ON THE CONVERGENCE OF THE NEWTON-GMBACK METHOD

Emil Cătinaş
ecatinas@ictp.acad.ro
Romanian Academy
"T. Popoviciu" Institute of Numerical Analysis, P.O. Box 68–1
Cluj-Napoca, Romania

Abstract

When the GMBACK solver is used in the Newton iterates, the iterates may be written either as inexact Newton iterates or as quasi-Newton iterates. In this paper present results which assure the local convergence of the iterates in the both settings.

1 NONLINEAR SYSTEMS

The Newton method for solving a nonlinear system

$$F(x) = 0, \quad F : D \subseteq \mathbb{R}^N \to \mathbb{R}^N$$

leads to the solving of a linear system at each iteration step:

$$\begin{aligned} &F'(x_k)\, s_k = -F(x_k) \\ &x_{k+1} = x_k + s_k, \qquad k = 0, 1, \ldots, \qquad x_0 \in D. \end{aligned}$$

Under the following conditions (which will be implicitly assumed throughout this paper) the Newton method converge locally at q-superlinear rate (see [9]):
- $\exists x^* \in \operatorname{int} D$ such that $F(x^*) = 0$
- the mapping F is Fréchet differentiable on $\operatorname{int} D$, with F' continuous at x^*;
- the Jacobian $F'(x^*)$ is invertible.

When the linear systems from each step are not solved exactly, we are lead to the inexact Newton method

$$\begin{aligned} &F'(x_k)\, s_k = -F(x_k) + r_k \\ &x_{k+1} = x_k + s_k, \quad k = 0, 1, \ldots, \quad x_0 \in D. \end{aligned}$$

The terms $r_k \in \mathbb{R}^N$ represent the residuals of the approximate solutions s_k. Dembo, Eisenstat and Steihaug characterized the superlinear convergence of the inexact Newton (IN) method above.

Theorem 1. *[4] Assume that the* IN *iterates converge to x^*. Then the convergence is superlinear if and only if*

$$\|r_k\| = o(\|F(x_k)\|), \quad \text{as } k \to \infty. \tag{1}$$

When one considers approximate Jacobians at each step, $F'(x_k) + \Delta_k \in \mathbb{R}^{N \times N}$, we are lead to the quasi-Newton (QN) iterates

$$\begin{aligned} &(F'(x_k) + \Delta_k) s_k = -F(x_k) \\ &x_{k+1} = x_k + s_k, \qquad k = 0, 1, \ldots, \qquad x_0 \in D. \end{aligned}$$

We have characterized the superlinear convergence of these iterates in the following result:

Theorem 2. *[3] Assume that the* QN *iterates converge to* x^*. *Then the convergence is superlinear if and only if*

$$\|\Delta_k s_k\| = o\big(\|F(x_k)\|\big), \quad \text{as } k \to \infty. \tag{2}$$

Remark 3. *1. Dennis and Moré [5] have previously obtained the following condition which characterizes the superlinear convergence of the QN iterates:*

$$\|\Delta_k s_k\| = o\big(\|s_k\|\big), \quad \text{as } k \to \infty. \tag{3}$$

2. In [3] we have obtained in fact that the condition

$$\|-\Delta_k s_k + \delta_k + \hat{r}_k\| = o\big(\,\|F\left(_k\right)\|\,\big), \quad \text{as } k \to \infty$$

characterizes the superlinear convergence of the more general iterations

$$\begin{aligned} &\big(F'\left(x_k\right) + \Delta_k\big)\, s_k = \left(-F\left(x_k\right) + \delta_k\right) + \hat{r}_k \\ &x_{k+1} = x_k + s_k, \qquad k = 0, 1, \ldots, \quad x_0 \in D. \end{aligned}$$

2 THE GMBACK METHOD

When the dimension N is large, the numerical solving of a linear system

$$Au = b, \quad A \in \mathbb{R}^{N\times N} \text{nonsingular}, b \in \mathbb{R}^N,$$

becomes a difficult task. The Krylov solvers are popular choices for accomplishing this task, since they may offer good approximations at low computational cost.

We shall consider here the GMBACK solver introduced by Kasenally in [6]. For a given subspace dimension $m \in \{1, \ldots, N\}$ and an initial approximation $u_0 \in \mathbb{R}^N$ having the residual $r_0 = b - Au_0$, it finds $u_m^{GB} \in u_0 + \mathcal{K}_m = u_0 + \operatorname{span}\{r_0, Ar_0, \ldots, A^{m-1}r_0\}$ by the following minimization property:

$$\left\|\Delta_m^{GB}\right\|_F = \min_{u_m \in u_0 + \mathcal{K}_m} \|\Delta_m\|_F \quad \text{w.r.t. } (A - \Delta_m)u_m = b.$$

Here $\|\cdot\|_F$ denotes the Frobenius norm of a matrix, $\|Z\|_F = tr(ZZ^t)^{1/2}$ while $\|\cdot\|_2$ will denote the Euclidean norm from $\mathbb{R}^N$ and its induced operator norm.

The following steps are performed for determining u_m^{GB}:

Arnoldi

- Determine $V_m = [v_1 \ \ldots \ v_m] \in \mathbb{R}^{N\times m}$ and the upper Hessenberg matrix $\bar{H}_m \in \mathbb{R}^{(m+1)\times m}$;

GMBACK

- Let $\beta = \|r_0\|_2$,
  $\hat{H}_m = \left[-\beta e_1 \quad \bar{H}_m\right] \in \mathbb{R}^{(m+1)\times(m+1)}$, $\hat{G}_m = \left[u_0 \quad V_m\right] \in \mathbb{R}^{N\times(m+1)}$,
  $P = \hat{H}_m^t \hat{H}_m \in \mathbb{R}^{(m+1)\times(m+1)}$ and $Q = \hat{G}_m^t \hat{G}_m \in \mathbb{R}^{(m+1)\times(m+1)}$;
- Determine an eigenvector v_{m+1} corresponding to the smallest eigenvalue λ_{m+1}^{GB} of the generalized eigenproblem $Pv = \lambda Qv$;

- If the first component $v_{m+1}^{(1)}$ is nonzero, compute the vector $y_m^{GB} \in \mathbb{R}^m$ by scaling v_{m+1} such that
$$\begin{bmatrix} 1 \\ y_m^{GB} \end{bmatrix} = \frac{1}{v_{m+1}^{(1)}} v_{m+1};$$
- Set $u_m^{GB} = x_0 + V_m y_m^{GB}$.

This algorithm may lead to two possible breakdowns, either in the Arnoldi method or in the scaling of v_{m+1}. The first is a happy breakdown, because the solution may be determined exactly using $\bar{H}_m$ and V_m. The second appears when all the eigenvectors associated to λ_{m+1}^{GB} have the first component zero, the inevitable divisions by zero leading to uncircumventible breakdowns. In such a case either m is increased or the algorithm is restarted with a different initial approximation u_0. We shall assume in the following analysis that u_m^{GB} exists.

Kasenally proved that for any $u_0 \in \mathbb{R}^N$ and $m \in \{1, \dots, N\}$, the backward error Δ_m^{GB} corresponding to the GMBACK solution satisfies
$$\|\Delta_m^{GB}\|_F = \sqrt{\lambda_{m+1}^{GB}}. \tag{4}$$

Regarding the induced operator Euclidean norm, the following inequality is known:
$$\|Z\|_2 \le \|Z\|_F, \quad \text{for all } Z \in \mathbb{R}^{N\times N}. \tag{5}$$

3 THE CONVERGENCE OF THE NEWTON-GMBACK METHOD

The Newton-GMBACK iterates may be written in two equivalent ways, taking into account the properties of the Krylov solutions. On one hand, we may use the inexact Newton model, write as
$$F'(y_k)s_k^{GB} = -F(y_k) + r_k^{GB}, \tag{6}$$
and control the convergence of the iterates with the aid of residuals, by Theorem 1.

We obtain:

Theorem 4. *Assume that the Newton-GMBACK iterates are well defined and converge to x^*. Then the convergence is superlinear if and only if*
$$\|r_k^{GB}\| = o(\|F(x_k)\|), \quad \textit{as } k \to \infty. \tag{7}$$

On the other hand, we may use the quasi-Newton model, write
$$\left(F'(y_k) - \Delta_k^{GB}\right) s_k^{GB} = -F(y_k), \tag{8}$$
and control the convergence of the iterates by Theorem 2 with the aid of the backward error. It is worth noting that the GMBACK algorithm does not provide the backward error itself, but we may use formulas (4) and (5) to evaluate its magnitude in the Euclidean norm.

We obtain:

Theorem 5. *Assume that the Newton-GMBACK iterates are well defined and converge to x^*. If*
$$\lambda_k^{GB} \to 0, \quad \textit{as } k \to \infty, \tag{9}$$
then they converge superlinearly.

It is interesting to compare which of these two ways of controlling the convergence is most efficient numerically (since the evaluation of the residual may not be a trivial computational task when the dimension N is large). This is the purpose of a future investigation.

4 ACKNOWLEDGEMENTS

This research has been supported by MEdC under grant 2CEx06-11-96.

References

[1] E. Cătinaş, *Inexact perturbed Newton methods and applications to a class of Krylov solvers*, J. Optim. Theory Appl. **108** (2001), 543–570.

[2] E. Cătinaş, *On the superlinear convergence of the successive approximations method*, J. Optim. Theory Appl. **113** (2002), 473–485.

[3] E. Cătinaş, *The inexact, inexact perturbed and quasi-Newton methods are equivalent models*, Math. Comp., 74 (2005) no. 249, pp. 291–301 .

[4] R. S. Dembo, S. C. Eisenstat and T. Steihaug, *Inexact Newton methods*, SIAM J. Numer. Anal. **19** (1982), 400–408.

[5] J. E. Dennis, Jr. and J. J. Moré, *A characterization of superlinear convergence and its application to quasi-Newton methods*, Math. Comp. **28** (1974), 549–560.

[6] E. M. Kasenally, *GMBACK: a generalised minimum backward error algorithm for nonsymmetric linear systems*, SIAM J. Sci. Comput. **16** (1995), 698–719.

[7] E. M. Kasenally and V. Simoncini, *Analysis of a minimum perturbation algorithm for nonsymmetric linear systems*, SIAM J. Numer. Anal. **34** (1997), 48–66.

[8] B. Morini, *Convergence behaviour of inexact Newton methods*, Math. Comp. **68** (1999), 1605–1613.

[9] J. M. Ortega and W. C. Rheinboldt, *Iterative Solution of Nonlinear Equations in Several Variables*, Academic Press, New York, 1970.

[10] F. A. Potra, *On Q-order and R-order of convergence*, J. Optim. Theory Appl. **63** (1989), 415–431.