PAMM · Proc. Appl. Math. Mech. **5**, 785–786 (2005) / **DOI** 10.1002/pamm.200510367

The relationship between the models of perturbed Newton iterations, with applications

Emil Cătinaş*

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

The quasi-Newton method and the inexact Newton method are classical models of Newton iterates which take into account certain error terms. We have recently introduced the inexact perturbed Newton method, which takes into account all the possible sources of perturbations. We show that these models are in fact equivalent, in the sense that each one may be used to characterize the high convergence orders of the other two.

1 The equivalence of the inexact and quasi-Newton methods

Let $F(x) = 0$ be a nonlinear system, with $F : D \subseteq \mathbb{R}^n \to \mathbb{R}^n$. The local convergence of the Newton iterates

$$F'(x_k)s_k = -F(x_k), \quad x_{k+1} = x_k + s_k, \qquad k = 0, 1, \ldots, \quad x_0 \in D,$$

is studied under the following usual conditions (which will be implicitly assumed hereafter):
- there 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.

Given a norm $\|\cdot\|$ on $\mathbb{R}^n$, these hypotheses assure the existence of a radius $r > 0$ such that the Newton iterates converge q-superlinearly to x^* for any initial approximation x_0 with $\|x_0 - x^*\| < r$ [8, Th.10.2.2] (see also [10, Th.4.4]):

$$\limsup_{k\to\infty} \frac{\|x_{k+1}-x^*\|}{\|x_k-x^*\|} = 0, \quad \text{(assuming } x_k \neq x^* \text{ for all } k \geq k_0\text{)},$$

also denoted by $\|x_{k+1} - x^*\| = o(\|x_k - x^*\|)$, as $k \to \infty$.

In the practical applications, there are used mainly two models of perturbed Newton iterates: the quasi-Newton (QN) and the inexact Newton (IN) method

$$B_k s_k = -F(x_k),$$
$$F'(x_k)s_k = -F(x_k) + r_k.$$

The first model assumes the use of perturbed Jacobians, while the second assumes that at each step the correction s_k is only approximately determined (i.e., having nonzero residual r_k). The superlinear convergence (as well as the q-orders $p \in (1, 2]$) of these methods was characterized by Dennis and Moré, respectively Dembo, Eisenstat and Steihaug.

Theorem 1.1 ([4]) *Consider a sequence $(B_k)_{k\geq 0} \subset \mathbb{R}^{n\times n}$ of invertible matrices and an initial approximation $x_0 \in D$. If the* QN *iterates converge to x^*, then they converge superlinearly iff*

$$\frac{\|(B_k - F'(x^*))(x_{k+1}-x_k)\|}{\|x_{k+1}-x_k\|} \to 0, \quad \text{as } k \to \infty. \tag{DM}$$

Theorem 1.2 ([3]) *If the* IN *iterates converge to x^*, then the convergence is superlinear iff*

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

Denoting $\Delta_k = B_k - F'(x_k)$, the quasi-Newton iterates are transcribed as (QN'):

$$(F'(x_k) + \Delta_k) s_k = -F(x_k).$$

The following two results show the connection between the conditions for the superlinear convergence of the two methods.

Theorem 1.3 ([2]) *If the* QN' *iterates are written as* IN *iterates,*

$$F'(x_k)s_k = -F(x_k) - \Delta_k s_k,$$

then the (DES) *condition for these resulted iterates is equivalent to the* (DM) *condition for the initial* QN' *iterates.*

* e-mail: ecatinas@ictp.acad.ro, web: www.ictp.acad.ro/catinas, Phone: +40 264 420 942, Fax: +40 264 420 942

Let $\|\cdot\|_2$ denote the Euclidean norm.

Theorem 1.4 ([2]) *If the* IN *are written as* QN' *iterates,*

$$\left(F'(x_k) - \frac{1}{\|s_k\|_2^2} r_k s_k^t\right) s_k = -F(x_k),$$

then the (DM) *condition for these resulted iterates is equivalent to the* (DES) *condition for the initial* IN *iterates.*

2 The inexact perturbed Newton method

In [1] we have considered the inexact perturbed Newton (IPN) method and characterized its superlinear convergence:

$$(F'(x_k) + \Delta_k) s_k = \big(-F(x_k) + \delta_k\big) + \hat{r}_k.$$

Previous results contained only sufficient conditions. The elements $\Delta_k \in \mathbb{R}^{n\times n}$ represent perturbations to the Jacobians, $\delta_k \in \mathbb{R}^n$ perturbations to the function evaluations, while $\hat{r}_k \in \mathbb{R}^n$ are the residuals of the approximate solutions s_k of the resulted perturbed linear systems. The perturbation terms may contain errors resulted from floating point representation/arithmetical operations, so this model may be conceived to contain all sources of perturbations. In [1] we have assumed that the perturbed Jacobians are invertible. This condition may be relaxed, and one can assume just that the perturbed linear systems are compatible (i.e., the iterations are well defined):

Theorem 2.1 ([2]) *If the* IPN *iterates are well defined and converge to x^*, then the convergence is superlinear iff*

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

As shown in [2], the (DM) and (DES) conditions may be retrieved from the above one; on the other hand, the IPN iterates may be regarded as IN iterates, so it follows that the conditions characterizing the high convergence orders of the QN, IN and IPN methods are equivalent.

The above result can be successfully applied to characterize the convergence of some Newton-Krylov methods. Let

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

and consider an initial approximation u_0 to the exact solution $u^* = A^{-1}b$. Denote $\mathcal{K}_m := \operatorname{span}\{r_0, Ar_0, \ldots, A^{m-1}r_0\}$, $1 \le m \le n$, where $r_0 := b - Au_0$.

The GMBACK algorithm [5] finds $u_m^{GB} \in u_0 + \mathcal{K}_m$ as argmin for (i.e., minimizes the backward error in A)

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

where $\|\cdot\|_F$ denotes the Frobenius norm. Depending on the parameters, the above problem may have a unique, several ones, or no solution at all. Therefore, Theorem 1.1 is not the best choice (whereas Theorem 2.1 is) for characterizing the superlinear convergence of the Newton-GMBACK iterates, when written in the QN form

$$\big(F'(x_k) - \Delta_{k,m_k}^{GB}\big) s_{k,m_k}^{GB} = -F(x_k).$$

The same holds for the MINPERT method [6] which finds u_m^{MP} as argmin for (i.e., minimizes the joint backward error)

$$\left\|[\Delta_m^{MP}\ \delta_m^{MP}]\right\|_F = \min_{x_m \in x_0 + \mathcal{K}_m} \left\|[\Delta_m\ \delta_m]\right\|_F \quad \text{w.r.t. } (A - \Delta_m)x_m = b + \delta_m.$$

Acknowledgements I would like to express my gratitude to the organizers, who made possible my participation to the conference.

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, *The inexact, inexact perturbed and quasi-Newton methods are equivalent models*, Math. Comp., **74** (2005) no. 249, pp. 291–301.
[3] R. S. Dembo, S. C. Eisenstat and T. Steihaug, *Inexact Newton methods*, SIAM J. Numer. Anal. **19** (1982), 400–408.
[4] 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.
[5] E. M. Kasenally, *GMBACK: a generalised minimum backward error algorithm for nonsymmetric linear systems*, SIAM J. Sci. Comput. **16** (1995), 698–719.
[6] E. M. Kasenally and V. Simoncini, *Analysis of a minimum perturbation algorithm for nonsymmetric linear systems*, SIAM J. Numer. Anal. **34** (1997), 48–66.
[7] C. T. Kelley, *Iterative Methods for Linear and Nonlinear Equations*, SIAM, Philadelphia, 1995.
[8] J. M. Ortega and W. C. Rheinboldt, *Iterative Solution of Nonlinear Equations in Several Variables*, Academic Press, New York, 1970.
[9] F. A. Potra, *Q-superlinear convergence of the iterates in primal-dual interior-point methods*, Math. Progr., **91** no. 1, Ser. A, pp. 99–115, 2001.
[10] W. C. Rheinboldt, *Methods for Solving Systems of Nonlinear Equations*, Second ed., SIAM, Philadelphia, PA, 1998.