Critical solutions of nonlinear equations*

A.F. Izmailov, A.S. Kurennoy, and M.V. Solodov
Moscow State University, Moscow, Russia,
Tambov State University, Tambov, Russia,
IMPA, Rio de Janeiro, Brazil

1. Introduction and the key assumption

We consider a nonlinear equation

$$\Phi(u) = 0,$$

where $\Phi : \mathbb{R}^p \to \mathbb{R}^p$ is a given smooth mapping. Even though here we deal solely with the case when the number of equations is the same as the number of variables, it will never be assumed that the solution in question is isolated, and moreover, the case of nonisolated solutions will be of special interest. Observe that any nonisolated solution $\bar{u}$ is necessarily singular in a sense that $\Phi'(\bar{u})$ is a singular matrix.

In [7], it was shown that a solution $\bar{u}$ of the unconstrained equation "survives" perturbations in large classes if Φ is smooth enough, and there exists $\bar{v} \in \ker \Phi'(\bar{u})$ such that Φ is 2-regular at $\bar{u}$ in the direction $\bar{v}$, the latter meaning that

$$\operatorname{im} \Phi'(\bar{u}) + \Phi''(\bar{u})[\bar{v}, \ker \Phi'(\bar{u})] = \mathbb{R}^p.$$

Importantly, such $\bar{v}$ may exist even if $\bar{u}$ is a nonisolated solution.

Furthermore, as demonstrated in [7], if $\bar{u}$ is a singular solution, the needed $\bar{v}$ cannot belong to $T_{\Phi^{-1}(0)}(\bar{u})$, and hence it can never exist if $T_{\Phi^{-1}(0)}(\bar{u}) = \ker \Phi'(\bar{u})$. The latter is one of the two ingredients of the concept of noncriticality of solution $\bar{u}$, as introduced in [7] The second ingredient is Clarke regularity of $\Phi^{-1}(0)$ at $\bar{u}$, and as demonstrated in [7], under the appropriate smoothness assumptions, this combination of properties is equivalent to the local Lipschitzian error bound

$$\operatorname{dist}(u, \Phi^{-1}(0)) = O(\|\Phi(u)\|) \quad \text{as } u \in \mathbb{R}^p \text{ tends to } \bar{u},$$

which is known to be equivalent to the following upper Lipschitzian property:

$$\operatorname{dist}(u(w), \Phi^{-1}(0)) = O(\|w\|) \quad \text{as } w \in \mathbb{R}^p \text{ tends to } 0,$$

*This research was supported in part by the Volkswagen Foundation, by the Russian Foundation for Basic Research Grant 17-01-00125, by CNPq Grant 303724/2015-3, and by FAPERJ Grant 203.052/2016.

where $u(w)$ is any solution of the perturbed equation

$$\Phi(u) = w,$$

close enough to $\bar{u}$. In addition, the results in [7] imply that singular noncritical solutions of the unconstrained equation can only be stable subject to very special perturbations. At the same time, critical solutions (i.e., those which are not noncritical), or, more precisely, those solutions for which $T_{\Phi^{-1}(0)}(\bar{u})$ is a proper subset of $\ker \Phi'(\bar{u})$, can naturally satisfy the 2-regularity condition with some $\bar{v} \in \ker \Phi'(\bar{u})$, and hence, be stable subject to wide classes of perturbations.

In this work, we demonstrate that 2-regularity in a direction $\bar{v} \in \ker \Phi'(\bar{u})$ (which is our key assumption, and which may never hold at noncritical singular solutions, as discussed above) makes $\bar{u}$ specially attractive for sequences generated by Newton-type methods. Apart form the basic Newton method (NM), we will consider some modifications of it, intended specially for tackling the case of nonisolated solution. Specifically, these are the Levenberg–Marquardt method (L-MM) and the LP-Newton method (LP-NM).

2. Perturbed Newton method

We define the perturbed Newton method (pNM) for equation in question as follows: for a current iterate $u^k \in \mathbb{R}^p$, the next iterate is $u^{k+1} = u^k + v^k$, with v^k computed as a solution of linear equation

$$\Phi(u^k) + (\Phi'(u^k) + \Omega(u^k))v = \omega(u^k),$$

where $\Omega : \mathbb{R}^p \to \mathbb{R}^{p \times p}$ and $\omega : \mathbb{R}^p \to \mathbb{R}^p$ characterizes perturbation.

The following can be regarded as an extension of [5, Lemma 5.1] from the basic NM to pNM.

Every $u \in \mathbb{R}^p$ is uniquely decomposed into the sum $u = u_1 + u_2$, $u_1 \in (\ker \Phi'(\bar{u}))^\perp$, $u_2 \in \ker \Phi'(\bar{u})$. Let Π be the orthogonal projector onto $(\operatorname{im} \Phi'(\bar{u}))^\perp$, and assume that the norm is Euclidian. Let $\mathbf{S}$ stand for the unit sphere in $\mathbb{R}^p$.

Theorem. *Let Φ be twice differentiable near $\bar{u} \in \mathbb{R}^p$, with its second derivative Lipschitz-continuous with respect to $\bar{u}$. Let $\bar{u}$ be a solution of the nonlinear equation in question, and assume that Φ is 2-regular at $\bar{u}$ in a direction $\bar{v} \in \ker \Phi'(\bar{u}) \cap \mathbf{S}$. Let $\Omega : \mathbb{R}^p \to \mathbb{R}^{p \times p}$ and $\omega : \mathbb{R}^p \to \mathbb{R}^p$ satisfy the estimates*

$$\Omega(u) = O(\|u - \bar{u}\|), \quad \Pi\Omega(u) = O(\|u_1 - \bar{u}_1\|) + O(\|u - \bar{u}\|^2),$$

$$\omega(u) = O(\|u - \bar{u}\|^2), \quad \Pi\omega(u) = O(\|u - \bar{u}\|\|u_1 - \bar{u}_1\|) + O(\|u - \bar{u}\|^3).$$

Then there exist $\varepsilon = \varepsilon(\bar{v}) > 0$ and $\delta = \delta(\bar{v}) > 0$ such that any starting point $u^0 \in \mathbb{R}^p \setminus \{\bar{u}\}$ satisfying

$$\|u^0 - \bar{u}\| \leq \varepsilon, \quad \left\| \frac{u^0 - \bar{u}}{\|u^0 - \bar{u}\|} - \bar{v} \right\| \leq \delta$$

uniquely defines the sequence $\{u^k\} \subset \mathbb{R}^p$ of the pNM, $u_2^k \neq \bar{u}_2$ for all k, the sequence $\{u^k\}$ converges to $\bar{u}$, and

$$\lim_{k \to \infty} \frac{\|u^{k+1} - \bar{u}\|}{\|u^k - \bar{u}\|} = \frac{1}{2}, \quad \|u_1^{k+1} - \bar{u}_1\| = O(\|u^k - \bar{u}\|^2).$$

Theorem above establishes the existence of a set with nonempty interior, which is star-like with respect to $\bar{u}$, and such that the pNM initialized at any point of this set converges linearly to $\bar{u}$. Moreover, if Φ is 2-regular at $\bar{u}$ in at least one direction $\bar{v} \in \ker \Phi'(\bar{u})$, then set of such $\bar{v}$ is open and dense in $\ker \Phi'(\bar{u}) \cap \mathbf{S}$: its complement is the null set of the nontrivial homogeneous polynomial $\det \Pi\Phi''(\bar{u})[\cdot]|_{\ker \Phi'(\bar{u})}$ considered on $\ker \Phi'(\bar{u}) \cap \mathbf{S}$. The union of convergence domains coming with all such $\bar{v}$ is also a star-like convergence domain with nonempty interior. In the case when $\Phi'(\bar{u}) = 0$ (full singularity) this domain is quite large. In particular, it is "asymptotically dense": its complement is "asymptotically thin", and the only excluded directions are those in which Φ is not 2-regular at $\bar{u}$, which is the null set of a nontrivial homogeneous polynomial.

The assumptions on perturbations in Theorem automatically hold if

$$\Omega(u) = O(\|\Phi(u)\|), \quad \omega(u) = O(\|u - \bar{u}\|\|\Phi(u)\|).$$

3. Levenberg–Marquardt method

The L-MM is a well-established tool for tackling possibly nonisolated solutions. The iteration subproblem of this method has the form

$$\text{minimize} \quad \frac{1}{2}\|\Phi(u^k) + \Phi'(u^k)v\|^2 + \frac{1}{2}\sigma(u^k)\|v\|^2, \quad v \in \mathbb{R}^p,$$

where $\sigma : \mathbb{R}^p \to \mathbb{R}_+$ defines the regularization parameter. In particular, from the results in [8] it follows that being initialized near a noncritical solution, the L-MM with $\sigma(u) = \|\Phi(u)\|^2$ generates a sequence which is quadratically convergent to a (nearby) solution.

The L-MM subproblem is equivalent to the linear system

$$(\Phi'(u^k))^{\mathrm{T}}\Phi(u^k) + ((\Phi'(u^k))^{\mathrm{T}}\Phi'(u^k) + \sigma(u^k)I)v = 0,$$

characterizing stationary points of that convex optimization problem.

From [5, Lemma 3.1] it can be seen that $\bar{v}$ in Theorem applied to *the basic NM* comes with a "conic neighborhood" such that for every u in it, $\Phi'(u)$ is invertible, and $(\Phi'(u))^{-1} = O(\|u-\bar{u}\|^{-1})$. Multiplying both sides of the iteration system by $(((\Phi'(u^k))^{\mathrm{T}})^{-1} = (((\Phi'(u^k))^{-1})^{\mathrm{T}}$, we now obtain

$$\Phi(u^k) + (\Phi'(u^k) + \sigma(u^k)(((\Phi'(u^k))^{-1})^{\mathrm{T}})v = 0,$$

which is the pNM iteration system with the perturbation terms

$$\Omega(u) = \sigma(u)(((\Phi'(u))^{-1})^{\mathrm{T}} = O(\|u-\bar{u}\|^{-1}\sigma(u)), \quad \omega \equiv 0$$

as $u \to \bar{u}$, and therefore, the needed requirements on Ω will hold, e.g., if $\sigma(u) = \|\Phi(u)\|^{\tau}$ with $\tau \geq 2$.

Moreover, in the case when $\Phi'(\bar{u}) = 0$ (full singularity) the appropriate values are all $\tau \geq 3/2$.

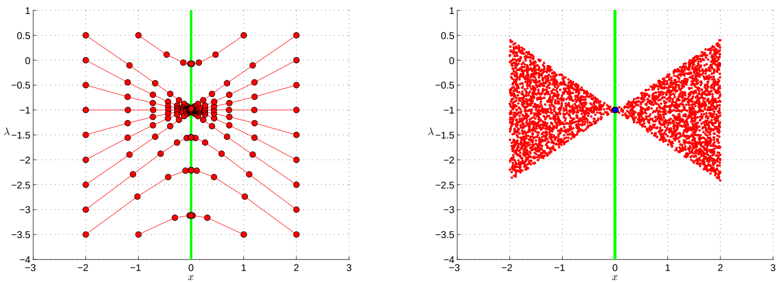

Fig. 1. Levenberg–Marquardt method with $\tau = 1$.

Example. Consider the equality-constrained optimization problem

$$\text{minimize} \quad x^2 \quad \text{subject to} \quad x^2 = 0.$$

Stationary points and associated Lagrange multipliers of this problem are characterized by the Lagrange optimality system which has the form

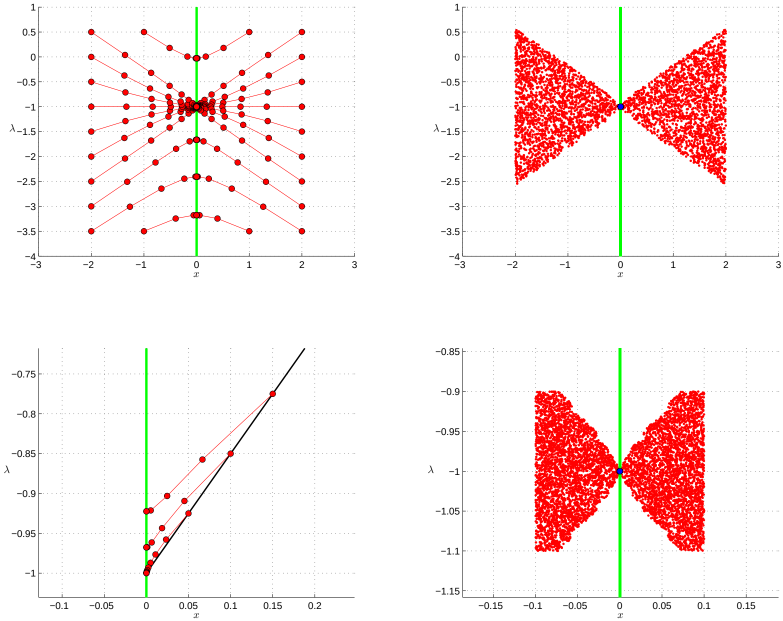

Fig. 2. Levenberg–Marquardt method with $\tau = 3/2$.

of a nonlinear equation with $\Phi : \mathbb{R}^2 \to \mathbb{R}^2$, $\Phi(u) = (2x(1+\lambda),\, x^2)$, where $u = (x,\, \lambda)$. The unique feasible point (hence, the unique solution, and the unique stationary point) of this problem is $\bar{x} = 0$, and the set of associated Lagrange multipliers is the entire $\mathbb{R}$. Therefore, the solution set of the Lagrange system (i.e., the primal-dual solution set) is $\{\bar{x}\} \times \mathbb{R}$. The unique critical solution is $\bar{u} = (\bar{x},\, \bar{\lambda})$ with $\bar{\lambda} = -1$, the one for which $\Phi'(\bar{u}) = 0$ (full singularity).

In Figures 1 and 2, the vertical gray line corresponds to the primal-dual solution set. These figures demonstrate some iterative sequences generated by the L-MM, and the domains from which convergence to the critical solution was detected.

4. LP-Newton method

A more recent approach, alternative to the L-MM, is the LP-NM proposed in [2]. The iteration subproblem of this method has the form

$$\begin{array}{ll} \text{minimize} & \gamma \\ \text{subject to} & \|\Phi(u^k) + \Phi'(u^k)v\| \leq \gamma \|\Phi(u^k)\|^2, \\ & \|v\| \leq \gamma \|\Phi(u^k)\|, \\ & (v,\, \gamma) \in \mathbb{R}^p \times \mathbb{R}. \end{array}$$

With l_∞ norm, this is a linear programming problem. As demonstrated in [2, 3], local convergence properties of the LP-NM (near noncritical solutions!) are the same as for L-MM.

The first constraint in the LP-NM subproblem can be interpreted as the pNM with $\Omega \equiv 0$ and some $\omega(\cdot)$, which will satisfy the assumptions in Theorem if the optimal value $\gamma(u)$ of the LP-NM subproblem with $u^k = u$ satisfies

$$\gamma(u) = O(\|\Phi(u)\|^{-1} \|u - \bar{u}\|)$$

as $u \to \bar{u}$.

In order to establish the needed estimate on $\gamma(\cdot)$, suppose again that u belongs to the "conic neighborhood" of $\bar{v}$ where *the basic NM* step $v(u)$ is uniquely defined, and $v(u) = O(\|u - \bar{u}\|)$. Then the point $(v,\, \gamma) = (v(u),\, \|v(u)\|/\|\Phi(u)\|)$ is feasible in the LP-N subproblem, and hence,

$$\gamma(u) \leq \gamma = \|\Phi(u)\|^{-1} \|v(u)\| = O(\|\Phi(u)\|^{-1} \|u - \bar{u}\|)$$

as $u \to \bar{u}$.

Figure 3 has the same meaning as Figure 2, but for LP-NM instead of L-MM, with the same conclusions.

A detailed exposition of these results can be found in [6]. The extensions of these and related results to constrained equations can be found in [1], [4].

References

1. A.V. Arutyunov and A.F. Izmailov Stability of possibly nonisolated solutions of constrained equations, with applications to complementarity and equilibrium problems. Set-Valued Var. Analys. 26 (2018), 327–352.
2. F. Facchinei, A. Fischer, and M. Herrich. An LP-Newton method: Nonsmooth equations, KKT systems, and nonisolated solutions. Math. Program. 146 (2014), 1–36.
3. F. Facchinei, A. Fischer, and M. Herrich. A family of Newton methods for nonsmooth constrained systems with nonisolated solutions. Math. Methods Oper. Res. 77 (2013), 433–443.

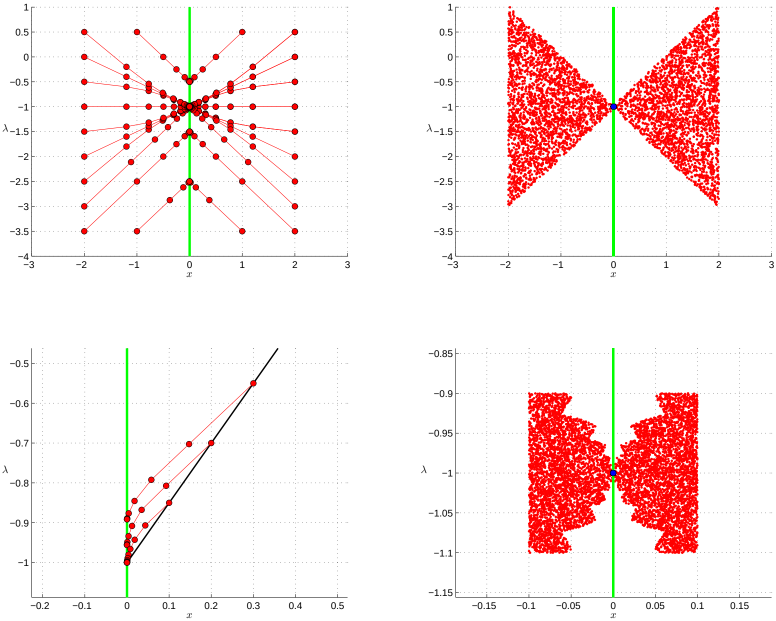

Fig. 3. LP-Newton method.

4. A. Fischer, A.F. Izmailov, and M.V. Solodov. Local attractors of Newton-type methods for constrained equations and complementarity problems with nonisolated solutions. J. Optim. Theory Appl. 2018. DOI 10.1007/s10957-018-1297-2.
5. A. Griewank. Starlike domains of convergence for Newton's method at singularities. Numer. Math. 35 (1980), 95–111.
6. A.F. Izmailov, A.S. Kurennoy, and M.V. Solodov. Critical solutions of nonlinear equations: local attraction for Newton-type methods. Math. Program. 167 (2018), 355–379.
7. A.F. Izmailov, A.S. Kurennoy, and M.V. Solodov. Critical solutions of nonlinear equations: stability issues. Math. Program. 168 (2018), 475–507.
8. N. Yamashita and M. Fukushima. On the rate of convergence of the Levenberg–Marquardt method. Computing. Suppl. 15 (2001), 237–249.