Math. Program., Ser. A 102: 223–248 (2005)

Digital Object Identifier (DOI) 10.1007/s10107-004-0523-x

Liqun Qi* · Alexander Shapiro · Chen Ling

Differentiability and semismoothness properties of integral functions and their applications

Received: July 22, 2003 / Accepted: April 5, 2004
Published online: 31 May 2004 – © Springer-Verlag 2004

Abstract. In this paper we study differentiability and semismoothness properties of functions defined as integrals of parameterized functions. We also discuss applications of the developed theory to the problems of shape-preserving interpolation, option pricing and semi-infinite programming.

Keywords: Smoothness – Semismoothness – Shape-preserving interpolation – Option price problem – Semi-infinite programming

1. Introduction

The integral function $F : \mathbb{R}^n \to \mathbb{R}$, defined by

$$F(x) := \int_a^b [g(x, s)]_+ p(s) ds, \tag{1.1}$$

where $\alpha_+ := \max\{0, \alpha\}$ and $p(s) \geq 0$ for all $s \in [a, b]$, arises from nonsmooth equation reformulations of the shape-preserving interpolation problem and the option price problem. It also arises in the aggregate reformulation of the semi-infinite program. Convergence analyses of numerical methods designed for solving such problems via their reformulations are highly related to differentiability properties of this integral function.

Differentiability properties and applications of the integral function (1.1) were discussed in recent publications by Dontchev, Qi and Qi [4],[5], Qi [16], Qi and Tseng [19], Qi and Yin [20], and Wang, Yin and Qi [33]. The aim of this paper is an investigation of smoothness, semismoothness, p-order semismoothness, strong semismoothness and SC^1 properties of a general class of integral functions which includes functions of the form (1.1) as a particular case. We also discuss applications of the derived results to the above mentioned problems.

L. Qi: Department of Applied Mathematics, The Hong Kong Polytechnic University, Hung Hom, Kowloon, Hong Kong. e-mail: maqilq@polyu.edu.hk

A. Shapiro: School of Industrial and Systems Engineering, Georgia Institute of Technology, Atlanta, Georgia 30332-0205, USA. e-mail: ashapiro@isye.gatech.edu

C. Ling: School of Information, Zhejiang University of Finance and Economics, Hangzhou, 310012, China. e-mail: linghz@hzcnc.com

Current address: Department of Applied Mathematics, The Hong Kong Polytechnic University, Hung Hom, Kowloon, Hong Kong. e-mail: 01902146r@polyu.edu.hk

* Supported by the Hong Kong Research Grant Council under grant PolyU 5296/02P.

Let us remark that semismoothness, p-order semismoothness, and strong semismoothness are the key conditions for superlinear, $(1 + p)$-order and quadratic convergence, respectively, of the generalized Newton method for solving a system of nonsmooth equations [14],[18],[17]. On the other hand, the SC^1 property is the key condition for superlinear convergence of the SQP method for solving one time, but not twice, differentiable nonlinear programming problems [15],[7],[12].

In the last decade, the semismooth Newton method became a powerful tool for solving large scale nonlinear complementarity and variational inequality problems. This may be seen in the fundamental monograph by Facchinei and Pang [8] and the abundant references in that book. In the recent five years, while there are still further research work on the semismooth Newton method for solving nonlinear complementarity and variational inequality problems, the semismooth Newton method has been further applied to semidefinite problems [26], operator equations [30], shape-preserving interpolation problems [4],[5] and option price problems [33].

In the applications of shape-preserving interpolation problems and option price problems, the integral function F, defined by (1.1), plays a central role. Its semismoothness was established in [4]. This proved superlinear convergence of a Newton-like method for solving the system of nonsmooth equations arising in that problem, which was a conjecture for 15 years [9]. In [5], strong semismoothness of a particular form of the integral function for that shape-preserving interpolation problem was established. This, further, established quadratic convergence of the Newton-like method. In [20], this result was generalized to a class of integral functions, which are still a special case of (1.1). A counterexample was also given there showing that an integral function defined by (1.1) may not be strongly semismooth. In [33], the semismooth Newton method was further applied to the option price problem. Based upon the results of [4], semismoothness of the integral function F, and hence superlinear convergence of the employed generalized Newton method, were established, but strong semismoothness of F and quadratic convergence of the method still remained a question.

It was discovered in [5] that though the function $[g(x, s)]_+ p(s)$ is piecewise smooth as long as g and p are smooth (continuously differentiable), the integral function F may not be piecewise smooth. Such function F, constructed in [5], is composed from two-dimensional functions, which are strongly semismooth everywhere except at the origin, where they are not differentiable. In [23], Rockafellar proved that a continuous n-dimensional function, where $n \geq 2$, is not piecewise smooth if it is smooth everywhere except at a point where it is not differentiable. Qi and Tseng [19] revealed that such functions belong to a class of nonsmooth functions, which is totally different from piecewise smooth functions. They call such functions almost smooth functions. Many familiar functions, such as p-norm functions $(1 < p < \infty)$, differentiable penalty functions, smoothing functions, are actually almost smooth functions. In [16], Qi suggested that the SQP method can be applied to the aggregate reformulation of the semi-infinite program. This gives a further motivation for studying the SC^1 properties of the integral functions.

The aim of this paper is to answer, at least partially, these questions and to apply the obtained results to the above mentioned applications. In the following sections 2–5 we give a basic analysis of differentiability and semismoothness properties of integral functions, while section 6 is devoted to applications. We apply our theoretical results to the option pricing problem. We show that the generalized Newton method for solving the no-arbitrage option price interpolation problem, proposed by Wang, Yin and Qi [33], has at least $\frac{4}{3}$-order convergence. We give conditions when this method has $\frac{3}{2}$-order or quadratic convergence. We also give a damped version of the generalized Newton method and show that it is globally convergent and the convergence order is at least $\frac{4}{3}$. Applications to shape-preserving interpolation and semi-infinite programs are also discussed.

2. Differentiability properties of integral functions

Consider an integral function $F : \mathbb{R}^n \to \mathbb{R}$, defined by

$$F(x) := \int_\Omega f(x,s)d\mu(s), \tag{2.1}$$

where $f : X \times \Omega \to \mathbb{R}$, X is an open subset of $\mathbb{R}^n$ and μ is a finite measure defined on a measurable space $(\Omega, \mathcal{F})$. We assume that for every $x \in X$, the function $f(x,\cdot)$ is $\mathcal{F}$-measurable and μ-integrable, i.e., $\int_\Omega |f(x,s)|d\mu(s) < +\infty$. This implies that the integral function $F(x)$ is well-defined and finite valued. Denote $f_s(\cdot) := f(\cdot, s)$ and let $x \in X$ be fixed. We say that a property holds for almost every (a.e.) $s \in \Omega$ if it holds for all $s \in \Omega$ except on a set of μ-measure zero. By $f_s'(x,h)$ we denote the directional derivative of $f_s(\cdot)$ at x in direction h. The following result is a consequence of the Lebesgue Dominated Convergence Theorem (e.g., [2, Proposition 5.108]).

Proposition 1. *Suppose that:* (i) *there exists an integrable function* $\kappa : \Omega \to \mathbb{R}_+$ *such that*

$$|f(x^1,s) - f(x^2,s)| \le \kappa(s)\|x^1 - x^2\| \text{ for all } x^1, x^2 \in X \text{ and a.e. } s \in \Omega, \tag{2.2}$$

(ii) *for* a.e. $s \in \Omega$, $f_s(\cdot)$ *is directionally differentiable at a point* $x \in X$. *Then* $F(\cdot)$ *is Lipschitz continuous on* X, *directionally differentiable at* x *and*

$$F'(x,h) = \int_\Omega f_s'(x,h)d\mu(s). \tag{2.3}$$

Condition (2.2) implies, of course, that for a.e. $s \in \Omega$ the function $f(\cdot, s)$ is Lipschitz continuous on X. Note that the results of the above proposition have a local nature and the set X can be reduced to a neighborhood of a considered point x. Note also that for locally Lipschitz continuous functions the concepts of Fréchet and Gâteaux directional differentiability do coincide (e.g., [25]). Hence, under the conditions of Proposition 1, we may simply discuss directional differentiability (or differentiability) of $F(\cdot)$ at x.

It immediately follows from (2.3) that $F'(x,h)$ is linear in h, i.e., $F(\cdot)$ is differentiable at x, if $f_s(\cdot)$ is differentiable at x for a.e. $s \in \Omega$. Moreover, we have the following result (e.g., [24, Chapter 2]).

Proposition 2. *Suppose that, in addition to the assumptions* (i) *and* (ii) *of Proposition* 1, $f_s'(x,\cdot)$ *is convex for* a.e. $s \in \Omega$. *Then* $F(\cdot)$ *is differentiable at* x *and*

$$\nabla F(x) = \int_\Omega \nabla f_s(x)d\mu(s) \tag{2.4}$$

if and only if $f_s(\cdot)$ *is differentiable at* x *for* a.e. $s \in \Omega$.

Suppose now that the function $f(x,s)$ is given as the maximum of a family of smooth functions $g_j : X \times \Omega \to \mathbb{R}$, $j \in J$. That is,

$$f(x,s) := \sup_{j\in J} g_j(x,s). \tag{2.5}$$

We make the following assumptions.

(A1) Ω is a compact metric space and $\mathcal{F}$ is its Borel sigma algebra.
(A2) For every $s \in \Omega$ and $j \in J$, the function $g_{js}(\cdot) := g_j(\cdot, s)$ is continuously differentiable on X.
(A3) $G_{js}(x) := \nabla g_{js}(x)$ is continuous on $X \times \Omega \times J$.
(A4) The set J is a compact metric space.

Of course, if the set J is finite, then the last assumption (A4) holds automatically.

By the Danskin theorem (e.g., [2, Theorem 4.13]) it follows from assumptions (A2)–(A4) that the max-function $f_s(\cdot)$, defined in (2.5), is directionally differentiable at every point $x \in X$ and

$$f'_s(x,h) = \sup_{j\in J^*_s(x)} h^T G_{js}(x). \tag{2.6}$$

Here $J^*_s(x)$ denotes the index set of active at $x \in X$ constraints,

$$J^*_s(x) := \arg\max_{j\in J} g_j(x,s). \tag{2.7}$$

Note that since it is assumed that the set J is compact and $g_j(x,s)$ is continuous in $j \in J$, the set $J^*_s(x)$ is nonempty and compact.

Let $V \subset X$ be a compact neighborhood of a point $\bar{x} \in X$. By the Mean Value theorem we have by assumptions (A2) and (A3) that for all $x^1, x^2 \in V$ and $\kappa_j(s) := \sup_{x\in V} \|G_{js}(x)\|$, the following holds

$$|g_j(x^1,s) - g_j(x^2,s)| \le \kappa_j(s)\|x^1 - x^2\|.$$

It follows that $f_s(\cdot)$ is Lipschitz continuous on V with the Lipschitz constant $\kappa(s) := \sup_{j\in J} \kappa_j(s)$. Because of the assumption (A3) and since the sets Ω and J are compact, the function $\kappa(s)$ is bounded on Ω, and hence is integrable.

We have by formula (2.6) that $f'_s(x,\cdot)$ is given by the maximum of linear functions and hence is convex. It also follows from (2.6) that $f_s(\cdot)$ is differentiable at x iff $G_{js}(x)$ is the same for all $j \in J^*_s(x)$, say $G_{js}(x) = G_s(x)$ for all $j \in J^*_s(x)$, in which case $\nabla f_s(x) = G_s(x)$. Consider the set

$$\Upsilon(x) := \left\{ s \in \Omega : \text{there exist } i, j \in J^*_s(x) \text{ such that } G_{is}(x) \neq G_{js}(x) \right\}. \tag{2.8}$$

The set $\Upsilon(x)$ is the set of those $s \in \Omega$ for which $f_s(\cdot)$ is not differentiable at x. The above discussion together with Propositions 1 and 2 imply the following result.

Theorem 1. *Consider the integral function $F(\cdot)$ defined in* (2.1) *with $f(\cdot,\cdot)$ defined in* (2.5). *Suppose that the assumptions* (A1)–(A4) *are satisfied. Then $F(\cdot)$ is locally Lipschitz continuous, directionally differentiable and formula* (2.3) *holds. Moreover, $F(\cdot)$ is differentiable at a point $x \in X$, and formula* (2.4) *holds, if and only if the set $\Upsilon(x)$ has μ-measure zero.*

Clearly, for $x \in X$, the set $\Upsilon(x)$ is included in the set of such $s \in \Omega$ that $J_s^*(x)$ is not a singleton. Therefore, it follows from the above theorem that if $J_s^*(x)$ is a singleton for a.e. $s \in \Omega$, then $F(\cdot)$ is differentiable at x.

Denote by X_F the set of such $x \in X$ that $F(\cdot)$ is differentiable at x. Since $F(\cdot)$ is locally Lipschitz continuous, we have by Rademacher's theorem that $F(\cdot)$ is differentiable almost everywhere, i.e., the set $X \setminus X_F$ has Lebesgue measure zero. We say that $F(\cdot)$ is *continuously differentiable* at a point $\bar{x} \in X$ if $\bar{x} \in X_F$ and

$$\lim_{X_F \ni x \to \bar{x}} \nabla F(x) = \nabla F(\bar{x}).$$

Note that it is assumed in the above that $F(\cdot)$ is differentiable at $\bar{x}$, but not necessarily at all x near $\bar{x}$.

Proposition 3. *Suppose that the set J is finite, assumptions* (A1)–(A3) *are satisfied and, for $\bar{x} \in X$, the set $J_s^*(\bar{x})$ is a singleton for* a.e. $s \in \Omega$. *Then $F(\cdot)$ is continuously differentiable at $\bar{x}$.*

Proof. By the above discussion we have that, under the assumptions (A1)–(A3) and since $J_s^*(\bar{x}) = \{j_s\}$ is a singleton for a.e. $s \in \Omega$, the integral function $F(\cdot)$ is differentiable at $\bar{x}$, i.e., $\bar{x} \in X_F$. Also since $G_{js}(\cdot)$ are continuous and J is finite, we have that if $J_s^*(\bar{x}) = \{j_s\}$ is a singleton for some $s \in \Omega$, then $J_s^*(x) = \{j_s\}$ for all x in a neighborhood (depending on s) of $\bar{x}$. For such x and s we have that $\nabla f_s(x) = G_{j_s s}(x)$. Since Ω is compact and for every $j \in J$, $G_{js}(x)$ is continuous on $X \times \Omega$, there exists a constant $L > 0$ such that $\|G_{js}(x)\| \le L$ for all $s \in \Omega$, x in a neighborhood of a point $\bar{x}$ and $j \in J$. Consequently, by the Lebesgue Dominated Convergence theorem we can take the following limit inside the integral

$$\lim_{X_F \ni x \to \bar{x}} \nabla F(x) = \int_\Omega \lim_{X_F \ni x \to \bar{x}} G_{j_s s}(x) d\mu(s).$$

Continuity of $\nabla F(x)$ then follows from the continuity of $G_{js}(\cdot)$. □

In the remainder of this section we discuss the following particular case of integral functions which is important for applications considered in section 6. Let $g : X \times \Omega \to \mathbb{R}$ and consider the integral function

$$F(x) := \int_\Omega [g(x, s)]_+ d\mu(s). \tag{2.9}$$

In particular, if $\Omega = [a, b]$ and $d\mu(s) = p(s)ds$, then the above integral function $F(\cdot)$ becomes the function defined in (1.1). Clearly the function $f_s(x) := [g(x, s)]_+$ can be written as the maximum of the function $g(x, s)$ and the identically zero function. The corresponding assumptions (A2) and (A3) take here the following form:

(A5) For every $s \in \Omega$ the function $g_s(\cdot) := g(\cdot, s)$ is continuously differentiable.
(A6) $G_s(x) := \nabla g_s(x)$ is continuous on $X \times \Omega$.

We have then that $f_s(\cdot)$ is directionally differentiable and

$$f'_s(x,h) = \begin{cases} [h^T G_s(x)]_+, & \text{if } s \in \Omega_0(x), \\ 0, & \text{if } s \in \Omega_-(x), \\ h^T G_s(x), & \text{if } s \in \Omega_+(x), \end{cases} \tag{2.10}$$

where

$$\begin{aligned} \Omega_0(x) &:= \{s \in \Omega : g_s(x) = 0\}, \\ \Omega_-(x) &:= \{s \in \Omega : g_s(x) < 0\}, \\ \Omega_+(x) &:= \{s \in \Omega : g_s(x) > 0\}. \end{aligned}$$

We have here that for a given $x \in X$ and any $s \in \Omega$, the function $f_s(\cdot)$ is differentiable at x iff either $s \in \Omega_-(x) \cup \Omega_+(x)$ or $s \in \Omega_0(x)$ and $G_s(x) = 0$. Therefore, the following result is a consequence of Propositions 1 and 2 and Theorem 1.

Corollary 1. *Consider the integral function $F(\cdot)$ defined in* (2.9). *Suppose that the assumptions* (A1),(A5) *and* (A6) *are satisfied. Then $F(\cdot)$ is locally Lipschitz continuous, directionally differentiable and formula* (2.3) *holds. Moreover, $F(\cdot)$ is differentiable at a point $x \in X$ if and only if*

$$\mu\big(\{s \in \Omega_0(x) : G_s(x) \neq 0\}\big) = 0, \tag{2.11}$$

in which case

$$\nabla F(x) = \int_{\Omega_+(x)} G_s(x) d\mu(s). \tag{2.12}$$

Note that because of the condition (2.11), the set $\Omega_+(x)$ can be replaced by the set $\Omega_+(x) \cup \Omega_0(x)$ without changing the value of the integral in the right hand side of (2.12). We have by Corollary 1 that, under the specified assumptions, the set X_F is formed by such $x \in X$ that condition (2.11) holds.

Let us denote by $\Omega_1(x,h)$ the set of such $s \in \Omega$ that $g_s(x+h)$ and $g_s(x)$ have the same sign, and by $\Omega_2(x,h) = \Omega \setminus \Omega_1(x,h)$ the set of such $s \in \Omega$ that $g_s(x+h)$ and $g_s(x)$ have different signs. (By definition we say that $g_s(x+h)$ and $g_s(x)$ have the same sign if one of these numbers is zero.)

Proposition 4. *Consider the integral function $F(\cdot)$ defined in* (2.9). *Suppose that the assumptions* (A1),(A5) *and* (A6) *hold and condition* (2.11) *is satisfied at a point $\bar{x} \in X$. Suppose, further, that the following condition holds*

$$\lim_{h \to 0} \mu\big(\Omega_2(\bar{x}, h)\big) = 0. \tag{2.13}$$

Then $F(\cdot)$ is continuously differentiable at $\bar{x}$.

Proof. By Corollary 1, formula (2.12) holds for all $x \in X_F$. Therefore, we have for all $x = \bar{x} + h \in X_F$ in a neighborhood of $\bar{x}$, that

$$\begin{aligned} \|\nabla F(\bar{x}+h) - \nabla F(\bar{x})\| \le & \int_\Omega \|G_s(\bar{x}+h) - G_s(\bar{x})\| d\mu(s) \\ & + \int_{\Omega_2(\bar{x},h)} (\|G_s(\bar{x}+h)\| + \|G_s(\bar{x})\|) d\mu(s). \end{aligned} \tag{2.14}$$

By the Lebesgue Dominated Convergence theorem we can take the following limit inside the integral, and hence

$$\lim_{x \to \bar{x}} \int_{\Omega} \|G_s(x) - G_s(\bar{x})\| d\mu(s) = \int_{\Omega} \lim_{x \to \bar{x}} \|G_s(x) - G_s(\bar{x})\| d\mu(s) = 0. \qquad (2.15)$$

We also have that the second integral in (2.14) is bounded by $2L\mu(\Omega_2(\bar{x}, h))$, where L is a constant bounding $\|G_s(x)\|$ for all $s \in \Omega$ and x in a neighborhood of $\bar{x}$. It follows then by (2.13) and (2.15) that

$$\lim_{X_F \ni x \to \bar{x}} \|\nabla F(x) - \nabla F(\bar{x})\| = 0, \qquad (2.16)$$

which proves that $F(\cdot)$ is continuously differentiable at $\bar{x}$. □

Let us note that condition (2.13) alone does not imply differentiability of $F(\cdot)$. Think, for example, about $g(x, s) \equiv g(x)$ independent of s and such that $g(\bar{x}) = 0$ while $\nabla g(\bar{x}) \neq 0$. In that case the set $\Omega_2(\bar{x}, h)$ is empty and hence condition (2.13) holds. On the other hand, $F(\cdot) = \mu(\Omega)([g(\cdot)]_+)$ is not differentiable at $\bar{x}$.

3. Semismoothness properties of integral functions

Consider the integral function $F(\cdot)$ defined in (2.1). Suppose that in addition to the assumptions (i) and (ii) of Proposition 1, $f_s(\cdot)$ is semismooth (Mifflin [11], see also [14],[18]) for a.e. $s \in \Omega$. That is, $f_s(\cdot)$ is directionally differentiable and

$$\left| f_s'(x + h, h) - f_s'(x, h) \right| \leq \varepsilon_s(h)\|h\|, \qquad (3.1)$$

where $\varepsilon_s(h) \to 0$ as $h \to 0$. Then

$$\begin{aligned} \left| F'(x + h, h) - F'(x, h) \right| &\leq \int_\Omega \left| f_s'(x + h, h) - f_s'(x, h) \right| d\mu(s) \\ &\leq \|h\| \int_\Omega \varepsilon_s(h) d\mu(s). \end{aligned} \qquad (3.2)$$

Suppose, further, that $\varepsilon_s(h)$ is dominated by an integrable function $\gamma(s)$ for all h in a neighborhood V of $0 \in \mathbb{R}^n$, i.e., $\sup_{h \in V} \varepsilon_s(h) \leq \gamma(s)$ for a.e. $s \in \Omega$ and $\int_\Omega \gamma(s) d\mu(s) < \infty$. Then by the Lebesgue Dominated Convergence theorem we can take the limit inside the integral, and hence

$$\lim_{h \to 0} \int_{\Omega} \varepsilon_s(h) d\mu(s) = \int_{\Omega} \lim_{h \to 0} \varepsilon_s(h) d\mu(s) = 0. \qquad (3.3)$$

It follows from (3.2) and (3.3) that $F(\cdot)$ is semismooth.

Proposition 5. *Suppose that $f(x, s)$ is the max-function, defined in* (2.5), *and that the assumptions* (A1)–(A4) *hold. Then the integral function $F(\cdot)$ is semismooth at every $x \in X$.*

Proof. Assumptions (A2)–(A4) imply that for every $s \in \Omega$ the function $f_s(\cdot)$ is semismooth, [11, Theorem 2]. Consider a point $x \in X$. By the above discussion we only have to verify that $\varepsilon_s(h)$ is dominated by an integrable function for all h in a neighborhood V of $0 \in \mathbb{R}^n$. We have that the constant

$$K := \sup \left\{ \|G_{js}(x+h)\| : j \in J, s \in \Omega, h \in V \right\}$$

is finite, provided that the neighborhood V is compact. Also by (2.6) we have that $|f_s'(x+h,h)| \le K\|h\|$ for all $s \in \Omega$, $j \in J$ and $h \in V$. It follows that $\varepsilon_s(h)$ is dominated by the constant function $\gamma(s) \equiv 2K$, and hence the proof is complete. □

Suppose now that for every $s \in \Omega$, $f_s(\cdot)$ is *p-order* semismooth [18], at a point $x \in X$, for $0 < p \le 1$. (Note that 1-order semismoothness was later called *strongly* semismooth [17].) That is, $\limsup_{h \to 0} c_s(h) < \infty$, where

$$c_s(h) := \frac{\left| f_s'(x+h,h) - f_s'(x,h) \right|}{\|h\|^{1+p}}, \quad h \ne 0.$$

We have that

$$\left| F'(x+h,h) - F'(x,h) \right| \le \|h\|^{1+p} \int_\Omega c_s(h) d\mu(s). \tag{3.4}$$

Therefore, in order to show that $F(\cdot)$ is p-order semismooth, at x, we need to verify that

$$\limsup_{h \to 0} \int_\Omega c_s(h) d\mu(s) < \infty. \tag{3.5}$$

As an example consider the integral function $F(\cdot)$ defined in (2.9), and suppose that the corresponding assumptions (A1), (A5) and (A6) hold. It follows by Proposition 5 that the corresponding integral function $F(\cdot)$ is semismooth.

Let us study now p-order semismoothness of $F(\cdot)$. We have that $c_s(h) \le q_s(h)$ with

$$q_s(h) := \begin{cases} \|h\|^{-p} \|G_s(x+h) - G_s(x)\|, & \text{if } s \in \Omega_1(x,h), \\ \|h\|^{-p} \max\{\|G_s(x+h)\|, \|G_s(x)\|\}, & \text{if } s \in \Omega_2(x,h). \end{cases} \tag{3.6}$$

Theorem 2. *Consider the integral function* $F(\cdot)$ *defined in* (2.9)*. Suppose that the assumptions* (A1), (A5) *and* (A6) *are satisfied. Then the integral function* $F(\cdot)$ *is semismooth. Suppose, further, that the following two conditions hold: there exists an integrable function* $\eta : \Omega \to \mathbb{R}_+$ *such that*

$$\left\| G_s(x^1) - G_s(x^2) \right\| \le \eta(s)\|x^1 - x^2\|^p, \quad \forall x^1, x^2 \in X \text{ and a.e. } s \in \Omega, \tag{3.7}$$

and

$$\mu(\Omega_2(x,h)) = O(\|h\|^p). \tag{3.8}$$

Then $F(\cdot)$ *is* p*-order semismooth at* x*.*

Proof. Semismoothness of $F(\cdot)$ follows by Proposition 5. In order to show p-order semismoothness of $F(\cdot)$ we need to verify (3.5). We have that

$$\int_{\Omega} c_s(h)d\mu(s) \leq \int_{\Omega_1(x,h)} q_s(h)d\mu(s) + \int_{\Omega_2(x,h)} q_s(h)d\mu(s). \tag{3.9}$$

By (3.6) and (3.7) we have

$$\int_{\Omega_1(x,h)} q_s(h)d\mu(s) \leq \int_{\Omega} \eta(s)d\mu(s) < \infty.$$

Since $\|G_s(x+h)\|$ and $\|G_s(x)\|$ are bounded for $s \in \Omega$, we have by (3.6) and (3.8) that the second integral in the right hand side of (3.9) is also bounded and the assertion follows. □

The above condition (3.7) holds, in particular, if $G_s(\cdot)$ is differentiable and $\nabla G_s(x)$ is continuous on $X \times \Omega$. Condition (3.8) is more delicate. It is clear that this condition implies condition (2.13). In the following we will analyze cases in which condition (3.8) holds. Before doing this, we remark that this condition for $p = 1$ may fail and the integral function F may not be strongly semismooth. An example of an integral function $F(\cdot)$, of the form (2.9), which is not strongly semismooth was given in Qi and Yin [20]. That example was simplified further by Ralph [21]. Ralph's example is as follows: F is defined by (2.9), $g(x,s) := s^2 - x$, $s \in [0,1]$, $x \in \mathbb{R}$ and $d\mu(s) = ds$. Here, for $x = 0$ and $h > 0$, $\Omega_2(0,h) = (0, \sqrt{h})$, and hence condition (3.8) does not hold for $p = 1$. And, indeed, the integral function is not strongly semismooth in this example.

We now assume that Ω is an interval $[a,b]$ of the real line and $d\mu(s) = p(s)ds$. For the sake of convenience, we first recall the concept of tensors and discuss their properties. We use $A_n^{(k)}$ to denote a k-th order n-dimensional tensor and use $A_{n,i_1\cdots i_k}^{(k)}$ to denote it elements. We assume that $i_l = 1, \cdots, n$ for $l = 1, \cdots, k$, and that $A_n^{(k)}$ is totally symmetric, i.e.,

$$A_{n,i_1\cdots i_k}^{(k)} = A_{n,j_1\cdots j_k}^{(k)}$$

if $j_1, \cdots, j_k$ is any reordering of $i_1, \cdots, i_k$. Let $x \in \mathbb{R}^n$. Denote

$$A_n^{(k)} x^k := \sum_{i_1,\cdots,i_k=1}^{n} A_{n,i_1\cdots i_k}^{(k)} x_{i_1} \cdots x_{i_k}.$$

Let $\|\cdot\|$ be the F-norm of the tensor space, that is,

$$\left\| A_n^{(k)} \right\| = \sqrt{\sum_{i_1,\cdots,i_k=1}^{n} (A_{n,i_1\cdots i_k}^{(k)})^2}.$$

Note that the above concept extends the F-norm concepts of matrices and vectors. The following proposition can be easily proved by mathematical induction.

Proposition 6. *Let $A_n^{(k)}$ be a k-th order n-dimensional tensor and $x \in \mathbb{R}^n$. Then*

$$\left| A_n^{(k)} x^k \right| \leq \left\| A_n^{(k)} \right\| \|x\|^k. \tag{3.10}$$

Since semismoothness of F at a point $\bar{x}$ is a local property which depends on only the status of F near $\bar{x}$, in order to study p-order semismoothness of F defined by (1.1), we first establish the following lemma which characterizes the perturbation property of the root $s(x)$ of $g(x,\cdot)=0$, where $g:\mathbb{R}^{n+1}\to\mathbb{R}$ has continuous m-order derivative. In this lemma we assume $s(x)$ (but maybe not unique) exists for any $x\in X$, where X is a certain open convex set containing $\bar{x}$. Note that X can be a neighborhood of $\bar{x}$.

Lemma 1. *Let $\bar{x}\in X$, where $X\subset\mathbb{R}^n$ is an open convex set. Suppose that for any $x\in X$, $g(x,\cdot)=0$ has at least one root on $[a,b]$, denoted as $s(x)$. Let $\bar{s}=s(\bar{x})$ and assume that the following condition holds:*

$$\begin{cases} \nabla_s^{(k)} g(\bar{x},\bar{s})=0, \quad k=1,2,\cdots,m-1,\\ \left|\nabla_s^{(m)} g(x,s)\right|>c,\ \forall x\in X,\ s\in[a,b], \end{cases} \tag{3.11}$$

where m is a positive integer and c is a positive number. Then $s(x)\to\bar{s}$ whenever $x\to\bar{x}$, and there exists a positive constant $L(\bar{x},\bar{s})$ such that

$$|s(x)-\bar{s}|^m\le L(\bar{x},\bar{s})\|x-\bar{x}\|, \tag{3.12}$$

and $\bar{s}$ is the unique root of $g(\bar{x},\cdot)=0$.

Proof. Write $x_{n+1}=s$, $z=(x,x_{n+1})$, $\bar{z}=(\bar{x},\bar{s})$ and

$$A^{(k)}_{n+1,i_1\cdots i_k}(z)=\left.\frac{\partial^k g}{\partial x_{i_1}\cdots\partial x_{i_k}}\right|_z,\ i_l=1,\cdots,n+1,\ l=1,\cdots,k. \tag{3.13}$$

It is clear that $A^{(k)}_{n+1}(z)$ is a totally symmetric tensor. By using (3.13), the standard Taylor theorem for multivariate functions [3] can be written as:

$$g(z)=g(\bar{z})+A^{(1)}_{n+1}(\bar{z})\Delta z+\frac{A^{(2)}_{n+1}(\bar{z})\Delta z^2}{2!}+\cdots+\frac{A^{(m-1)}_{n+1}(\bar{z})\Delta z^{m-1}}{(m-1)!}+\frac{A^{(m)}_{n+1}(z(\xi))\Delta z^m}{m!}, \tag{3.14}$$

where $\Delta z=z-\bar{z}$, and $z(\xi)$ is a point in the segment connecting z and $\bar{z}$. We obtain, by direct computation, that

$$\begin{aligned} A^{(k)}_{n+1}(z)\Delta z^k &= \nabla_s^{(k)} g(z)\Delta s^k + C_k^{k-1}\big(B^{(1)}_{n,k}(z)\Delta x\big)\Delta s^{k-1}\\ &\quad + C_k^{k-2}\big(B^{(2)}_{n,k}(z)\Delta x^2\big)\Delta s^{k-2}+\cdots\\ &\quad + C_k^{1}\big(B^{(k-1)}_{n,k}(z)\Delta x^{k-1}\big)\Delta s\\ &\quad + C_k^{0}\big(B^{(k)}_{n,k}(z)\Delta x^{k}\big),\ k=1,\cdots,m, \end{aligned} \tag{3.15}$$

where

$$C_k^j=\frac{k!}{(k-j)!\,j!},\quad j=0,\cdots,k$$

and $B^{(l)}_{n,k}(z)$ is an lth order n-dimensional tensor with components

$$B^{(l)}_{n,i_1\cdots i_l}(z)=\left.\frac{\partial^k g}{\partial x_{i_1}\cdots\partial x_{i_l}\partial s^{k-l}}\right|_z,\quad l=1,\cdots,k.$$

From (3.15) and the condition given, the formula (3.14) can be rewritten as follows:

$$\begin{aligned}
-\frac{1}{m!}\nabla_s^{(m)} g(z(\xi))\Delta s^m = &\Big[B_{n,1}^{(1)}(\bar{z})\Delta x + \frac{1}{2!}C_2^1\big(B_{n,2}^{(1)}(\bar{z})\Delta x\big)\Delta s + \cdots \\
&+\frac{1}{(m-1)!}C_{m-1}^{m-2}\big(B_{n,m-1}^{(1)}(\bar{z})\Delta x\big)\Delta s^{m-2} \\
&+C_m^{m-1}\big(B_{n,m}^{(1)}(z(\xi))\Delta x\big)\Delta s^{m-1}\Big] \\
&+\Big[\frac{1}{2!}C_2^0\big(B_{n,2}^{(2)}(\bar{z})\Delta x^2\big) + \frac{1}{3!}C_3^1\big(B_{n,3}^{(2)}(\bar{z})\Delta x^2\big)\Delta s + \cdots \\
&+\frac{1}{(m-1)!}C_{m-1}^{m-3}\big(B_{n,m-1}^{(2)}(\bar{z})\Delta x^2\big)\Delta s^{m-3} \\
&+\frac{1}{m!}C_m^{m-2}\big(B_{n,m}^{(2)}(z(\xi))\Delta x^2\big)\Delta s^{m-2}\Big] + \cdots \\
&+\Big[\frac{1}{(m-1)!}C_{m-1}^1\big(B_{n,m-1}^{(m-1)}(\bar{z})\Delta x^{m-1}\big) \\
&+\frac{1}{m!}C_m^1\big(B_{n,m}^{(m-1)}(z(\xi))\Delta x^{m-1}\big)\Delta s\Big] \\
&+\frac{1}{m!}C_m^0\big(B_{n,m}^{(m)}(z(\xi))\Delta x^m\big). \qquad (3.16)
\end{aligned}$$

Since $|\nabla_s^{(m)} g(x,s)| > c$, by Proposition 6, we have that

$$\begin{aligned}
|\Delta s|^m \le m!\big|\nabla_s^{(m)} g(z(\xi))\big|^{-1}\Big\{&\Big[\big\|B_{n,1}^{(1)}(\bar{z})\big\| + \frac{1}{2!}C_2^1\big\|B_{n,2}^{(1)}(\bar{z})\big\||\Delta s| + \cdots \\
&+\frac{1}{(m-1)!}C_{m-1}^{m-2}\big\|B_{n,m-1}^{(1)}(\bar{z})\big\||\Delta s|^{m-2} + C_m^{m-1}\big\|B_{n,m}^{(1)}(z(\xi))\big\||\Delta s|^{m-1}\Big] \\
&+\Big[\frac{1}{2!}C_2^0\big\|B_{n,2}^{(2)}(\bar{z})\big\| + \cdots + \frac{1}{(m-1)!}C_{m-1}^{m-3}\big\|B_{n,m-1}^{(2)}(\bar{z})\big\||\Delta s|^{m-3} \\
&+\frac{1}{m!}C_m^{m-2}\big\|B_{n,m}^{(2)}(z(\xi))\big\||\Delta s|^{m-2}\Big]\|\Delta x\| + \cdots \\
&+\Big[\frac{1}{(m-1)!}C_{m-1}^1\big\|B_{n,m-1}^{(m-1)}(\bar{z})\big\| + \frac{1}{m!}C_m^1\big\|B_{n,m}^{(m-1)}(z(\xi))\big\||\Delta s|\Big]\|\Delta x\|^{m-2} \\
&+\frac{1}{m!}C_m^0\big\|B_{n,m}^{(m)}(z(\xi))\big\|\|\Delta x\|^{m-1}\Big\}\|\Delta x\|. \qquad (3.17)
\end{aligned}$$

It is clear that all coefficients of $\|\Delta x\|^k$ and $|\Delta s|^k$, $(k = 1, 2, \cdots, m-1)$ are bounded. Consequently, from the fact that $\big|\nabla_s^{(m)} g(z(\xi))\big|^{-1}$ is bounded, there exists a positive constant $\bar{L}(\bar{x}, \bar{s})$ such that

$$|\Delta s|^m \le \bar{L}(\bar{x}, \bar{s})\|\Delta x\|.$$

This shows that $s(x) \to \bar{s}$ whenever $x \to \bar{x}$, and $\bar{s}$ is the unique root of $g(\bar{x}, \cdot) = 0$. The proof is completed. □

Remark. The second item in (3.11), i.e., $\left|\nabla_s^{(m)} g(x,s)\right| > c, \forall x \in X, s \in [a,b]$ characterizes the uniform sharpness of the curve family $\{g(x,\cdot) : x \in X\}$ on interval $[a,b]$ in a sense. For instance, in the case $m = 2$, since the value of $\nabla_s^{(2)} g(x,s)$ characterizes the "convexity degree" of curve $g(x,\cdot)$, the second item in (3.11) shows the curve family $\{g(x,\cdot) : x \in X\}$ has at least "convexity degree" c on interval $[a,b]$.

Now, we discuss the p-order semismoothness property of the integral function $F(\cdot)$ defined by (1.1). We give a sufficient condition under which F is p-order semismooth at a given point $\bar{x}$. Note, this condition is easier to check than (3.8).

Theorem 3. *Consider the integral function $F(\cdot)$, defined by* (1.1)*, at a point $\bar{x} \in \mathbb{R}^n$. Suppose that:* (i) *$g_s(x) = g(x,s)$ is m-order continuously differentiable, jointly in x and s, where m is a certain positive integer,* (ii) *$\Omega_0(\bar{x})$ is a singleton set, denoted as $\{\bar{s}\}$,* (iii) *$\bar{s}$ is an m-th order root of $g(\bar{x},\cdot) = 0$, that is,*

$$\begin{cases} \nabla_s^{(k)} g(\bar{x},\bar{s}) = 0, & k = 1,2,\cdots,m-1, \\ \nabla_s^{(m)} g(\bar{x},\bar{s}) \neq 0, \end{cases} \tag{3.18}$$

(iv) *there exists an integrable function $\eta(s)$ such that*

$$\|G_s(x+h) - G_s(x)\| \leq \eta(s)\|h\|^{\frac{1}{m}}.$$

Then $F(\cdot)$ is $\frac{1}{m}$-order semismooth at $\bar{x}$.

Proof. By Theorem 2, we only need to check if (3.8) holds. Since $\nabla_s^{(m)} g(x,s)$ is continuous at $(\bar{x},\bar{s})$ and $\bar{d} := \nabla_s^{(m)} g(\bar{x},\bar{s}) \neq 0$, there exist a neighborhood U of $\bar{x}$ and a subinterval $[\bar{s}-\delta, \bar{s}+\delta]$ such that

$$\min_{x \in U, s \in [\bar{s}-\delta, \bar{s}+\delta]} \left|\nabla_s^{(m)} g(x,s)\right| > \left|\frac{\bar{d}}{2}\right|.$$

Take any $h \in \mathbb{R}^n$ with $\bar{x}+h \in U$. There are four cases: Case (i) $\bar{d} < 0$ and m is an even number; Case (ii) $\bar{d} < 0$ and m is an odd number; Case (iii) $\bar{d} > 0$ and m is an even number; Case (iv) $\bar{d} > 0$ and m is an odd number. We now only discuss cases (i) and (ii), the proof for the other two cases is similar.

Case (i). In this case, it is not difficult to know that

$$g(\bar{x},s) \leq g(\bar{x},\bar{s}) = 0, \ \forall s \in [a,b]. \tag{3.19}$$

We assume, without loss of generality, that $a < \bar{s} < b$. If $\Omega_0(\bar{x}+h) = \emptyset$, then $\Omega_-(\bar{x}+h) = [a,b]$. Hence, $\Omega_2(\bar{x}+h) = \emptyset$. Now, we assume that $\Omega_0(\bar{x}+h) \neq \emptyset$. Write

$$\hat{s}(h) := \sup\{s : s \in \Omega_0(\bar{x}+h)\},$$

$$\tilde{s}(h) := \inf\{s : s \in \Omega_0(\bar{x}+h)\}.$$

We know, because of the continuity of g, that $\Omega_0(\bar{x}+h)$ is a closed set. Consequently, $\hat{s}(h), \tilde{s}(h) \in \Omega_0(\bar{x}+h)$. We assume, shrinking U if necessary, that $s(x) \in [\bar{s}-\delta, \bar{s}+\delta]$

for all $s(x) \in \Omega_0(x)$ and $x \in U$. We also conclude that for any $s \in (\hat{s}(h), b]$ or $s \in [a, \tilde{s}(h))$,

$$g(\bar{x} + h, s) < 0.$$

In fact, if there exists, without loss of generality, an $s' \in (\hat{s}(h), b]$ such that

$$g(\bar{x} + h, s') > 0,$$

then, since $g(\bar{x}, b) < 0$, $g(\bar{x} + h, b) < 0$ whenever $\|h\|$ is small enough. By the Mean-Value Theorem, there exists an s'' on the open line segment from s' to b such that

$$g(\bar{x} + h, s'') = 0,$$

contradicting the definition of $\hat{s}(h)$. So,

$$\Omega_2(\bar{x}, h) \subseteq [\tilde{s}(h), \hat{s}(h)].$$

Further, since $\hat{s}(h), \tilde{s}(h) \in [\bar{s} - \delta, \bar{s} + \delta]$, applying Lemma 1 to the case that $X = U$ and $[a, b] = [\bar{s} - \delta, \bar{s} + \delta]$, we have

$$\mu(\Omega_2(\bar{x}, h)) \leq \Delta\hat{s}(h) + \Delta\tilde{s}(h) = O(\|h\|^{\frac{1}{m}}),$$

where $\Delta\hat{s}(h) = |\hat{s}(h) - \bar{s}|$ and $\Delta\tilde{s}(h) = |\tilde{s}(h) - \bar{s}|$.

Case (ii). In this case, we have

$$g(\bar{x}, s) > 0, \ \forall s \in [a, \bar{s}),$$
$$g(\bar{x}, s) < 0, \ \forall s \in (\bar{s}, b].$$

That is, $\Omega_+(\bar{x}) = [a, \bar{s})$ and $\Omega_-(\bar{x}) = (\bar{s}, b]$. Hence, for any $h \in \mathbb{R}^n$, $\Omega_0(\bar{x} + h) \neq \emptyset$, whenever $\|h\|$ is small enough, and

$$g(\bar{x} + h, s) > 0, \ \forall s \in [a, \tilde{s}(h)),$$
$$g(\bar{x} + h, s) < 0, \ \forall s \in (\hat{s}(h), b].$$

So,

$$\Omega_2(\bar{x}, h) \subseteq [\tilde{s}(h), \bar{s}] \bigcup [\bar{s}, \hat{s}(h)].$$

Similarly to case (i), we have also

$$\mu(\Omega_2(\bar{x}, h)) \leq O(\|h\|^{\frac{1}{m}}).$$

The proof is completed. □

Remark. In Theorem 3, the condition (ii) is not essential since the sum of a finite number of $p-$order semismooth functions is still a $p-$order semismooth function. Suppose the set $\Omega_0(\bar{x}) = \{s \in [a, b] : g_s(\bar{x}) = 0\}$ is finite and the highest order of roots is m, then by separating $[a, b]$ into a certain number of subintervals such that every subinterval contains only a single root of $g_s(\bar{x}) = 0$, we know that F is the sum of the corresponding integral functions defined on these subintervals. By Theorem 3, these integral functions are at least $\frac{1}{m}$-order semismooth at $\bar{x}$ and so is F.

With the above results and the results of [19] at hand, we may identify whether a particular integral function F is smooth, (p-order) semismooth, piecewise smooth, almost smooth or none of the above.

4. SC^1 properties of integral functions

A smooth function $F(\cdot)$ is called an LC^1 (SC^1) function if $\nabla F(\cdot)$ is locally Lipschitz continuous (semismooth) [7],[12],[15]. Hence, we may establish SC^1 properties of $F(\cdot)$ based upon the results of the last two sections.

Proposition 7. *Consider the integral function $F(\cdot)$ defined in* (2.9). *Suppose that: the assumptions* (A1),(A5),(A6) *and condition* (3.7) *hold, condition* (2.11) *is satisfied at every $x \in X$, and there exists a constant $K > 0$ such that*

$$\mu\big(\Omega_2(x,h)\big) \le K\|h\|, \quad \forall x, x+h \in X. \tag{4.1}$$

Then $F(\cdot)$ is differentiable and $\nabla F(\cdot)$ is Lipschitz continuous on X.

Proof. By Corollary 1, assumptions (A1),(A5) and (A6) and condition (2.11) imply that $F(\cdot)$ is differentiable on X and formula (2.12) holds. By using (2.14) and (3.7) we obtain that for $x, x+h \in X$,

$$\|\nabla F(x+h) - \nabla F(x)\| \le \|h\| \int_\Omega \eta(s)d\mu(s) + 2L\mu(\Omega_2(x,h)),$$

where L is such that $\|G_s(x+h)\| \le L$ for all $s \in \Omega$ and h in a neighborhood of 0. Together with (4.1) this completes the proof. □

Note that condition (4.1) implies condition (3.8) "uniformly in x". Note also that one can formulate a local version of the above theorem by assuming that conditions (3.7) and (4.1) hold locally.

As an example, consider $g(x,s) := v(x) + s^T u(x)$, where $v : X \to \mathbb{R}$ and $u : X \to \mathbb{R}^p$, and suppose that μ is a Borel measure on $\mathbb{R}^p$ having density $p(\cdot)$ (i.e., $d\mu(s) = p(s)ds$) with support $\Omega \subset \mathbb{R}^p$. In that case

$$F(x) = \int_\Omega \Big[v(x) + s^T u(x)\Big]_+ p(s)ds. \tag{4.2}$$

Proposition 8. *Consider the integral function $F(\cdot)$ defined in* (4.2). *Suppose that:* (i) *the functions $v(\cdot)$ and $u(\cdot)$ are continuously differentiable, and* (ii) *the set Ω is bounded. Denote $V(x) := \nabla v(x)$ and $U(x) := \nabla u(x)$. Then $F(\cdot)$ is differentiable at a point $x \in X$ and*

$$\nabla F(x) = \int_{\Omega_+(x)} \big[V(x) + U(x)s\big]p(s)ds, \tag{4.3}$$

if and only if either $|v(x)| + \|u(x)\| \ne 0$, or $v(x) = 0$, $u(x) = 0$ and $V(x) = 0$, $U(x) = 0$. Suppose that, in addition to the above assumptions (i) *and* (ii), *$V(\cdot)$ and $U(\cdot)$ are locally Lipschitz continuous and the density function $p(\cdot)$ is bounded. Then $F(\cdot)$ is strongly semismooth, and if, moreover, $F(\cdot)$ is differentiable on X, then $H(\cdot) := \nabla F(\cdot)$ is locally Lipschitz continuous.*

Proof. We can apply Corollary 1 to the present case. Since Ω is a support set of a measure, it is closed. Therefore, since it is assumed that Ω is bounded, it is compact. The assumptions (A1),(A5) and (A6) then hold. We have here that $\Omega_0(x) = \Omega \cap L_0(x)$, where

$$L_0(x) := \left\{ s \in \mathbb{R}^p : v(x) + s^T u(x) = 0 \right\}$$

is an affine subspace of $\mathbb{R}^p$. Consequently, $L_0(x)$ has Lebesgue measure zero unless $L_0(x)$ coincides with the whole space $\mathbb{R}^p$. It follows that if $u(x) \neq 0$, then $\mu(\Omega_0(x)) = 0$, and hence the corresponding integral function $F(\cdot)$ is differentiable at x and formula (2.12) holds. If $v(x) = 0$ and $u(x) = 0$, then clearly $L_0(x) = \mathbb{R}^p$ and $\Omega_0(x) = \Omega$. In that case $F(\cdot)$ is differentiable at x iff the set $\Omega \setminus L_1(x)$, where $L_1(x) := \{s \in \mathbb{R}^p : V(x) + U(x)s = 0\}$, has μ-measure zero. Again, since $L_1(x)$ is an affine (may be empty) subspace of $\mathbb{R}^p$ it has Lebesgue measure zero if $\|V(x)\| + \|U(x\| \neq 0$, in which case $\mu(\Omega \setminus L_1(x)) = \mu(\Omega) \neq 0$, or $L_1(x) = \mathbb{R}^p$ if $V(x) = 0$ and $U(x) = 0$, in which case $\Omega \setminus L_1(x) = \emptyset$.

Suppose, further, that $V(\cdot)$ and $U(\cdot)$ are locally Lipschitz continuous and the density function $p(\cdot)$ is bounded. Then condition (3.7) holds (locally) with $\eta(s) := \ell_V + \ell_U \|s\|$, where ℓ_V and ℓ_U are local Lipschitz constants of V and U. Moreover, since Ω is bounded we have that the Lebesgue measure of the set $\Omega_2(x, h)$ is of order $O(\|h\|)$ if $u(x) \neq 0$. Indeed, we have that the set $\Omega_2(x, h)$ is given by such $s \in \Omega$ that $a_1 + s^T b_1$ and $a_2 + s^T b_2$ have different signs, where $a_1 := v(x)$, $b_1 := u(x)$, $a_2 := v(x+h)$ and $b_2 := u(x+h)$. Suppose that $u(x) \neq 0$, i.e., $b_1 \neq 0$. Since, by continuity of $u(\cdot)$, we only need to consider b_2 sufficiently close to b_1, we can assume that $b_2 \neq 0$ as well. If $b_1 = b_2$, then the set $\Omega_2(x, h)$ is formed by $s \in \Omega$ which lie between two hyperplanes defined by the equations $a_1 + s^T b_1 = 0$ and $a_2 + s^T b_1 = 0$, and since Ω is bounded it follows in that case that $\Omega_2(x, h) \leq c\|h\|$ for some constant c independent of h. So let $b_1 \neq b_2$. If $p = 1$, i.e., s is one dimensional, then the assertion clearly holds. So suppose that $p \geq 2$. Let $\bar{s}$ be a solution of the equations $a_1 + s^T b_1 = 0$ and $a_2 + s^T b_2 = 0$. By making the transformation $s \mapsto \bar{s} + s$ we can assume that $a_1 = a_2 = 0$. The angle between vector b_1 and b_2 is of order $O(\|b_2 - b_1\|)$, and $\|b_2 - b_1\| = O(\|h\|)$ since $u(\cdot)$ is continuously differentiable. Again since Ω is bounded this implies that $\Omega_2(x, h) = O(\|h\|)$.

Since Ω is bounded, we have that if $u(x) = 0$ and $v(x) \neq 0$, then $\Omega_2(x, h)$ is empty for all h sufficiently close to 0. If $u(x) = 0$ and $v(x) = 0$, then the set $\Omega_2(x, h)$ is empty. In any case we obtain that condition (3.8) holds, and hence $F(\cdot)$ is strongly semismooth at x by Theorem 2.

In fact, we have then that condition (4.1) holds locally, and hence if $F(\cdot)$ is differentiable, then $H(\cdot)$ is locally Lipschitz continuous by Proposition 7. □

Unfortunately, it is not possible to take second order derivatives of $g_s(\cdot)$ inside the integral. Consider, for example, $g(x, s) := x - s$, with $x, s \in \mathbb{R}$, and $\Omega := [0, 1]$. Then, for $x \in [0, 1]$, we have that $F(x) = x^2/2$ while $\partial^2 g(x, s)/\partial x^2 \equiv 0$.

If we assume that, for some $x \in X$, the μ-measure of the set $\{s \in \Omega : g(x, s) = 0\}$ is zero, then by Corollary 1 we have that $F(\cdot)$ is differentiable at x and

$$\nabla F(x) = \int_{g(x,s) \geq 0} G_s(x) d\mu(s). \tag{4.4}$$

General formulas for derivatives of functions of the form given in the right hand side of (4.4) are quite involved (see [31],[32]). Therefore, we restrict our discussion of the SC^1 property of $F(\cdot)$ to the case (1.1) described in the beginning of the paper: $x \in \mathbb{R}^n$, $s \in \mathbb{R}$, $\Omega := [a, b]$ and $d\mu(s) = p(s)ds$. Actually, this case is very useful in practice. In the next section, we discuss this case in detail.

5. The one dimensional case

In this section we discuss properties of the integral function $F : \mathbb{R}^n \to \mathbb{R}$ defined in (1.1). That is, we assume that Ω is an interval $[a, b]$ of the real line and $\mu(s) = p(s)ds$. We consider two special cases of (1.1).

The first case is $g(x, s) := v(x) + su(x)$, where $v, u : \mathbb{R}^n \to \mathbb{R}$ are assumed to be continuously differentiable. Then

$$F(x) = \int_a^b [v(x) + su(x)]_+ p(s)ds. \tag{5.1}$$

Of course, the above is a particular case of (4.2) with $\Omega := [a, b] \subset \mathbb{R}$. Denote $V(x) := \nabla v(x)$, $U(x) := \nabla u(x)$ and $c(x) := -\frac{v(x)}{u(x)}$ if $u(x) \neq 0$. By Proposition 8 we have that the function $F(\cdot)$ is differentiable at a point x iff either $|v(x)| + |u(x)| \neq 0$ or $u(x) = v(x) = 0$ and $V(x) = U(x) = 0$. If $F(\cdot)$ is differentiable at x, then for $H(x) := \nabla F(x)$ formula (4.3) can be written as follows. If $u(x) > 0$ and $c(x) \leq b$, then

$$H(x) = \int_{\max\{a,c(x)\}}^b [V(x) + sU(x)]p(s)ds. \tag{5.2}$$

If $u(x) < 0$ and $c(x) \geq a$, then

$$H(x) = \int_a^{\min\{b,c(x)\}} [V(x) + sU(x)]p(s)ds. \tag{5.3}$$

If $u(x) = 0$ and $v(x) > 0$, then

$$H(x) = \int_a^b V(x)p(s)ds. \tag{5.4}$$

If $u(x) > 0$ and $c(x) > b$, or if $u(x) < 0$ and $c(x) < a$, or if $u(x) = 0$ and $v(x) < 0$, or if $u(x) = v(x) = 0$ and $U(x) = V(x) = 0$, then

$$H(x) = 0. \tag{5.5}$$

Clearly in all above cases $H(x)$ is continuous at x.

Proposition 9. *If $u(\cdot)$ and $v(\cdot)$ are twice continuously differentiable and $p(\cdot)$ is bounded, then $F(\cdot)$ is a strongly semismooth function.*

Proof. Clearly assumptions (A1),(A5),(A6) and (3.7) of Theorem 2 are satisfied here. Therefore, we only need to verify condition (3.8). We have that if $u(x) = 0$, then the set $\Omega_2(x, h)$ is empty for all h sufficiently close to 0, and if $u(x) \neq 0$, then the Lebesgue measure of the set $\Omega_2(x, h)$ is of order $O(|h|)$. Since $p(\cdot)$ is bounded, condition (3.8) follows. □

In fact, if $u(x) \neq 0$, then in order to verify strong semismoothness of $F(\cdot)$ at x it suffices to assume in the above proposition boundedness of $p(\cdot)$ in a neighborhood of $c(x)$.

We now further discuss the SC^1 property of $F(\cdot)$. This will be useful in investigation of semi-infinite programming problems. Recall that $H(\cdot)$ is semismooth at x if for any $h \in \mathbb{R}^n$, the limit $\lim_{t\downarrow 0} H'(x + th, h)$ exists (cf., [14],[18]). If this holds, then this limit is equal to $H'(x, h)$. Note that if $v(\cdot)$ and $u(\cdot)$ are differentiable, then, of course, $c(\cdot) := -v(\cdot)/u(\cdot)$ is also differentiable and $c'(x, h) = h^T \nabla c(x)$ provided that $u(x) \neq 0$.

Theorem 4. *Suppose that the functions $u(\cdot)$ and $v(\cdot)$ are twice continuously differentiable. If $|u(x)| + |v(x)| \neq 0$, then $H(\cdot)$ is semismooth at x, i.e., $F(\cdot)$ is SC^1 at x. If, moreover, $U(\cdot)$ and $V(\cdot)$ are strongly semismooth at x, then $H(\cdot)$ is strongly semismooth at x. Furthermore, the following cases hold:*

(A) *If either $u(x) > 0$ and $c(x) < a$, or $u(x) < 0$ and $c(x) > b$, or $u(x) = 0$ and $v(x) > 0$, then H is continuously differentiable at x and*

$$\nabla H(x) = \nabla U(x) \int_a^b sp(s)ds + \nabla V(x) \int_a^b p(s)ds. \tag{5.6}$$

(B) *If either $u(x) > 0$ and $c(x) > b$, or $u(x) < 0$ and $c(x) < a$, or $u(x) = 0$ and $v(x) < 0$, then H is continuously differentiable at x and*

$$\nabla H(x) = 0. \tag{5.7}$$

(C) *If $u(x) > 0$ and $a < c(x) < b$, then H is continuously differentiable at x and*

$$\begin{aligned}\nabla H(x) = \nabla U(x) \textstyle\int_{c(x)}^b sp(s)ds + \nabla V(x) \int_{c(x)}^b p(s)ds \\ -\nabla c(x)[c(x)U(x) + V(x)]^T p(c(x)).\end{aligned} \tag{5.8}$$

(D) *If $u(x) < 0$ and $a < c(x) < b$, then H is continuously differentiable at x and*

$$\begin{aligned}\nabla H(x) = \nabla U(x) \textstyle\int_a^{c(x)} sp(s)ds + \nabla V(x) \int_a^{c(x)} p(s)ds \\ +\nabla c(x)[c(x)U(x) + V(x)]^T p(c(x)).\end{aligned} \tag{5.9}$$

(E) *If $u(x) > 0$ and $c(x) = a$, then for all h,*

$$\begin{aligned}H'(x, h) = U'(x, h) \textstyle\int_a^b sp(s)ds + V'(x, h) \int_a^b p(s)ds \\ -[c'(x, h)]_+[aU(x) + V(x)]^T p(a).\end{aligned} \tag{5.10}$$

(F) *If $u(x) > 0$ and $c(x) = b$, then for all h,*

$$H'(x, h) = [-c'(x, h)]_+[bU(x) + V(x)]^T p(b). \tag{5.11}$$

(G) *If* $u(x) < 0$ *and* $c(x) = a$*, then for all* h*,*

$$H'(x, h) = [c'(x, h)]_+ [aU(x) + V(x)]^T p(a). \tag{5.12}$$

(H) *If* $u(x) < 0$ *and* $c(x) = b$*, then for all* h*,*

$$\begin{aligned} H'(x, h) = U'(x, h) \textstyle\int_a^b sp(s)ds + V'(x, h) \int_a^b p(s)ds \\ +[c'(x, h)]_+ [bU(x) + V(x)]^T p(b). \end{aligned} \tag{5.13}$$

Proof. In cases (A),(B),(C) or (D) the set of points x defined by the corresponding constraints forms an open subset of $\mathbb{R}^n$. Then continuous differentiability of $H(\cdot)$, in these four cases, and formulas (5.6)–(5.9) follow by the twice continuous differentiability of $u(\cdot)$ and $v(\cdot)$, formulas (5.2)–(5.5) and the Newton-Leibniz integration formula.

Consider case (E). If $c'(x, h) > 0$, we have $c(x + th) > a$ for $t > 0$ small enough. By (5.8), we have

$$\begin{aligned} \lim_{t \downarrow 0} H'(x + th, h) &= \lim_{t \downarrow 0} \nabla H(x + th)^T h \\ &= \lim_{t \downarrow 0} \Bigg[U'(x + th, h) \int_{c(x+th)}^b sp(s)ds \\ &\quad + V'(x + th, h) \int_{c(x+th)}^b p(s)ds \\ &\quad - c'(x + th, h)[c(x + th)U(x) + V(x)]^T p(c(x + th)) \Bigg] \\ &= U'(x, h) \int_a^b sp(s)ds + V'(x, h) \int_a^b p(s)ds \\ &\quad - [c'(x, h)]_+ [aU(x) + V(x)]^T p(a). \end{aligned}$$

If $c'(x, h) < 0$, we have $c(x + th) < a$ for small $t > 0$. By (5.6), we have

$$\begin{aligned} \lim_{t \downarrow 0} H'(x + th, h) &= \lim_{t \downarrow 0} \nabla H(x + th)^T h \\ &= \lim_{t \downarrow 0} \Bigg[U'(x + th, h) \int_{c(x+th)}^b sp(s)ds \\ &\quad + V'(x + th, h) \int_{c(x+th)}^b p(s)ds \Bigg] \\ &= U'(x, h) \int_a^b sp(s)ds + V'(x, h) \int_a^b p(s)ds \\ &\quad - [c'(x, h)]_+ [aU(x) + V(x)]^T p(a). \end{aligned}$$

If $c'(x, h) = 0$, we have

$$\begin{aligned} H'(x + th, h) &= \nabla H(x + th)^T h \\ &= U'(x + th, h) \int_{c(x+th)}^b sp(s)ds + V'(x + th, h) \int_{c(x+th)}^b p(s) \\ &\quad - c'(x + th, h)[c(x + th)U(x) + V(x)]^T p(c(x + th)) \end{aligned}$$

if $c(x+th) > a$ and

$$H'(x+th,h) = \nabla H(x+th)^T h$$
$$= U'(x+th,h)\int_{c(x+th)}^{b} sp(s)ds + V'(x+th,h)\int_{c(x+th)}^{b} p(s)$$

if $c(x+th) \le a$. Taking limit, we still have

$$\lim_{t\downarrow 0} H'(x+th,h) = U'(x,h)\int_a^b sp(s)ds + V'(x,h)\int_a^b p(s)ds$$
$$-[c'(x,h)]_+[aU(x)+V(x)]^T p(a).$$

This shows that H is semismooth at x. This also prove (5.10). Similarly, we can prove (5.11)–(5.13) and that $H(\cdot)$ is semismooth at x in (F)-(H). If $U(\cdot)$ and $V(\cdot)$ are strongly semismooth, we may prove strong semismoothness of H at x similarly. □

We now consider the second case of (1.1).

Theorem 5. *Consider the integral function $F(\cdot)$ defined by* (1.1)*, at a point $\bar{x} \in \mathbb{R}^n$. Suppose that:* (i) *$g_s(x) = g(x,s)$ is continuously differentiable, jointly in x and s,* (ii) *$g_s(\cdot)$ is twice differentiable and $\nabla^2 g_s(x)$ is continuous on $\mathbb{R}^n \times [a,b]$,* (iii) *the set $\Omega_0(\bar{x}) = \{s \in [a,b] : g_s(\bar{x}) = 0\}$ is finite, and* (iv) *$\nabla_s g(\bar{x},s) \ne 0$, for all $s \in \Omega_0(\bar{x})$.*

Then $F(\cdot)$ is differentiable and $H(\cdot) := \nabla F(\cdot)$ is Lipschitz continuous in a neighborhood of $\bar{x}$. If, moreover, $p(\cdot)$ is continuous in a neighborhood of every $s \in \Omega_0(\bar{x})$, then $H(\cdot)$ is semismooth at $\bar{x}$. If furthermore $a, b \notin \Omega_0(\bar{x})$, then $F(\cdot)$ is twice continuously differentiable in a neighborhood of $\bar{x}$.

Proof. Consider first the case where $a, b \notin \Omega_0(\bar{x})$. By the assumption (iii) we have that the set $\Omega_0(\bar{x})$ is finite, say $\Omega_0(\bar{x}) = \{\alpha_1, ..., \alpha_m\}$ with $a < \alpha_1 < ... < \alpha_m < b$. Because of the assumption (iv) we have that $\nabla_s g(\bar{x}, \alpha_i) \ne 0$, $i = 1, ..., m$. Suppose that $\nabla_s g(\bar{x}, \alpha_1) > 0$ (if $\nabla_s g(\bar{x}, \alpha_1) < 0$ the analysis is similar). Then $g(\bar{x}, a) < 0$, the function $g(\bar{x}, \cdot)$ changes signs at $\alpha_1, ..., \alpha_m$, and by the Implicit Function Theorem, for all x sufficiently close to $\bar{x}$ we have that $\Omega_0(x) = \{\alpha_1(x), ..., \alpha_m(x)\}$ with $\alpha_i(\cdot)$ being continuously differentiable in a neighborhood of $\bar{x}$ and

$$\nabla \alpha_i(x) = -[\nabla_s g(x, \alpha_i(x))]^{-1} G(x, \alpha_i(x)). \tag{5.14}$$

We also have that for all x near $\bar{x}$,

$$F(x) = \int_{\alpha_1(x)}^{\alpha_2(x)} g(x,s)p(s)ds + \int_{\alpha_3(x)}^{\alpha_4(x)} g(x,s)p(s)ds + ..., \tag{5.15}$$

$F(\cdot)$ is differentiable at x and and

$$H(x) = \int_{\alpha_1(x)}^{\alpha_2(x)} G(x,s)p(s)ds + \int_{\alpha_3(x)}^{\alpha_4(x)} G(x,s)p(s)ds + ... \tag{5.16}$$

Since $\alpha_i(x)$ are Lipschitz continuous and $G(x,s)$ are bounded for all x in a neighborhood of $\bar{x}$ and all $s \in [a,b]$, it follows from (5.16) that $H(\cdot)$ is Lipschitz continuous in a neighborhood of $\bar{x}$. (If $\alpha_1 = a$ or $\alpha_m = b$, Lipschitz continuity of $H(\cdot)$ still holds basically by the same arguments.)

Suppose, further, that the function $p(\cdot)$ is continuous near every α_i. Then for all x in a neighborhood of $\bar{x}$, the second order derivatives $\nabla^2 F(x)$ can be calculated by using formula

$$\nabla_x \left[\int_{\alpha_i(x)}^{\alpha_{i+1}(x)} G(x,s)p(s)ds\right] = -\frac{p(\alpha_{i+1}(x))}{\nabla_s g(x,\alpha_{i+1}(x))} G(x,\alpha_{i+1}(x))G(x,\alpha_{i+1}(x))^T + \frac{p(\alpha_i(x))}{\nabla_s g(x,\alpha_i(x))} G(x,\alpha_i(x))G(x,\alpha_i(x))^T + \int_{\alpha_i(x)}^{\alpha_{i+1}(x)} \nabla G_s(x)p(s)ds, \tag{5.17}$$

which follows by the chain rule of differentiation from (5.14) and the Newton-Leibniz formula. We obtain twice continuous differentiability of $F(\cdot)$ in a neighborhood of $\bar{x}$. If $\alpha_1 = a$ or $\alpha_m = b$, semismoothness of $H(\cdot)$ still holds basically by the same arguments as the proof of Theorem 5. □

Note that Proposition 9 and Theorem 4 have overlaps with Theorem 5, but do not cover each other.

6. Applications

In this section, we discuss applications of our theoretical results. The application for the shape-preserving interpolation problem is related with Theorem 4. We give a brief summary on this application in Subsection 6.1. Then we discuss the application to the option pricing problem in detail in Subsection 6.2. We show that the generalized Newton method for solving the no-arbitrage option price interpolation problem, proposed by Wang, Yin and Qi in [33], has at least $\frac{4}{3}$-order convergence. We give conditions when this method has $\frac{3}{2}$-order or quadratic convergence. We also outline a damped version of the generalized Newton method and show that it is globally convergent and the convergence order is at least $\frac{4}{3}$. The main tool there is Theorem 3 for the p-order semismoothness of integral functions discussed in Section 3. In Subsection 6.3, we briefly discuss an application of Theorems 4 and 5 to semi-infinite programs.

6.1. Shape-preserving interpolation

The constrained approximation problem comes from practical applications in computer aided geometric design where one has not only to approximate data points but also to achieve a desired shape of a curve or a surface. This is also called *shape preserving approximation*.

Examples of a desired shape property include convexity and monotonicity. A special case of shape preserving approximation is *shape preserving interpolation*. That is, to find a function, whose graph has a desired shape, to interpolate given points.

Consider the following *convex best interpolation* problem:

$$\begin{aligned} &\text{Min } \|f''\|_2 \\ &\text{s.t. } f(s_i) = y_i, \quad i = 1, \cdots, n+2, \\ &\qquad f \text{ is convex on } [a,b], \quad f \in W^{2,2}[a,b], \end{aligned} \tag{6.1}$$

where $a = s_1 < s_2 < \cdots < s_{n+2} = b$ and y_i, $i = 1, \ldots, n+2$ are given numbers, $\|\cdot\|_2$ is the Lebesgue $L^2[a, b]$ norm, and $W^{2,2}[a, b]$ denotes the Sobolev space of functions with absolutely continuous first derivatives and second derivatives in $L^2[a, b]$, and equipped with the norm being the sum of the $L^2[a, b]$ norms of the function, its first, and its second derivatives.

Employing the normalized B-splines B_i of order two associated with (s_i, y_i) and the corresponding second divided differences d_i, and using the Lagrange duality theory, the interpolation conditions can be equivalently written [4] as a system of *nonsmooth equations*

$$F(x) = d, \tag{6.2}$$

where $d = (d_1, \cdots, d_n)$ and the i-th component of F is defined by

$$F_i(x) = \int_a^b \left(\sum_{l=1}^n x_l B_l(s) \right)_+ B_i(s) ds. \tag{6.3}$$

We see that F_i is an integral function of the form (1.1), with

$$g(x, s) = \sum_{l=1}^n x_l B_l(s)$$

and

$$p(s) = B_i(s).$$

Irvine, Marin and Smith [9] proposed in 1986 a Newton-type method for solving the equation (6.2). By monitoring the decrease of the norm of the residual $F(\lambda) - d$, they observed fast convergence in their numerical experiments and raised the question of theoretically estimating the rate of convergence. They wrote: *"Although we have not established rigorous convergence results for Newton's method we have been very encouraged by numerical experiments...."*.

We may view the Newton-type method of Irvine, Marin and Smith as a generalized Newton method [14],[18]. According to [14],[18], if F is semismooth at the solution and all the matrices in the generalized Jacobian of F at the solution are nonsingular, we get superlinear convergence of the method. If further more, F is strongly semismooth at the solution, we get quadratic convergence of the method. The integral function F defined by (6.3) was proved to be semismooth in [4], and further proved to be strongly semismooth in [5]. The results of [4] and [5] was generalized and applied to the edge convex minimum norm network interpolation problem in [20]. Now, in this paper, Theorem 4 gives a complete analysis of differentiability of F defined by (5.1). This gives a more coherent view/proof of the semismoothness and strong semismoothness property of a certain generalized form for the convex interpolation problem.

6.2. Option pricing problem

Wang, Yin and Qi [33] developed an interpolation method to preserve the shape of the *option price* function. The interpolation is optimal in terms of minimizing the distance between the implied risk-neutral density and a prior approximation function in L^2-norm, which is very important when only a few observations are available.

Since the seminal paper of Black-Scholes [1], numerous theoretical and empirical studies have been done on the no-arbitrage pricing theory, see Duffie [6] and the references therein. If the uncertainty of nature can be described by a stochastic process q_t, then the absence of arbitrage opportunities implies that there exists a state-price density (SPD) or risk-neutral density, which is denoted by $p(q_{t_2}|F_{t_1})$, where t_2 is any time after time t_1, F_{t_1} is all the information available at time t_1. The price of any financial security can be expressed as the expected net present value of future payoffs, where the expectation is taken with respect to the risk-neutral density. In the call option pricing case, the underlying asset price S_t can be used as the state variable, the risk-free rate is considered as constant. So the price at time t is

$$C(S_t, s, \tau, r_{t,\tau}) = e^{-r_{t,\tau}\tau} \int_0^\infty (S_T - s)_+ p(S_T|S_t, \tau, r_{t,\tau}) dS_T, \tag{6.4}$$

where S_t is the underlying asset price at time t, s is the strike price of the option contract, τ is the time-to-expiration, $T = t + \tau$ is the expiration time, $r_{t,\tau}$ is the risk free rate from time t to $T = t + \tau$. No matter what kind of process of the underlying asset price S_t is, and whether the market is complete or not, the equation above always holds.

Wang, Yin and Qi [33] also proved that the option price function is convex with respect to the strike price s.

Without loss of the generality, we assume that $0 < a = s_0 < s_1 < \cdots < s_{n+2} = b < +\infty$ and consider the following constrained interpolation problem:

$$\begin{aligned} &\text{Min } \|f''(s) - h(s)\|_2 \\ &\text{s.t. } f(s_i) = y_i,\ i = 1, 2, \cdots, n+2, \\ &\qquad f''(s) \geq 0 \text{ for } a.e.\ s \in [a, b],\ f \in W_2^2[a, b], \end{aligned} \tag{6.5}$$

where

$$h(s) = e^{-r_{t,\tau}\tau} \frac{1}{x\sigma\sqrt{2\pi\tau}} exp\left\{-\frac{(\log s - \log S_t - r_{t,\tau}\tau + \sigma^2\tau/2)^2}{2\sigma^2\tau}\right\}. \tag{6.6}$$

By using the duality theory and Lagrange multipliers, Wang, Yin and Qi [33] converted the minimization problem (6.5) to a system of *nonsmooth equations*

$$F(x) = d, \tag{6.7}$$

where $d = (d_1, d_2, \cdots, d_n)^T$, $F = (F_1, F_2, \cdots, F_n)^T : \mathbb{R}^n \to \mathbb{R}^n$ and the $i-$th component of F is defined by

$$F_i(x) = \int_a^b \left(\sum_{l=1}^n x_l B_l(s) + h(s)\right)_+ B_i(s) ds. \tag{6.8}$$

Again, F_i is an integral function of the form (1.1), with

$$g(x,s) = \sum_{l=1}^{n} x_l B_l(s) + h(s)$$

and

$$p(s) = B_i(s).$$

Wang, Yin and Qi [33] applied the following generalized Newton method to solve (6.7):

$$M(x^k)x^{k+1} = d - \int_a^b h(s)B(s)dx \quad k = 1, 2, \cdots , \tag{6.9}$$

where $B(s) = (B_1(s), B_2(s), \cdots , B_n(s))^T$, and $M(x) \in \mathbb{R}^{n\times n}$ with components

$$M_{ij}(x) = \int_a^b \mathbf{1}_{(0,+\infty)} \left(\sum_{l=1}^{n} x_l B_l(s) + h(s) \right) B_i(s)B_j(s)ds,$$

where $\mathbf{1}_{(0,+\infty)}(\cdot)$ is the characteristic function of the set $(0, +\infty)$, i.e., $\mathbf{1}_{(0,+\infty)}(x) = 1$ for $x > 0$ and $\mathbf{1}_{(0,+\infty)}(x) = 0$ for $x \le 0$.

By applying the results of [4], Wang, Yin and Qi [33] established semismoothness of the integral function F defined by (6.8), and hence superlinear convergence of the generalized Newton method (6.9). However, Wang, Yin and Qi [33] has not established p-order semismoothness of the integral function F defined by (6.8), and hence has not established a convergence rate of (6.9), higher than superlinear convergence.

Now, we apply Theorem 3 to the integral function $F(\cdot)$ defined by (6.8). Recall that the B-spline B_i is given by

$$B_i(s) = \begin{cases} \alpha_i(s - s_i), & \text{for } s \in [s_i, s_{i+1}], \\ \bar{\alpha}_i(s_{i+2} - s), & \text{for } s \in [s_{i+1}, s_{i+2}], \\ 0, & \text{otherwise}, \end{cases}$$

where we denote

$$\alpha_i = 2/((s_{i+2} - s_i)(s_{i+1} - s_i)), \quad \bar{\alpha}_i = 2/((s_{i+2} - s_i)(s_{i+2} - s_{i+1})).$$

In the sequel we study the following functions:

$$\begin{aligned}
\Phi_1(x_1) &= \int_{s_1}^{s_2} (x_1 B_1(s) + h(s))_+ B_1(s)ds, \\
\Phi_2(x_n) &= \int_{s_{n+1}}^{s_{n+2}} (x_n B_n(s) + h(s))_+ B_n(s)ds, \\
\Gamma_i(x_{i-1}, x_i) &= \int_{s_i}^{s_{i+1}} (x_{i-1}B_{i-1}(s) + x_i B_i(s) + h(s))_+ B_i(s)ds, \quad i = 2, \ldots , n, \\
\Psi_i(x_i, x_{i+1}) &= \int_{s_{i+1}}^{s_{i+2}} (x_i B_i(s) + x_{i+1}B_{i+1}(s) + h(s))_+ B_i(s)ds, \quad i = 1, \ldots , n - 1.
\end{aligned}$$

Then

$$F_1(x) = \Phi_1(x_1) + \Psi_1(x_1, x_2),$$
$$F_i(x) = \Gamma_i(x_{i-1}, x_i) + \Psi_i(x_i, x_{i+1}), \quad i = 2, \dots, n-1,$$
$$F_n(x) = \Gamma_n(x_{n-1}, x_n) + \Phi_2(x_n).$$

We may use Theorem 3 to establish p-order semismoothness of Φ_i, Γ_i and Ψ_i. This implies p-order semismoothness of F. Since the $h(\cdot)$ possesses a very special structure, it has at most two inflection points. For any $\bar{x} = (\bar{x}_1, \bar{x}_2, \cdots, \bar{x}_n)^T \in \mathbb{R}^n$, we assume, separating $[a, b]$ more finely if necessary, that $\sum_{l=1}^n \bar{x}_l B_l(s) + h(s) = 0$ has at most a root on $[s_i, s_{i+1}]$. So, the position relation between the line segment $y = \sum_{l=1}^n \bar{x}_l B_l(s)$ on $[s_i, s_{i+1}]$ and the curve $y = -h(s)$ has four cases: (1) not intersected, (2) intersected but not tangent, (3) tangent at a convex or concave arc point of the curve $y = -h(s)$, (4) tangent at an inflection point of the curve $y = -h(s)$. On $[s_i, s_{i+1}]$, the first two cases happen if and only if the root of $\sum_{l=1}^n \bar{x}_l B_l(s) + h(s) = 0$ is simple on this subinterval. The third case happens if and only if $\sum_{l=1}^n \bar{x}_l B_l(s) + h(s) = 0$ has a 2-order root on this subinterval. The fourth case happens if and only if the order of root of $\sum_{l=1}^n \bar{x}_l B_l(s) + h(s) = 0$ on this subinterval is 3. Therefore, the following result immediately follows Theorem 3. This result strengthens Theorem 4.6 of [33].

Theorem 6. *Consider the integral function $F(\cdot)$ defined by* (6.8). *For any $\bar{x} \in \mathbb{R}^n$, the following three cases hold:* (1) *If the roots of $\sum_{l=1}^n \bar{x}_l B_l(s) + h(s) = 0$ are simple on every subinterval $[s_i, s_{i+1}] \subseteq [a, b]$, $i = 1, 2, \cdots, n+1$, then F is* 1*-order (strongly) semismooth at $\bar{x}$.* (2) *If there exists a certain subinterval $[s_i, s_{i+1}] \subseteq [a, b]$ such that the highest order of roots of $\sum_{l=1}^n \bar{x}_l B_l(s) + h(s) = 0$ on $[s_i, s_{i+1}]$ is* 2*, then F is $\frac{1}{2}$-order semismooth at $\bar{x}$.* (3) *If there exists a certain subinterval $[s_i, s_{i+1}] \subseteq [a, b]$ such that the highest order of roots of $\sum_{l=1}^n \bar{x}_l B_l(s) + h(s) = 0$ on $[s_i, s_{i+1}]$ is* 3*, then F is $\frac{1}{3}$-order semismooth at $\bar{x}$.*

Let

$$L(x) = \frac{1}{2} \int_a^b \left[\sum_{l=1}^n x_l B_l(s) + h(s) \right]_+^2 ds - \sum_{l=1}^n x_l d_l.$$

The following algorithm is a “damped" globalization of Newton’s method based on regularization controlled by the residual.

Algorithm 10 (Damped Newton method). (S.0)*(Initialization) Choose $x^0 \in \mathbb{R}^n$, $\rho \in (0, 1)$, $\sigma \in (0, \frac{1}{2})$, and tolerance* tol$> 0$. $k := 0$.
(S.1)*(Termination criterion) If $\varepsilon_k = \|F(x^k) - d\| \le$ tol then stop. Otherwise, go to* (S.2).
(S.2)*(Direction generation) Let s^k be a solution of the following linear system*

$$(M(x^k) + \varepsilon_k I)s = -\nabla L(x^k).$$

(S.3)*(Line search) choose m_k as the smallest nonnegative integer m satisfying*

$$L(x^k + \rho^m s^k) - L(x^k) \le \sigma \rho^m \nabla L(x^k)^T s^k.$$

(S.4) *(Update) Set $x^{k+1} = x^k + \rho^{m_k} s^k$, $k := k + 1$, return to step* (S.1).

Using the technique employed in [12], we may establish the convergence results of this algorithm as follows. Since these arguments are routine in the literature for semismooth Newton methods, we do not go to details.

Theorem 7. *Let $x^0 \in \mathbb{R}^n$ and $\{x^k\}$ be generated by Algorithm* 10. *Then the sequence $\{x^k\}$ converges to the solution x^* of* (6.7), *and the convergence is at least of order $\frac{4}{3}$.*

6.3. Semi-Infinite Programming

Consider the following semi-infinite programming (SIP) problem:

$$\begin{aligned} \text{Min}\ & f(x) \\ \text{s.t.}\ & h_j(x) \le 0, \quad j = 1, \cdots, p, \\ & g_j(x, s) \le 0, \quad s \in [a, b] \quad j = 1, \cdots, m, \end{aligned} \tag{6.10}$$

where $h_j(x) \le 0, \quad j = 1, \cdots, p$ are conventional inequality constraints, while $g_j(x, s) \le 0, \quad s \in [a, b] \quad j = 1, \cdots, m$ are infinite functional constraints, g is continuously differentiable (smooth) in x and s.

Such a SIP problem has wide applications [13],[22]. In 1989–1993, Teo and his collaborators [10],[27],[28],[29] proposed to aggregate the functional constraints to

$$G_j(x) := \int_a^b [g_j(x, s)]_+ \, ds = 0, \quad j = 1, \cdots, m.$$

Then the SIP problem (6.10) is converted to a nonlinear programming problem:

$$\begin{aligned} \text{Min}\ & f(x) \\ \text{s.t.}\ & h_j(x) \le 0, \quad j = 1, \cdots, p, \\ & G_j(x) \le 0, \quad j = 1, \cdots, m, \end{aligned} \tag{6.11}$$

where G_j may be nonsmooth. Teo and his collaborators [10],[27],[28],[29] proposed a smoothing method to solve (6.11). We see that G_j is an integral function (1.1), with $p(s) \equiv 1$.

Now, Theorems 4 and 5 give two cases in which G_j are SC^1. We then may apply SQP methods to solve (6.11) and get superlinear convergence and global convergence by the results in [15],[12].

References

1. Black, F., Scholes, M.: The pricing of options and corporate liabilities. J. Political Economy **81**, 637–659 (1973)
2. Bonnans, J.F., Shapiro, A.: Perturbation Analysis of Optimization Problems. Springer, New York, 2000
3. Corwin, L.J., Szczarba, R.H.: Multivariable Calculus. Marcel Dekker, Inc. New York, 1979
4. Dontchev, A.L., Qi, H., Qi, L.: Convergence of Newton's method for convex best interpolation. Numer. Math. **87**, 435–456 (2001)
5. Dontchev, A.L., Qi, H.-D., Qi, L.: Quadratic convergence of Newton's method for convex interpolation and smoothing. Constr. Approx. **19**, 123–143 (2003)
6. Duffie, D.: Dynamic Asset Pricing Theory. Princeton University Press, Princeton, 1996
7. Facchinei, F.: Minimization of SC^1 functions and the Maratos effect. Oper. Res. Lett. **17**, 131–137 (1995)

8. Facchinei, F., Pang, J.S.: Finite-Dimensional Variational Inequalities and Complementarity Problems. Springer, New York, 2003
9. Irvine, L.D., Marin, S.P., Smith, P.W.: Constrained interpolation and smoothing. Constr. Approx. **2**, 129–151 (1986)
10. Jennings, L.S., Teo, K.L.: A computational algorithm for functional inequality constrained optimization problems. Automatica **26**, 371–375 (1990)
11. Mifflin, R.: Semismooth and semiconvex functions in constrained optimization. SIAM J. Control Optim. **15**, 957–972 (1977)
12. Pang, J.S., Qi, L.: A globally convergent Newton method for convex SC^1 minimization problems. J. Optim. Theory Appl. **85**, 633–648 (1995)
13. Polak, E.: Optimization: Algorithms and Consistent Approximation. Springer-Verlag, New York, 1997
14. Qi, L.: Convergence analysis of some algorithms for solving nonsmooth equations. Math. Oper. Res. **18**, 227–253 (1993)
15. Qi, L.: Superlinearly convergent approximate Newton methods for LC^1 optimization problems. Math. Prog. **64**, 277–294 (1994)
16. Qi, L.: Semismoothness properties and applications of an integral function. In: Tamura, A., Ito, H. (eds.), Proceedings of The Fourteenth RAMP Symposium, RAMP, Kyoto, 2002, pp. 103–114
17. Qi, L., Jiang, H.: Semismooth Karush-Kuhn-Tucker equations and convergence analysis of Newton and quasi-Newton methods for solving these equations. Math. Oper. Res. **22**, 301–325 (1997)
18. Qi, L., Sun, J.: A nonsmooth version of Newton's method. Math. Prog. **58**, 353–367 (1993)
19. Qi, L., Tseng, P.: An analysis of piecewise smooth functions and almost smooth functions. Department of Applied Mathematics, The Hong Kong Polytechnic University, Hong Kong, 2002
20. Qi, L., Yin, H.: A strongly semismooth integral function and its application. Comput. Optim. Appl. **25**, 223–246 (2003)
21. Ralph, D.: Private Communication, 2002
22. Reemsten, R., Rückmann, J.: Semi-Infinite Programming. Kluwer, Boston, 1998
23. Rockafellar, R.T.: Some properties of piecewise smooth functions. Comput. Optim. Appl. **25**, 247–250 (2003)
24. Ruszczyński, A., Shapiro, A., eds.: Stochastic Programming. In: Handbooks in OR & MS, Vol. 10, North-Holland Publishing Company, Amsterdam, 2003
25. Shapiro, A.: On concepts of directional differentiability. J. Optim. Theory Appl. **66**, 477–487 (1990)
26. Sun, J., Sun, D., Qi, L.: A squared smoothing Newton method for nonsmooth matrix equations and its applications in semidefinite optimization problems. SIAM J. Optim. **14**, 783–806 (2004)
27. Teo, K.L., Jennings, L.S.: Nonlinear optimal control problems with continuous state inequality constraints. J. Optim. Theory Appl. **63**, 1–22 (1989)
28. Teo, K.L., Goh, C.J., Wong, K.H.: A Unified Computational Approach to Optimal Control Problems. Longman Scientific and Technical, 1991
29. Teo, K.L., Rehbock, V., Jennings, L.S.: A new computational algorithm for functional inequality constrained optimization problems. Automatica **29**, 789–792 (1993)
30. Ulbrich, M.: Semismooth Newton methods for operator equations in function spaces. SIAM. J. Optim. **13**, 805–842 (2003)
31. Uryasev, S.: A Differentiation formula for integrals over sets given by inclusion. Numer. Func. Anal. Optim. **10**, 827–841 (1989)
32. Uryasev, S.: Derivatives of probability functions and integrals over sets given by inequalities. J. Comput. Applied Math. **56**, 197–223 (1994)
33. Wang, Y., Yin, H., Qi, L.: No-Arbitrage interpolation of the option price function and its reformulation. J. Optim. Theory Appl. **120**, 629–649 (2004)