A CONJUGATE DIRECTIONS-TYPE PROCEDURE FOR QUADRATIC MULTIOBJECTIVE OPTIMIZATION

ELLEN H. FUKUDA, L. M. GRAÑA DRUMMOND, AND ARIANE M. MASUDA

Abstract. We propose an extension of the real-valued conjugate directions method for unconstrained quadratic multiobjective problems. As in the single-valued counterpart, the procedure requires a set of directions that are simultaneously conjugate with respect to the positive definite matrices of all quadratic objective components. Likewise, the multicriteria version computes the steplength by means of the unconstrained minimization of a single-variable strongly convex function at each iteration. When it is implemented with a weakly-increasing (strongly-increasing) auxiliary function, the scheme produces weak Pareto (Pareto) optima in finitely many iterations.

Keywords: Multiobjective optimization; weak Pareto optimality; Pareto optimality; conjugate directions method

1. Introduction

Solving multicriteria problems is not always a straightforward task. Some very popular procedures have severe drawbacks. Widely used schemes such as the weighting method may lead to unbounded problems. This procedure consists in minimizing a convex combination of the objectives, and the difficulty is that we do not know a priori the suitable coefficients. In order to overcome these flaws, extensions to the multiobjective setting of classical methods for scalar optimization have been proposed in the last two decades (see [2, 3, 4, 6, 5, 7, 13], for instance).

We propose a conjugate directions-type procedure for unconstrained multiobjective quadratic problems. As far as we know, this is the first attempt to adapt the scalar method to the multicriteria setting.

The proposed extension is suitable for quadratic multiobjectives that satisfy a strong condition, namely, the existence of a Hamel basis *simultaneously* conjugate with respect to the symmetric positive definite matrices that determine the quadratic part of the components. This condition extends the essential requirement of the classical conjugate directions method, i.e., the existence of an orthogonal basis for $\mathbb{R}^n$ with respect to the inner product induced by the quadratic objective function's matrix. In the scalar case one can use the Gram-Schmidt procedure to obtain such a basis. For the multiobjective case we exhibit a large class of quadratic problems whose Hamel basis can be explicitly constructed.

Our procedure requires solving finitely many single-variable unconstrained strongly convex minimization real-valued problems. In each subproblem the unique solution is the steplength of the next iterate. The objective functions are essentially the composition of the quadratic multiobjective's components with an auxiliary function taken from quite large families.

Date: July 28, 2020.

As expected, the convergence results of the extension are not better than the ones for the scalar counterpart. Regardless of how poor the initial guess is, the scheme implemented with a weakly (strongly) increasing auxiliary function produces a weak Pareto (Pareto) optimal solution in finitely many iterations. Recall that for general nonlinear scalar-valued optimization problems, the related approach (called *conjugate gradient method*) does not guarantee convergence with a finite number of iterations. This also happens with its multiobjective version, proposed in [11]. Our paper actually clarifies what happens when more general directions are considered in multiobjective quadratic problems.

The outline of this work is as follows. In Section 2 we introduce some notations and exhibit a large class of problems suitable for the application of the method. More specifically, for any finite set of symmetric positive definite matrices, we give a sufficient condition for the existence of a Hamel basis for $\mathbb{R}^n$ whose elements are simultaneously conjugate with respect to all these matrices. In Section 3 we formally define the quadratic problem for an m-multifunction of n variables and the auxiliary functions to be used in the procedure. We also recall a well-known result relating monotonicity and optimality that is essential for the convergence analysis. In Section 4 we present and analyze the main algorithm, which starts from an initial point x^0 and produces a sequence $\{x^1, \ldots, x^n\}$ in $\mathbb{R}^n$. In particular, in the real-valued case we retrieve the classical conjugate directions method. In Section 5 we define the procedure which consists in running the algorithm $n+1$ times with $n+1$ suitable initial points. In Section 6 we analyze the convergence of the procedure. We apply the multiobjective conjugate directions-type method for a simple instance of the problem in Section 7. In this example we show that all Pareto optima can be found by varying a parameter of the auxiliary function. We end with some final remarks in Section 8.

2. Preliminaries

We let $x \in \mathbb{R}^n$ stand for the column vector with coordinates $x_1, \ldots, x_n$ and its transpose given by $x^\top = (x_1, \ldots, x_n) \in \mathbb{R}^{1\times n}$. For the sake of simplicity, we also write the vector x as $[x_i]_{i=1}^n$. We denote sequences in $\mathbb{R}$ and $\mathbb{R}^n$ (for $n \geq 2$) by $\{x_k\}$ and $\{x^k\}$, respectively. For $i = 1, \ldots, n$, the canonical vector $e^i \in \mathbb{R}^n$ has 1 in the i^{th} position and 0 elsewhere, i.e., $(e^i)_j = \delta_{ij}$, where δ_{ij} is the Kronecker δ defined by $\delta_{ij} = 1$ if $i = j$, and 0 otherwise. The transpose of an $n \times n$ matrix $A = (a_{ij})$ is $A^\top = (a_{ji})$. The canonical matrix $e^{ij} \in \mathbb{R}^{n\times n}$ has 1 in the position ij and 0 elsewhere. A diagonal matrix $D \in \mathbb{R}^{n\times n}$ is a matrix such that $d_{ij} = 0$ for all $i \neq j$. It is represented by $\operatorname{diag}(d_{11}, \ldots, d_{nn})$; in particular, the identity matrix I can be written as $\operatorname{diag}(1, \ldots, 1)$. A symmetric positive definite matrix Q induces an inner product and a norm over $\mathbb{R}^n$ given by $\langle x, y\rangle_Q = x^\top Q y$ and a $\|x\|_Q = \sqrt{\langle x, x\rangle_Q}$, respectively. When $Q = I$, we obtain the canonical inner product and the Euclidean norm. In this case we simply write $\langle \cdot, \cdot\rangle$ and $\|\cdot\|$.

Given $M \subset \mathbb{R}^n$, $\operatorname{span} M$ represents the subspace generated by the elements in M. For a subspace $S \subset \mathbb{R}^n$, $S^{\perp_Q}$ stands for its orthogonal complement with respect to $\langle \cdot, \cdot\rangle_Q$, i.e., $S^{\perp_Q} = \{x \in \mathbb{R}^n \colon \langle x, y\rangle_Q = 0 \text{ for all } y \in S\}$. The Paretian cone is $\mathbb{R}^n_+ := \{x \in \mathbb{R}^n \colon x_i \geq 0 \text{ for } i = 1, \ldots, n\}$. For a function $h \colon \mathbb{R}^n \to \mathbb{R}$ and $W \subset \mathbb{R}^n$, the set of minimizers of h restricted to W is denoted by $\operatorname{argmin}_{x\in W} h(x)$. If this set has a single element, say x^*, we write $x^* = \operatorname{argmin}_{x\in W} h(x)$.

In unconstrained scalar optimization, the so-called conjugate directions method seeks to achieve a faster convergence rate than the steepest descent's one with a smaller computational cost than Newton's method's one [1, 8]. The conjugate directions method was originally developed for convex quadratic objective functions $f \colon \mathbb{R}^n \to \mathbb{R}$. When f is also strongly convex, it produces the unconstrained minimizer in at most n iterations. For a large n, one obtains a good approximation of the optimum with fewer iterations. The conjugate directions method can be used to solve linear systems of equations $Qx = b$ for a nonsingular matrix Q. We may assume Q to be symmetric positive definite; otherwise, we consider the equivalent system $Q^\top Qx = Q^\top b$, which can be solved by applying a conjugate directions method to $\min_{x\in\mathbb{R}^n} \|Q^\top Qx - Q^\top b\|^2/2$ in finitely many steps. In nonquadratic optimization problems, although the conjugate directions method does not have finite termination in general, it has nice convergence properties.

Our goal is to extend the conjugate directions method to the multiobjective setting. For a quadratic vector-valued function f with strongly convex components f_i, we propose a conjugate direction-type scheme by using a Hamel basis of conjugate directions with respect to the matrices associated to all f_i's. Before establishing a sufficient condition for the existence of this basis, we recall the notion of conjugacy.

Definition 2.1. *Let $Q \in \mathbb{R}^{n\times n}$ be symmetric positive definite. A set $\{u^1, \dots, u^r\} \subset \mathbb{R}^n \setminus \{0\}$ is said to be Q-*conjugate*, if its elements are pairwise orthogonal with respect to $\langle\cdot,\cdot\rangle_Q$.*

The following result concerns sets of conjugate directions.

Lemma 2.2. *Let $Q \in \mathbb{R}^{n\times n}$ be symmetric positive definite. If $\{u^1, \dots, u^r\}$ is Q-conjugate, then it is linearly independent. In particular, it is a Hamel basis for $\mathbb{R}^n$ when $r = n$.*

Proof. See [8, Lemma 3.4.1]. □

In order to establish an existence condition for the Hamel basis, we point out the following fact.

Remark 2.3. Consider a symmetric positive definite matrix Q in $\mathbb{R}^{n\times n}$. Then there exist an $n \times n$ orthogonal matrix P and $D = \operatorname{diag}(\lambda_1, \dots, \lambda_n)$ with $\lambda_j > 0$ for all j such that $Q = PDP^\top$; see [12, Theorem 1.13]. Since P is orthogonal and the canonical basis $\{e^1, \dots, e^n\}$ is D-conjugate, it follows that

$$(Pe^i)^\top QPe^j = (e^i)^\top De^j = \lambda_i \delta_{ij}.$$

So $\{Pe^1, \dots, Pe^n\}$ is a Q-conjugate basis for $\mathbb{R}^n$. Moreover, every Pe^j is an eigenvector of Q associated to the eigenvalue λ_j.

Given symmetric positive definite matrices $Q_1, \dots, Q_m$ in $\mathbb{R}^{n\times n}$, we show how Remark 2.3 can lead us to obtain a Q_i-conjugate Hamel basis for $i = 1, \dots, m$.

First, recall that $\sigma\colon J_n \to J_n$ is a *permutation* on $J_n := \{1, \dots, n\}$, if σ is injective. By setting $\sigma_i := \sigma(i)$, we can represent σ by $(\sigma_1, \dots, \sigma_n)$. We denote by I_σ an $n \times n$ matrix whose j^{th} column coincides with the σ_j^{th} column of I, for all $j = 1, \dots, n$. Since I and I_σ have the same columns, the matrix I_σ is nonsingular.

Lemma 2.4. *Let σ be a permutation on J_n and $Q, P, D \in \mathbb{R}^{n\times n}$ be such that Q is symmetric positive definite, P is orthogonal, $D = \operatorname{diag}(\lambda_1, \dots, \lambda_n)$ and $Q = PDP^\top$. The following statements hold true:*

- (i) *the j^{th} column of PI_σ coincides with the σ_j^{th} column of P for all j; in particular, P and PI_σ have the same columns;*
- (ii) *$I_\sigma = e^{\sigma_1 1} + e^{\sigma_2 2} + \dots + e^{\sigma_n n}$;*
- (iii) *$De^{k\ell} = \lambda_k e^{k\ell}$ for any $e^{k\ell} \in \mathbb{R}^{n\times n}$;*
- (iv) *$e^{ij}De^{k\ell} = \delta_{kj}\lambda_k e^{i\ell}$ for any $e^{k\ell} \in \mathbb{R}^{n\times n}$; in particular, $e^{ij}e^{k\ell} = \delta_{jk}e^{i\ell}$;*

(v) $I_\sigma^\top D I_\sigma = \mathrm{diag}(\lambda_{\sigma_1}, \ldots, \lambda_{\sigma_n})$*; in particular,* I_σ *is orthogonal and so is* PI_σ*;*
(vi) $I_\sigma D I_\sigma^\top = \mathrm{diag}(\lambda_{\zeta_1}, \ldots, \lambda_{\zeta_n})$ *for a certain permutation* ζ *on* J_n*.*

Proof. (i) By definition, the j^{th} column of I_σ coincides with the canonical (column) vector e^{σ_j} for all j.
(ii) The result follows immediately, since by (i) the j^{th} column of I_σ coincides with e^{σ_j} for all j.
(iii) Immediate.
(iv) The first part is immediate. By taking $D = I$, we obtain the second one.
(v) By using (ii) and (iv), we see that

$$\begin{aligned} I_\sigma^\top D I_\sigma &= (e^{1\sigma_1} + e^{2\sigma_2} + \cdots + e^{n\sigma_n}) D (e^{\sigma_1 1} + e^{\sigma_2 2} + \cdots + e^{\sigma_n n}) \\ &= \lambda_{\sigma_1} e^{11} + \lambda_{\sigma_2} e^{22} + \cdots + \lambda_{\sigma_n} e^{nn} \\ &= \mathrm{diag}(\lambda_{\sigma_1}, \ldots, \lambda_{\sigma_n}). \end{aligned}$$

(vi) By using (ii), (iv) and the injectivity of σ, we have

$$\begin{aligned} I_\sigma D I_\sigma^\top &= (e^{\sigma_1 1} + e^{\sigma_2 2} + \cdots + e^{\sigma_n n}) D (e^{1\sigma_1} + e^{2\sigma_2} + \cdots + e^{n\sigma_n}) \\ &= \lambda_1 e^{\sigma_1\sigma_1} + \lambda_2 e^{\sigma_2\sigma_2} + \cdots + \lambda_n e^{\sigma_n\sigma_n} \\ &= \mathrm{diag}(\lambda_{\zeta_1}, \ldots, \lambda_{\zeta_n}), \end{aligned}$$

for an adequate permutation ζ.

□

Definition 2.5. *Let* σ *be a permutation on* J_n*, and let* $Q_1, Q_2, P_1, D_1, D_2 \in \mathbb{R}^{n\times n}$ *be such that* Q_1, Q_2 *are symmetric positive definite,* P_1 *is orthogonal and* D_1, D_2 *are diagonal. We say that* Q_1 *and* Q_2 *are* σ-related *if* $Q_1 = P_1 D_1 P_1^\top$ *and* $Q_2 = (P_1 I_\sigma) D_2 (P_1 I_\sigma)^\top$*.*

Lemma 2.6. *Let* $Q, \tilde{Q}$ *be symmetric positive definite matrices in* $\mathbb{R}^{n\times n}$ *such that* $Q = PDP^\top$ *and* $\tilde{Q} = \tilde{P}\tilde{D}\tilde{P}^\top$*, where* P*,* $\tilde{P}$ *are orthogonal and* D*,* $\tilde{D}$ *are diagonal.*
(i) *The matrices* Q *and* $\tilde{Q}$ *are* σ*-related for some permutation* σ *on* J_n *if and only if* Q *and* $\tilde{Q}$ *have the same eigenvectors. In this case the set of eigenvectors is* Q*- and* $\tilde{Q}$*-conjugate.*
(ii) *Let* Q *and* $\tilde{Q}$ *be* σ*-related. Then* $Q\tilde{Q} = \tilde{Q}Q$*.*

Proof. (i) By Remark 2.3, the sets $\{Pe^1, \ldots, Pe^n\}$ and $\{\tilde{P}e^1, \ldots, \tilde{P}e^n\}$ consist of eigenvectors of Q and $\tilde{Q}$, respectively. By hypothesis, $\tilde{P} = PI_\sigma$, so $\{\tilde{P}e^1, \ldots, \tilde{P}e^n\} = \{Pe^1, \ldots, Pe^n\}$ by Lemma 2.4(i). Conversely, by Remark 2.3 it follows that the set of eigenvectors of Q is $\{Pe^1, \ldots, Pe^n\}$, which coincides with $\{\tilde{P}e^1, \ldots, \tilde{P}e^n\}$. Therefore P and $\tilde{P}$ have the same columns. By Lemma 2.4(i), there exists a permutation σ on J_n such that $\tilde{P} = PI_\sigma$. By using Remark 2.3 once again, this common set of eigenvectors is Q- and $\tilde{Q}$-conjugate.
(ii) Assume $\tilde{P} = PI_\sigma$. Then

$$\begin{aligned} Q\tilde{Q} &= PDP^\top (PI_\sigma)\tilde{D}(PI_\sigma)^\top \\ &= PD(I_\sigma \tilde{D} I_\sigma^\top)P^\top \\ &= P(I_\sigma \tilde{D} I_\sigma^\top)DP^\top \\ &= PI_\sigma \tilde{D} I_\sigma^\top (P^\top P) DP^\top \\ &= (PI_\sigma)\tilde{D}(PI_\sigma)^\top PDP^\top \end{aligned}$$

$$= \tilde{Q}Q,$$

where we use the orthogonality of P in the second and fourth equalities, and Lemma 2.4(vi) combined with the fact that diagonal matrices commute in the third one. □

Proposition 2.7. *Let $Q_1, \dots, Q_m \in \mathbb{R}^{n\times n}$ be symmetric positive definite matrices such that $Q_1 = P_1 D_1 P_1^\top$, where P_1 is orthogonal and D_1 is diagonal. Let $\sigma^1 = (1, \dots, n)$ and σ^i be a permutation on J_n for all $i = 2, \dots, m$. If Q_1 is σ^i-related to Q_i for every i, then there exists a Q_i-conjugate Hamel basis for every i.*

Proof. We have $Q_i = (P_1 I_{\sigma^i}) D_i (P_1 I_{\sigma^i})^\top$ with D_i diagonal for $i = 1, \dots, m$. Consider $\{w^0, \dots, w^{n-1}\}$ where $w^{j-1} = P_1 e^j$ for $j = 1, \dots, n$. The result follows by Lemma 2.6(i). □

Remark 2.8. Let $i = 1, \dots, m$. If $Q_1 = P_1 D_1 P_1^\top$ is σ^i-related to Q_i, then $Q_i = P_1 (I_{\sigma^i} D_i I_{\sigma^i}^\top) P_1^\top$. By Lemma 2.4(vi), the matrix $I_{\sigma^i} D_i I_{\sigma^i}^\top$ is diagonal. Therefore, Q_1 is σ^i-related to Q_i for every i if and only if $Q_1, \dots, Q_m$ are simultaneously diagonalizable. The commutativity of $Q_1, \dots, Q_m$ is not a criterion for determining whether or not they are simultaneously diagonalizable. In fact, the converse of Lemma 2.6(ii) is not true. For instance, take $\tilde{Q} = I$ and any $n \times n$ symmetric positive definite matrix Q whose set of eigenvectors is not the canonical basis, so $\tilde{Q}Q = Q\tilde{Q}$. However, by Lemma 2.6(i), the matrices Q and $\tilde{Q}$ are not σ-related for any $\sigma \in J_n$. Given $Q_1, \dots, Q_m$, if $Q_i Q_j \neq Q_j Q_i$ for some $i \neq j$, then there does not exist a basis for $\mathbb{R}^n$ comprised of common eigenvectors. Therefore, for symmetric positive definite matrices $Q_1, \dots, Q_m \in \mathbb{R}^{n\times n}$, the existence of a Q_i-conjugate Hamel basis for $i = 1, \dots, m$ requires checking that these matrices are simultaneously diagonalizable, or equivalently, that they have the same eigenvectors.

3. The quadratic multicriteria problem

Let $f\colon \mathbb{R}^n \to \mathbb{R}^m$ be a quadratic function given by

$$f_i(x) = \frac{1}{2} x^\top Q_i x + (q^i)^\top x,$$

where $Q_i \in \mathbb{R}^{n\times n}$ is symmetric positive definite and $q^i \in \mathbb{R}^n$ for $i = 1, \dots, m$. Note that f is $\mathbb{R}^m_+$-strongly convex, i.e., each f_i is strongly convex. We consider the unconstrained multiobjective problem,

$$\min_{x\in\mathbb{R}^n} f(x) \tag{1}$$

in the weak-Pareto or Pareto sense.

A point $\bar{x} \in \mathbb{R}^n$ is a *weak-Pareto* optimum for Problem (1), if there does not exist $x \in \mathbb{R}^n$ such that

$$f(x) < f(\bar{x}),$$

and *Pareto* optimum for Problem (1), if there does not exist $x \in \mathbb{R}^n$ such that

$$f(x) \leq f(\bar{x}) \quad \text{and} \quad f(x) \neq f(\bar{x});$$

the inequalities $<$ and $\leq$ are defined componentwise.

There is no loss of generality in assuming that the quadratic objective components f_i's in Problem (1) have no zero-degree terms and that the corresponding Q_i's are symmetric. Indeed, we observe that $\bar{x} \in \mathbb{R}^n$ is a weak Pareto (Pareto) optimum for Problem (1) if and only if $\bar{x}$ is a weak Pareto (Pareto) optimum for $\min_{x\in\mathbb{R}^n} \tilde{f}(x)$, where $\tilde{f}_i(x) = f_i(x) + r_i$ with $r_i \in \mathbb{R}$ for $i = 1,\dots,m$. If $Q_i \in \mathbb{R}^{n\times n}$ is not symmetric, then $\bar{Q}_i = \left(Q_i + Q_i^\top\right)/2$ is symmetric and $x^\top \bar{Q}_i x = x^\top Q_i x$ for all i. Therefore solving Problem (1) is equivalent to solving $\min_{x\in\mathbb{R}^n} \bar{f}(x)$, where $\bar{f}_i(x) = (1/2)x^\top \bar{Q}_i x + (q^i)^\top x + r_i$ for $i = 1,\dots,m$.

We recall the following notions of monotonicity for functions from $\mathbb{R}^m$ to $\mathbb{R}$; see [9].

Definition 3.1. *A function $\Phi\colon \mathbb{R}^m \to \mathbb{R}$ is* weakly-increasing *(*w-increasing*) if*

$$u < v \quad \Rightarrow \quad \Phi(u) < \Phi(v),$$

and strongly-increasing *(*s-increasing*) if*

$$u \le v \text{ and } u \ne v \quad \Rightarrow \quad \Phi(u) < \Phi(v).$$

We may write $v > u$ and $v \ge u$ to mean $u < v$ and $u \le v$, respectively.

Every s-increasing function is w-increasing. Moreover, if Φ is w-increasing and continuous, then

$$u \le v \quad \Rightarrow \quad \Phi(u) \le \Phi(v), \tag{2}$$

and by the above observation this also holds when Φ is s-increasing and continuous.

Now we state a well-known result that is used in the sequel.

Proposition 3.2. *Let $g\colon \mathbb{R}^n \to \mathbb{R}^m$, $\Phi\colon \mathbb{R}^m \to \mathbb{R}$, $M \subset \mathbb{R}^n$, and $\bar{x} \in \operatorname{argmin}_{x\in M} \Phi(g(x))$.*

(i) *If Φ is w-increasing, then $\bar{x}$ is a weak Pareto solution to* $\min_{x\in M} g(x)$.
(ii) *If Φ is s-increasing, then $\bar{x}$ is a Pareto solution to* $\min_{x\in M} g(x)$.

Proof. See [9, Lemmas 5.14 and 5.24]. $\square$

4. The multiobjective algorithm

In this section we propose a conjugate directions-type algorithm for Problem (1). From now on, we assume that the following condition holds.

Assumption $\mathcal{E}$: There exists a Q_i-conjugate Hamel basis $\{w^0,\dots,w^{n-1}\} \subset \mathbb{R}^n$, where $Q_i \in \mathbb{R}^{n\times n}$ corresponds to f_i in Problem (1) for $i = 1,\dots,m$.

This condition is not vacuous. As we have seen in Section 2, for a certain type of problems, Assumption $\mathcal{E}$ holds for $w^j = P_1 e^j$, where $Q_i = P_1 I_{\sigma_i} D_i (P_1 I_{\sigma_i})^\top$, P_1 is orthogonal, D_i is diagonal, and σ_i is a permutation on J_n for $i = 1,\dots,m$.

Now we present the algorithm, which, as we will see, has to be performed $n+1$ times in order to reach an optimal solution to Problem (1).

Algorithm 1: Multiobjective Conjugate Directions Algorithm (weak version)

Step 1: Take a Q_i-conjugate set $\{w^0, \dots, w^{n-1}\} \subset \mathbb{R}^n$ for $i = 1, \dots, m$ and a w-increasing continuous auxiliary function $\Phi\colon \mathbb{R}^m \to \mathbb{R}$ such that $\Phi \circ f$ is strongly convex. Choose $x^0 \in \mathbb{R}^n$.

Step 2: For $k = 0, \dots, n-1$, let

$$t_k = \operatorname*{argmin}_{t\in\mathbb{R}} \Phi(f(x^k + tw^k)), \tag{3}$$

and set

$$x^{k+1} = x^k + t_k w^k. \tag{4}$$

The strong version of Algorithm 1 is similar to the weak one: it suffices to take an s-increasing continuous auxiliary function Φ in Step 1.

Some comments are in order. First, Step 1 is well-defined in view of Assumption $\mathcal{E}$. Step 2 is also well-defined. In fact, since the function $t \mapsto \Phi(f(x^k + tw^{k+1}))$ is strongly convex [1], Subproblem (3) has a unique minimizer t_k for all k. Therefore, starting from an arbitrary initial point x^0, the algorithm generates $x^1, \dots, x^n \in \mathbb{R}^n$ whose $\Phi \circ f$ values decrease. Indeed, by (3) and (4), we have

$$\Phi(f(x^{k+1})) = \Phi(f(x^k + t_k w^k)) \le \Phi(f(x^k)) \quad \text{for } k = 0, \dots, n-1. \tag{5}$$

At each iteration, Algorithm 1 requires the minimization of a (possibly nonsmooth) continuous scalar-valued function defined over the real line. If the unconstrained convex subproblems happen to be nonsmooth, in some cases they may be rewritten in such a way that the lost smoothness is retrieved at a certain cost. For instance, consider the w-increasing continuous auxiliary function defined by $\Phi(u) = \max_{i=1,\dots,m}\{u_i\}$. Then the Subproblem (3) can be expressed as

$$\min_{(t,\lambda)\in G} \psi_k(t, \lambda) = \lambda,$$

where $G = \{(t, \lambda) \in \mathbb{R} \times \mathbb{R}\colon f_i(x^k + tw^k) \le \lambda \text{ for } i = 1, \dots, m\}$.

For $m = 1$, Problem (1) consists in minimizing a strongly convex real-valued quadratic function $f(x) = (1/2)x^\top Qx + q^\top x$ over $\mathbb{R}^n$. So Assumption $\mathcal{E}$ is an empty condition: in Step 1 the set $\{w^0, \dots, w^{n-1}\}$ is Q-conjugate. Besides the directions $w^j = Pe^j$ from Proposition 2.7, we can also use the Gram-Schmidt procedure to obtain a Q-conjugate set of n elements; see [10, Theorem 1.17] or [1, page 120]. Moreover, the function $\Phi\colon \mathbb{R} \to \mathbb{R}$ is strictly increasing, so is immaterial in scalar minimization. Thus (3) is the same k^{th} subproblem of the classical scalar procedure. In other words, we retrieve the conjugate directions method.

In the following section we define a procedure that requires running Algorithm 1 with different initial points $n + 1$ times. Among the last iterates of the generated $n + 1$ sequences, we find an optimal point. In the scalar optimization case, performing Algorithm 1 yields the same optimum each of the $n + 1$ times.

[1] We point out that in [5, Subsection 3.2] there are examples of w- or s-increasing continuous auxiliary functions Φ such that $\Phi \circ f$ is strongly convex. For instance, $\Phi(u) = \max_{i=1,\dots,m}\{u_i\}$ is a w-increasing continuous function such that $\Phi \circ f$ is strongly convex.

One could ask why not simply apply the scalar-valued conjugate directions method to $\Phi \circ f$, obtain a minimizer of $\Phi \circ f$ in at most n iterations, and then apply Proposition 3.2 in order to reach an optimum for Problem (1). The answer is simple: the function $\Phi \circ f$ may not be quadratic.

Now we need to study some single-variable convex quadratic functions derived from the multiobjective f. For the initial guess $x^0 \in \mathbb{R}^n$ and the Q_i-conjugate Hamel basis $\{w^0, \dots, w^{n-1}\}$ for all i chosen in Step 1 of Algorithm 1, we consider the strongly convex quadratic function $\Delta_i^{(j)} \colon \mathbb{R} \to \mathbb{R}$ given by

$$\Delta_i^{(j)}(\tau) = \Delta_i^{(j)}(x^0, \tau) := \frac{1}{2}\big[(w^j)^\top Q_i w^j\big]\tau^2 + \big[(Q_i x^0 + q^i)^\top w^j\big]\tau, \tag{6}$$

for $i = 1, \dots, m$ and $j = 0, \dots, n-1$. We define

$$\mu_- = \mu_-(x^0) := \min\{\tau \colon \Delta_i^{(j)}(\tau) = 0 \text{ and } \tau < 0 \text{ for some } i = 1, \dots, m, j = 0, \dots, n-1\} \tag{7}$$

and

$$\mu_+ = \mu_+(x^0) := \max\{\tau \colon \Delta_i^{(j)}(\tau) = 0 \text{ and } \tau \geq 0 \text{ for some } i = 1, \dots, m, j = 0, \dots, n-1\}. \tag{8}$$

Then $\Delta_i^{(j)}(\tau) \geq 0$ for $\tau \in (-\infty, \mu_-] \cup [\mu_+, +\infty)$, $i = 1, \dots, m$ and $j = 0, \dots, n-1$. For the sake of simplicity, take

$$\mu = \mu(x^0) := \max\{|\mu_-|, \mu_+\}. \tag{9}$$

If the set in (7) is empty, we take $\mu = \mu_+$. The set in (8) is never empty, because $\tau = 0$ is a root of $\Delta_i^{(j)}$ for all i, j. Finally, by taking τ_j such that $|\tau_j| \geq \mu$ for $j = 0, \dots, n-1$, we guarantee that

$$\sum_{j=0}^{n-1} \Delta_i^{(j)}(\tau_j) \geq 0 \quad \text{for } i = 1, \dots, m. \tag{10}$$

5. Multiobjective Conjugate Directions Procedure (*MCDP*)

We begin this section by analyzing the behavior of the last iterate x^n generated by Algorithm 1. We show that $\Phi\left(f(x^n)\right)$ is a lower bound for $\Phi\left(f(\cdot)\right)$ outside a certain set $C \subset \mathbb{R}^n$, which depends on x^0 and μ. We also show that C is a union of the open coordinate hyperplanes with respect to the basis $\{w^0, \dots, w^{n-1}\}$ enlarged by a factor of 2μ in their corresponding orthogonal directions.

Proposition 5.1. *Suppose that $\{x^1, \dots, x^n\}$ is the sequence produced by Algorithm 1, implemented with the Q_i-conjugate basis $\{w^0, \dots, w^{n-1}\}$ for all i, the w- or s-increasing continuous auxiliary function $\Phi \colon \mathbb{R}^m \to \mathbb{R}$ and the initial guess $x^0 = x_0^0 w^0 + \dots + x_{n-1}^0 w^{n-1} \in \mathbb{R}^n$. Let $\tau_0, \dots, \tau_{n-1} \in \mathbb{R}$ be such that $|\tau_j| \geq \mu$ for $j = 0, \dots, n-1$, where μ is given by (9). Then*

$$\Phi(f(x^n)) \leq \Phi\Big(f(x^0 + \sum_{j=0}^{n-1} \tau_j w^j)\Big). \tag{11}$$

Moreover, if $\mu = 0$, then $x^n \in \operatorname{argmin}_{x \in \mathbb{R}^n} \Phi(f(x))$.

Proof. By using (5), the definition of f_i, the Q_i-conjugacy of $\{w^0, \dots, w^{n-1}\}$ for all i, the symmetry of Q_i, (6), and (10) combined with (2), we see that

$$\Phi(f(x^n)) \leq \Phi(f(x^0))$$

$$
\begin{aligned}
&= \Phi\Big([\frac{1}{2}(x^0 + \sum_{j=0}^{n-1} \tau_j w^j - \sum_{j=0}^{n-1} \tau_j w^j)^\top Q_i (x^0 + \sum_{j=0}^{n-1} \tau_j w^j - \sum_{j=0}^{n-1} \tau_j w^j) + \\
&\quad + (q^i)^\top (x^0 + \sum_{j=0}^{n-1} \tau_j w^j - \sum_{j=0}^{n-1} \tau_j w^j)]_{i=1}^m \Big) \\
&= \Phi\Big(\big[f_i(x^0 + \sum_{j=0}^{n-1} \tau_j w^j) - (x^0 + \sum_{j=0}^{n-1} \tau_j w^j)^\top Q_i (\sum_{j=0}^{n-1} \tau_j w^j) + \\
&\quad + \frac{1}{2} (\sum_{j=0}^{n-1} \tau_j w^j)^\top Q_i (\sum_{j=0}^{n-1} \tau_j w^j) - (q^i)^\top (\sum_{j=0}^{n-1} \tau_j w^j) \big]_{i=1}^m \Big) \\
&= \Phi\Big(\big[f_i(x^0 + \sum_{j=0}^{n-1} \tau_j w^j) - \sum_{j=0}^{n-1} \Delta_i^{(j)}(\tau_j) \big]_{i=1}^m \Big) \\
&\leq \Phi\Big(\big[f_i(x^0 + \sum_{j=0}^{n-1} \tau_j w^j) \big]_{i=1}^m \Big).
\end{aligned}
$$

If $\mu = 0$, then (11) holds for any $\tau_j \in \mathbb{R}$ with $j = 0, \dots, n-1$. By Lemma 2.2, the last iterate x^n is an unconstrained minimizer of $\Phi \circ f$. $\square$

We have just seen that $\Phi(f(x^n)) \leq \Phi(f(x))$ for $x \in \mathbb{R}^n \setminus C := \{x^0 + \sum_{j=0}^{n-1} \tau_j w^j : |\tau_j| \geq \mu \text{ for } j = 0, \dots, n-1\}$. We do not know whether or not x^n is a minimizer of $\Phi \circ f$ in $\mathbb{R}^n \setminus C$, since it may belong to C. The open set C consists of all affine combinations $x^0 + \sum_{j=0}^{n-1} \tau_j w^j$ with at least one $\tau_j \in (-\mu, \mu)$. The set of these sums is the union of the coordinate hyperplanes $\text{span}\{w^j\}^{\perp_{Q_1}}$, added by x^0 and enlarged by a factor of 2μ in the direction of $\text{span}\{w^j\}$, without the boundary, for $j = 0, \dots, n-1$. In other words, we have

$$C = C(x^0, \mu) = \bigcup_{j=0}^{n-1} H_\mu^j(x^0),$$

where

$$H_\mu^j(x^0) = \mathbb{R}^j \times (x_j^0 - \mu, x_j^0 + \mu) \times \mathbb{R}^{n-j-1}. \tag{12}$$

In the case $n = 2$ the set C is a 2μ-width cross, centered at $x^0 = x_0^0 w^0 + x_1^0 w^1$, unbounded in both senses of the left-right and up-down directions as Figure 1 indicates.

Our goal is to find a minimizer of $\Phi \circ f$ in the affine manifold $\{x^0 + \sum_{j=0}^{n-1} \tau_j w^j : \tau_j \in \mathbb{R} \text{ for } j = 0, \dots, n\}$, which coincides with $\mathbb{R}^n$ by Lemma 2.2. We plan to do this by performing Algorithm 1 $n+1$ times with cleverly chosen different initial points, and then comparing the $\Phi \circ f$ values of the respective last iterates. We now present the weak version of this scheme.

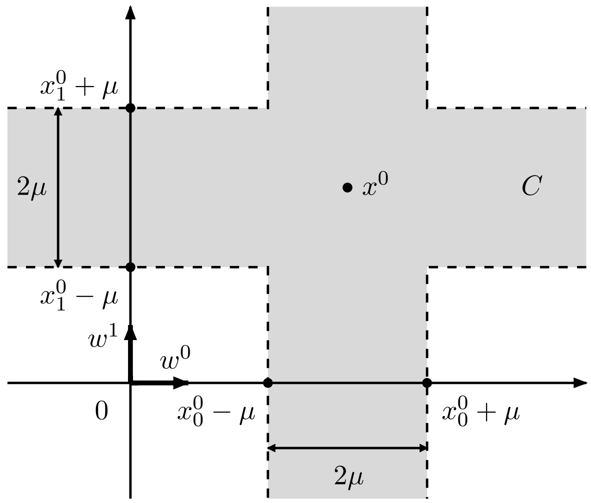

FIGURE 1. $C = \left[(x_0^0-\mu, x_0^0+\mu)\times\mathbb{R}\right] \cup \left[\mathbb{R}\times(x_1^0-\mu, x_1^0+\mu)\right]$.

Procedure 1: Multiobjective Conjugate Directions Procedure - MCDP(weak version)

Step 1: Take a Q_i-conjugate set $\{w^0,\dots,w^{n-1}\}\subset\mathbb{R}^n$ for $i=1,\dots,m$ and a w-increasing continuous auxiliary function $\Phi\colon\mathbb{R}^m\to\mathbb{R}$ such that $\Phi\circ f$ is strongly convex. Choose $x^0\in\mathbb{R}^n$.

Step 2: For $\ell=0,\dots,n$, apply Algorithm 1 with the following initial point

$$x^{\ell,0} = x^0 + \ell(2\bar{\mu}+1)\sum_{j=0}^{n-1} w^j, \tag{13}$$

where $\bar{\mu}=\max\{\mu(x^{\ell,0})\colon \ell=0,\dots,n\}$ and $\mu(x^{\ell,0})$ is given by (9). The sequence $\{x^{\ell,1},\dots,x^{\ell,n}\}$ is then produced.

Step 3: Set $\ell^* = \underset{\ell=0,\dots,n}{\operatorname{argmin}}\{\Phi(f(x^{\ell,n}))\}$ and $x^*=x^{\ell^*,n}$.

The strong version of Procedure 1 is similar to the weak one: it suffices to take an s-increasing continuous auxiliary function Φ in Step 1.

Let $\ell=0,\dots,n$. For each initial point $x^{\ell,0}$, there are $n+1$ sets of mn real-valued quadratic functions $\Delta_i^{(j)}=\Delta_i^{(j)}(x^{\ell,0},\cdot)$ for $i=1,\dots,m$ and $j=0,\dots,n-1$. Substituting μ by $\bar{\mu}$ gives that (10) holds for the $mn(n+1)$ quadratic functions $\Delta_i^{(j)}(x^{\ell,0},\cdot)$ for all i,j,ℓ. Whence, replacing both μ by $\bar{\mu}$ and $\{x^1,\dots,x^n\}$ by $\{x^{\ell,1},\dots,x^{\ell,n}\}$ in Proposition 5.1, we have

$$\Phi(f(x^{\ell,n}) \le \Phi(f(x)) \quad \text{for } x\in\mathbb{R}^n\setminus C^\ell \text{ and } \ell=0,\dots,n, \tag{14}$$

where

$$C^\ell := C(x^{\ell,0},\bar{\mu}) = \bigcup_{j=0}^{n-1} H_{\bar{\mu}}^j(x^{\ell,0}) \tag{15}$$

with $H^j_\mu(x^{\ell,0}) = \mathbb{R}^j \times (x^{\ell,0}_j - \mu, x^{\ell,0}_j + \mu) \times \mathbb{R}^{n-j-1}$.

6. Convergence analysis of *MCDP*

In this section we show that x^* from Step 3 of *MCDP* is an unconstrained optimum of $\Phi \circ f$. For the benefit of the reader, we first concentrate on the cases $n = 1$ and 2. By Proposition 5.1, from now on we assume $\bar{\mu} > 0$.

Let $n = 1$. By taking $w_0 = 1$, $\Phi\colon \mathbb{R}^m \to \mathbb{R}$, and the point $x_0 \in \mathbb{R}$, *MCDP* generates the initial points $x_{0,0} = x_0$ and $x_{1,0} = x_0 + 2\bar{\mu} + 1$ by (13), and then the sequences $\{x_{0,1}\}$ and $\{x_{1,1}\}$. Without loss of generality, we may assume $x^* = x_{0,1}$ in Step 3, so

$$\Phi(f(x^*)) \leq \Phi(f(x_{1,1})).$$

By using that

$$\Phi(f(x^*)) \leq \Phi(f(x)) \quad \text{ for } x \in \mathbb{R} \setminus C^0$$

and

$$\Phi(f(x_{1,1})) \leq \Phi(f(x)) \quad \text{ for } x \in \mathbb{R} \setminus C^1,$$

we obtain

$$\Phi(f(x^*)) \leq \Phi(f(x))$$

for $x \in (\mathbb{R} \setminus C^0) \cup (\mathbb{R} \setminus C^1) = \mathbb{R} \setminus (C^0 \cap C^1) = \mathbb{R}$, since

$$\begin{aligned} C^0 \cap C^1 &= (x_{0,0} - \bar{\mu}, x_{0,0} + \bar{\mu}) \cap (x_{1,0} - \bar{\mu}, x_{1,0} + \bar{\mu}) \\ &= (x_0 - \bar{\mu}, x_0 + \bar{\mu}) \cap (x_0 + \bar{\mu} + 1, x_0 + 3\bar{\mu} + 1) \\ &= \emptyset. \end{aligned}$$

So *MCDP* yields an unconstrained minimizer of $\Phi \circ f$ in at most two iterations of Algorithm 1.

Now let $n = 2$. By taking a Q_i-conjugate basis $\{w^0, w^1\}$ for every i, $\Phi\colon \mathbb{R}^m \to \mathbb{R}$ and $x^0 = x^0_0 w^0 + x^0_1 w_1 \in \mathbb{R}^2$ in Step 1, *MCDP* gives the initial points $x^{0,0}$, $x^{1,0}$ and $x^{2,0}$, and then the sequences $\{x^{0,1}, x^{0,2}\}$, $\{x^{1,1}, x^{1,2}\}$ and $\{x^{2,1}, x^{2,2}\}$. Without loss of generality, we may assume $x^* = x^{0,2}$ in Step 3, so

$$\Phi(f(x^*)) \leq \Phi(f(x^{\ell,2})) \quad \text{ for } \ell = 1, 2. \tag{16}$$

Since

$$\Phi(f(x^*)) \leq \Phi(f(x)) \quad \text{ for } x \in \mathbb{R}^2 \setminus C^0$$

and

$$\Phi(f(x^{\ell,2})) \leq \Phi(f(x)) \quad \text{ for } x \in \mathbb{R}^2 \setminus C^\ell \text{ and } \ell = 1, 2,$$

by (16) we conclude that

$$\Phi(f(x^*)) \leq \Phi(f(x))$$

for $x \in (\mathbb{R}^2 \setminus C^0) \cup (\mathbb{R}^2 \setminus C^1) \cup (\mathbb{R}^2 \setminus C^2) = \mathbb{R}^2 \setminus (C^0 \cap C^1 \cap C^2) = \mathbb{R}^2$; see Figures 2 and 3.

We are now in the position to state and prove that *MCDP* ends up with an unconstrained minimizer of $\Phi \circ f$ in at most $n(n+1)$ iterations of Algorithm 1 in the general case.

Theorem 6.1. *Let $\ell = 0, \ldots, n$. Suppose that $\{x^{\ell,1}, \ldots, x^{\ell,n}\}$ is the ℓ^{th} sequence produced by* MCDP*, implemented with the Q_i-conjugate basis $\{w^0, \ldots, w^{n-1}\}$ for all i, the w- or s-increasing continuous auxiliary function $\Phi\colon \mathbb{R}^m \to \mathbb{R}$, and $x^0 = x^0_0 w^0 + \cdots + x^0_{n-1} w^{n-1} \in \mathbb{R}^n$. Then*

$$x^* \in \operatorname*{argmin}_{x \in \mathbb{R}^n} \Phi(f(x)),$$

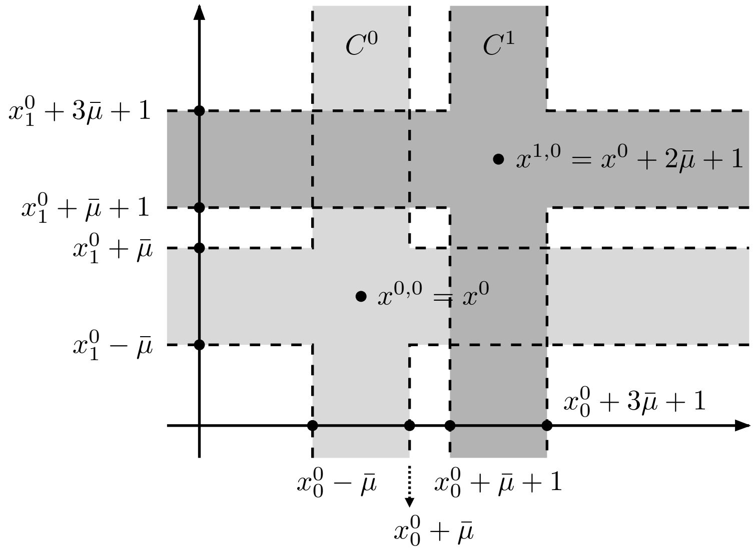

FIGURE 2. $C^0 \cap C^1$ is the union of two squares of sides $2\bar{\mu}$ centered at $(x_0^{0,0}, x_1^{1,0})$ and $(x_0^{1,0}, x_1^{0,0})$.

where x^* *is given by Step* 3 *of* MCDP.

Proof. We begin by studying the intersections of the sets C^ℓ. Recall that $C^0 = \bigcup_{j=0}^{n-1} H_{\bar{\mu}}^j(x^{0,0})$, where $H_{\bar{\mu}}^j(x^{0,0})$ is the n-dimensional affine manifold (without the border) parallel to span$\{w^j\}^{\perp_{Q_1}}$ passing through $x^{0,0} = x^0$, enlarged by a factor of $2\bar{\mu}$ in the (only bounded) direction span$\{w^j\}$; see (15).

The first intersection is $C^0 \cap C^1 = \bigcup_{j=0}^{n-1}(H_{\bar{\mu}}^j(x^{0,0}) \cap C^1)$. We claim that this is a union of the $n(n-1)$ open n-dimensional unbounded parallelepipeds $H_{\bar{\mu}}^j(x^{0,0}) \cap C^1$, each one bounded in exactly two directions. Indeed, every $H_{\bar{\mu}}^j(x^{0,0}) \cap C^1$ is formed by $n-1$ unbounded parallelepipeds which are the intersections of $H_{\bar{\mu}}^j(x^{0,0})$ with the nonparallel enlarged coordinate hyperplanes $H_{\bar{\mu}}^i(x^{1,0})$. More precisely, these parallelepipeds are given by cartesian products with $(x_j^{0,0} - \bar{\mu}, x_j^{0,0} + \bar{\mu})$ in the j^{th} position, $(x_i^{1,0} - \bar{\mu}, x_i^{1,0} + \bar{\mu})$ in the i^{th} position $(i \neq j)$, and $\mathbb{R}$ in the remaining positions. Since there are $n-1$ such possible cartesian products and n sets $H_{\bar{\mu}}^j(x^{0,0})$, the number of parallelepipeds in $C^0 \cap C^1$ is $n(n-1)$.

Similarly, $C^0 \cap C^1 \cap C^2$ is a union of $n(n-1)(n-2)$ unbounded open parallelepipeds. Indeed, every $(H_{\bar{\mu}}^j(x^{0,0}) \cap C^1) \cap C^2$ is a union of open parallelepipeds given by cartesian products with $(x_j^{0,0} - \bar{\mu}, x_j^{0,0} + \bar{\mu})$ in the j^{th} position, $(x_i^{1,0} - \bar{\mu}, x_i^{1,0} + \bar{\mu})$ in the i^{th} position $(i \neq j)$, $(x_k^{2,0} - \bar{\mu}, x_k^{2,0} + \bar{\mu})$ in the k^{th} position $(k \neq i, j)$, and $\mathbb{R}$ elsewhere. There are $n-2$ such possible cartesian products and $n(n-1)$ sets $H_{\bar{\mu}}^j(x^{0,0}) \cap C^1$, so we conclude that $C^0 \cap C^1 \cap C^2$ is a union of $n(n-1)(n-2)$ parallelepipeds.

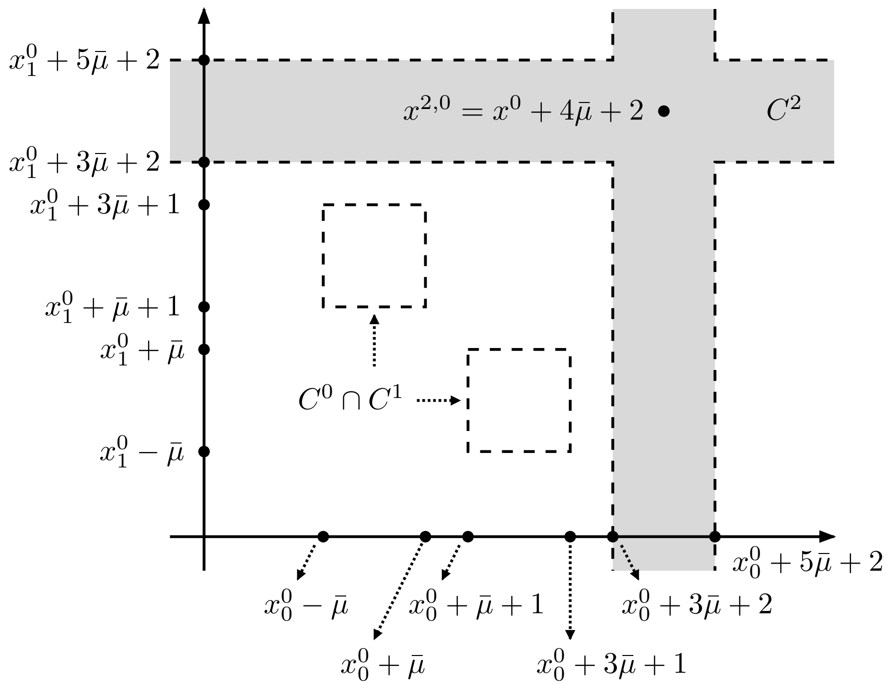

FIGURE 3. $C^0\cap C^1\cap C^2=\emptyset$.

Proceeding in the same way, we see that $C^0\cap\cdots\cap C^{n-1}$ is a union of $n!$ open bounded parallelepipeds. Explicitly, we have

$$C^0\cap\cdots\cap C^{n-1}=\bigcup_{\substack{\theta=(\theta_0,\ldots,\theta_{n-1})\\ \text{permutation on }\{0,\ldots,n-1\}}}(x_0^{\theta_0,0}-\bar{\mu},x_0^{\theta_0,0}+\bar{\mu})\times\cdots\times(x_{n-1}^{\theta_{n-1},0}-\bar{\mu},x_{n-1}^{\theta_{n-1},0}+\bar{\mu}). \quad (17)$$

We finally show that *MCDP* solves $\min_{x\in\mathbb{R}^n}\Phi(f(x))$. The procedure generates the sequence $\{x^{\ell,1},\ldots,x^{\ell,n}\}$ for $\ell=0,\ldots,n$. Each one of the $n+1$ last iterates satisfies

$$\Phi(f(x^{\ell,n}))\leq\Phi(f(x))\quad\text{ for } x\in\mathbb{R}^n\setminus C^\ell,$$

for $\ell=0,\ldots,n$; see (14)-(15). By Step 3 of *MCDP*, it follows that

$$\Phi(f(x^*))\leq\Phi(f(x))\quad\text{ for } x\in\left(\mathbb{R}^n\setminus C^0\right)\cup\cdots\cup(\mathbb{R}^n\setminus C^n)=\mathbb{R}^n\setminus\bigcap_{\ell=0}^{n}C^\ell. \quad (18)$$

In view of (13), $x^{\ell,0}=x^{\ell-1,0}+(2\bar{\mu}+1)\sum_{j=0}^{n-1}w^j$, so $x_j^{\ell,0}=x_j^{\ell-1,0}+2\bar{\mu}+1$ for $j=0,\ldots,n-1$ and $\ell=1,\ldots,n$. Thus

$$x_j^{n,0}-\bar{\mu}>x_j^{\ell,0}+\bar{\mu}\quad\text{ for } j=0,\ldots,n-1 \text{ and } \ell=0,\ldots,n-1.$$

Since C^n is the union of $H_\mu^j(x^{n,0})=\mathbb{R}^j\times(x_j^{n,0}-\mu,x_j^{n,0}+\mu)\times\mathbb{R}^{n-j-1}$ with $j=0,\ldots,n-1$, the $n!$ hypercubes that appear in (17) do not overlap with C^n, i.e.,

$$C^0\cap\cdots\cap C^{n-1}\cap C^n=\emptyset.$$

The result now follows from (18). □

We conclude the convergence analysis by showing that *MCDP* implemented with a w-increasing (s-increasing) Φ yields a w-Pareto (Pareto) optimum. In other words, the scheme produces a solution to Problem (1) after at most $n(n+1)$ iterations of Algorithm 1.

Corollary 6.2. *Let $\ell = 0, \dots, n$. Suppose that $\{x^{\ell,1}, \dots, x^{\ell,n}\}$ is the ℓ^{th} sequence produced by* MCDP*, implemented with the Q_i-conjugate basis $\{w^0, \dots, w^{n-1}\}$ for all i, the w- or s-increasing continuous auxiliary function $\Phi\colon \mathbb{R}^m \to \mathbb{R}$, and $x^0 = x^0_0 w^0 + \cdots + x^0_{n-1} w^{n-1} \in \mathbb{R}^n$. If Φ is w-increasing (s-increasing), then x^* is a w-Pareto (Pareto) optimal solution to Problem* (1).

Proof. It suffices to combine Theorem 6.1 with Proposition 3.2. □

In summary, under Assumption $\mathcal{E}$, by applying *MCDP* implemented with a fixed Q_i-conjugate Hamel basis for all i, a w-increasing (s-increasing) continuous function $\Phi\colon \mathbb{R}^m \to \mathbb{R}$, and $x^0 \in \mathbb{R}^n$, we obtain a weak Pareto (Pareto) optimal solution to Problem (1). This means that in at most $n(n+1)$ iterations of Algorithm 1 we can find a weak Pareto or a Pareto optimum to Problem (1).

7. An example

We present an ad hoc example of Problem (1) for which, by varying a parameter in the auxiliary function, *MCDP* furnishes the whole Pareto optimal set and consequently the whole Pareto frontier.

Example 7.1. Consider $f\colon \mathbb{R} \to \mathbb{R}^2$ defined by

$$f(x) = (x^2, x^2 - 2x)^\top.$$

For $f_1(x) = x^2$, we have $Q_1 = 1 \in \mathbb{R}^{1\times 1}$ and $q_1 = 0 \in \mathbb{R}^1$. For $f_2(x) = x^2 - 2x$, we have $Q_2 = Q_1 \in \mathbb{R}^{1\times 1}$ and $q_2 = -2 \in \mathbb{R}^1$. By taking the basis $\{w_0\}$ for $\mathbb{R}$ with $w_0 = 1$, Assumption $\mathcal{E}$ is satisfied. Since f_1 and f_2 decrease from $-\infty$ to 0, both increase from 1 to $+\infty$, and f_1 increases while f_2 decreases on $[0,1]$, the set of Pareto optima is given by $S_P = [0,1]$. We consider the family of auxiliary functions $\{\Psi_\omega\colon \mathbb{R}^2 \to \mathbb{R}\}_{\omega\in[0,2]}$ defined by $\Psi_\omega(u) = \max_{i=1,2}\{(u - \omega e^1)_i\}$. The function Ψ_ω is w-increasing and continuous, and $\Psi_\omega \circ f\colon \mathbb{R} \to \mathbb{R}$ is strongly convex. We apply *MCDP* with $\{w_0\}$, Ψ_ω for $\omega \in [0,2]$, and $x_0 \in \mathbb{R}$. By Corollary 6.2, we obtain a family $\{x^*_\omega\}_{\omega\in[0,2]}$ of weak Pareto optima. We see that $\{x^*_\omega\}_{\omega\in[0,2]} = [0,1]$ and, since there is no weak Pareto solution that is not Pareto optimum, we obtain all Pareto solutions to the problem. Starting from $x_{0,0} = x_0 \in \mathbb{R}$, Algorithm 1 produces the sequence $\{x_{0,1}\}$. We compute the steplength

$$\begin{aligned} t_{0,0} &= \operatorname*{argmin}_{t\in\mathbb{R}} \Psi_\omega(f(x_0 + t w_0)) \\ &= \operatorname*{argmin}_{t\in\mathbb{R}} \max_{i=1,2}\{f_i(x_0 + t) - \omega e^1\} \\ &= \operatorname*{argmin}_{t\in\mathbb{R}}\{(x_0 + t)^2 - \omega, (x_0 + t)^2 - 2(x_0 + t)\}, \end{aligned}$$

which is the argument, where $f_1(x_0 + t) - \omega$ crosses $f_2(x_0 + t)$. By solving $(x_0 + t)^2 - \omega = (x_0 + t)^2 - 2(x_0 + t)$, we get $t_{0,0} = \omega/2 - x_0$, so

$$x_{0,1} = x_{0,0} + t_{0,0} w_0 = x_0 + \left(\frac{\omega}{2} - x_0\right) 1 = \frac{\omega}{2}.$$

Hence the element generated by Algorithm 1 does not depend on the initial point, and the next application with the initial point $x_{1,0} = x_0 + 2\mu + 1$ yields $x_{1,1} = \omega/2$. So, for each auxiliary function Ψ_ω we get the Pareto optimum $x^*_\omega = \omega/2$. By applying *MCDP* with all $\omega \in [0, 2]$, we obtain $S_P = [0, 1]$.

8. Final remarks

We propose a conjugate directions-type method for unconstrained quadratic multiobjective problems. Essentially, the strategy consists in substituting the unconstrained multicriteria problem by a finite sequence of single-variable unconstrained scalar-valued convex optimization problems, all with a single optimal solution. Depending on the chosen auxiliary function, the procedure yields a weak Pareto or a Pareto optimum.

Example 7.1 suggests that it may be worth to investigate which classes of quadratic multiobjective problems are such that the scheme produces all optima by varying a parameter on the auxiliary function. The fact that *MCDP* furnishes an unconstrained minimizer of $\Phi \circ f$ (an apparent limitation) may be a good starting point for characterizing the classes of problems whose efficient frontier can be entirely computed. Another possible research direction is to explore applications of the scheme to multiobjective problems that satisfy weaker conditions.

9. Funding

The first author received support by the Grant-in-Aid for Scientific Research (C) (19K11840) from Japan Society for the Promotion of Science. The second author received support from FAPERJ/CNPq through PRONEX-Optimization E-26/010.001247/2016. The third author received support by a PSC-CUNY grant (611157-00 49), jointly funded by The Professional Staff Congress and The City University of New York. This paper was completed in part with support from A Room of One's Own.

References

[1] Bertsekas, D. P. *Nonlinear programming*, first ed. Athena Scientific Optimization and Computation Series. Athena Scientific, Belmont, MA, 1995.

[2] Da Cruz Neto, J. X., Da Silva, G. J. P., Ferreira, O. P., and Lopes, J. O. A subgradient method for multiobjective optimization. *Comput. Optim. Appl. 54*, 3 (2013), 461–472.

[3] Fliege, J., Graña Drummond, L. M., and Svaiter, B. F. Newton's method for multiobjective optimization. *SIAM J. Optim. 20*, 2 (2009), 602–626.

[4] Fliege, J., and Svaiter, B. F. Steepest descent methods for multicriteria optimization. *Math. Meth. Oper. Res. 51*, 3 (2000), 479–494.

[5] Fukuda, E. H., Graña Drummond, L. M., and Raupp, F. M. P. An external penalty-type method for multicriteria. *TOP 24*, 2 (2016), 493–513.

[6] Fukuda, E. H., Graña Drummond, L. M., and Raupp, F. M. P. A barrier-type method for multiobjective optimization. *Optimization 0* (2019), 1–17.

[7] Fukuda, E. H., and Graña Drummond, L. M. A survey on multiobjective descent methods. *Pesq. Oper. 34*, 3 (2014), 585–620.

[8] Izmailov, A., and Solodov, M. *Otimização. Vol. 2*, first ed. Instituto Nacional de Matemática Pura e Aplicada (IMPA), Rio de Janeiro, 2007. Métodos computacionais. [Computational methods].

[9] Jahn, J. *Vector Optimization – Theory, Applications and Extensions.* Springer, Erlangen, 2003.

[10] Larrotonda, A. R. *Álgebra lineal y geometría.* Eudeba, Buenos Aires, 1977.

[11] Lucambio Pérez, L. R., and Prudente, L. F. Nonlinear conjugate gradient methods for vector optimization. *SIAM J. Optim. 28*, 3 (2018), 2690–2720.

[12] Ortega, J. M. *Numerical analysis*, second ed., vol. 3 of *Classics in Applied Mathematics.* Society for Industrial and Applied Mathematics (SIAM), Philadelphia, PA, 1990. A second course.

[13] Tanabe, H., Fukuda, E. H., and Yamashita, N. Proximal gradient methods for multiobjective optimization and their applications. *Comput. Optim. Appl. 72*, 2 (2019), 339–361.

Department of Applied Mathematics and Physics, Graduate School of Informatics, Kyoto University, Kyoto 606-8501, Japan
E-mail address: `ellen@i.kyoto-u.ac.jp`

Faculdade de Administração e Ciências Contábeis, Universidade Federal do Rio de Janeiro, FACC-UFRJ, Rio de Janeiro, Brazil
E-mail address: `bolsigeno@gmail.com`

Department of Mathematics, New York City College of Technology, The City University of New York (CUNY), 300 Jay Street, Brooklyn, NY 11201 USA
E-mail address: `amasuda@citytech.cuny.edu`