Research Article

Proximal Point Algorithms for Vector DC Programming with Applications to Probabilistic Lot Sizing with Service Levels

Ying Ji and Shaojian Qu

Business School, University of Shanghai for Science and Technology, Shanghai, China

Correspondence should be addressed to Shaojian Qu; qushaojian@163.com

We present a new algorithm for solving vector DC programming, where the vector function is a function of the difference of C-convex functions. Because of the nonconvexity of the objective function, it is difficult to solve this class of problems. We propose several proximal point algorithms to address this class of problems, which make use of the special structure of the problems (i.e., the DC structure). The well-posedness and the global convergence of the proposed algorithms are developed. The efficiency of the proposed algorithm is shown by an application to a multicriteria model stemming from lot sizing problems.

1. Introduction

Vector-valued optimization arises from multiobjective programming, multicriteria decision-making, statistics, and cooperative game theory [1]. This class of problems has been widely applied and studied in various decision-making contexts and many methods have been developed to address such problems [2–4]. Denote C as a convex, closed, and pointed cone with $\mathrm{int}(C) \neq \emptyset$. The partial order $\preceq$ can be defined as follows: for any $u, v \in R^m$, $u \preceq v \Leftrightarrow v - u \in C$. The relation $\prec$ is defined as $u \prec v \Leftrightarrow v - u \in \mathrm{int}(C)$. This paper studies the following vector DC optimization problems with linear constraints:

$$\min_C \left\{ f(x) := \psi(x) - \varphi(x) : \text{s.t. } x \in S \right\}, \tag{1}$$

where functions φ and ψ: $R^n \rightarrow R^m$ are C-convex (see the definition in Section 2); $S \subset R^n$ is a closed and convex set. Note that "$\min_C$" implies the multiobjective minimization over the cone C, that is, the weak Pareto efficiency or Pareto efficiency with respect to the cone C. When $C := R^m_+$, the objective function f reduces to the multiobjective DC function and problem (1) is the multiobjective DC program which is studied by Qu et al. [5]. Similar to the adoption in [5], we utilize the following convention: $\infty - \infty = \infty$.

Though many researchers have widely studied vector DC program, they analyze mainly the theoretical aspects. For example, the duality theories and optimal conditions for multiobjective optimization were presented by Qu et al. [5]; the optimal conditions for DC vector optimization were also studied by Gadhi et al. [6]. However, to the best of our knowledge, there are few studies on how to efficiently solve this class of problems. This paper presents several proximal point algorithms for solving this class of problems.

Proximal point algorithms have an extensive application in solving scalar optimization problems [7], which originates from the work of [8]. It has been shown that proximal point algorithms can efficiently solve scalar optimization problems [9, 10]. As an extension, recently, the proximal point algorithm has been utilized to solve multicriteria convex optimization problems [11, 12]. However, to the best of our knowledge, few studies have been paid on extending proximal point algorithms to multicriteria optimization problems with DC objective functions. This paper proposes the exact and inexact proximal point algorithms to solve vector DC optimization, where the special structure, the DC structure of this class of problems, has been considered.

Our contributions can be concluded as follows. The proximal point methods in both exact and inexact forms are extended to address the vector DC optimization. The well-definedness of the proposed methods is presented. The global convergence of the proposed algorithms is proved. An

application in lot sizing problems is presented to show the efficiency of the proposed algorithms.

We organize the rest of this paper as follows. Preliminaries are presented in Section 2. The algorithms are proposed in Section 3. The theoretical analysis of the well-definedness and global convergence are also discussed in Section 3. An application of the algorithms to probabilistic lot sizing with service levels is presented in Section 4. The conclusion is given in Section 5.

2. Preliminaries

Function $\phi : R^n \to R^m$ is denoted as $C - convex$ iff, for any $x, y \in R^n$ and any $\lambda \in [0, 1]$,

$$\phi \left(\lambda x + (1 - \lambda) y \right) \preceq \lambda \phi (x) + (1 - \lambda) \phi (y) . \tag{2}$$

Function ϕ is denoted as strictly $C - convex$ iff, for any $x, y \in R^n$ and any $\lambda \in (0, 1)$,

$$\phi \left(\lambda x + (1 - \lambda) y \right) \prec \lambda \phi (x) + (1 - \lambda) \phi (y) . \tag{3}$$

Given $\epsilon \geq 0$, the ϵ-subdifferential $\partial_\epsilon \phi(\overline{x})$ of ϕ at $\overline{x} \in \operatorname{dom} \phi$ is a possibly empty, closed, and convex set and can be defined as

$$\partial_\epsilon \phi (x) := \left\{ V \in R^{m \times n} : \phi (x) + V (y - x) \preceq \phi (y) \right. \\ \left. + \epsilon e, \ \forall y \in R^n \right\}, \quad \epsilon \geq 0, \tag{4}$$

where $e := (1, \ldots, 1)^T \in R^m$. We denote $\partial_\epsilon \phi(\overline{x}) := \emptyset$ for $\overline{x} \notin \operatorname{dom} \phi$ and $\partial \phi_\epsilon(\overline{x})$ is nonempty convex and compact (see Theorem 4.12 of [13]) if $x \in \operatorname{dom} \phi$. When $m = 1$, $\epsilon = 0$, and $C = R_+$, $\partial \phi_0(\overline{x})$ reduces to the usual classical subdifferential [14]. The indicator function of a set $S \subset R^n$ can be named as $\delta(x; S) := 0$ if $x \in S$ and $\delta(x; S) := +\infty$, if $x \notin S$. We denote $\overline{x} \in S$ as the normal cone to the set S which can be described as

$$N (\overline{x}; S) := \partial \delta (\overline{x}; S) \\ = \left\{ x^* \in R^n \mid (x - \overline{x})^T x^* \leq 0, \ \forall x \in S \right\}. \tag{5}$$

To facilitate the discussion in this paper, we need the following information about refined epigraphical and subdifferential rules in convex analysis presented in [15].

Lemma 1. *Assume that $\phi_i : \Omega \to \overline{R}$, $i = 1, 2$, are l.s.c. and proper. If $\operatorname{dom} \phi_1 \cap \operatorname{dom} \phi_2 \neq \emptyset$, then the following conclusions are equivalent:*

(i) *The set $epi\phi_1^* + epi\phi_2^*$ is weak* closed in $\Omega^* \times R$.*

(ii) *The refined conjugate epigraphical rule holds: $epi(\phi_1 + \phi_2)^* = epi\phi_1^* + epi\phi_2^*$.*

Throughout of this paper, the equivalent conditions in Lemma 1 are supposed to hold which means that the following subdifferential sum rule holds, $\partial(\phi_1 + \phi_2)(\overline{x}) = \partial \phi_1(\overline{x}) + \partial \phi_2(\overline{x})$, $\forall x \in \operatorname{dom} \phi_1 \cap \operatorname{dom} \phi_2$ [15].

In multiobjective problem, no unique solution can minimize all of objective functions simultaneously; therefore, the decision-maker has to trade-off solutions. That is, if he wants to improve some objectives, he has to give up other objectives. So the optimal concept has to be replaced by the Pareto optimal concept, which can be defined as follows: $x^* \in S$ is called a Pareto optimum (PO) or a weak Pareto optimum (WPO) of (1) if there is no strategy $x \in S$ such that

$$f (x) \preceq f (x^*),$$
$$f (x) \neq f (x^*) \tag{6}$$
$$\text{or } f (x) \prec f (x^*).$$

x^* is named as a local PO or a local WPO of (1) if and only if $\exists N(x^*)$ as a neighborhood of x^* such that x^* is a PO or a WPO in $N(x^*) \cap S$.

To analyze the optimization conditions of (1), we put forward the following assumption.

(A1) Assume that functions φ and ψ: $R^n \to R^m$ are C-convex. Assume that $\overline{x}$ is a locally Pareto efficient solution of (1) and is limited to $S \cap \operatorname{dom} \psi \neq \emptyset$.

Under the assumption that $S \cap \operatorname{dom} \psi \neq \emptyset$, it is guaranteed that $\partial \psi(\overline{x}) \neq \emptyset$. Assumption (A1) will be used to prove Theorem 2.

Theorem 2. *Suppose that the assumption (A1) holds and $\overline{x}$ is a locally WPO to (1), then $\exists u \in C^*$, such that*

$$\partial \varphi (\overline{x})^T u \subset \partial \psi (\overline{x})^T u + N (\overline{x}; S), \quad \|u\| = 1. \tag{7}$$

Proof. This theorem can be proved by considering the two cases about $\overline{x} \in S \cap \operatorname{dom} \psi$: either $\overline{x} \notin \operatorname{dom} \varphi$ or $\overline{x} \in \operatorname{dom} \varphi$. The first one means that $\partial \varphi(\overline{x}) = \emptyset$, which leads to (7). For the other, define $F_u(x) := u^T f(x)$, $\forall u \in C^*$ with $\|u\| = 1$ and $\forall x \in S$. Given $\overline{u} \in C^*$ with $\|\overline{u}\| = 1$, then we can prove that $\overline{x}$ is a local optimal solution to the following scalar optimization problem:

$$\min_{x \in S} \quad F_{\overline{u}} (x). \tag{8}$$

We prove it by contradiction; that is, the conclusion does not hold. Then $F_u(\overline{x}) > F_u(\tilde{x})$ holds for at least one $\tilde{x} \in S \cap B_\varepsilon(\overline{x})$ with any given neighborhood $B_\varepsilon(\overline{x}) \subset S$ of $\overline{x}$, $\varepsilon > 0$, and any $u \in C^*$ with $\|u\| = 1$. It is easy to show that $f(\tilde{x}) \prec f(\overline{x})$ for the choice of u, which contradicts the assumption about $\overline{x}$ and this implies that $\overline{x}$ is also a local optimum to (8). To show that (7) is true, we only need to prove that $V^T \overline{u} \in \partial \psi(\overline{x})^T \overline{u} + N(\overline{x}; S)$ holds $\forall V \in \partial \varphi(\overline{x}) \subseteq R^{m \times n}$. According to the definition of subdifferential in $\overline{x}$, it is easy to learn that $\overline{x}$ is also globally solving the following convex optimization problem:

$$\min \left\{ \overline{F} (x) := \overline{u}^T \left(\psi (x) - \varphi (\overline{x}) \right) - \overline{u}^T V (x - \overline{x}) : x \right. \\ \left. \in S \right\}. \tag{9}$$

Then, from the necessary optimal conditions of (9), the assertion of this theorem can be followed directly. □

Usually, it is difficult to utilize the necessary condition in the above theorem to design the solution algorithms for

(1). Hence, the other optimal condition should be presented and this paper looks for finding a critical point of (1), that is, finding a point $\overline{x}$, such that

$$\overline{x} \in X := \left\{ x \in R^n : \exists u \in C^*, \text{ s.t., } 0 \in \partial \psi (x)^T u \right.$$

$$\left. - \partial \varphi (x)^T u, \ \|u\| = 1 \right\}. \tag{10}$$

Define $U := \{ u \in C^* : \|u\| = 1 \}$. Then, an alternative characterization of criticality is given in the following results which is useful in the following discussions.

Theorem 3. $\overline{x}$ *is a critical point to (1) if and only if there is* $\overline{V} \in \partial \varphi(\overline{x})$ *such that* $\overline{x}$ *globally solves the following problem:*

$$\min_{x \in S} \max_{u \in U} \quad u^T \psi (x) - u^T \overline{V} (x - \overline{x}). \tag{11}$$

Proof. It is obvious that problem (11) is convex. Hence, the proof can be reduced to prove that $\overline{x}$ is a critical point if and only if $\overline{x}$ is an optimum of (11). We first prove that $\overline{x} \in X$ iff $\exists \overline{V} \in \partial \varphi(\overline{x})$ such that $\overline{x}$ is a (globally) WPO to the following convex vector optimization problem:

$$\min_{x \in S} \quad \psi (x) - \overline{V} (x - \overline{x}). \tag{12}$$

According to the necessary optimal conditions that if $\overline{x} \in S$ is a WPO of (12), there exists some $u \in C^*$ with $\|u\| = 1$, such that $0 \in \partial \psi(\overline{x})^T u - \overline{V}^T u$. This implies that the above conclusion is true. Moreover, it follows from the definition of U that the cone generated by its convex hull is C^*, which means that there are a integer $q > 0$, $u_i \in U$, and $\lambda_i > 0$ with $i = 1, \dots, q$ such that $u = \sum_{i=1}^{q} \lambda_i u_i$ and $\sum_{i=1}^{q} \lambda_i = 1$. Hence, a WPO $\overline{x} \in S$ satisfies

$$0 \in \sum_{i=1}^{q} \lambda_i \left(\partial \psi (\overline{x})^T u_i - \overline{V}^T u_i \right),$$

$$\sum_{i=1}^{q} \lambda_i = 1, \ u_i \in U, \ \lambda_i > 0, \ i = 1, \dots, q. \tag{13}$$

Now the necessary optimality of (11) is analyzed. According to the formula about the subdifferential of a maximum of convex functions and the assumption about U that if $\overline{x}$ is an optimal solution to (11), there are a positive integer $\overline{q}$ and $\overline{u}_k \in U$ and $\overline{\lambda}_k > 0$ with $1 \le k \le \overline{q}$ such that

$$\overline{u}_k^T \psi (x) - \overline{u}_k^T \overline{V} (x - \overline{x})$$

$$= \max_{u \in U} u^T \psi (x) - u^T \overline{V} (x - \overline{x}), \quad k = 1, \dots, \overline{q},$$

$$0 \in \sum_{k=1}^{\overline{q}} \overline{\lambda}_k \left(\partial \psi (\overline{x})^T \overline{u}_k - \overline{V}^T \overline{u}_k \right), \quad \sum_{k=1}^{\overline{q}} \overline{\lambda}_k = 1. \tag{14}$$

The above result together with (13) leads to the assertion of this theorem. $\square$

According to the results of Theorem 3, the solution of a critical point to (1) can be equivalently transformed into solving a min–max problem. Therefore from this point, proximal point algorithms are presented for solving (1) in the following section.

3. Main Results

3.1. Algorithm 4. First, the algorithm for solving (1) is proposed which is globally convergent and well-defined.

Algorithm 4.

Step 0. A small enough constant $\varepsilon > 0$ and an initial point $x^0 \in S$ are chosen, respectively. Select a positive constant $\theta_0 > 0$. Let $k := 0$.

Step 1. At the kth iteration, calculate $V^k \in \partial \varphi(x^k)$. Solve the following problem and let x^{k+1} be an optimum:

$$\min_{x \in S} \max_{u \in U} \quad \left\{ u^T \psi (x) - u^T V^k \left(x - x^k \right) \right\}$$

$$+ \frac{\theta_k}{2} \left\| x - x^k \right\|^2. \tag{15}$$

Step 2. If $\|x^{k+1} - x^k\| \le \varepsilon$, then stop; else, update θ_k as θ_{k+1}, set $k := k + 1$, and go to Step 1.

In Step 1, the subproblem is convex and is solved at every iteration. We present the optimality conditions of this subproblem as follows:

$$x^{k+1} = P_k \left(\left(V^k \right)^T u^k + \theta_k x^k \right), \quad x^{k+1} \in S, \tag{16}$$

where $u^k \in U$ and $P_k := (\partial \psi^T u^k + \theta_k I)^{-1}$. We note that the algorithm does not rely on scalarization approaches which are usually used in solving vector optimization; that is, the algorithm does not use a priori chosen weighting parameters for the vector in different objective function.

We note that Algorithm 4 aims at obtaining one Pareto optimal solution. In recent years, the classical iteration methods (e.g., decent-direction-type and proximal point-type methods) for scalar optimization have been extended to address the vector problems to obtain one solution [16–18]. In this respect, the method proposed in this paper is similar to these ideas. However, when a need to obtain the Pareto surface (the Pareto optimal solution set) is necessary, this method may fail. It can be proved by the following example given in Antoni and Giannessi [19].

Example 5. Consider the following bilevel optimization problem (upper level):

$$\min_x \quad (x_1 - 2)^2 + \left(x_2 - \frac{5}{6} \right)^2,$$

$$\text{s.t.} \quad x \in K^0, \tag{17}$$

where K^0 is the Pareto optimal solution set of the following vector optimization problem:

$$\min_{R_+^2} \quad f(x),$$

$$\text{s.t.} \quad x \in K := \left\{ x \in R^2 : g(x) \ge 0 \right\}, \tag{18}$$

with

$$f(x) := \begin{pmatrix} f_1(x) := 2x_1 - x_2 \\ f_2(x) := -x_1 + 2x_2 \end{pmatrix},$$

$$g(x) := \begin{pmatrix} g_1(x) := 2x_1 + x_2 - 1 \\ g_2(x) := x_1 + 2x_2 - 1 \end{pmatrix}. \tag{19}$$

"$\min_{R_+^2}$" marks the Pareto optimum with respect to R_+^2. By Theorem 2.1 of [12] and Proposition 13 of [20], the Pareto optimum x^* of the lower level problem satisfies $x^* = (1/3, 1/3)$, or $x_1^* = 1 - 2x_2^*$, $0 \leq x_2^* < 1/3$, or $x_2^* = 1 - 2x_1^*$, $0 \leq x_1^* < 1/3$. Then, the optimal solution to the corresponding bilevel problem (17) is $x^* = (1, 0)$ and the optimal value is $61/36$. When we use Algorithm 4 to obtain K^0, it can be found that $K^0 = \{(0.3333, 0.3333)\}$. Therefore, the optimal solution to the corresponding bilevel problem (17) is $x^* = (0.3333, 0.3333)$ and the corresponding optimal value is 3.0279, which is larger than $61/36$. Since Algorithm 4 can not obtain the entire solution set K^0, the method fails to get the optimum of the bilevel problem.

For Algorithm 4, we have the following conclusion.

Theorem 6. *Assume that $\{x^k\}$ is generated by Algorithm 4. If one sets $\varepsilon = 0$, then the following conclusions about $\{x^k\}$ hold:*

(i) *either Algorithm 4 stops at a critical point of problem (1)*

(ii) *or there is $u_k \in U$ such that $u_k^T(f(x^{k+1}) - f(x^k)) < 0$.*

Proof. If $x^{k+1} = x^k$, then it follows that x^k is a critical point of (1) from the definition for criticality. According to the generation of x^{k+1}, the following inequality holds:

$$\max_{u \in U} \left\{ u^T \psi(x^{k+1}) - u^T V^k(x^{k+1} - x^k) \right\}$$
$$+ \frac{\theta_k}{2} \left\| x^{k+1} - x^k \right\|^2 \leq \max_{u \in U} u^T \psi(x^k) \tag{20}$$

which together with the assumption about U implies that $\exists u_k \in U$ such that

$$u_k^T \psi(x^{k+1}) - u_k^T V^k(x^{k+1} - x^k) + \frac{\theta_k}{2} \left\| x^{k+1} - x^k \right\|^2$$
$$\leq u_k^T \psi(x^k). \tag{21}$$

The subgradient that, for any $V^k \in \partial\varphi(x^k)$,

$$\varphi(x^k) + V^k(x^{k+1} - x^k) \leq \varphi(x^{k+1}). \tag{22}$$

Summing (21) and u_k^T (22) leads to

$$u_k^T \left(f(x^{k+1}) - f(x^k) \right) \leq -\frac{\theta_k}{2} \left\| x^{k+1} - x^k \right\|^2. \tag{23}$$

The above result together with $x^{k+1} \neq x^k$ and the positivity of θ_k implies that the assertion is true. □

Theorem 6 means that Algorithm 4 either generates a descent sequence satisfying the second conclusion or stops at a critical point. For proving the global convergence, the following assumptions are proposed:

(A2) The set $S \subset R^n$ is bounded and the iterations do not stop finitely.

(A3) For a large enough k, $\exists u \in U$ such that

$$x^{k+1} = \widehat{P}_k \left(\left(V^k \right)^T u + \theta_k x^k \right), \quad x^{k+1} \in S, \tag{24}$$

where $\widehat{P}_k := (\partial\psi^T u + \theta_k I)^{-1}$.

Actually, under assumption (A2), the set S is compact and convex. ϕ and ψ being C–convex imply that

$$\forall u \in C^* : \emptyset \neq \arg\min \left\{ u^T f(x) : x \in S \right\}$$
$$\subseteq \arg\min_w \left\{ f(x) : x \in S \right\}, \tag{25}$$

where $\arg\min_w$ denotes WPO set. So, to find a WPO for this problem at this case, it is sufficient to resolve $\min\{u^T f(x) : x \in S\}$, for any $u \in C^*$. Therefore, at each iteration of Algorithm 4, we can replace subproblem (15) with the following problem with fixed $u \in C^*$:

$$\min_{x \in S} \left\{ u^T \psi(x) - u^T V^k (x - x^k) \right\} + \frac{\theta_k}{2} \left\| x - x^k \right\|^2, \tag{26}$$

which means that assumption (A3) holds at this case.

Assumption (A3) means that a descent sequence for function $u^T f(x)$ on S is generated by Algorithm 4 after finite iterations. Hence, the proof of global convergence of Algorithm 4 is equivalent to showing that the sequence generated by Algorithm 4 is global convergence to a critical point of function $u^T f(x)$. For this purpose, from the conclusion of [20], it is sufficient to show that the sequence generated by Algorithm 4 satisfies three properties: descent, closedness, and boundedness. First, assumption (A3) implies that the sequence generated by Algorithm 4 is the descent directions for function $u^T f(x)$. Second, the compactness of S means the boundedness of $\{x^k\}$. Finally, we prove that the map $(\partial\varphi^T u + \theta_k I)^{-1} \circ (\sum_{i \in I} u_j \partial\psi_i + \theta_k I)$ of the algorithm is closed. The conclusion can be derived by the theorem on the composition of closed point-to-set maps proposed by Zangwill [21] and the assumption on functions φ and ψ. Then this implies that the following theorem is true.

Theorem 7. *Assume that $\{x^k\}$ is generated by Algorithm 4. If assumptions (A2) and (A3) are true, then any accumulation point of $\{x^k\}$ is a critical point of problem (1).*

3.2. Algorithm 8. We now present another algorithm for solving (1). Both the well-definedness and global convergence of Algorithm 8 are also established.

Algorithm 8.

Step 0. A sufficiently small constant $\varepsilon > 0$ and an initial point $x^0 \in S$ are presented, respectively. Choose a constant $r_0 \in C \setminus \{0\}$. Let $k := 0$.

Step 1. At kth iteration, calculate $V^k \in \partial\varphi(x^k)$. Solve the following problem and let x^{k+1} be an optimum:

$$\min_{x \in S} \max_{u \in U} \quad \left\{ u^T \psi(x) - u^T V^k (x - x^k) + \frac{r_k^T u}{2} \|x - x^k\|^2 \right\}. \quad (27)$$

Step 2. If $\|x^{k+1} - x^k\| \le \varepsilon$, then stop; else, update r_k, set $k := k + 1$, and go to Step 1.

At every iteration of Algorithm 8, similar to Algorithm 4, a convex subproblem is presented and the enclosed form expression for the optimum can be computed as

$$x^{k+1} = P_k \left(\left(V^k \right)^T u^k + \left(r_k^T u^k \right) x^k \right), \quad x^{k+1} \in S, \quad (28)$$

$$\max_{u \in U} \quad \left\{ u^T \psi \left(x^{k+1} \right) - u^T V^k \left(x^{k+1} - x^k \right) + \frac{r_k^T u}{2} \left\| x^{k+1} - x^k \right\|^2 \right\} \le \max_{u \in U} u^T \psi \left(x^k \right). \quad (29)$$

The above inequality means that $\exists u_k \in U$ such that

$$u_k^T \psi \left(x^{k+1} \right) - u_k^T V^k \left(x^{k+1} - x^k \right) + \frac{r_k^T u}{2} \left\| x^{k+1} - x^k \right\|^2$$
$$\le u_k^T \psi \left(x^k \right). \quad (30)$$

The subgradient definition leads to that, for any $V^k \in \partial\varphi(x^k)$,

$$\varphi \left(x^k \right) + \left(x^{k+1} - x^k \right)^T V^k \le \varphi \left(x^{k+1} \right). \quad (31)$$

The following inequality comes from summing (30) and u^T (31):

$$u_k^T \left(f \left(x^{k+1} \right) - f \left(x^k \right) \right) \le -\frac{r_k^T u}{2} \left\| x^{k+1} - x^k \right\|^2. \quad (32)$$

This together with $r_k^T u_k > 0$ and $x^{k+1} \ne x^k$ implies that the assertion is true. □

Theorem 9 implies that the sequence generated by Algorithm 8 either satisfies the second conclusion or stops at a critical point. The following assumption is proposed for proving the global convergence:

(A4) For large enough k, $\exists u \in U$ such that

$$x^{k+1} = \widetilde{P}_k \left(\left(V^k \right)^T u + r_k^T u x^k \right), \quad x^{k+1} \in S, \quad (33)$$

where $\widetilde{P}_k := ((\partial\varphi^T u + r_k^T u I))^{-1}$.

Similar to the proof of the globally convergence analysis for Algorithm 4, Algorithm 8 is also globally convergent.

Theorem 10. *Assume that $\{x^k\}$ is generated by Algorithm 8. If one supposes that assumptions (A2) and (A4) hold, then any accumulation point of $\{x^k\}$ is a critical point of problem (1).*

where $u^k \in U$ and $P_k := (\partial\varphi^T u^k + (r_k^T u^k)I)^{-1}$. The following two conclusions hold for Algorithm 8.

Theorem 9. *If one sets $\varepsilon = 0$ and choose r_k with $r_k^T u > 0$, $\forall u \in U$, then the sequence $\{x^k\}$ generated by Algorithm 8 satisfies*

(i) *either the algorithm stops at a critical point of problem (1)*

(ii) *or there is $u_k \in U$ such that $u_k^T(f(x^{k+1}) - f(x^k)) < 0$.*

Proof. If $x^{k+1} = x^k$, from Theorem 2, it follows that x^k is a critical point of (1). According to the generation of x^{k+1}, the following inequality holds:

3.3. ϵ-Proximal Algorithm.

We propose an inexact version for the above algorithms, that is, an ϵ-proximal algorithm for obtaining an ϵ-critical point of problem (1) by utilizing the ϵ-subdifferential. For this purpose we define the ϵ-critical point of (1) as follows: find a point $\overline{x}$ such that

$$\overline{x} \in X_\epsilon := \left\{ x \in R^n : \exists u \in U, \text{ s.t., } 0 \in \partial_\epsilon \psi(x)^T u \right.$$
$$\left. - \partial_\epsilon \varphi(x)^T u \right\}, \quad (34)$$

where ∂_ϵ is the ϵ-subdifferential defined above. This algorithm is in connection with inexact proximal point algorithms proposed in [22].

ϵ-Proximal Algorithm

Step 0. A small enough constant $\varepsilon > 0$ and an initial point $x^0 \in S$ are chosen, respectively. Let $\epsilon_0, \theta_0 > 0$ be constants. Set $k := 0$.

Step 1. At the kth iteration, calculate $V_{\epsilon_k}^k \in \partial_{\epsilon_k}\varphi(x^k)$. $x^{k+1} \in S$ satisfies the following equation:

$$x^{k+1} = P_{\epsilon_k}^k \left(\left(V_{\epsilon_k}^k \right)^T u^k + \theta_k x^k \right), \quad (35)$$

where $u^k \in U$ and $P_{\epsilon_k}^k := (\partial_{\epsilon_k} \psi^T u^k + \theta_k I)^{-1}$.

Step 2. Stop if $\|x^{k+1} - x^k\| \le \varepsilon$; else, update θ_k, set $k := k + 1$, and go to Step 1.

This is an approximate algorithm to Algorithm 4 by using the ϵ-subdifferential instead of the subdifferential. We can similarly propose an approximate algorithm for Algorithm 8 by replacing the subdifferential with the ϵ-subdifferential and the corresponding properties can also be found. The next two theorems present the descent and globally convergent properties of the approximate algorithm, respectively.

Theorem 11. *Assume that $\{x^k\}$ is generated by ϵ-proximal algorithm, $\varepsilon = 0$ and $\epsilon_k \le \epsilon$, $\forall k$. If the algorithm do not stop*

finitely, then $\{f(x^k)\}$ is a 2ϵ-descent sequence, which implies that $\exists u^k \in U$ such that

$$\left(f\left(x^{k+1}\right) - f\left(x^k\right)\right)^T u^k < 2\epsilon, \quad \forall k. \tag{36}$$

Proof. The conclusion comes from Theorem 5.4 in [23] and Theorem 6. □

To show the global convergent, the following assumptions are presented:

(A5) For large enough k, $\exists u \in U$ such that

$$x^{k+1} = \widehat{P}_{\epsilon_k}^k \left(\left(V_{\epsilon_k}^k\right)^T u + \theta_k x^k\right), \quad x^{k+1} \in S, \tag{37}$$

where $\widehat{P}_{\epsilon_k}^k := (\partial_{\epsilon_k} \varphi^T u + \theta_k I)^{-1}$;

(A6) For any $u \in U$, $u^T f(x)$ is bounded below for any $x \in S$; there is $0 < \underline{\theta} \leq \overline{\theta} < +\infty$ such that $\theta_k \in [\underline{\theta}, \overline{\theta}]$, $\forall k$; the parameters ϵ_k, $\forall k$, satisfy $\sum_{k=0}^{+\infty} \epsilon_k < +\infty$.

Hence, similar to Theorem 2.1 in [24], we can prove the approximate algorithm has the following convergent results.

Theorem 12. *Suppose that $\{x^k\}$ is generated by the ϵ_k-proximal algorithm. If assumptions (A2), (A5), and (A6) hold, then any accumulation point of $\{x^k\}$ is a critical point of problem (1) and $\{x^k\}$ is asymptotically regular in the following sense:* $\lim_{k \to +\infty} \theta_k \|x^k - x^{k+1}\| = 0$.

4. Numerical Tests

4.1. Probabilistic Lot Sizing with Service Levels. In probabilistic lot sizing with service levels, the order quantities, x_t, $t = 1, \ldots, n$, are determined at the beginning of the planning horizon by considering to simultaneously minimize the expected total cost and maximize the guarantee of both the inventory balance and service levels. The model can be described as follows [25]:

$$\min \quad E\left(\sum_{j=1}^{n} \left(\mu_j x_j + \nu_j s_j \left(\xi_j\right) + \gamma_j w_j\right)\right)$$

$$\max \quad \mathbb{P}\begin{pmatrix} x_1 \geq \xi_1 \\ x_1 + x_2 \geq \xi_2 \\ \vdots \\ x_1 + x_2 + x_3 + \cdots + x_n \geq \xi_n \end{pmatrix}$$

$$\text{s.t.} \quad 0 \leq x_t \leq M_t w_t, \quad t \in [1, n], \tag{38}$$

$$s_t \left(\xi_t\right) \geq \sum_{j=1}^{t} x_j - \xi_t, \quad t \in [1, n-1],$$

$$s_n \left(\xi_n\right) = \sum_{t=1}^{n} x_t - \xi_n,$$

$$s_t \left(\xi_t\right) \geq 0,$$

$$w_t^2 + \left(1 - w_t\right)^2 = 1,$$

$$t \in [1, n],$$

where μ_j, ν_j, and γ_j are parameters, x_j, s_j, and w_j are decision variables, and ξ_t, $t = 1, \ldots, n$ are stochastic variables.

Define $g_t(x, \xi) := \xi_t - (x_1 + \cdots + x_t)$, $t = 1, \ldots, n$. A fundamental difficulty in solving (38) is that the evaluation of function $\mathbb{P}[g(x, \xi) \leq 0]$ is nontrival and it is difficult to have its closed form, where $g := (g_1, \ldots, g_n)$. Therefore, many approximate methods have been proposed. CVaR approximation presented by Rockafellar and Uryasev is one of those methods, which is the best convex conservative approximation [26]. However, Hong et al. point out that the CVaR approximation is not a good approximation [27]. Therefore, Hong et al. present a DC approach, which can approximate $p(x)$ better than CVaR approximation [27]. Define $p(x) := 1 - \mathbb{P}[g(x, \xi) \leq 0]$. The DC approximation of $p(x)$ can be stated as follows:

$$\widetilde{p}(x, t) := \frac{\pi_1(x, t) - \pi_2(x)}{t}, \tag{39}$$

where $\pi_1(x, t) := E[t + \widetilde{g}(x, \xi)]^+$, $\pi_2(x) := \pi_1(x, 0)$, and $\widetilde{g}(x, \xi) := \max\{g_1(x, \xi), \ldots, g_n(x, \xi)\}$. Note that $[a]^+ := \max\{a, 0\}$.

We use DC approximation (39) to replace the probability function in (38). We assume that the stochastic variables ξ are defined on the sample space Ω. For an integrable function $\psi : \Omega \to R$, the Monte Carlo sampling estimate for $E[\cdot]$ is obtained by taking independently and identically distributed random samples $[\xi_1, \ldots, \xi_L]$ from Ω and letting $E[\psi(\omega)] \approx (1/L) \sum_{l=1}^{L} \rho(\omega_l) \psi(\omega_l)$, where $\rho(\omega_l)$ is the probability. Define

$$c(x, \xi_l) := \max_{1 \leq t \leq n} \frac{c_j(x, \xi_{tl})}{L}, \quad l = 1, \ldots, L, \tag{40}$$

where $c_t(x, \xi_{tl}) := \xi_{tl} - (x_1 + x_2 + \cdots + x_t)$, $t = 1, \ldots, n$, $l = 1, \ldots, L$. In this paper, we use

$$f_2(x) := \frac{\sum_{l=1}^{L} \left([0.05 + c(x, \xi_l)]^+ - [c(x, \xi_l)]^+\right)}{0.05} \tag{41}$$

as the approximation function to (39), where $\xi_l \in [1, 2]$, $l = 1, \ldots, L$ are stochastically generated, and $[\cdot]^+$ denotes that $[x]^+ = x$, if $x > 0$; otherwise $[x]^+ = 0$, $(x \leq 0)$. Define the feasible set

$$S := \begin{pmatrix} 0 \leq x_t \leq M_t w_t, \ t \in [1, n], \\ s_t \geq \sum_{j=1}^{t} x_j - \frac{\sum_{l=1}^{L} \xi_{tl}}{L}, \ t \in [1, n-1], \\ s_n = \sum_{j=1}^{n} x_j - \frac{\sum_{l=1}^{L} \xi_{nl}}{L}, \\ s_t \geq 0, \ w_t^2 + \left(1 - w_t\right)^2 = 1, \ t \in [1, n] \end{pmatrix}. \tag{42}$$

With the above notations, we can describe Algorithm 4 with $U := R_+^m$ for solving Problem (38) as follows.

Algorithm 13.

Step 0. Give $w^0 := (w_1^0, w_2^0, \ldots, w_n^0) \in \{0, 1\}^n$; $x_j^0 \in [0, M_t w_t]$, $j = 1, \ldots, n$; $s_t^0 := \sum_{j=1}^{t} x_j^0 - \sum_{k=1}^{K} \xi_{tk}/K$, $t = 1, \ldots, n$. Let $k := 0$.

TABLE 1: Numerical results.

n	K	Algorithm 4		Algorithm 8		ϵ-Algorithm 4		ϵ-Algorithm 8	
		Iter	CPU	Iter	CPU	Iter	CPU	Iter	CPU
10	500	15	0.02	17	0.02	17	0.03	27	0.04
10	1000	16	0.03	17	0.02	18	0.03	27	0.03
20	500	21	0.09	24	0.09	26	0.1	32	0.15
20	1000	21	0.12	34	0.22	28	0.25	34	0.28
50	500	37	0.47	48	0.75	50	0.55	50	0.56
50	1000	37	0.47	49	0.77	59	0.62	59	0.67
100	500	59	0.99	78	0.81	75	1.33	78	1.47
100	1000	87	1.55	98	1.89	124	3.96	170	4.39
150	500	341	57.2	500	207.13	500	292.12	500	287.58
150	1000	500	321.22	500	210.35	500	299.27	500	294.77

Step 1. Compute the following optimization problem:

$$
\begin{aligned}
\min \quad & \eta \\
\text{s.t} \quad & f_1(x,s,w) + \frac{1}{2}\left\|(x,s,w) - \left(x^k, s^k, w^k\right)\right\|^2 \leq \eta \\
& \frac{\sum_{l=1}^{L}\left(\left[0.05 + c\left(x,\xi_l\right)\right]^+ - \left[\nabla_x c_{j*}\left(x^p, \xi_l\right) I_{(0,+\infty)}\left(c\left(x^p, \xi_l\right)\right)\right]^T \left(x - x^k\right)\right)}{0.05} + \frac{\theta_k}{2}\left\|x - x^k\right\|^2 \leq \eta \\
& (x,s,w) \in S,
\end{aligned}
\tag{43}
$$

where $I_{(0,+\infty)}(\cdot)$ is the indicator function defined as $I_{(0,+\infty)}(y) = 1$, if $y \in (0,+\infty)$; otherwise, 0; $j*$ satisfies $c(x,\xi_l) := \max_{1 \leq j \leq n}(c_j(x,\xi_{jl})/L)$, $l = 1,\ldots,L$. $(x^{k+1}, s^{k+1}, w^{k+1})$ is the optimum of the above subproblem.

Step 2. If $\|(x^{k+1}, s^{k+1}, w^{k+1}) - (x^k, s^k, w^k)\| \leq 10^{-6}$, then stop; otherwise, $k := k + 1$ and go to Step 1.

Similarly, we can present Algorithm 8 and ϵ–approximate algorithm for solving problem (38).

4.2. Numerical Results. To show the efficiency of the proposed method, this section presents several numerical tests; that is, use Algorithms 4 and 8 and ϵ-proximal algorithm for solving problem (38). We note that the codes are written in Matlab 7.10 with built-in solver "fmincon" to solve the convex subproblems. A DELL computer is used to conduct the tests where the computer is with 4.00 GB of memory and Intel(R)Core(TM)i5-2400 processor (3.10 GHz).

The parameters used in the tests are set as follows: θ_{jk} and θ_k are stochastically generated in the set $[1,2]$; the initial point is stochastically chosen from the feasible region of (38); the terminating rule $\varepsilon := 1 \times 10^{-6}$; for $k \geq 1$, ϵ_k is set as $1/k^2$; ϵ_0 is set as zero; we assume that $M_t \in [10, 20]$, and $\mu_t, \nu_t, \gamma_t, \xi_t \in [1,2]$, $t = 1,\ldots,n$, are stochastically generated. The numerical reports are presented in Table 1, where computing time in seconds ("CPU") and the number of iterations ("Iter") are given. We note that 500 is set as the

maximal number of iterations. From Table 1, we conclude that Algorithm 8 under performs Algorithm 4 with respect to CPU time and the number of iterations. The ϵ-approximation of Algorithm 8 also underperforms the corresponding ϵ-approximation of Algorithm 4 with respect to CPU time and the number of iterations. These results mean that the theoretical results are true and the proposed algorithms are effective in solving (38).

5. Conclusion

This paper proposes both exact and inexact proximal point algorithms for solving vector DC optimization problems. The proposed algorithms enjoy both well-posedness and global convergence under suitable assumptions. An application to a probabilistic lot sizing with service levels is considered. We first show that this model can be equivalently cast as a multiobjective DC optimization problem. Then, the proposed algorithms are utilized to solve the resulting problem. The numerical results show that the methods are efficient. For future work, we can discuss the local convergence of the proposed algorithms. The construction of methods for solving the vector DC optimization problems with nonsmooth objectives is also an interesting topic for future work.

Conflicts of Interest

The authors declare that there are no conflicts of interest.
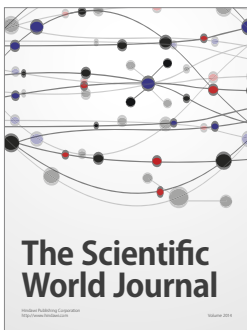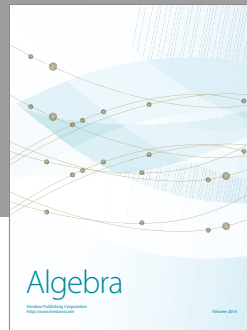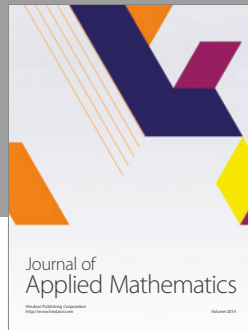
Authors' Contributions

Acknowledgments

References

[1] J. Jahn, *Vector Optimization: Theory, Applications, And Extensions*, Spring, Erlangen, Germany, 2003.

[2] Y. Xing and S. Ding, "Higher integrability of Green's operator and homotopy operator," *Journal of Mathematical Analysis and Applications*, vol. 446, no. 1, pp. 648–662, 2017.

[3] W. Song and G. M. Yao, "Homotopy method for a general multiobjective programming problem," *Journal of Optimization Theory and Applications*, vol. 138, no. 1, pp. 139–153, 2008.

[4] Y. Xing and S. Ding, "Global estimates for the maximal operator and homotopy operator," *Mathematical Inequalities & Applications*, vol. 15, no. 1, pp. 27–35, 2012.

[5] S. Qu, M. Goh, S.-Y. Wu, and R. De Souza, "Multiobjective DC programs with infinite convex constraints," *Journal of Global Optimization*, vol. 59, no. 1, pp. 41–58, 2014.

[6] N. Gadhi, M. Laghdir, and A. Metrane, "Optimality conditions for D.C. vector optimization problems under reverse convex constraints," *Journal of Global Optimization*, vol. 33, no. 4, pp. 527–540, 2005.

[7] Y.-J. Liu, D. Sun, and K.-C. Toh, "An implementable proximal point algorithmic framework for nuclear norm minimization," *Mathematical Programming*, vol. 133, no. 1-2, Ser. A, pp. 399–436, 2012.

[8] J.-J. Moreau, "Proximate' et dualite' dans un espace Hilbertien," *Bulletin de la Société Mathématique de France*, vol. 93, pp. 273–299, 1965.

[9] R. T. Rockafellar, "Monotone operators and the proximal point algorithm," *SIAM Journal on Control and Optimization*, vol. 14, no. 5, pp. 877–898, 1976.

[10] R. E. Bruck and S. Reich, "Nonexpansive projections and resolvents of accretive operators in Banach spaces," *Houston Journal of Mathematics*, vol. 3, no. 4, pp. 459–470, 1977.

[11] K. D. Villacorta and P. R. Oliveira, "An interior proximal method in vector optimization," *European Journal of Operational Research*, vol. 214, no. 3, pp. 485–492, 2011.

[12] S.-J. Qu, M. Goh, R. De Souza, and T.-N. Wang, "Proximal point algorithms for convex multi-criteria optimization with applications to supply chain risk management," *Journal of Optimization Theory and Applications*, vol. 163, no. 3, pp. 949–956, 2014.

[13] R. S. Burachik and V. Jeyakumar, "A dual condition for the convex subdifferential sum formula with applications," *Journal of Convex Analysis*, vol. 12, no. 2, pp. 279–290, 2005.

[14] D. T. Luc, N. X. Tan, and P. N. Tinh, "Convex vector functions and their subdifferential," *Acta Mathematica Vietnamica*, vol. 23, no. 1, pp. 107–127, 1998.

[15] R. T. Rockafellar, *Convex Analysis*, Princeton Mathematical Series, No. 28, Princeton University Press, Princeton, NJ, USA, 1970.

[16] S. Qu, M. Goh, and F. T. Chan, "Quasi-Newton methods for solving multiobjective optimization," *Operations Research Letters*, vol. 39, no. 5, pp. 397–399, 2011.

[17] S. Qu, M. Goh, and B. Liang, "Trust region methods for solving multiobjective optimisation," *Optimization Methods & Software*, vol. 28, no. 4, pp. 796–811, 2013.

[18] H. Bonnel, A. N. Iusem, and B. F. Svaiter, "Proximal methods in vector optimization," *SIAM Journal on Optimization*, vol. 15, no. 4, pp. 953–970, 2005.

[19] C. Antoni and F. Giannessi, "Some remarks on bi-level vector extremum problems," in *Constructive nonsmooth analysis and related topics*, vol. 87 of *Springer Optim. Appl.*, pp. 137–157, Springer, New York, 2014.

[20] F. Giannessi, G. Mastroeni, and L. Pellegrini, "On the theory of vector optimization and variational inequalities. Image space analysis and separation," in *Vector variational inequalities and vector equilibria*, vol. 38 of *Nonconvex Optim. Appl.*, pp. 153–215, Kluwer Acad. Publ., Dordrecht, 2000.

[21] W. I. Zangwill, *Nonlinear Programming: A Unified Approach*, Prentice-Hall, New York, NY, USA, 1969.

[22] O. Nevanlinna and S. Reich, "Strong convergence of contraction semigroups and of iterative methods for accretive operators in Banach spaces," *Israel Journal of Mathematics*, vol. 32, no. 1, pp. 44–58, 1979.

[23] W.-y. Sun, R. J. Sampaio, and M. A. Candido, "Proximal point algorithm for minimization of DC function," *Journal of Computational Mathematics*, vol. 21, no. 4, pp. 451–462, 2003.

[24] A. Moudafi and P.-E. Mainge', "On the convergence of an approximate proximal method for DC functions," *Journal of Computational Mathematics*, vol. 24, no. 4, pp. 475–480, 2006.

[25] S. Goel, *Dynamic probabilistic lot-sizing with service level constraints, Thesis*, The Ohio State University, 2011.

[26] R. T. Rockafellar and S. Uryasev, "Optimization of conditional value-at-risk," *The Journal of Risk*, vol. 2, no. 3, pp. 21–41, 2000.

[27] L. J. Hong, Y. Yang, and L. Zhang, "Sequential convex approximations to joint chance constrained programs: a Monte Carlo approach," *Operations Research*, vol. 59, no. 3, pp. 617–630, 2011.