl., 12 (1973), 223–245.

[7] L. Donghui and M. Fukushima, A globally and superlinearly convergent Gauss-Newton-based BFGS method for symmetric nonlinear equations, SIAM J. Numer. Anal., **37** (1999), 152–172.

[8] J. E. Dennis and J. J. Morè, Quasi-Newton methods, motivation and theory, SIAM Rev., **19** (1997), 46–89.

[9] G. Strang, Linear Algebra and Its Applications (4th Ed.), Thomson Learning, 2005.

**Received: August 31, 2006**