ARTICLE IN PRESS

Computers and Mathematics with Applications ▮ (▮▮▮▮) ▮▮▮–▮▮▮

Contents lists available at ScienceDirect

Computers and Mathematics with Applications

journal homepage: www.elsevier.com/locate/camwa

Differentiability of the objective in a class of coefficient inverse problems[☆]

Maciej Smołka

AGH University of Science and Technology, Al. Mickiewicza 30, 30-059 Kraków, Poland

ARTICLE INFO

Article history:
Received 10 November 2016
Received in revised form 3 March 2017
Accepted 10 April 2017
Available online xxxx

Keywords:
Coefficient inverse problem
Objective functional
Fréchet differentiability

ABSTRACT

In this paper we study the Fréchet differentiability of the objective functional for a quite general class of coefficient inverse problems. We present sufficient conditions for the existence of arbitrary-order differentials as well as formulae for the calculation of first and second order derivatives. The general case theorems are then applied to two real-world geophysical inverse problems.

1. Introduction

We consider the following class of inverse problems: find the value of a parameter $\gamma^* \in \mathcal{D}$ such that

$$\mathcal{C}(\gamma^*) = \min_{\gamma \in \mathcal{D}} \mathcal{C}(\gamma). \tag{1}$$

Here

$$\mathcal{C} : \mathcal{D} \ni \gamma \longmapsto f(d_o, \mathcal{Q}(u_\gamma)) \in \mathbb{R}_+ \tag{2}$$

is the objective functional (or simply the objective), $\mathcal{D}$ is a set of admissible values, i.e. a model space (cf. [1]),

$$A_\gamma : V \longrightarrow V' \text{ or } \overline{V}'$$

is a γ-dependent direct problem operator between a Hilbert space V and its dual (or the complex conjugate to its dual when V is complex), $u_\gamma \in V$ is the solution of the direct problem

$$A_\gamma(u_\gamma) = 0,$$

i.e., the direct solution corresponding to a parameter $\gamma \in \mathcal{D}$,

$$\mathcal{Q} : V \longrightarrow \mathcal{O}$$

is an observation operator mapping the direct solution to an observable quantity (called *quantity of interest*), $\mathcal{O}$ is a data space (cf. [1]), $d_o \in \mathcal{O}$ denotes actual observation data and

$$f : \mathcal{O} \times \mathcal{O} \longrightarrow \mathbb{R}_+$$

☆ This work has been partially supported by the National Science Centre, Poland grant No. 2015/17/B/ST6/01867.

E-mail address: smolka@agh.edu.pl.
URL: http://www.ki.agh.edu.pl/en/staff/smolka-maciej/.

http://dx.doi.org/10.1016/j.camwa.2017.04.011

is a *misfit* functional. In this paper we assume that the operator A_γ has the following form

$$A_\gamma(u)(v) = b(\gamma; u, v) - L(v) \tag{3}$$

for every $\gamma \in \mathcal{D}$, $u, v \in V$, i.e., the direct problem in (1) is to find $u_\gamma \in V$ such that

$$b(\gamma; u_\gamma, v) = L(v) \quad \text{for every } v \in V. \tag{4}$$

Hence, we consider a class of coefficient inverse problems (cf. [2]) formulated as global optimization problems using the quasi-solution method due to Ivanov (cf. [3]).

Since our problem is put into the optimization framework, it is natural to apply fast gradient-based optimization methods (like Newton, Gauss–Newton, conjugate gradients, BFGS, etc., cf. [4]) to solve it. And, indeed, these methods have been a common choice in the inverse problem community, see e.g. [5–9] and the references therein. Although the straightforward application of gradient methods can be problematic because of the typical ill-posedness of problems (1), even when using the regularization [10], the objective derivatives can be used in the construction of effective inversion strategies. In one of such, called the Hierarchic Memetic Strategy (HMS) [11] they are already used or are planned to be used in the following areas:

- hybridizing the gradient methods with a stochastic evolutionary global optimization solver to increase the accuracy and reduce the computational cost of the latter;
- constructing a local approximation of the objective that shall be more accurate than the one utilizing only objective values; such an approximation can be then used for rough detection of solution attraction basins and plateaus in the objective landscape;
- constructing so-called utility function in a special type of selection operator based on the idea of multiwinner voting, that has the capability of good plateau coverage [12].

A fully justified application of gradient methods requires the the first-order (sometimes second-order) Fréchet differentiability of the objective. Furthermore, successful applications of the regularization method are based on conditions involving Fréchet derivatives, such as Theorem 1.9.1.2 from the book [13]. Hence, the existence of Fréchet differentials (and not only, e.g., Gâteaux ones) of the objective is crucial from the applicational point of view. Despite that, authors of many papers pay little attention to formal proving the differentiability and providing formulae for the computation of derivatives. Nevertheless, there are also quite a few papers that deal with those formal aspects, like [14–16,7,17–21]. Typically, such papers concentrate on the objective gradient, however, Kwon and Yazici [22] show formulae for the computation of higher-order derivatives as well. Many papers deal with particular classes of inverse problems, like inverse scattering [16], electrical impedance tomography [14] or vertical electrical sounding [17]. In contrast, Dierkes et al. [23] consider quite a general case of coefficient inverse problems and obtain results applicable to a few various concrete problems. In this paper we propose a different and in a sense yet more general approach that seems to be applicable to a quite broad class of coefficient inverse problems. We prove the existence of arbitrary-order Fréchet derivatives under some natural assumptions and show formulae to compute the Jacobian and the Hessian in a general case, and then we apply the method to two geophysical coefficient inverse problems related to the conductivity (or resistivity) logging measurements. Let us note that the presented real-world problems were already solved numerically using HMS [24,25].

2. Some remarks on Banach algebra analysis

For the convenience of the reader let us start with recalling the definition of the Fréchet differential. Let U and W be Banach spaces and let G be an open subset of U. We say that a mapping

$$F : G \longrightarrow W$$

is *Fréchet differentiable* at $u \in G$ if there exists a linear and continuous mapping

$$D_uF : U \longrightarrow W$$

such that for every $h \in U$ such that $u + h \in G$ we have

$$\frac{\|F(u+h) - F(u) - D_uF(h)\|_W}{\|h\|_U} \to 0 \quad \text{as } h \to 0.$$

The mapping D_uF is called *the differential of* F *in* u. The value of D_uF on a vector h, i.e. $D_uF(h)$, is *the derivative of* F *in* u *in direction* h.

If F is differentiable at every $u \in G$, then the following mapping is defined

$$DF : G \ni u \longmapsto D_uF \in \mathcal{L}(U; W),$$

where $\mathcal{L}(U; W)$ denotes the space of linear and continuous mappings from U into W. Since the latter is itself a Banach space, we can define higher differentials and derivatives in the same way as above. In particular, the second differential of F at $u \in G$ (if it exists) is the following linear and continuous mapping

$$D_u^2F = D_uF(DF) : U \longrightarrow \mathcal{L}(U; W).$$