Applied Mathematics and Computation 296 (2017) 168–181

Contents lists available at ScienceDirect

Applied Mathematics and Computation

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

On modeling and global solutions for d.c. optimization problems by canonical duality theory

Zhong Jin [a], David Y. Gao [b,*]

[a] Department of Mathematics, Shanghai Maritime University, Shanghai 201306, PR China
[b] Faculty of Science and Technology, Federation University Australia, Victoria 3353, Australia

ARTICLE INFO

MSC:
90C26
90C30
90C46

Keywords:
Global optimization
Canonical duality theory
D.C. programming
Mathematical modeling

ABSTRACT

This paper presents a canonical d.c. (difference of canonical and convex functions) programming problem, which can be used to model general global optimization problems in complex systems. It shows that by using the canonical duality theory, a large class of nonconvex minimization problems can be equivalently converted to a unified concave maximization problem over a convex domain, which can be solved easily under certain conditions. Additionally, a detailed proof for triality theory is provided, which can be used to identify local extremal solutions. Applications are illustrated and open problems are presented.

1. Problems and motivation

It is known that in Euclidean space every continuous global optimization problem on a compact set can be reformulated as a d.c. optimization problem, i.e. a nonconvex problem which can be described in terms of *d.c. functions* (difference of convex functions) and *d.c. sets* (difference of convex sets) [37]. By the fact that any constraint set can be equivalently relaxed by a nonsmooth indicator function, general nonconvex optimization problems can be written in the following standard d.c. programming form

$$\min\{f(x) = g(x) - h(x) \mid \forall x \in \mathcal{X}\}, \tag{1}$$

where $\mathcal{X} = \mathbb{R}^n$, $g(x)$, $h(x)$ are convex proper lower-semicontinuous functions on $\mathbb{R}^n$, and the d.c. function $f(x)$ to be optimized is usually called the “objective function” in mathematical optimization. A more general model is that $g(x)$ can be an arbitrary function [37]. Clearly, this d.c. programming problem is artificial. Although it can be used to “model” a very wide range of mathematical problems [24] and has been studied extensively during the last thirty years (cf. [25,34,39]), it comes at a price: it is impossible to have an elegant theory and powerful algorithms for solving this problem without detailed structures on these arbitrarily given functions. As the result, even some very simple d.c. programming problems are considered as NP-hard [37]. This dilemma is mainly due to the existing gap between mathematical optimization and mathematical physics.

1.1. Objectivity and multi-scale modeling

Generally speaking, the concept of *objectivity* used in our daily life means the state or quality of being true even outside of a subject’s individual biases, interpretations, feelings, and imaginings (see Wikipedia at

* Corresponding author.
E-mail addresses: asdjz1234@163.com (Z. Jin), d.gao@federation.edu.au, davidygao@gmail.com (D. Y. Gao).

http://dx.doi.org/10.1016/j.amc.2016.10.010
0096-3003/

https://en.wikipedia.org/wiki/Objectivity_(philosophy)). In science, the objectivity is often attributed to the property of scientific measurement, as the accuracy of a measurement can be tested independent from the individual scientist who first reports it, i.e. an objective function does not depend on observers. In Lagrange mechanics and continuum physics, a real-valued function $W : \mathcal{X} \to \mathbb{R}$ is said to be objective if and only if (see [9], Chapter 6)

$$W(x) = W(Rx) \;\; \forall x \in \mathcal{X}, \;\; \forall R \in \mathcal{R}, \tag{2}$$

where $\mathcal{R}$ is a special rotation group such that $R^{-1} = R^T$, $\det R = 1$, $\forall R \in \mathcal{R}$.

Geometrically, an objective function does not depend on the rotation, but only on certain measure of its variable. The simplest measure in $\mathbb{R}^n$ is the ℓ_2 norm $\|x\|$, which is an objective function since $\|Rx\|^2 = (Rx)^T(Rx) = x^T R^T R x = \|x\|^2$ for all special orthogonal matrix $R \in SO(n)$. By Cholesky factorization, any positive definite matrix has a unique decomposition $C = D^*D$. Thus, any convex quadratic function is objective. It was emphasized by Ciarlet in his recent nonlinear analysis book [4] that the objectivity is not an assumption, but an axiom. Indeed, the objectivity is also known as the *axiom of frame-invariance* in continuum physics (see p. 8 in [27] and p. 42 in [35]). Although the objectivity has been well-defined in mathematical physics, it is still subjected to seriously study due to its importance in mathematical modeling (see [30–32]).

Based on the original concept of objectivity, a multi-scale mathematical model for general nonconvex systems was proposed by Gao in [9,17]:

$$(\mathcal{P}) : \quad \inf\{\Pi(x) = W(Dx) - F(x) \mid \forall x \in \mathcal{X}\}, \tag{3}$$

where $\mathcal{X}$ is a feasible space; $F : \mathcal{X} \to \mathbb{R} \cup \{-\infty\}$ is a so-called *subjective function*, which is linear on its effective domain $\mathcal{X}_a \subset \mathcal{X}$, wherein, certain "geometrical constraints" (such as boundary/initial conditions, etc.) are given; correspondingly, $W : \mathcal{Y} \to \mathbb{R} \cup \{\infty\}$ is an *objective function* on its effective domain $\mathcal{Y}_a \subset \mathcal{Y}$, in which, certain physical constraints (such as constitutive laws, etc.) are given; $D : \mathcal{X} \to \mathcal{Y}$ is a linear operator which assign each decision variable in configuration space $\mathcal{X}$ to an internal variable $y \in \mathcal{Y}$ at different scale. By Riesz representation theorem, the subjective function can be written as $F(x) = \langle x, \bar{x}^* \rangle \; \forall x \in \mathcal{X}_a$, where $\bar{x}^* \in \mathcal{X}^*$ is a given input (or source), the bilinear form $\langle x, x^* \rangle : \mathcal{X} \times \mathcal{X}^* \to \mathbb{R}$ puts $\mathcal{X}$ and $\mathcal{X}^*$ in duality. Additionally, the positivity conditions $W(y) \geq 0 \; \forall y \in \mathcal{Y}_a$, $F(x) \geq 0 \; \forall x \in \mathcal{X}_a$ and coercivity condition $\lim_{\|y\| \to \infty} W(y) = \infty$ are needed for the target function $\Pi(x)$ to be bounded below on its effective domain $\mathcal{X}_c = \{x \in \mathcal{X}_a | \; Dx \in \mathcal{Y}_a\}$ [17]. Therefore, the extremality condition $0 \in \partial \Pi(x)$ leads to the equilibrium equation [9]

$$0 \in D^* \partial W(Dx) - \partial F(x) \;\; \Leftrightarrow \;\; D^* y^* - x^* = 0 \;\; \forall x^* \in \partial F(x), \;\; y^* \in \partial W(y). \tag{4}$$

In this model, the objective duality relation $y^* \in \partial W(y)$ is governed by the constitutive law, which depends only on mathematical modeling of the system; the subjective duality relation $x^* \in \partial F(x)$ leads to the input $\bar{x}^*$ of the system, which depends only on each given problem. Thus, $(\mathcal{P})$ can be used to model general problems in multi-scale complex systems.

1.2. Real-world problems

In management science the variable $x \in \mathcal{X}_a \subset \mathbb{R}^n$ could represent the products of a manufacture company. Its dual variable $\bar{x}^* \in \mathbb{R}^n$ can be considered as market price (or demands). Therefore, the subjective function $F(x) = x^T \bar{x}^*$ in this example is the total income of the company. The products are produced by workers $y \in \mathbb{R}^m$. Due to the cooperation, we have $y = Dx$ and $D \in \mathbb{R}^{m \times n}$ is a matrix. Workers are paid by salary $y^* = \partial W(y)$, therefore, the objective function $W(y)$ in this example is the cost. Thus, $\Pi(x) = W(Dx) - F(x)$ is the *total loss or target* and the minimization problem $(\mathcal{P})$ leads to the equilibrium equation $D^T \partial_y W(Dx) = \bar{x}^*$. The cost function $W(y)$ could be convex for a very small company, but usually nonconvex for big companies.

In Lagrange mechanics, the variable $x \in \mathcal{X} = \mathcal{C}^1[I; \mathbb{R}^n]$ is a continuous vector-valued function of time $t \in I \subset \mathbb{R}$, its components $\{x_i(t)\} (i = 1, \ldots, n)$ are known as the Lagrange coordinates. The subjective function in this case is a linear functional $F(x) = \int_I x(t)^T \bar{x}^*(t) dt$, where $\bar{x}^*(t)$ is a given external force field. While $W(Dx)$ is the so-called action:

$$W(Dx) = \int_I L(x, \dot{x}) dt, \;\; L = T(\dot{x}) - V(x), \tag{5}$$

where T is the kinetic energy density, V is the potential density, and $L = T - V$ is the standard *Lagrangian density* [29]. The linear operator $Dx = \{\partial_t, 1\} x = \{\dot{x}, \; x\}$ is a vector-valued mapping. The kinetic energy T must be an objective function of the velocity (quadratic for Newton's mechanics and convex for Einstein's relativistic theory) [9], while the potential density V could be either convex or nonconvex, depending on each problem. Together, $\Pi(x) = W(Dx) - F(x)$ is called *total action*. The extremality condition $\partial \Pi(x) = 0$ leads to the well-known Euler–Lagrange equation

$$D^* \partial W(Dx) = \partial_t^* \frac{dT(\dot{x})}{d\dot{x}} - \frac{dV(x)}{dx} = \bar{x}^*, \tag{6}$$

where ∂_t^* is an adjoint operator of ∂_t. For convex Hamiltonian systems, both T and V are convex, thus, the least action principle leads to a typical d.c. minimization problem

$$\inf\{\Pi(x) = K(\partial_t x) - P(x)\}, \;\; K(y) = \int_I T(y) dt, \;\; P(x) = \int_I [V(x) + x^T \bar{x}^*] dt, \tag{7}$$