Proceedings of the Institute of Mathematics and Mechanics,
National Academy of Sciences of Azerbaijan
Volume 48, Number 1, 2022, Pages 50–62
https://doi.org/10.30546/2409-4994.48.1.2022.50

INFIMAL CONVOLUTION AND DUALITY IN CONVEX MATHEMATICAL PROGRAMMING

ELIMHAN N. MAHMUDOV AND MISIR J. MARDANOV

Abstract. In the paper it is considered a convex programming problem (CPP) with functional and non-functional constraints. In contrast to previous works, in the study of convex optimization problems, we do not deal with the classical approach of perturbations. In particular, thanks to the new representation of the indicator function on a convex set, the successful use of the infimal convolution method in this work plays a key role in proving duality results for problem CPP. Also, we consider a convex mathematical programming problem with inequality and linear equality constraints given by some matrix. In this case, it turns out that the dual cone to the cone of tangent directions coincides with the set of the image of the points of transposed matrix, taken with a minus sign.

1. Introduction

Convex optimization of mathematical programming [1, 2, 3, 4, 7, 10, 11, 22, 23, 24, 28, 29, 30, 31] has wide applications in automatic control systems, signal evaluation and processing, data analysis and modelling, finance electronic circuit design, statistics and structural optimization, etc. With the latest advances in computing and optimization algorithms, convex programming is almost as easy as linear programming. The work [2] studies linear programming problems on time scales. After a brief introduction to time scales, both the basic and dual models of linear programming of time scales are formulated. Further, a weak duality theorem and a theorem on optimality conditions for arbitrary time scales are established and proved. In the work [1], quadratic programming problems were formulated and solved using the time-scale approach. This approach combines discrete and continuous quadratic programming models and extends them to other cases. The formulation of the primal as well as the dual time scales quadratic programming models has been successfully constructed on arbitrary time scales. The new formulation provides an exact optimal solution for quadratic programming models using isolated time scales. In the note [29], a convex mathematical programming problem is formulated, in which the usual definition of the

2010 *Mathematics Subject Classification.* 51M16, 49N15, 47N10, 90C46.

Key words and phrases. Conjugate, duality, dual cone, infimal convolution, indicator function, Lagrange function.

feasible region is replaced by an essentially different strategy. Instead of specifying the feasible region by a set of convex inequalities $f_i(x) \leq b_i, i = 1, \dots, m$, the feasible region is defined via set containment. In accordance with the Pshenichny-Rockafellar Lemmas [11], the necessary and sufficient optimality condition is valid whenever the regularity condition is satisfied and an element $a \in dom f \cap C$ is the optimal solution to minimize the function $f(x)$ under the condition $x \in C$ if and only if $0 \in \partial f(a) - K_C^*(a)$. In the paper [8] the solution of nonlinear programming problems by a Sequential Quadratic Programming trust-region algorithm is considered. The aim of the present work is to promote global convergence without the need to use a penalty function.

Besides, various algorithms have been developed for solving the convex programming problem, which iteratively go to the optimum by constructing a cutting plane through the centre of the polyhedral approximation to the optimum. As a result, a sequence of primal feasible points is generated, the limit points of which satisfy the Kuhn-Tucker conditions of the problem. Additionally, are presented simple, effective rule for dropping prior cuts, an easily calculated bound on the objective function, and a rate of convergence. During the last decade, the field of interior polynomial point techniques has become one of the dominant areas of convex optimization. The goal of [22] is to present a general theory of algorithms for polynomial interior points in convex programming. The theory makes it possible to explain all known methods of this type and to extend them from the original field of interior point technique - linear and quadratic programming - to a wide range of essentially nonlinear classes of convex programs. In [31], to study the behavior of interior point methods in very large-scale linear programming problems, the application of such methods to continuous semi-infinite linear programming problems in both primal and dual form is considered. Considering various ways of discretizing such problems, the author arrives at a certain invariance property for (finite-dimensional) interior point methods. One of the main questions in convex programming theory is what happens to the best cost function when the constraint limits are changed from zero. It turns out that the Lagrange multipliers in the optimal design provide information to answer the sensitivity question. The study of this issue leads to a physical interpretation of the Lagrange multipliers, which can be very useful in practical applications; with relatively large values, they will have a significant impact on the optimal cost if the corresponding constraints are changed.

Along with these the duality theory plays a fundamental role in the analysis of optimization and variational problems. The reader can refer to [1, 8, 9, 25, 27] and their references for more details on this topic. It not only provides a powerful theoretical tool in the analysis of these problems, but also paves the way to designing new algorithms for solving them. Often, duality is associated with convex problems, yet it turns out that duality theory also has a fundamental impact even on the analysis of nonconvex problems. Note that the duality in convex optimization, in addition to the known ones, can be interpreted as a concept of the sensitivity of the optimization problem to perturbations of its data. It should be noted that in the works of Pshenichnyi [24], Bot [6, 7], Bonans [4], and many others, the perturbed problem is used for further research. Pshenichnyi's book [24] provides an overview of some recent and significant developments in the theory of perturbed optimization problems. Particular attention is paid to methods based on upper and lower estimates of the objective function of perturbed problems. Consequently, the proof of dual problems is based on not simple calculations associated with additional auxiliary set-valued mappings. In recent works by the authors Mahmudov and Mardanov [19], the concept of infimal convolution has been applied to an evolutionary dynamical problem with second-order differential inclusions. For the first time in the presented article, it is shown that due to the result of the representability of indicator functions described by convex inequalities the concept of infimal convolution can also be successfully applied in convex/linear programming problems. In the papers [12, 13, 14, 15, 16, 17, 18, 19, 20, 21], for optimal control problems of discrete processes and differential inclusions the necessary and sufficient conditions of optimality are formulated and the dual problems are constructed. The work [18] investigates the Mayer's problem with evolutionary differential inclusions and functional constraints of the theory of convex optimization; For this, first used an interesting auxiliary problem with second-order discrete time and discrete approximate inclusions. Necessary and sufficient conditions are proved, including the Euler-Lagrange inclusion, Hamiltonian inclusion, transversality and additional slackness conditions. This approach and results serve as a bridge between the problem of convex optimization theory with differential inclusion and mathematical programming problems with constraints in finite-dimensional spaces. The article [27] investigates the theory of optimal control of second-order polyhedral discrete and differential inclusions with state constraints. In terms of Euler-Lagrange inclusions and special conditions of "transversality" are formulated optimality conditions for posed problems. Despite the external dissimilarities, by nature the convex problem of mathematical programming and the problems associated with polyhedral inclusions are closely related to each other.

In the present work, the optimality conditions for a CPP are formulated without perturbed approach [5, 6, 7, 24]. Also,without going over to the perturbed problem we treat the dual results according to the dual operations of addition and infimal convolution of convex functions [12, 13, 14, 17, 21, 26]. Note that the paper [5] presents an overview of some recent, and significant, progress in the theory of optimization problems with perturbations.

The duality approach for posed problems are new. The paper is organized in the following order:

In Section 2, the needed facts and supplementary results from the book of Mahmudov [13] are given; infimal convolution of proper convex functions, conjugate function, etc. are introduced and the CPP with functional and non-functional constraints is formulated.

In Section 3 necessary and sufficient conditions of optimality for the problem CPP are formulated and new proofs are offered for this. In particular, we consider the convex optimization problem (PDL) with inequality and linear equality constraints defined by $C : \mathbb{R}^n \to \mathbb{R}^m$. In this case, it turns out that in the conditions of Theorem 3.1 obtained for problems CPP, the dual cone to the cone of tangent directions coincides with the image set of points of the transposed matrix C^*, taken with a minus sign. Many applications including numerical algorithms to solve problem CPP require so-called effective Slater's constraint qualification

conditions that ensure $y_0^* = 1$ in Theorem 3.1.
In Section 4 the duality theorems are proved and duality relation is established. Use of infimal convolution throughout this work plays a key role in proofs of duality results for problems CPP. The aim of the work is to establish conditions under which strong duality can be guaranteed. To this purpose, convexity is a compulsory requirement over the involved functional constraints and set in the primal problem. At the end of the section, we establish a dual theorem for the classical linear programming problem. In turn, this succeeds in new representations of the indicator function on the convex set. In what follows, we prove that if α and α^* are the values of primary and dual problems, respectively, then there is weak duality, i.e., $\alpha \geq \alpha^*$ for all feasible solutions. Moreover, if there is a "strictly feasible point", then the above statement can be strengthened, and the existence of a solution to one of these problems implies the existence of a solution to another problem in which strong duality holds, i.e., $\alpha = \alpha^*$, and in the case of finiteness of α , the dual problem has a solution.

2. Needed Facts and Problem Statement

Further, for the convenience of the reader, all the necessary concepts, definitions of a convex analysis can be found in the book of Mahmudov [13]. Let $\mathbb{R}^n$ be a n-dimensional Euclidean space, $\langle x, y\rangle$ be an inner product of elements $x, y \in \mathbb{R}^n$ and (x, y) be a pair of x, y; $x_k, y_k (k = 1, \dots, n)$ are the components of vectors x and y, respectively. The convex cone $K_A(x^0)$, is called the cone of tangent directions at a point $x^0 \in A$ to the set A if from $\bar{x} \in K_A(x^0)$ it follows that $\bar{x}$ is a tangent vector to the set A at point $x^0 \in A$, i.e., there exists such function $\eta : \mathbb{R}^1 \to \mathbb{R}^n$ that $x^0 + \lambda\bar{x} + \eta(\lambda) \in A$ for sufficiently small $\lambda > 0$ and $\lambda^{-1}\eta(\lambda) \to 0$, as $\lambda \downarrow 0$. In the convexity case of the set A, it is easy to see that

$$K_A(x^0) = \{\bar{x} : \bar{x} = \lambda(x - x^0),\ x \in A,\ \lambda > 0\}$$

is a cone of tangent directions at a point $x^0 \in A$. To do this, it suffices to set $\eta(\lambda) \equiv 0$ in the definition of the cone $K_A(x^0)$. A function φ is called a proper function if it does not assume the value $-\infty$ and is not identically equal to $+\infty$. Obviously, φ is proper if and only if $dom\varphi \neq \varnothing$ and $\varphi(x)$ is finite for $x \in dom\varphi = \{x : \varphi(x) < +\infty\}$.

Definition 2.1. A function $\varphi(x)$ is said to be a closure if its epigraph $epi\varphi = \{(x_0, x) : x_0 \geq \varphi(x), x_0 \in \mathbb{R}\}$ is a closed set.

Definition 2.2. The function $\varphi^*(x^*) = \sup_x \left\{\langle x, x^*\rangle - \varphi(x)\right\}$ is called the conjugate of φ. It is clear to see that the conjugate function is closed and convex.

Definition 2.3. For two proper functions f_1 and f_2, we associate the function f by the formula

$$\begin{aligned} f(x) &= inf\left\{f_1(x^1) + f_2(x^2) : x^1 + x^2 = x,\ x^1,\ x^2 \in \mathbb{R}^n\right\} \\ &= \inf_{x^1 \in \mathbb{R}^n} \left\{f_1(x^1) + f_2(x - x^1)\right\} \end{aligned}$$

and call the function f the infimal convolution of the functions f_1, f_2 and denote by $f = f_1 \Box f_2$.

In the mathematical literature, sometimes the infimal convolution is also denoted as $f = f_1 \oplus f_2$ or $f = f_1 \bigtriangledown f_2$. Throughout this article, we will adhere to the notation $f = f_1 \Box f_2$. The terminology infimal convolution arises from the classical formula for integral convolution. Note that the operation $\Box$ is commutative and associative, i.e., $f_1 \Box f_2 = f_2 \Box f_1$ and $(f_1 \Box f_2) \Box f_3 = f_1 \Box f_2 \Box f_3$, respectively. The infimal convolution $f_1 \Box f_2$ is said to be exact provided the infimum above is attained for every $x \in \mathbb{R}^n$. One has $\mathrm{dom}(f_1 \Box f_2) = dom f_1 + dom f_2$. Besides for a proper convex closed functions $f_i, i = 1, 2$ their infimal convolution $f_1 \Box f_2$ is convex and closed (but not necessarily proper). If $f_i, i = 1, 2$ are functions not identically equal to $+\infty$, then $(f_1 \Box f_2)^* = f_1^* + f_2^*$. Thus, the conjugate of infimal convolution is the sum of the conjugates and this holds without any requirement on the convex functions. The operations $+$ and $\Box$ are thus dual to each other with respect to taking conjugates. Section 3 is concerned with the following convex programming problem CPP labelled as (CPP):

$$\text{infimum } \varphi(x), \tag{2.1}$$

$$\text{(CPP)} \qquad \text{subject to} \quad f_k(x) \le 0, \; k = 1, \dots, N, \; x \in M, \tag{2.2}$$

where objective function φ and $f_k, k = 1, \dots, N$ are convex functions, M is a convex set and $M \subset dom f_k, k = 1, \dots, N, M \subset dom\varphi$. It is required find a point x^0 to a problem CPP, satisfying (2.2) and minimizing φ. In what follows, to this end our further strategy is as follows: first to derive necessary and sufficient conditions of optimality for problem CPP and then to derive duality results for them.

3. The optimality conditions for a convex CPP

Let us introduce the Lagrange function for problem CPP depending only the cost and functional constraints by

$$L(x, y^*) = \varphi(x) + \langle y^*, f(x) \rangle, \; y^* = (y_1^*, \dots, y_N^*) \in \mathbb{R}^N$$

and the active index set denoting by $I(x) = \{k \in \{1, \dots, N\} : f_k(x) = 0\}$.

Theorem 3.1. *Let x^0 be an optimal solution to CPP. Then there are multipliers $y_0^* \ge 0$ and $y^* = (y_1^*, \dots, y_N^*) \in \mathbb{R}^N$, not all equal to zero simultaneously, such that $y_k^* \ge 0, k = 0, \dots, N$,*

$$0 \in y_0^* \partial\varphi(x^0) + \sum_{k=1}^{N} y_k^* \partial f_k(x^0) - K_M^*(x^0),$$

where $K_M^*(x^0)$ is the dual cone to the cone of tangent directions $K_M(x^0)$ at a point $x^0 \in M$ and $y_k^* f_k(x^0) = 0, k = 1, \dots, N$.

Proof. Let us construct a function

$$f(x) = \sup\{\varphi(x) - \varphi(x^0), f_1(x), \dots, f_N(x)\},$$

for which obviously x^0 solves the following problem with a non-functional constraint

$$\text{infimum} \quad f(x) \quad \text{subject to} \quad x \in M.$$

By Theorem 3.2 [13] $\partial f(x^0)\cap K_M^*(x^0)\neq\varnothing$ or
$$0\in\partial f(x^0)-K_M^*(x^0). \tag{3.1}$$
Now, according to Theorem 1.32 [13] we have
$$\partial f(x^0)=co\Big[\partial\varphi(x^0)\cup\Big(\cup_{k\in I(x^0)}\partial f_k(x^0)\Big)\Big],$$
where $I(x^0)=\{k: f_k(x^0)=0\}$ is a set of active indexes. Then from (3.1) we derive that
$$0\in co\Big[\partial\varphi(x^0)\cup\Big(\cup_{k\in I(x^0)}\partial f_k(x^0)\Big)\Big]-K_M^*(x^0).$$
It follows that there are $y_0^*\geq 0$ and $y_k^*\geq 0, k\in I(x^0)$ such that $y_0^*+\sum_{k\in I(x^0)}y_k^*=1$, and
$$0\in y_0^*\partial\varphi(x^0)+\sum_{k=1}^{N}y_k^*\partial f_k(x^0)-K_M^*(x^0). \tag{3.2}$$
Setting $y_k^*=0, k\notin I(x^0)$ we obtain $y_k^*f_k(x^0)=0, k=1,\ldots,N$, where $(y_0^*,y_k^*)\neq 0$. □

Corollary 3.1. *Suppose we are considering the following problem without functional constraints*
$$\textit{infimum}\quad\varphi(x),\quad\textit{subject to}\quad x\in M.$$
Then the condition of Theorem 3.1 is converted as follows
$$\partial\varphi(x^0)\cap K_M^*(x^0)\neq\varnothing.$$

Proof. In fact, the absence of functional constraints is ensured, for example, in the case when $f_k, k=1\ldots,N$ are functions with negative values in the entire space. Then the conditions $y_k^*f_k(x^0)=0, k=1,\ldots,N$ is ensured, if the Lagrange multipliers are equals to zero, i.e., $y_k^*=0$ for all $k=1,\ldots,N$. On the other hand, since not all $y_0^*,y_1^*,\ldots,y_n^*$ equal to zero simultaneously, then $y_0^*>0$. As a result, by condition of Theorem 3.1 we have $0\in y_0^*\partial\varphi(x^0)-K_M^*(x^0)$. Dividing the left and right sides of this relationship by y_0^* and considering that $(1/y_0^*)K_M^*(x^0)=K_M^*(x^0)$, we derive $0\in\partial\varphi(x^0)-K_M^*(x^0)$ or $\partial\varphi(x^0)\cap K_M^*(x^0)\neq\varnothing$. □

Theorem 3.2. *Let's in problem CPP the Slater's constraint qualification holds, i.e. there exists $\bar{x}\in M$ such that $f_k(\bar{x})<0$ for all $k=1,\ldots,N$. Then x^0 is an optimal solution to CPP if and only if there exist nonnegative Lagrange multipliers $(y_1^*,\ldots,y_N^*)\in\mathbb{R}^N$ such that*
$$0\in\partial\varphi(x^0)+\sum_{k=1}^{N}y_k^*\partial f_k(x^0)-K_M^*(x^0). \tag{3.3}$$
$y_k^*f_k(x^0)=0, k=1,\ldots,N.$

Proof. To establish the necessary condition, it suffices to show that in (3.2) $y_0^*\neq 0$. Suppose on the contrary that $y_0^*=0$ and find $(y_1^*,\ldots,y_N^*)\neq 0, x^{k*}\in\partial f_k(x^0)$, and $x^*\in K_M^*(x^0)$ satisfying
$$\sum_{k=1}^{N}y_k^*x^{k*}-x^*=0$$

or

$$\sum_{k=1}^{N} y_k^* \langle x^{k*}, x - x^0 \rangle - \langle x^*, x - x^0 \rangle = 0.$$

Then since $f_k(x) - f_k(x^0) \geq \langle x^{k*}, x - x^0 \rangle$ and $\langle x^*, x - x^0 \rangle \geq 0, \forall x \in M$, immediately we have

$$0 = \sum_{k=1}^{N} y_k^* \langle x^{k*}, x - x^0 \rangle - \langle x^*, x - x^0 \rangle \leq \sum_{k=1}^{N} y_k^* \big(f_k(x) - f_k(x^0) \big),$$

from which it follows that

$$\sum_{k=1}^{N} y_k^* f_k(x) \geq 0, \ \forall x \in M. \tag{3.4}$$

On the other hand, since by Slater's condition

$$\sum_{k=1}^{N} y_k^* f_k(\bar{x}) < 0, \ \bar{x} \in M.$$

The inequality (3.4) contradicts the Slater condition.
Let us prove the sufficiency of condition (3.3) of theorem; choose $x^{0*} \in \partial \varphi(x^0)$, $x^{k*} \in \partial f_k(x^0), x^* \in K_M^*(x^0)$ such that $x^{0*} + \sum_{k=1}^{N} y_k^* x^{k*} - x^* = 0, \ y_k^* f_k(x^0) = 0$, with $y_k^* \geq 0, k = 1, \dots, N$. We now show, that optimality of x^0 in the problem CPP follows immediately from the definitions of cone of tangent directions and subdifferential. Actually, for any $x \in A$ we have

$$\begin{aligned} 0 = \textstyle\sum_{k=1}^{N} y_k^* \langle x^{k*}, x - x^0 \rangle - \langle x^*, x - x^0 \rangle \leq \textstyle\sum_{k=1}^{N} y_k^* \big(f_k(x) - f_k(x^0) \big), \\ + \varphi(x) - \varphi(x^0) = \textstyle\sum_{k=1}^{N} y_k^* f_k(x) + \varphi(x) - \varphi(x^0) \leq \varphi(x) - \varphi(x^0), \end{aligned}$$

that is $\varphi(x) - \varphi(x^0) \geq 0$ or $\varphi(x) \geq \varphi(x^0)$ for all $x \in A$. The proof of the theorem is complete. □

The previous theorems allow for further detailing if we concretize the way of specifying the set M . On this occasion we consider the convex optimization problem (PDL) with inequality and linear equality constraints

$$(PDL) \qquad \begin{aligned} &\text{infimum} \quad \varphi(x), \\ &\text{subject to} \quad f_k(x) \leq 0, \ k = 1, \dots, N, \\ &\qquad\qquad Cx = d, \end{aligned}$$

where $C : \mathbb{R}^n \to \mathbb{R}^m$ is an $m \times n$ matrix and $d \in \mathbb{R}^n$.

Theorem 3.3. *Suppose that Slater's constraint qualification holds for problem CPP with the set*

$$M = \{x \in \mathbb{R}^n : Cx - d = 0\}.$$

Then, for x^0 to be an optimal solution of (PDL), it is necessary and sufficient that there exist nonnegative multipliers $y_1^*, \dots, y_N^*$ such that

$$0 \in \partial \varphi(x^0) + \sum_{k=1}^{N} y_k^* \partial f_k(x^0) + imC^* \ \text{and} \quad y_k^* f_k^*(x^0) = 0, \ k = 1, \dots, N.$$

Proof. First of all we prove that

$$-K_M^*(x^0) = imC^* = \{x^* \in \mathbb{R}^n : x^* = C^*y,\ y \in \mathbb{R}^m\}, \tag{3.5}$$

where the adjoint mapping $C^* : \mathbb{R}^m \to \mathbb{R}^n$ defined by $\langle Cx^*, y\rangle = \langle x^*, C^*, y\rangle, x^* \in \mathbb{R}^n, y \in \mathbb{R}^m$ corresponds to the matrix transposition.). For all $x \in M$ we have

$$K_M(x^0) = \{\bar{x} : \bar{x} = \lambda(x - x^0), Cx = d, \lambda > 0\}$$
$$= \{\bar{x} : \bar{x} = \lambda(x - x^0), C(x - x^0) = 0, \lambda > 0\} = \{\bar{x} : C\bar{x} = 0\}.$$

In order to verify the inclusion " $\supset$ " in (3.5) take any $x^* \in \mathbb{R}^n$ with $x^* = C^*y$ for some $y \in \mathbb{R}^m$. Then considering $K_M(x^0) = \{\bar{x} : C\bar{x} = 0\}$, we have $\langle x^*, x - x^0\rangle = \langle C^*y, x - x^0\rangle = \langle y, Cx - Cx^0\rangle = 0$, from which it follows that $\bar{x} \in K_M(x^0)$ and consequently, $x^* \in -K_M^*(x^0)$ for all $\bar{x} \in -K_M(x^0)$. Now we need to check the opposite inclusion " $\subset$ ". Take $x^* \in -K_M^*(x^0)$ and get $\langle -x^*, x - x^0\rangle \geq 0$ for all x such that $Cx = d$. Fixing any $\bar{x} \in kerC = \{\bar{x} \in \mathbb{R}^n : C\bar{x} = 0\}$, it is easy to see that $C(x^0 - \bar{x}) = Cx^0 = d$, which yields $\langle x^*, \bar{x}\rangle \geq 0$ as $\bar{x} \in kerC$. Obviously, $-\bar{x} \in kerC$ and it follows that $\langle x^*, \bar{x}\rangle = 0$ for all $\bar{x} \in kerC$. Arguing by contradiction, assume that there is no $y \in \mathbb{R}^m$ with $x^* = C^*y$. Hence, $x^* \notin \Omega = C^*\mathbb{R}^m \subset \mathbb{R}^n$, where Ω is nonempty, closed, and convex set. Then by separation Theorem 1.5 [13] there exists nonzero point $\bar{x}^0 \in \mathbb{R}^n$ satisfying

$$\sup\left\{\langle \bar{x}^0, v\rangle : v \in \Omega\right\} < \langle \bar{x}^0, x^*\rangle.$$

But since $0 \in \Omega$ it follows that $\langle \bar{x}^0, x^*\rangle > 0$. On the other hand

$$\gamma\langle \bar{x}^0, C^*y\rangle = \langle \bar{x}^0, C^*(\gamma y)\rangle < \langle \bar{x}^0, x^*\rangle,\ \gamma \in \mathbb{R},\ y \in \mathbb{R}^m.$$

Therefore $\langle \bar{x}^0, C^*y\rangle = 0$ or $\langle C\bar{x}^0, y\rangle = 0$ as $y \in \mathbb{R}^m$, i.e., $C\bar{x}^0 = 0$. As a result $\bar{x}^0 \in kerC$, while $\langle \bar{x}^0, x^*\rangle > 0$. This contradiction justifies the inclusion $x^* \in imC^*$ in (3.5) or what is the same " $\subset$ ". This completes the proof of (3.5). Now we return to the proof of Theorem 3.3; the required proof follows immediately from Theorem 3.1 and formula (3.5), where the cone of tangent directions is calculated in terms of the image of the adjoint operator C^* taken with a minus sign. □

The result below is a classical version of the Lagrange multiplier rule for convex problems with differentiable properties.

Theorem 3.4. *Suppose that φ and $f_k, k = 1, \dots, N$ are differentiable functions at x^0 and that the gradient vectors $\left\{f_k'(x^0) : k \in I(x^0)\right\}$ are linearly independent. Then, for x^0 to be an optimal solution to problem CPP, with $M = \mathbb{R}^n$ it is necessary and sufficient that there exist nonnegative multipliers $y_1^*, \dots, y_N^*$ such that*

$$0 = \varphi'(x^0) + \sum_{k=1}^{N} y_k^* f_k'(x^0) = L_x(x^0, y^*) \text{ and } y_k^* f_k(x^0) = 0,\ k = 1, \dots, N.$$

Proof. Since $M = \mathbb{R}^n$ it follows that $K_M(x^0) = \mathbb{R}^n$ and as a result $K_M^*(x^0) = \{0\}$. Then the formula (3.3) implies $y_0^* = 0$, which contradicts the linear independence of the gradient vectors $\left\{f_k'(x^0) : k \in I(x^0)\right\}$. □

4. Infimal Convolution and Duality for a convex CPP

First, we formulate the following proposition, with the help of which we construct the dual problem and prove the duality theorems. We recall from convex analysis that the indicator function of a set is defined as follows:

$$\delta_A(x) = \begin{cases} 0, & x \in A, \\ +\infty, & x \notin A. \end{cases}$$

Proposition 4.1. *Let $f(x) = (f_1(x), f_2(x), \dots, f_N(x))$ be vector-function and $A = \{f_k(x) \le 0, k = 1, \dots, N, x \in M\}$. Then the indicator function $\delta_A(x)$ of the set can be represented as follows*

$$\delta_A(x) = \sup_{y^* \ge 0} \langle y^*, f(x) \rangle.$$

Proof. Let us denote $D = \{f_k(x) \le 0, k = 1, \dots, N\}$. Obviously, $A = D \cap M$. Then, if $x \in A$, we have $x \in D = \{f_k(x) \le 0, k = 1, \dots, N\}$ and $x \in M$. Hence, the supremum of the inner product $\langle y^*, f(x) \rangle$ is attained at $y^* = 0$. It means that $\delta_A(x) = 0, \forall x \in A$. Suppose now $x \notin A$. Then there are two cases: (2.1) either $x \notin D$ or $x \notin M$, (2.2) $x \notin D$ and $x \notin M$. In both the first and second cases, if $x \notin D$, then there is at least one k_0 for which $f_k(x) > 0$. Therefore, $\delta_A(x) = \sup_{y_k^* \ge 0} \sum_{k \ne k_0} y_k^* f_k(x) + y_k^* f_{k_0}(x)$ and recalling that $f_k(x) \le 0, k \ne k_0$ by assumption, we derive that $y_k^* = 0, k \ne k_0$ and $\delta_A(x) = y_{k_0}^* f_{k_0}(x)$. Tending now $y_{k_0}^*$ to $+\infty$ here, we have $\sup_{y^* \ge 0} \langle y^*, f(x) \rangle = +\infty$ i.e., $\delta_A(x) = +\infty$ when $x \notin A$. Note that in the case $x \in D, x \notin M$ as above the inner product $\langle y^*, f(x) \rangle$ is attained at $y^* = 0$. On the definition of the indicator function the required formula is proved. □

According to the results of convex analysis, it is known that the operations of addition and infimal convolution of convex functions are dual to each other [13, 26]. To this end, if there exists a point $x^0 \in A$ where φ is continuous (φ is continuous on the relative interior *ridom*φ, however, φ may have a point of discontinuity in its boundary), the problem (2.1), (2.2) can be converted as follows

$$\begin{aligned} \inf_{x \in A} \varphi(x) &= inf_{x \in A} \big\{ \varphi(x) + \delta_A(x) \big\} = -\sup_{x \in \mathbb{R}^n} \big\{ -\varphi(x) - \delta_A(x) \big\} \\ &= -\sup_{x \in \mathbb{R}^n} \big\{ \langle x, 0 \rangle - [\varphi(x) + \delta_A(x)] \big\} \\ &= -(\varphi + \delta_A)^*(0) = -\big(\varphi^* \Box \delta_A^* \big)(0) = -\inf_{x^*} \big\{ \varphi^*(x^*) + \delta_A^*(-x^*) \big\} \\ &= \sup_{x^*} \big\{ -\varphi^*(x^*) - \delta_A^*(-x^*) \big\} \end{aligned} \tag{4.1}$$

where $\delta_A(\cdot)$ is the indicator function of A. In general, it can be noticed that $(\varphi + \delta_A)^*(0) \le (\varphi^* \Box \delta_A^*)(0)$ and so

$$\inf_{x \in A} \varphi(x) \ge \sup_{x^*} \{ -\varphi^*(x^*) - \delta_A^*(-x^*) \}.$$

Then it is reasonable to announce that the dual problem to the primary problem (2.1), (2.2) has the form

$$\sup_{x^*} \{ -\varphi^*(x^*) - \delta_A^*(-x^*) \}. \tag{4.2}$$

In addition, if the value of the primal problem CPP is finite, then the supremum in the problem (4.2) is attained for all x^*. Thus, first of all, to ensure strong duality, it is necessary to check the fulfilment of the following equality

$$(\varphi + \delta_A)^*(0) = (\varphi^* \Box \delta_A^*)(0) \tag{4.3}$$

Denote now by $G(y^*)$ the infimum of the Lagrange function, i.e.,

$$G(y^*) = \inf_{x \in M} L(x, y^*).$$

The problem of maximizing $G(y^*)$ over all $y^* \geq 0$ is called the dual problem.

Theorem 4.1. *(Duality theorem) Suppose that φ and $f_k, k = 1, \dots, N$ are closed, proper and convex functions and all functions except possibly one are continuous at x^0 and these functions are finite at x^0. In addition, suppose that M is closed set. Then*

$$\inf_x \{\varphi(x) : x \in A\} = \sup_{y^*} \{G(y^*) : y^* \geq 0\},$$

that is, the exact lower bound in the primal problem coincides with the exact upper bound of the objective function in the dual problem.

Proof. From (4.1) we see that

$$\sup_{x^*} \{-\varphi^*(x^*) - \delta_A^*(-x^*)\} = -(\varphi^* \Box \delta_A^*)(0). \tag{4.4}$$

On the other hand, it is easy to see that under the conditions of theorem, δ_A is closed proper convex function and according to Theorem 3.15 [13] the equality (4.4) holds and, as a consequence (see (4.3)),

$$\sup_{x^*} \{-\varphi^*(x^*) - \delta_A^*(-x^*)\} = -(\varphi + \delta_A)^*(0). \tag{4.5}$$

Therefore, for each x^* the supremum is attained, i.e.,

$$(\varphi + \delta_A)^*(0) = \sup_x \left\{ \langle x, 0 \rangle - \big(\varphi(x) + \delta_A(x)\big) \right\} = -\inf_x \{\varphi(x) + \delta_A(x)\}. \tag{4.6}$$

Now applying Proposition 4.1 from (4.5), (4.6), according to the definition of the Lagrange function we obtain

$$\begin{aligned} sup_{x^*} \big\{ -\varphi^*(x^*) - \delta_A^*(-x^*) \big\} &= \inf_x \{\varphi(x) + \delta_A(x)\} \\ \inf_x \Big\{ \varphi(x) + &\sup_{y^* \geq 0} \langle y^*, f(x) \rangle \Big\} \\ = \sup_{y^* \geq 0} \inf_x \Big\{ \varphi(x) + \langle y^*, f(x) \rangle \Big\} &= \sup_{y^* \geq 0} G(y^*). \end{aligned} \tag{4.7}$$

In (4.7) considering that

$$\inf_x \{\varphi(x) + \delta_A(x)\} = \inf_x \{\varphi(x) : x \in A\}$$

we have the desired result, i.e.,

$$\inf_x \{\varphi(x) : x \in A\} = \sup_{y^* \geq 0} \big\{ G(y^*) : y^* \geq 0 \big\}.$$

□

Remark 4.1. Note that in the case of $M \neq \mathbb{R}^n$, in the dual problem (4.2) we have to take, $A = D \cap M$, where $D = \{f_k(x) \leq 0, k = 2, \dots, N\}$, and calculate $\delta_A^*(-x^*)$. But given the well-known fact that $\delta_A^*(-x^*)$ is a support function of i.e., $\delta_A^*(x^*) = \sup\{\langle x, x^* \rangle : x \in A\} \equiv W_A(x^*)$ and $W_A(x^*) \leq \min\{W_D(x^*), W_M(x^*)\}$, we can express the dual problem (4.2) in terms of support functions, which is typical for representing dual problems (see, for example [12, 14, 16, 17, 19, 20]).

Corollary 4.1. *The duality relation, i.e., the strong duality of Theorem 4.1 is equivalent to*

$$\inf_{x\in M}\sup_{y^*\geq 0} L(x,y^*) = \sup_{y^*\geq 0}\inf_{x\in M} L(x,y^*).$$

Proof. It is not hard to see that

$$\sup_{y^*\geq 0} L(x,y^*) = \begin{cases} \varphi(x), & \text{if} \quad f_k(x)\leq 0,\ k=1,\dots,N, \\ +\infty, & \text{if} \quad f_k(x)>0, \text{ for some } k \end{cases}$$

and the relation $\inf_x\{\varphi(x) : x\in A\} = \sup_{y^*}\{G(y^*) : y^*\geq 0\}$, of Theorem 4.1 is satisfied if and only if

$$\inf_{x\in M}\sup_{y^*\geq 0} L(x,y^*) = \sup_{y^*\geq 0}\inf_{x\in M} L(x,y^*).$$

□

Definition 4.1. A vector $y^{0*}\geq 0$ is called a Kuhn-Tucker vector, if the relation

$$\inf_x\{\varphi(x) : x\in A\} = \inf_{x\in M} L(x,y^{0*}).$$

holds.

Corollary 4.2. *Let the problem CPP has a vector Kuhn-Tucker y^{0*}. Then the duality relations of Theorem 4.1 hold and, moreover,*

$$\inf_x\{\varphi(x) : x\in A\} = G(y^{0*}) = \sup_{y^*\geq 0} G(y^*) = \inf_{x\in M} L(x,y^{0*}).$$

In particular, if x^0 is a solution of problem CPP, then $\varphi(x^0)\leq L(x,y^{0*})$.

Proof. It was shown above that $\varphi(x)\geq G(y^*)$ for all $y^*\geq 0$ and $x\in A$. Therefore,

$$G(y^{0*}) = \inf_x\{\varphi(x) : x\in A\}\geq G(y^*),\ y^*\geq 0.$$

□

Suppose B is an $m\times n$ matrix and $c\in\mathbb{R}^m, d\in\mathbb{R}^n$. Setting $\varphi(x)=\langle d,x\rangle, A=\{x : Bx\leq c\}, M=\mathbb{R}^n$, instead of CPP, we have the following linear programming problem,

$$\text{minimize} \quad \langle d,x\rangle, \text{ subject to } Bx\leq c. \tag{4.8}$$

Proposition 4.2. *The indicator function δ_A of the set $A=\{x : Bx\leq c\}$ has the form*

$$\delta_A(x) = \sup_{y^*\geq 0}\langle y^*, Bx-c\rangle.$$

Proof. An elementary exercise. □

Return to the linear programming problem (4.8). Clearly

$$\varphi^*(x^*) = \sup_x\langle x^*-d,x\rangle = \begin{cases} 0, & \text{if} \quad x^*=d, \\ +\infty, & \text{if} \quad x^*\neq d. \end{cases}$$

On the other hand, by Proposition 4.2, it is not hard to see that

$$\begin{aligned}\delta_A^*(-x^*) &= \sup_x\left\{\langle x,-x^*\rangle - \delta_A(x)\right\} = \sup_x\left\{\langle x,-x^*\rangle - \sup_{y^*\geq 0}\langle y^*,Bx-c\rangle\right\} \\ &= \sup_x\inf_{y^*\geq 0}\left\{\langle x,-x^*-B^*y^*\rangle + \langle y^*,c\rangle\right\}.\end{aligned}$$

Therefore,

$$\delta_A^*(-x^*) = \begin{cases} \inf_{y^*\geq 0}\langle y^*, c\rangle & \text{if} \quad x^* + B^*y^* = 0, \\ +\infty, & \text{if} \quad x^* + B^*y^* \neq 0. \end{cases}$$

Now, since $dom(\varphi^* \square \delta_A^*) = dom\varphi^* + dom\delta_A^*$, it follows that if $0 \in dom(\varphi^* \square \delta_A^*)$, i.e., $x^* + B^*y^* = 0$, then

$$\sup_{x^*} \Big\{ -\varphi^*(x^*) - \delta_A^*(-x^*) \Big\} = -\delta_A^*(-d) = \sup_{y^*\geq 0} \Big\{ \langle -y^*, c\rangle : d + B^*y^* = 0 \Big\}.$$

In conclusion, the relation (4.1) can be rewritten as the following known duality relation:

$$\inf_x \Big\{ \langle d, x\rangle : Bx \leq c \Big\} = \sup_{y^*\geq 0} \Big\{ \langle -y^*, c\rangle : d + B^*y^* = 0 \Big\}.$$

References

[1] R. Al-Salih, B. Martin, Quadratic programming on time scales, *Appl. Comput. Math.*, **19**(2) (2020), 205-219.

[2] R. Al-Salih, B. Martin, Linear programming problems on time scales, *Appl.Anal. Discrete Math.*, **12**(1) (2018), 192-204.

[3] A. Auslender, M. Teboulle, Interior projection-like methods for monotone variational inequalities, *Math. Program.*, **104**(1) (2003), 39-68.

[4] A. Berman, Cones, *Matrices and Mathematical Programming.* Springer, 1973.

[5] J.F. Bonnans, A. Shapiro, Optimization Problems with Perturbations: A Guided Tour, *Soc. Ind. Appl. Math.*, **40** (2) (1998), 228-264.

[6] R.I. Bot, G. Kassay, G. Wanka, Duality for almost convex optimization problems via the perturbation approach, *J. Glob. Optim.*, **42** (2008), 385-399.

[7] R.I. Bot, E.R. Csetnek, G.Wanka, Sequenial Optimality Conditions in Convex Programming via Perturbation Approach, CODE-2007 : Conference de la SMAI sur l'Optimisation et la Decision Institut Henri Poincare, Paris 18-20 April 2007

[8] R. Fletcher, S. Leyffer, Nonlinear programming without a penalty function, *Math. Program.*, **91** (2002), 239-270.

[9] K.H. Hoffmann, H.J. Kornstaedt, Higher-order necessary conditions in abstract mathematical programming. *J. Optim. Theory Appl.*, **26** (1978), 533-568.

[10] R. Horst, P. Pardalos. *Handbook of Global Optimization.* Kluwer, 1994.

[11] V. Jeyakumar, Z.Y. Wu, A qualification free sequential Pshenichnyi - Rockafellar Lemma and convex semidefinite programming, *J. Conv. Anal.*, **13**(3-4) (2006), 773-784.

[12] E.N.Mahmudov, B.N. Pshenichnyi, The optimality principle for discrete and differential inclusions of parabolic type with distributed parameters and duality, *Izvestiya: Math.*, **42** (2) (1994), 299.

[13] E.N. Mahmudov, *Approximation and Optimization of Discrete and Differential Inclusions.* Boston, MA, USA, Elsevier, 2011.

[14] E.N. Mahmudov, On duality in problems of optimal control described by convex differential inclusions of Goursat-Darboux type, *J. Math. Anal. Appl.*, **307** (2005), 628-640.

[15] E.N. Mahmudov, Sufficient conditions for optimality for differential inclusions of parabolic type and duality, *J. Glob. Optim.*, **41** (2008), 31-42.

[16] E.N. Mahmudov, On duality in second-order discrete and differential inclusions with delay, *J. Dyn. Contr. Syst.*, **26** (2020), 733-760.

[17] E.N. Mahmudov, Infimal convolution and duality in problems with third-order discrete and differential inclusions, *J. Optim. Theory Appl.*, **184** (2020), 781-809.

[18] E.N. Mahmudov, Optimal control of higher order differential inclusions with functional constraints, *ESAIM: Contr. Optim. Calcul. Variat.*, doi: https://doi.org/10.1051 /cocv/2019018.

[19] E.N. Mahmudov, M.J. Mardanov, On duality in optimal control problems with second-order differential inclusions and initial-point constraints. *Proceed. Inst. Math. Mech., Nation. Acad. Sci. Azerb.*, **46** (2020), 115-128.

[20] E.N. Mahmudov, Optimization of higher-order differential inclusions with endpoint constraints and duality, *Adv. Math. Models Appl.*, **6** (1) (2021), 5-21.

[21] E.N. Mahmudov, Infimal convolution and duality in convex optimal control problems with second order evolution differential inclusions, *Evol. Equ. Contr. Theory*, **10** (2021), 37-59.

[22] Y.E. Nesterov, A. Nemirovskii, *Interior-Point Polynomial Algorithms in Convex Programming*, SIAM, Stud. Appl. Numer. Math., 1994.

[23] Y.E. Nesterov, M.J. Todd, Self-Scaled Barriers and Interior-Point Methods for Convex Programming, *Math. Oper. Research*, **22** (1997), 1-4.

[24] B.N. Pshenichnyi, *Convex Analysis and Extremal Problems*, Moscow, "Nauka",1980 (In Russian).

[25] R. T. Rockafellar. Lagrange multipliers and optimality. *SIAM Review*, **35** (1993),183-283.

[26] R.T. Rockafellar, *Conjugate Duality and Optimization*, Philadelphia: Soc. Indust. App. Math., 1974.

[27] S.D. Sağlam, E.M. Mahmudov, Polyhedral optimization of second-order discrete and differential inclusions with delay. *Turkish J. Math.*, **45** (1) (2021), 244-263.

[28] A.L. Soyster, Convex Programming with Set-Inclusive Constraints and Applications to Inexact Linear Programming, *Oper. Research*, **21** (2015), 1154-1157.

[29] J. Stoer, C. Witzgall, *Convexity and optimization in finite dimensions*, 1, Springer, 1970.

[30] M.J. Todd, Y.Ye, On Adaptive Step Primal-Dual Interior-Point Algorithms for Linear Programming, *Math. Oper. Research*, **18** (1993), 964-981.

[31] M.J. Todd, Interior-point algorithms for semi-infinite programming (8/91). *Math. Program.*, **65** (1994), 217-245.

Elimhan N. Mahmudov

Department of Mathematics, Istanbul Technical University, 34469 Maslak, Istanbul, Turkey.

Institute of Control Systems, Azerbaijan National Academy of Sciences, AZ1141, Baku, Azerbaijan.

E-mail address: `elimhan22@yahoo.com`

Misir J. Mardanov

Institute of Mathematics and Mechanics, Azerbaijan National Academy of Sciences, AZ1141, Baku, Azerbaijan.

Baku State University, Baku, Azerbaijan.

E-mail address: `misirmardanov@yahoo.com`

Received: December 24, 2021; Revised: February 7, 2022; Accepted: February 11, 2022