Nonlinear Analysis 75 (2012) 1290–1304

Contents lists available at SciVerse ScienceDirect

Nonlinear Analysis

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

The optimal value and optimal solutions of the proximal average of convex functions

Rafal Goebel [a], Warren Hare [b,*], Xianfu Wang [b]

[a] Department of Mathematics and Statistics, Loyola University, Chicago, IL 60626, USA

[b] Mathematics, Irving K. Barber School, University of British Columbia Okanagan, Kelowna, British Columbia V1V 1V7, Canada

ARTICLE INFO

Communicated by Ravi Agarwal

MSC:
primary 90C25
secondary 52A41
26B25
47H05
47H10
49N60

Keywords:
Convex function
Proximal average
Moreau envelope
Parametric optimization
Minima
Sharp minima

ABSTRACT

The proximal average of a finite collection of convex functions is a parameterized convex function that provides a continuous transformation between the convex functions in the collection. This paper analyzes the dependence of the optimal value and the minimizers of the proximal average on the weighting parameter. Concavity of the optimal value is established and implies further regularity properties of the optimal value. Boundedness, outer semicontinuity, single-valuedness, continuity, and Lipschitz continuity of the minimizer mapping are concluded under various assumptions. Sharp minimizers are given further attention. Several examples are given to illustrate our results.

1. Introduction

In this paper, we consider a specific example of a parametric optimization problem:

$$\varphi(\lambda) := \inf_x F(x, \lambda).$$

Such problems arise naturally in a variety of settings [1,2], most notably multi-objective optimization [3]. In this paper, $F(x, \lambda)$ is the proximal average of a finite number of convex functions.

For a finite collection of convex functions, the proximal average provides a novel technique of averaging the functions, while maintaining and creating strong analytic properties. Unlike the arithmetic average, the proximal average results in a well-defined function regardless of the intersection of the effective domains of the averaged functions, and is bounded below and above, respectively, by the arithmetic and the epigraphical average. A striking property of the proximal average is that it is self-dual with respect to Fenchel conjugacy. For general definitions and a systematic study, see [4]; for earlier special cases or further extensions, see [5–11].

Whenever researching analytically the properties of parameterized functions, it is natural to examine the question of how the minimum and minimizers of the function behave as the parameters vary. For the purposes of optimization, one

* Corresponding author. Tel.: +1 250 807 9378.
E-mail addresses: rgoebel1@luc.edu (R. Goebel), warren.hare@ubc.ca (W. Hare), shawn.wang@ubc.ca (X. Wang).

doi:10.1016/j.na.2011.06.017

hopes to conclude that the minimum and minimizers are stable under small perturbations in parameter values. In the best case, one hopes that the minimum and minimizers behave in a predictable manner (e.g., linearly) under perturbations in the parameter values. In this paper we show that the proximal average has these properties.

On a side note, like it is the case for all convex functions, the minimizer set of the proximal average is precisely the fixed point set of its proximal point mapping. In the case of the proximal average, this proximal point mapping corresponds to the convex combination of the proximal point mappings of the averaged functions. Averaged mappings play a very important role in fixed point algorithms [5,12,13], therefore the research in this paper also provides interesting insight into the field of fixed point theory.

The remainder of this paper is organized as follows. In Section 2 we recall the basic notions, state the definition of the proximal average and of the basic objects of study in the paper, and gather several properties of the proximal average and other background on convex functions. In Section 3 we give results on concavity, boundedness, and continuity of the optimal value function, and on boundedness, outer semicontinuity, single-valuedness, continuity, and Lipschitz continuity of the optimal solution mapping. In Section 4 we turn our attention to functions with sharp minimizers, and show when the optimal value behaves affinely, and the optimal solution of the average is the average of sharp minimizers. In the final section, we provide examples and useful techniques.

2. Preliminaries

The set of all proper lower semicontinuous convex functions on $\mathbb{R}^N$ is denoted by $\Gamma(\mathbb{R}^N)$. We use

$$\mathfrak{q} := \frac{1}{2}\|\cdot\|^2, \qquad \Delta^{n-1} := \left\{(\lambda_1,\ldots,\lambda_n)\in\mathbb{R}^n : 0\le\lambda_i\le 1, \sum_{i=1}^n \lambda_i = 1\right\},$$

$e_i := (0,\ldots,0,1,0,\ldots,0)$ to represent the standard unit vector in $\mathbb{R}^n$ for $i = 1,\ldots,n$, Id to represent the identify mapping, and $\mathbb{B}_\delta(x)$ to denote the *closed* ball with radius $\delta > 0$ centered at $x \in \mathbb{R}^N$. For a set $C \subset \mathbb{R}^N$, its *support function* is defined by $\sigma_C(x) := \sup\langle x, C\rangle$ and its *indicator function* is defined by $\iota_C(x) := 0$ if $x \in C$ and $+\infty$ if $x \notin C$. We use $\operatorname{int} C$, $\operatorname{ri} C$, $\operatorname{affine} C$, $\operatorname{conv} C$ for its interior, relative interior, affine hull and convex hull of C respectively. In particular,

$$\operatorname{ri}\Delta^{n-1} = \left\{(\lambda_1,\ldots,\lambda_n) : \sum_{i=1}^n \lambda_i = 1, 0<\lambda_i<1,\ i=1,\ldots,n\right\}.$$

Given $f \in \Gamma(\mathbb{R}^N)$, the *Fenchel conjugate* f^* is the function defined by

$$f^*(y) := \sup_{x\in\mathbb{R}^N}(\langle x, x^*\rangle - f(x)),$$

and the *subdifferential* of f is the set-valued mapping $\partial f : \mathbb{R}^N \rightrightarrows \mathbb{R}^N$ defined by

$$\partial f(x) := \{x^*\in\mathbb{R}^N : f(y)\ge f(x)+\langle y-x, x^*\rangle\ \forall y\in\mathbb{R}^N\}.$$

The *Moreau envelope* of f is defined by

$$e_\mu f(x) := \inf_y\{f(y)+\mu^{-1}\mathfrak{q}(x-y)\} \quad \text{for all } x\in\mathbb{R}^N,$$

and the *proximal mapping* of f, with parameter μ, is the minimizer in the definition of $e_\mu f$, equivalently, $\operatorname{Prox}_\mu f := (\mathrm{Id} + \mu\partial f)^{-1}$. When $\mu = 1$, we write $\operatorname{Prox}_\mu f$ as $\operatorname{Prox} f$. A function f is called *essentially strictly convex* if f is strictly convex on every convex subset of $\operatorname{dom}\partial f := \{x | \partial f(x) \neq \emptyset\}$. Essentially strictly convex functions have at most a single minimizer. Indeed, any minimizer x of such f satisfies $0 \in \partial f$; the set of minimizers is convex; and, if it is nonempty, strict convexity implies it is a singleton. A set-valued mapping $S : \mathbb{R}^N \rightrightarrows \mathbb{R}^N$ is *outer semicontinuous (osc)* at $\bar{x}$ if $\limsup_{x\to\bar{x}} S(x) \subset S(\bar{x})$, or equivalently $\limsup_{x\to\bar{x}} S(x) = S(\bar{x})$. For details on these and other basic concepts of set-valued and convex analysis, see [1,14]. To shorten the notation, we will write $\boldsymbol{f} := (f_1,\ldots,f_n)$, $\lambda := (\lambda_1,\ldots,\lambda_n)$, and $\boldsymbol{f}^* := (f_1,\ldots,f_n)^* := (f_1^*,\ldots,f_n^*)$.

2.1. Main objects of study

Given a collection of functions $\{f_1, f_2,\ldots,f_n\}$, where $f_i \in \Gamma(\mathbb{R}^N)$ for $i = 1,\ldots,n$ and a proximal parameter $\mu > 0$, the *proximal average* is a parameterized function $p_\mu(\boldsymbol{f})$ defined by

$$\begin{aligned} p_\mu(\boldsymbol{f}) \quad &: \quad \mathbb{R}^N\times\Delta^{n-1}\to\,]-\infty,+\infty] \\ &: \quad (x,\lambda)\mapsto\left(\lambda_1(f_1+\mu^{-1}\mathfrak{q})^*+\cdots+\lambda_n(f_n+\mu^{-1}\mathfrak{q})^*\right)^* - \mu^{-1}\mathfrak{q}. \end{aligned}$$

It is important to note that for every $\lambda \in \Delta^{n-1}$, $x \mapsto p_\mu(\boldsymbol{f})(x,\lambda)$ is a convex function. Further properties are collected in Section 5.4.