Nonlinear-Programming Reformulation of the Order-Value Optimization problem *

R. Andreani† C. Dunder‡ J. M. Martínez§

August 16, 2004: 13.56

Abstract

Order-value optimization (OVO) is a generalization of the minimax problem motivated by decision-making problems under uncertainty and by robust estimation. New optimality conditions for this nonsmooth optimization problem are derived. An equivalent mathematical programming problem with equilibrium constraints is deduced. The relation between OVO and this nonlinear-programming reformulation is studied. Particular attention is given to the relation between local minimizers and stationary points of both problems.

Keywords: Order-value optimization, optimality conditions, nonlinear-programming, equilibrium constraints, optimization algorithms.

1 Introduction

Assume that $f_1, \dots, f_m$ are real-valued functions defined on an arbitrary set Ω. For each $x \in \Omega$ the values $f_1(x), \dots, f_m(x)$ are ordered in such a way that

$$f_{i_1(x)}(x) \leq f_{i_2(x)}(x) \leq \dots \leq f_{i_m(x)}(x).$$

*This research was supported by CNPq and FAPESP (PT 2001-04597-4)

†Department of Applied Mathematics, IMECC-UNICAMP, University of Campinas, CP 6065, 13081-970 Campinas SP, Brazil. This author was supported by FAPESP (Grant 01/05492-1) and CNPq (Grant 301115/96-6). E-Mail: andreani@ime.unicamp.br.

‡Department of Applied Mathematics, IME-USP, University of Sao Paulo, CEP 05508-900 Sao Paulo, SP, Brazil. E-mail: dunder@ime.usp.br

§Department of Applied Mathematics, IMECC-UNICAMP, University of Campinas, CP 6065, 13081-970 Campinas SP, Brazil. E-Mail: martinez@ime.unicamp.br

For all $p \in I \equiv \{1, \dots, m\}$, the $p-$order-value function $f : \Omega \to I\!R$ is defined by

$$f(x) = f_{i_p(x)}(x).$$

The function f is well defined, despite the fact that the set of indices $\{i_1(x), \dots, i_m(x)\}$ is not univocally defined. If $p = 1$, $f(x) = \min\{f_1(x), \dots, f_m(x)\}$ and, if $p = m$, $f(x) = \max\{f_1(x), \dots, f_m(x)\}$.

The OVO problem consists in minimizing the $p-$order-value function. In [2] a primal method with guaranteed convergence to points that satisfy a weak optimality condition was introduced. One of the motivations invoked in [2] for solving OVO was the estimation of parameters in situations where large and systematic errors are present. See [12]. In those cases the OVO technique seems to be useful to eliminate the influence of outliers.

When x is a vector of portfolio positions and $f_i(x)$ is the predicted loss of the decision x under the scenario i, the order-value function is the discrete Value-at-Risk (VaR) function, largely used in risk evaluations (see, for example, [13]). The relationship between the order-value function and the VaR function was unknown to the authors at the time they wrote [2]. Nevertheless, in [2] the application of OVO to decision making was mentioned.

This paper is organized as follows. In Section 2 we prove new optimality conditions for the OVO problem. In Section 3 we introduce the reformulation as a nonlinear-programming problem. In Section 4 we prove that stationary points of the sum of squares of infeasibilities are feasible points. In Section 5 we show that local minimizers of the OVO problem are KKT points of the reformulation. Conclusions are stated in Section 6.

Throughout this paper we assume that $\|\cdot\|$ denotes the Euclidian norm, although in many cases it can be replaced by an arbitrary norm in the finite dimensional space under consideration. We denote $e = (1, \dots, 1)$ and $I\!N = \{0, 1, 2, \dots\}$. As usually, we denote $\#A$ the number of elements of the set A.

2 Optimality conditions

In this section, assuming smoothness of the functions f_i, we derive optimality conditions for the OVO problem. The conditions derived here are stronger than the one used in [2]. First-order optimality conditions will be used in forthcoming sections in connection to the reformulation.

Assume that $f_i : \Omega \to I\!R$ for all $i = 1, \ldots, m$ and define, as in the introduction,

$$f(x) = f_{i_p(x)}(x)$$

for all $x \in \Omega$, where

$$f_{i_1(x)}(x) \leq \ldots \leq f_{i_m(x)}(x).$$

The OVO problem considered here is

$$\text{Minimize } f(x) \quad \text{s.t. } x \in \Omega. \tag{1}$$

From now on we assume that $\Omega \subset I\!R^n$ and all the functions f_i are continuous on Ω. In this case the $p-$order function f is continuous (see [2]).

The objective of this section is to prove optimality conditions for the OVO problem.

For all $x \in \Omega$ we define:

$$L(x) = \{i \in \{1, \ldots, m\} \mid f_i(x) < f(x)\}, \tag{2}$$

$$E(x) = \{i \in \{1, \ldots, m\} \mid f_i(x) = f(x)\}, \tag{3}$$

and

$$G(x) = \{i \in \{1, \ldots, m\} \mid f_i(x) > f(x)\}. \tag{4}$$

The sets $L(x), E(x)$ and $G(x)$, as well as the function $f(x)$, depend on the choice of p. However, we do not make this dependence explicit in order to simplify the notation.

Clearly, for all $x \in \Omega$,

$$\#L(x) < p \leq \#[L(x) \cup E(x)].$$

We say that a sequence $\{x^k\}$ is *feasible* if $\{x^k\} \subset \Omega$. Given $x \in \Omega$, a feasible sequence is said to be a *descent sequence* for the function ϕ if

$$\lim_{k \to \infty} x^k = x$$

and there exists $k_0 \in I\!N$ such that

$$\phi(x^k) < \phi(x) \;\; \forall \;\; k \geq k_0.$$

In the following theorem we give a characterization of local minimizers which, in turn, will be useful to prove optimality conditions.

Theorem 2.1. *Assume that $x \in \Omega$. Then, x is a local minimizer of the OVO problem (1) if, and only if, for all feasible sequences $\{x^k\}$ that converge to x,*

$$\#\{i \in E(x) \mid \{x^k\} \text{ is a descent sequence for } f_i\} < p - \#L(x). \tag{5}$$

Proof. Assume that x is a local minimizer and that (5) does not hold for all feasible sequences $\{x^k\}$ that converge to x. Then, there exists a feasible sequence $\{x^k\}$ that is a descent sequence for all $i \in D \subset E(x)$, where

$$\#D \geq p - \#L(x).$$

By continuity, there exists $\varepsilon > 0$ such that

$$f_i(y) < f(x) \ \ \forall \ \ i \in L(x), \ \ \|y - x\| \leq \varepsilon.$$

Moreover, there exists $k_0 \in I\!N$ such that

$$f_i(x^k) < f(x) \ \ \forall \ \ k \geq k_0, \ \ i \in D.$$

If $k_1 \geq k_0$ is large enough and $k \geq k_1$, $\|x^k - x\| \leq \varepsilon$, therefore

$$f_i(x^k) < f(x) \ \ \forall \ \ k \geq k_1, \ \ i \in D \cup L(x).$$

But $\#D \cup L(x) \geq p$, so

$$f(x^k) = f_{i_p(x^k)}(x^k) < f(x) \ \ \forall \, k \geq k_1.$$

This implies that x is not a local minimizer.

Conversely, assume that x is not a local minimizer of (1). Therefore, there exists a feasible sequence $\{x^k\}$ with $\lim_{k\to\infty} x^k = x$ such that

$$f(x^k) < f(x) \text{ for } k \text{ large enough.}$$

So, there exists $k_2 \in I\!N$ such that

$$f_{i_1(x^k)}(x^k) \leq \ldots \leq f_{i_p(x^k)}(x^k) = f(x^k) < f(x) \tag{6}$$

for all $k \geq k_2$.

Since there exist a finite number of sets of the form $\{i_1(x^k), \ldots, i_p(x^k)\}$, at least one of them is repeated infinitely many times in (6). This set will be called $\{i_1, \dots, i_p\}$. Thus, taking an appropriate subsequence (which is also a feasible sequence), we have:

$$f_{i_1}(x^k) \leq \dots \leq f_{i_p}(x^k) = f(x^k) < f(x)$$

for all $k \geq k_2$.

Since $f_i(x) > f(x)$ for all $i \in G(x)$, the continuity of the functions implies that the set $\{i_1, \dots, i_p\}$ does not contain elements of $G(x)$. So,

$$\{i_1, \dots, i_p\} \subset L(x) \cup E(x).$$

Therefore, for at least $p - \#L(x)$ elements of $E(x)$ we have that

$$f_i(x^k) < f(x)$$

if k is large enough. Thus, the sequence x^k is a descent sequence for at least $p - \#L(x)$ functions from the set $E(x)$. This completes the proof. $\square$

We say that $d \in I\!R^n$ is an *unitary tangent direction* to the set Ω at the point $x \in \Omega$ if there exists a feasible sequence $\{x^k\}$ such that

$$\lim_{k \to \infty} x^k = x$$

and

$$d = \lim_{k \to \infty} \frac{x^k - x}{\|x^k - x\|}.$$

The following theorems state optimality conditions related to tangent directions. We are going to assume that the functions f_i have continuous first derivatives. Under this assumption, although not necessarily differentiable, the function f is locally Lipschitzian.

Theorem 2.2. *Assume that x is a local minimizer of (1) and f_i has continuous first derivatives in a neighborhood of x for all $i \in E(x)$. Then, for all unitary tangent directions d,*

$$\#\{i \in E(x) \mid \langle d, \nabla f_i(x) \rangle < 0\} < p - \#L(x).$$

Proof. Assume that the thesis is not true. Then, there exists an unitary tangent direction d and a set $D_1 \subset E(x)$, such that $\#D_1 \geq p - \#L(x)$ and

$$\langle d, \nabla f_i(x) \rangle < 0$$

for all $i \in D_1$. Let $\{x^k\}$ be a feasible sequence that converges to x and such that

$$d = \lim_{k\to\infty} \frac{x^k - x}{\|x^k - x\|}.$$

By the differentiability of f_i, for all $i \in D_1$ we have that:

$$f_i(x^k) = f_i(x) + \langle \nabla f_i(x), x^k - x\rangle + o(\|x^k - x\|).$$

Therefore,

$$\frac{f_i(x^k) - f_i(x)}{\|x^k - x\|} = \left\langle \nabla f_i(x), \frac{x^k - x}{\|x^k - x\|} \right\rangle + \frac{o(\|x^k - x\|)}{\|x^k - x\|}.$$

Taking limits on the right-hand side, we have that for k large enough,

$$\frac{f_i(x^k) - f_i(x)}{\|x^k - x\|} \le \frac{\langle \nabla f_i(x), d\rangle}{2} < 0.$$

Therefore, for k large enough and for all $i \in D_1$,

$$f_i(x^k) < f_i(x).$$

This contradicts Theorem 2.1. $\square$

Theorem 2.2 justifies the following definition of first-order stationary points.

First-order stationary points

Assume that all the functions f_i that define the OVO problem have continuous first derivatives in an open set that contains Ω. We say that $x \in \Omega$ is a first-order stationary point for (1) if, for all unitary tangent directions d,

$$\#\{i \in E(x) \mid \langle d, \nabla f_i(x)\rangle < 0\} < p - \#L(x).$$

In the next theorems of this section we prove second-order necessary conditions and a sufficient condition for local minimizers. Although these results may be useful for future developments they will not be used in connection with the reformulation of OVO.

Theorem 2.3. *Assume that x is a local minimizer of (1) and f_i has continuous first and second derivatives in a neighborhood of x for all $i \in E(x)$. For all unitary tangent directions d, define*

$$D'(d) = \{i \in E(x) \mid \langle d, \nabla f_i(x) \rangle < 0\}$$

and

$$D''(d) = \{i \in E(x) \mid \nabla f_i(x) = 0 \text{ and } d^T \nabla^2 f_i(x) d < 0\}.$$

Then, for all unitary tangent direction d,

$$\#(D'(d) \cup D''(d)) < p - \#L(x).$$

Proof. Assume that the thesis is not true. Then, there exists $d \in I\!R^n$, an unitary tangent direction such that

$$\#(D'(d) \cup D''(d)) \geq p - \#L(x).$$

Let $\{x^k\}$ a feasible sequence that converges to x and

$$d = \lim_{k \to \infty} \frac{x^k - x}{\|x^k - x\|}.$$

If $i \in D'(d)$, the reasoning of the proof of Theorem 2.2 shows that $\{x^k\}$ is a descent sequence for f_i. Assume, now, that $i \in D''(d)$.

Since f_i has continuous second derivatives, we have that

$$f_i(x^k) = f_i(x) + \langle \nabla f_i(x), x^k - x \rangle + \frac{1}{2}(x^k - x)^T \nabla^2 f(x)(x^k - x) + o(\|x - x^k\|^2)$$

$$= f_i(x) + \frac{1}{2}(x^k - x)^T \nabla^2 f(x)(x^k - x) + o(\|x - x^k\|^2).$$

So,

$$\frac{[f_i(x^k) - f_i(x)]}{\|x - x^k\|^2} = \frac{1}{2} \frac{(x^k - x)^T}{\|x - x^k\|} \nabla^2 f(x) \frac{x^k - x}{\|x - x^k\|} + \frac{o(\|x - x^k\|^2)}{\|x - x^k\|^2}.$$

Taking limits, we have that, for k large enough, $f_i(x^k) < f_i(x)$. Therefore, $\{x^k\}$ is a descent sequence for f_i. This contradicts Theorem 2.1. □

Theorem 2.4. *Assume that $x \in \Omega$ and f_i has continuous second derivatives for all $i \in E(x)$. For all unitary tangent direction d define $S(d) = S_1(d) \cup S_2(d) \subset E(x)$ by*

$$S_1(d) = \{i \in E(x) \mid \langle d, \nabla f_i(x) \rangle > 0\},$$

and

$$S_2(d) = \{i \in E(x) \mid \nabla f_i(x) = 0 \quad and \quad d^T \nabla^2 f_i(x) d > 0\}.$$

Assume that, for all unitary tangent direction d,

$$\#S(d) > \#L(x) + \#E(x) - p,$$

Then, x is a local minimizer.

Proof. Assume that x is not a local minimizer. Then, by Theorem 2.1, there exists a descent sequence $\{x^k\}$ for at least $p - \#L(x)$ functions of the set $E(x)$. Define $S_3 \subset E(x)$ by

$$i \in S_3 \text{ iff } \{x^k\} \text{ is a descent sequence for } f_i.$$

Then $\#S_3 \geq p - \#L(x)$.

Take a convergent subsequence of $(x^k - x)/\|x^k - x\|$ and a subsequence of $\{x^k\}$ so that, for this subsequence,

$$\lim_{k\to\infty} \frac{x^k - x}{\|x^k - x\|} = d.$$

Then, d is an unitary tangent direction. Consider the sets $S_1(d)$ and $S_2(d)$ associated to d.

If $i \in S_1(d)$ then $\langle d, \nabla f_i(x)\rangle > 0$, so:

$$f_i(x^k) - f_i(x) = \langle \nabla f_i(x^k), x^k - x\rangle + o(\|x^k - x\|)$$

and

$$\frac{f_i(x^k) - f_i(x)}{\|x^k - x\|} = \left\langle \nabla f_i(x), \frac{x - x^k}{\|x - x^k\|} \right\rangle + \frac{o(\|x^k - x\|)}{\|x^k - x\|}.$$

Therefore, taking limits we see that $f_i(x^k) > f_i(x)$ for k large enough. Therefore, $\{x^k\}$ is not a descent sequence for f_i.

Analogously, if $i \in S_2(d)$,

$$f_i(x^k) - f_i(x) = \langle \nabla f_i(x), x^k - x\rangle + \frac{1}{2}(x^k - x)^T \nabla^2 f_i(x)(x^k - x) + o(\|x^k - x\|^2)$$

$$= \frac{1}{2}(x^k - x)^T \nabla^2 f_i(x)(x^k - x) + o(\|x^k - x\|^2).$$

Dividing by $\|x^k - x\|^2$ and taking limits, we obtain that $f_i(x^k) > f_i(x)$ for k large enough. So, $\{x^k\}$ is not a descent sequence.

Therefore $(S_1 \cup S_2) \cap S_3 = \emptyset$. So,

$$\#E(x) \geq \#(S_1 \cup S_2) + \#S_3 > \#L(x) + \#E(x) - p + p - \#L(x) = \#E(x),$$

which is a contradiction. □

3 Nonlinear-programming reformulation

The optimization problem (1) is a nonsmooth nonconvex minimization problem. In this section we transform it into a smooth nonlinear-programming problem. The constraints of this particular nonlinear programming problem are equilibrium constraints. See [16]. The use of nonlinear programming algorithms for solving mathematical programming problems with equilibrium constraints has been justified in recent papers [3, 7, 9, 10].

The following lemma prepares the theorem that justifies the equivalence result.

Lemma 3.1. *Assume that $z_1, \ldots, z_m$ are real numbers such that*

$$z_1 \leq z_2 \leq \ldots \leq z_m.$$

Then, for all $p \in \{1, \ldots, m\}$, there exist $r', u', w' \in I\!R^m$ such that (r', u', w', z_p) is a solution of the following problem:

$$\text{Minimize } z$$
$$s.t. \begin{cases} \sum_{i=1}^m r_i w_i = 0 \\ \sum_{i=1}^m (1 - r_i) u_i = 0 \\ \sum_{i=1}^m r_i = p \\ u_i - z + z_i - w_i = 0, \quad i = 1, \ldots, m \\ u \geq 0, \quad 0 \leq r \leq e, w \geq 0. \end{cases} \tag{7}$$

Proof. Define u', r' e w' by

$$\begin{array}{lll} r'_i = 1, & i = 1, \ldots, p, & \\ r'_i = 0, & i = p+1, \ldots, m, & \\ u'_i = z_p - z_i, & i = 1, \ldots, p, & \\ u'_i = 0, & i = p+1, \ldots, m, & \\ w'_i = 0, & i = 1, \ldots, p & \text{and} \\ w'_i = z_i - z_p, & i = p+1, \ldots, m & \end{array}$$

Clearly, (r', u', w', z_p) is a feasible point of (7) for which the objective function value is z_p. Now, assume that (r, u, w, z) is a feasible point such that $z < z_p$. But, by feasibility,

$$w_i = u_i + z_i - z \;\; \forall \;\; i = 1, \ldots, m.$$

Since $z < z_p \leq z_{p+1} \leq \ldots \leq z_m$ we have that

$$z_i - z > 0 \; \forall \; i = p, \ldots, m.$$

Therefore, since $u_i \geq 0$,

$$w_i = u_i + z_i - z > 0 \; \forall \; i = p, \ldots, m.$$

Thus, since $r_i w_i = 0$ for all $i = 1, \ldots, m$,

$$r_i = 0, \; \forall \; i = p, \ldots, m.$$

So, since $r \leq e$ and $p = \sum_{i=1}^m r_i$,

$$p = \sum_{i=1}^{m} r_i = \sum_{i=1}^{p-1} r_i \leq p - 1.$$

This is a contradiction. Therefore, (r, u, w, z) cannot be feasible.

This means that, for all feasible (r, u, w, z), we have that $z \geq z_p$. Since (r', u', w', z_p) is feasible, the proof is complete. $\square$

Now we are able to prove an equivalence result. In the next theorem we show that solving (1) is equivalent to solve a nonlinear-programming problem.

Theorem 3.1. *The point $x \in \Omega$ is a solution of the OVO problem (1) if, and only if, there exist $r', u', w' \in I\!R^m$ and $z' \in I\!R$ such that (x, r', u', w', z') is a solution of*

$$\textit{Minimize } z$$

$$s.t. \begin{cases} \sum_{i=1}^m r_i w_i = 0 \\ \sum_{i=1}^m (1 - r_i) u_i = 0 \\ \sum_{i=1}^m r_i = p \\ u_i - z + f_i(x) - w_i = 0 \quad i = 1, \ldots, m \\ u \geq 0, \;\; 0 \leq r \leq e, \;\; w \geq 0, \;\; x \in \Omega. \end{cases} \tag{8}$$

In that case, $z' = f(x)$.

Proof. By Lemma 3.1, given $x \in \Omega$, $f(x)$ is the value of z that solves

$$\text{Minimize } z$$

$$\text{s.t.} \begin{cases} \sum_{i=1}^m r_i w_i = 0 \\ \sum_{i=1}^m (1 - r_i) u_i = 0 \\ \sum_{i=1}^m r_i = p \\ u_i - z + f_i(x) - w_i = 0 \quad i = 1, \dots, m \\ u \geq 0, \;\; 0 \leq r \leq e, \;\; w \geq 0. \end{cases}$$

The desired result follows trivially from this fact. □

It is easy to see that in the case $p = m$, which corresponds to the minimax problem, the reformulation (8) reduces to the classical nonlinear programming reformulation of minimax problems:

$$\text{Minimize } z \quad \text{s. t. } z \geq f_i(x) \;\; i = 1, \dots, m, \; x \in \Omega.$$

So far, the global solutions of the OVO problem have been identified with the global solutions of the nonlinear-programming problem (8). Now we prove that such identification also exists between the local minimizers of both problems.

In a preparatory lemma we will prove that feasible points of (8) necessarily satisfy a set of simple relations. Before proving this lemma, and remembering the definition (3), we give three additional definitions.

If (x, r, u, w, z) is a feasible point of (8), we define

$$E_1(x, r, u, w, z) = \{i \in E(x) \mid r_i = 1\}, \tag{9}$$

$$E_0(x, r, u, w, z) = \{i \in E(x) \mid r_i = 0\}, \tag{10}$$

and

$$E_+(x, r, u, w, z) = \{i \in E(x) \mid 0 < r_i < 1\}. \tag{11}$$

In order to simplify the notation, we will write

$$E_1(x) = E_1(x, r, u, w, z),$$

$$E_0(x) = E_0(x, r, u, w, z),$$

$$E_+(x) = E_+(x, r, u, w, z),$$

Lemma 3.2. *Let (x, r, u, w, z) be a feasible point of (8). Then,*

$$z > f(x) \Rightarrow \#[L(x) \cup E(x)] = p, \tag{12}$$

$$f(x) \leq z \leq \min_{i \in G(x)} f_i(x), \tag{13}$$

$$u_i > 0 \; \forall \, i \in L(x), \tag{14}$$

$$r_i = 1 \;\; \forall \;\; i \in L(x), \tag{15}$$

$$w_i = 0 \;\; \forall \;\; i \in L(x). \tag{16}$$

Moreover, if $z = f(x)$, we have:

$$w_i > 0 \;\; \forall \;\; i \in G(x), \tag{17}$$

$$r_i = 0 \;\; \forall \;\; i \in G(x), \tag{18}$$

$$u_i = 0 \;\; \forall \;\; i \in G(x), \tag{19}$$

$$u_i = w_i = 0 \;\; \forall \;\; i \in E(x), \tag{20}$$

$$\sum_{i \in E(x)} r_i = p - \#L(x), \tag{21}$$

and

$$\#[E_1(x) \cup E_+(x)] \geq p - \#L(x). \tag{22}$$

Proof. Suppose that $z > f(x)$ and $\#L(x) + \#E(x) > p$. Then, by feasibility,

$$u_i = z - f_i(x) + w_i > 0 \;\; \forall \, i \in E(x) \cup L(x). \tag{23}$$

Since $u_i(1 - r_i) = 0$, (23) implies that

$$r_i = 1 \;\; \forall \, i \in E(x) \cup L(x).$$

This contradicts the fact that $\sum_{i=1}^{m} r_i = p$. Therefore, (12) is proved.

The fact that $f(x) \leq z$ is a direct consequence of Lemma 3.1.

Assume that $z > \min_{i \in G(x)} f_i(x)$. Then, $z > f(x)$. So, by (12),

$$\#[L(x) \cup \#E(x)] = p.$$

Then,

$$u_i = z - f_i(x) + w_i > 0 \tag{24}$$

for all $i \in E(x) \cup L(x)$ and for at least an additional index belonging to $G(x)$. Therefore, the inequality (24) holds for at least $p + 1$ indices. As in the proof of (12), this contradicts the fact that $\sum_{i=1}^m r_i = p$. Therefore, (13) is proved.

If $i \in L(x)$ we have that $f_i(x) < f(x) \leq z$. So, since $w_i \geq 0$,

$$u_i = w_i + z - f_i(x) > 0.$$

Thus, (14) is proved. Therefore, since $u_i(1 - r_i) = 0$, we deduce (15) and, since $r_i w_i = 0$, we obtain (16).

If $i \in G(x)$ and $z = f(x)$, we have that $f_i(x) > f(x) = z$. So, since $u_i \geq 0$, we obtain (17) and, since $r_i w_i = 0$, (18) is deduced. Then, since $(1 - r_i)u_i = 0$, we get (19).

If $i \in E(x)$, we have that $f_i(x) = f(x) = z$, therefore, since $u_i = w_i + z - f_i(x)$, we get

$$u_i = w_i \;\; \forall \;\; i \in E(x), \tag{25}$$

But

$$0 = w_i r_i = (1 - r_i)u_i,$$

then, by (25),

$$0 = w_i r_i = w_i(1 - r_i) \;\; \forall \;\; i \in E(x).$$

This implies (20).

By (15) and (18), since $\sum_{i=1}^m r_i = p$, we obtain (21). So, (22) also holds. □

In Lemma 3.2 we proved that, if (x, r, u, w, z) is a feasible point of the nonlinear-programming reformulation then $z \geq f(x)$. The possibility $z > f(x)$ is not excluded at feasible points of (8). However, in the following lemma we prove that, at local minimizers of (8), the identity $z = f(x)$ necessarily holds.

Lemma 3.3. *Assume that $(x_*, r_*, u_*, w_*, z_*)$ is a local minimizer of (8). Then, $z_* = f(x_*)$.*

Proof. By (13), since $(x_*, r_*, u_*, w_*, z_*)$ is feasible, we have that

$$f(x_*) \leq z_* \leq \min_{i \in G(x_*)} f_i(x_*).$$

Suppose that $z_* > f(x_*)$. By (12), $\#[E(x_*) \cup L(x_*)] = p$. Then, by the feasibility of $(x_*, r_*, u_*, w_*, z_*)$, we have that:

$$[r_*]_i = \begin{cases} 1 & i \in E(x_*) \cup L(x_*) \\ 0 & i \in G(x_*), \end{cases}$$

$$[u_*]_i = \begin{cases} z_* - f_i(x_*) & i \in E(x_*) \cup L(x_*) \\ 0 & i \in G(x_*) \end{cases}$$

and

$$[w_*]_i = \begin{cases} 0 & i \in E(x_*) \cup L(x_*) \\ f_i(x_*) - z_* & i \in G(x_*). \end{cases}$$

Define $\delta = z_* - f(x_*) > 0$ and, for all $k \in I\!N$,

$$z_k = z_* - \frac{\delta}{2(k+1)} < z_*,$$

$$[u^k]_i = \begin{cases} [u_*]_i - (z_* - z_k) & i \in E(x_*) \cup L(x_*) \\ 0 & i \in G(x_*), \end{cases}$$

$$r^k = r_*$$

and

$$[w^k]_i = \begin{cases} [w_*]_i + (z_* - z_k) & i \in G(x_*) \\ 0 & i \in E(x_*) \cup L(x_*). \end{cases}$$

Let us show that $\{(x_*, r^k, u^k, w^k, z_k)\}_{k \in I\!N}$ is feasible. Clearly,

$$\sum_{i=1}^{m} [r^k]_i [w^k]_i = 0, \quad \sum_{i=1}^{m} (1 - [r^k]_i)[u^k]_i = 0, \quad \sum_{i=1}^{m} [r^k]_i = p, \quad 0 \leq [r^k]_i \leq e.$$

Moreover:

(i) If $i \in L(x_*) \cup E(x_*)$,

$$[u^k]_i = [u_*]_i - (z_* - z_k) = [u_*]_i - \frac{\delta}{2(k+1)} \geq \delta - \frac{\delta}{2(k+1)} > 0$$

and

$$[u^k]_i - z_k + f_i(x_*) - [w^k]_i = [u_*]_i - (z_* - z_k) - z_k + f_i(x_*) =$$

$$= [u_*]_i - z_* + f_i(x_*) = 0.$$

(ii) If $i \in G(x_*)$,

$$[w^k]_i = [w_*]_i + (z_* - z_k) > 0$$

and

$$[u^k]_i - z_k + f_i(x_*) - [w^k]_i = -z_k + f_i(x_*) - ([w_*]_i + (z_* - z_k)) =$$

$$= -z_* + f_i(x_*) - [w_*]_i = 0.$$

Then, the sequence $\{(x_*, r^k, u^k, w^k, z_k)\}_{k \in I\!N}$ is feasible and converges to $(x_*, r_*, u_*, w_*, z_*)$. However, $z_k < z_*$ for all k, then $(x_*, r_*, u_*, w_*, z_*)$ is not a local minimizer. □

The following theorem states the relations between local minimizers of the OVO problem and its reformulation. Essentially, a local minimizer of (1) induces a natural local minimizer of (8). The reciprocal property needs and additional hypothesis which, in turn, will be shown to be unavoidable.

Theorem 3.2. *Assume that $x \in \Omega$ is a local minimizer of (1) and that $(x, r, u, w, f(x))$ is a feasible point of (8). Then, $(x, r, u, w, f(x))$ is a local minimizer of (8). Reciprocally, if, for some $z \in I\!R$, we have that (x, r, u, w, z) is a local minimizer of (8) whenever (x, r, u, w, z) is feasible, then $z = f(x)$ and x is a local minimizer of (1).*

Proof. Assume that $x \in \Omega$ is a local minimizer of (1) and that $(x, r, u, w, f(x))$ is a feasible point of (8). Suppose, by contradiction, that $(x, r, u, w, f(x))$ is not a local minimizer of (8). Therefore, there exists a sequence of feasible points $\{(x^k, r^k, u^k, w^k, z_k)\}$ such that

$$\lim_{k \to \infty} (x^k, r^k, u^k, w^k, z_k) = (x, r, u, w, f(x))$$

and

$$z_k < f(x) \;\; \forall \;\; k \in I\!N. \tag{26}$$

But, by Lemma 3.1, $f(x^k)$ is a minimum value of z among the points (r, u, w, z) that satisfy

$$\begin{cases} \sum_{i=1}^m r_i w_i = 0 \\ \sum_{i=1}^m (1 - r_i) u_i = 0 \\ \sum_{i=1}^m r_i = p \\ u_i - z + f_i(x^k) - w_i = 0, \;\; i = 1, \dots, m \\ u \geq 0, \;\; 0 \leq r \leq e, w \geq 0. \end{cases}$$

Moreover, by the feasibility of $(x^k, r^k, u^k, w^k, z_k)$,

$$\begin{cases} \sum_{i=1}^m r_i^k w_i^k = 0 \\ \sum_{i=1}^m (1 - r_i^k) u_i^k = 0 \\ \sum_{i=1}^m r_i^k = p \\ u_i^k - z_k + f_i(x^k) - w_i^k = 0, \;\; i = 1, \dots, m \\ u^k \ge 0, \;\; 0 \le r^k \le e, w^k \ge 0. \end{cases}$$

Therefore $f(x^k) \le z_k$. So, by (26),

$$f(x^k) < f(x) \;\; \forall \;\; k \in I\!N.$$

This implies that x is not a local minimizer of (1).

Conversely, let us assume that for some $z \in I\!R$, (x, r, u, w, z) is a local minimizer of (8) whenever (x, r, u, w, z) is feasible. By Lemma 3.3, this implies that $z = f(x)$. Assume, by contradiction, that x is not a local minimizer of (1). Then, there exists a sequence $\{x^k\} \subset \Omega$ such that

$$\lim_{k \to \infty} x^k = x$$

and

$$f(x^k) < f(x). \tag{27}$$

For all $k \in I\!N$ let us define $r^k, u^k, w^k \in I\!R^m$, $z_k \in I\!R$ by:

$$z_k = f(x^k),$$

$$r^k_{i_j(x^k)} = 1, \; j = 1, \dots, p,$$

$$r^k_{i_j(x^k)} = 0, \; j = p+1, \dots, m,$$

$$u^k_{i_j(x^k)} = f(x^k) - f_{i_j(x^k)}(x^k), \; j = 1, \dots, p,$$

$$u^k_{i_j(x^k)} = 0, \; j = p+1, \dots, m,$$

$$w^k_{i_j(x^k)} = 0,\ j = 1, \ldots, p,$$

$$w^k_{i_j(x^k)} = f_{i_j(x^k)}(x^k) - f(x^k),\ j = p+1, \ldots, m.$$

Clearly, $(x^k, r^k, u^k, w^k, z_k)$ is a feasible point of (8). Moreover, $z_k < f(x)$ for all k,

$$\lim_{k\to\infty} x^k = x \text{ and } \lim_{k\to\infty} z_k = f(x). \tag{28}$$

Since the set of permutations of $\{1, \ldots, m\}$ is finite, there exists one of them (say $(i_1, \ldots, i_m)$) such that

$$i_1 = i_1(x^k), \ldots, i_m = i_m(x^k)$$

infinitely many times. Taking the corresponding subsequence of the original $\{x^k\}$, we have that:

$$z_k = f(x^k),$$

$$r^k_{i_j} = 1,\ j = 1, \ldots, p,$$

$$r^k_{i_j} = 0,\ j = p+1, \ldots, m,$$

$$u^k_{i_j} = f(x^k) - f_{i_j}(x^k),\ j = 1, \ldots, p,$$

$$u^k_{i_j} = 0,\ j = p+1, \ldots, m,$$

$$w^k_{i_j} = 0,\ j = 1, \ldots, p,$$

$$w^k_{i_j} = f_{i_j}(x^k) - f(x^k),\ j = p+1, \ldots, m.$$

for all the indices of the new sequence. By the continuity of the functions f_i, we can take limits in the above equations, and we get that

$$\lim_{k\to\infty} (x^k, r^k, u^k, w^k, z_k) = (x, r, u, w, z),$$

where

$$z = f(x),$$

$$r_{i_j} = 1,\ j = 1, \ldots, p,$$

$$r_{i_j} = 0,\ j = p+1, \ldots, m,$$

$$u_{i_j} = f(x) - f_{i_j}(x),\ j = 1, \ldots, p,$$
$$u_{i_j} = 0,\ j = p + 1, \ldots, m,$$
$$w_{i_j} = 0,\ j = 1, \ldots, p,$$
$$w_{i_j} = f_{i_j}(x) - f(x),\ j = p + 1, \ldots, m.$$

By continuity, (x, r, u, w, z) is a feasible point of (8) and, by (27) and (28) it is not a local minimizer of (8). This completes the proof. $\square$

Remark. In the previous theorem we proved the identity between local minimizers of (1) and (8) in the following sense. On one hand, if x is a local minimizer of (1) then $(x, r, u, w, f(x))$ is a local minimizer of (8) for all feasible choices of r, u, w. On the other hand, if $(x, r, u, w, f(x))$ is a local minimizer of (8) *for all* feasible choices of r, u, w, then x is a local minimizer of the OVO problem. A natural question remains: if x is not a local minimizer of (1), is it possible that, *for a particular choice* of r, u, w, the point $(x, r, u, w, f(x))$ is a local minimizer of (8)? The following example shows that, in fact, this possibility exists. So, the "for all" assumption in the converse proof of Theorem 3.2 cannot be eliminated.

Let us consider the OVO problem defined by $n = 1$, $p = 2$ and

$$f_1(x) = x, \quad f_2(x) = 2x, \quad f_3(x) = 3x \quad \forall\, x \in I\!R.$$

In this case, the reformulation (8) is:

$$\text{Minimize } z$$
$$\begin{array}{lll} & (a) & r_1 w_1 + r_2 w_2 + r_3 w_3 = 0 \\ & (b) & (1 - r_1) u_1 + (1 - r_2) u_2 + (1 - r_3) u_3 = 0 \\ & (c) & r_1 + r_2 + r_3 = 2 \\ s.t. & (d) & u_1 - z + x - w_1 = 0 \\ & (e) & u_2 - z + 2x - w_2 = 0 \\ & (f) & u_3 - z + 3x - w_3 = 0 \\ & (g) & u_i \geq 0, \ \ 0 \leq r_i \leq e \ \ w_i \geq 0, \ \ i = 1, 2, 3 \end{array} \tag{29}$$

Clearly, $\bar{x} = 0$ is not a local minimizer of the OVO problem. Moreover, it is not a first-order stationary point.

However, defining

$$\bar{y} = (\bar{x}, \bar{r}, \bar{u}, \bar{w}, \bar{z}) = (0,\ (1, 0, 1),\ (0, 0, 0),\ (0, 0, 0),\ 0),$$

it is easy to verify that $\bar{y}$ is a local minimizer of (29).

4 Feasible points of the reformulation

In this section we assume that $\Omega \subset I\!R^n$ and the functions f_i are continuously differentiable.

In Theorem 4.1, we prove a practical important property of the feasible set of (8). This property says that stationary points of the sum of squares of infeasibilities are feasible points. This is an important result if one is planning to solve (8) using nonlinear-programming algorithms, since most reasonable nonlinear-programming methods converge to stationary points of the sum of squares of infeasibilities.

Suppose that the set Ω is defined by

$$\Omega = \{x \in I\!R^n \text{ such that } h(x) = 0, l_b \leq x \leq u_b\}, \tag{30}$$

where, perhaps, some bounds are infinite and h is a continuously differentiable vector-valued function. Consider the problem of minimizing the sum of squares of infeasibilities:

$$\text{Minimize } [\sum_{i=1}^{m} r_i w_i]^2 + [\sum_{i=1}^{m} (1 - r_i) u_i]^2 \tag{31}$$

$$+[(\sum_{i=1}^{m} r_i) - p]^2 + \sum_{i=1}^{m} [u_i - z + f_i(x) - w_i]^2 + \sum_{i=1}^{m} h_i^2(x) \tag{32}$$

$$\text{s.t. } u \geq 0, \;\; 0 \leq r \leq e, \;\; w \geq 0, \;\; l_b \leq x \leq u_b. \tag{33}$$

We want to know whether stationary points of (31-33) represent feasible points of (8). In the following lemma we state a simple property that will be useful for our further analysis.

Lemma 4.1. *Assume that $x \in \Omega$. Then, (r, u, w, z) satisfies the first-order optimality conditions of*

$$\textit{Minimize } [\sum_{i=1}^{m} r_i w_i]^2 + [\sum_{i=1}^{m} (1 - r_i) u_i]^2 \tag{34}$$

$$+[(\sum_{i=1}^{m} r_i) - p]^2 + \sum_{i=1}^{m} [u_i - z + f_i(x) - w_i]^2 \tag{35}$$

$$\textit{s.t. } u \geq 0, \;\; 0 \leq r \leq e, \;\; w \geq 0. \tag{36}$$

if, and only if, (x, r, u, z, w) satisfies the first-order optimality conditions of (31-33).

Proof. Write the optimality conditions of (31-33) and compare them with the ones of (34-36) using the fact that $x \in \Omega$. □

Due to Lemma 4.1, our question is whether, given $x \in \Omega$, the optimality conditions of (34-36) imply the fulfillment of the constraints of (8). In other words, we want to know if, with a feasible $x \in \Omega$ and a stationary (r, u, z, w) with respect to (34-36), we can be sure that $z = f_{i_p(x)}(x)$. The answer is positive, and is stated in the following theorem.

Theorem 4.1. *Assume that Ω is given by (30), the functions f_i and h are continuously differentiable, $x^* \in \Omega$ and $(x^*, r^*, u^*, w^*, z_*)$ is a stationary (KKT) point of (31-33). Then, $(x^*, r^*, u^*, w^*, z_*)$ satisfies the constraints of (8).*

Proof. Define

$$\theta_1 = \sum_{i=1}^{m} [r^*]_i [w^*]_i,$$

$$\theta_2 = \sum_{i=1}^{m} (1 - [r^*]_i)[u^*]_i,$$

$$\theta_3 = \sum_{i=1}^{m} [r^*]_i - p.$$

Since $(x^*, r^*, u^*, w^*, z_*)$ is a KKT point of (31-33), there exist $\mu_1, \mu_2, \mu_3, \mu_4 \in I\!R^m$ such that

$$2\theta_1[w^*]_i - 2\theta_2[u^*]_i + 2\theta_3 - [\mu_1]_i + [\mu_2]_i = 0, \quad i = 1, \ldots, m, \tag{37}$$

$$2\theta_2(1 - [r^*]_i) + 2([u^*]_i - z_* + f_i(x^*) - [w^*]_i) - [\mu_3]_i = 0, \quad i = 1, \ldots, m, \tag{38}$$

$$-2\sum_{i=1}^{m} ([u^*]_i - z_* + f_i(x^*) - [w^*]_i) = 0, \tag{39}$$

$$2\theta_1[r^*]_i - 2([u^*]_i - z_* + f_i(x^*) - [w^*]_i) - [\mu_4]_i = 0, \quad i = 1, \ldots, m, \tag{40}$$

$$[\mu_1]_i[r^*]_i = [\mu_2]_i(1 - [r^*]_i) = [\mu_3]_i[u^*]_i = [\mu_4]_i[w^*]_i = 0, \quad i = 1, \ldots, m, \tag{41}$$

$$u^* \geq 0, \; 0 \leq r^* \leq e, \; w^* \geq 0, \mu_1 \geq 0, \mu_2 \geq 0, \mu_3 \geq 0, \mu_4 \geq 0. \tag{42}$$

We consider four possibilities:

(a) Suppose that $\theta_1 = \theta_2 = 0$. By (38),(40) and (42), we have that

$$0 \le [u^*]_i - z_* + f_i(x^*) - [w^*]_i \le 0, \quad \text{for all } i.$$

This implies that $[u^*]_i - z_* + f_i(x^*) - [w^*]_i = 0$ for all i. Then (37)-(42) represent the KKT conditions of the following convex problem:

$$\text{Minimize } \ [(\sum_{i=1}^{m} r_i) - p]^2 \ \text{ s.t. } \ 0 \le r \le e.$$

Since $p \in [1, m]$, we have that $\theta_3 = 0$, and the desired result is proved.

(b) Suppose that $\theta_1 = 0$. Then, by (40) and (42), we obtain:

$$2([u^*]_i - z_* + f_i(x^*) - [w^*]_i) = -[\mu_4]_i \le 0, \ \ i = 1, \ldots, m.$$

By (39),

$$[u^*]_i - z_* + f_i(x^*) - [w^*]_i = 0, \ \ i = 1, \ldots, m.$$

Multiplying (38) by $[u^*]_i$ and using (41), we get:

$$2\theta_2 [u^*]_i (1 - [r^*]_i) = 0, \ \ i = 1, \ldots, m.$$

So, by the definition of θ_2 and by (42), we obtain $\theta_2 = 0$. Therefore, the feasibility result follows from **(a)**.

(c) Suppose that $\theta_2 = 0$. By (38) and (42),

$$2([u^*]_i - z_* + f_i(x^*) - [w^*]_i) = [\mu_3]_i \ge 0, \ \ i = 1, \ldots, m.$$

By (39),

$$[u^*]_i - z_* + f_i(x^*) - [w^*]_i = 0, \ \ i = 1, \ldots, m.$$

Multiplying (40) by $[w^*]_i$ and using (41), we get:

$$2\theta_1 [w^*]_i [r^*]_i = 0, \ \ i = 1, \ldots, m.$$

Then, by the definition of θ_1 and (42) we have that $\theta_1 = 0$. Therefore, the feasibility result also follows from **(a)**.

(d) Suppose that $\theta_1 > 0$ e $\theta_2 > 0$.

By the definition of θ_1, there exists an index k such that $[r^*]_k > 0$ and $[w^*]_k > 0$. By (40) and (41),

$$2\theta_1[r^*]_k = 2([u^*]_k - z_* + f_i(x^*) - [w^*]_k) > 0$$

and $[\mu_1]_k = 0$.

By (38) and the fact that $\theta_2 > 0$, we have that

$$0 \leq 2\theta_2(1 - [r^*]_k) = -2([u^*]_k - z_* + f_k(x^*) - [w^*]_k) + [\mu_3]_k.$$

So, $[\mu_3]_k > 0$. Therefore, by (41), we obtain that $[u^*]_k = 0$. So,

$$[\mu_1]_k = 0, [w^*]_k > 0, \text{ and } [u^*]_k = 0.$$

Thus, by (37),

$$2\theta_1[w^*]_k + 2\theta_3 + [\mu_2]_k = 0.$$

Therefore,

$$\theta_3 < 0. \tag{43}$$

On the other hand, by the definition of θ_2, there exists k such that $1 - [r^*]_k > 0$ and $[u^*]_k > 0$. By (38) and (41),

$$-2\theta_2(1 - [r^*]_k) = ([u^*]_k - z_* + f_i(x^*) - [w^*]_k) < 0$$

and $[\mu_2]_k = 0$. By (40) and since $\theta_1 > 0$, we get:

$$0 \leq 2\theta_1[r^*]_k = 2([u^*]_k - z_* + f_k(x^*) - [w^*]_k) + [\mu_4]_k.$$

This implies that $[\mu_4]_k > 0$. So, by (41), $[w^*]_k = 0$ and, therefore,

$$[w^*]_k = 0 \text{ and } [\mu_2]_k = 0.$$

So, by (37),

$$-2\theta_2[u^*]_k + 2\theta_3 - [\mu_1]_k = 0.$$

Therefore, $\theta_3 > 0$. This contradicts (43). So, the proof is complete. $\square$

5 KKT points of the reformulation

In this section we assume that $\Omega = I\!R^n$. The reformulation (8) is a smooth minimization problem with nonlinear (equilibrium-like) constraints. It includes the complementarity constraints

$$r_i w_i = 0, r_i \geq 0, w_i \geq 0$$

and

$$(1 - r_i)u_i = 0, r_i \leq 1, u_i \geq 0.$$

Complementarity constraints are responsible for the fact that no feasible point satisfies the Mangasarian-Fromovitz constraint qualification [17]. See [4, 22]. Mathematical programming problems with equilibrium constraints share this difficulty. See [3, 9, 16]. Therefore, minimizers of the problem might not satisfy the KKT optimality conditions of nonlinear programming and this might represent a difficulty for nonlinear programming algorithms. The main result of this section is that, at least when $\Omega = I\!R^n$, this possible drawback does not exist. We will prove that local minimizers of (1) generate KKT points of (8) regardless of the lack of regularity of the points.

Observe that, given $x \in \Omega$, it is easy to define r, u, w such that $(x^*, r, u, w, f(x^*))$ is a feasible point of (8). In fact, we may set

$$r_i = \begin{cases} 1 & i \in L(x) \\ \frac{p - \#L(x)}{\#E(x)} & i \in E(x) \\ 0 & i \in G(x), \end{cases}$$

$$u_i = \begin{cases} f(x) - f_i(x) & i \in L(x) \\ 0 & i \in E(x) \cup G(x), \end{cases}$$

$$w_i = \begin{cases} f_i(x) - f(x) & i \in G(x) \\ 0 & i \in L(x) \cup E(x). \end{cases}$$

Therefore, the only essential assumption of Theorem 5.1 below is that x^* is a first-order stationary point of OVO.

Theorem 5.1. *Assume that $\Omega = I\!R^n$ and let $x^* \in \Omega$ be a first-order stationary point of the OVO problem. Let $r, u, w \in I\!R^m$ such that (x^*, r, u, w, z_*)*

is a feasible point of (8) with $z_* = f(x^*)$. Then, (x^*, r, u, w, z_*) is a KKT point of (8).

Proof. We must prove that there exist multipliers $\gamma, \beta, \rho \in I\!R$, $\lambda, \theta, \pi, \mu_1, \mu_2 \in I\!R^m$ such that

$$\sum_{i=1}^{m} \lambda_i = 1, \tag{44}$$

and, for all $i = 1, \ldots, m$,

$$\sum_{i=1}^{m} \lambda_i \nabla f_i(x^*) = 0, \tag{45}$$

$$w_i \gamma - u_i \beta + \rho - [\mu_1]_i + [\mu_2]_i = 0, \tag{46}$$

$$(1 - r_i)\beta + \lambda_i - \theta_i = 0, \tag{47}$$

$$r_i \gamma - \lambda_i - \pi_i = 0, \tag{48}$$

$$u_i \theta_i = w_i \pi_i = r_i [\mu_1]_i = (1 - r_i)[\mu_2]_i = 0, \tag{49}$$

with

$$\theta \geq 0, \pi \geq 0, \mu_1 \geq 0, \mu_2 \geq 0. \tag{50}$$

Let us define $L(x^*), G(x^*), E(x^*), E_1(x^*), E_+(x^*), E_0(x^*)$ by (2)–(3) and (9)–(11).

By Lemma 3.2, the feasibility of (x^*, r, u, w, z_*) and the fact that $z_* = f(x^*)$, the possible values of r_i, u_i, w_i for indices i in the disjoint sets $L(x^*)$, $E_1(x^*)$, $E_0(x^*)$, $E_+(x^*)$ are the ones given in Table 1.

	$L(x^*)$	$E_1(x^*)$	$E_+(x^*)$	$E_0(x^*)$	$G(x^*)$
r_i	1	1	$\in (0, 1)$	0	0
u_i	> 0	0	0	0	0
w_i	0	0	0	0	> 0

Table 1. Possible values of r_i, u_i, w_i.

Since x^* is a first-order stationary point and, by (22), $\#[E_1(x^*) \cup E_+(x^*)] \geq p - \#L(x^*)$, it turns out that the set of $d \in I\!R^n$ such that

$$\langle \nabla f_i(x^*), d \rangle < 0 \ \ \forall i \in E_1(x^*) \cup E_+(x^*)$$

is empty. Therefore, $d = 0$ is a solution of the linear-programming problem

$$\text{Minimize } \ y$$

$$\text{s. t. } \langle \nabla f_i(x^*), d\rangle \le y \;\; \forall \;\; i \in E_1(x^*) \cup E_+(x^*).$$

Writing the KKT conditions for this problem, we obtain that, for all $i \in E_1(x^*) \cup E_+(x^*)$ there exists

$$\lambda_i^* \ge 0 \tag{51}$$

such that

$$\sum_{i \in E_+(x^*) \cup E_1(x^*)} \lambda_i^* = 1, \qquad \sum_{i \in E_+(x^*) \cup E_1(x^*)} \lambda_i^* \nabla f_i(x^*) = 0 \tag{52}$$

Let us define the multipliers

$$\lambda_i = \begin{cases} 0 & \text{if } i \in L(x^*) \cup G(x^*) \cup E_0(x^*) \\ \lambda_i^* & \text{if } i \in E_+(x^*) \cup E_1(x^*), \end{cases} \tag{53}$$

$$\gamma = \max_{i \in E_+(x^*) \cup E_1(x^*)} \frac{\lambda_i}{r_i}, \tag{54}$$

$$\beta = 0, \tag{55}$$

$$\rho = 0. \tag{56}$$

Let us define $[\mu_1]_i, [\mu_2]_i, \theta_i$ and π_i by Table 2.

	$i \in L(x^*)$	$i \in E_+(x^*)$	$i \in E_1(x^*)$	$i \in E_0(x^*)$	$i \in G(x^*)$
$[\mu_1]_i =$	0	0	0	0	$w_i\gamma$
$[\mu_2]_i =$	0	0	0	0	0
$\theta_i =$	0	λ_i	λ_i	0	0
$\pi_i =$	γ	$r_i\gamma - \lambda_i$	$\gamma - \lambda_i$	0	0

Table 2. Definition of multipliers.

Observe that, by (51) and (53) we have that $\pi \ge 0$ and $\theta \ge 0$. By (54) we have that $\mu_1 \ge 0$, then (50) is satisfied.

Now, we show that with these definitions of the multipliers, the equations (44)–(49) are satisfied. Clearly, (44) and (45) follow from (51), (52) and (53).

By Tables 1 and 2 it is straightforward to verify that the complementarity conditions (49) and are verified.

Let us analyze the remaining equations.

(a) Equation (46):

- $i \in L(x^*)$

$$w_i\gamma - u_i\beta + \rho - [\mu_1]_i + [\mu_2]_i = 0\gamma - u_i(0) + 0 - 0 + 0 = 0$$

- $i \in E(x^*)$

$$w_i\gamma - u_i\beta + \rho - [\mu_1]_i + [\mu_2]_i = (0)\gamma - 0 + 0 - 0 + 0 = 0$$

- $i \in G(x^*)$

$$w_i\gamma - u_i\beta + \rho - [\mu_1]_i + [\mu_2]_i = w_i\gamma - 0 + 0 - w_i\gamma + 0 = 0$$

(b) Equation (47):

- $i \in L(x^*)$

$$(1 - r_i)\beta + \lambda_i - \theta_i = (0)(0) + (0) - (0) = 0$$

- $i \in E_+(x^*) \cup E_1(x^*)$

$$(1 - r_i)\beta + \lambda_i - \theta_i = (1 - r_i)(0) + \lambda_i - \lambda_i = 0$$

- $i \in E_0(x^*)$

$$(1 - r_i)\beta + \lambda_i - \theta_i = (1 - 0)(0) + 0 - 0 = 0$$

- $i \in G(x^*)$

$$(1 - r_i)\beta + \lambda_i - \theta_i = (1 - 0)(0) + 0 - 0 = 0$$

(c) Equation (48):

- $i \in L(x^*)$

$$r_i\gamma - \lambda_i - \pi_i = \gamma - 0 - \gamma = 0$$

- $i \in E_+(x^*) \cup E_1(x^*)$

$$r_i\gamma - \lambda_i - \pi_i = r_i\gamma - \lambda_i - (r_i\gamma - \lambda_i) = 0$$

- $i \in E_0(x^*) \cup G(x^*)$

$$r_i\gamma - \lambda_i - \pi_i = (0)\gamma - 0 - 0 = 0$$

Therefore, the theorem is proved. $\square$

¿From the statement of Theorem 5.1, a natural question arises about the existence of KKT points of (8) (with $\Omega = I\!R^n$) such that $z_* > f(x^*)$. In the following, we prove that those points do not exist. So, all the KKT points of (8) satisfy $z_* = f(x^*)$.

Theorem 5.2. *Assume that $\Omega = I\!R^n$ and (x^*, r, u, w, z_*) is a KKT point of (8). Then $z_* = f(x^*)$.*

Proof. We have already proved that the feasibility of (x^*, r, u, w, z_*) implies that

$$f(x^*) \le z_* \le \min_{i \in G(x^*)} f_i(x^*).$$

Assume that $z_* > f(x^*)$. Then

$$z_* - f_i(x^*) > 0 \;\forall\, i \in L(x^*) \cup E(x^*).$$

But, by feasibility,

$$z_* - f_i(x^*) = u_i - w_i,$$

then

$$u_i > w_i \ge 0 \;\forall\, i \in L(x^*) \cup E(x^*). \tag{57}$$

So, by the complementarity condition $(1 - r_i)u_i = 0$,

$$r_i = 1 \;\forall\, i \in L(x^*) \cup E(x^*) \tag{58}$$

and, by the complementarity condition $r_i w_i = 0$,

$$w_i = 0 \;\forall\, i \in L(x^*) \cup E(x^*).$$

Then, by (47),

$$\lambda_i = \theta_i \;\forall\, i \in L(x^*) \cup E(x^*).$$

But, by (49) and (57),

$$\theta_i = 0 \;\forall\, i \in L(x^*) \cup E(x^*),$$

therefore,

$$\lambda_i = 0 \;\forall\, i \in L(x^*) \cup E(x^*). \tag{59}$$

Now, since $\#[L(x^*) \cup E(x^*)] \geq p$, by (58) and $\sum_{i=1}^m r_i = p$, we have that

$$r_i = 0 \; \forall \; i \in G(x^*).$$

Then, by (48),

$$\lambda_i = -\pi_i \leq 0 \; \forall \; i \in G(x^*). \tag{60}$$

So, by (59) and (60),

$$\sum_{i=1}^m \lambda_i \leq 0.$$

This contradicts (44). Therefore, (x^*, r, u, w, z_*) is not a KKT point. □

Remark. The inspection of the proof of Theorem 5.1 shows that we have proved something stronger than the KKT thesis. In fact, we proved that, when x^* is a first-order stationary point of the OVO problem and $(x^*, r, u, w, f(x^*))$ is feasible for (8), there exist multipliers $\gamma \in I\!R$, $\lambda, \pi \in I\!R^m$ such that

$$\sum_{i=1}^m \lambda_i = 1, \tag{61}$$

$$\lambda + \pi = \gamma r, \tag{62}$$

$$\sum_{i=1}^m \lambda_i \nabla f_i(x^*) = 0 \tag{63}$$

and, for all $i = 1, \ldots, m$,

$$u_i \lambda_i = w_i \pi_i = 0, \tag{64}$$

with

$$\lambda \geq 0, \pi \geq 0, \gamma \geq 0. \tag{65}$$

It is easy to verify, using Tables 1 and 2, that the conditions (61)–(65), together with the feasibility conditions, imply the KKT conditions (44)–(50).

6 Final remarks

The reliability and efficiency of many nonlinear-programming algorithms is closely related to the fulfillment of KKT conditions at the solution. Sequential quadratic programming methods, for example, can be thought as modifications of the Newton method applied to the KKT nonlinear system.

See [6, 8, 11]. In feasible and semifeasible nonlinear-programming methods [1, 15, 19, 18, 20, 23] efficiency is linked to the possibility of decreasing a good approximation of the Lagrangian function on a linear approximation of the feasible region. Of course, a good approximation of the Lagrangian is only possible if Lagrange multipliers at the solution exist. Primal-dual interior point methods (see, for example, [5]) are also based on simultaneous approximation of the primal and the dual (Lagrangian) solution of the problem. (It is worthwhile to mention that the application of nonlinear-programming algorithms to problems whose feasible points do not satisfy regularity conditions has been considered in recent papers [3, 7, 9, 10].)

Therefore, the fact that minimizers are KKT points may be invoked as a strong argument to try ordinary optimization methods for solving the OVO problem. Our preliminary experience (with an augmented Lagrangian method [14]) has been encouraging in the sense that we realised that the main difficulty is associated to the existence of many local minimizers of the problem. No visible stability problems appeared in spite of the lack of fulfillment of the Mangasarian-Fromovitz [17] constraint qualification.

In [2] we introduced a primal method, without additional variables, that deal directly with the nonsmoothness of (1). Primal methods have the advantage of dealing with a smaller number of variables, but their convergence is guaranteed to points that satisfy a weaker optimality condition than the one considered in this paper. On the other hand, primal methods as the one introduced in [2] are complementary to methods based in the approach of the present paper in the following sense: If a weak stationary point found by the primal method is not a KKT point of the nonlinear programming problem, a suitable nonlinear programming algorithm can be applied starting from this point until a feasible point with a smaller objective function value is found. Since the functional values obtained by the primal method are strictly decreasing, this guarantees that the spurious weak stationary point will not be found again.

The development of specific algorithms for (1) and (8) is a matter of future research. The relationship between the general OVO problem and the minimax problem must be exploited. As mentioned before, the minimax problem corresponds to OVO with $p = m$. Since many effective algorithms for minimax exist (see, for example, [21]) suitable generalizations of these algorithms are likely to be effective for solving the OVO problem.

The applicability of the OVO problem to important practical problems is, of course, linked to the effectiveness of general or specific methods for its solution. The difficulties of minimizing the Value-at-Risk (one of the most stimulating practical problems related with OVO) are mentioned in many papers that can be found in the web site www.gloriamundi.org.

Acknowledgement

We are indebted to an anonymous referee whose comments helped us a lot to improve the first version of this paper.

References

[1] J. Abadie and J. Carpentier. Generalization of the Wolfe Reduced-Gradient Method to the Case of Nonlinear Constraints. In *Optimization.* Edited by R. Fletcher, Academic Press, New York. (1968) 37-47.

[2] R. Andreani, C. Dunder and J. M. Martínez. Order-Value Optimization: formulation and solution by means of a primal Cauchy method. *Mathematical Methods of Operations Research* 58, pp. 387-399 (2003).

[3] R. Andreani and J. M. Martínez. On the solution of mathematical programming problems with equilibrium constraints. *Mathematical Methods of Operations Research* (54) (2001) 345-358.

[4] D. Bertsekas. *Nonlinear Programming*, 2nd edition, Athena Scientific, Belmont, Massachusetts, 1998.

[5] R. H. Byrd, M. E. Hribar and J. Nocedal. An interior point algorithm for large-scale nonlinear programming, *SIAM Journal on Optimization* (9) (1999) 877-900.

[6] J. E. Dennis, M. El-Alem and M. C. Maciel. A Global Convergence Theory for General Trust-Region-Based Algorithms for Equality Constrained Optimization. *SIAM Journal on Optimization* (7) (1997) 177-207.

[7] M. C. Ferris, S. P. Dirkse and A. Meeraus. *Mathematical programs with equilibrium constraints: Automatic reformulation and solution via constrained optimization.* Numerical Analysis Group Research Report NA 02-11, Oxford University Computing Laboratory, Oxford University, 2002.

[8] R. Fletcher. *Practical Methods Optimization.* Wiley, New York, 1987.

[9] R. Fletcher and S. Leyffer. *Numerical experience with solving MPECs as NLPs*. University of Dundee Report NA 210, 2002.

[10] R. Fletcher, S. Leyffer, D. Ralph and S. Scholtes. *Local Convergence of SQP Methods for Mathematical Programs with Equilibrium Constraints*, University of Dundee Report NA 209, 2002.

[11] F. M. Gomes, M. C. Maciel, J. M. Martínez. Nonlinear programming algorithms using trust regions and augmented Lagrangians with nonmonotone penalty parameters. *Mathematical Programming* (84) (1999) 161-200.

[12] P. J. Huber. *Robust statistics*. Wiley, New York, 1981.

[13] P. Jorion. *Value at risk: the new benchmark for managing financial risk*, 2nd edition. Mc Graw-Hill, New York, 2001.

[14] N. Krejić, J. M. Martínez, M. P. Mello and E. A. Pilotta. Validation of an Augmented Lagrangian algorithm with a Gauss-Newton Hessian approximation using a set of hard-spheres problems. *Computational Optimization and Applications* (16) (2000) 247-263.

[15] L. S. Lasdon. Reduced Gradient Methods. In *Nonlinear Optimization 1981*. Edited by M. J. D. Powell. Academic Press, New York.(1982) 235-242.

[16] Z-Q. Luo, J-S. Pang and D. Ralph. *Mathematical programs with Equilibrium Constraints*. Cambridge University Press, Cambridge, 1996.

[17] O. L. Mangasarian and S. Fromovitz. Fritz-John necessary optimality conditions in presence of equality and inequality constraints. *Journal of Mathematical Analysis and Applications* (17) (1967) 37-42.

[18] J. M. Martínez. Inexact-restoration method with Lagrangian tangent decrease and new merit function for nonlinear programming. *Journal of Optimization Theory and Applications* (111) (2001) 39-58.

[19] J. M. Martínez and E. A. Pilotta. Inexact restoration algorithm for constrained optimization, *Journal of Optimization Theory and Applications* (104) (2000) 135-163.

[20] A. Miele, H. Y. Huang and J. C. Heideman. Sequential Gradient-Restoration Algorithm for the Minimization of Constrained Functions, Ordinary and Conjugate Gradient Version. *Journal of Optimization Theory and Applications* (4)(1969) 213-246.

[21] E. Polak. *Optimization: Algorithms and Consistent Approximations*, Springer-Verlag, New York, 1997.

[22] R. T. Rockafellar. Lagrange multipliers and optimality. *SIAM Review* (35) (1993) 183-238.

[23] J. B. Rosen. The Gradient Projection Method for Nonlinear Programming, Part 2, Nonlinear Constraints. *SIAM Journal on Applied Mathematics* (9) (1961) 514-532.