Workshop on Nonsmooth Optimization and Applications NOPTA 8-12 April 2024 University of Antwerp, Belgium

Newton-type Methods in Abstract Metric Spaces

Titus Pinta
Institut für Numerische und Angewandte Mathematik, University of Göttingen
t.pinta@gmail.com

GEORG-AUGUST-UNIVERSITÄT GÖTTINGEN IN PUBLICA COMMODA SEIT 1737

Abstract

This poster presents a novel approach to the differentiation of vector valued functions defined on abstract metric spaces. Using this machinery, a family of Newton-type methods is proposed. Such a method aims to provide an efficient algorithm for solving $F(x) = 0$ where $F : \mathbb{M} \to \mathbb{R}^n$ and $\mathbb{M}$ is a complete metric space. We then establish sufficient conditions for the super-linear convergence of the methods. Further, a test for the existence of a zero for F is developed, yielding a Newton-Kantorovich type theorem. The proposed framework can easily find real world applications, for instance in the study of network flow problems and in optimization problems with manifold constrains.

Algebra

Consider a complete metric space $(\mathbb{M}, \mathrm{dist})$ and a mapping $H : \mathbb{M} \times \mathbb{M} \to \mathbb{R}^n$ such that

$$\forall x \in \mathbb{M} \quad H(x, x) = 0. \tag{1}$$

We denote the set of all mappings that satisfy (1) with $S(\mathbb{M})$.

Definition 1 *A $H \in S(\mathbb{M})$ is called* invertably compatible *if there exists $m \in \mathbb{R}$ such that*

$$\forall x, y \in M, \forall v, w \in \mathbb{R}^n \quad d(H^{-1}(x, v), H^{-1}(y, w)) \le m\|v - w - H(x, y)\|. \tag{2}$$

We denote $\|H^{-1}\| := \inf\{m \mid (2) \text{ holds}\}$. The set of all such mappings is denoted by $GS(\mathbb{M})$

Newton Differentiability

Definition 2 *Let $\mathbb{M}$ be a complete metric space. A function $F : \mathbb{M} \to \mathbb{R}^n$ is called* weakly pointwise Newton differentiable at $\bar{x}$ *if there exists a set valued mapping $\mathcal{H}F : \mathbb{M} \rightrightarrows S(\mathbb{M})$ such that*

$$\lim_{x \to \bar{x} \in V} \sup_{H \in \mathcal{H}F(x)} \frac{\|F(x) - F(\bar{x}) - H(x, \bar{x})\|}{\mathrm{dist}(x, \bar{x})} < \infty.$$

Furthermore, if

$$\lim_{x \to \bar{x} \in V} \sup_{H \in \mathcal{H}F(x)} \frac{\|F(x) - F(\bar{x}) - H(x, \bar{x})\|}{\mathrm{dist}(x, \bar{x})} = 0.$$

the function F is called pointwise Newton differentiable at $\bar{x}$.

The set valued mapping $\mathcal{H}F$ is called a Newton differential of F.

Definition 3 *Let $F : \mathbb{M} \to \mathbb{M}$ Newton differentiable at $\bar{x}$ with Newton differential $\mathcal{H}F$. The fixed point iteration of the set-valued operator $\mathcal{N}_{\mathcal{H}F} : \mathbb{M} \rightrightarrows \mathbb{M}$,*

$$\mathcal{N}_{\mathcal{H}F} x = \{H^{-1}(x, -F(x)) \mid H \in \mathcal{H}F(x) \cap GS(\mathbb{M})\},$$

$$x^{k+1} \in \mathcal{N}_{\mathcal{H}F} x^k \tag{3}$$

is called a Newton-type method.

Main Result

Theorem 1 *Let $F : U \subseteq \mathbb{M} \to \mathbb{R}^n$ be pointwise Newton differentiable at $\bar{x}$ with $F(\bar{x}) = 0$. and Newton differential $\mathcal{H}F$. Assume that the Newton-type iteration is proper (nowhere empty) i.e. for all x^k produced by the iteration, $\mathcal{H}F(x^k) \cap GS(\mathbb{M}) \neq \emptyset$. Furthermore, assume that the set $\bigcup_{x \in U}\{\|H^{-1}\| \mid H \in \mathcal{H}F(x) \cap GS(\mathbb{M})\}$ is bounded. Then any sequence $\{x^k\}_{k \in \mathbb{N}}$ generated by (3) with x^0 near $\bar{x}$ converges super-linearly to $\bar{x}$.*

Kantorovich Theorem

For the super-linear convergence of Newton-type methods, the existence of a zero is required. The next theorem provides a test, evaluated only at x^0, for the existence of such a zero. The first step in stating the theorem consists in extending the notion of Hölder continuity.

Definition 4 *A mapping $H : \mathbb{M} \to GS(\mathbb{M})$ is called* pointwise h-smooth at x^0 *if there exists $\kappa > 0$ and $\alpha > 0$ such that*

$$\forall x \in \mathbb{M}, \forall y, z \in \mathbb{M} \quad \mathrm{dist}(x, H(x)^{-1}(x, H(x^0)(z, y))) \le (1 + \kappa \mathrm{dist}(x, x^0)^{\alpha}) \,\mathrm{dist}(y, z).$$

Theorem 2 *Let $\mathbb{M}$ be a complete metric spaces and $F : U \subseteq \mathbb{M} \to \mathbb{R}^n$ be pointwise Newton differentiable on U with the Newton differential, $\mathcal{H}F$, having a h-smooth selection denote by HF i.e. HF is h-smooth and $\forall x \in U, HF(x) \in \mathcal{H}F(x)$. Let $\kappa > 0$ and $\gamma \in [1, 2]$ be such that*

$$\forall x \in \mathbb{M}, \forall y, z \in \mathbb{M} \quad \mathrm{dist}(x, HF(x)^{-1}(x, HF(x^0)(z, y))) \le (1 + \kappa \mathrm{dist}(x, x^0)^{\gamma - 1}) \,\mathrm{dist}(y, z).$$

and let $L > 0$ be such that

$$\forall x, y \in U : \quad \|F(x) - F(y) - HF(x)(x, y)\| \le \frac{L}{2} \mathrm{dist}(x, y)^{\gamma}.$$

Let $x^0 \in U$ and assume there exists $B < \infty$ such that

$$B = \sup_{x \neq y \in \mathbb{M}} \frac{\|HF(x^0)(x, y)\|}{d(x, y)}$$

and $\eta := d(x^0, HF(x^0)^{-1}(x^0, F(x^0)))$. Suppose further that there exist a constant M^* such that for all $x \in U$

$$\|HF(x)^{-1}\| \le M^*.$$

Assume that for all $x \in U$,

$$HF(x)^{-1}(x, -F(x)) \in U.$$

Furthermore assume there exists $\bar{t} \in (0, \infty)$ and a function $f \in \mathcal{C}^2[0, \bar{t}]$ with Lipschitz continuous second derivative such that for all $t < \bar{t}$

(a)

$$\frac{LB}{2}\left(\frac{f(t)}{f'(t)}\right)^{\gamma} \le f\left(t - \frac{f(t)}{f'(t)}\right), \tag{4}$$

(b)

$$f(0) = \eta, \quad f(t) > 0, \quad f(\bar{t}) = 0.$$

(c)

$$-1 < -1 + \kappa t^{\gamma - 1} \le f'(t) < 0,$$

(d)

$$f''(t) > 0$$

Then there exists $\bar{x} \in U$ such that $F(\bar{x}) = 0$.

The theorem requires similar smoothness assumptions as the theorem proving super-linear convergence. Further assumptions require only information regarding the function and its Newton differential at x^0, transforming this theorem into a numerically implementable test for the existence of a zero. In order to implement this test, an optimization based solver for the differential inclusion in (4) can be employed.

Euclidean Spaces

In the setting of $\mathbb{M} = \mathbb{R}^n$ with the Euclidean metric, the presented formalism recovers known results about nonsmooth Newton-type methods.

Example 1 *Let $T \in \mathbb{R}^{n \times n}$. Then the mapping $H : \mathbb{R}^n \times \mathbb{R}^n \to \mathbb{R}^n$ defined by*

$$H(x, y) = T(y - x)$$

is in $S(\mathbb{R}^n)$. Furthermore, if T is invertible, then H is invertably compatible with

$$H^{-1}(x, v) = x + T^{-1}v.$$

This can be seen because

$$\mathrm{dist}(H^{-1}(x, v), H^{-1}(y, w)) = \|y + T^{-1}w - x - T^{-1}v\| = \|T^{-1}\|\|v - w - T(y - x)\|$$

Definition 5 *A function $F : U \subseteq \mathbb{R}^n \to \mathbb{R}^m$ has a directional derivative at $\bar{x} \in U$ where U is a neighbourhood of $\bar{x}$ if for all $y \in U$*

$$\lim_{t \downarrow 0} \frac{F(x + td) - F(x)}{t} = F'(x; d) \tag{5}$$

exists and is finite.

Definition 6 *A Lipschitz continuous mapping $F : U \subseteq \mathbb{R}^n \to \mathbb{R}^m$ is called weak semi-smooth at $\bar{x} \in U$ if it has a directional derivative at all $x \in U$ and*

$$\lim_{x \to \bar{x}} \frac{\|F'(x; x - \bar{x}) - F'(\bar{x}; x - \bar{x})\|}{\|x - \bar{x}\|} < \infty, \tag{6}$$

where D is the full measure where F is differentiable. Further, if

$$\lim_{x \to \bar{x}} \frac{\|F'(x; x - \bar{x}) - F'(\bar{x}; x - \bar{x})\|}{\|x - \bar{x}\|} = 0, \tag{7}$$

F is called semi-smooth at $\bar{x} \in U$.

Proposition 1 *A (weakly) semi-smooth mapping at $\bar{x}$, $F : U \subseteq \mathbb{R}^n \to \mathbb{R}^m$ is (weakly) Newton differentiable at $\bar{x}$ with $\mathcal{H}F(x) = \{(x_1, x_2) \mapsto F'(x; x_2 - x_1)\}$.*

Gradient Descent

The classic convergence rate result of the gradient descent algorithm applied to a strongly convex function represents an essential example of the application of the Newton differentiability theory.

Proposition 2 *Let $f : \mathbb{R}^n \to \mathbb{R}$ a bounded from below, strongly convex function with L-Lipschitz continuous gradient, then ∇f is weak Newton differentiable near the unique global minimum $\bar{x}$ with a Newton differential given by $\mathcal{H}\nabla f(x) = \{\alpha \mathrm{I}\}$ for any choice of step size α.*

Riemannian Manifolds

The framework developed can also apply to Riemannian manifolds. Let $F : U \subseteq \mathcal{M} \to \mathbb{R}^n$ a smooth mapping and $\mathrm{d}F(x) : T_x\mathcal{M} \to \mathbb{R}^n$ its differential. Let $\bar{x} \ni U \subseteq \mathcal{M}$ such that $F(\bar{x}) = 0$. On the neighbourhood U, we can define the mappings $\log(x) : \mathcal{M} \to T_x\mathcal{M}$, $\exp(x) : T_x\mathcal{M} \to \mathcal{M}$ and a local frame bundle yielding isomorphisms $O(x) : T_x\mathcal{M} \to \mathbb{R}^n$. The parallel transport map is denoted by $\Gamma(x, y) : T_x\mathcal{M} \to T_y\mathcal{M}$.

Proposition 3 *Let $\mathcal{M}$ be a complete n dimensional Riemannian manifold, with the Riemannian metric. Let $F : \mathcal{M} \to \mathbb{R}^m$ a smooth mapping. Then F is pointwise Newton Differentiable at any point x^0 with Newton Differential $\mathcal{H}F : \mathcal{M} \to S(\mathcal{M})$ given by*

$$\mathcal{H}F(x)(y, z) = \{O(x)(\mathrm{d}F(x)\Gamma(y, z)(\log(y)(z)))\}.$$

Proposition 4 *Let $F : U \subseteq \mathcal{M} \to \mathbb{R}^n$ a smooth mapping and let L such that*

$$\forall x, y \in U \quad \frac{\|F(x) - F(y) - O(x)\,\mathrm{d}F(x)\Gamma(y, x)(\log(y)(x))\|}{\mathrm{dist}(x, y)^2} \le \frac{L}{2}.$$

Denote by $B := \|\,\mathrm{d}F(x^0)\|$, and $\eta := \|(\mathrm{d}F(x^0)O(x))^{-1}F(x^0)\|$. Furthermore assume that $LB\eta \le 1/2$. Then there exists $\bar{x} \in U$ such that $F(\bar{x}) = 0$.

Conclusions

The main aim of the present work is to develop a framework for understanding the behaviour of Newton-type methods in abstract metric spaces. The notion of invertible linear maps has to be adapted to the non-linear setting of metric spaces. This definition is informed by the requirements for the convergence of the Newton-type methods. Similarly, the concept of differentiability materializes itself as Newton-differentiability in the non-linear context. With this definitions, we can formulate super-linearly convergent Newton-type methods. Finally the classic Kantorovich theorem can be adapted to the setting of metric spaces.

Acknowledgements

The author would like to express his gratitude to Russell Luke for his guidance and advice.
This work was carried out with the support of DFG RTG2088.

References

[1] R. Bartle. Newton's Method in Banach Spaces. 1955.

[2] R. Cibulka, A. Dontchev, J. Preininger, V. Veliov, and T. Roubal. Kantorovich-type Theorems for Generalized Equations. *Journal of Convex Analysis*, 25:459–486, 01 2018.

[3] L. V. Kantorovich. Functional Analysis and Applied Mathematics. *Uspehi Matem. Nauk (N.S.)*, 3(6(28)):89–185, 1948.

[4] J. M. Ortega. The Newton-Kantorovich Theorem. *American Mathematical Monthly*, 75:658, 1968.

[5] L. Qi and J. Sun. A Nonsmooth Version of Newton's Method. *Math. Program.*, 58:353–367, 01 1993.

Download

POSTER