Newton’s method and its use in optimization

B.T. Polyak *

Institute for Control Science, Profsojuznaya 65, Moscow 117997, Russia

Received 30 September 2004; accepted 20 June 2005
Available online 25 April 2006

Abstract

Newton’s method is a basic tool in numerical analysis and numerous applications, including operations research and data mining. We survey the history of the method, its main ideas, convergence results, modifications, its global behavior. We focus on applications of the method for various classes of optimization problems, such as unconstrained minimization, equality constrained problems, convex programming and interior point methods. Some extensions (non-smooth problems, continuous analog, Smale’s results, etc.) are discussed briefly, while some others (e.g., versions of the method to achieve global convergence) are addressed in more details.

© 2006 Elsevier B.V. All rights reserved.

Keywords: Nonlinear programming; Newton’s method; Convergence; Global behavior; Interior-point methods

1. Introduction

Newton’s method is one of the fundamental tools in numerical analysis, operations research, optimization and control. It has numerous applications in management science, industrial and financial research, data mining. Its role in optimization cannot be overestimated: the method is the basis for the most effective procedures in linear and nonlinear programming. For instance, the polynomial time interior point algorithms in convex optimization are based on Newton’s method.

The present paper surveys the key ideas of the method in the historical perspective as well as modern investigations and applications. Our goal is to address the most challenging problems in this field of research and to provide suitable references for a reader.

The paper is organized as follows. Section 2 presents the basic idea of Newton’s method and the history of its development. Main mathematical results on the convergence are addressed in Section 3. Newton’s method in its basic form possesses just local convergence, its global behavior and modifications to achieve global convergence are discussed in Sections 4 and 5. The case of underdetermined systems is worth of special consideration (Section 6). In its original form, Newton’s method is destined for solving equations. However, it can be easily tailored for unconstrained (Section 7) and constrained (Section 8) optimization. Some extensions of the method and directions for future research are described in Section 9.
2. Idea and history of the method

The basic idea of Newton’s method is very simple—it is linearization. Suppose $F : \mathbb{R}^1 \rightarrow \mathbb{R}^1$ is a differentiable function, and we are solving the equation

$$F(x) = 0.$$  \hfill (1)

Starting from an initial point $x_0$ we can construct the linear approximation of $F(x)$ in the neighborhood of $x_0$:

$$F(x_0 + h) \approx F(x_0) + F'(x_0)h$$

and solve the arising linear equation $F(x_0) + F'(x_0)h = 0$. Thus, we arrive at recurrent method

$$x_{k+1} = x_k - F'(x_k)^{-1} F(x_k), \quad k = 0, 1, \ldots$$  \hfill (2)

This is the method proposed by Newton in 1669. To be more precise, Newton dealt with polynomials only; in the expression of $F(x + h)$ he discarded higher order terms in $h$. The method was illustrated on the example $F(x) = x^3 - 2x - 5 = 0$. The starting approximation for the root was $x = 2$. Then $F(2 + h) = h^3 + 6h^2 + 10h - 1$, neglecting higher order terms Newton got the linear equation $10h - 1 = 0$. Thus the next approximation is $x = 2 + 0.1 = 2.1$ and the process can be repeated for this point. Fig. 1 demonstrates that the convergence to the root of $F(x)$ is very fast. It was J. Raphson, who proposed in 1690 the general form of method (2) (i.e., $F(x)$ was not assumed to be a polynomial only and the notion of a derivative was exploited), this is why the method is often called the Newton–Raphson method.

The progress in development of the method is linked with such famous mathematicians as Fourier, Cauchy and others. For instance, Fourier proved in 1818 that the method converges quadratically in the neighborhood of a root, while Cauchy (1829, 1847) provided the multidimensional extension of (2) and used the method to prove the existence of a root of an equation. Important early contributions to the investigation of the method are due to Fine [17] and Bennet [3]; their papers are published in one volume of *Proceedings of National Academy of Sciences USA* in 1916. Fine proved the convergence in the $n$-dimensional case with no assumption on the existence of a solution. Bennet extended the result for the infinite-dimensional case; this is a surprising attempt because the foundations of the functional analysis were not created at that moment. In 1948, L.V. Kantorovich published a seminal paper [25] where an extension of Newton’s method for functional spaces was provided (Newton–Kantorovich method). The results were also included in the survey paper [26]; for further developments of the method by Kantorovich, see [27–31].

The basic results on Newton’s method and numerous references can be found in the books by Ostrowski [47], Ortega and Rheinboldt [48], Kantorovich and Akilov [32,33]. A more recent bibliography is available in the books [50,60,13], the
3. Convergence results

Kantorovich [25] analyzes the same equation as (1):
\[ F(x) = 0, \]
but now \( F : X \to Y \), where \( X, Y \) are Banach spaces. The proposed method reads as (2)
\[ x_{k+1} = x_k - F'(x_k)^{-1}F(x_k), \quad k = 0, 1, \ldots, \]
where \( F'(x_k) \) is (Frechet) derivative of the nonlinear operator \( F(x) \) at the point \( x_k \) and \( F'(x_k)^{-1} \) – its inverse. The main convergence result from [25] looks as follows.

**Theorem 1.** Suppose \( F \) is defined and twice continuously differentiable on the ball \( B = \{ x : \| x - x_0 \| \leq r \} \), that the linear operator \( F'(x_0) \) is invertible, \( \| F'(x_0)^{-1}F(x_0) \| \leq \eta, \) \( \| F'(x_0)^{-1}F'(x) \| \leq K, \) \( x \in B \) and
\[ h = Kr < \frac{1}{2}, \quad r \geq 1 - \sqrt{1 - 2h \eta}, \]
Then Eq. (3) has a solution \( x^* \in B \), the process (4) is well defined and converges to \( x^* \) with a quadratic rate:
\[ \| x_k - x^* \| \leq \frac{\eta}{h^2} (2h)^{k^2}. \]

The proof of the Theorem is simple enough; the main novelty of Kantorovich’s contribution were not technicalities, but the general problem formulation and the use of the appropriate techniques of functional analysis. Until Kantorovich’s works [25–27] it was not understood that numerical analysis should be considered in the framework of functional analysis (note the title of [26]: “Functional analysis and applied mathematics”). Another peculiarity of the above theorem is the lack of assumption on the existence of a solution – the theorem is not only a convergence result for a specific method, but simultaneously an existence theorem for nonlinear equations.

These properties of the approach ensured a wide range of applications. Numerous nonlinear problems – nonlinear integral equations, ordinary and partial differential equations, variational problems – could be cast in the framework of (3) and method (4) could be applied. Various examples of such applications are presented in the monographs [32,33]. Moreover, the Newton–Kantorovich method as a tool for existence results was immediately used in the classical works of Kolmogorov, Arnold and Moser (see, e.g., [1]) on KAM-theory in mechanics. Actually many classical results in the functional analysis are proved by use of Newton-like methods; the typical example is Ljusternik’s theorem on tangent spaces, see the original paper [38] or modern research [22].

Later Kantorovich [28,29] obtained another proof of Theorem 1 and its versions, based on the method of majorants. The idea is to compare iterations (4) with scalar iterations, which in a sense majorize them and which possess convergence properties. This approach provides more flexibility; the original proof is a particular case with a quadratic majorant.

There exist numerous versions of Theorem 1, which differ in assumptions and results. We mention just one of them due to Mysovskikh [43].

**Theorem 2.** Suppose \( F \) is defined and twice continuously differentiable on the ball \( B = \{ x : \| x - x_0 \| \leq r \} \), the linear operator \( F'(x) \) is invertible on \( B \) and \( \| F'(x)^{-1} \| \leq \beta, \) \( \| F'(x) \| \leq K, x \in B, \) \( \| F(x_0) \| \leq \eta \) and
\[ h = K\beta^2 \eta < 2, \quad r \geq \beta \eta \sum_{n=0}^{\infty} \left( \frac{h}{2} \right)^n < 1. \]
Then Eq. (3) has a solution \( x^* \in B \) and the process (4) converges to \( x^* \) with quadratic rate:
\[ \| x_k - x^* \| \leq \frac{\beta \eta (h/2)^{k^2}}{1 - (h/2)^{2k^2}}. \]

The difference to Theorem 1 is the assumption on the invertibility of \( F'(x) \) on \( B \) (while in Theorem 1 it was assumed to be invertible in the initial point \( x_0 \) only) and the slighter assumption on \( h \) (\( h < 2 \) instead of \( h < 1/2 \)). Other versions can be found in the books [32,33,47,48,35,10].

Newton’s method (4) requires to calculate derivatives \( F'(x) \) and to solve the linear equation at each iteration; it can be a pretty hard computational burden. The modified Newton method
\[ x_{k+1} = x_k - F'(x_0)^{-1}F'(x_k), \quad k = 0, 1, \ldots \]
avoids this difficulty – the matrix \( F'(x_0) \) is calculated and inverted at the initial point only. The method converges under the same conditions as (4), however the price we pay for simplification is high – method (9) converges linearly, not quadratically. There are
some tradeoffs – for instance, one can update the derivatives after several iterations.

4. Global behavior

The critical condition for convergence is (5) or (7). They mean that at the initial approximation $x_0$ the function $\|F(x_0)\|$ should be small enough, that is $x_0$ should be close to the solution. Thus, Newton’s method is locally convergent. Very simple one-dimensional examples demonstrate the lack of the global convergence even for smooth monotone $F(x)$. There are many ways to modify the method to achieve its global convergence (we will discuss them later), but the problem of interest is the global behavior of iterations. Obviously there are many simple situations – say, there is a neighborhood $S$ of a solution such that $x_0 \in S$ implies convergence to the solution (such a set is called basin of attraction) while trajectories starting outside $Q$ do not converge (e.g., tend to infinity). However, in the case of non-uniqueness of a solution the structure of basins of attractions may be very complicated and exhibit the fractal nature. It was Cayley who formulated this problem as early as 1879.

Let us consider Cayley’s example: solve the equation $z^3 = 1$ by Newton’s method. Thus we take $F(z) = z^3 - 1$ and apply method (2):

$$z_{k+1} = z_k - \frac{z_k^3 - 1}{3z_k^2} = \frac{2z_k}{3} + \frac{1}{3z_k^2}.$$

It is worth to mention that we have formulated Newton’s method in real spaces but it is as well applicable in complex spaces; in the above example we take $z_k \in \mathbb{C}$. The equation has three roots $z^*_1 = 1$, $z^*_2, 3 = -\frac{1}{2} \pm \frac{i\sqrt{3}}{2}$, and it is natural to expect that the entire plane $\mathbb{C}$ is partitioned into three basins of attraction

$$S_m = \{z_0 : z_k \rightarrow z^*_m\}, \quad m = 1, 2, 3$$

located around the corresponding roots. However, the true picture is much more involved. First, there is a single point $z = 0$ where the method is not defined. It has three preimages – points $z_0$ such that $z_1 = 0$, they are $-\rho, \rho(1/2 \pm i\sqrt{3}/2)$, where $\rho = 1/\sqrt{2}$. Again, each of them has three preimages and so on. Thus there are $3^k$ points $z_0$ which are mapped after $k$ iterations into the point $z_k = 0$, and they generate a countable set of such initial points, that the method is not applicable at some iteration

$$S_0 = \{z_0 : z_k = 0 \text{ for some } k\}.$$

It can be proved that for all points $z_0 \notin S_0$ the method converges to one of the solutions (note that if $|z_k| > 1$ then $|z_{k+1}| < |z_k|$) and we have

$$\mathbb{C} = \bigcup_{m=0}^{3} S_m.$$

![Fig. 2. Basin of attraction for $x^* = 1$.](image-url)
The sets $S_m$ have a fractal structure: $S_0$ is the boundary of each $S_m$, $m = 1, 2, 3$ and in any neighborhood of any point $z \in S_0$ there exist points from $S_m$, $m = 1, 2, 3$. The set $S_1$ is displayed in Fig. 2, the set $S_0$ in Fig. 3. The sets of initial points which possess no convergence (like the set $S_0$) for iterations of general rational maps were studied by Julia [24] and are called now Julia sets. A lot of examples linked with Newton’s method can be found in numerous books on fractals [2,39], papers [12,49] and web materials [23,14] (we mention just few of the sources available). Some of these examples exhibit much more complicated behavior than for Cayley’s one. For instance, one can meet periodic or chaotic trajectories of the iterations.

5. Overcoming the local nature of the method

In Cayley’s example Newton’s method converged for almost all initial points (exceptions were the countable number of points $S_0$); the complex structure of basins of attraction was caused by the existence of several roots. However if (1) has a single root, the method usually has local convergence only. For instance, take $F(x) = \arctan x$, which is smooth, monotone and has the single root $x^* = 0$. It is easy to check that (2) converges if and only if $|x_0| < x^*$ ($x^* > 0$ being a root of $2x = (1 + x^2)\arctan x$), for $|x_0| = x^*$ we have periodic behavior of iterations $x_0 = -x_1 = x_2 = -x_3 = \ldots$, while for $|x_0| > x^*$ the iterations diverge: $|x_k| \to \infty$.

There are several ways to modify the basic Newton method to achieve global convergence. The first one is to introduce a regulated step-size to avoid too large steps; this is the so called damped Newton’s method:

$$x_{k+1} = x_k - \alpha_k F'(x_k)^{-1}F(x_k), \quad k = 0, 1, \ldots, \tag{10}$$

where the step-size $0 < \alpha_k \leq 1$ is chosen to achieve monotone decrease of $||F(x)||$, that is condition $||F(x_{k+1})|| < ||F(x_k)||$ holds. We will discuss the particular algorithms for choosing $\alpha_k$ later for minimization problems. The main goal in constructing such algorithms is to preserve the balance between convergence and rate of convergence, i.e., one should take $\alpha_k < 1$ when $x_k$ is beyond a basin of attraction of “pure Newton’s method” and switch to $\alpha_k = 1$ inside this basin.

The second approach is the Levenberg-Marquardt method [36,40]:

$$x_{k+1} = x_k - (\alpha_k I + F'(x_k))^{-1}F(x_k), \quad k = 0, 1, \ldots \tag{11}$$

For $\alpha_k = 0$ the method converts into pure Newton’s method, while for $\alpha_k \gg 1$ it is close to a strongly damped gradient method, which usually converges globally. There are various strategies for adjusting parameters $\alpha_k$, they are described, for instance, in [48]. Method (11) works even when the pure Newton’s method does not – in a situation with the degenerate operator $F'(x_k)$. As we shall see, for
minimization problems method (11) is very promising, because it does not assume the matrix \( F(x_k) \) to be positive definite.

One more strategy is to modify Newton’s method to prevent large steps; it can be done not by choosing step-sizes as in the damped Newton’s method (10), but introducing trust regions, where a linear approximation of \( F(x) \) is valid. Such an approach (originated in [19] and widely developed in [11]) will be discussed later in the connection with optimization problems.

6. Underdetermined systems

In the above analysis of Newton’s method it was assumed that the linear operator \( F(x_k) \) is invertible. However there are situations when this is definitely wrong. For instance suppose that \( F: \mathbb{R}^m \rightarrow \mathbb{R}^m \), where \( m < n \). That is we solve an underdetermined system of \( m \) equations with \( n > m \) variables; of course the rectangular matrix \( F(x_k) \) has no inverse. Nevertheless an extension of Newton’s method for this case can be provided, this is due to Graves [21], who exploited the method to prove existence of solutions of nonlinear mappings. We present the result from [51] where emphasis on the method is done with more accurate estimates.

**Theorem 3.** Suppose \( F: X \rightarrow Y \) is defined and differentiable on the ball \( B = \{ x : \| x - x_0 \| \leq r \} \), its derivative satisfies a Lipschitz condition on \( B \)

\[ ||F'(x) - F'(z)|| \leq L ||x - z||, \quad x, z \in B, \]

\( F(x) \) maps \( X \) onto \( Y \) and the following estimate holds:

\[ ||F'(x)^*y|| \geq \mu \| y \|, \quad \forall y \in Y, \]

with \( \mu > 0 \) (star denotes conjugation). Introduce the function

\[ H_n(t) = \sum_{k=n}^{\infty} t^k \]

and suppose that

\[ h = \frac{L \mu^2 \| F(x_0) \|}{2} < 1, \quad \rho = \frac{2H_0(h)}{L \mu} \leq r. \]  \hspace{1cm} (13)

Then the method

\[ x_{k+1} = x_k - y_k, \quad F'(x_k)y_k = F(x_k), \]

\[ \| y_k \| \leq \frac{1}{\mu} \| F(x_k) \|, \quad k = 0, 1, \ldots \]

is well defined and converges to a solution \( x^* \) of the equation \( F(x) = 0, \| x^* - x_0 \| \leq \rho \) with the rate

\[ \| x_k - x^* \| \leq \frac{2H_k(h)}{L \mu}. \]  \hspace{1cm} (15)

Thus at each iteration of the method one should solve the linear equation \( F'(x_k)y = F(x_k) \), where the linear operator \( F'(x_k) \) in general does not have the inverse; however it maps \( X \) onto \( Y \) and this equation has a solution (may be not a single one). Among the solutions there exists the solution \( y_k \) with the property \( \| y_k \| \leq (1/\mu) \| F(x_k) \| \); this one is exploited in the method. In the finite dimensional case \( (X = \mathbb{R}^n, \ Y = \mathbb{R}^m, n > m) \) such solution is provided by the formula \( y_k = F(x_k)^+ F(x_k) \), where \( A^+ \) denotes the pseudoinverse of the matrix \( A \).

We consider an application of **Theorem 3** to convex analysis. The following result on convexity of the nonlinear image of a small ball in Hilbert spaces holds [52].

**Theorem 4.** Suppose \( X, Y \) are Hilbert spaces, \( F: X \rightarrow Y \) is defined and differentiable on the ball \( B = \{ x : \| x - a \| \leq r \} \), its derivative satisfies a Lipschitz condition on \( B \)

\[ ||F'(x) - F'(z)|| \leq L ||x - z||, \quad x, z \in B, \]

\( F'(a) \) maps \( X \) onto \( Y \) and the following estimate holds:

\[ ||F'(a)^*y|| \geq \mu \| y \|, \quad \forall y \in Y, \]

with \( \mu > 0 \) and \( r < \mu(2L) \). Then the image of the ball \( B \) under the map \( F \) is convex, i.e., the set \( S = \{ F(x) : x \in B \} \) is a convex set in \( Y \).

This theorem has numerous applications in optimization [52], linear algebra [53], optimal control [54]. For instance, the pseudospectrum of a \( n \times n \) matrix (a set of all eigenvalues of perturbed matrices for perturbations bounded in Frobenius norm) happens to be the union of \( n \) convex sets in the complex plane provided that the nominal matrix has all eigenvalues distinct and perturbations are small enough. Another result is the convexity of the reachable set of a nonlinear system for \( L_2 \) bounded controls [54].

7. Unconstrained optimization

Consider the simplest unconstrained minimization problem in a Hilbert space \( H \):

\[ \min f(x), \quad x \in H. \]  \hspace{1cm} (17)

Assuming that \( f \) is twice differentiable, we can get Newton’s method for minimization by two different approaches.
First, the necessary (and sufficient for $f$ convex) condition for minimization is Fermat’s condition
$$\nabla f(x) = 0,$$
that is, we should solve Eq. (3) with $F(x) = \nabla f(x)$. Applying Newton’s method for this equation we arrive at Newton’s method for minimization:
$$x_{k+1} = x_k - (\nabla^2 f(x_k))^{-1}\nabla f(x_k), \quad k = 0, 1, \ldots,$$
where $\nabla^2 f$ denotes the second Frechet derivative (the Hessian matrix in a finite dimensional case).

Second, we can approximate $f(x)$ in the neighborhood of a point $x_k$ by three terms of its Taylor series:
$$f(x_k + h) \approx f_k(h) = f(x_k) + (\nabla f(x_k), h) + \frac{1}{2}(\nabla^2 f(x_k)h, h).$$

Then, minimizing the quadratic function $f_k(h)$ we get the same method (18). Both these interpretations would be exploited for constructing Newton methods for more general optimization problems.

The theorems on convergence of Newton’s method for equations can be immediately adopted to the unconstrained minimization case (just replace $F(x)$ by $F(x)$ and $F'(x)$ by $\nabla^2 f(x)$). The main feature of the method – fast local convergence – remains unchanged. However, there are some specific properties. The most important: Newton’s method in its pure form does not distinguish minima, maxima and saddle points – when started from a neighborhood of a nonsingular critical point (i.e., a point $x^*$ with $F(x^*) = 0$, $\nabla^2 f(x^*)$ invertible) the method converges to it, making no difference between minimum, maximum or saddle points.

There are numerous ways to convert the method into globally convergent one, they are similar to modifications discussed in Section 5 (damped Newton, Levenberg-Marquardt, trust region). We consider an important version due to Nesterov and Nemirovski [44], in this version its complexity (the number of iterations to achieve the desired accuracy) can be estimated. Nesterov and Nemirovski introduce the class of self-concordant functions. These are three times differentiable convex functions defined on a convex set $D \subset \mathbb{R}^n$, satisfying the property
$$|\nabla^3 f(x)(h, h, h)| \leq 2(\nabla^2 f(x)h, h)^{3/2},$$
$$\forall x \in D, \ h \in \mathbb{R}^n.$$

The above formula includes the third and the second derivatives of $f$ and their action on a vector $h \in \mathbb{R}^n$. In simpler form it can be presented as the relation between the third and the second derivatives of the scalar function $\varphi(t) = f(x + th)$:
$$|\varphi''(0)| \leq 2(\varphi''(0))^{3/2}$$
for all $x \in D, h \in \mathbb{R}^n$. For instance, the function $f(x) = -\ln x, x > 0, x \in \mathbb{R}^1$ is self-concordant, while the function $f(x) = 1/x, x > 0$ is not. For $x_k \in D$ define Newton’s decrement
$$\delta_k = \sqrt{(\nabla f(x_k))^T(\nabla^2 f(x_k))^{-1}\nabla f(x_k)}.$$

Now the Nesterov–Nemirovski version of the damped Newton’s method reads
$$x_{k+1} = x_k - \alpha_k(\nabla^2 f(x_k))^{-1}\nabla f(x_k), \quad k = 0, 1, \ldots,$$
$$x_k = 1, \quad \text{if } \delta_k \leq \frac{1}{4},$$
$$x_k = \frac{1}{1 + \delta_k}, \quad \text{if } \delta_k > \frac{1}{4}. \quad (21)$$

**Theorem 5.** If $f(x)$ is self-concordant and $f(x) \geq f^*$, $\forall x \in D$ then for the method (19)–(21) with $x_0 \in D, \varepsilon > 0$ one has
$$f(x_k) - f^* \leq \varepsilon$$
for
$$k = c_1 + c_2 \log \log \left(\frac{1}{\varepsilon}\right) + c_3(f(x_0) - f^*). \quad (22)$$

Here $c_1, c_2, c_3$ are some absolute constants.

The idea of the proof is simple enough. The minimization procedure consists of two stages. At Stage 1, method (19), (21) is applied, and the function decreases monotonically: $f(x_{k+1}) \leq f(x_k) - \gamma$, where $\gamma$ is a positive constant. Obviously, this stage terminates after a finite number of iterations (because $f(x)$ is bounded from below). At stage 2, the pure Newton’s method (19), (20) works and it converges with the quadratic rate: $2\delta_{k+1} \leq (2\delta_k)^2$. Note that the Newton decrement provides a convenient tool to express the rate of convergence: there are no constants like $L, K, \eta$ in Theorems 1–3.

For equality (22) simple values of the constants can be determined. Note that $\log \log (1/\varepsilon)$ is not large even if $\varepsilon$ is small enough; for all reasonable $\varepsilon$ the following estimate holds
$k = 5 + 11(f(x_0) - f^*)$.

However, numerous simulation results for various types of optimization problems of different dimensions [8] validate the following empirical formula for the number of steps for method (19)–(21):

$$k = 5 + 0.6(f(x_0) - f^*).$$

The serious restriction of the above analysis was the convexity assumption. Recently [45] another version of Newton’s method has been proposed, where global convergence and complexity results were obtained for general smooth functions. Suppose that $f \in C^{2,1}$, that is $f$ is twice differentiable on $\mathbb{R}^n$ and its second derivative satisfies a Lipschitz condition with constant $L$ on the set $\{x : f(x) \leq f(x_0)\}$. The following version of Newton’s method is considered in [45]:

$$x_{k+1} = \arg \min_{x \in \mathbb{R}^n} f_k(x), \quad h = x - x_k,$$

$$f_k(x) = f(x_k) + (\nabla f(x_k), h) + \frac{1}{2}(\nabla^2 f(x_k)h, h) + \frac{L}{6}\|h\|^3. \quad (23)$$

Thus at each iteration we solve the unconstrained minimization problem with the same quadratic term as in the pure Newton’s method, but regularized via the cubic term. This problem looks hard and non-convex, however it can be reduced to one-dimensional convex optimization (see details in [45]). Surprisingly, the proposed method has many advantages if compared with pure Newton. Firstly, it converges globally, for an arbitrary initial point. Secondly, it does not converge to maximum points or saddle points in contrast to Newton’s method. Thirdly, its complexity for various classes of functions can be estimated. There are implementable versions of method (23), where the value $L$ is not known apriori.

8. Constrained optimization

We start with some particular cases of constrained optimization problems.

The first one is optimization subject to simple constraints:

$$\min f(x), \quad x \in Q, \quad (24)$$

where the set $Q$ in a Hilbert space $H$ is “simple” in the sense that (24) with a quadratic function $f$ can be solved explicitly. For instance, $Q$ may be a ball, a linear subspace, etc. The extension of Newton’s method for this case is based on the second interpretation of the method for unconstrained minimization:

$$x_{k+1} = \arg \min_{x \in Q} f_k(x),$$

$$f_k(x) = (\nabla f(x_k), x - x_k) + \frac{1}{2}(\nabla^2 f(x_k)(x - x_k), (x - x_k)). \quad (25)$$

The method converges under the same assumptions as for the unconstrained case: if $f$ is convex and twice differentiable with Lipschitz second derivatives on $Q$, $Q$ is closed convex, $f$ attains its minimum on $Q$ in a point $x^*$, $\nabla^2 f(x^*) > 0$, then sequence (25) converges locally to $x^*$ with a quadratic rate. This result has been obtained in [37], see more details and examples in [55,6].

Another simple situation is equality constrained optimization:

$$\min f(x), \quad g(x) = 0, \quad (26)$$

where $f : X \rightarrow \mathbb{R}^1$, $g : X \rightarrow Y$; $X$, $Y$ are Hilbert spaces. If a solution $x^*$ of the problem exists and is a regular point ($g(x^*)$ maps $X$ onto $Y$), then there exists a Lagrange multiplier $y^*$ such that the pair $x^*$, $y^*$ is a stationary point of the Lagrangian

$$L(x, y) = f(x) + (y, g(x))$$

that is the solution of the nonlinear equation

$$L_x(x, y) = 0, \quad L_y(x, y) = 0. \quad (27)$$

Hence, we can apply Newton’s method for solving this equation. Under natural assumptions it converges locally to $x^*$, $y^*$, see rigorous results in [56,6,5]. Various implementations of the method can be also found in these references.

There are other versions of Newton’s method for solving (26) which do not involve the dual variables $y$. Historically, the first application of the Newton-like method for finding the largest eigenvalue of a matrix $A = A^T$ by reducing it to the constrained optimization problem:

$$\max(Ax, x), \quad \|x\|^2 = 1$$

is due to Rayleigh (1899). Later Kantorovich [27] has proposed a pure Newton method for the above problem. Another simple situation, where calculations can be simplified, is the case of linear constraints $g(x) = Cx - d$. Then the method is equivalent to Newton’s method for unconstrained minimization of the restriction of $f$ to the affine subspace $Cx = d$.

Now we proceed to applications of Newton’s method for convex constrained optimization.
problems – the area, where the method plays the key role in constructing the most effective optimization algorithms. The basic scheme of interior-point methods looks as follows [44,66]. For the convex optimization problem
\[
\min f(x), \quad x \in Q, 
\]
with a convex self-concordant \( f \) and a convex \( Q \in \mathbb{R}^n \) we construct a self-concordant barrier \( F(x) \), defined on the interior of \( Q \) and growing to infinity when a point approaches the boundary of \( Q \):
\[
F: \text{int} Q \to \mathbb{R}, \quad F(x) \to \infty \text{ for } x \to \partial Q.
\]
Such barriers exist for numerous examples of the constraints, for instance, if \( Q = \{ x : x \geq 0 \} \), then the logarithmic barrier has the desired properties:
\[
F(x) = \sum_{i=1}^n - \ln x_i.
\]
Using the barriers, we take the function
\[
f_k(x) = t_k f(x) + F(x)
\]
depending on the parameter \( t_k > 0 \). Under natural assumptions it can be proved that \( f_k(x) \) has a minimum point \( x_k^* \) on \( \text{int} Q \) and \( f(x_k^*) \to f^* \) (the minimal value in (28)) when \( t_k \to \infty \) (this is so called central path). However, there is no need to obtain precise values of \( x_k^* \), it suffices to make one step of the damped Newton method (19)–(21) and then to vary \( t_k \). There exists such way to adjust the parameters \( a_k, t_k \) that the method yields polynomial-time complexity, see details and rigorous results in [44,46,4,8]. The theoretical result on polynomial-time complexity is very important, however the practical simulation results on implementation of interior-point methods are not less important. They demonstrate very high effectiveness of the method for a broad spectrum of convex optimization problems – from linear programming to semi-definite programming. For instance, the method is the successful competitor to the classical simplex method for linear programs.

Similar ideas are exploited for non-convex constrained optimization problems, where so called sequential quadratic programming (SQP) is the basis for efficient algorithms, see, e.g., [11,20,7,66,61,34].

9. Some extensions

Newton’s method has numerous extensions, we consider below just few of them.

- **Relaxed smoothness assumptions.** Newton’s method can be applied in many situations where equations (or functions to be minimized) are not smooth enough. The simplest example is solving Eq. (1) with piece-wise linear non-smooth convex function \( F \). Then, if a solution exists, method (2) (naturally extended) finds it after a finite number of iterations.

For the general case there exist numerous non-smooth versions of Newton’s method, beginning with the concept of semi-smoothness by Mifflin [42]. Many researchers have contributed to the area, see [9,34,57,58,62] and references therein; the book [34] contains many basic ideas and historical remarks. A related approach is based on smoothing the equations, consult [18,9].

- **Multiple roots.** We analyzed Newton’s method in a neighborhood of a simple root \( x^* \). In the case of a multiple root Newton’s method either remains convergent, but looses its fast rate of convergence, or diverges. There is a modification of the method, which preserves quadratic convergence; it is due to Schroder (1870). We provide it for the one-dimensional case (1):
\[
x_{k+1} = x_k - pF'(x_k)^{-1}F(x_k), \quad k = 0, 1, \ldots , \quad (29)
\]
here \( p \) is the multiplicity of the root. Unfortunately, we should know in advance this multiplicity, moreover the situation in the multi-dimensional case can be much more complicated.

- **Higher-order methods.** The local rate of convergence of Newton’s method is fast enough; however, some researchers construct methods with still higher rate of convergence. This problem looks a bit artificial (actually very few iterations of Newton’s method are required to obtain high precision when we achieve its convergence domain), so there is no need to accelerate the method.

- **Continuous version.** Instead of discrete iterations in Newton’s method (4) one can apply its continuous analog
\[
\dot{x}(t) = -F'(x(t))^{-1}F(x(t)). \quad (30)
\]
The simplest difference approximation of (30) leads to the damped Newton method (10) with constant step-size \( \varepsilon_k = \varepsilon \). As we know, the damped Newton method can exhibit global convergence, thus one can expect the same property for the continuous version. Indeed, the global convergence for (30) has been validated by Smale [63], and its rate of convergence is analyzed in
[59]. Of course, the numerical value of continuous methods is arguable, because their implementation demands their discretization.

• **Data at one point.** All results on the convergence of Newton’s method include assumptions, which are valid in some neighborhood of the solution or in some prescribed ball. In contrast, Smale [64] provides a convergence theorem based on data available at the single initial point. However, these data include bounds for all derivatives.

• **Solving complementarity and equilibrium problems.** There are many works where Newton’s method is applied to problems, which cannot be casted into equation solving, the typical examples are complementarity, variational inequalities and equilibrium-type problems. For references see, e.g., [16,34].

• **Implementation issues.** We are unable to discuss implementation issues of various versions of Newton’s method, which are indeed important for their practical application. Many details can be found in the recent book [13], while codes of corresponding algorithms can be downloaded from [15]. For convex optimization problems such issues are discussed in [8].

• **Complexity.** There exist very deep results on the complexity of basic problems of numerical analysis (e.g., finding all roots of a polynomial), closely related to Newton’s method (some modification of the method is usually proved to achieve the best possible result). The interested reader can consult [65].

References


[29] L.V. Kantorovich, Some further applications of principle of majorants, Doklady AN SSSR 80 (6) (1951) 849–852.
