Noname manuscript No.
(will be inserted by the editor)

Solving underdetermined nonlinear equations by Newton-like method

Boris Polyak · Andrey Tremba

Received: date / Accepted: date

Abstract Newton method is one of the most powerful methods for finding solution of nonlinear equations. In its classical form it is applied for systems of n equations with n variables. However it can be modified for underdetermined equations (with $m < n$, m being the number of equations). Theorems on solvability of such equations as well as conditions for convergence and rate of convergence of Newton-like methods are addressed in the paper. The results are applied to systems of quadratic equations, one-dimensional equations and inequalities.

Keywords nonlinear equations · quadratic equations · existence of solution · Newton method · underdetermined equations · feasibility

1 Introduction

Consider nonlinear equation

$$g(x) = y, \tag{1}$$

written via vector function $g : \mathbb{R}^n \to \mathbb{R}^m$. There exists the huge bunch of literature on solvability of such equations and numerical methods for their solution, see e.g. the classical monographs [15,3]. One of the most powerful methods is *Newton method*, going back to such giants as Newton, Cauchy, Fourier. The general form of the method is due to Kantorovich [6,7]; on history

This work was supported by Russian Science Foundation (project 16-11-10015).

B. Polyak
Institute for Control Sciences, Profsoyuznaya 65, 117997 Moscow, Russia;
Skolkovo Institute of Science and Technology, Skolkovo Innovation Center Building 3, 143026 Moscow, Russia, E-mail: boris@ipu.ru

A. Tremba
Institute for Control Sciences, Profsoyuznaya 65, 117997 Moscow, Russia, E-mail: atremba@ipu.ru

and references see [8,23,19]. The basic version of Newton method for (1) is applicable when $g(x)$ is differentiable and $g'(x)$ is invertible (this implies $m = n$):

$$x^{k+1} = x^k - g'(x^k)^{-1}(g(x^k) - y) \tag{2}$$

The method converges under some natural conditions, moreover it can be used for obtaining existence theorems for the solution (see references cited above). Unfortunately Newton method converges only locally: it requires good initial approximation x^0 (so called “hot start”). Convergence conditions can be relaxed for *damped Newton method*

$$x^{k+1} = x^k - \alpha g'(x^k)^{-1}(g(x^k) - y)$$

with $0 < \alpha < 1$.

However the case of underdetermined equations ($m < n$) attracted much less attention. The pioneering result is due to Graves [5] in more general setting of Banach spaces. Graves' theorem for finite-dimensional case claims, that if condition

$$||g(x^a) - g(x^b) - A(x^a - x^b)|| \le C||x^a - x^b||$$

holds in the ball B of radius ρ for a matrix A with minimal singular value $\mu > C > 0$, then solution of the equation (1) exists provided $||y||$ is small enough, namely $||y|| \le \rho(\mu - C)$ and it can be found via a version of modified Newton method, where next iteration requires solution of the linear equation with matrix A, see [4,12] for details.

In explicit form Newton method for $m \ne n$ has been written by Ben-Israel [1]:

$$x^{k+1} = x^k - g'(x^k)^{\dagger}(g(x^k) - y),$$

where $A^{\dagger}$ stands for Moore-Penrose pseudoinverse of A. However the results in [1] are mostly oriented on overdetermined systems, and the assumptions of the theorems in [1] are hard to check.

In the paper [16] results on solvability of nonlinear equations in Banach spaces and on application of Newton-like methods have been formulated in different form. One of the results from [16] adopted to our notation and finite-dimensional case claims that if $g(0) = 0, g'(x)$ exists and is Lipschitz on B and estimate $||g'(x)^T h|| \ge \mu||h||,\ \mu > 0,\ \forall h$ holds on B, then equation (1) has a solution x^* provided $||y|| < \frac{\rho}{\mu}$. Another result deals with convergence of Newton method; however the method is not provided in explicit form.

The main contribution of the present paper is the analysis of the novel version of Newton method for solving the underdetermined equation (1) with $m < n$. It has the form

$$\begin{array}{l} x^{k+1} = x^k - \alpha_k z^k, \quad k = 0, 1, \ldots \\ z^k = \arg\min_z \{||z|| : g'(x^k)z = y - g(x^k)\}. \end{array} \tag{3}$$

If $m = n$ and $g'(x^k)^{-1}$ exists, the method coincides with classical Newton for $\alpha_k = 1$ and damped Newton for $\alpha_k = \alpha < 1$. Starting at $x^0 = 0$ the latter method converges to x^* under some additional constraints on $||y||$, L, ρ, μ, α (see Theorem 1 below on rigorous conditions). Notice that norms in $\mathbb{R}^n, \mathbb{R}^m$ can be chosen arbitrarily, and they imply different forms of Newton method (3) and various conditions on solvability and convergence.

The first goal of the present paper is to give explicit expressions of the method (3) for various norms and to provide simple, easily checkable conditions for convergence of the method. This also provides existence theorems: what is *the feasible set* Y such that $y \in Y$ implies solvability of (1).

The second goal is to develop constructive algorithms for choosing stepsizes α_k to achieve fast and *as global as possible* convergence. We suggest different strategies for constructing algorithms and study their properties.

We also examine some special cases of the nonlinear equations. One of them is the quadratic case, when all components of g are quadratic functions:

$$g_i(x) = \frac{1}{2}(A_i x, x) + (b_i, x), \quad A_i = A_i^T, \quad b_i \in \mathbb{R}^n. \tag{4}$$

In this case we try to specify above results and design the algorithms to check feasibility of a vector $y \in R^m$. The next important case is the scalar one, i.e. $m = 1$. We specify general results for scalar equations and inequalities; the arising algorithms have much in common with unconstrained minimization methods. Finally we discuss nonlinear equations having some special structure. Then convergence results can be strongly enhanced.

Few words on comparison with known results. The paper, which contains the closest results to ours, is [14]. Nesterov addresses the same problem (1) and his method (in our notation) has the form

$$\begin{aligned} &x^{k+1} = x^k - z^k, \quad k = 0, 1, \dots \\ &z^k = \arg\min_z \{||g(x^k) - y + g'(x^k)z|| + M||z||^2\}, \end{aligned}$$

where M is a scalar parameter to be adjusted at each iteration. Nesterov's assumptions are close to ours and his results on solvability of equations and on convergence of the method are similar. The main difference is the method itself; it is not clear how to solve the auxiliary optimization problem in Nesterov's method, while finding z^k in our method can be implemented in explicit form. Other papers on underdetermined equations mentioned above either do not specify the technique for solving the linearized auxiliary equation, or restrict analysis with Euclidean norm and pure Newton stepsize $\alpha_k = 1$, see e.g. [16, 20, 21, 9].

The rest of the paper is organized as follows. In next section we remind few notions and results, and discuss explicit (or half-explicit) solutions for optimization problem in (3). Next (Section 3) we prove simple solvability conditions for (1). In Section 4 we propose few variants of general Newton algorithm (3) and estimate their convergence rate. Some particular cases (scalar equations and inequalities, quadratic equations, problems with special structure) are treated in Section 5. Results of numerical simulation are exhibited in Section 6. Conclusion part finalizes the paper (Section 7).

2 Preliminaries

First of all let us specify subproblem of finding vector z^k in (3) for different norms of $x \in \mathbb{R}^n$.

1. For $||x|| = ||x||_1$ vector z^k is a solution of LP problem
$$\min\{||z||_1 : g'(x^k)z = y - g(x^k)\},$$
its dual is LP problem with one scalar constrain
$$\min\{||g'(x^k)^T h||_\infty : (y - g(x^k), h) = 1\}.$$
2. For $||x|| = ||x||_\infty$ vector z^k is a solution of LP problem
$$\min\{||z||_\infty : g'(x^k)z = y - g(x^k)\},$$
its dual is LP problem with one scalar constrain too
$$\min\{||g'(x^k)^T h||_1 : (y - g(x^k), h) = 1\}.$$
3. For $||x|| = ||x||_2$ vector z^k can be written explicitly
$$z^k = g'(x^k)^\dagger (g(x^k) - y).$$
For $m \leq n$ Moore-Penrose pseudo-inverse of a matrix A is written as $A^\dagger = A^T(AA^T)^{-1}$, if A has full row rank.

Thus in these (most important) cases algorithm (3) can be implemented effectively. Also solution of first two problems may be non-unique.

Note that linear constraint
$$Az = b, \quad b \in \mathbb{R}^m, \ z \in \mathbb{R}^n \tag{5}$$
in the primal optimization problems above describes either linear subspace, either empty set. The classical result below (which goes back to Banach, see [7,12,14]) guarantees solvability of the linear equation (5) and the estimates of its solutions. We suppose that spaces $\mathbb{R}^n, \mathbb{R}^m$ are equipped with some norms, the dual norms are denoted $||\cdot||_*$ (for linear functional c, associated with vector of same dimension $\|c\|_* = \sup_{x:\|x\|=1}(c,x)$). Operator norms are subordinate with the vector norms, e.g. for $A : X \to Y$ we have $\|Ax\|_Y \leq \|A\|_{X,Y}\|x\|_X$. In most cases we do not specify vector norms; dual and operator norms are obvious from the context.

Lemma 1 *If $A \in \mathbb{R}^{m\times n}$ satisfies condition*

$$\|A^T h\|_* \geq \mu_0 \|h\|_*, \;\; \mu_0 > 0, \tag{6}$$

for all $h \in \mathbb{R}^m$, then equation (5) has a solution for all $b \in \mathbb{R}^m$, and all solutions of optimization problem

$$\widehat{z} = \arg\min\{\|z\| : Az = b\}$$

have bounded norms $\|\widehat{z}\| \leq \frac{1}{\mu_0}\|b\|_*$.

The Lemma is claiming that matrix A has full row rank equal to m provided (6) holds. It is another way to say that the mapping $A : \mathbb{R}^n \to \mathbb{R}^m$ is *onto* mapping, i.e. covering all image space. In case of Euclidean norms parameter μ_0 is the smallest singular value of the matrix $\mu_0 = \sigma_m(A)$ (we denote singular values of a matrix in $\mathbb{R}^{m\times n}$ in decreasing order as $\sigma_1 \geq \sigma_2 \geq \ldots \geq \sigma_m$).

Finally we introduce sum of double exponential functions $H_k \colon [0, 1) \to \mathbb{R}_+$, $H_k(\delta) = \sum_{\ell=k}^{\infty} \delta^{(2^\ell)}$ (cf. [16]) and inverse function for the first of them $\Delta(\cdot) \colon \mathbb{R}_+ \to [0, 1)$, such that of $\Delta(H_0(\delta)) \equiv \delta, \;\; \delta \in [0, 1)$. All functions $H_k(\delta)$ are monotonically increasing on δ. In results below we also use specific constant

$$c = H_0\Big(\frac{1}{2}\Big) \approx 0.8164, \;\; \text{s.t. } \Delta(c) = \frac{1}{2}.$$

Following upper and lower approximations appear to be useful for simplifying Theorems' resulting expressions

$$H_k(\delta) \leq \frac{\delta^{(2^k)}}{1 - \delta^{(2^k)}} = \frac{1}{\delta^{-(2^k)} - 1}, \;\; \Delta(H) \geq \frac{H}{1 + H}. \tag{7}$$

3 Solvability of underdetermined equations

Below the problem of solvability of equation (1) is addressed. We apply algorithm (3) with small α and prove that iterations converge while the limit point is a solution. This techniques follows the idea from [16]. Remind that $\mathbb{R}^n, \mathbb{R}^m$ are equipped with some norms, the dual norms are denoted $||\cdot||_*$.

Assumptions.

A. $g(0) = 0, \; g(x)$ is differentiable in the ball $B = \{x \in \mathbb{R}^n : \|x\| \leq \rho\}$ and $g'(x)$ satisfies Lipschitz condition in B:

$$\|g'(x^a) - g'(x^b)\| \leq L\|x^a - x^b\|.$$

B. The following inequality holds for all $x \in B$ and some $\mu > 0$:

$$\|g'(x)^T h\|_* \geq \mu\|h\|_*, \;\; \forall h \in \mathbb{R}^m.$$

C. $\|y\| < \mu\rho$.

Theorem 1 *If conditions* $\mathbf{A}, \mathbf{B}, \mathbf{C}$ *hold then there exists a solution x^* of (1), and $\|x^*\| \leq \frac{\|y\|}{\mu}$.*

Proof We apply algorithm (3) with $\alpha > 0$ small enough and $x^0 = 0$. The algorithm is well defined — condition **B** and Lemma 1 imply existence of solutions z^k provided that $x^k \in B$; this is true for $k = 0$ and will be validated recurrently. Standard formula

$$g(x+z) = g(x) + \int_0^1 g'(x+tz)zdt$$

combined with condition **A** provides for $x = x^k,\ z = -\alpha z^k$ and $u_k = \|g(x^k) - y\|$ recurrent relation

$$u_{k+1} \le |1-\alpha|u_k + \frac{L\alpha^2}{2}\|z^k\|^2.$$

Now condition **B** and Lemma 1 transform this estimate into

$$u_{k+1} \le |1-\alpha|u_k + \frac{L\alpha^2 u_k^2}{2\mu^2}.$$

Choose $\alpha = \varepsilon\frac{2\mu^2}{Lu_0}(1 - \frac{u_0}{\mu\rho})$ with small $\varepsilon < 1$ satisfying $0 < \alpha < 1$; it is possible due to condition **C**. From the above inequality we get $u_{k+1} \le u_k(1 - \alpha + \alpha\varepsilon\frac{u_k}{u_0}(1 - \frac{u_0}{\mu\rho}))$. For $k = 0$ this implies $u_1 < u_0$ and recurrently $u_{k+1} < u_k$. We also get $u_{k+1} \le qu_k,\ q = (1 - \alpha + \alpha\varepsilon(1 - \frac{u_0}{\mu\rho})) < 1$. Thus $u_k \le q^k u_0$ and $u_k \to 0$ for $k \to \infty$.

On the other hand we have $\|x^{k+1} - x^k\| = \alpha\|z_k\| \le \frac{\|g(x^k)-y\|}{\mu} = \frac{u_k}{\mu} \le q^k\frac{u_0}{\mu}$. Hence for any k, s and for $k \to \infty$

$$\|x^{k+s} - x^k\| \le \sum_{i=k}^{k+s-1} \|x^{i+1} - x^i\| \le q^k \frac{u_0}{(1-q)\mu} \to 0.$$

It means that x^k is a Cauchy sequence and converges to some point $x^*(\varepsilon)$. We had $g(x^k) \to y$, thus continuity reasons imply $g(x^*(\varepsilon)) = y$. Now, for all iterations we got $\|x^k - x^0\| = \|x^k\| \le \sum_{j=0}^{k-1}\|x^{j+1} - x^j\| \le \alpha\frac{u_0}{\mu(1-q)} < \rho$. Hence all iterations x^k remain in the ball B and our reasoning was correct. Finally taking $\varepsilon \to 0$ we have $\|x^k\| \le \frac{u_0}{\mu}$ and its limit point $x^*(\varepsilon) \to x^*$ which is a solution as well and $\|x^*\| \le \frac{u_0}{\mu}$. □

Corollary 1 *If $\rho = \infty$ (that is conditions* $\mathbf{A}, \mathbf{B}$ *hold on the entire space $\mathbb{R}^n$) then equation* (1) *has a solution for arbitrary right-hand side y.*

It is worth noting that if we apply pure Newton method (i.e. take $\alpha_k \equiv 1$), the conditions of its convergence are more restrictive: we need $\|y\| \le \frac{2\mu^2}{L}$, that is we guarantee only local convergence even for $\rho = \infty$.

Corollary 2 *If $m = n$ and Condition* **B** *is replaced with $\|g'(x)^{-1}\| \le \frac{1}{\mu}, x \in B$, then the statement of Theorem 1 holds true.*

In this case our method (3) reduces to classical Newton method (2).

For $m < n$ equation (1) in general case has many solutions, we constructed just one of them, and we can not guarantee that the solution is the closest to the initial point $x^0 = 0$.

Among assumptions of Theorem 1 the main difficulty admits Assumption **B**. In many examples L can be estimated relatively easy; it is not the case for μ. Much simpler is to check its analog in a single point $x = 0$. Let us modify the solvability result in this way, with assumptions

B′. The following inequality holds for some $\mu_0 > 0$:

$$\|g'(0)^T h\|_* \geq \mu_0 \|h\|_*, \ \ \forall h \in \mathbb{R}^m.$$

C′. $\|y\| < \frac{\mu_0^2}{4L}, \ \frac{\mu_0}{2L} \leq \rho$.

Theorem 2 *If conditions* $\mathbf{A}, \mathbf{B}', \mathbf{C}'$ *hold then there exists a solution* x^* *of* (1)*, and* $\|x^*\| \leq \frac{2\|y\|}{\mu_0}$.

Proof For any linear operator A and its dual A^* (coinciding with A^T in real-valued vector spaces) holds $\|A\| = \|A^*\|_*$, where *dual operator* norm is induced by *dual* norms [7]. Then from condition **A** follows same Lipschitz constant for transposed derivative operator $g'(\cdot)^T$:

$$\|(g'(x^a) - g'(x^b))^T\|_* = \|g'(x^a) - g'(x^b)\| \leq L\|x^a - x^b\|, \ x^a, x^b \in B.$$

Denote $r = \frac{\mu_0}{2L}$, due to second part of Condition **C** ball $B_r = \{x : \|x\| \leq r\}$ lays within B. In B_r evidently holds condition **A**. Also in B_r we have $\|g'(x)^T h\|_* \geq \|g'(0)^T h\|_* - \|(g'(x) - g'(0))^T h\|_* \geq (\mu_0 - L\|x\|)\|h\|_* \geq (\mu_0 - Lr)\|h\|_* \geq \frac{\mu_0 \|h\|_*}{2}$, thus condition **B** holds with $\mu = \frac{\mu_0}{2}$. Condition **C** is also satisfied on this ball due to first part of Condition **C′** as $\|y\| < \frac{\mu_0^2}{4L} = r\mu$ and Theorem 1 holds (with r instead of ρ in its statement). □

It is worth noting that the result of Theorem 2 is always local, even for $\rho = \infty$ (compare with Corollary 1). From the proof it also follows that second condition $\frac{\mu_0}{2L} \leq \rho$ in Assumption **C′** is non-restrictive.

Corollary 3 *Assumption* **C′** *can be replaced with* $\|y\| \leq (\mu_0 - Lr^*)\, r^*$*, resulting in theorem's statement* $\|x^*\| \leq r^*$*, where* $r^* = \min\{\rho, \frac{\mu_0}{2L}\}$.

4 Main algorithms

For our purposes it is more convenient to write the main equation not in the form (1) (where we assumed $g(0) = 0, \ x^0 = 0$) but as

$$P(x) = 0, \ \ P : \mathbb{R}^n \to \mathbb{R}^m. \tag{8}$$

Then via trivial change of variables the Newton method becomes

$$\begin{aligned} z^k &= \arg \min_{P'(x^k) z = P(x^k)} \|z\|, \\ x^{k+1} &= x^k - \alpha_k z^k, \ \ k = 0, 1, \ldots \end{aligned}$$

with x^0 not necessarily equal to 0. Conditions on y ensuring solvability of (1) are essentially transformed into conditions on $P(x^0)$ guaranteeing solution of (8) and vice-versa.

In previous Section we proved solvability of equation by use of the algorithm with constant $\alpha_k \equiv \alpha > 0$; choosing α smaller we obtained larger solvability domain. However in this Section our goal is different — to reach the fastest convergence to a solution. For this purpose different strategies for design of step-sizes are needed. The basic policy is as follows. First, we rewrite assumptions in new notation.

A$''$. $P(x)$ is differentiable in the ball $B = \{x \in \mathbb{R}^n : \|x - x^0\| \le \rho\}$ and $P'(x)$ satisfies Lipschitz condition in B:

$$\|P'(x^a) - P'(x^b)\| \le L\|x^a - x^b\|.$$

B$''$. The following inequality holds for all $x \in B$ and some $\mu > 0$:

$$\|P'(x)^T h\|_* \ge \mu\|h\|_*, \ \ \forall h \in \mathbb{R}^m.$$

If $\mathbf{A}'', \mathbf{B}''$ hold true, we have same recurrent inequalities for $u_k = \|P(x_k)\|$:

$$u_{k+1} \le |1 - \alpha_k| u_k + \frac{L\alpha_k^2\|z_k\|^2}{2}, \tag{9}$$

$$u_{k+1} \le |1 - \alpha_k| u_k + \frac{L\alpha_k^2 u_k^2}{2\mu^2}, \tag{10}$$

the second one being just continuation of the first one based on estimate $\|z_k\|^2 \le \frac{u_k}{\mu}$, compare calculations in the proof of Theorem 1. Now we can minimize right-hand sides of these inequalities over α_k; it is natural to expect that such choice of step-size imply the fastest convergence of u_k to zero and thus the fastest convergence of iterations x_k to a solution. If one applies such policy based on inequality (10), optimal α depends on μ, L (Algorithm 1 below). Its value is hard to estimate in most applications, thus the method would be hard for implementation. Fortunately, we can modify the algorithm using parameter adjustment (Algorithm 2). On the other hand the same policy based on (9) requires just the value L, which is commonly available (Algorithm 3).

Thus we arrive to an algorithm which we call *Newton method* while in fact it is blended *pure Newton* with *damped Newton* with special rule for damping. In some relation it reminds *Newton method* for minimization of self-concordant functions [13].

4.1 Newton method with known constants

If both constants L and μ are known, then $\alpha_k = \min\{1, \frac{\mu^2}{L\|P(x^k)\|}\}$ is taken as minimizer of right-hand part (10).

Algorithm 1 (Basic Newton method)

$$z^k = \arg \min_{P'(x^k)z = P(x^k)} \|z\|,$$

$$x^{k+1} = x^k - \min\Big\{1, \frac{\mu^2}{L\|P(x^k)\|}\Big\} z^k, \; k \geq 0. \tag{11}$$

The algorithm is well-defined, as soon $\|P(x^k)\| = 0$ means that a solution is already found (formally $z^k = 0$, $\alpha_k = 1$ thereafter). We remind that in calculation of z^k any vector norm in $\mathbb{R}^n$ can be used, also any vector norm in $\mathbb{R}^m$ can used for $\|P(x^k)\|$, and constants L, μ must comply with these norms.

The update step in (11) can be written in less compact but more illustrative form:

$$\begin{array}{lll} x^{k+1} = x^k - \frac{\mu^2}{L\|P(x^k)\|} z^k, & \text{if } \|P(x^k)\| \geq \frac{\mu^2}{L} & \text{(Stage 1 step)}, \\ x^{k+1} = x^k - z^k, & \text{otherwise} & \text{(Stage 2 step)}. \end{array}$$

The latter case is a pure Newton step while the primal one is damped Newton step. Direction z^k calculation is the same in both stages. The result on convergence and rate of convergence is given below. We use upper ($\lceil\cdot\rceil$) and lower ($\lfloor\cdot\rfloor$) rounding to integer; function $\varDelta(\cdot)$ and constant $c \approx 0.8614$ were introduced in the end of Section 2.

Theorem 3 *Suppose that Assumptions $\mathbf{A}''$, $\mathbf{B}''$ hold and*

$$\|P(x^0)\| \leq \frac{\mu^2}{L} \times \begin{cases} 2\varDelta\left(\dfrac{L\rho}{2\mu}\right), & \textit{if } \dfrac{L\rho}{2\mu} \leq c, \qquad (12a) \\ 1 + \dfrac{1}{2}\left\lfloor \dfrac{L\rho}{\mu} - 2c \right\rfloor, & \textit{if } \dfrac{L\rho}{2\mu} > c, \qquad (12b) \end{cases}$$

then the sequence $\{x^k\}$ generated by Algorithm 1 converges to a solution $x^* : P(x^*) = 0$.

Function $\|P(x^k)\|$ is monotonically decreasing, and there are not more than

$$k_{\max} = \max\{0, \left\lceil \frac{2L}{\mu^2}\|P(x^0)\| \right\rceil - 2\} \tag{13}$$

iterations at Stage 1. At k-th step following estimates for the rate of convergence hold:

$$\|P(x^k)\| \leq \|P(x^0)\| - \frac{\mu^2}{2L}k, \qquad k < k_{\max}, \tag{14a}$$

$$\|x^k - x^*\| \leq \frac{\mu}{L}(k_{\max} - k + 2c), \qquad k < k_{\max}, \tag{14b}$$

$$\|P(x^k)\| \leq \frac{2\mu^2}{L} 2^{-(2^{(k-k_{\max})})}, \qquad k \geq k_{\max}, \tag{14c}$$

$$\|x^k - x^*\| \leq \frac{2\mu}{L} H_{k-k_{\max}}\Big(\frac{1}{2}\Big), \qquad k \geq k_{\max}. \tag{14d}$$

Proof Assume that all $x_k \in B,\ k \geq 0$. Below we state condition enabling this assumption. Using $u_k = \|P(x^k)\|$ and denoting $\beta = \frac{\mu^2}{L}$, we rewrite (10) with generic[1] stepsize α as

$$u_{k+1} \leq |1-\alpha| u_k + \frac{1}{2\beta}\alpha^2 u_k^2.$$

Its optimum over α is at $\alpha_k = \frac{\beta}{u_k} < 1$, if $\frac{\beta}{u_k} < 1$ (i.e. $u_k > \beta$); and $\alpha_k = 1$ otherwise.

During Stage 1 of damped Newton steps ($\alpha_k < 1$) target functional monotonically decreases as

$$u_{k+1} \leq u_k - \frac{\beta}{2}. \tag{15}$$

There are at most $k_{\max} = \max\{0,\ \lceil \frac{2u_0}{\beta} \rceil - 2\}$ iterations in the phase, say $\overline{k}$ ones, resulting in $u_{\overline{k}} \leq \beta$. As soon u_k reaches threshold β, the algorithm switches to Stage 2 (pure Newton steps). Then recurrent relation (15) becomes

$$u_{k+1} \leq \frac{1}{2\beta} u_k^2 = 2\beta \left(\frac{u_k}{2\beta}\right)^2,\ k \geq \overline{k}.$$

so we can write

$$u_{\overline{k}+\ell} \leq 2\beta \left(\frac{u_{\overline{k}}}{2\beta}\right)^{(2^\ell)} \leq 2\beta \left(\frac{1}{2}\right)^{(2^\ell)},\ \ell \geq 0.$$

For the second phase $\|x^{i+1} - x^i\| = \|z^i\| \leq \frac{1}{\mu} u_i$ due Lemma 1, and for $\ell_2 \geq \ell_1 \geq 0$ holds

$$\|x^{\overline{k}+\ell_2} - x^{\overline{k}+\ell_1}\| \leq \sum_{i=\ell_1}^{\ell_2-1} \|x^{i+1} - x^i\| \leq \frac{2\beta}{\mu}\Big(H_{\ell_1}\Big(\frac{u_{\overline{k}}}{2\beta}\Big) - H_{\ell_2}\Big(\frac{u_{\overline{k}}}{2\beta}\Big)\Big). \tag{16}$$

The $\{x^k\}$ sequence is a Cauchy sequence because $H_j(\frac{u_{\overline{k}}}{2\beta}) \leq H_j(\frac{1}{2}) \to_{j\to\infty} 0$. It converges to a point $x^* : \|P(x^*)\| = \lim_{k\to\infty} u_k = 0$ due to continuity of P, with

$$\|x^{\overline{k}+\ell} - x^*\| \leq \frac{2\beta}{\mu} H_\ell \left(\frac{u_{\overline{k}}}{2\beta}\right) \leq \frac{2\beta}{\mu} H_\ell\Big(\frac{1}{2}\Big),\ \ell \geq 0. \tag{17}$$

Next we are to estimate distance from points x^k in Stage 1 to the limit solution point x^*. One-step distance is bounded by constant $\|x^{k+1} - x^k\| \leq \frac{\alpha_k}{\mu} u_k = \frac{\beta}{\mu},\ \ k < \overline{k}$, and

$$\|x^k - x^*\| \leq \|x^{\overline{k}} - x^*\| + \sum_{i=k}^{\overline{k}-1} \|x^{i+1} - x^i\| \leq \frac{\beta}{\mu}(\overline{k} - k + 2c),\ k < \overline{k}. \tag{18}$$

Note that the formula also coincides with upper bound (17) at $k = \overline{k}$. Exact number $\overline{k}$ of steps in first phase is not known, but we can replace it with upper

[1] The relation is very alike to Newton method analysis of [16], with $\gamma = |1-\alpha|, \lambda = \frac{\alpha}{\mu}$.

bound $k_{\max}$ in all estimates (15)–(18). Substituting $\beta = \frac{\mu^2}{L}$ back we arrive to Theorem 3 bounds (14).

Finally we are to ensure our primal assumption of algorithm-generated points x^k being within B. This is guaranteed by one of two conditions, depending on whether the Algorithm starts from Stage 1 or Stage 2 steps.

In the first case $\|P(x^0)\| \ge \frac{\mu^2}{L}$, and $\|x^0 - x^k\|$ can be bounded similarly to (16)

$$\|x^0 - x^k\| \le \sum_{i=0}^{k-1} \|x^{i+1} - x^i\| \le \sum_{i=0}^{\overline{k}-1} \|x^{i+1} - x^i\| + \sum_{i=\overline{k}}^{\infty} \|x^{i+1} - x^i\| \le$$

$$\le \frac{\mu}{L}(\overline{k} + 2c) \le \frac{\mu}{L}(k_{\max} + 2c) = \frac{\mu}{L}\Big(\left\lceil \frac{2L}{\mu^2}\|P(x^0)\| \right\rceil - 2 + 2c\Big).$$

It is sufficient to satisfy $\|P(x^0)\| \le \frac{\mu^2}{L}(1 + \frac{1}{2}\lfloor \frac{L\rho}{\mu} - 2c \rfloor)$ for guaranteeing $\|x^0 - x^k\| \le \rho$. As we assumed $\|P(x^0)\| \ge \frac{\mu^2}{L}$ in this case, we conclude that this condition may hold only if $\frac{L\rho}{\mu} \ge 2c$. This results in (12b).

In the second case we have $\|P(x^0)\| < \frac{\mu^2}{L}$, and the algorithm makes pure Newton steps with $\alpha_k \equiv 1$ from the beginning. Then $\overline{k} = 0$, and from (16) follows

$$\|x^k - x^0\| \le \frac{2\mu}{L}(H_0\Big(\frac{\|P(x^0)\|}{2\beta}\Big) - H_k\Big(\frac{\|P(x^0)\|}{2\beta}\Big)) \le \frac{2\mu}{L} H_0\Big(\frac{L\|P(x^0)\|}{2\mu^2}\Big), \; k \ge 0.$$

The inequality $\|x^0 - x^k\| \le \rho, \; k \ge 0$ is satisfied if

$$\|P(x^0)\| \le \frac{2\mu^2}{L} \mathit{\Delta}\left(\frac{L\rho}{2\mu}\right).$$

In order to have $\|P(x^0)\| < \frac{\mu^2}{L}$, argument value $\frac{L\rho}{2\mu}$ must be less than c. This results in condition (12a). Altogether (12) ensures $x^k \in B, \; k \ge 0$ and bounds (14). □

Result on the rate of convergence means, roughly speaking, that after no more than $k_{\max}$ iterations one has very fast (quadratic) convergence. For good initial approximations $k_{\max} = 0$, and pure Newton method steps are performed from the very start.

If we use approximation bounds (7), then condition (12a) can be replaced with the simpler one:

$$\|P(x^0)\| \le \frac{2\mu^2}{L}\left(1 + \frac{2\mu}{L\rho}\right)^{-1}, \; \mathit{if} \; \frac{L\rho}{2\mu} \le c. \tag{19}$$

also (14d) can be roughly estimated as

$$\|x^k - x^*\| \le \frac{2\mu}{L}\frac{1}{2^{(2^{k-k_{\max}})} - 1}, \;\; k \ge k_{\max},$$

or even simpler bound $\|x^k - x^*\| \le 2.32 \cdot 2^{-2^{(k-k_{\max}-1)}} \frac{\mu}{L}$.

Corollary 4 *If $\rho = \infty$ (that is conditions $\mathbf{A}'', \mathbf{B}''$ hold on the entire space $\mathbb{R}^n$) then Algorithm 1 converges to a solution of* (8) *for any $x^0 \in \mathbb{R}^n$.*

4.2 Adaptive Newton method

Presented Algorithm 1 explicitly uses two constants μ and L but both enter into the algorithm as one parameter $\beta = \frac{\mu^2}{L}$. There is a simple modification allowing adaptively change an estimate of the parameter.

Input of the algorithm is an initial point x^0, approximation β_0 and parameter $0 < q < 1$.

```
Algorithm 2 (Adaptive Newton method)
1. Calculate
       $z^k = \arg \min_{P'(x^k)z = P(x^k)} \|z\|,$
       $\alpha_k = \min\{1, \frac{\beta_k}{\|P(x^k)\|}\},$
       $u_{k+1} = P(x^k - \alpha_k z_k).$
2. If either
       $\alpha_k < 1$ and $u_{k+1} < (1 - \frac{\alpha_k}{2}) u_k,$
   or
       $\alpha_k = 1$ and $u_{k+1} < \frac{1}{2} u_k,$
   holds, then go to Step 4. Otherwise
3. apply update rule $\beta_k \leftarrow q\beta_k$ and return to Step 1 without increasing counter.
4. Take
       $x^{k+1} = x^k - \alpha_k z_k,$
   set $\beta_{k+1} = \beta_k$, increase counter $k \leftarrow k + 1$, and go to Step 1.
```

Properties of Algorithm 2 are similar to Algorithm 1. We omit the formal proof of convergence; it follows the lines of the proof of Theorem 3 with respect to properties:

- Algorithm 2 does real steps at Step 4 and some number of fictitious steps resulting in update rule Step 3;
- β_k is non-increasing sequence;
- if $\beta_k < \beta$, then Step 3 won't appear and β_k won't decrease anymore. It means that there is maximum $\widehat{k} = \max\{0, \lceil \log_{1/q}(\frac{\beta_0}{\beta}) \rceil\}$ check steps. Minimal possible value of β_k is $\beta_{\min} = q^{\widehat{k}}\beta$ then, and number of Stage 1 steps is limited by $\widehat{k}_{\max} = \max\{0, \left\lceil 2\frac{\|P(x^0)\|}{\beta_{\min}} \right\rceil - 2\}$ as well;

– if Step 4 is made with $\beta_k > \beta$ due to validity of a condition in Step 2, then $\|P(x^{k+1})\|$ decrease *more* than at a corresponding step with "optimal" step-size $\alpha_k = \min\{1, \frac{\beta}{\|P(x^k)\|}\}$.

Let us mention two other versions of adaptive Newton method. First one uses *increasing* update (e.g. $\beta_{k+1} = q_2\beta_k$ with $q_2 > 1$) in the end of Step 4, thus adapting the constant to current x^k. Also other decrease policies can be applied to β_k in Step 3.

Second option is to use line-search or Armijo-like rules for choosing step-size α_k to minimize objective function $\|P(x^k - \alpha z^k)\|$ directly. Rigorous validation of the algorithms can be provided.

4.3 Method for L known

As we mentioned in the beginning of the section, we can use better approximation (9) instead of (10). It results to an algorithm using Lipschitz constant only, and it differs in update step-size.

Algorithm 3 (L-Newton method)

$$z^k = \arg \min_{P'(x^k)z = P(x^k)} \|z\|,$$

$$x^{k+1} = x^k - \min\left\{1, \frac{\|P(x^k)\|}{L\|z^k)\|^2}\right\} z^k, \ k \geq 0.$$

The algorithm is well-defined, as due to Assumption $\mathbf{B}''$ condition $\|z^k\| = 0$ holds only if $P(x^k) = 0$, i.e. the solution was found on previous step. Then $z^k = 0$, $\alpha_k = 0$ and $x^{k+1} = x^k$ thereafter.

Theorem 3 is valid for the Algorithm 3 with two exceptions. First, condition (12b) should be replaced with condition

$$\|P(x^0)\| \leq \frac{\mu^2}{2L} \left\lfloor \frac{-1 + \sqrt{25 + 16(\frac{L\rho}{\mu} - 2c)}}{2} \right\rfloor, \ \textit{if} \ \frac{L\rho}{2\mu} > c. \tag{20}$$

Second, upper bound (14b) should be replaced with

$$\|x^k - x^*\| \leq \frac{\mu}{L} \left(\frac{(k_{\max} - k)(k_{\max} - k + 5)}{4} + 2c \right), \ k < k_{\max}. \tag{21}$$

Corollary 1 on everywhere convergence for Algorithm 3 is valid as well. Proof is following same lines as of Theorem 3, and is omitted for brevity. The only notable difference of the Algorithms 3 is (possible) interlacing Stage 1 and Stage 2 steps, which may lead to *larger* step-size $\|x^{k+1} - x^k\| \leq \frac{u_k}{\mu}$ for $k \leq k_{\max}$, (cf. to a formulae prior to (18)). These large steps result to bound on distance (21), and eventually lead to more conservative condition (20). Surprisingly enough, in practice the algorithm (and its adaptive variant) sometimes converges faster than Algorithm 1, because direction-wise (along z^k) Lipschitz constant is less or equal than uniform Lipschitz constant of Assumption $\mathbf{A}''$, and convergence rate can be better.

The idea of adaptive algorithm with estimates L_k works as well for Algorithm 3; including its modifications with increasing L_k and/or line-search over α.

4.4 Pure Newton method

For comparison specify convergence property of pure Newton method ($\alpha_k = 1$).

Theorem 4 *Let conditions $\mathbf{A}'', \mathbf{B}''$ hold. If $\delta = \frac{L}{2\mu^2}\|P(x^0)\| < 1$ and $\frac{2\mu}{L}H_0(\delta) \le \rho$, then pure Newton method converges to a solution x^* of (8), and*

$$\|P(x^k)\| \le \frac{2\mu^2}{L}\delta^{(2^k)}, \quad \|x^k - x^*\| \le \frac{2\mu}{L}H_k(\delta).$$

It coincides with Corollary 1 of [16], proven in Banach space setup (a misprint in [16] is corrected here). In $m = n$ case the result is minor extension of Mysovskikh's theorem [7].

4.5 μ in a single point known

Most of the theorems above require quite strict Assumption $\mathbf{B}''$ or similar ones. In fact, usually we can estimate μ in one point, say, x^0, but not for entire ball B. Initially Kantorovich-type results on Newton method were proved for such "single-point condition" setup, and we proceed to such results as corollary of more general case.

Let μ_0 satisfy condition (6) with $A = P'(x^0)$. Following same reasoning as in Theorem 2 proof, we can estimate constants μ on balls B_r with variable radius

$$\mu(r) \ge \mu_0 - rL.$$

We added dependence on r to indicate that for arbitrary $r \in [0, \frac{\mu_0}{L})$ there is a corresponding constant $\mu > 0$.

If assumption of type $\mathbf{A}$ or $\mathbf{A}''$ hold on B_ρ, then Theorem 3 and their extensions can be rewritten in terms of $\mu(r)$ and r instead of ρ. It is done similarly to Corollary 3. Optimization over interval $r \in [0, \frac{\mu_0}{L}) \cap [0, \rho]$ gives maximal allowed range of $\|P(x^0)||$. The simplest choice is $r = \frac{\mu_0}{2L}$, then we can take $\mu = \frac{\mu_0}{2}$. We omit obvious versions of the Algorithms and convergence results for this case. Two examples of such theorems are given in Subsection 5.4, for the case where $\mathbf{A}'', \mathbf{B}''$ hold everywhere. We encountered that both Algorithm 1 and Algorithm 3 in such case are subject to the same upper bound, and, in fact, shall start with Stage 2 of pure Newton method.

5 Special Cases

In the section we outline few important cases in more detail, namely solving equations with special structure, solving scalar equations or inequalities, solvability of quadratic equations.

5.1 Structured problems

The problem is to solve $P(x) = y$ where $P(x)_i = \varphi(c_i^T x),\ c_i \in R^n, i = 1, \dots m$. Here $\varphi(t)$ is twice differentiable scalar function,

$$|\varphi'(t)| \geq \mu_\varphi > 0,\ \ |\varphi''(t)| \leq M,\ \ \forall t$$

It is not hard to see that Assumptions $\mathbf{A}'', \mathbf{B}''$ hold on the entire space $\mathbb{R}^n$ and Algorithm 1 converges, with Theorem 3 and Corollary 1 providing rate of convergence. The rate of convergence depends on estimates for L, μ which can be written as functions of μ_φ, M and singular values of matrix C with rows c_i. However the special structure of the problem allows to get much sharper results.

Indeed $P'(x) = D(x)C,\ D(x) = \operatorname{diag}(\varphi'(c_i^T x))$ and the main inequality (16) can be proved to be

$$u_{k+1} \leq (1-\alpha)u_k + \gamma\frac{\alpha^2 u_k^2}{2}, \gamma = \frac{M}{\mu_\varphi^2}.$$

Hence $u_{k+1} \leq u_k - \frac{1}{2\gamma}$ at Stage 1, thus this inequality does not depend on C! As the result we get estimates for the rate of convergence which are the same for ill-conditioned and well-conditioned matrices C.

This example is just an illustrating one (explicit solution of the problem can be found easily), but it emphasizes the role of special structure in equations to solve.

5.2 One-dimensional case

Suppose we solve one equation with n variables:

$$f(x) = 0,\ \ f : \mathbb{R}^n \to \mathbb{R}.$$

Here 0 is *not a minimal value* of f, thus it is not a minimization problem! Nevertheless our algorithms will remind some minimization methods. This case has some specific features compared with arbitrary m. For instance calculation of z^k may be done explicitly. Norm in image space is absolute value $|\cdot|$, and ℓ_p norms in pre-image space $\mathbb{R}^n$, $p \in \{1, 2, \infty\}$ can be considered. Then

$$\begin{aligned} z^k &= \frac{f(x^k)}{\|\nabla f(x^k)\|_\infty} e^i,\ i = \arg\max_i |\nabla f(x^k)_i|, && \text{in case of } \ell_1\text{-norm},\\ z^k &= \frac{f(x^k)}{\|\nabla f(x^k)\|_2^2} \nabla f(x^k), && \text{in case of Euclidean norm},\\ z^k &= \frac{f(x^k)}{\|\nabla f(x^k)\|_1} \operatorname{sign}(\nabla f(x^k)), && \text{in case of } \ell_\infty\text{-norm}, \end{aligned}$$

where $e^j = (0, \dots, 0, 1, 0, \dots, 0)^T$ is j-th orth vector, and $\operatorname{sign}(\cdot)$ function is coordinate-wise sign function, $\operatorname{sign} : \mathbb{R}^n \to \{-1, 1\}^n$.

Constant μ (and μ_0) are also calculated explicitly via conjugate (dual) vector norm $\mu = \min_{x \in B} \|\nabla f(x)\|_*$, $\mu_0 = \|\nabla f(x^0)\|_*$. For any norms $\|z^k\| = |f(x^k)|/\|\nabla f(x^k)\|_*$, and damped Newton step is performed iff $\|\nabla f(x^k)\|_*^2 < L|f(x^k)|$, otherwise pure Newton step is made.

If we choose ℓ_1 norm, the method becomes coordinate-wise one. Thus, if we start with $x^0 = 0$ and perform few steps (e.g. we are in the domain of attraction of pure Newton algorithm) we arrive to a *sparse* solution of the equation.

In Euclidean case a Stage 1 step (damped Newton) of Algorithm 1 is

$$x^{k+1} = x^k - \frac{1}{L} \operatorname{sign}(f(x^k)) \nabla f(x^k),$$

which is exactly gradient minimization step for function $|f(x^k)|$. Stage 2 (pure Newton) step is

$$x^{k+1} = x^k - \frac{f(x^k)}{\|\nabla f(x^k)\|_2^2} \nabla f(x^k).$$

This reminds well-known subgradient method for minimization of convex functions. However in our case we do not assume any convexity properties, and the direction may be either gradient or anti-gradient in contrast with minimization methods!

5.3 Solving an inequality

One-dimensional inequality

$$f(x) \le 0, \quad f : \mathbb{R}^n \to \mathbb{R}.$$

can be efficiently solved as well. Denote the set of points where the inequality is violated as $S = \{x : f(x) > 0\}$. Suppose that $\mu = \min_{x \in S} \|\nabla f(x)\|_* > 0$ and L is Lipshitz constant for $\nabla f(x)$ on S. Then Algorithm 1 can be applied for $x^k \in \mathbb{R}^n \setminus S$ with the only change — if $f(x^k) \le 0$, then we have obtained the solution and algorithm stops. Thus we arrive to the method (for ℓ_2 norm):

1. If $f(x^k) \le 0$, then stop, a solution is found;

2. If $\|\nabla f(x)\|_2^2 < Lf(x^k)$, then $x^{k+1} = x^k - \frac{1}{L} f(x^k)\nabla f(x^k)$.
3. Otherwise $x^{k+1} = x^k - \frac{f(x^k)}{\|\nabla f(x^k)\|_2^2}\nabla f(x^k)$, increase k and return to Step 1.

Again Stage 2 is similar to well-known method for solving convex inequalities, the main difference is the necessity of Stage 1 and the lack of convexity assumption. The method globally converges under above formulated assumptions.

Note that solving inequality $f(x) - f^* - \varepsilon \le 0$ is equivalent to minimization of function f with known minima f^* and desired accuracy ε, again without convexity assumption. We remind that convergence rate is quadratic for the algorithms.

5.4 Quadratic equations

Proceed to a specific nonlinear equation, namely the quadratic one. Then the function $g(x)$ may be written componentwise as (4), with gradients

$$\nabla g_i(x) = A_i x + b_i \in \mathbb{R}^n, \;\; i = 1, \dots, m.$$

Obviously $g(0) = 0$, the question is solvability of $g(x) = y$. There are some results on on construction of the entire set of feasible points $Y = \{y : g(x) = y\} = g(\mathbb{R}^n)$, including its convexity, see e.g. [17]. We focus on local solvability, trying to derive the largest ball inscribed in Y.

The derivative matrix $g'(x)$ is formed row-wise as

$$g'(x) = \begin{bmatrix} \nabla g_1(x)^T \\ \vdots \\ \nabla g_m(x)^T \end{bmatrix} = \begin{bmatrix} x^T A_1 + b_1^T \\ \vdots \\ x^T A_m + b_m^T \end{bmatrix} \in \mathbb{R}^{m \times n}.$$

One has $g'(0) = H$, H being $m \times n$ matrix with rows b_i. We suppose H has rank m (recall that $m \le n$), then its smallest singular value $\sigma_m(H) > 0$ serves as μ_0.

The derivative $g'(x)$ is linear on x, meaning it has uniform Lipschitz constant L on $\mathbb{R}^n$, thus assumption **A** holds everywhere. There are several estimates for the Lipschitz constants, for example (for ℓ_2 norm)

$$L \le L_1 = \sqrt{\lambda_{\max}\left(\sum_{i=1}^m A_i^T A_i\right)}$$

from [18], where $\lambda_{\max}$ is the maximal eigenvalue of a matrix. Other estimates can be obtained via elaborate convex semidefinite optimization problem (SDP), cf. [22] for details. We obtain the following consequence of Theorem 2 for Euclidean norms.

Theorem 5 *Suppose that matrix H has rank m, and $\mu_0 > 0$ is its smallest singular value. For quadratic function g equation $g(x) = y$ has a solution $x^*(y)$ for all*

$$\|y\| < \frac{\mu_0^2}{4L}, \tag{22}$$

with $\|x^*(y)\| \leq \frac{\mu_0}{2L}$.

Quadratic equations play significant role in power system analysis, because power flow equations are quadratic, see [10, 11]. It is of interest to compare our estimate (22) with some known results on solvability of power flow equations [2, 24].

As for Algorithms 1–3, we can evaluate bounds explicitly likewise Theorem 2 and Corollary 3.

Theorem 6 *Suppose that matrix H has rank m, and $\mu_0 > 0$ is its smallest singular value. Then Algorithms 1, 3 converge to a solution of* (1) *if*

$$\|y\| \leq s_1 \frac{\mu_0^2}{L}, \quad s_1 \approx 0.1877178,$$

with $\|x^*(y)\| \leq t_1 \frac{\mu_0}{L}$, $\quad t_1 \approx 0.40100511$, where the constants s_1 and r_1 are maxima and maximizer of function $S(t) = 2(1-t)^2 \Delta(\frac{t}{2(1-t)})$, $\ t \in [0, \frac{1}{2}]$.

Proof As $\rho = \infty$, the Theorem 3 is valid for any $r \leq \frac{L}{\mu_0}$ with $\mu(r) = \mu_0 - Lr$. We are to estimate a value of r with maximal allowed bound (12) on $\|P(x^0)\|$.

First assume that $r \leq \frac{\mu_0}{L}\frac{2c}{2c+1} \approx 0.62\frac{\mu_0}{L}$. Then function $S(t)$ is a representation of upper bound (12a) normalized to multiplier $\frac{\mu^2}{L}$ and written via variable $t = \frac{L}{\mu} r$. The function is unimodal with maximum $s_1 = S(t_1)$. Direct check of second case $r \in (\frac{2c}{2c+1}\frac{L}{\mu_0}, \frac{L}{\mu_0}]$ and corresponding bounds (12b), (20) with respect to substitution $\mu \leftarrow \mu(r)$, $\rho \leftarrow r$ reveal lesser than $s_1 \frac{\mu_0^2}{L}$ bound on $\|P(x^0)\|$.

Optimization over approximation (19) instead of (12a) results in smaller constant $s_2 = 5\sqrt{5} - 11 \approx 0.18034$.

We compared all constants and noticed that numerically value s_1 is slightly greater than $\frac{3}{4}$ of maximal range (22) of Theorem 5. As result we propose estimate solvability of (1) in case of quadratic functions by inequality

$$\|y\| \leq \frac{3}{16}\frac{\mu_0^2}{L}.$$

5.5 Solving systems of inequalities

We have discussed above how Newton-like methods can be applied for solving one inequality. Below we address some tricks to convert systems of inequalities into systems of equations.

First, if one seeks a solution of a system of inequalities

$$g_i(x) \le 0,\ i = 1, \dots, m,\ \ x \in \mathbb{R}^\ell,$$

with $m \le 2n$ then by introducing slack variables the problem is reduced to solution of underdetermined system of equations

$$g_i(x) + x_{\ell+i}^2 = 0,\ i = 1, \dots, m,\ \ x \in \mathbb{R}^n, n = \ell + m.$$

Similarly finding a feasible point for linear inequalities $x \ge 0, Ax = b,\ \ x \in \mathbb{R}^n,\ b \in \mathbb{R}^m$ can be transformed to underdetermined system

$$\sum_{j=1}^{n} A_{i,j} z_i^2 = b_i,\ i = 1, \dots, m,\ \ z \in \mathbb{R}^n.$$

The efficiency of such reductions is unclear a priori and should be checked by intensive numerical study.

6 Numerical tests

At present we have numerous results on numerical simulation for various test problems. We plan to present them in a separate publication. Here we restrict ourselves with the single example to demonstrate how the methods work for medium-size problems ($n = 60, m = 21$). The equations have special structure as in Section 5.1:

$$P_i(x) = \phi((c_i, x) - b_i) - y_i, x \in \mathbb{R}^n, y \in \mathbb{R}^m$$

where

$$\phi(t) = \frac{t}{1 + e^{-|t|}},\ \ \phi'(t) = \frac{1 + (1 + |t|)e^{-|t|}}{(1 + e^{-|t|})^2}.$$

Matrix C with rows c_i, vectors b, y were generated randomly. For function $\phi(t)$ we have $\mu_\phi = \max_t \phi'(t) \ge 0.5,\ M = \max_t |\phi''(t)| \le 2$ for all t. Thus if we do not pay attention to the special structure of the problem we have $\mu \ge 0.5\sigma_{\min}(C),\ L \le 2\sigma_{\max}(C)$ and the convergence can be slow, because matrix C can be ill-conditioned. On the other hand if we take into account the structure and replace $\frac{\mu^2}{L}(= 0.0031$ in example) in Algorithm 3 (see Subsection 4.3) with $\frac{\mu_\phi^2}{M} = 0.125$ much faster convergence is achieved. Simulation results on Figure 1 confirm this conclusion. Most iterations are generally spent in first phase of the Algorithm, and the number is close to theoretic bound $N \approx k_{\max}(\beta)$, (13). We also tested adaptive Algorithm 2 on same initial point, and it performed better for bigger initial $\beta_0 = 5$, as shown on Figure 2.

Fig. 1 $\|P(x^k)\|$ of Algorithm 1 using constants μ, L (left) and constants μ_φ, M (right).

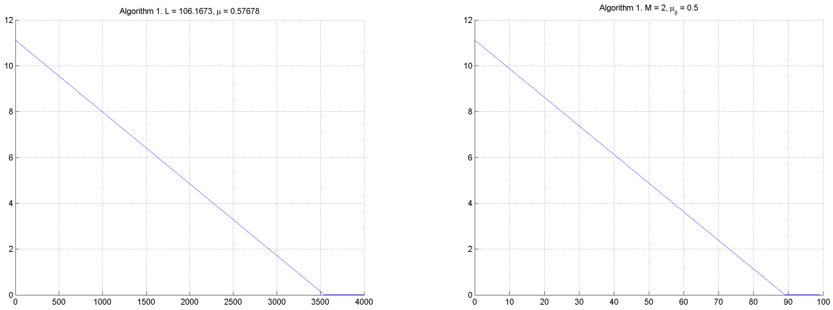
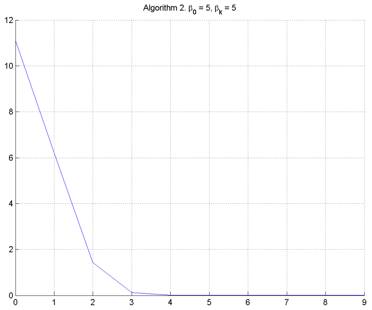


Fig. 2 $\|P(x^k)\|$ of Algorithm 2 using initial approximation $\beta_0 = 5$.



7 Conclusions and future research

New solvability conditions for underdetermined equations (with wider solvability set) are proposed. The algorithms for finding a solution are easy to implement, they combine weaker assumptions on initial approximations and fast convergence rate. No convexity assumptions are required. The algorithms have large flexibility in using prior information, various norms and problem structure. It is worth mentioning that we do not try to convert the problem into optimization one. Combination of damped/pure Newton method is a contribution for solving classic $n = m$ problems as well.

There are numerous directions for future research.

1. We suppose that the auxiliary optimization problem for finding direction z^k is solved exactly. Of course an approximate solution of the sub-problem suffices.
2. The algorithms provide a solution of the initial problem which is not specified apriory. Sometimes we are interested in the solution closest to x^0, i.e. $\min_{P(x)=0} \|x - x^0\|$. An algorithm for this purpose is of interest.
3. More general theory of structured problems (Section 5.1) is needed.

4. It is not obvious how to introduce regularization techniques into the algorithms.

Acknowledgements Authors thank Yuri Nesterov for helpful discussions and references.

References

1. Ben-Israel, A.: A Newton-Raphson method for the solution of systems of equations. J. Mathematical Analysis and Applications. 15, 243–252 (1966)
2. Bolognani, S., Zampieri, S.: On the existence and linear approximation of the power flow solution in power distribution networks. IEEE Transactions on Power Systems. 31(1), 163–172 (2016)
3. Dennis, J.E., Schnabel, R.B.: Numerical Methods for Unconstrained Optimization and Nonlinear Equations. SIAM, Philadelphia (1996)
4. Dontchev, A.L.: The Graves theorem revisited. J. of Convex Analysis. 3(1), 45–53 (1996)
5. Graves, L.M.: Some mapping theorems. Duke Math. J. 17, 111–114 (1950)
6. Kantorovich, L.V.: The method of successive approximations for functional analysis. Acta. Math. 71, 63–97 (1939)
7. Kantorovich, L.V., Akilov, G.P.: Functional Analysis. 2nd ed. Pergamon Press, Oxford (1982)
8. Kelley, C.T.: Solving Nonlinear Equations with Newton's Method. SIAM, Philadelphia (2003)
9. Levin, Y., Ben-Israel, A.: A Newton method for systems of m equations in n variables. Nonlinear Analysis. 47, 1961–1971 (2001)
10. Low, S.H.: Convex relaxation of optimal power flow, part I: Formulations and equivalence. IEEE Trans. on Control of Network Systems. 1(1), 15–27 (2014). Part II: Exactness. ibid. 1(2), 177–189 (2014)
11. Machowski, J., Bialek, J., Bumby, J.: Power System Dynamics. Stability and Control. 2nd ed. John Wiley & Sons Ltd. (2012)
12. Magaril-Il'yaev, G.G., Tikhomirov, V.M.: Newton's method, differential equations and the Lagrangian principle for necessary extremum conditions. Proc. Steklov Inst. Math. 262, 149–169 (2008)
13. Nesterov, Yu., Nemirovskii, A.: Interior-point Polynomial Algorithms in Convex Programming. SIAM, Philadelphia (1994)
14. Nesterov, Yu.: Modified Gauss-Newton scheme with worst case guarantees for global performance. Optimization Methods and Software. 22(3), 469–483 (2007)
15. Ortega, J.M., Rheinboldt, W.C.: Iterative Solution of Nonlinear Equations in Several Variables. SIAM, Philadelphia (2000)
16. Polyak, B.T.: Gradient methods for solving equations and inequalities. USSR Computational Mathematics and Mathematical Phys. 4(6), 17–32 (1964)
17. Polyak, B.T.: Quadratic transformations and their use in optimization. J. of Optimization Theory and Applications. 99(3), 553-583 (1998)
18. Polyak, B.T.: Convexity of nonlinear image of a small ball with applications to optimization. Set-Valued Analysis. 9, 159–168 (2001)
19. Polyak, B.T.: Newton-Kantorovich method and its global convergence. J. Mathematical Sciences. 133(4), 1513–1523 (2006)
20. Prusinska, A., Tret'yakov, A.A.: On the existence of solutions to nonlinear equations involving singular mappings with non-zero p-kernel. Set Valued Analysis. 19, 399–416 (2011)
21. Walker, H.F.: Newton-like methods for underdetermined systems. In: Allgower, E.L., Georg K. (eds.) Computational Solution of Nonlinear Systems of Equations, Lecture Notes in Applied Mathematics. 26, pp. 679–699, AMS, Providence, RI (1990)
22. Xia, Y.: On local convexity of quadratic transformations. J. of the Operations Research Society of China. 8(2), 341–350 (2014)
23. Yamamoto, T.: Historical developments in convergence analysis for Newton's and Newton-like methods. J. Computational Appl. Math. 124, 1–23 (2000)

24. Yu, S., Nguyen, H.D., Turitsyn, K.S.: Simple certificate of solvability of power flow equations for distribution systems. Power & Energy Society General Meeting, IEEE (2015)