

Sub-Homogeneous Optimization*

Shota Yamanaka[†] and Nobuo Yamashita[†]

August 29, 2017

Abstract

We consider an optimization problem with sub-homogeneous functions in its objective and constraint functions. Examples of such sub-homogeneous functions include the absolute value function and the p -norm function, where p is a positive real number. The problem, which is not necessarily convex, extends the absolute value optimization proposed in [O. L. Mangasarian, Absolute value programming, Computational Optimization and Applications 36 (2007) pp. 43–53]. In this work, we propose a dual formulation that, differently from the Lagrangian dual approach, has a closed-form and some interesting properties. In particular, we discuss the relation between the Lagrangian duality and the one proposed here, and give some sufficient conditions under which these dual problems coincide. Finally, we show that some well-known problems, e.g., sum of norms optimization and the group Lasso-type optimization problems, can be reformulated as sub-homogeneous optimization problems.

Keywords: Sub-homogeneous functions, duality, nonconvex optimization

1 Introduction

Recently, the so-called *absolute value equations* (AVE) and *absolute value optimization* (AVO) problems have been attracted much attention. The AVE were introduced in 2004 by Rohn [22]. Basically, if \tilde{A}, \tilde{B} are given matrices, and \tilde{b} is a given vector, one should find a vector x that satisfies $\tilde{A}x + \tilde{B}|x| = \tilde{b}$, where $|x|$ is a vector whose i -th entry is the absolute value of the i -th entry of x . It is known that AVE are equivalent to the *linear complementarity problems* (LCP) [10, 17, 21], which include many real-world applications. As an extension of AVE, Mangasarian [15] proposed in 2007 the AVO problems, which

*This work was supported in part by a Grant-in-Aid for Scientific Research (C) (17K00032) from Japan Society for the Promotion of Science.

[†]Department of Applied Mathematics and Physics, Graduate School of Informatics, Kyoto University, Kyoto 606-8501, Japan (shota@amp.i.kyoto-u.ac.jp, nobuo@i.kyoto-u.ac.jp).

have the absolute value of variables in their objective and constraint functions. More precisely, the AVO problem considered is given by

$$\begin{aligned} \min \quad & \tilde{c}^T x + \tilde{d}^T |x| \\ \text{s.t.} \quad & \tilde{A}x + \tilde{B}|x| = \tilde{b}, \\ & \tilde{H}x + \tilde{K}|x| \geq \tilde{p}, \end{aligned}$$

where $\tilde{A}, \tilde{B}, \tilde{H}, \tilde{K}$ are given matrices, and $\tilde{c}, \tilde{d}, \tilde{b}, \tilde{p}$ are vectors with appropriate dimensions. Since AVE and LCP are equivalent, the AVO include the *mathematical programs with linear complementarity constraints* [13], which are one of the formulations of equilibrium problems. As another application of AVO, Yamanaka and Fukushima [27] presented facility location problems.

Since 2007, some methods for solving AVE have been presented in the literature. For example, Rohn [23] considered an iterative algorithm using the sign of variables for the case that \tilde{A} and \tilde{B} are square matrices. For more general \tilde{A} and \tilde{B} , Mangasarian [15] provided a method involving successive linearization techniques. Another methods include a concave minimization approach, given by Mangasarian [14], and Newton-type methods, proposed by Caccetta et al. [4], Mangasarian [16], and Zhang and Wei [29]. Some generalizations of AVE were also proposed. For example, Hu et al. [11] considered an AVE involving the absolute value of variables associated to the second-order cones. Miao et al. [19] investigated an AVE with the so-called circular cones. In both papers, quasi-Newton based algorithms were used.

As for AVO problems, Yamanaka and Fukushima [27] proposed to use a branch-and-bound technique. In the branching procedure, two subproblems are generated by fixing the sign of a variable as nonnegative or nonpositive. In the bounding procedure, the dual information are considered. However, to the best of our knowledge, there is no other method that can find a global solution of AVO. When comparing to AVE, the research associated to AVO problems is insufficient and one of these reasons is the difficulty for obtaining feasible solutions of the problems. In fact, their constraints include AVE, which are known to be NP-hard [15].

Another optimization problem that is related to AVO was recently investigated by Friedlander et al. [9] and Aravkin et al. [2]. It is called *gauge optimization*, which basically consists in an optimization problem with the so-called gauge function. However, differently from AVO, this problem does not consider multiple constraints, but only one gauge constraint. In [2, 9], the authors showed that the Lagrange dual of gauge optimization problems can be written in a closed-form by using the polar of the gauge functions.

In this paper, similarly to [2, 9], we introduce a generalized AVO problem, and show that it has a wider practical application comparing to AVO problems. It is also more general than gauge optimization problems, because multiple constraints can be considered here. The generalization is done by replacing absolute value functions with *sub-homogeneous* functions. The concept of sub-homogeneous functions varies from each literature, but here we consider the one given in [3], with the term Λ as the set of non-negative real numbers. So, the problem uses not only absolute value terms but also, for instance, p -norm functions with $p \in (0, \infty]$. This generalized problem is referred here as

sub-homogeneous optimization (SHO).

Here, we introduce the SHO dual problem and compare it with the Lagrange dual. We also show that the weak duality theorem holds, similarly to the AVO problems [15]. In addition, we investigate the relation between the sub-homogeneous duality and the Lagrange duality, proving that these dual problems are equivalent under some conditions. In this case, the Lagrange dual of a sub-homogeneous problem can be written in a closed-form. We point out that the gauge functions are special cases of the sub-homogeneous functions, which are not necessarily convex, differently from the gauge. Moreover, the proposed problems here have linear and sub-homogeneous terms in their objective functions and constraints, which is different from the problem considered in [2, 9] that has only one gauge term. Here, we also give some applications for the sub-homogeneous problems, which include p -order cone optimization, sum of norms optimization and group Lasso-type optimization problems, and we show that their Lagrange dual can be written in a closed-form even without convexity assumptions.

The paper is organized as follows. In Section 2, we give the definition of sub-homogeneous functions as well as its dual, showing some of their properties. In Section 3, we define the SHO problems, and we prove that weak duality holds. In Section 4, the relation between the Lagrangian dual and the sub-homogeneous dual is discussed. We give some applications for SHO problems in Section 5. We conclude the paper in Section 6, with final remarks and some future works.

We consider the following notations throughout the paper. Let $x \in \mathbb{R}^n$ be a n -dimensional column vector, and $A \in \mathbb{R}^{n \times m}$ be a matrix with dimension $n \times m$. We use T to denote transpose. For two vectors x and y , we denote the vector $(x^T, y^T)^T$ as $(x, y)^T$ for simplicity. If $x \in \mathbb{R}^n$, then its i -th entry is denoted by x_i , so $x = (x_1, \dots, x_n)^T$. Moreover, if $I \subseteq \{1, \dots, n\}$, then x_I corresponds to the subvector of x with entries x_i , $i \in I$. The notation $\#J$ denotes the number of elements of a set J . The identity matrix with dimension n is given by $E_n \in \mathbb{R}^{n \times n}$. Also, we denote by $\|\cdot\|_p$ and $\|\cdot\|_\infty$ the p -norm with $p > 0$ and the supremum norm, respectively. If no distinction is made for the norm, we just use the notation $\|\cdot\|$.

2 Sub-homogeneous functions

In this section, we first introduce the definitions of sub-homogeneous and vector sub-homogeneous functions. They are special cases of Λ -sub-homogeneous functions originally given in [3], with Λ as the set of nonnegative real numbers \mathbb{R}_+ . Then, we define their dual, which will be used to describe the dual of SHO problems. Moreover, we show some properties associated to these functions.

Definition 1. (*Sub-homogeneous functions*) A function $\psi: \mathbb{R}^n \rightarrow \mathbb{R}$ is sub-homogeneous if the following inequality holds:

$$\psi(\lambda x) \leq \lambda \psi(x) \quad \text{for all } x \in \mathbb{R}^n, \lambda \in \mathbb{R}_+.$$

Definition 2. (*Vector sub-homogeneous functions*) A mapping $\Psi: \mathbb{R}^n \rightarrow \mathbb{R}^m$ is a vector sub-homogeneous function if the following property holds:

$$\Psi(x) = \begin{bmatrix} \psi_1(x_{I_1}) \\ \vdots \\ \psi_m(x_{I_m}) \end{bmatrix} \quad \text{for all } x \in \mathbb{R}^n,$$

where $\psi_i: \mathbb{R}^{n_i} \rightarrow \mathbb{R}$ is a sub-homogeneous function for all $i = 1, \dots, m$, $n = n_1 + \dots + n_m$, $I_i \subseteq \{1, \dots, m\}$ is a set of indices satisfying

$$I_i \cap I_j = \emptyset, \quad i \neq j, \quad \text{and} \quad \#I_i = n_i,$$

and $x_{I_i} \in \mathbb{R}^{n_i}$ is a disjoint subvector of x .

The above definition basically says that Ψ is vector sub-homogeneous if its block components are all sub-homogeneous. We now introduce the dual function of ψ , which can be seen as a generalization of the dual norm. Similarly, we also define the dual of vector sub-homogeneous functions.

Definition 3. (*Dual sub-homogeneous functions*) Let $\psi: \mathbb{R}^n \rightarrow \mathbb{R}$ be a sub-homogeneous function. Then, $\psi^*: \mathbb{R}^n \rightarrow \mathbb{R}$ defined by

$$\psi^*(y) := \sup\{x^T y \mid \psi(x) \leq 1\} \quad \text{for all } x \in \mathbb{R}^n$$

is called the dual sub-homogeneous function of ψ .

Note that ψ^* is convex from definition. In fact, for all $y, z \in \mathbb{R}^n$ and $\alpha \in (0, 1)$, we have

$$\begin{aligned} \psi^*(\alpha y + (1 - \alpha)z) &= \sup\{x^T(\alpha y + (1 - \alpha)z) \mid \psi(x) \leq 1\} \\ &\leq \alpha \sup\{x^T y \mid \psi(x) \leq 1\} + (1 - \alpha) \sup\{x^T z \mid \psi(x) \leq 1\} \\ &= \alpha \psi^*(y) + (1 - \alpha) \psi^*(z). \end{aligned}$$

Definition 4. (*Dual vector sub-homogeneous functions*) Let $\Psi: \mathbb{R}^n \rightarrow \mathbb{R}^m$ be a vector sub-homogeneous function. A function $\Psi^*: \mathbb{R}^n \rightarrow \mathbb{R}^m$ is a dual vector sub-homogeneous function associated to Ψ if the following property holds:

$$\Psi^*(y) = \begin{bmatrix} \psi_1^*(y_{I_1}) \\ \vdots \\ \psi_m^*(y_{I_m}) \end{bmatrix}, \quad i = 1, \dots, m, \quad \text{for all } y \in \mathbb{R}^n$$

where $\psi_i^*: \mathbb{R}^{n_i} \rightarrow \mathbb{R}$ is the dual of sub-homogeneous function ψ_i for each $i = 1, \dots, m$.

In this paper, we assume two conditions for sub-homogeneous functions.

Assumption 1. Let $\Psi: \mathbb{R}^n \rightarrow \mathbb{R}^m$ be a vector sub-homogeneous function as in Definition 2. Then, for all $i = 1, \dots, m$, the sub-homogeneous function ψ_i satisfies the following conditions:

1. $\psi_i(x_{I_i}) \geq 0$ for all $x_{I_i} \in \mathbb{R}^{n_i}$,
2. If $x_{I_i} \neq 0$, then $\psi_i(x_{I_i}) > 0$.

From the definition of sub-homogeneous functions, we observe that $\psi_i(0) = 0$. In fact, if $\lambda = 0$ then $\psi_i(0) \leq 0$. Moreover, if $x = 0$ and $\lambda = 2$, then $\psi_i(0) \geq 0$. Moreover, the second condition of the above assumption shows that zero is the only point that satisfies $\psi_i(x) = 0$. We also observe that if ψ_i is taken as the usual vector norm, then it satisfies these assumptions. We now show an important property satisfied by vector sub-homogeneous functions and their dual.

Proposition 1. *Let Ψ and Ψ^* be a vector sub-homogeneous function and its dual, respectively. Suppose that Assumption 1 holds. Then, the following inequalities hold:*

$$\begin{aligned}\Psi^*(y) &\geq 0, \\ \Psi(x)^T \Psi^*(y) &\geq x^T y\end{aligned}$$

for any $x, y \in \mathbb{R}^n$.

Proof. For simplicity, we take an arbitrary index i and denote ψ_i and x_{I_i} as ψ and x , respectively. From Definition 1, we have $\psi(0) = 0$. Using this result and Definition 3, we obtain

$$\psi^*(y) = \sup\{x^T y \mid \psi(x) \leq 1\} \geq 0 \quad \text{for all } y \in \mathbb{R}^n.$$

This shows that $\Psi^*(y) \geq 0$ for all $y \in \mathbb{R}^n$ from Definition 4.

If $x = 0$, then the second inequality of this proposition clearly holds. If $x \neq 0$, then $\psi(x) > 0$ from Assumption 1 and so

$$\psi\left(\frac{x}{\psi(x)}\right) \leq \frac{1}{\psi(x)}\psi(x) = 1$$

holds once again from Definition 1. Therefore, we obtain

$$\psi^*(y) \geq \left(\frac{x}{\psi(x)}\right)^T y \quad \text{for all } y \in \mathbb{R}^n.$$

Then, for all $x, y \in \mathbb{R}^n$, we have

$$\psi(x)\psi^*(y) \geq x^T y,$$

which indicates that

$$\Psi(x)^T \Psi^*(y) = \sum_{i=1}^m \psi_{I_i}(x) \psi_{I_i}^*(y) \geq \sum_{i=1}^m x_{I_i}^T y_{I_i} = x^T y.$$

□

3 Sub-homogeneous optimization problems

We consider the following *sub-homogeneous optimization* (SHO) problem:

$$(P) \quad \begin{aligned} \min \quad & c^T x + d^T \Psi(x) \\ \text{s.t.} \quad & Ax + B\Psi(x) = b, \\ & Hx + K\Psi(x) \geq p, \end{aligned}$$

where $c \in \mathbb{R}^n, d \in \mathbb{R}^m, b \in \mathbb{R}^k, p \in \mathbb{R}^\ell, A \in \mathbb{R}^{k \times n}, B \in \mathbb{R}^{k \times m}, H \in \mathbb{R}^{\ell \times n}$ and $K \in \mathbb{R}^{\ell \times m}$ are given constant vectors and matrices, and $\Psi: \mathbb{R}^n \rightarrow \mathbb{R}^m$ is a vector sub-homogeneous function satisfying Assumption 1.

Now we give the Lagrangian dual of the problem (P) as follows:

$$(D_{\mathcal{L}}) \quad \sup_{\substack{u \\ v \geq 0}} \omega(u, v),$$

where $\omega: \mathbb{R}^k \times \mathbb{R}^\ell \rightarrow \mathbb{R}$ is given by

$$(1) \quad \omega(u, v) := \inf_x \mathcal{L}(x, u, v),$$

and $\mathcal{L}: \mathbb{R}^n \times \mathbb{R}^k \times \mathbb{R}^\ell \rightarrow \mathbb{R}$ is the Lagrangian function of (P) defined by

$$\begin{aligned} \mathcal{L}(x, u, v) &:= c^T x + d^T \Psi(x) + u^T (b - Ax - B\Psi(x)) + v^T (p - Hx - K\Psi(x)) \\ &= b^T u + p^T v - (A^T u + H^T v - c)^T x + (d - B^T u - K^T v)^T \Psi(x), \end{aligned}$$

with $u \in \mathbb{R}^k$ and $v \in \mathbb{R}^\ell$ as the *Lagrange multipliers* associated to the equality and inequality constraints, respectively. Notice that it is difficult to write concretely the objective function of the problem $(D_{\mathcal{L}})$ because it is, in general, not convex with respect to x .

In order to obtain a closed-form dual problem, we consider a convex relaxation of the original problem (P) and its Lagrangian dual. For simplicity, we investigate the case where $\Psi(x) = |x| := (|x_1|, \dots, |x_n|)^T$, and (P) has a linear objective function and only inequality constraints. More precisely, we analyze the following problem:

$$(P_a) \quad \begin{aligned} \min \quad & c^T x \\ \text{s.t.} \quad & Ax + B|x| \geq b. \end{aligned}$$

If we set $x = x^+ - x^-$ and $|x| = x^+ + x^-$, where $x_i^+ = \max\{0, x_i\}$ and $x_i^- = \max\{0, -x_i\}$, then we can write (P_a) as

$$\begin{aligned} \min \quad & [c^T | - c^T] \begin{bmatrix} x^+ \\ x^- \end{bmatrix} \\ \text{s.t.} \quad & [A | -A] \begin{bmatrix} x^+ \\ x^- \end{bmatrix} + [B | B] \begin{bmatrix} x^+ \\ x^- \end{bmatrix} \geq b, \end{aligned}$$

which is equivalent to the following problem:

$$\begin{aligned}
\min \quad & [c^T | -c^T] \begin{bmatrix} y_1 \\ y_2 \end{bmatrix} \\
\text{s.t.} \quad & [A | -A] \begin{bmatrix} y_1 \\ y_2 \end{bmatrix} + [B | B] \begin{bmatrix} y_1 \\ y_2 \end{bmatrix} \geq b, \\
& y_1, y_2 \geq 0, \\
& y_1^T y_2 = 0,
\end{aligned}$$

where $y_1, y_2 \in \mathbb{R}^n$. Notice that the above problem is not convex due to the complementarity constraint $y_1^T y_2 = 0$. Therefore, we remove it from the problem and obtain the following relaxed one:

$$\begin{aligned}
\min \quad & [c^T | -c^T] y \\
\text{s.t.} \quad & [A + B | -A + B] y \geq b, \\
& y \geq 0,
\end{aligned}$$

where $y = (y_1, y_2)^T$. This problem is just a linear programming, then its Lagrangian dual can be written easily as

$$\begin{aligned}
\max \quad & b^T u \\
\text{s.t.} \quad & \begin{bmatrix} A^T + B^T \\ -A^T + B^T \end{bmatrix} u \leq \begin{bmatrix} c \\ -c \end{bmatrix}, \\
& u \geq 0.
\end{aligned}$$

Observing that the first constraint is equivalent to $|A^T u - c| + B^T u \leq 0$, we finally obtain the following closed-form dual problem:

$$\begin{aligned}
\max \quad & b^T u \\
\text{(D}_a\text{)} \quad & \text{s.t.} \quad |A^T u - c| + B^T u \leq 0, \\
& u \geq 0.
\end{aligned}$$

In fact, the problem (D_a) is the AVO dual of (P_a) proposed by Mangasarian in [15], and the weak duality clearly holds in this case.

Let us return to the general problem (P). Inspired by the above AVO dual problem (D_a), we consider the following problem as the sub-homogeneous dual problem:

$$\begin{aligned}
\max \quad & b^T u + p^T v \\
\text{(D)} \quad & \text{s.t.} \quad \Psi^*(A^T u + H^T v - c) + B^T u + K^T v \leq d, \\
& v \geq 0,
\end{aligned}$$

where Ψ^* is the dual vector sub-homogeneous function associated to Ψ . Note that (D) is a convex optimization problem since each component ψ_i^* of Ψ^* is a convex function.

The theorem below shows that the proposed dual problem (D) is reasonable, in the sense that the weak duality holds between (P) and (D).

Theorem 2. (*Weak duality*) For problems (P) and (D), the following inequality holds:

$$c^T x + d^T \Psi(x) \geq b^T u + p^T v$$

for all feasible points $x \in \mathbb{R}^n$ and $(u, v) \in \mathbb{R}^k \times \mathbb{R}^\ell$ of (P) and (D), respectively.

Proof . Let $x \in \mathbb{R}^n$ and $(u, v) \in \mathbb{R}^k \times \mathbb{R}^\ell$ be feasible for (P) and (D), respectively. Then, we have

$$\begin{aligned} c^T x + d^T \Psi(x) &\geq c^T x + (\Psi^*(A^T u + H^T v - c) + B^T u + K^T v)^T \Psi(x) \\ &= c^T x + \Psi^*(A^T u + H^T v - c)^T \Psi(x) + u^T B \Psi(x) + v^T K \Psi(x), \end{aligned}$$

where the inequality holds from the first constraint of (D) and the nonnegativity of Ψ . From the second inequality of Proposition 1, we also obtain:

$$\begin{aligned} c^T x + d^T \Psi(x) &\geq c^T x + (A^T u + H^T v - c)^T x + u^T B \Psi(x) + v^T K \Psi(x) \\ &= u^T (Ax + B \Psi(x)) + v^T (Hx + K \Psi(x)). \end{aligned}$$

Finally, the constraints of (P) gives

$$c^T x + d^T \Psi(x) \geq b^T u + p^T v,$$

which completes the proof. \square

The weak duality theorem itself is a powerful theoretical result, but it does not mention how large the duality gap between (P) and (D) is. And the duality gap can be large depending on problems, then the dual problem (D) may be useless. Therefore, in the next section, we investigate the relation between the Lagrangian dual problem $(D_{\mathcal{L}})$ and the one (D) proposed here. As a result, surprisingly, we find that $(D_{\mathcal{L}})$ and (D) are equivalent.

4 The sub-homogeneous duality and the Lagrangian duality

In this section, we consider the relation between the sub-homogeneous duality and the more traditional Lagrangian duality of the problem (P), investigating conditions under which the Lagrangian dual problem $(D_{\mathcal{L}})$ and the sub-homogeneous dual problem (D) are equivalent. Recalling (1), we first show a condition that makes $\omega(\bar{u}, \bar{v})$, the objective function of $(D_{\mathcal{L}})$, unbounded from below for some (\bar{u}, \bar{v}) .

Lemma 3. *Let ψ_i^* be the dual of the sub-homogeneous functions ψ_i for $i = 1, \dots, m$. Suppose that Assumption 1 holds. Also, assume that there exists (\bar{u}, \bar{v}) and an index i_0 satisfying*

$$\psi_{i_0}^*(\alpha_{I_{i_0}}) > \beta_{i_0},$$

where $\alpha := A^T \bar{u} + H^T \bar{v} - c \in \mathbb{R}^n$, and $\beta := d - B^T \bar{u} - K^T \bar{v} \in \mathbb{R}^m$. Then, there exists a sequence $\{x^k\}$ such that $\|x^k\| \rightarrow +\infty$ and $\mathcal{L}(x^k, \bar{u}, \bar{v}) \rightarrow -\infty$ as $k \rightarrow +\infty$. Therefore, $\omega(\bar{u}, \bar{v})$ is unbounded from below.

Proof . Firstly, we denote $\bar{\alpha}$ and $\bar{\alpha}(\lambda)$ as follows:

$$\begin{aligned} \bar{\alpha} &:= (\alpha_{I_1}, \alpha_{I_2}, \dots, \alpha_{I_{i_0}}, \dots, \alpha_{I_m}) \in \mathbb{R}^n, \\ \bar{\alpha}(\lambda) &:= (\alpha_{I_1}, \alpha_{I_2}, \dots, \lambda \hat{x}, \dots, \alpha_{I_m}) \in \mathbb{R}^n, \end{aligned}$$

where $\lambda \in \mathbb{R}_+$ and $\hat{x} \in \mathbb{R}^n$ is defined as the supreme point of the following problem:

$$\sup\{x^T \alpha_{I_{i_0}} \mid \psi_{i_0}(x) \leq 1\}.$$

From the definition of \hat{x} , we obtain $\psi_{i_0}(\hat{x}) \leq 1$ because the objective function $x^T \alpha_{I_{i_0}}$ is linear. Then, from Definition 3, we have

$$\hat{x}^T \alpha_{I_{i_0}} = \psi_{i_0}^*(\alpha_{I_{i_0}}) \geq \psi_{i_0}(\hat{x}) \psi_{i_0}^*(\alpha_{I_{i_0}}).$$

The above equality and the definition of the Lagrangian function give

$$\begin{aligned} \mathcal{L}(\bar{\alpha}(\lambda), \bar{u}, \bar{v}) &= b^T \bar{u} + p^T \bar{v} - \bar{\alpha}^T \bar{\alpha}(\lambda) + \beta^T \Psi(\bar{\alpha}(\lambda)) \\ &= b^T \bar{u} + p^T \bar{v} - \sum_{i \neq i_0} \alpha_{I_i}^T \alpha_{I_i} - \lambda \hat{x}^T \alpha_{I_{i_0}} + \sum_{i \neq i_0} \beta_i \psi_i(\alpha_{I_i}) + \beta_{i_0} \psi_{i_0}(\lambda \hat{x}) \\ &= \gamma - \lambda \hat{x}^T \alpha_{I_{i_0}} + \beta_{i_0} \psi_{i_0}(\lambda \hat{x}) \\ &\leq \gamma - \lambda \psi_{i_0}(\hat{x}) \psi_{i_0}^*(\alpha_{I_{i_0}}) + \beta_{i_0} \psi_{i_0}(\lambda \hat{x}), \end{aligned}$$

where $\gamma := b^T \bar{u} + p^T \bar{v} - \sum_{i \neq i_0} \alpha_{I_i}^T \alpha_{I_i} + \sum_{i \neq i_0} \beta_i \psi_i(\alpha_{I_i}) \in \mathbb{R}$ is constant with respect to λ . Moreover, Definition 1 shows that

$$\begin{aligned} \mathcal{L}(\bar{\alpha}(\lambda), \bar{u}, \bar{v}) &\leq \gamma - \lambda \psi_{i_0}(\hat{x}) \psi_{i_0}^*(\alpha_{I_{i_0}}) + \lambda \beta_{i_0} \psi_{i_0}(\hat{x}) \\ &= \gamma + \lambda \psi_{i_0}(\hat{x}) (\beta_{i_0} - \psi_{i_0}^*(\alpha_{I_{i_0}})) \\ &= \gamma + \lambda (\beta_{i_0} - \psi_{i_0}^*(\alpha_{I_{i_0}})). \end{aligned}$$

Therefore, $\mathcal{L}(\bar{\alpha}(\lambda), \bar{u}, \bar{v})$ converges to minus infinity when λ increases. Finally, if we set $x^k = \bar{\alpha}(\lambda^k)$ where $\lambda^k \rightarrow +\infty$ as $k \rightarrow +\infty$, then $\mathcal{L}(x^k, \bar{u}, \bar{v}) \rightarrow -\infty$ and we complete the proof. \square

We now show that the sub-homogeneous dual problem (D) and the Lagrangian one ($D_{\mathcal{L}}$) are equivalent under some conditions.

Lemma 4. *Suppose that Assumption 1 holds. Assume also that the sub-homogeneous dual problem (D) has a feasible solution $(\bar{u}, \bar{v}) \in \mathbb{R}^k \times \mathbb{R}^\ell$, and that there exist $x^* \in \mathbb{R}^n$ satisfying the following equality:*

$$(2) \quad (d - B^T \bar{u} - K^T \bar{v})^T \Psi(x^*) - (A^T \bar{u} + H^T \bar{v} - c)^T x^* = 0.$$

Then, the sub-homogeneous dual problem (D) and the Lagrangian dual problem ($D_{\mathcal{L}}$) are equivalent.

Proof. From Lemma 3, the function ω is unbounded from below if there exists an index i_0 such that $\psi_{i_0}^*(\alpha_{I_{i_0}}) > \beta_{i_0}$, where $\alpha := A^T \bar{u} + H^T \bar{v} - c \in \mathbb{R}^n$, and $\beta := d - B^T \bar{u} - K^T \bar{v} \in \mathbb{R}^m$. Therefore, the problem ($D_{\mathcal{L}}$) is equivalent to

$$(D'_{\mathcal{L}}) \quad \begin{aligned} &\sup \quad \omega(u, v) \\ &\text{s.t.} \quad \Psi^*(A^T u + H^T v - c) \leq d - B^T u - K^T v, \\ &\quad \quad v \geq 0. \end{aligned}$$

Let $(\bar{u}, \bar{v}) \in \mathbb{R}^k \times \mathbb{R}^\ell$ be the feasible solution of $(D'_\mathcal{L})$. From the definition of the Lagrangian function, we obtain:

$$\begin{aligned}\mathcal{L}(x, \bar{u}, \bar{v}) &= c^T x + d^T \Psi(x) + \bar{u}^T (b - Ax - B\Psi(x)) + \bar{v}^T (p - Hx - K\Psi(x)) \\ &= b^T \bar{u} + p^T \bar{v} - (A^T \bar{u} + H^T \bar{v} - c)^T x + (d - B^T \bar{u} - K^T \bar{v})^T \Psi(x).\end{aligned}$$

Then, taking $x^* \in \mathbb{R}^n$ that satisfies (2), we have

$$\mathcal{L}(x^*, \bar{u}, \bar{v}) = b^T \bar{u} + p^T \bar{v}.$$

Notice that x^* is the solution of the problem

$$\inf_x \mathcal{L}(x, \bar{u}, \bar{v}),$$

because $\mathcal{L}(x, \bar{u}, \bar{v}) \geq b^T \bar{u} + p^T \bar{v}$ holds from Proposition 1. Therefore, the problem $(D'_\mathcal{L})$ can be described as follows:

$$\begin{aligned}\sup \quad & b^T u + p^T v \\ \text{s.t.} \quad & \Psi^*(A^T u + H^T v - c) \leq d - B^T u - K^T v, \\ & v \geq 0,\end{aligned}$$

which is equivalent to the sub-homogeneous dual problem (D). \square

As a consequence of the above theorem, we obtain the following result.

Theorem 5. *Suppose that the Lagrangian dual problem $(D_\mathcal{L})$ has a feasible solution. Assume also that the vector sub-homogeneous function Ψ satisfies Assumption 1. Then, the sub-homogeneous dual problem (D) is equivalent to the Lagrangian dual problem $(D_\mathcal{L})$.*

Proof. From Definition 1 and Assumption 1, we have $\Psi(0) = 0$. It means that equation (2) holds at $x^* = 0$. Thus, the problems (D) and $(D_\mathcal{L})$ are equivalent. \square

The above theorem shows that the Lagrangian dual problem $(D_\mathcal{L})$ can be written in a closed-form when the function Ψ is sub-homogeneous and satisfies Assumption 1. The paper [15] does not show that the same property holds for the AVO problem. We now give it as a direct consequence of Theorem 5.

Corollary 1. *If the dual of an AVO problem has a feasible solution, then it is equivalent to the Lagrangian dual problem $(D_\mathcal{L})$.*

Proof. It holds from Theorem 5 and the fact that the absolute value function is sub-homogeneous and satisfies Assumption 1. \square

Corollary 2. *If the optimal value of an AVO primal problem and its Lagrangian dual problem $(D_\mathcal{L})$ are equivalent, then the strong duality holds between the AVO primal and the AVO dual problem.*

Proof. It holds straightforward from Corollary 1. \square

From the above result, AVO can be applied to solve 0-1 integer optimization problems. To solve such problems, their Lagrangian dual are often considered, which is, in general, nondifferentiable due to the integer constraints. On the other hand, a 0-1 integer constraint, that is $x \in \{0, 1\}$, is equivalent to $|2x - 1| = 1$. Then, 0-1 integer optimization problems can be reduced to AVO, and we obtain their AVO dual, which are actually *linear programming* (LP) problems. These LP dual problems are much easier to solve compared to the nondifferentiable ones. Therefore, it might be worth considering AVO dual problems from the computational point of view.

5 Examples of sub-homogeneous optimization problems

In this section, we present several applications that are formulated as SHO, and show their closed-form dual problems.

First, we observe that any p -norm function with $p \in [1, \infty)$ is sub-homogeneous. So, if ψ is the p -norm, then ψ^* becomes the q -norm, where $1/p + 1/q = 1$. Therefore, if ψ is taken as $\|\cdot\|_1, \|\cdot\|_2, \|\cdot\|_\infty$, then ψ^* becomes $\|\cdot\|_\infty, \|\cdot\|_2, \|\cdot\|_1$, respectively. Moreover, in the case that $p \in (0, 1)$, the dual function ψ^* is equal to $\|\cdot\|_\infty$ for all $p \in (0, 1)$, which is proved in Proposition 6 of Appendix A. From the result, we can consider any p -norm functions as ψ in SHO problems. And, even if such functions are nonconvex with $p \in (0, 1)$, the Lagrangian dual problem can be written in a closed-form from Theorem 5.

We now show some sub-homogeneous problems using these p -norm functions. The first example is the so-called linear second-order cone optimization problem [1], which is one of the famous convex optimization problem.

Example 1. Let $x = (x_1, x_2)^T \in \mathbb{R} \times \mathbb{R}^{n-1}$. Then, we consider the linear second-order cone optimization problem written by

$$(P_1) \quad \begin{aligned} \min \quad & c^T x \\ \text{s.t.} \quad & Ax = b, \\ & x_1 - \|x_2\|_2 \geq 0, \end{aligned}$$

where $c \in \mathbb{R}^n$, $A \in \mathbb{R}^{m \times n}$ and $b \in \mathbb{R}^m$. The above problem can be written in SHO form as

$$\begin{aligned} \min \quad & c^T x + 0^T \Psi(x) \\ \text{s.t.} \quad & Ax + 0\Psi(x) = b, \\ & Hx + K\Psi(x) \geq 0, \end{aligned}$$

with $H = (1, 0, \dots, 0) \in \mathbb{R}^{1 \times n}$, $K = (0, -1) \in \mathbb{R}^{1 \times 2}$ and $\Psi: \mathbb{R}^n \rightarrow \mathbb{R}^2$, $\Psi(x) = (|x_1|, \|x_2\|_2)^T$. Then, recalling (D), its dual problem is given by

$$\begin{aligned} \max \quad & b^T u \\ \text{s.t.} \quad & \Psi^*(A^T u + H^T v - c) + K^T v \leq 0, \\ & v \geq 0, \end{aligned}$$

where Ψ^* is identical to Ψ in this case. Then, from the definition of Ψ , we have

$$\begin{aligned} \max \quad & b^T u \\ \text{s.t.} \quad & |(A^T u)_1 + v - c_1| \leq 0, \\ & \|(A^T u)_2 - c_2\|_2 \leq v, \\ & v \geq 0, \end{aligned}$$

with $(A^T u)_1$ as the first component of $A^T u$, $(A^T u)_2$ is the rest of it, and $c = (c_1, c_2)^T \in \mathbb{R} \times \mathbb{R}^{n-1}$. The first constraint of the above problem shows that

$$v = c_1 - (A^T u)_1,$$

and $v \geq 0$ automatically holds from the second constraint. Then, we obtain

$$\begin{aligned} \max \quad & b^T u \\ \text{s.t.} \quad & \|(A^T u)_2 - c_2\|_2 \leq c_1 - (A^T u)_1 \end{aligned}$$

as the dual problem of (P_1) . In fact, the above problem is the standard dual of the linear second-order cone optimization problem [1].

Although we use the 2-norm in the above example, any p -norm function with $p \in (0, \infty]$ can be considered. In this case, if $p \in [1, \infty]$, then the primal and dual problems are p -order cone and q -order cone optimization problems, respectively, where $1/p + 1/q = 1$ [26]. If $p \in (0, 1)$, then the dual is ∞ -order cone optimization problem.

In the next example, we consider a gauge optimization problem, which is also a convex problem with multiple gauge functions in its objective and constraint functions. Here, we recall that f is a gauge function if and only if it is nonnegative, convex, positively homogeneous and satisfies $f(0) = 0$ [8]. For such a problem, we introduce its dual in SHO form.

Example 2. Let $x \in \mathbb{R}^n$. We consider the following problem:

$$\begin{aligned} (P_2) \quad \min \quad & \sum_{i=1}^s \alpha_i f_i(A_i x - a_i) \\ \text{s.t.} \quad & g_j(B_j x - b_j) \leq \beta_j, \quad j = 1, \dots, t, \end{aligned}$$

where $\alpha_i, \beta_j \in \mathbb{R}_+$, $A_i \in \mathbb{R}^{m_i \times n}$, $B_j \in \mathbb{R}^{k_j \times n}$, $a_i \in \mathbb{R}^{m_i}$ and $b_j \in \mathbb{R}^{k_j}$ are given for all $i = 1, \dots, s$ and $j = 1, \dots, t$, and $f_i: \mathbb{R}^{m_i} \rightarrow \mathbb{R}$ and $g_j: \mathbb{R}^{k_j} \rightarrow \mathbb{R}$ are gauge functions. Letting $y_i := A_i x - a_i$ and $z_j := B_j x - b_j$, (P_2) can be written as

$$\begin{aligned} \min \quad & \sum_{i=1}^s \alpha_i f_i(y_i) \\ \text{s.t.} \quad & g_j(z_j) \leq \beta_j, \quad j = 1, \dots, t, \\ & A_i x - y_i = a_i, \quad i = 1, \dots, s, \\ & B_j x - z_j = b_j, \quad j = 1, \dots, t. \end{aligned}$$

The above problem does not have a gauge function defined for the variable x , so we introduce such a gauge function $x \mapsto \psi(x)$ and rewrite the problem into the following way:

$$\begin{aligned} \min \quad & 0 \times \psi(x) + \sum_{i=1}^s \alpha_i f_i(y_i) + 0 \times \sum_{j=1}^t g_j(z_j) \\ \text{s.t.} \quad & 0 \times \psi(x) \leq 0, \\ & 0 \times f_i(y_i) \leq 0, \quad i = 1, \dots, s, \\ & g_j(z_j) \leq \beta_j, \quad j = 1, \dots, t, \\ & A_i x - y_i = a_i, \quad i = 1, \dots, s, \\ & B_j x - z_j = b_j, \quad j = 1, \dots, t. \end{aligned}$$

Note that $\psi: \mathbb{R}^n \rightarrow \mathbb{R}$ is a dummy gauge function with x as its domain.

Let

$$\hat{x} := (x, y_1, \dots, y_s, z_1, \dots, z_t) \in \mathbb{R}^{n + \sum_{i=1}^s m_i + \sum_{j=1}^t k_j}$$

and

$$\Psi(\hat{x}) := (\psi(x), f_1(y_1), \dots, f_s(y_s), g_1(z_1), \dots, g_t(z_t))^T.$$

Then the above problem can be rewritten as

$$\begin{aligned} \min \quad & d^T \Psi(\hat{x}) \\ \text{s.t.} \quad & K \Psi(\hat{x}) \leq p, \\ & \hat{A} \hat{x} = \hat{b}, \end{aligned}$$

where $d = (0, \alpha_1, \dots, \alpha_s, 0, \dots, 0)^T \in \mathbb{R}^{1+s+t}$, $p = (0, \dots, 0, \beta_1, \dots, \beta_t)^T \in \mathbb{R}^{1+s+t}$,

$$K = \begin{bmatrix} 0 & 0 \\ 0 & E_t \end{bmatrix}, \hat{A} = \left[\begin{array}{c|cc} A_1 & -E_{m_1} & \\ \vdots & & \ddots \\ A_s & & -E_{m_s} \\ \hline B_1 & & -E_{k_1} \\ \vdots & 0 & \ddots \\ B_t & & -E_{k_t} \end{array} \right], \text{ and } \hat{b} = \begin{bmatrix} a_1 \\ \vdots \\ a_s \\ b_1 \\ \vdots \\ b_t \end{bmatrix}.$$

Moreover, its sub-homogeneous dual problem is given by

$$\begin{aligned} \max \quad & \hat{b}^T u - p^T v \\ \text{s.t.} \quad & \Psi^*(\hat{A}^T u) - K^T v \leq d, \\ & v \geq 0. \end{aligned}$$

For simplification, let $u = (u_{11}, \dots, u_{1s}, u_{21}, \dots, u_{2t})^T$ with $u_{1i} \in \mathbb{R}^{m_i}, i = 1, \dots, s$ and $u_{2j} \in \mathbb{R}^{k_j}, j = 1, \dots, t$. Then the above problem is rewritten as

$$\begin{aligned} \max \quad & \sum_{i=1}^s a_i^T u_{1i} + \sum_{j=1}^t b_j^T u_{2j} - \sum_{\ell=1}^t \beta_\ell v_{1+s+\ell} \\ \text{(D}_2\text{)} \quad \text{s.t.} \quad & \sum_{i=1}^s A_i^T u_{1i} + \sum_{j=1}^t B_j^T u_{2j} = 0, \\ & f_i^*(-u_{1i}) \leq \alpha_i, \quad i = 1, \dots, s, \\ & g_j^*(-u_{2j}) \leq v_{1+s+j}, \quad j = 1, \dots, t. \end{aligned}$$

Notice that the last constraint implies $v \geq 0$ because g_j^* is also a gauge function. Moreover, (D₂) does not include the dual function ψ^* of the dummy gauge function ψ .

The next example is the group Lasso-type problems [18, 28], which is a special case of (P₂) and consist in unconstrained minimizations of the sum of certain norms. Such problems have many applications, in particular they appear in compressed sensing area [7, 24], where the sparsity of solutions are important. As an example, we consider a primal problem with p_1 -norm and p_2 -norm where $p_1, p_2 \in \mathbb{R}_+$, which are used in the regularization terms.

Example 3. Let $x \in \mathbb{R}^n$ and $p_1, p_2 \in \mathbb{R}_+$. We consider the following problem:

$$(P_3) \quad \min \quad \|Ax - b\|_2 + \lambda_1 \sum_{i=1}^{m'} \|x_{I_i}\|_{p_1} + \lambda_2 \sum_{i=m'+1}^m \|x_{I_i}\|_{p_2}$$

where $\lambda_1, \lambda_2 \in \mathbb{R}_+$, $b \in \mathbb{R}^m$, $A \in \mathbb{R}^{m \times n}$ and $0 < m' < m$.

Notice that the first term of the objective function of group Lasso-type problems are usually the square of 2-norm functions. However, it is not sub-homogeneous, so we removed the square and considered just the 2-norm functions.

We obtain the above problem by setting, in (P₂), $s = m + 1$,

$$\alpha_i = \begin{cases} \lambda_1, & \text{if } i = 1, \dots, m', \\ \lambda_2, & \text{if } i = m' + 1, \dots, m, \\ 1, & \text{if } i = m + 1, \end{cases}$$

$$A_i = \begin{cases} E_{I_i}, & \text{if } i = 1, \dots, m, \\ A, & \text{if } i = m + 1, \end{cases}$$

where E_{I_i} is a submatrix of E_n with E_j , $j \in I_i$ as its rows,

$$a_i = \begin{cases} 0, & \text{if } i = 1, \dots, m, \\ b, & \text{if } i = m + 1, \end{cases}$$

and

$$f_i(\cdot) = \begin{cases} \|\cdot\|_{p_1}, & \text{if } i = 1, \dots, m', \\ \|\cdot\|_{p_2}, & \text{if } i = m' + 1, \dots, m, \\ \|\cdot\|_2, & \text{if } i = m + 1. \end{cases}$$

Then, recalling (P₂) and (D₂), the dual of (P₃) can be written as

$$\begin{aligned} \max \quad & b^T u_{1(m+1)} \\ \text{s.t.} \quad & \sum_{i=1}^m E_{I_i}^T u_{1i} + A^T u_{1(m+1)} = 0, \\ & \|\cdot - u_{1i}\|_{q_1} \leq \lambda_1, \quad i = 1, \dots, m', \\ & \|\cdot - u_{1i}\|_{q_2} \leq \lambda_2, \quad i = m' + 1, \dots, m, \\ & \|\cdot - u_{1(m+1)}\|_2 \leq 1, \end{aligned}$$

where $q_i, i = 1, 2$ are obtained by

$$(3) \quad q_i = \begin{cases} \frac{p_i}{p_i - 1}, & \text{if } p_i > 1, \\ \infty, & \text{if } p_i \in (0, 1], \end{cases}$$

from Proposition 6 of Appendix A. Notice that the first equality constraint can be rewritten as

$$u_{1i} + (A^T)_{I_i} u_{1(m+1)} = 0, \quad i = 1, \dots, m.$$

Then, the above problem is described as

$$\begin{aligned} \max \quad & b^T u \\ \text{s.t.} \quad & \|(A^T)_{I_i} u\|_{q_1} \leq \lambda_1, \quad i = 1, \dots, m', \\ & \|(A^T)_{I_i} u\|_{q_2} \leq \lambda_2, \quad i = m' + 1, \dots, m, \\ & \| -u \|_2 \leq 1, \end{aligned}$$

where we denote $u_{1(m+1)}$ as u for simplicity.

The next example is also a Lasso-type problem. In this case, the objective function is a gauge, because the sum of gauge functions is also gauge. In order to obtain the dual of a gauge optimization problem, the polar of the objective function should be considered [2, 9]. However, it may be difficult to obtain the polar of a sum of gauge functions. To overcome this drawback, we use here the SHO framework.

Example 4. Let $x \in \mathbb{R}^n$ and $p_1, p_2 \in \mathbb{R}_+$. We consider the following problem:

$$(P_4) \quad \begin{aligned} \min \quad & \lambda_1 \|x\|_{p_1} + \lambda_2 \|x\|_{p_2} \\ \text{s.t.} \quad & \|Ax - b\|_2 \leq \beta, \end{aligned}$$

where $\lambda_1, \lambda_2, \beta \in \mathbb{R}_+$, $A \in \mathbb{R}^{m \times n}$ and $b \in \mathbb{R}^m$. The above problem can be obtained if we set, in (P_2) , $s = 2$, $t = 1$, $\alpha_1 = \lambda_1$, $\alpha_2 = \lambda_2$, $A_1 = A_2 = E_n$, $a_1 = a_2 = 0$, $B_1 = A$, $b_1 = b$, $f_1(\cdot) = \|\cdot\|_{p_1}$, $f_2(\cdot) = \|\cdot\|_{p_2}$, $g_1(\cdot) = \|\cdot\|_2$. Then, recalling (D_2) , the dual of (P_4) is written by

$$\begin{aligned} \max \quad & b^T u_{21} - \beta v_4 \\ \text{s.t.} \quad & u_{11} + u_{12} + A^T u_{21} = 0, \\ & \| -u_{11} \|_{q_1} \leq \lambda_1, \\ & \| -u_{12} \|_{q_2} \leq \lambda_2, \\ & \| -u_{21} \| \leq v_4, \end{aligned}$$

which is finally rewritten as

$$\begin{aligned} \max \quad & b^T u_2 - \beta v \\ \text{s.t.} \quad & \|u_1 + A^T u_2\|_{q_1} \leq \lambda_1, \\ & \| -u_1 \|_{q_2} \leq \lambda_2, \\ & \| -u_2 \| \leq v, \end{aligned}$$

where we set u_{12}, u_{21} and v_4 as u_1, u_2 and v , respectively, and q_1 and q_2 are defined in (3).

In order to control the sparsity of the solutions of the above Lasso-type problems, we can use any combination of p -norm functions, with $p \in (0, \infty]$, as the regularization terms. Especially, it is reported that the p -norm functions with $p \in (0, 1)$ in (P₃) is useful because they give sparser solutions than 1-norm functions [5, 6, 20].

We now give another example: the sum of norms optimization problems, which are generally nonconvex. Such problems have applications, for example, in facility location, where locations of new facilities should be decided by analyzing the distance between the new and the existing facilities [25]. Moreover, the problem of the following example can be applied not only to the minimization of the distance but also maximization of it by taking the constant λ_i as $-\lambda_i$. Such a situation can be found for instance in locating obnoxious facilities in residential areas.

Example 5. Let $x \in \mathbb{R}^n$. We consider the following problem:

$$(P_5) \quad \begin{aligned} \min \quad & \sum_{i=1}^s \lambda_i f_i(A_i x - a_i), \\ \text{s.t.} \quad & Bx \leq b, \end{aligned}$$

where $\lambda_i \in \mathbb{R}$, $A_i \in \mathbb{R}^{m_i \times n}$, $B \in \mathbb{R}^{k \times n}$, $a_i \in \mathbb{R}^{m_i}$ and $b \in \mathbb{R}^k$ are given, and $f_i: \mathbb{R}^{m_i} \rightarrow \mathbb{R}$, $i = 1, \dots, s$ are sub-homogeneous functions. We now introduce its sub-homogeneous dual by taking almost the same procedure as in Example 2. Let $y_i := A_i x - a_i$, then (P₅) is equivalent to

$$\begin{aligned} \min \quad & \sum_{i=1}^s \lambda_i f_i(y_i) \\ \text{s.t.} \quad & A_i x - y_i = a_i, \quad i = 1, \dots, s, \\ & Bx \leq b. \end{aligned}$$

By introducing additional constraints, we consider the following problem:

$$\begin{aligned} \min \quad & \sum_{i=1}^s \lambda_i f_i(y_i) \\ \text{s.t.} \quad & A_i x - y_i = a_i, \quad i = 1, \dots, s, \\ & Bx \leq b, \\ & c_i f_i(y_i) \leq d_i, \quad i = 1, \dots, s, \end{aligned}$$

where c_i and d_i are strictly positive constants. Notice that the additional constraints ensure the boundedness of the each term of the objective function especially when λ_i is strictly negative. Without such constraints, (P₅) can be unbounded depending on the linear constraint, and then its dual becomes infeasible. Note that the additional constraints do not change solutions, when we choose c_i and d_i so that the constraint $c_i f_i(y_i) \leq d_i$ will include reasonable solutions.

Let $\hat{x} := (x, y_1, \dots, y_s)^T \in \mathbb{R}^{n + \sum_{i=1}^s m_i}$ and $\Psi(\hat{x}) := (\psi(x), f_1(y_1), \dots, f_s(y_s))^T \in \mathbb{R}^{1+s}$, where $\psi(\cdot)$ is a dummy sub-homogeneous function. Then the above problem can be described as

$$\begin{aligned} \min \quad & d^T \Psi(\hat{x}) \\ \text{s.t.} \quad & \hat{A} \hat{x} = \hat{a}, \\ & H \hat{x} + K \Psi(\hat{x}) \geq p, \end{aligned}$$

where $d = (0, \lambda_1, \dots, \lambda_s)^T$, $\hat{a} = (a_1, \dots, a_s)^T$, $p = (-b, -d_1, \dots, -d_s)^T$,

$$\hat{A} = \begin{bmatrix} A_1 & -E_{m_1} & & 0 \\ \vdots & & \ddots & \\ A_s & 0 & & -E_{m_s} \end{bmatrix}, H = \begin{bmatrix} -B & 0 \\ 0 & 0 \end{bmatrix}, \text{ and } K = \left[\begin{array}{c|ccc} 0 & & & 0 \\ \hline & -c_1 & & \\ 0 & & \ddots & \\ & & & -c_s \end{array} \right].$$

Then, recalling the sub-homogeneous dual (D), the dual of the above problem can be written as

$$\begin{aligned} \max \quad & \hat{a}^T u + p^T v \\ \text{s.t.} \quad & \Psi^*(\hat{A}^T u + H^T v) \leq d - K^T v, \\ & v \geq 0, \end{aligned}$$

which is rewritten by

$$\begin{aligned} \max \quad & \sum_{i=1}^s a_i^T u_i - b^T v_1 - \sum_{i=1}^s d_i^T v_{i+1} \\ \text{s.t.} \quad & \sum_{i=1}^s A_i^T u_i - B^T v_1 = 0, \\ & f_i^*(-u_i) \leq \lambda_i + c_i, \quad i = 1, \dots, s, \\ & v \geq 0, \end{aligned}$$

where $v = (v_1, \dots, v_{s+1})^T$.

6 Conclusion

In this paper, we proposed an optimization problem with sub-homogeneous functions, which we call sub-homogeneous optimization problem. We also introduced its dual problem and showed the weak duality theorem between these problems. Moreover, we gave sufficient conditions for the equivalency between the proposed dual and the Lagrangian dual problems. Finally, we presented some examples of sub-homogeneous problems to show their value in real-world applications. One natural future work will be to propose methods that obtain approximate solutions of sub-homogeneous optimization problems. We believe the theoretical results described here are essential for that.

Acknowledgements The authors are grateful to Prof. Ellen. H. Fukuda for helpful comments and suggestions. This work was supported in part by a Grant-in-Aid for Scientific Research (C) (17K00032) from Japan Society for the Promotion of Science.

A Appendix

The following proposition shows that the dual of the p -norm function is the ∞ -norm even when p is less than 1.

Proposition 6. *Suppose that $p \in (0, 1)$. Then, the dual of the p -norm function is equal to the ∞ -norm.*

Proof. Let $y \in \mathbb{R}^n$ be an arbitrary vector. If $y = 0$, this proposition clearly holds. If $y \neq 0$, from Definition 3, we obtain

$$\begin{aligned}
\|y\|_p^* &= \sup\{x^T y \mid \|x\|_p \leq 1\} \\
&\leq \sup\{|x^T y| \mid \|x\|_p \leq 1\} \\
&\leq \sup\left\{\sum_{i=1}^n |x_i| |y_i| \mid \|x\|_p \leq 1\right\} \\
&\leq \max_j |y_j| \left(\sup\left\{\sum_{i=1}^n |x_i| \mid \|x\|_p \leq 1\right\}\right) \\
&= \max_j |y_j| \left(\sup\{\|x\|_1 \mid \|x\|_p \leq 1\}\right).
\end{aligned}$$

Since $p \in (0, 1)$, we note that $\|x\|_1 \leq \|x\|_p$ holds [12]. Then, we have

$$\|y\|_p^* \leq \max_j |y_j| \left(\sup\{\|x\|_p \mid \|x\|_p \leq 1\}\right) = \max_j |y_j| = \|y\|_\infty.$$

Now, take an arbitrary $i_0 \in \operatorname{argmax}_i |y_i|$, and define \bar{x}_i as follows:

$$\bar{x}_i = \begin{cases} \operatorname{sign}(y_{i_0}), & \text{if } i = i_0, \\ 0, & \text{otherwise,} \end{cases}$$

where

$$\operatorname{sign}(y_{i_0}) = \begin{cases} 1, & \text{if } y_{i_0} > 0, \\ 0, & \text{if } y_{i_0} = 0, \\ -1, & \text{if } y_{i_0} < 0. \end{cases}$$

Then, $\|\bar{x}\|_p = 1$ and we have

$$\|y\|_p^* = \sup\{x^T y \mid \|x\|_p \leq 1\} \geq \bar{x}^T y = \max_i |y_i| = \|y\|_\infty,$$

which completes the proof. □

Reference

- [1] F. Alizadeh and D. Goldfarb. Second-order cone programming. *Mathematical programming*, 95(1):3–51, 2003.
- [2] A. Y. Aravkin, J. V. Burke, D. Drusvyatskiy, M. P. Friedlander, and K. MacPhee. Foundations of gauge and perspective duality. *arXiv preprint arXiv:1702.08649*, 2017.

- [3] P. Burai and A. Száz. Relationships between homogeneity, subadditivity and convexity properties. *Publikacije Elektrotehničkog fakulteta. Serija Matematika*, 16:77–87, 2005.
- [4] L. Caccetta, B. Qu, and G. Zhou. A globally and quadratically convergent method for absolute value equations. *Computational Optimization and Applications*, 48(1):45–58, 2011.
- [5] R. Chartrand. Exact reconstruction of sparse signals via nonconvex minimization. *IEEE Signal Processing Letters*, 14(10):707–710, 2007.
- [6] R. Chartrand and W. Yin. Iteratively reweighted algorithms for compressive sensing. In *Acoustics, speech and signal processing, 2008. ICASSP 2008. IEEE international conference on*, pages 3869–3872. IEEE, 2008.
- [7] Y. C. Eldar and M. Mishali. Robust recovery of signals from a structured union of subspaces. *IEEE Transactions on Information Theory*, 55(11):5302–5316, 2009.
- [8] R. M. Freund. Dual gauge programs, with applications to quadratic programming and the minimum-norm problem. *Mathematical Programming*, 38(1):47–67, 1987.
- [9] M. P. Friedlander, I. Macedo, and T. K. Pong. Gauge optimization and duality. *SIAM Journal on Optimization*, 24(4):1999–2022, 2014.
- [10] S. Hu and Z. Huang. A note on absolute value equations. *Optimization Letters*, 4(3):417–424, 2010.
- [11] S. Hu, Z. Huang, and Q. Zhang. A generalized Newton method for absolute value equations associated with second order cones. *Journal of Computational and Applied Mathematics*, 235(5):1490–1501, 2011.
- [12] U. S. Kirmaci, M. K. Bakula, M. E. Özdemir, and J. E. Pecaric. On some inequalities for p -norms. *Journal of Inequalities in Pure & Applied Mathematics*, 9(1):1–8, 2008.
- [13] Z.-Q. Luo, J.-S. Pang, and D. Ralph. *Mathematical Programs with Equilibrium Constraints*. Cambridge University Press, 1996.
- [14] O. L. Mangasarian. Absolute value equation solution via concave minimization. *Optimization Letters*, 1(1):3–8, 2007.
- [15] O. L. Mangasarian. Absolute value programming. *Computational Optimization and Applications*, 36(1):43–53, 2007.
- [16] O. L. Mangasarian. A generalized Newton method for absolute value equations. *Optimization Letters*, 3(1):101–108, 2009.
- [17] O. L. Mangasarian and R. R. Meyer. Absolute value equations. *Linear Algebra and Its Applications*, 419(2):359–367, 2006.

- [18] L. Meier, S. van de Geer, and P. Bühlmann. The group Lasso for logistic regression. *Journal of the Royal Statistical Society: Series B (Statistical Methodology)*, 70(1):53–71, 2008.
- [19] X. Miao, J. Yang, and S. Hu. A generalized Newton method for absolute value equations associated with circular cones. *Applied Mathematics and Computation*, 269:155–168, 2015.
- [20] N. Mourad and J. P. Reilly. Minimizing nonconvex functions for sparse vector reconstruction. *IEEE Transactions on Signal Processing*, 58(7):3485–3496, 2010.
- [21] O. Prokopyev. On equivalent reformulations for absolute value equations. *Computational Optimization and Applications*, 44(3):363–372, 2009.
- [22] J. Rohn. A theorem of the alternatives for the equation $ax + b|x| = b$. *Linear and Multilinear Algebra*, 52(6):421–426, 2004.
- [23] J. Rohn. An algorithm for solving the absolute value equation. *Electronic Journal of Linear Algebra*, 18(5):589–599, 2009.
- [24] M. Stojnic, F. Parvaresh, and B. Hassibi. On the reconstruction of block-sparse signals with an optimal number of measurements. *IEEE Transactions on Signal Processing*, 57(8):3075–3085, 2009.
- [25] G. W. Wolf. *Facility Location: Concepts, Models, Algorithms and Case Studies*. Taylor & Francis, 2011.
- [26] G. Xue and Y. Ye. An efficient algorithm for minimizing a sum of p -norms. *SIAM Journal on Optimization*, 10(2):551–579, 2000.
- [27] S. Yamanaka and M. Fukushima. A branch-and-bound method for absolute value programs. *Optimization*, 63(2):305–319, 2014.
- [28] M. Yuan and Y. Lin. Model selection and estimation in regression with grouped variables. *Journal of the Royal Statistical Society: Series B (Statistical Methodology)*, 68(1):49–67, 2006.
- [29] C. Zhang and Q. J. Wei. Global and finite convergence of a generalized Newton method for absolute value equations. *Journal of Optimization Theory and Applications*, 143(2):391–403, 2009.