REVUE D'ANALYSE NUMÉRIQUE ET DE THÉORIE DE L'APPROXIMATION
Rev. Anal. Numér. Théor. Approx., vol. 28 (1999) no. 2, pp. 125–132
ictp.acad.ro/jnaat

ON THE HIGH CONVERGENCE ORDERS OF THE NEWTON-GMBACK METHODS*

EMIL CĂTINAŞ

Abstract. The high convergence orders of the Newton-GMBACK methods can be characterized applying three different existing results. In this note we show by some direct computations that these characterizations are equivalent.

AMS Subject Classification: 65H10, 65F10.

1. INTRODUCTION

The GMBACK method introduced by Kasenally [11], and which we shall describe in the following section, is a Krylov method for solving large linear systems. When it is used in a Newton method, the convergence orders of the resulted iterations may be characterized applying three existing results.

The first result was given by Dennis and Moré in [7], but before enouncing it we review the common setting. Given $F : D \subseteq \mathbb{R}^N \to \mathbb{R}^N$, the local convergence of different Newton-type methods is usually studied under the following standard assumptions:

(C1) there exists $y^* \in D$ such that $F(y^*) = 0$;
(C2) the mapping F is differentiable in a neighborhood of y^*, with the derivative F' continuous at y^*;
(C3) the Jacobian $F'(y^*)$ is nonsingular: $\exists F'(y^*)^{-1} \in \mathbb{R}^{N\times N}$.

We shall denote hereafter by $\|\cdot\|$ an arbitrary fixed norm on $\mathbb{R}^N$ or its induced operator norm. The symbol $\|\cdot\|_2$ stands for the euclidean norm and $\|\cdot\|_F$ denotes the Frobenius norm. For definitions and results concerning the convergence orders we refer to [15, ch.9] (see also [19], [18]).

THEOREM 1.1. [7]. *Let $F : D \to \mathbb{R}^N$ be differentiable in the open convex set $D \subseteq \mathbb{R}^N$ and assume that for some $y^* \in D$ the mapping F satisfies* (C2) *and* (C3). *Let $(B_k)_{k\geq 0} \subset \mathbb{R}^{N\times N}$ be a sequence of nonsingular matrices and*

*Work supported by the Romanian Academy under grant GAR 97/1999.

suppose that for some $y_0 \in D$ the quasi-Newton iterates given by

$$y_{k+1} = y_k - B_k^{-1} F(y_k), \quad k = 0, 1, \dots$$

remain in D and converge to y^*. Then $(y_k)_{k\geq 0}$ converges q-superlinearly to y^* and $F(y^*) = 0$ if and only if

$$\lim_{k\to\infty} \frac{\|(B_k - F'(y^*))(y_{k+1} - y_k)\|}{\|y_{k+1} - y_k\|} = 0.$$

The second result was obtained by Dembo, Eisenstat and Steihaug in [6].

Theorem 1.2. [6]. *Assume the mapping F satisfies the standard assumptions and suppose that for some $y_0 \in D$ the sequence of the inexact Newton iterates given by*

$$\begin{aligned} F'(y_k)\, s_k &= -F(y_k) + r_k \\ y_{k+1} &= y_k + s_k, \qquad k = 0, 1, \dots \end{aligned}$$

remains in D and converges to y^*. Then $(y_k)_{k\geq 0}$ converges q-superlinearly if and only if the residuals r_k satisfy

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

Martinez, Parada and Tapia [14] obtained some results for the sequences of damped and perturbed quasi-Newton methods

$$y_{k+1} = y_k - \alpha_k B_k^{-1} (F(y_k) + r_k),$$

where $0 < \alpha_k \leq 1$, $r_k \in \mathbb{R}^N$, $k = 0, 1, \dots$ and $y_0 \in \mathbb{R}^N$. Though, those results do not fully characterize the convergence orders of the above sequences. On the other hand, we shall see that if we want to fit in a direct manner the Newton-GMBACK iterates in this frame, we must reduce the above iterations either to the classical quasi-Newton or to the IN ones.

We have recently extended Theorem 1.2.

Theorem 1.3. [5]. *Assume the mapping F satisfies the standard assumptions and consider the following elements: $(\Delta_k)_{k\geq 0} \subset \mathbb{R}^{N\times N}$ (perturbations in the Jacobians), $(\delta_k)_{k\geq 0} \subset \mathbb{R}^N$ (perturbations in the function evaluations) and $(\hat{r}_k)_{k\geq 0} \subset \mathbb{R}^N$ (residuals of the approximate solutions s_k to the perturbed linear systems $(F'(y_k) + \Delta_k)\, s = -F(y_k) + \delta_k$). If for some $y_0 \in D$ the sequence of the inexact perturbed Newton iterates given by*

$$\begin{aligned} (F'(y_k) + \Delta_k)\, s_k &= (-F(y_k) + \delta_k) + \hat{r}_k \\ y_{k+1} &= y_k + s_k, \qquad k = 0, 1, \dots \end{aligned}$$

is well defined (i.e. the matrices $F'(y_k) + \Delta_k$ are nonsingular and the iterates y_k remain in D) and converges to y^*, then the convergence is q-superlinear if and only if

$$\left\| \Delta_k (F'(y_k) + \Delta_k)^{-1} F(y_k) + \left(I - \Delta_k \left(F'(y_k) + \Delta_k\right)^{-1}\right)(\delta_k + \hat{r}_k) \right\| = o\left(\|F(y_k)\|\right),$$

as $k \to \infty$.

Remark. Modest additional continuity conditions imposed on the derivative F' allow the characterizations of the q-convergence orders $1+p$, $p \in (0,1]$ of the three methods considered above. For these extensions we refer the reader to [8], [6] and resp. [5]. □

In Section 2 we briefly describe the GMBACK method and we deduce some results, while in Section 3 we show that when applying the above three theorems for the characterization of the q-superlinear convergence of the Newton-GMBACK method we are led to some equivalent results.

2. THE GMBACK METHOD

Consider the linear nonsingular system

$$Ax = b,$$

with $A \in \mathbb{R}^{N \times N}$ and $b \in \mathbb{R}^N$. The GMBACK algorithm introduced by Kasenally in [11] belongs to the class of Krylov methods for solving such systems when the dimension N is large. Given the initial approximation $x_0 \in \mathbb{R}^N$ of the true solution x^* and a number $m \in \{1, \dots, N\}$, by a modified Gram-Schmidt procedure there is constructed an orthonormal basis $\{v_1, \dots, v_m\}$ in the Krylov subspace $\mathcal{K}_m = \mathcal{K}_m(A, r_0) = span\{r_0, Ar_0, \dots, A^{m-1}r_0\}$, where $r_0 = b - Ax_0$ is the residual of the initial approximation. Finally, GMBACK determines an approximation $x_m^{GB} \in x_0 + \mathcal{K}_m$ which solves the following minimization problem:

$$\min_{x_m \in x_0 + \mathcal{K}_m} \|\Delta_A\|_F \quad \text{subject to} \quad (A - \Delta_A)\, x_m = b.$$

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

Arnoldi

- Let $r_0 = b - Ax_0$, $\beta = \|r_0\|_2$ and $v_1 = \frac{1}{\beta} r_0$;
- For $j = 1, \ldots, m$ do

$$h_{ij} = (Av_j, v_i), \quad i = 1, \ldots, j$$
$$\hat{v}_{j+1} = A\hat{v}_j - \sum_{j=1}^{i} h_{ij} v_i$$
$$h_{j+1,j} = \|\hat{v}_{j+1}\|_2$$
$$v_{j+1} = \frac{1}{h_{j+1,j}} \hat{v}_{j+1}$$

- Form the Hessenberg matrix $\bar{H}_m \in \mathbb{R}^{(m+1)\times m}$ with the (possible) nonzero elements h_{ij} computed above and the matrix $V_m \in \mathbb{R}^{N\times m}$ having on columns the vectors v_j: $V_m = [v_1 \ldots v_m]$;

GMBACK

- Let $\hat{H}_m = \begin{bmatrix} -\beta e_1 & \bar{H}_m \end{bmatrix} \in \mathbb{R}^{(m+1)\times(m+1)}$, $\hat{G}_m = [x_0 \quad V_m] \in \mathbb{R}^{N\times(m+1)}$
  $P = \hat{H}_m^t \hat{H}_m \in \mathbb{R}^{(m+1)\times(m+1)}$ and $Q = \hat{G}_m^t \hat{G}_m \in \mathbb{R}^{(m+1)\times(m+1)}$;
- Determine an eigenvector u_{m+1} corresponding to the smallest eigenvalue λ_{m+1}^{GB} of the generalized eigenproblem $Pu = \lambda Qu$;
- If the first component $u_{m+1}^{(1)}$ is nonzero, compute the vector $y_m^{GB} \in \mathbb{R}^m$ by scaling u_{m+1} such that

$$\begin{bmatrix} 1 \\ y_m^{GB} \end{bmatrix} = \frac{1}{u_{m+1}^{(1)}} u_{m+1};$$

- Set $x_m^{GB} = x_0 + V_m y_m^{GB}$.

This algorithm may lead to two possible breakdowns, either in the Arnoldi method or in the scaling of u_{m+1}. The first one is as for GMRES a happy breakdown, because the solution may be determined exactly using $\bar{H}_m$ and V_m. The second one 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 x_0. We shall assume in the following analysis that x_m^{GB} exist.

It is worth noting that the algorithm may be used in the restarted version, by taking after the m steps the computed solution x_m^{GB} as the new initial approximation.

We shall prove the following result:

PROPOSITION 2.1. *Consider some arbitrary elements $x_0 \in \mathbb{R}^N$ and $m \in \{1, \ldots, N\}$. If there exists a GMBACK solution x_m^{GB}, then its corresponding* *backward error* $\Delta_{A,m}^{GB} \in \mathbb{R}^{N\times N}$ *satisfies*

$$\left\|\Delta_{A,m}^{GB} \cdot x_m^{GB}\right\|_2 = \left\|r_m^{GB}\right\|_2.$$

Proof. Kasenally [11] proved that the backward error $\Delta_{A,m}^{GB}$ corresponding to x_m^{GB} is given by

$$\Delta_{A,m}^{GB} = V_{m+1}\left(\bar{H}_m y_m^{GB} - \beta e_1\right)\frac{\left(x_m^{GB}\right)^t}{\|x_m^{GB}\|_2^2}. \tag{2.1}$$

The matrices V_{m+1} and $\bar{H}_m$ computed in the Arnoldi algorithm satisfy the following known relation (see for example [22]):

$$AV_m = V_{m+1}\bar{H}_m,$$

which shows that

$$\begin{aligned}\left\|V_{m+1}\left(\bar{H}_m y_m^{GB} - \beta e_1\right)\right\|_2 &= \left\|AV_m y_m^{GB} - r_0\right\|_2\\ &= \left\|AV_m y_m^{GB} + Ax_0 - b\right\|_2\\ &= \left\|Ax_m^{GB} - b\right\|_2\\ &= \left\|r_m^{GB}\right\|_2.\end{aligned}$$

Taking into account (2.1), we are led to the stated result. □

3. THE SUPERLINEAR CONVERGENCE OF THE NEWTON-GMBACK METHOD

The Newton-GMBACK iterates may be written in two equivalent ways

$$\left(F'(y_k) - \Delta_{A_k}^{GB}\right)s_k^{GB} = -F(y_k) \tag{3.1}$$

$$F'(y_k)\,s_k^{GB} = -F(y_k) + r_k^{GB}, \quad k = 0, 1, \dots, \tag{3.2}$$

where for the first writting we have considered the linear systems $A_k s = b_k$ with $A_k = F'(y_k)$ and $b_k = -F(y_k)$.

Applying Theorem 1.1 of Dennis and Moré for the first writing of the iterates we get

THEOREM 3.1. *Assume the standard conditions hold and that for a given element* $y_0 \in D$, *the sequence of Newton-GMBACK iterates is well defined and converges to* y^*. *Then the convergence is q-superlinear if and only if*

$$\left\|r_k^{GB}\right\|_2 = o\left(\left\|s_k^{GB}\right\|_2\right), \qquad \text{as } k \to \infty.$$

Proof. We first observe that under the notations of formula (3.1), $B_k = F'(y_k) - \Delta_{A_k}^{GB}$ such that $B_k - F'(y_k) = -\Delta_{A_k}^{GB}$. Next, Theorem 1.1 and Proposition 2.1 lead to the stated affirmation. □

Theorem 1.2 of Dembo, Eisenstat and Steihaug applied to the writting (3.2) yields

THEOREM 3.2. *Under the same assumptions of Theorem* 3.1, *the convergence of the Newton-GMBACK iterates is q-superlinear if and only if*

$$\left\| r_k^{GB} \right\| = o\left(\left\| F(y_k) \right\|\right), \quad \text{as } k \to \infty.$$

One can see that the damped and perturbed quasi-Newton method of Martinez, Parada and Tapia must be reduced either to the quasi-Newton or to the IN method if we want to fit the Newton-GMBACK iterates in this frame.

Regarding the Newton-GMBACK iterates as IPN iterates we get

THEOREM 3.3. *Under the same assumptions of Theorem* 3.1, *the convergence of the Newton-GMBACK iterates is q-superlinear if and only if*

$$\left\| r_k^{GB} \right\| = o\left(\left\| F(y_k) \right\|\right), \quad \text{as } k \to \infty.$$

Proof. Taking $\Delta_k = -\Delta_{A_k}^{GB}$, $\delta_k = 0$ and $\hat{r}_k = 0$, $k = 0, 1, \ldots$ in Theorem 3.3 we are led by the writting (3.1) and by Proposition 2.1 to the stated affirmation. □

REMARK. Since the proofs from our IPN model relied on the IN one, Theorems 3.2 and 3.3 were expected to yield the same results. □

In order to complete our initial assertion we see that it suffices to show that $(\|F(y_k)\|)_{k\geq 0}$ and $(\|s_k\|)_{k\geq 0}$ have the same rate of convergence. This is true by the following considerations.

Walker proved for an arbitrary sequence $(y_k)_{k\geq 0} \subset \mathbb{R}^N$ that it converges q-superlinearly only at the same time with $(y_{k+1} - y_k)_{k\geq 0}$.

LEMMA 3.4. [24]. *Let* $(y_k)_{k\geq 0} \subset \mathbb{R}^N$ *be a convergent sequence and denote* $s_k = y_{k+1} - y_k$, $k = 0, 1, \ldots$ *Then* $(y_k)_{k\geq 0}$ *converges q-superlinearly if and only if* $(s_k)_{k\geq 0}$ *converges q-superlinearly.*

The fact that $(y_k)_{k\geq 0}$ and $(s_k)_{k\geq 0}$ have precisely the same rate when converging q-superlinearly is known for a longer time, by a result of Dennis and Moré.

LEMMA 3.5. [7]. *In the hypotheses of the above Lemma, if* $(y_k)_{k\geq 0}$ *converges q-superlinearly at* y^*, *then*

$$\lim_{k\to\infty} \frac{\|y^* - y_k\|}{\|y_{k+1} - y_k\|} = 1.$$

The connection between the rates of $\|y^* - y_k\|$ and $\|F(y_k)\|$ is expressed by the following result obtained by Dembo, Eisenstat and Steihaug:

LEMMA 3.6. [6]. *Under the standard assumptions on* F *there exists* $\varepsilon, \beta > 0$ *such that*

$$\tfrac{1}{\beta} \|y^* - y\| \leq \|F(y)\| \leq \beta \|y^* - y\| \quad \textit{for all } \|y - y^*\| \leq \varepsilon.$$

The equivalence of the characterizations is now completed.

REFERENCES

[1] P.N. BROWN, *A theoretical comparison of the Arnoldi and GMRES algorithms*, SIAM J. Sci. Stat. Comput., 12 (1991) no. 1, pp. 58–78.

[2] P.N. BROWN and H.F. WALKER, *GMRES on (nearly) singular systems*, SIAM J. Matrix Anal. Appl., 18 (1997) no. 1, pp. 37–51.

[3] E. CĂTINAŞ, *A note on inexact secant methods*, Rev. Anal. Numér. Théor. Approx., 25 (1996) nos. 1–2, pp. 33–41.

[4] E. CĂTINAŞ, *Newton and Newton-Krylov methods for solving nonlinear systems in* $\mathbb{R}^n$, Ph.D. Thesis, Cluj-Napoca, Romania, (to be defended).

[5] E. CĂTINAŞ, *Inexact perturbed Newton methods and some applications for a class of Krylov solvers*, J. Optim. Theory Appl., submitted.

[6] R.S. DEMBO, S.C. EISENSTAT, and T. STEIHAUG, *Inexact Newton methods*, SIAM J. Numer. Anal., 19 (1982) no. 2, pp. 400–408.

[7] J.E. DENNIS, JR. and J.J. MORÉ, *A characterization of superlinear convergence and its application to quasi-Newton methods*, Math. Comput., 28 (1974) no. 126, pp. 549–560.

[8] J.E. DENNIS, JR. and J.J. MORÉ, *Quasi-Newton methods, motivation and theory*, SIAM Review, 19 (1977) no. 1, pp. 46–89.

[9] J.E. DENNIS, JR. and R.B. SCHNABEL, *Numerical Methods for Unconstrained Optimization and Nonlinear Equations*, Prentice-Hall Series in Computational Mathematics, Englewood Cliffs, NJ, 1983.

[10] S.C. EISENSTAT and H.F. WALKER, *Choosing the forcing terms in an inexact Newton method*, SIAM J. Sci. Comput., 17 (1996) no.1, pp. 16–32.

[11] E.M. KASENALLY, *GMBACK: a generalised minimum backward error algorithm for nonsymmetric linear systems*, SIAM J. Sci. Comput., 16 (1995) no. 3, pp. 698–719.

[12] E.M. KASENALLY and V. SIMONCINI, *Analysis of a minimum perturbation algorithm for nonsymmetric linear systems*, SIAM J. Numer. Anal, 34 (1997) no. 1, pp. 48–66.

[13] C.T. Kelley, *Iterative Methods for Linear and Nonlinear Equations*, SIAM, Philadelphia, 1995.

[14] H.J. Martinez, Z. Parada and R.A. Tapia, *On the characterization of Q-superlinear convergence of quasi-Newton interior-point methods for nonlinear programming*, Bol. Soc. Mat. Mexicana, 1 (1995) no. 3, pp. 137–148.

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

[16] I. Păvăloiu, *Introduction in the Approximation of the Solutions of Equations Theory*, Ed. Dacia, Cluj-Napoca, 1976 (in Romanian).

[17] F.A. Potra and V. Pták, *Nondiscrete Induction and Iterative Processes*, Pitman, London, 1984.

[18] F.A. Potra, *On Q-order and R-order of convergence*, J. Optim. Theory Appl., 63 (1989) no. 3, pp. 415–431.

[19] W.C. Rheinboldt, *Methods for Solving Systems of Nonlinear Equations*, SIAM, Philadelphia, 1996.

[20] Y. Saad, *Krylov subspace methods for solving large unsymmetric linear systems*, Math. Comp., 37 (1981) no. 155, pp. 105–126.

[21] Y. Saad, *Iterative Methods for Sparse Linear Systems*, PWS Pub. Co., Boston, 1996.

[22] Y. Saad and M.H. Schultz, *GMRES: a generalized minimal residual algorithm for solving nonsymmetric linear systems*, SIAM J. Sci. Stat. Comput., 7 (1986) no. 3, pp. 856–869.

[23] G.W. Stewart and J.-G. Sun, *Matrix Perturbation Theory*, Academic Press, New York, 1990.

[24] H.F. Walker, *An approach to continuation using Krylov subspace methods*, Research Report 1/97/89, Dept. of Math., Utah State University, submitted to Computational Science in the 21st Century.

Received March 3, 1998

"T. Popoviciu" Institute of Numerical Analysis
Romanian Academy
P.O. Box 68, 3400 Cluj-Napoca 1
Romania