Kantorovich's Majorants Principle for Newton's Method

O. P. Ferreira* B. F. Svaiter†

January 17, 2006

Abstract

We prove Kantorovich's theorem on Newton's method using a convergence analysis which makes clear, with respect to Newton's Method, the relationship of the majorant function and the non-linear operator under consideration. This approach enable us to drop out the assumption of existence of a second root for the majorant function, still guaranteeing Q-quadratic convergence rate and to obtain a new estimate of this rate based on a directional derivative of *the derivative* of the majorant function. Moreover, the majorant function does not have to be defined beyond its first root for obtaining convergence rate results.
AMSC: 49M15, 90C30.

1 Introduction

Kantorovich's Theorem on Newton's Method guarantee convergence of that method to a solution using semi-local conditions. It does not require *a priori* existence of a solution, proving instead the existence of the solution and its uniqueness on some region[5]. For a current historical perspective, see [7].

Recently, Kantorovich's Theorem has been the subject of many new research, see [2, 4, 8, 10]. It has been improved by relaxing the (original) assumption of Lipschitz continuity of the derivative of the non-linear functional in question, see [1, 4, 10] and its references. These

*IME/UFG, Campus II- Caixa Postal 131, CEP 74001-970 - Goiânia, GO, Brazil (E-mail:orizon@mat.ufg.br). The author was supported in part by FUNAPE/UFG, PADCT-CNPq, PRONEX–Optimization(FAPERJ/CNPq) and IMPA.

†IMPA, Estrada Dona Castorina, 110, Jardim Botânico, CEP 22460-320 - Rio de Janeiro, RJ, Brazil (E-mail:benar@impa.br). The author was supported in part by CNPq Grant 301200/93-9(RN) and by PRONEX–Optimization(FAPERJ/CNPq).

new versions of Kantorovich's Theorem has also been used to prove many particular results on Newton Method, previously unrelated, see [2, 10].

The aim of this paper is twofold. We shall present a new convergence analysis for Kantorovich's Theorem which makes clear, with respect to Newton's Method, the relationship of the majorant function and the non-linear operator under consideration. Instead of looking only to the generated sequence, we identify regions where Newton Method, for the nonlinear ptoblem, is well behaved, as compared with Newton Method applyed to the majorant function. This new analysis was introduced in [3], to generalize Kantorovich's Theorem on Newton's Method to Riemannian manifolds, and has been also used by [2] in the same context. Now, in a simpler context, this convergence analysis allow us relax the assumptions for guaranteeing Q-quadratic convergence of the method, and obtain a new estimate of the Q-quadratic convergence, based on a directional derivative of *the derivative* of the majorant function. We drop out the assumption of existence of a second root for the majorant function, still guaranteeing Q-quadratic convergence. Moreover, the majorant function even don't need to be defined beyond its first root.

The organization of our paper is as follows. In Subsection 1.1, we list some notations and auxiliary results used in our presentation. In Section 2 the main result is stated and proved and we given some remarks about applications of this result in Section 3.

1.1 Notation and auxiliary results

The following notation is used throughout our presentation. Let X, Y be a Banach spaces. The open and closed ball at x are denoted, respectively by

$$B(x,r) = \{y \in X;\ \|x-y\| < r\} \quad \text{and} \quad B[x,r] = \{y \in X;\ \|x-y\| \leqslant r\}.$$

The following auxiliary results of elementary convex analysis will be needed:

Proposition 1. *Let $I \subset \mathbb{R}$ be an interval, and $\varphi : I \to \mathbb{R}$ be convex.*

1. *For any $u_0 \in \operatorname{int}(I)$, the application*

$$u \mapsto \frac{\varphi(u_0) - \varphi(u)}{u_0 - u}, \qquad u \in I, u \neq u_0,$$

is increasing and there exist (in $\mathbb{R}$)

$$D^-\varphi(u_0) = \lim_{u \to u_0^-} \frac{\varphi(u_0) - \varphi(u)}{u_0 - u} = \sup_{u<u_0} \frac{\varphi(u_0) - \varphi(u)}{u_0 - u}.$$

2. *If $u, v, w \in I$, $u < w$, and $u \le v \le w$ then*

$$\varphi(v) - \varphi(u) \le [\varphi(w) - \varphi(u)] \frac{v-u}{w-u}.$$

2 Kantorovich's Theorem

Our goal is to states and prove the Kantorovich's theorem on Newton's method. The first things that we will do is to prove that this theorem holds for a real majorant function. Then, we will prove well definedness of Newton's Method and convergence, also uniqueness in the suitable region and convergence rates will be established. The statement of the theorem is:

Theorem 2. *Let X be a Banach space, $C \subseteq X$ and $F : C \to Y$ a continuous function, continuously differentiable on int(C). Take $x_0 \in \mathrm{int}(C)$ with $F'(x_0)$ non-singular. Suppose that there exist $R > 0$ and a continuously differentiable function $f : [0, R) \to \mathbb{R}$ such that, $B(x_0, R) \subseteq C$,*

$$\|F'(x_0)^{-1}\,[F'(y) - F'(x)]\,\| \leq f'(\|y - x\| + \|x - x_0\|) - f'(\|x - x_0\|), \qquad (1)$$

for $x, y \in B(x_0, R)$, $\|x - x_0\| + \|y - x\| < R$,

$$\|F'(x_0)^{-1}F(x_0)\| \leq f(0)\,, \qquad (2)$$

and

h1) *$f(0) > 0$, $f'(0) = -1$;*

h2) *f' is convex and strictly increasing;*

h3) *$f(t) = 0$ for some $t \in (0, R)$.*

Then f has a smallest zero $t_* \in (0, R)$, the sequences generated by Newton's Method for solving $f(t) = 0$ and $F(x) = 0$ with starting point $t_0 = 0$ and x_0, respectively,

$$t_{k+1} = t_k - f'(t_k)^{-1}f(t_k), \qquad x_{k+1} = x_k - F'(x_k)^{-1}F(x_k), \qquad k = 0, 1, \ldots\,, \qquad (3)$$

are well defined, $\{t_k\}$ is strictly increasing, is contained in $[0, t_*)$, and converges to t_*, $\{x_k\}$ is contained in $B(x_0, t_*)$ and converges to a point $x_* \in B[x_0, t_*]$ which is the unique zero of F in $B[x_0, t_*]$,

$$\|x_* - x_k\| \leq |t_* - t_k|, \qquad \|x_* - x_{k+1}\| \leq \frac{t_* - t_{k+1}}{(t_* - t_k)^2}\,\|x_* - x_k\|^2, \qquad k = 0, 1, \ldots\,, \qquad (4)$$

and the sequences $\{t_k\}$ and $\{x_k\}$ converge Q-linearly as follows

$$\|x_* - x_{k+1}\| \leq \frac{1}{2}\|x_* - x_k\|, \qquad t_* - t_{k+1} \leq \frac{1}{2}(t_* - t_k) \qquad k = 0, 1, \ldots\,. \qquad (5)$$

If, additionally,

h4) $f'(t_*) < 0$,

then the sequences $\{t_k\}$ *and* $\{x_k\}$ *converge Q-quadratically as follows*

$$\|x_* - x_{k+1}\| \leqslant \frac{D^- f'(t_*)}{-2f'(t_*)}\|x_* - x_k\|^2, \qquad t_* - t_{k+1} \leq \frac{D^- f'(t_*)}{-2f'(t_*)}(t_* - t_k)^2, \quad k = 0, 1, \dots\,, \tag{6}$$

and x_* *is the unique zero of* F *in* $B(x_0, \bar{\tau})$, *where* $\bar{\tau} > t_*$ *is defined as*

$$\bar{\tau} = \sup\{t \in [t_*, R) \ : \ f(t) \leq 0\}.$$

Remark 1. *Under Theorem's 2 assumptions* **h1**-**h3** *on* $f : [0, R) \to \mathbb{R}$,

1. $f(t) = 0$ *has at most one root on* (t_*, R);
2. *condition* **h4** *is implied by any one of the following alternative conditions on* f:

   h4-a) $f(t_{**}) = 0$ *for some* $t_{**} \in (t_*, R)$,

   h4-b) $f(t) < 0$ *for some* $t \in (t_*, R)$,

where t_* *is the smallest root of* f *in* $[0, R)$.

In the usual versions of Kantorovich's Theorem, to guarantee R-quadratic convergence of the sequence $\{x_k\}$ and $\{t_k\}$, condition **h4-a** is used. As we discussed, this condition is more restrictive that condition **h4**.

From now on, we assume that the hypotheses of Theorem 2 hold, with the exception of **h4**, which will be considered to hold only when explicitly stated.

2.1 Newton's Method applied to the majorant function

In this subsection we will study the majorant function f and prove all results regarding only the sequence $\{t_k\}$.

Proposition 3. *The function* f *has a smallest root* $t_* \in (0, R)$, *is strictly convex, and*

$$f(t) > 0, \quad f'(t) < 0, \qquad t < t - f(t)/f'(t) < t_*, \qquad\qquad \forall\, t \in [0, t^*). \tag{7}$$

Moreover, $f'(t_*) \leq 0$ *and*

$$f'(t_*) < 0 \iff \exists\, t \in (t_*, R),\ f(t) \leq 0. \tag{8}$$

Proof. As f is continuous in $[0, R)$ and have a zero there(**h3**), it must have a smallest zero t_*, which is greater than 0 because $f(0) > 0$ (**h1**). Since f' is strictly increasing(**h2**), f is strictly convex.

The first inequality in (7) follows from the assumption $f(0) > 0$ and the definition of t_* as the smallest root of f. Since f is strictly convex,

$$0 = f(t_*) > f(t) + f'(t)(t_* - t), \qquad t \in [0, R),\ t \neq t_*. \tag{9}$$

If $t \in [0, t_*)$ then $f(t) > 0$ and $t_* - t > 0$, which, combined with (9) yields the second inequality in (7). The third inequality in (7) follows form the first and the second ones. The last inequality in (7) is obtained by division of the inequality on (9) by $-f'(t)$ (which is strictly positive) and direct algebraic manipulations of the resulting inequality.

As $f > 0$ in $[0, t_*)$ and $f(t_*) = 0$, we must have $f'(t_*) \leq 0$. In (8), the implication $\Rightarrow$ holds trivially. To prove the implication $\Leftarrow$, interchange t and t_* in (9) and note that $f(t_*) = 0$. $\square$

In view of the first inequality in (7), Newton iteration is well defined in $[0, t_*)$. Let us call it

$$\begin{array}{rcl} n_f : [0, t_*) & \to & \mathbb{R} \\ t & \mapsto & t - f(t)/f'(t). \end{array} \tag{10}$$

Proposition 4. *Newton iteration n_f is strictly increasing, maps $[0, t^*)$ in $[0, t^*)$, and*

$$t_* - n_f(t) \leqslant \frac{1}{2}(t_* - t), \qquad \forall\, t \in [0, t_*). \tag{11}$$

If f also satisfies **h4**, *i.e., $f'(t_*) < 0$, then*

$$t_* - n_f(t) \leq \frac{D^- f'(t_*)}{-2f'(t_*)}(t_* - t)^2, \qquad \forall\, t \in [0, t_*). \tag{12}$$

Proof. The first two statements of the proposition follows trivially for the last inequalities in (7).

To prove (11) take some $t \in [0, t_*)$. Note that $f(t_*) = 0$ (Prop. 3). Using also (10) and the continuity of f' we have

$$\begin{aligned} t_* - n_f(t) &= \frac{1}{f'(t)}\left[f'(t)(t_* - t) + f(t)\right] \\ &= \frac{1}{f'(t)}\left[f'(t)(t_* - t) + f(t) - f(t_*)\right] = \frac{1}{-f'(t)} \int_t^{t_*} f'(u) - f'(t)\, du. \end{aligned}$$

As f' is convex and $t < t_*$, it follows from Proposition 1 that

$$f'(u) - f'(t) \leqslant [f'(t_*) - f'(t)]\,\frac{u-t}{t_* - t}, \qquad \forall\, u \in [t, t^*].$$

Taking in account the positivity of $-1/f'(t)$ (second inequality in (7)) and combining the two above equations we have

$$t_* - n_f(t) \leqslant (-1/f'(t)) \int_t^{t_*} [f'(t_*) - f'(t)]\,\frac{u-t}{t_* - t}\, du.$$

Direct integration of the last term of the above equation yields

$$t_* - n_f(t) \leqslant \frac{1}{2}\left(\frac{f'(t_*) - f'(t)}{-f'(t)}\right)(t_* - t). \tag{13}$$

Therefore, above inequality together $f'(x_*) \leq 0$ and $f'(t) < 0$ imply (11).

Finally, we assume that f satisfies assumption **h4**. Take $t \in [0, t_*)$. As f' is increasing, $f'(x_*) \leq 0$, and $f'(t) < 0$, we obtain

$$\frac{f'(t_*) - f'(t)}{-f'(t)} \leq \frac{f'(t_*) - f'(t)}{-f'(t_*)} = \frac{1}{-f'(t_*)}\,\frac{f'(t_*) - f'(t)}{t_* - t}\,(t_* - t) \leq \frac{D^- f'(t_*)}{-f'(t_*)}\,(t_* - t),$$

where the last inequality follows from Proposition 1. Combining the above inequality with (13) we conclude that (12) holds. □

The definition of $\{t_k\}$ in Theorem 2 is equivalent to the following one

$$t_0 = 0, \quad t_{k+1} = n_f(t_k), \qquad k = 0, 1, \ldots . \tag{14}$$

Therefore, using also Proposition 4 we conclude that

Corollary 5. *The sequence $\{t_k\}$ is well defined, strictly increasing and is contained in $[0, t_*)$. Moreover, it satisfies* (5) *(second inequality) and converges Q-linearly to t_*.*

If f also satisfies assumption **h4**, *then $\{t_k\}$ satisfies the second inequality in* (6) *and converges Q-quadratically.*

So, all statements involving only $\{t_k\}$ on Theorem 2 are valid.

2.2 Convergence

In this subsection we will prove well definedness and convergence of the sequence $\{x_k\}$ specified on (3) in Theorem 2, i.e., the sequence generated by Newton's Method to solve $F(x) = 0$ with the starting point x_0.

Proposition 6. *If* $\|x - x_0\| \leq t < t_*$ *then* $F'(x)$ *is non-singular and*

$$\|F'(x)^{-1}F'(x_0)\| \leqslant -1/f'(t).$$

In particular, F' *is non-singular in* $B(x_0, t_*)$.

Proof. Take $x \in B[x_0, t]$, $0 \leq t < t_*$. Using the assumptions (1), **h2**, **h1** of Theorem 2 and the second inequality of (7) in Proposition 3 we obtain

$$\begin{aligned} \|F'(x_0)^{-1}F'(x) - I\| = \|F'(x_0)^{-1}[F'(x) - F'(x_0)]\| &\leqslant f'(\|x - x_0\|) - f'(0) \\ &\leq f'(t) - f'(0) \\ &= f'(t) + 1 < 1. \end{aligned}$$

Using Banach's Lemma and the above equation we conclude that $F'(x_0)^{-1}F'(x)$ is non-singular and

$$\|F'(x)^{-1}F'(x_0)\| \leqslant \frac{1}{1 - (f'(t) + 1)} = \frac{1}{-f'(t)}.$$

Finally, as $F'(x_0)^{-1}F'(x)$ is non-singular, $F'(x)$ is also non-singular. □

Newton iteration at a point happens to be a zero of the linearization of F at such point, which is also the first-order Taylor expansion of F. So, we study the error in the linearization of F at point in $B(x_0, t)$

$$E(x, y) := F(y) - [F(x) + F'(x)(y - x)], \qquad y \in C, x \in B(x_0, R). \tag{15}$$

We will bound this error by the error in the linearization on the majorant function f.

$$e(t, v) := f(v) - [f(t) + f'(t)(v - t)], \qquad t, v \in [0, R). \tag{16}$$

Lemma 7. *Take*

$$x, y \in B(x_0, R) \quad \text{and} \quad 0 \leq t < v < R.$$

If $\|x - x_0\| \leqslant t$ *and* $\|y - x\| \leqslant v - t$, *then*

$$\|F'(x_0)^{-1}E(x, y)\| \leqslant e(t, v)\frac{\|y - x\|^2}{(v - t)^2}.$$

Proof. As $x, y \in B(x_0, R)$ and the ball is convex

$$x + u(y - x) \in B(x_0, R) \quad \text{for} \quad 0 \le u \le 1.$$

Hence, F being continuously differentiable in $B(x_0, R)$, (15) is equivalent to

$$E(x, y) = \int_0^1 [F'(x + u(y - x)) - F'(x)](y - x)\, du,$$

which, combined with assumption (1) in Theorem 2 gives

$$\begin{aligned}
\|F'(x_0)^{-1}E(x, y)\| &\le \int_0^1 \left\|F'(x_0)^{-1}[F'(x + u(y - x)) - F'(x)]\right\| \, \|y - x\| \; du \\
&\le \int_0^1 [f'(\|x - x_0\| + u\,\|y - x\|) - f'(\|x - x_0\|)]\, \|y - x\|\, du\,.
\end{aligned}$$

Now, using the convexity of f', the hypothesis $\|x - x_0\| < t$, $\|y - x\| < v - t$, $v < R$ and Proposition 1 we have, for any $u \in [0, 1]$

$$\begin{aligned}
f'(\|x - x_0\| + u\|y - x\|) - f'(\|x - x_0\|) &\le f'(t + u\|y - x\|) - f'(t) \\
&\le [f'(t + u(v - t)) - f'(t)]\frac{\|y - x\|}{v - t}\,.
\end{aligned}$$

Combining the two above equations we obtain

$$\|F'(x_0)^{-1}E(x, y)\| \le \int_0^1 [f'(t + u(v - t)) - f'(t)]\frac{\|y - x\|^2}{v - t}\, du,$$

which, after performing the integration yields the desired result. □

Proposition 6 guarantee non-singularity of F', and so well definedness of Newton iteration map for solving $F(x) = 0$, in $B(x_0, t_*)$. Let us call N_F the Newton iteration map (for F) in that region

$$\begin{array}{rcl}
N_F : B(x_0, t_*) & \to & Y \\
x & \mapsto & x - F'(x)^{-1}F(x).
\end{array} \tag{17}$$

One can apply a *single* Newton iteration on any $x \in B(x_0, t_*)$ to obtain $N_F(x)$ which may not belong to $B(x_0, t_*)$, or even may not belong to the domain of F. So, this is enough to guarantee, on $B(x_0, t_*)$, well definedness of only one iteration. To ensure that Newton iterations may be repeated indefinitely in x_0, we need some additional results.

First, define some subsets of $B(x_0, t_*)$ in which, as we shall prove, Newton iteration (17) is "well behaved".

$$K(t) := \left\{ x \in B[x_0, t] \ : \ \|F'(x)^{-1}F(x)\| \leqslant -\frac{f(t)}{f'(t)} \right\}, \qquad t \in [0, t_*)\,, \tag{18}$$

$$K := \bigcup_{t \in [0, t_*)} K(t). \tag{19}$$

In (18), $0 \leqslant t < t_*$, therefore, $f'(t) \neq 0$ and F' is non-singular in $B[x_0, t] \subset B[x_0, t_*)$ (Proposition 6). So, the definitions are consistent.

Lemma 8. *For each $t \in [0, t_*)$, $K(t) \subset B(x_0, t_*)$ and*

$$N_F\left(K(t)\right) \subset K\left(n_f(t)\right).$$

As a consequence, $K \subset B(0, t_*)$ and $N_F(K) \subset K$.

Proof. The first inclusion follows trivially from the definition of $K(t)$.

Take $t \in [0, t_*)$, $x \in K(t)$. Using definition (18) and the first two statements in Proposition 4 we have

$$\|x - x_0\| \leq t, \qquad \|F'(x)^{-1}F(x)\| \leq -f(t)/f'(t), \quad t < n_f(t) < t_*. \tag{20}$$

Therefore

$$\begin{aligned} \|N_F(x) - x_0\| &\leqslant \|x - x_0\| + \|N_F(x) - x\| = \|x - x_0\| + \|F'(x)^{-1}F(x)\| \\ &\leqslant t - f(t)/f'(t) = n_f(t) < t_*\,, \end{aligned}$$

and

$$N_F(x) \in B[x_0, n_f(t)] \subset B(x_0, t_*). \tag{21}$$

Since $N_F(x)$, $n_f(t)$ belongs to the domain of F and f, respectively, using the definitions of Newton iterations on (10), (17) and linearization erros (15) and (16), we obtain

$$\begin{aligned} f(n_f(t)) &= f(n_f(t)) - [f(t) + f'(t)(n_f(t) - t)] \\ &= e(t, n_f(t)) \end{aligned}$$

and

$$\begin{aligned} F(N_F(x)) &= F(N_F(x)) - [F(x) + F'(x)(N_F(x) - x)] \\ &= E(x, N_F(x))\,. \end{aligned}$$

From the two latter equations, (20) and Lemma 7 we have

$$\begin{aligned}\|F'(x_0)^{-1}F(N_F(x))\| &= \|F'(x_0)^{-1}E(x, N_F(x))\| \\ &\leq e(t, n_f(t))\frac{\|F'(x)^{-1}F(x)\|^2}{(f(t)/f'(t))^2} \leq e(t, n_f(t)) = f(n_f(t)).\end{aligned}$$

As $\|N_F(x) - x_0\| \leq n_f(t)$, it follows from Proposition 6 that $F'(N_F(x))$ is non-singular and

$$\|F'(N_F(x))^{-1}F'(x_0)\| \leq -1/f'(n_f(t)).$$

Combining the two above inequalities we conclude

$$\begin{aligned}\|F'(N_F(x))^{-1}F(N_F(x))\| &\leqslant \|F'(N_F(x))^{-1}F'(x_0)\| \, \|F'(x_0)^{-1}F(N_F(x))\| \\ &\leqslant -f(n_f(t))/f'(n_f(t)).\end{aligned}$$

This result, together with (21) show that $N_F(x) \in K(n_f(t))$, which proofs the second inclusion.

The next inclusion (first on the second sentence), follows trivially from definitions (18) and (19). To verify the last inclusion, take $x \in K$. Then $x \in K(t)$ for some $t \in [0, t_*)$. Using the first part of the lemma, we conclude that $N_F(x) \in K(n_f(t))$. To end the proof, note that $n_f(t) \in [0, t_*)$ and use the definition of K. □

Finally, we are ready to prove the main result of this section which is an immediate consequence of the latter result. First note that the sequence $\{x_k\}$ (see (3)) satisfies

$$x_{k+1} = N_F(x_k), \qquad k = 0, 1, \dots , \tag{22}$$

which is indeed an equivalent definition of this sequence.

Corollary 9. *The sequence $\{x_k\}$ is well defined, is contained in $B(x_0, t_*)$, converges to a point $x_* \in B[x_0, t_*]$,*

$$\|x_* - x_k\| \leq t_* - t_k, \qquad k = 0, 1, \dots ,$$

and $F(x_*) = 0$.

Proof. Form (2) and assumption **h1** of the main theorem, we have

$$x_0 \in K(0) \subset K,$$

where the second inclusion follows trivially from (19). Using the above equation, the inclusions $N_F(K) \subset K$ (Lemma 8) and (22) we conclude that the sequence $\{x_k\}$ is well defined and rests in K. From the first inclusion on second part of the Lemma 8 we have trivially that $\{x_k\}$ is contained in $B(x_0, t_*)$.

We will prove, by induction that

$$x_k \in K(t_k), \qquad k = 0, 1, \dots . \tag{23}$$

The above inclusion, for $k = 0$ is the first result on this proof. Assume now that $x_k \in K(t_k)$. Thus, using Lemma 8, (22) and (14) we conclude that $x_{k+1} \in K(t_{k+1})$, which completes the induction proof of (23).

Now, using (23) and (18), we have

$$\|F'(x_k)^{-1} F(x_k)\| \leq -f(t_k)/f'(t_k), \qquad k = 0, 1, \dots ,$$

which, by (3), is equivalent to

$$\|x_{k+1} - x_k\| \leq t_{k+1} - t_k, \qquad k = 0, 1, \dots . \tag{24}$$

Since $\{t_k\}$ converges to t_*, the above inequalities implies that

$$\sum_{k=k_0}^{\infty} \|x_{k+1} - x_k\| \leq \sum_{k=k_0}^{\infty} t_{k+1} - t_k = t_* - t_{k_0} < +\infty,$$

for any $k_0 \in \mathbb{N}$. Hence, $\{x_k\}$ is a Cauchy sequence in $B(x_0, t_*)$ and so, converges to some $x_* \in B[x_0, t_*]$. The above inequality also implies that $\|x_* - x_k\| \leq t_* - t_k$, for any k.

It remains to prove that $F(x_*) = 0$. First, observe that

$$\begin{aligned} \|F'(x_k)\| &\leqslant \|F'(x_0)\| + \|F'(x_k) - F'(x_0)\| \\ &\leq \|F'(x_0)\| + \|F'(x_0)\| \|F'(x_0)^{-1} [F'(x_k) - F'(x_0)]\|. \end{aligned}$$

As $\|x_k - x_0\| \leq t_k$ and $t_k < t_* < R$,

$$\|F'(x_0)^{-1} [F'(x_k) - F'(x_0)]\| \leq f'(\|x_k - x_0\|) - f'(0) \leq f'(t_*) - f'(0).$$

Combining the two above inequalities we have that $\{\|F'(x_k)\|\}$ is bounded. On the other hand, it follows from (3) and (24) that

$$\begin{aligned} \|F(x_k)\| &\leqslant \|F'(x_k)\| \, \|F'(x_k)^{-1} F(x_k)\| \\ &\leqslant \|F'(x_k)\| \, (t_{k+1} - t_k). \end{aligned}$$

Due the fact that $\{\|F'(x_k)\|\}$ is bounded and $\{t_k\}$ converges, we can take limit in the last inequality to conclude that

$$\lim_{k\to\infty} F(x_k) = 0.$$

Since F is continuous in $B[x_0, t_*]$, $\{x_k\} \subset B(x_0, t_*)$ and $\{x_k\}$ converges to x_*, we also have

$$\lim_{k\to\infty} F(x_k) = F(x_*).$$

□

2.3 Uniqueness and Convergence Rate

So far we have proved that the sequence $\{x_k\}$ converges to a solution x_* of $F(x) = 0$ and $x_* \in B[x_0, t_*]$. Now, we are going to prove that this convergence to x_* is at least Q-linearly and x_* is the unique solution of $F(x) = 0$ in the region $B[x_0, t_*]$. Furthermore, by assuming that f satisfies **h4**, we also prove that $\{x_k\}$ converges Q-quadratically to x_* and the uniqueness region increase from $B[x_0, t_*]$ to $B(x_0, \bar{\tau})$. The results will be obtained as a consequence of the following lemma.

Lemma 10. *Take $x, y \in B(x_0, R)$ and $0 \le t < v < R$. If*

$$t < t_*, \qquad \|x - x_0\| \le t, \qquad \|y - x\| \le v - t, \qquad f(v) \le 0, \quad \text{and} \quad F(y) = 0,$$

then,

$$\|y - N_F(x)\| \leqslant [v - n_f(t)] \frac{\|y - x\|^2}{(v - t)^2}.$$

Proof. Direct algebraic manipulation yields

$$\begin{aligned} y - N_F(x) &= y - x + F'(x)^{-1}F(x) - F'(x)^{-1}F(y) \\ &= -F'(x)^{-1}\left[F(y) - F(x) - F'(x)(y - x)\right] = -F'(x)^{-1}E(x, y), \end{aligned}$$

with the assumption $F(y) = 0$ being used in the first equality and definition (15) in the last equality. From the above equation we trivially have

$$y - N_F(x) = \left[-F'(x)^{-1}F'(x_0)\right]\left[F'(x_0)^{-1}E(x, y)\right].$$

Taking the norm on both sides of this equality and using Proposition 6, Lemma 7 together with the assumptions of the lemma we obtain

$$\|y - N_F(x)\| \le (-1/f'(t)) \;\; e(t, v)\, \frac{\|y - x\|^2}{(v - t)^2}.$$

As $0 \le t < t_*$, $f'(t) < 0$. Using also (16) and the assumptions $f(v) \le 0$ we have

$$\begin{aligned}(-1/f'(t)) \quad e(t,v) &= v - t + f(t)/f'(t) - f(v)/f'(t)\\ &\le v - t + f(t)/f'(t) = v - n_f(t).\end{aligned}$$

To end the proof, combine the two above equations. $\square$

Corollary 11. *The sequences $\{x_k\}$ and $\{t_k\}$ satisfy*

$$\|x_* - x_{k+1}\| \le \frac{t_* - t_{k+1}}{(t_* - t_k)^2} \|x_* - x_k\|^2, \qquad \textit{for} \quad k = 0, 1, \ldots . \tag{25}$$

In particular,

$$\|x_* - x_{k+1}\| \leqslant \frac{1}{2}\|x_* - x_k\|, \qquad \textit{for} \quad k = 0, 1, \ldots . \tag{26}$$

Additionally, if f satisfies **h4** *then*

$$\|x_* - x_{k+1}\| \leqslant \frac{D^- f'(t_*)}{-2f'(t_*)}\|x_* - x_k\|^2, \qquad \textit{for} \quad k = 0, 1, \ldots . \tag{27}$$

Proof. Take an arbitrary k and apply Lemma 10 with $x = x_k$, $y = x_*$, $t = t_k$ and $v = t_*$, to obtain

$$\|x_* - N_F(x_k)\| \le [t_* - n_f(t_k)] \, \frac{\|x_* - x_k\|^2}{(t_* - t_k)^2}.$$

Equation (25) follows from the above inequality, (22) and (14).

Note that, by (11) in Proposition 4, (14) and Corollary 9, for any k

$$\frac{t_* - t_{k+1}}{t_* - t_k} \le 1/2 \quad \text{and} \quad \frac{\|x_* - x_k\|}{t_* - t_k} \le 1.$$

Combining these inequalities with (25) we have (26).

Now, assume that **h4** holds. Then, by Corollary 5 the second inequality on (6) holds, which combined with (25) imply (27). $\square$

Corollary 12. *The limit x_* of the sequence $\{x_k\}$ is the unique zero of F in $B[x_0, t_*]$.*

Furthermore, if f satisfies **h4** *then x_* is the unique zero of F in $B(x_0, \bar{\tau})$, where $\bar{\tau}$ is defined as in Theorem 2, i.e.,*

$$\bar{\tau} = \sup\{t \in [t_*, R) \ : \ f(t) \le 0\}.$$

Proof. Let y_* be a zero of F in $B[x_0, t_*]$:

$$\|y_* - x_0\| \leq t_*, \qquad F(y_*) = 0.$$

We will prove by induction that

$$\|y_* - x_k\| \leq t_* - t_k, \qquad k = 0, 1, \dots . \tag{28}$$

For $k = 0$ the above inequality holds trivially, because $t_0 = 0$. Now, assume that the inequality holds for some k. As we also have $\|x_k - x_0\| \leq t_k$, we may apply Lemma 10 with $x = x_k$ $y = y_*$, $t = t_k$ and $v = t_*$ to obtain

$$\|y_* - N_F(x_k)\| \leq [t_* - n_f(t_k)] \frac{\|y_* - x_k\|^2}{(t_* - t_k)^2}.$$

Using latter inequality, the inductive hypothesis (to estimate the quotient in the last term), (14) and (22) we obtain that (28) also holds for $k + 1$. This completes the induction proof. Because $\{x_k\}$ converges to x_* and $\{t_k\}$ converges to t_*, from (28) we conclude $y_* = x_*$. Therefore, x_* is the unique zero of F in $B[x_0, t_*]$.

Now, suppose that f satisfies **h4** and so, equivalently, $t_* < \bar{\tau}$. We already know that if $y_* \in B[x_0, t_*]$ then $y_* = x_*$. It remains to prove that F does not have zeros in $B(x_0, \bar{\tau}) \setminus B[x_0, t_*]$. For proving this fact by contradiction, assume that F does have a zero there, i.e., there exists $y_* \in X$,

$$t_* < \|y_* - x_0\| < \bar{\tau}, \qquad F(y_*) = 0.$$

We will prove that the above assumptions can not hold. First, using Lemma 7 with $x = x_0$, $y = y_*$, $t = 0$ and $v = \|y_* - x_0\|$ we obtain that

$$\|F'(x_0)^{-1} E(x_0, y_*)\| \leqslant e(0, \|y_* - x_0\|) \frac{\|y_* - x_0\|^2}{\|y_* - x_0\|^2} = e(0, \|y_* - x_0\|).$$

As we are assuming that $F(y_*) = 0$, using also (15) and (2) we conclude

$$\begin{aligned}
\|F'(x_0)^{-1} E(x_0, y_*)\| &= \|F'(x_0)^{-1}[-F(x_0) - F'(x_0)(y_* - x_0)]\| \\
&= \|y_* - x_0 + F'(x_0)^{-1} F(x_0)\| \\
&\geq \|y_* - x_0\| - \|F'(x_0)^{-1} F(x_0)\| \\
&= \|y_* - x_0\| - f(0).
\end{aligned}$$

From (16) and assumption **h1** we have that

$$e(0, \|y_* - x_0\|) = f(\|y_* - x_0\|) - f(0) + \|y_* - x_0\|.$$

Now, combining this equality with two above inequalities it easy to see that

$$f(\|y_* - x_0\|) - f(0) + \|y_* - x_0\| \geq \|y_* - x_0\| - f(0),$$

or equivalently, $f(\|y_* - x_0\|) \geq 0$. Thus f, being strictly convex, is strictly positive in the interval $(\|y_* - x_0\|, R)$. So, $\bar{\tau} \leq \|y_* - x_0\|$, in contradictions with the above assumptions. Therefore, F does not have zeros in $B[x_0, \bar{\tau}) \setminus B[x_0, t_*]$ and x_* is the unique zero of F in $B(x_0, \bar{\tau})$. □

Therefore, it follows from Corollary 5, Corollary 9, Corollary 11 and Corollary 12 that all statements in Theorem 2 are valid.

Remark 2. *The proof of the second part of Corollary 12 is essentially the same one presented in [3, Lemma 3.1]*

2.4 Limit Case for Kantorovich's Theorem

For proving convergence and estimating its rate, only the regions $[0, t_*]$ and $B[x_0, t_*]$ were considered. Indeed, for obtaining Q-quadratic convergence, the behavior of f beyond t_* was not used. So, we give now a formulation which involves only the regions above mentioned.

Theorem 13. *Let X be a Banach space, $C \subseteq X$ and $F : C \to Y$ a continuous function, continuously differentiable on $int(C)$.*

Take $x_0 \in \operatorname{int}(C)$ with $F'(x_0)$ non-singular. Suppose that there exist $t_* > 0$ and a continuously differentiable function $f : [0, t_*] \to \mathbb{R}$ such that, $B[x_0, t*] \subseteq C$,

$$\|F'(x_0)^{-1} [F'(y) - F'(x)] \| \leq f'(\|y - x\| + \|x - x_0\|) - f'(\|x - x_0\|),$$

for $x, y \in B(x_0, t_*)$, $\|x - x_0\| + \|y - x\| < t_*$,

$$\|F'(x_0)^{-1} F(x_0)\| \leq f(0) \, ,$$

and

h1') $f(0) > 0$, $f'(0) = -1$*;*

h2') f' *is convex and strictly increasing;*

h3') $f(t) > 0$ *in* $[0, t_*)$ *and* $f(t_*) = 0$*.*

Then the sequences generated by Newton's Method for solving $f(t) = 0$ and $F(x) = 0$ with starting point $t_0 = 0$ and x_0, respectively,

$$t_{k+1} = t_k - f'(t_k)^{-1} f(t_k), \qquad x_{k+1} = x_k - F'(x_k)^{-1} F(x_k), \qquad k = 0, 1, \dots ,$$

are well defined, $\{t_k\}$ is strictly increasing, is contained in $[0, t_*)$, and converges to t_*, $\{x_k\}$ is contained in $B(x_0, t_*)$ and converges to a point $x_* \in B[x_0, t_*]$ which is the unique zero of F in $B[x_0, t_*]$,

$$\|x_* - x_k\| \leq |t_* - t_k|, \qquad \|x_* - x_{k+1}\| \leq \frac{t_* - t_{k+1}}{(t_* - t_k)^2} \|x_* - x_k\|^2, \qquad k = 0, 1, \dots ,$$

and the sequences $\{t_k\}$ and $\{x_k\}$ converge Q-linearly as follows

$$\|x_* - x_{k+1}\| \leq \frac{1}{2} \|x_* - x_k\|, \qquad t_* - t_{k+1} \leq \frac{1}{2}(t_* - t_k) \qquad k = 0, 1, \dots .$$

If, additionally,

h4') $f'(t_*) < 0$,

then the sequences $\{t_k\}$ and $\{x_k\}$ converge Q-quadratically as follows

$$\|x_* - x_{k+1}\| \leqslant \frac{D^- f'(t_*)}{-2f'(t_*)} \|x_* - x_k\|^2, \qquad t_* - t_{k+1} \leq \frac{D^- f'(t_*)}{-2f'(t_*)} (t_* - t_k)^2, \quad k = 0, 1, \dots .$$

3 Final Remarks

Kantorovich's Theorem was used in [10] to prove Smale's Theorem [9], and it was used in [2] to prove Nesterov-Nemirovskii's Theorem [6]. We present these proofs here, for the sake of ilustration.

Let we start with a definition.

Definition 1. *Let X be a Banach space, $C \subseteq X$ and $F : C \to Y$ a continuous function, continuously differentiable on $int(C)$. Take $x_0 \in \mathrm{int}(C)$ with $F'(x_0)$ non-singular. A continuously differentiable function $f : [0, R) \to \mathbb{R}$ is said to be a majorant function to F in x_0 if $B(x_0, R) \subseteq C$ and* (1), (2), **h1**, **h2**, **h3**, **h4** *are satisfied.*

The next results gives a condition more easy to check then condition (1), when de functions under consideration are two times continuously differentiable.

Lemma 14. *Let X be a Banach space, $C \subseteq X$ and $F : C \to Y$ a continuous function, two times continuously differentiable on $int(C)$. Let $f : [0, R) \to \mathbb{R}$ be a two times continuously differentiable function with derivative f' convex. Then F satisfies* (1) *if, only if,*

$$\|F(x_0)^{-1}F''(x)\| \leqslant f''(\|x - x_0\|), \tag{29}$$

for all $x \in C$ such that $\|x - x_0\| < R$.

Proof. If F satisfies (1) then (29) holds trivially.

Reciprocally, taking $x, y \in C$ such that $\|x - x_0\| + \|y - x\| < R$, we obtain that

$$\|F(x_0)^{-1}[F'(y) - F'(x)]\| \leqslant \int_0^1 \|F''(x + \tau(y - x))\|\|y - x\|d\tau.$$

Now, as f satisfies (29) and f' is convex, we obtain from last inequality that

$$\begin{aligned}\|F(x_0)^{-1}[F'(y) - F'(x)]\| &\leqslant \int_0^1 f''(\|(x - x_0) + \tau(y - x)\|)\|y - x\|d\tau \\ &\leqslant \int_0^1 f''(\|x - x_0\| + \tau\|y - x\|)\|y - x\|d\tau \\ &= f'(\|x - x_0\| + \|y - x\|) - f'(\|x - x_0\|),\end{aligned}$$

which implies that F satisfies (1), and the lemma is proved. □

Theorem 15. *(Smale's Theorem). Let X be a Banach space, $C \subseteq X$ and $F : C \to Y$ a continuous function, analytic on $int(C)$. Take $x_0 \in int(C)$ with $F'(x_0)$ non-singular and define*

$$\gamma := \sup_{k>1} \left\| \frac{F'(x_0)^{-1}F^{(k)}(x_0)}{k!} \right\|^{\frac{1}{k-1}}.$$

Suppose that $B(x_0, 1/\gamma) \subseteq C$ and there exists $\beta \geq 0$ such that

$$\left\|F'(x_0)^{-1}F(x_0)\right\| \leqslant \beta,$$

and $\alpha := \beta\gamma \leqslant 3 - 2\sqrt{2}$. Then sequence generated by Newton's Method for solving $F(x) = 0$ with starting x_0

$$x_{k+1} = x_k - F'(x_k)^{-1}F(x_k), \qquad k = 0, 1, \dots,$$

is well defined, $\{x_k\}$ is contained in $B(x_0, t_*)$ and converges to a point x_* which is the unique zero of F in $B[x_0, t_*]$, where $t_* := (\alpha + 1 - \sqrt{(\alpha + 1)^2 - 8\alpha})/(4\gamma)$. Moreover, $\{x_k\}$ converges Q-linearly as follows

$$\|x_* - x_{k+1}\| \leqslant \frac{1}{2}\|x_* - x_k\|, \qquad k = 0, 1, \dots.$$

Additionally, if $\alpha < 3 - 2\sqrt{2}$ *then* $\{x_k\}$ *converges Q-quadratically as follows*

$$\|x_* - x_{k+1}\| \leqslant \frac{\gamma}{(1-\gamma t_*)[2(1-\gamma t_*)^2 - 1]}\|x_* - x_k\|^2, \qquad k = 0, 1, \dots ,$$

and x_* *is the unique zero of* F *in* $B(x_0, t_{**})$*, where* $t_{**} := (\alpha + 1 + \sqrt{(\alpha+1)^2 - 8\alpha})/(4\gamma)$.

Proof. Use Lemma 14 to prove that $f : [0, 1/\gamma) \to \mathbb{R}$ defined by $f(t) = t/(1-\gamma t) - 2t + \beta$, is a majorant function to F in x_0, with roots equal to t_* and t_{**}, see [10]. So, the result follows from Theorem 2. $\square$

Theorem 16. *(Nesterov-Nemirovskii's Theorem). Let* $C \subset \mathbb{R}^n$ *be a open convex set and let* $g : C \to \mathbb{R}$ *be a strictly convex function, three times continuously differentiable* $int(C)$*. Take* $x_0 \in int(C)$ *with* $g''(x_0)$ *non-singular. Define the norm*

$$\|u\|_{x_0} := \sqrt{\langle u, u\rangle_{x_0}}, \qquad \forall\, u \in \mathbb{R}^n,$$

where $\langle u, v\rangle_{x_0} = a^{-1}\langle g''(x_0)u, v\rangle$*, for all* $u, v \in \mathbb{R}^n$ *and some* $a > 0$*. Suppose that* g *is a-self-concordant, i.e., satisfies*

$$|g'''(x)[h,h,h]| \leqslant 2a^{-1/2}(g''(x)[h,h])^{3/2}, \qquad \forall\, x \in C,\ h \in \mathbb{R}^n,$$

$W_1(x_0) = \{x \in \mathbb{R}^n : \|x - x_0\|_{x_0} < 1\} \subset C$ *and there exists* $\beta \geq 0$ *such that*

$$\|g''(x_0)^{-1}g'(x_0)\|_{x_0} \leqslant \beta \leqslant 3 - 2\sqrt{2}.$$

Then the sequence generated by Newton method to solve $g'(x) = 0$ *(or equivalently, to minimizer* g*) with starting point* x_0

$$x_{k+1} = x_k - g''(x_k)^{-1}g'(x_k), \qquad k = 0, 1, \dots ,$$

is well defined, $\{x_k\}$ *is contained in* $W_{t_*}(x_0) = \{x \in \mathbb{R}^n : \|x - x_0\|_{x_0} < t_*\}$ *and converges to a point* x_* *which is the unique minimizer of* g *in* $W_{t_*}[x_0] = \{x \in \mathbb{R}^n : \|x - x_0\|_{x_0} \leqslant t_*\}$*, where* $t_* := (\beta + 1 - \sqrt{(\beta+1)^2 - 8\beta})/4$*. Moreover,* $\{x_k\}$ *converges Q-linearly as follows*

$$\|x_* - x_{k+1}\| \leqslant \frac{1}{2}\|x_* - x_k\|, \qquad k = 0, 1, \dots .$$

Additionally, if $\beta < 3 - 2\sqrt{2}$ *then* $\{x_k\}$ *converges Q-quadratically as follows*

$$\|x_* - x_{k+1}\| \leqslant \frac{1}{(1-t_*)[2(1-t_*)^2 - 1]}\|x_* - x_k\|^2, \qquad k = 0, 1, \dots ,$$

and x_* *is the unique minimizer of* g *in* $W_{t_{**}}(x_0) = \{x \in \mathbb{R}^n : \|x - x_0\|_{x_0} < t_{**}\}$*, where* $t_{**} := (\beta + 1 + \sqrt{(\beta+1)^2 - 8\beta})/4$.

Proof. Let $X := (\mathbb{R}^n, \|.\|_{x_0})$ be a Banach space. Use Lemma 14 to prove that $f : [0, 1) \to \mathbb{R}$ defined by $f(t) = t/(1-t) - 2t + \beta$, is a majorant function to g' in x_0, with roots equal to t_* and t_{**}, see [2]. So, the result follows from Theorem 2. $\square$

References

[1] Appel,J., De Pascale, E., Lysenko, J.V., and Zabrejko, P.P., *New results on Newton-Kantorovich approximations with applications to nonlinear integral equations*, Numer. funct. Anal. and Optimiz. , 18(1& 2) (1997), 1-17.

[2] Alvarez, F., Botle, J. and Munier, J., *A Unifying Local Convergence Result for Newton's Method in Riemannian Manifolds* INRIA, Rapport de recherche, N. 5381, (2004).

[3] Ferreira, O.P. and Svaiter, B. F., *Kantorovich's Theorem on Newton's method in Riemannian Manifolds* Journal of Complexity, 18, (2002), 304–329.

[4] Huang, Z., *Newton Method under weak Lipschitz continuous derivative in Banach Spaces*, Appl. Math. and Comput., 140, (2003), 115-126.

[5] Kantorovich, L.V., and Akilov, G.P., *Functional analysis in normed spaces*, Oxford, Pergamon (1964).

[6] Nesterov, Y. and Nemirovskii, A. *Interior-Point Polynomial Algorithms in Convex Programming*, SIAM Studies in Applied Mathematics, 13, Philadelphia, PA, 1994.

[7] Polyak, B. T. *Newton-Kantorovich method and its global convergence.* Zap. Nauchn. Sem. S.-Peterburg. Otdel. Mat. Inst. Steklov. (POMI) 312 (2004).

[8] Potra, Florian A., *The Kantorovich Theorem and interior point methods* Mathematical Programming, 102, 1 (2005) 47–70.

[9] Smale, S. , *Newton method estimates from data at one point*, The Merging of Disciplines: New Directions in Pure, Applied and Computational Mathematics (Ewing, R., Gross, K., and Martin, C. eds.), Springer-Verlag, New York, (1986), pp. 185–196.

[10] Wang, X., *Convergence of Newton's method and inverse function theorem in Banach space*, Math. Comp. **68**, 225 (1999), pp.169-186.