A VARIANT OF THE TOPKIS-VEINOTT METHOD FOR SOLVING INEQUALITY CONSTRAINED OPTIMIZATION PROBLEMS [1]

J.R. BIRGE
Department of Industrial and Operations engineering
University of Michigan, Ann Arbor, MI 48109, USA

L. QI and Z. WEI
School of Mathematics,
University of New South Wales, Sydney NSW 2052, Australia

Technical Report 97-11
June 18, 1997

A Variant of the Topkis-Veinott Method for Solving Inequality Constrained Optimization Problems [1]

J. R. Birge
Department of Industrial and Operations Engineering
University of Michigan, Ann Arbor, MI 48109, USA

L. Qi and Z. Wei
School of Mathematics,
University of New South Wales, Sydney NSW 2052, Australia

June 18, 1997

Abstract. In this paper, we give a variant of the Topkis-Veinott method for solving inequality constrained optimization problems. This method uses a linearly constrained positive semi-definite quadratic problem to generate a feasible descent direction at each iteration. Under mild assumptions, the algorithm is shown to be globally convergent in the sense that every accumulation point of the sequence generated by the algorithm is a Fritz-John point of the problem. We introduce a Fritz-John (FJ) function, an FJ1 strong second-order sufficiency condition (FJ1-SSOSC) and an FJ2 strong second-order sufficiency condition (FJ2-SSOSC), and then show, without any constraint qualification (CQ), that (i) if an FJ point z satisfies the FJ1-SSOSC, then there exists a neighborhood $N(z)$ of z such that for any FJ point $y \in N(z) \setminus \{z\}$, $f_0(y) \neq f_0(z)$, where f_0 is the objective function of the problem; (ii) if an FJ point z satisfies the FJ2-SSOSC, then z is a strict local minimum of the problem. The result (i) implies that the entire iteration point sequence generated by the method converges to an FJ point. We also show that if the parameters are chosen large enough, a unit step length can be accepted by the proposed algorithm.

Key Words. Constrained optimization, feasibility, global convergence, unit step, stochastic programming.

AMS subject classification. 65K10, 90C30, 52A41.

[1]This material is based on work supported by the National Science Foundation under Award Number DDM-9215921 and DMII-9523275 and the Australian Research Council.

1. Introduction

Consider the optimization problem

$$\min\{f_0(x) \mid x \in X\}, \tag{1.1}$$

where $X = \{x \in \Re^n \mid f_i(x) \leq 0, i = 1, ..., m\}$ and $f_i : \Re^n \to \Re, i = 0, 1, ..., m$, are continuously differentiable. Assume that $X \neq \emptyset$. Denote $I = \{1, ..., m\}$ and $I^0 = I \bigcup \{0\}$.

Let $x \in X$ be a given feasible point of (1.1). Denote

$$I(x) = \{i \mid i \in I, f_i(x) = 0\}.$$

Recall that x is said to be a Fritz-John point of (1.1) if there is a vector $(\tau, u) \in \Re \times \Re^m$ such that $(\tau, u) \geq 0, (\tau, u) \neq 0$ and

$$\begin{cases} \tau \nabla f_0(x) + \sum_{i=1}^m u_i \nabla f_i(x) = 0; \\ u_i f_i(x) = 0, \quad \text{for } \; i \in I; \\ f_i(x) \leq 0, \quad \text{for } \; i \in I. \end{cases} \tag{1.2}$$

We call such a pair (τ, u) an FJ multiplier of x, and denote the set of all possible FJ multipliers associated with x by $M(x)$. If $\tau \neq 0$, then such an FJ point x is also a Karush-Kuhn-Tucker (KKT) point of (1.1).

Given $x^1 \in X$, methods of feasible direction [1, 19, 24, 26, 35] for solving (1.1) construct a sequence $\{x^k\}_{k=1}^{\infty}$ such that, for all k,

$$x^k \in X \tag{1.3}$$

and

$$f_0(x^{k+1}) \leq f_0(x^k). \tag{1.4}$$

These two conditions turn out to be important in their own right in many contexts, e.g., (i) when the objective function is not well defined outside the feasible set X or (ii) in real-time applications, when it is crucial that a feasible solution is available at the next "stopping time" (see [22] for discussion of optimization problems arising from design problems).

Some methods for solving (1.1), such as SQP type algorithms, stop at a KKT point of (1.1) [1, 5, 10, 13, 14, 18, 24, 25, 26, 27, 34]. In the following example, no KKT point exists, though the problem has a unique solution, FJ point, and bounded level sets. Methods that only stop at KKT points may not work for this example.

Example 1.1 Consider the problem (P1):

$$\begin{aligned} \min \quad & f_0(x) := (x_1)^2 + x_2 + (x_3)^2 \\ \text{s.t.} \quad & f_1(x) := x_1 + 5(x_3)^2 \leq 0, \\ & f_2(x) := -x_1 - (x_2)^5 \leq 0. \end{aligned}$$

It is not hard to verify the following properties:

(a1) (P1) has a unique solution $x^* = (0, 0, 0)^T$;

(b1) (P1) has a unique FJ point $x^* = (0,0,0)^T$;

(c1) (P1) does not have a KKT point;

(d1) for any $C_0 > 0$, the level set $L(C_0) = \{x \in \Re^3 \mid f_0(x) \leq C_0, x \in X\}$ is bounded;

(e1) for each $x \in X \setminus \{x^*\}$, the linear independence constraint qualification (LICQ) holds.

The following example shows that even if (1.1) has a KKT point, the solution of (1.1) is only an FJ point.

Example 1.2 Consider the following problem (P2)

$$\begin{array}{ll} \min & f_0(x) := (x_1)^2 + x_2 \\ \text{s.t.} & f_1(x) := -x_1 - (x_2)^5 \leq 0, \\ & f_2(x) := x_1 + (x_2)^2 + (x_3)^2 \leq 0. \end{array}$$

It is not hard to verify the following properties:

(a2) (P2) has a unique solution $x^* = (0,0,0)^T$;

(b2) $x^* = (0,0,0)^T$ is an FJ point of (P2), but it is not a KKT point of (P2);

(c2) (P2) has a KKT point $\bar{x} = (-1,1,0)^T$;

(d2) for any $C_0 > 0$, the level set $L(C_0) = \{x \in \Re^3 \mid f_0(x) \leq C_0, x \in X\}$ is bounded.

From properties (a2) and (c2) of this example, even if the SQP type algorithms [1, 5, 10, 13, 14, 18, 24, 25, 26, 27, 34] work for (P2), they may not be able to find its solution.

Given an estimate $x \in X$ of the solution x^* to the problem (1.1), Topkis and Veinott [35] generated a search direction using the following linear programming

$$\begin{array}{ll} \min & v \\ \text{s.t.} & \nabla f_0(x)^T d \leq v, \\ & f_i(x) + \nabla f_i(x)^T d \leq v, \quad i \in I, \\ & -1 \leq d_j \leq 1, \ j = 1, \ldots, n. \end{array} \tag{1.5}$$

Like other feasible direction methods, once a feasible descent direction is obtained, a line search is needed for the Topkis-Veinott method. From [1] and [35], the Topkis-Veinott method works well for (P1) (also (P2)) though the SQP type methods mentioned above do not. However, the subproblem (1.5) is also a source of disadvantages associated with the Topkis-Veinott method. Due to linearity, (1.5) always produces extreme point solutions, and the directions obtained may therefore depend more on properties of the feasible set X than on properties of the objective function f_0. Consequently, the convergence of the algorithm is, in general, slow and characterized by zigzagging.

The approach suggested in this paper is based on the above two examples and a modification of the subproblem (1.5) of the Topkis-Veinott method. In order to avoid nonuniqueness of the direction, the extreme point solution in (1.5) and the slow convergence behaviour of the Topkis-Veinott method, we introduce a penalty term in the objective function of (1.5) and some weights in the constraints of (1.5). More precisely, we use the following positive semi-definite quadratic problem to generate a search direction d at an estimate $x \in X$ of the solution x^* to the problem (1.1),

$$\begin{array}{ll} \min & v + \frac{1}{2} d^T H(x) d \\ \text{s.t.} & \nabla f_0(x)^T d \leq c_0 v, \\ & f_i(x) + \nabla f_i(x)^T d \leq c_i v, \quad i \in I, \end{array} \tag{1.6}$$

where $H(x)$ is a symmetric positive definite $n \times n$ matrix. We present the algorithm in Section 2 and establish its global convergence in Section 3.

It is known that the Topkis-Veinott method is globally convergent in the sense that every accumulation point of the iteration point sequence is an FJ point of the problem without any CQ. However, two questions remain: (I) Does the entire iteration point sequence converge to an FJ point of (1.1)? (II) Is such an FJ point a strict local minimum of (1.1)? Robinson [33] proved that if a KKT point satisfies the LICQ, the second-order sufficiency conditions (SOSC) and the strict complementarity slackness (SCS), then it is an isolated KKT point. This result has been widely used to prove the convergence of the entire iteration point sequence generated by algorithms for general nonlinear programming [33, 24, 25, 26, 27]. In 1982, Robinson [34] proved that if a KKT point satisfies the Mangasarian-Fromovitz constraint qualification (MFCQ) and the strong second-order sufficiency conditions (SSOSC), then it is an isolated KKT point. From this result and the fact that every FJ point of (1.1) is a KKT point if the MFCQ holds, we can easily deduce that if the MFCQ and the SSOSC hold, then the above question (I) has a positive answer. Recently, Qi and Wei [30] proved that the Robinson conditions can be reduced to the the positive linear independent regular condition (PLIRC) and the SSOSC. For question (II), Qi [29] proved, for LC^1 optimization problem, that if a KKT point satisfies the SSOSC, then it is a strict local minimum.

In Section 4, we introduce an FJ function, an FJ1-SSOSC and an FJ2-SSOSC, and then prove, without any CQ, that (1) if an FJ point z satisfies the FJ1-SSOSC, then there exists a neighborhood $N(z)$ of z such that for any FJ point $y \in N(z) \setminus \{z\}$, $f_0(y) \neq f_0(z)$; (2) if an FJ point z satisfies the FJ2-SSOSC, then it is a strict local minimum of (1.1). These results improve (a) the conditions for the convergence of the entire iteration point sequence generated by algorithms for general nonlinear programming [33, 24, 25, 26, 27] and (b) the conditions of a strict local minimum [29].

Consider the following two-stage stochastic programs with fixed recourse [2, 3, 4, 6, 7, 16, 31, 32, 36]

$$\begin{aligned} &\min\ P(x) + \phi(x) \\ &\text{s.t. } Ax \leq b, \end{aligned}$$

where P is a twice continuously differentiable convex function from $X \to \Re$,

$$\phi(x) = \int_{\Re^n} \psi(\omega - Tx)\mu(\omega)d\omega,$$

$$\psi(\omega - Tx) = \max\{-\frac{1}{2}y^T Hy + y^T(\omega - Tx) \mid Wy \leq q\},$$

$H \in \Re^{m\times m}$ is a symmetric positive definite matrix, $c \in \Re^n, A \in \Re^{r\times n}, b \in \Re^r, T \in \Re^{m\times n}, q \in \Re^{m_1}$ and $W \in \Re^{m_1\times m}$ are fixed matrices or vectors, $\omega \in \Re^m$ is a random vector and $\mu : \Re^m \to \Re_+$ is a continuously differentiable probability density function.

Since it is impossible generally to demand the exact evaluation of the function ϕ and its gradient, we always consider approximate problems of the form

$$\begin{aligned} &\min\ P(x) + \phi_N(x) \\ &\text{s.t. } Ax \leq b, \end{aligned} \tag{1.7}$$

where

$$\phi_N(x) = \sum_{i=1}^{N} \alpha_i \psi(\omega_i - Tx)\mu(\omega_i). \tag{1.8}$$

The weights $\{\alpha_i\}_{i=1}^N$ and points $\{\omega_i\}_{i=1}^N$ are generated by a form of multidimensional numerical integration rule. For the above problem (1.7), the calculation of the value of the objective function $P(x) + \phi_N(x)$ at x^k requires solving N quadratic programs (i=1, ..., N)

$$\psi(\omega_i - Tx^k) = \max\{-\frac{1}{2}y^T Hy + y^T(\omega - Tx^k) \mid Wy \leq q\}.$$

This implies that the following line search procedure

$$f_0(x^k + \rho^j d^k) - f_0(x^k) \leq \sigma\rho^j \nabla f_0(x^k)^T d^k$$

is very costly, where

$$f_0(x) = P(x) + \phi_N(x).$$

it is, therefore, an interesting issue to avoid a line search for such problems.

In Section 5, we prove that the unit step can be accepted by the proposed algorithm if the parameters $c_i, (i \in I^0)$ are chosen large enough under the assumption that $\nabla f_i (i \in I^0)$ are globally Lipschitz on $\Re^n$. This may be useful for solving the problems when $f_i(x)$ and $\nabla f_i(x)$, $i \in I^0$, are extremely difficult to calculate at each point x.

Throughout the paper, we denote the Euclidean norm of a vector w by $\|w\|$ and supp(w)= $\{i \mid w_i \neq 0\}$.

2. The algorithm

Algorithm 1

(Initialization). Let $\sigma \in (0,1), \rho \in (0,1), c_i^1 > 0, (i \in I^0)$; choose $x^1 \in X$ and a symmetric positive positive matrix $H_1 \in \Re^{n \times n}$; set $k = 1$.

Step 1: Solve the following subproblem

$$\begin{aligned} \min \quad & v + \frac{1}{2}d^T H_k d \\ \text{s.t.} \quad & \nabla f_0(x^k)^T d \leq c_0^k v, \\ & f_i(x^k) + \nabla f_i(x^k)^T d \leq c_i^k v, \quad i \in I. \end{aligned} \tag{2.1}$$

Let $((d^k)^T, v_k)^T$ be one of its solutions.

Step 2: If $((d^k)^T, v_k)^T = 0$, x^k is a Fritz-John (FJ) point for (1.1), stop.

Step 3: Let j_k be the smallest nonnegative integer j such that

$$f_0(x^k + \rho^j d^k) - f_0(x^k) \leq \sigma\rho^j \nabla f_0(x^k)^T d^k \tag{2.2}$$

and

$$f_i(x^k + \rho^j d^k) \leq 0, \quad i \in I. \tag{2.3}$$

Set $t_k = \rho^{j_k}$.

Step 4: Generate $c_i^{k+1} > 0$, for $i \in I^0$. Compute a new symmetric positive definite approximation matrix H_{k+1} to the Hessian of the Lagrangian. Set $x^{k+1} = x^k + t_k d^k, k := k+1$ and return to Step 1.

Remark 2.1 There are many methods to solve (2.1). One may use the dual quadratic programming method of Kiwiel [17] and the interior point algorithms for convex quadratic programming of Daya and Shetty [9], Hertog, Roos and Terlaky [15], Goldfarb and Liu [12], Monteiro and Adler [20], Monteiro, Adler and Resende [21], and Ye and Tse [38]. In fact, if we let

$$P_k = (\nabla f_0(x^k), \nabla f_1(x^k), ..., \nabla f_m(x^k))^T$$

and

$$a^k = (0, -f_1(x^k), ..., -f_m(x^k))^T,$$

then the dual problem of (2.1) is

$$\begin{array}{ll} \min & \frac{1}{2} y^T (P_k^T H_k^{-1} P_k) y + (a^k)^T y \\ \text{s.t.} & \sum_{i=0}^m c_i^k y_i = 1, \\ & y \geq 0, \end{array} \tag{2.4}$$

where $y \in \Re^{m+1}$. It is clear that (2.4) is a standard-form positive semi-definite convex quadratic programming problem. Hence, we can use the interior point algorithms mentioned above to solve (2.4).

Suppose that y^k is one of the solutions of (2.4). Then

$$d^k = -H_k^{-1} P_k y^k$$

is a solution of (2.1). Notice that (2.4) is also a least-squares problem. One can also use other methods such as the methods given in [11] to solve (2.4).

The following proposition indicates that Algorithm 1 is well-defined.

Proposition 2.1 *For any given k, if $((d^k)^T, v_k)^T = 0$, then x^k is an FJ point of (1.1). If $((d^k)^T, v_k)^T \neq 0$ and $x^k \in X$, then there exists $\tau_k > 0$ such that for any $t \in [0, \tau_k]$,*

$$f_0(x^k + td^k) - f_0(x^k) \leq \sigma t \nabla f_0(x^k)^T d^k$$

and

$$f_i(x^k + td^k) \leq 0, \quad i \in I.$$

Proof: Since $((d^k)^T, v_k)^T$ is a solution of (2.1), we have the following Fritz-John conditions

(Theorem 4.2.6 of [1]),

$$\begin{cases} \bar{u}_{-1}^k H_k d^k + \sum_{i=0}^m \bar{u}_i^k \nabla f_i(x^k) = 0; \\ \sum_{i=0}^m c_i^k \bar{u}_i^k = \bar{u}_{-1}^k; \\ \bar{u}_0^k(\nabla f_0(x^k)^T d^k - c_0^k v_k) = 0; \\ \bar{u}_i^k(f_i(x^k) + \nabla f_i(x^k)^T d^k - c_i^k v_k) = 0, \quad \text{for} \quad i \in I; \\ \nabla f_0(x^k)^T d^k \leq c_0^k v_k; \\ f_i(x^k) + \nabla f_i(x^k)^T d^k \leq c_i^k v_k, \quad \text{for} \quad i \in I; \end{cases} \tag{2.5}$$

where $\bar{u}_{-1}^k \geq 0$ and $\bar{u}_i^k \geq 0$ are some multipliers with $(\bar{u}_{-1}^k, \bar{u}^k) \neq 0$. From the second equality of (2.5), $\bar{u}_{-1}^k \neq 0$. Hence, if we let

$$u_i^k = \frac{\bar{u}_i^k}{\bar{u}_{-1}^k}, \quad i \in I^0,$$

then, the following Karush-Kuhn-Tucker (KKT) conditions for (2.1) hold,

$$\begin{cases} H_k d^k + \sum_{i=0}^m u_i^k \nabla f_i(x^k) = 0; \\ \sum_{i=0}^m c_i^k u_i^k = 1; \\ u_0^k(\nabla f_0(x^k)^T d^k - c_0^k v_k) = 0; \\ u_i^k(f_i(x^k) + \nabla f_i(x^k)^T d^k - c_i^k v_k) = 0, \quad \text{for} \quad i \in I; \\ \nabla f_0(x^k)^T d^k \leq c_0^k v_k; \\ f_i(x^k) + \nabla f_i(x^k)^T d^k \leq c_i^k v_k, \quad \text{for} \quad i \in I. \end{cases} \tag{2.6}$$

If $((d^k)^T, v_k)^T = 0$, then the above KKT conditions imply that there exist some multipliers $u_i^k \geq 0$ such that

$$\begin{cases} \sum_{i=0}^m u_i^k \nabla f_i(x^k) = 0; \\ u_i^k f_i(x^k) = 0, \quad \text{for} \quad i \in I; \\ f_i(x^k) \leq 0, \quad \text{for} \quad i \in I; \end{cases}$$

therefore, x^k is an FJ point for (1.1) and the proof of the first part of this proposition is completed.

We prove now the second part of this proposition. From Step 1, since $(d^T, v) = (0, 0)$ is a feasible solution of (2.1), we have

$$\nabla f_0(x^k)^T d^k \leq c_0^k v_k \leq -\frac{c_0^k}{2}(d^k)^T H_k d^k, \tag{2.7}$$

which implies that $d^k = 0$ if and only if $v_k = 0$. If $((d^k)^T, v_k)^T \neq 0$, then $d^k \neq 0$ and

$$v_k \leq -\frac{1}{2}(d^k)^T H_k d^k < 0.$$

Therefore,

$$\nabla f_0(x^k)^T d^k \leq -\frac{c_0^k}{2}(d^k)^T H_k d^k < 0 \tag{2.8}$$

and

$$f_i(x^k) + \nabla f_i(x^k)^T d^k \leq -\frac{c_i^k}{2}(d^k)^T H_k d^k < 0, \text{ for } i \in I. \tag{2.9}$$

If we let

$$q_k(t) = f_0(x^k + td^k) - f_0(x^k) - \sigma t \nabla f_0(x^k)^T d^k,$$

then, we obtain, from (2.8), that

$$\begin{aligned} q_k(t) &= t\nabla f_0(x^k)^T d^k + o(t) - \sigma t \nabla f_0(x^k)^T d^k \\ &= (1-\sigma)t\nabla f_0(x^k)^T d^k + o(t) \\ &\leq -(\frac{(1-\sigma)c_0^k}{2}(d^k)^T H_k d^k)t + o(t), \end{aligned}$$

which implies that there exists $\tau_0^k > 0$ such that for any $t \in [0, \tau_0^k]$,

$$q_k(t) \leq 0. \tag{2.10}$$

Similarly, for $i \in I_k = \{i : f_i(x^k) = 0, i \in I\}$, from (2.9) and

$$\begin{aligned} f_i(x^k + td^k) &= f_i(x^k) + t\nabla f_i(x^k)^T d^k + o(t) \\ &= t\nabla f_i(x^k)^T d^k + o(t) \\ &\leq -(\frac{c_i^k}{2}(d^k)^T H_k d^k)t + o(t), \end{aligned}$$

we have $\tau_i^k > 0$ such that for any $t \in [0, \tau_i^k]$,

$$f_i(x^k + td^k) \leq 0. \tag{2.11}$$

For $i \in I \setminus I_k$, from the continuity of f_i and $f_i(x^k) < 0$, we have $\tau_i^k > 0$ such that for any $t \in [0, \tau_i^k]$,

$$f_i(x^k + td^k) \leq 0. \tag{2.12}$$

Let $\tau_k = \min\{\tau_i^k \mid i \in I^0\}$, the conclusion follows from (2.10), (2.11) and (2.12). Q.E.D.

Remark 2.2 It is worth noting that, for any given FJ point x of (1.1), if the MFCQ holds, i.e.,

(H) there is a vectors $z \in \Re^n$ such that

$$\nabla f_i(x)^T z < 0, \quad \text{for} \quad i \in I(x),$$

then such an FJ point x is also a KKT point of (1.1). Furthermore, if $f_i, i \in I^0$ are convex, then the KKT point is also an optimal point of (1.1). In this paper, we do not assume (H) holds. Therefore, Algorithm 1 generates only an FJ point of (2.1).

3. Global convergence

Throughout the sequel, the following conditions are assumed to hold.

(A1) For $i \in I^0$,

$$0 < \liminf_{k\to\infty} c_i^k \leq \limsup_{k\to\infty} c_i^k < +\infty;$$

(A2) the Hessian estimates $\{H_k\}_{k=0}^{\infty}$ are bounded, i.e., there exists a scalar $C_1 > 0$ such that for all k,

$$\|H_k\| \leq C_1; \tag{3.1}$$

(A3) there exists a scalar $C_2 > 0$ such that, for all k, the Hessian estimates satisfy

$$d^T H_k d \geq C_2 \|d\|^2, \quad \text{for any} \quad d \in \Re^n. \tag{3.2}$$

Theorem 3.1 *Assume that (A1)-(A3) hold. Then Algorithm 1 described in Section 2 either stops at an FJ point or generates a sequence $\{x^k\}_{k=1}^{\infty}$ for which each accumulation point is an FJ point of (1.1).*

Proof: The first statement is obvious by using Proposition 2.1. Thus, suppose that there exists an infinite index set $\mathcal{K}$ such that $\{x^k\}_{k\in\mathcal{K}} \to x^*$. From (2.7), we have for $k \in \mathcal{K}$,

$$-\|\nabla f_0(x^k)\|\|d^k\| \leq c_0^k v_k \leq -\frac{1}{2} C_2 c_0^k \|d^k\|^2,$$

which implies that for $k \in \mathcal{K}$,

$$\|d^k\| \leq \frac{2}{C_2 \liminf_{k\in\mathcal{K}} c_0^k} \|\nabla f_0(x^k)\| \tag{3.3}$$

and

$$|v_k| \leq \frac{2}{C_2 (\liminf_{k\in\mathcal{K}} c_0^k)^2} \|\nabla f_0(x^k)\|^2. \tag{3.4}$$

The second equality of the KKT conditions for (2.1), (A1), (3.3) and (3.4) imply that we may assume, without loss of generality, that

$$\lim_{k\in\mathcal{K}, k\to\infty} d^k = d^*; \qquad \lim_{k\in\mathcal{K}, k\to\infty} v_k = v_*;$$

$$\lim_{k\in\mathcal{K}, k\to\infty} u_i^k = u_i^*; \qquad \lim_{k\in\mathcal{K}, k\to\infty} c_i^k = c_i^*, \quad \text{for} \quad i \in I^0.$$

It is clear that $d^* = 0$ if and only if $v_* = 0$. We assume by contradiction that x^* is not an FJ point of (1.1). From the KKT conditions for (2.1), we can deduce that $d^* \neq 0$. On the other hand, from (2.8) and (2.9), we have

$$\nabla f_0(x^*)^T d^* \leq -\frac{C_2 c_0^*}{2} \|d^*\|^2 < 0$$

and

$$f_i(x^*) + \nabla f_i(x^*)^T d^* \leq -\frac{C_2 c_i^*}{2} \|d^*\|^2 < 0, \quad \text{for} \quad i \in I^0.$$

Therefore, there exists $\delta > 0$ such that for all $k \in \mathcal{K}$, k large enough,

$$\nabla f_0(x^k)^T d^k \leq -\delta, \tag{3.5}$$

$$\nabla f_i(x^k)^T d^k \leq -\delta, \quad \text{for} \quad i \in I(x^*) \tag{3.6}$$

and

$$f_i(x^k) \leq -\delta, \quad \text{for} \quad i \in I \setminus I(x^*). \tag{3.7}$$

Similar to the proof of Proposition 2.1, using (3.5), (3.6) and (3.7), we can prove that there exists $\tau_* > 0$, such that for $k \in \mathcal{K}$, k large enough,

$$t_k \geq \tau_* > 0. \tag{3.8}$$

From (2.2), (3.5) and (3.8), we have, for $k \in \mathcal{K}$, k large enough,

$$f_0(x^{k+1}) - f_0(x^k) \leq -\sigma\delta\tau_*,$$

which contradicts the fact that $f_0(x^{k+1}) - f_0(x^k) \to 0$ as $k \to \infty$. The proof is completed. Q.E.D.

The following proposition shows that the existence of an accumulation point in the sequence generated by Algorithm 1 induces some regularity properties on this sequence.

Proposition 3.1 *Assume that (A1)-(A3) hold. Suppose that the sequence $\{x^k\}_{k=1}^{\infty}$ generated by Algorithm 1 has an accumulation point. Then*

$$\lim_{k\to\infty} \|x^{k+1} - x^k\| = 0. \tag{3.9}$$

Proof: Since $f_0(x^k)$ is monotonically decreasing, existence of an accumulation point of $\{x^k\}_{k=1}^{\infty}$ and continuity of f_0 imply that the sequence $\{f_0(x^k)\}_{k=1}^{\infty}$ is convergent. From

$$f_0(x^k + t_k d^k) - f_0(x^k) \leq \sigma t_k \nabla f_0(x^k)^T d^k,$$

$$\nabla f_0(x^k)^T d^k \leq -\frac{C_2 c_0^k}{2} \|d^k\|^2$$

and (A1), we have

$$\lim_{k\to\infty} t_k \|d^k\|^2 = 0,$$

which implies that

$$\lim_{k\to\infty} t_k\|d^k\| = 0. \tag{3.10}$$

Since

$$\|x^{k+1} - x^k\| \le t_k\|d^k\|,$$

the claim follows (3.10). Q.E.D.

Proposition 3.2 *Let Z denote the set of all possible accumulation points of $\{x^k\}_{k=1}^{\infty}$. If $Z \neq \emptyset$, then either Z is a singleton or Z is a connected set.*

Proof: The claim follows (3.9) and Remark 14.1.1 in [23]. Q.E.D.

4. Local structure of FJ points and the convergence of the entire iteration point sequence

In order to study the convergence of the sequence $\{x^k\}_{k=1}^{\infty}$, it suffices to show that Z is a singleton. In doing so, we need some strong regularity assumptions on the functions involved in problem (1.1).

For any $x \in \Re^n$ and any $(\tau, u) \in \Re \times \Re^m$, denote by $FJ(x, \tau, u)$ the FJ function

$$FJ(x,\tau,u) = \tau f_0(x) + \sum_{i=1}^{m} u_i f_i(x).$$

Let

$$F_{\tau,u}(x) \equiv \nabla_x FJ(x,\tau,u) = \tau\nabla f_0(x) + \sum_{i=1}^{m} u_i \nabla f_i(x).$$

In what follows, in addition to (A1)-(A3), we assume that the following condition holds:

(A4) For each $i \in I^0$, ∇f_i is locally Lipschitzian, i.e., f_i is an LC^1 function.

In general, assume that $F : \Re^n \to \Re^m$ is locally Lipschitzian. By Rademacher's Theorem, F is differentiable almost everywhere. Let D_F be the set where F is differentiable. Then the generalized Jacobian of F at x in the sense of Clarke [8] is

$$\partial F(x) = co\{\lim_{\substack{x^k\in D_F\\ x^k\to x}} \nabla F(x^k)\}. \tag{4.1}$$

In [28], the author introduced

$$\partial_B F(x) = \{\lim_{\substack{x^k\in D_F\\ x^k\to x}} \nabla F(x^k)\}.$$

F is said to be semismooth at $x \in \Re^n$ if F is Lipschitz continuous in an open neighbourhood of x and the limit

$$\lim_{\substack{D\in\partial F(x+td')\\ d'\to d, t\to 0^+}} Dd$$

exists for any vector $d \in \Re^n$.

Definition 4.1 Let z be an FJ point of (1.1). We say that the point z satisfies the *FJ1 strong second-order sufficiency conditions* (FJ1-SSOSC) if for any FJ multiplier $(\tau, u) \in M(z)$, all $V \in \partial F_{\tau,u}(z)$ are positive definite on the following set

$$G_1(z,\tau,u) = \quad \{d \in \Re^n \mid \nabla f_0(z)^T d = 0, \;\; \nabla f_i(z)^T d = 0 \;\text{ for }\; i \in supp(u),$$

$$\nabla f_i(z)^T d \leq 0, \;\text{ for }\; i \in (I(z) \setminus supp(u))\}.$$

Definition 4.2 Let z be an FJ point of (1.1). We say that the point z satisfies the *FJ2 strong second-order sufficiency conditions* (FJ2-SSOSC) if for any FJ multiplier $(\tau, u) \in M(z)$, all $V \in \partial F_{\tau,u}(z)$ are positive definite on the following set

$$G_2(z,\tau,u) = \quad \{d \in \Re^n \mid \nabla f_0(z)^T d \leq 0, \;\; \tau\nabla f_0(z)^T d = 0, \; \nabla f_i(z)^T d = 0$$

$$\text{for } \;\; i \in supp(u), \;\; \nabla f_i(z)^T d \leq 0, \;\text{ for }\; i \in (I(z) \setminus supp(u))\}.$$

It is clear that if an FJ point z satisfies the FJ2-SSOSC, then z satisfies the FJ1-SSOSC. If $\tau \neq 0$, then $G_1 = G_2$, the FJ1-SSOSC and the FJ2-SSOSC become the SSOSC introduced by Robinson [34].

The following second-order mean value theorem for LC^1 functions can be found in [29].

Lemma 4.1 *Suppose that $f : W \to \Re$ is an LC^1 function on W, where W is an open subset of $\Re^n$. Let $P = \nabla f$. Then for any $x, y \in W$, there exists $t \in [0,1]$ and $V \in \partial P(x + t(x-y))$ such that*

$$f(y) - f(x) - \nabla f(x)^T (y-x) = \frac{1}{2}(y-x)^T V (y-x).$$

Theorem 4.1 *Suppose that z is an FJ point of (1.1). If z satisfies the FJ1-SSOSC, then there exists a neighborhood $N(z)$ of z, such that for any FJ point $y \in N(z) \setminus \{z\}$,*

$$f_0(z) \neq f_0(y). \tag{4.2}$$

Proof. Suppose, by contradiction, that such a neighborhood $N(z)$ of z does not exist. Then there exists an FJ point sequence $\{z^k\}_{k=1}^{\infty}$ such that $z^k \neq z$,

$$\lim_{k\to\infty} z^k = z \tag{4.3}$$

and

$$f_0(z^k) = f_0(z). \tag{4.4}$$

For any given k, since z^k is an FJ point of (1.1), we have $(\bar{\tau}_k, \bar{u}^k) \in M(z^k)$. Let

$$(\tau_k, u^k) = \frac{(\bar{\tau}_k, \bar{u}^k)}{\|(\bar{\tau}_k, \bar{u}^k)\|}.$$

Then $(\tau_k, u^k) \in M(z^k)$ and $\|(\tau_k, u^k)\| = 1$. We may assume, without loss of generality, that

$$\lim_{k\to\infty} (\tau_k, u^k) = (\tau, u). \tag{4.5}$$

From the FJ conditions at z^k and $(\tau_k, u^k) \in M(z^k)$, we have, for all k,

$$\begin{cases} \tau_k \nabla f_0(z^k) + \sum_{i\in I} u_i^k \nabla f_i(z^k) = 0; \\ u_i^k f_i(z^k) = 0, \quad \text{for} \ \ i \in I; \\ f_i(z^k) \le 0, \quad \text{for} \ \ i \in I. \end{cases}$$

which combining with (4.3) and (4.5) implies

$$\begin{cases} \tau \nabla f_0(z) + \sum_{i\in I} u_i \nabla f_i(z) = 0; \\ u_i f_i(z) = 0, \quad \text{for} \ \ i \in I; \\ f_i(z) \le 0, \quad \text{for} \ \ i \in I. \end{cases}$$

Hence, $(\tau, u) \in M(z)$ and

$$F_{\tau,u}(z) = 0. \tag{4.6}$$

Since z^k is an FJ point of (1.1), we have $f_i(z^k) \le 0$ for $i \in I$, which combining with (4.4) and the facts that $u_i f_i(z) = 0$ and $(\tau, u) \ge 0$, implies, for all k,

$$FJ(z,\tau,u) = \tau f_0(z) + \sum_{i\in I} u_i f_i(z) \ge \tau f_0(z^k) + \sum_{i \in I} u_i f_i(z^k) = FJ(z^k, \tau, u). \tag{4.7}$$

By Lemma 4.1, we have

$$FJ(z^k,\tau,u) = FJ(z,\tau,u) + F_{\tau,u}(z)(z^k - z) + \frac{1}{2}(z^k - z)^T V_l (z^k - z), \tag{4.8}$$

where $V_k \in \partial F_{\tau,u}(z + t_k(z^k - z)), 0 \le t_l \le 1$. By (4.6), (4.7) and (4.8), we have for all k,

$$(z^k - z)V_k(z^k - z) \le 0. \tag{4.9}$$

By (4.1) and the Carathéodory Theorem, V_k can be approximated by a convex combination of $\nabla F_{\tau,u}(z_j^k)$, where $z_j^k \in D_{F_{\tau,u}}$ are close to $z + t_k(z^k - z)$ as we wish, $j = 0, ..., n$. Thus, we may choose $z_j^k \in D_{F_{\tau,u}}, z_j^k \to z$ as $l \to \infty$ for $j = 0, ..., n$ such that

$$\|V^k - \sum_{j=0}^{n} \lambda_j^k \nabla F_{\tau,u}(z_j^k)\| \le \frac{1}{k},$$

where $\sum_{j=0}^{n} \lambda_j^k = 1, \lambda_j^k \ge 0$ for all k and $j = 0, 1, ..., n$. By (4.9), we have

$$\sum_{j=0}^{n} \lambda_j^k \left(\frac{z^k - z}{\|z^k - z\|}\right) \nabla F_{\tau,u}(z_j^k) \left(\frac{z^k - z}{\|z^k - z\|}\right) \le O\left(\frac{1}{k}\right). \tag{4.10}$$

Since $F_{\tau,u}$ is locally Lipschitzian, $\nabla F_{\tau,u}$ is locally bounded in $D_{F_{\tau,u}}$. Without loss of generality, by passing to a subsequence, we may assume that $F_{\tau,u}(z_j^k) \to V^j$ for $V^j \in \partial_B F_{\tau,u}(z)$ and

$\lambda_j^k \to \lambda_j$ as $k \to \infty$. Then $\sum_{j=0}^n \lambda_j = 1, \lambda_j \geq 0$. Since $\|\frac{z^k-z}{\|z^k-z\|}\| = 1$, we may also assume that $d = \lim_{k\to\infty} \frac{z^k-z}{\|z^k-z\|}$. With this consideration, letting $k \to \infty$ in (4.10), we have

$$\sum_{j=1}^n d^T V^j d \leq 0. \tag{4.11}$$

On the other hand, for any $i \in supp(u)$, since $u_i^k \to u_i$, there is k_i such that for all $k > k_i$, $u_i^k > 0$. Hence, for all $k > l_i$, $i \in supp(u^k)$. Letting $\bar{k} = max\{k_i \mid i \in supp(u)\}$, we have, for all $k > \bar{k}$, $supp(u) \subseteq supp(u^k)$, and hence, for $k > \bar{k}$, that

$$f_i(z^k) - f_i(z) = 0, \text{ for } i \in \ supp(u). \tag{4.12}$$

By Taylor's theorem, we have

$$0 = f_i(z^k) - f_i(z) = \nabla f_i(z)^T(z^k - z) + o(\|z^k - z\|).$$

Dividing the above equality by $\|z^k - z\|$ and letting $k \to \infty$, we have

$$\nabla f_i(z)^T d = 0, \text{ for } i \ \in supp(u). \tag{4.13}$$

Similarly, from (4.4), we have

$$\nabla f_0(z)^T d = 0. \tag{4.14}$$

For $i \in I(z), f_i(z) = 0$. Thus, we have, from

$$\begin{aligned} 0 \geq f_i(z^k) &= f_i(z) + \nabla f_i(z)^T(z^k - z) + o(\|z^k - z\|) \\ &= \nabla f_i(z)^T(z^k - z) + o(\|z^k - z\|), \end{aligned}$$

i.e.,

$$\nabla f_i(z)^T(z^k - z) + o(\|z^k - z\|) \leq 0$$

that, for $i \in I(z)$

$$\nabla f_i(z)^T d \leq 0. \tag{4.15}$$

The relations, (4.11), (4.13), (4.14) and (4.15), contradict the assumption that z satisfies the FJ1-SSOSC. Hence the claim holds. Q.E.D.

Corollary 4.1 *Assume that $\{x^k\}_k^\infty$ is generated by Algorithm 1 and $Z \neq \emptyset$, $x^* \in Z$. If x^* satisfies the FJ1-SSOSC, then Z is a singleton and*

$$\lim_{k\to\infty} x^k = x^*. \tag{4.16}$$

Proof. Suppose that Z is not a singleton. From Proposition 3.2, we have $y^l \in Z$ such that $y^l \neq x^*$ and $\lim_{l\to\infty} y^l = x^*$. Since for all k, $f_0(x^{k+1}) < f_0(x^k)$ and there is a set $\mathcal{K}$ such that $\lim_{k\in\mathcal{K}} x^k = x^*$, we have $f_0(x^k) \to f_0(x^*)$, which implies that for all l,

$$f_0(y^l) = f_0(x^*),$$

which contradicts Theorem 4.1. Hence Z is a singleton and (4.16) holds. Q.E.D.

We now show that the FJ2-SSOSC implies a strict local minimum of (1.1).

Theorem 4.2 *Assume that an FJ point z of (1.1) satisfies the FJ2-SSOSC. Then z is a strict local minimum of (1.1).*

Proof: Suppose not. Then there is a sequence $\{z^k\}_{k=1}^{\infty}$ such that $z^k \to z, z^k \neq z, f_i(z^k) \leq 0$ for $i \in I$ and

$$f_0(z^k) - f_0(z) \leq 0. \tag{4.17}$$

Without loss of generality, we may assume that

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

Let $(\tau, u) \in M(z)$. Then

$$F_{\tau,u}(z) = 0.$$

Similar to the discussions of (4.6)- (4.11), we have

$$\sum_{j=1}^{n} d^T V^j d \leq 0. \tag{4.18}$$

Here V^j has the same meaning as in the proof of Theorem 3.2.

For $i \in I$,

$$0 \geq f_i(z^k) = f_i(z) + \nabla f_i(z)^T (z^k - z) + o(\|z^k - z\|).$$

For $i \in I(z), f_i(z) = 0$. Thus, we have

$$\nabla f_i(z)^T (z^k - z) + o(\|z^k - z\|) \leq 0.$$

Dividing the above expression by $\|z^k - z\|$ and let $k \to \infty$, we have

$$\nabla f_i(z)^T d \leq 0, \tag{4.19}$$

for $i \in I(z)$.

From (4.17), we have

$$\nabla f_0(z)^T d \leq 0. \tag{4.20}$$

On the other hand, from the FJ conditions. we have

$$\tau \nabla f_0(z) + \sum_{i \in supp(u)} u_i \nabla f_i(z) = 0,$$

which combining with (4.19). (4.20) and the fact that $supp(u) \subseteq I(z)$, yields that

$$\nabla f_i(z)^T d = 0. \quad \text{for } i \in supp(u), \tag{4.21}$$

and

$$\tau \nabla f_0(z)^T d = 0. \tag{4.22}$$

(4.18)-(4.22) contradict the assumption that z satisfies the FJ2-SSOSC. Hence, the claim holds. Q.E.D.

By using Theorem 3.1, Corollary 4.1 and Theorem 4.2, we have the following result.

Corollary 4.2 *Suppose that (A1)-(A4) hold, $\{x^k\}_{k=1}^{\infty}$ is generated by Algorithm 1. If $\{x^k\}_{k=1}^{\infty}$ has an accumulation point x^* satisfying the FJ2-SSOSC, then x^* is a strict local minimum of (1.1) and $x^k \to x^*$.*

Example 4.1 Let $\mathcal{N} = \{1, 2, ...\}$ and $p \in \mathcal{N}$. Consider the following example (P3) with parameter p,

$$\begin{aligned} \min \quad & f_0(x) := (x_1)^2 + x_2 \\ \text{s.t.} \quad & f_1(x) := -x_1 - (x_2)^5 \le 0, \\ & f_2(x) := x_1 + (x_2)^{2p} + (x_3)^2 \le 0. \end{aligned}$$

It is not hard to verify the following properties.

(a3) (P3) has a unique solution $x^* = (0,0,0)^T$;

(b3) $x^* = (0,0,0)^T$ is an FJ point of (P3), but it is not a KKT point of (P3);

(c3) (P3) has a KKT point $\bar{x} = (-1,1,0)^T$ if $p \in \{1,2\}$

(d3) for any $c_0 > 0$, the level set $L(c_0) = \{x \in \Re^3 \mid f_0(x) \le c_0, x \in X\}$ is bounded.

Simple calculations yield that

$$\nabla f_0(x^*) = (0,1,0)^T, \quad \nabla f_1(x^*) = (-1,0,0)^T, \quad \nabla f_2(x^*) = (1,0,0)^T,$$

$$M(x^*) = \{(0,a,a)^T \mid a > 0\},$$

$$G_1(x^*, \tau_*, u^*) = \{(0,0,d_3)^T \mid d_3 \in \Re\}$$

and

$$G_2(x^*, \tau_*, u^*) = \{(0, d_2, d_3)^T \mid d_2 \le 0, \}.$$

For any $(\tau_*, (u^*)^T)^T := (0, a, a)^T \in M(x^*)$,

$$\nabla F_{\tau^*, u^*}(x^*) = 2a \begin{pmatrix} 0 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}, \quad \text{if } p = 1$$

and

$$\nabla F_{\tau^*, u^*}(x^*) = 2a \begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 1 \end{pmatrix}, \quad \text{if } p \in \mathcal{N} \setminus \{1\}.$$

Hence, if $p = 1$, then $\nabla F_{\tau^*, u^*}(x^*)$ is positive definite in $G_1(x^*, \tau_*, u^*)$ and $G_2(x^*, \tau_*, u^*)$ for all FJ multipliers $(\tau_*, (u^*)^T)^T \in M(x^*)$. Therefore, x^* satisfies the FJ1-SSOSC and the FJ2-SSOSC. Similarly, if $p \in \mathcal{N} \setminus \{1\}$, then x^* satisfies the FJ1-SSOSC, but not the FJ2-SSOSC.

We consider $\bar{x}$ now. Keep in mind $p \in \{1, 2\}$ in this case. Simple calculations yield that

$$\nabla f_0(\bar{x}) = (-2, 1, 0)^T, \quad \nabla f_1(\bar{x}) = (-1, -5, 0)^T, \quad \nabla f_2(x^*) = (1, 2p, 0)^T,$$

$$M(\bar{x}) = \{(\frac{5-2p}{11}a, \frac{4p+1}{11}a, , a)^T \mid a > 0\}$$

and

$$G_1(\bar{x}, \bar{\tau}, \bar{u}) = \{(0,0,d_3)^T \mid d_3 \in \Re\}.$$

Since for any $(\bar{\tau}, (\bar{u})^T)^T \in M(\bar{x})$, $\bar{\tau} \neq 0$, we have

$$G_2(\bar{x}, \bar{\tau}, \bar{u}) = G_1(\bar{x}, \bar{\tau}, \bar{u}).$$

Since

$$\nabla^2 f_0(\bar{x}) = \begin{pmatrix} 2 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix}$$

$$\nabla^2 f_1(\bar{x}) = \begin{pmatrix} 0 & 0 & 0 \\ 0 & -10 & 0 \\ 0 & 0 & 0 \end{pmatrix}$$

and

$$\nabla^2 f_2(\bar{x}) = \begin{cases} \begin{pmatrix} 0 & 0 & 0 \\ 0 & 2 & 0 \\ 0 & 0 & 2 \end{pmatrix} & \text{if } \ p = 1, \\ \\ \begin{pmatrix} 0 & 0 & 0 \\ 0 & 12 & 0 \\ 0 & 0 & 2 \end{pmatrix} & \text{if } \ p = 2, \end{cases}$$

we have, for any $(\bar{\tau}, (\bar{u})^T)^T \in M(\bar{x})$, that

$$\nabla F_{\bar{\tau},\bar{u}}(\bar{x}) = \begin{cases} a \begin{pmatrix} \frac{6}{11} & 0 & 0 \\ 0 & -\frac{28}{11} & 0 \\ 0 & 0 & 2 \end{pmatrix} & \text{if } \ p = 1, \\ \\ a \begin{pmatrix} \frac{2}{11} & 0 & 0 \\ 0 & \frac{42}{11} & 0 \\ 0 & 0 & 2 \end{pmatrix} & \text{if } \ p = 2. \end{cases}$$

Hence, if $p \in \{1, 2\}$, then $\bar{x}$ satisfies the FJ1-SSOSC and the FJ2-SSOSC.

From the above discussions, we can see that, if we solve (P3) by Algorithm 1, then Algorithm 1 either stops at an FJ point or generates a sequence $\{x^k\}_{k=1}^{\infty}$ such that the sequence converges to an FJ point of (P3).

5. Unit steps

In Section 3, we have established global convergence for Algorithm 1. In this section, we will show that if we choose the parameters c_i^k large enough, then a unit step will always be acceptable by (2.2) and (2.3). We also obtain some new results for (1.1).

In order to study the unit step, we need the following blanket assumption.

(A5) For each $i \in I^0$, there exists $L_i > 0$, such that for any $x \in \Re^n$ and any $y \in \Re^n$,

$$\|\nabla f_i(x) - \nabla f_i(y)\| \leq L_i \|x - y\|.$$

Theorem 5.1 *Suppose that (A1)-(A5) hold and* $\{x^k\}_{k=1}^{\infty}$ *is generated by Algorithm 1. If*

$$\liminf_{k\to\infty} c_0^k \geq \frac{2L_0}{C_2(1-\sigma)} \tag{5.1}$$

and

$$\liminf_{k\to\infty} c_i^k \geq \frac{2L_i}{C_2}, \quad \text{for } \ i \in I, \tag{5.2}$$

then for $k \geq 1$,

$$f_0(x^k + d^k) - f_0(x^k) \leq \sigma \nabla f_0(x^k)^T d^k, \tag{5.3}$$

and

$$f_i(x^k + d^k) \leq 0, \quad \text{for } \ i \in I. \tag{5.4}$$

Proof: We first prove (5.3), i.e., $q_k(1) \leq 0$. For $k \geq 1$, from (2.8) and (A5), we have

$$\begin{aligned} q_k(1) &= \int_0^1 \nabla f_0(x^k + td^k)^T d^k dt - \sigma \nabla f_0(x^k)^T d^k \\ &= (1-\sigma)\nabla f_0(x^k)^T d^k + \int_0^1 (\nabla f_0(x^k + td^k) - \nabla f_0(x^k))^T d^k dt \\ &\leq -\frac{(1-\sigma)}{2} c_0^k (d^k)^T H_k d^k + L_0 \|d^k\|^2 \\ &\leq -\frac{(1-\sigma)C_2 c_0^k}{2} \|d^k\|^2 + L_0 \|d^k\|^2, \end{aligned}$$

which implies that if $c_0^k \geq \dfrac{2L_0}{C_2(1-\sigma)}$, $q_k(1) \leq 0$. Therefore, (5.3) follows (5.1).

We prove (5.4) now. For $k \geq 1$, from (2.9) and (A3) we have for each $i \in I$,

$$\begin{aligned} f_i(x^k + d^k) &= (f_i(x^k) + \nabla f_i(x^k)^T d^k) + \int_0^1 (\nabla f_i(x^k + td^k) - \nabla f_i(x^k))^T d^k dt \\ &\leq c_i^k v_k + L_i \|d^k\|^2 \\ &\leq -\frac{c_i^k}{2} (d^k)^T H_k d^k + L_i \|d^k\|^2 \\ &\leq -\frac{C_2 c_i^k}{2} \|d^k\|^2 + L_i \|d^k\|^2, \end{aligned}$$

which implies that if $c_i^k \geq 2L_i$, then $f_i(x^k + d^k) \leq 0$. Hence (5.4) follows (5.2). Q.E.D.

Let X^0 denote the interior point set of X and X_{FJ} denote the FJ point set of (1.1). From the proof of Theorem 5.1, we have the following result.

Proposition 5.1 *Suppose that* $f_i, i \in I$, *satisfy (A5). If* $X \neq X_{FJ}$, *then* X^0 *is nonempty.*

Proof: Let $z \in X - X_{FJ}$. For each $i \in I$, choose c_i satisfying the following relation

$$c_i \geq 2(L_i + 1). \tag{5.5}$$

Consider the following minimization problem

$$\begin{aligned} \min \ \ & v + \frac{1}{2} d^T d, \\ \text{s.t.} \ \ & \nabla f_0(z)^T d \leq v, \\ & f_i(z) + \nabla f_i(z)^T d \leq c_i v, \quad i \in I. \end{aligned}$$

Let d^z be one of the solution. Since z is not an FJ point for (1.1), then $d^z \neq 0$ by Theorem 3.1. Similar to the proof of Theorem 5.1, we have, for $i \in I$, that

$$f_i(z + d^z) \le -\frac{c_i}{2}\|d^z\|^2 + L_i\|d^z\|^2$$

which implies that $f_i(z + d^z) < 0$ by using $d^z \neq 0$ and (5.5), i.e., $X^0 \neq \emptyset$. Q.E.D.

By Proposition 5.1, we have the following corollary.

Corollary 5.1 *Suppose that $f_i, i \in I$, satisfy (A5). If $X^0 = \emptyset$, then $X = X_{FJ}$.*

By Theorem 5.1, we can present a variant of Algorithm 1 which does not employ a line search. The variant (Algorithm 2) is also globally convergent if the assumptions (A1)-(A5) hold.

Algorithm 2 Initialization, Step 1 and Step 2 are the same as those given in Algorithm 1.

Step 3: If

$$f_0(x^k + d^k) - f_0(x^k) \le \sigma \nabla f_0(x^k)^T d^k \tag{5.6}$$

and

$$f_i(x^k + d^k) \le 0, \quad i \in I, \tag{5.7}$$

go to Step 4. Otherwise, set

$$c_0^k := \begin{cases} c_0^k & \text{if (5.6) is satisfied,} \\ 2c_0^k & \text{if (5.6) is not satisfied} \end{cases}$$

and for $i \in I$,

$$c_i^k := \begin{cases} c_i^k & \text{if (5.7) is satisfied,} \\ 2c_i^k & \text{if (5.7) is not satisfied.} \end{cases}$$

Go to Step 1.

Step 4. Set $x^{k+1} = x^k + d^k, c_i^{k+1} = c_i^k, (i \in I^0)$ and $k := k + 1$, return to Step 1.

Remark 5.1 It is worth pointing out that for all $k \ge 1$, $t_k = 1$ does not imply that $\{x^k\}_{k=1}^{\infty}$ superlinearly converges to an FJ point of (1.1). In fact, d^k in general is not a Newton direction. Therefore, the proposed algorithm does not possess a superlinear convergence rate. For the feasible direction methods with superlinear convergence rate, we refer readers to [24, 26, 30].

References

[1] M. S. Bazaraa and C. M. Shetty, *Nonlinear Programming - Theory and Algorithms*, Wiley, New York, 1979.

[2] J. R. Birge and L. Qi, Subdifferentials in approximation for stochastic programming, *SIAM Journal on Optimization* 5 (1995) 436-453.

[3] J. R. Birge and L. Qi, Continuous approximation schemes for stochastic programs, *Annals of Operations Research* 56 (1995) 251-285.

[4] J. R. Birge and R. J -B. Wets, Designing approximation schemes for stochastic optimization problems, in particular, for stochastic programs with recourse, *Mathematical Programming Study* 27 (1986) 54-102.

[5] J. V. Burke and S. P. Han, A robust sequential quadratic programming method, *Mathematical Programming* 43 (1989) 277-303.

[6] X. Chen, L. Qi, and R. Womersley, Newton's method for quadratic stochastic programs with recourse, *Journal of Computational and Applied Mathematics* 60 (1995) 29-46.

[7] X. Chen and R. S. Womersley, A parallel inexact Newton method for stochastic programs with recourse, *Annals of Operations Research* 64 (1996) 113-141.

[8] F. H. Clarke, *Optimization and Nonsmooth Analysis*, Wiley, New York, 1983.

[9] M. B. Daya and C. M. Shetty, Polynomial barrier function algorithms for convex quadratic programming, *Arabian Journal for Science and Engineering* 15 (1990) 657-670.

[10] M. Fukushima, A successive quadratic programming algorithm with global and superlinear convergence properties, *Mathematical Programming* 35 (1986) 253-264.

[11] P. E. Gill, S. J. Hammarling, W. Murray, M. A. Saunders and M. H. Wright, User's guide for LSSOL (version 1.0): A Fortran package for constrained least-squares and convex quadratic programming, Report SOL 86-1, Department of Operations Research, Standford University, 1986.

[12] D. Goldfarb and S. Liu, An $O(n^3L)$ primal interior point algorithm for convex quadratic programming, *Mathematical Programming* 49 (1991) 325-340.

[13] S. P. Han, A globally convergent method for nonlinear programming, *Journal of Optimization Theory and Applications* 22 (1977) 297-309.

[14] J. Herskovits, A two-stage feasible directions algorithm for nonlinear constrained optimization, *Mathematical Programming* 36 (1986) 19-38.

[15] D. D. Hertog, C. Roos and T. Terlaky, A polynomial method of weighted centers for convex quadratic programming, *Journal of Information & Optimization Sciences* 12 (1991) 187-205.

[16] P. Kall and S. W. Wallace, *Stochastic Programming*, Wiley, New York, 1994.

[17] K. C. Kiwiel, A dual method for solving certain positive semi-definite quadratic programming problems, *SIAM Journal on Scientific and Statistical Computing* 10 (1989) 175-186.

[18] D. Q. Mayne and E. Polak, A superlinearly convergent algorithm for constrained optimization problems, *Mathematical Programming Study* 16 (1982) 45-61.

[19] A. Migdalas, A regularization of the Frank-Wolf method and unification of certain nonlinear programming methods, *Mathematical Programming* 65 (1994) 331-345.

[20] R. D. C. Monteiro and I. Adler, Interior path following primal-dual algorithms, part II: convex quadratic programming, *Mathematical Programming* 44 (1989) 43-66.

[21] R. D. C. Monteiro, I. Adler and M. G. C. Resende, A polynomial-time primal-dual affine scaling algorithm for linear and convex quadratic programming and its power series extension, *Mathematics of Operations Research* 15 (1990) 191-214.

[22] W. T. Nye and A. L. Tits, An application-oriented, optimization-based methodology for interactive design of engineering systems, *Internat. J. Control* 43 (1986) 1693-1721.

[23] J. M. Ortega and W. C. Rheinboldt, *Iterative Solution of Nonlinear Equations in Several Variables*, Academic Press, New York, 1970.

[24] E. R. Panier and A. L. Tits, A superlinearly convergent feasible method for the solution of inequality constrained optimization problems, *SIAM Journal on Control and Optimization* 4 (1987) 934-950.

[25] E. R. Panier and A. L. Tits, Avoiding the Maratos effect by means of a nonmonotone line search I. general constrained problems, *SIAM Journal on Numerical Analysis* 4 (1991) 1183-1195.

[26] E. R. Panier and A. L. Tits, On combining feasibility, descent and superlinear convergence in inequality constrained optimization, *Mathematical Programming* 59 (1993) 261-276.

[27] E. R. Panier, A. L. Tits and J. N. Herskovits, A QP-free globally convergent, locally superlinearly convergent algorithm for inequality constrained optimization, *SIAM Journal on Control and Optimization* 4 (1988) 788-811.

[28] L. Qi, Convergence analysis of some algorithms for solving nonsmooth equations, *Mathematics of Operations Research* 18 (1993) 227-244.

[29] L. Qi, Superlinearly convergent approximate Newton methods for LC^1 optimization problems, *Mathematical Programming* 64 (1994) 277-294.

[30] L. Qi and Z. Wei, Constant positive linear independence, KKT points and convergence of feasible SQP methods, AMR 97/4, Applied Mathematics Report, University of New South Wales, 1997.

[31] R. T. Rockafellar and R. J. -B. Wets, A Lagrangian finite-generation techniques for solving linear-quadratic problems in stochastic programming, *Mathematical Programming Study* 28 (1986) 63-93.

[32] R. T. Rockafellar and R. J. -B. Wets, Linear-quadratic problems with stochastic penalties: the finite generation algorithm, in: V. I. Arkin, A. Shiraev and R. J. -B. Wets, eds., *Stochastic Optimization* Lecture Notes in Control and Information Sciences 81, Springer-Verlag, Berlin, 1987, pp. 55-69.

[33] S. M. Robinson, A quadratically convergent algorithm for general nonlinear-programming algorithms, *Mathematical Programming* 3 (1972), 145-156.

[34] S. M. Robinson, Generalized equations and their solutions, part II: Applications to nonlinear programming, *Mathematical Programming Study* 19 (1982) 200-221.

[35] D. M. Topkis and A. F. Veinott, On the convergence of some feasible direction algorithms for nonlinear programming, *Journal on SIAM Control* 5 (1967) 268-279.

[36] J. Wang, Approximate nonlinear programming algorithms for solving stochastic programs with recourse, *Annals of Operations Research* 30 (1991) 371-384.

[37] R. J. -B. Wets, Stochastic programming : Solution techniques and approximation schemes, in: A. Bachem, M. Grötschel and B. Korte eds., *Mathematical Programming, The State of the Art - Bonn 1982* Springer-Verlag, Berlin, 1983, pp. 566-603.

[38] Y. Ye and E. Tse, An extension of Karmarkar's projective algorithm for convex quadratic programming, *Mathematical Programming* 44 (1989) 157-179.