

An Improved Exact Penalty Result for Mathematical Programs with Vanishing Constraints

Qingjie Hu^{1,2*}, Haiqi Zhang^{1,2}, Yu Chen³ and Ming Tang^{1,2}

¹Guangxi Colleges and Universities Key Laboratory of Data Analysis and Computation

²School of Mathematics and Computing Science, Guilin University of Electronic Technology, Guilin, P.R. China

³School of Mathematics and Statistics, Guangxi Normal University, Guilin, P.R. China

Email: hqj0715@126.com

Abstract In this paper, by introducing a Fritz-John type result for mathematical program with vanishing constraints (MPVC), we present some new constraint qualifications which are strictly weaker than the MPVC-Mangasarian Fromovitz constraint qualification (MPVC-MFCQ). We show that the MPVC-tailored penalty function which was introduced in [1] is still exact for MPVC under the MPVC-generalized pseudonormality. Our exact penalty result improves the one in [1].

Keywords: Mathematical programs with vanishing constraints, Fritz-John type result, Constraint qualification, Exact penalty

1. Introduction

We consider the following constrained optimization problem

$$\begin{aligned} \min & f(x) \\ \text{s.t.} & g_i(x) \leq 0, \quad i = 1, 2, \dots, p; \\ & h_j(x) = 0, \quad j = 1, 2, \dots, l; \\ & H_i(x) \geq 0, \quad i = 1, 2, \dots, m; \\ & G_i(x)H_i(x) \leq 0, \quad i = 1, 2, \dots, m \end{aligned} \tag{1.1}$$

that we call a *mathematical program with vanishing constraints*, where $f : R^n \rightarrow R$, $g : R^n \rightarrow R^p$, $h : R^n \rightarrow R^l$ and $G, H : R^n \rightarrow R^m$ are all continuously differentiable functions. Throughout this paper, X denotes the feasible region of Problem (1.1).

The MPVC originated from optimization problem of topology of mechanical structures, and it plays an important role in some other fields such as a robot path-finding problem with logic communication constraints in robot motion planning [18], scheduling problems with disjoint feasible regions in power generation dispatch [19] and mixed-integer nonlinear optimal control problems [20]. The major difficulty in solving problem (1.1) is that it does not satisfy most of the standard constraint qualifications, including linearly independent constraint qualification (LICQ) and MFCQ, at any feasible point of interest. The MPVC has attracted much attention in the recent years. Several theoretical properties and different numerical approaches for MPVC can be found in ([1]- [11]). In [1], a new MPVC-tailored penalty function for MPVC has been developed, i.e.,

$$P_\alpha^0(x) = f(x) + \alpha \left[\sum_{i=1}^p \text{dist}_{(-\infty, 0]}(g_i(x)) + \sum_{i=1}^l \text{dist}_{\{0\}}(h_i(x)) + \sum_{i=1}^m \text{dist}_C(G_i(x), H_i(x)) \right],$$

where $\text{dist}_A(x) = \inf\{\|y - x\| : y \in A\}$, $C = \{(a, b) : b \geq 0, ab \leq 0\}$, $\alpha > 0$ is a penalty parameter. This new penalty function $P_\alpha^0(x)$ is shown to be exact under the MPVC-MFCQ. A natural question is whether this penalty function is exact under the mild condition which is weaker than the MPVC-MFCQ.

In this paper, we give a positive answer to the question. Firstly, we prove a Fritz-John type result for MPVC by exploiting the special structure of the vanishing constraints. Based on the Fritz-John type condition, we introduce some new constraint qualifications which are strictly weaker than the MPVC-MFCQ. We show that the MPVC-tailored penalty function $P_\alpha^0(x)$ which was introduced in [1] is still exact for MPVC under the MPVC-generalized pseudonormality. Our exact penalty result is an improved version of the one in [1].

2. Some Constraint Qualifications for MPVC

Motivated by similar ideas for the standard nonlinear programs and the MPEC from [12, 13], we can obtain the following Fritz-Johns type result for the MPVC.

Theorem 2.1 *Let x^* be a local minimum of the MPVC. Then, there are multipliers $\alpha, \lambda, \mu, \gamma, \nu$ such that:*

$$(i) \alpha \nabla f(x^*) + \sum_{i=1}^p \lambda_i \nabla g_i(x^*) + \sum_{i=1}^l \mu_i \nabla h_i(x^*) - \sum_{i=1}^m \gamma_i \nabla H_i(x^*) + \sum_{i=1}^m \nu_i \nabla G_i(x^*) = 0.$$

(ii) $\alpha \geq 0, \lambda_i \geq 0$ for all $i \in I_g(x^*), \lambda_i = 0$ for all $i \notin I_g(x^*), \gamma_i = 0$ for all $i \in I_+(x^*), \gamma_i \geq 0$ for all $i \in I_{0-}(x^*), \gamma_i$ is free for all $i \in I_{0+}(x^*) \cup I_{00}(x^*), \nu_i = 0$ for all $i \in I_{+-}(x^*) \cup I_{0-}(x^*) \cup I_{0+}(x^*), \nu_i \geq 0$ for all $i \in I_{+0}(x^*) \cup I_{00}(x^*), \gamma_i \nu_i = 0$ for all $i \in I_{00}(x^*)$.

(iii) $\alpha, \lambda, \mu, \gamma, \nu$ are not all equal to zero.

(iv) If $\lambda, \mu, \gamma, \nu$ are not all equal to zero, then there is a sequence $\{x^k\} \rightarrow x^*$ such that for all k :

$$f(x^k) < f(x^*),$$

if $\lambda_i > 0$, then $\lambda_i g_i(x^k) > 0$, if $\mu_i \neq 0$, then $\mu_i h_i(x^k) > 0$,

if $\gamma_i > 0$, then $\gamma_i H_i(x^k) < 0$, if $\nu_i > 0$, then $\nu_i G_i(x^k) > 0$.

The technique of proof for the above theorem follows the one of Theorem 3.1 in [13], and is omitted.

Based on Theorem 3.1 and the related discussions in [12, 13], we now define some new constraint qualifications for MPVC as follows.

Definition 2.1 *A vector $x^* \in X$ is said to satisfy the MPVC-generalized MFCQ, if there is no multiplier $(\lambda, \mu, \gamma, \nu) \neq 0$ such that*

(i)

$$\sum_{i=1}^p \lambda_i \nabla g_i(x^*) + \sum_{i=1}^l \mu_i \nabla h_i(x^*) - \sum_{i=1}^m \gamma_i \nabla H_i(x^*) + \sum_{i=1}^m \nu_i \nabla G_i(x^*) = 0.$$

(ii) $\lambda_i \geq 0$ for all $i \in I_g(x^*), \lambda_i = 0$ for all $i \notin I_g(x^*), \gamma_i = 0$ for all $i \in I_+(x^*), \gamma_i \geq 0$ for all $i \in I_{0-}(x^*), \gamma_i$ is free for all $i \in I_{0+}(x^*) \cup I_{00}(x^*), \nu_i = 0$ for all $i \in I_{+-}(x^*) \cup I_{0-}(x^*) \cup I_{0+}(x^*), \nu_i \geq 0$ for all $i \in I_{+0}(x^*) \cup I_{00}(x^*), \gamma_i \nu_i = 0$ for all $i \in I_{00}(x^*)$.

Definition 2.2 *A vector $x^* \in X$ is said to satisfy the MPVC-generalized pseudonormality, if there is no multiplier $(\lambda, \mu, \gamma, \nu)$ such that*

(i) and (ii) hold in Definition 3.1.

(iii) there is a sequence $\{x^k\} \rightarrow x^*$ such that the following is true for all k :

$$\sum_{i=1}^p \lambda_i g_i(x^k) + \sum_{i=1}^l \mu_i h_i(x^k) - \sum_{i=1}^m \gamma_i H_i(x^k) + \sum_{i=1}^m \nu_i G_i(x^k) > 0.$$

Definition 2.3 *A vector $x^* \in X$ is said to satisfy the MPVC-generalized quasinormality, if there is no multiplier $(\lambda, \mu, \gamma, \nu)$ such that*

(i) and (ii) hold in Definition 3.1.

(iii) $(\lambda, \mu, \gamma, \nu) \neq (0, 0, 0, 0)$.

(iv) there is a sequence $\{x^k\} \rightarrow x^*$ such that the following is true for all k : for all $\lambda_i > 0, \lambda_i g_i(x^k) > 0$, for all $\mu_i \neq 0, \mu_i h_i(x^k) > 0$, for all $\gamma_i > 0, \gamma_i H_i(x^k) < 0$, for all $\nu_i > 0, \nu_i G_i(x^k) > 0$.

Subsequently, we give some properties about the three constraint qualifications.

Proposition 2.1 *MPVC-MFCQ \Rightarrow MPVC-generalized MFCQ \Rightarrow MPVC-generalized pseudonormality \Rightarrow MPVC-generalized quasinormality.*

The proof of the above property is obvious, and is omitted.

Remark 2.1 *In fact, the MPVC-generalized MFCQ is strictly weaker than the MPVC-MFCQ. This can be illustrated by the following example. Consider the following 2-dimension example with linear constraints*

$$\begin{aligned} \min \quad & f(x) = x_1^2 + x_2^2 \\ \text{s.t.} \quad & g(x) = x_2 \leq 0, \\ & H(x) = x_1 + x_2 \geq 0, \\ & G(x)H(x) = x_1(x_1 + x_2) \leq 0, \end{aligned}$$

$x^* = (0, 0)$ is a local minimum for the above problem. Obviously, the MPVC-MFCQ is violated at x^* by using Lemma 2.1. On the other hand, it is easy to see that the MPVC-generalized MFCQ is satisfied from Definition 3.1.

3. An Improved Exact Penalty Result for MPVC

In order to derive the exact penalty result, let us first rewrite our MPVC equivalently as

$$\begin{aligned} \min \quad & f(x) \\ \text{s.t.} \quad & F(x) \in \wedge, \end{aligned} \tag{3.2}$$

where $F(x) = \begin{pmatrix} g_i(x)_{i=1,2,\dots,p} \\ h_i(x)_{i=1,2,\dots,l} \\ \begin{pmatrix} G_i(x) \\ H_i(x) \end{pmatrix}_{i=1,2,\dots,m} \end{pmatrix}$, $\wedge = \begin{pmatrix} (-\infty, 0]^p \\ \{0\}^l \\ C^m \end{pmatrix}$, where C is given by

$$C = \{(a, b) \in R^2 | b \geq 0, ab \leq 0\}. \tag{3.3}$$

The penalty function associated to (3.2) is $P_\alpha(x) = f(x) + \alpha \text{dist}_\wedge(F(x))$, where $\text{dist}_\wedge(F(x)) = \inf\{\|y - F(x)\| : y \in \wedge\}$. In this paper, we will restrict ourselves to the l_1 -norm. In this case, $P_\alpha(x)$ is of the form

$$P_\alpha(x) = f(x) + \alpha \left[\sum_{i=1}^p \text{dist}_{(-\infty, 0]}(g_i(x)) + \sum_{i=1}^l \text{dist}_{\{0\}}(h_i(x)) + \sum_{i=1}^m \text{dist}_C(G_i(x), H_i(x)) \right].$$

Our next goal is to prove that the penalty function $P_\alpha(x)$ is exact at a local minimum x^* of (1.1) satisfying a suitable constraint qualification, i.e., that there is a $\bar{\alpha} \geq 0$ such that x^* is a local minimum of $P_\alpha(x)$ for any $\alpha \geq \bar{\alpha}$.

Firstly, we give the following auxiliary results.

Lemma 3.1 [1] *Let C be given by (3.3). Then for $a, b \in R$, we have*

$$\begin{aligned} \text{dist}_{(-\infty, 0]}(a) &= \max\{a, 0\}, \quad \text{dist}_{\{0\}}(a) = |a|, \\ \text{dist}_C(a, b) &= \max\{0, -b, \min\{a, b\}\} = \begin{cases} \min\{a, b\}, & \text{if } a, b \geq 0; \\ 0, & \text{if } a \leq 0, b \geq 0; \\ -b, & \text{if } b \leq 0. \end{cases} \end{aligned}$$

By using Lemma 4.1 and Lemma 5.1 in [1], similar to the proof of Lemma 4.2 in [13], we can get the following lemma.

Lemma 3.2 *Let C be given by (3.3). Then for $a, b \in R$ we have*

$$\partial \text{dist}_{(-\infty, 0]}(a) = \begin{cases} \{0\}, & \text{if } a < 0; \\ [0, 1], & \text{if } a = 0; \\ \{1\}, & \text{if } a > 0. \end{cases} \quad \partial \text{dist}_{\{0\}}(a) = \begin{cases} \{-1\}, & \text{if } a < 0; \\ [-1, 1], & \text{if } a = 0; \\ \{1\}, & \text{if } a > 0. \end{cases}$$

$$\partial dist_C(a, b) = \begin{cases} \{(0, -1)^T\}, & \text{if } b < 0 \\ \{(0, 1)^T\}, & \text{if } a > b > 0; \\ \{(1, 0)^T\}, & \text{if } b > a > 0; \\ \{(0, 1)^T\} \cup \{(1, 0)^T\}, & \text{if } a = b > 0; \\ \{(0, 0)^T\}, & \text{if } b > 0, a < 0; \\ conv\{(0, 0)^T, (1, 0)^T\}, & \text{if } b > 0, a = 0; \\ conv\{(0, -1)^T, (0, 1)^T\}, & \text{if } b = 0, a > 0; \\ conv\{(0, -1)^T, (0, 0)^T\}, & \text{if } b = 0, a < 0; \\ conv\{(0, -1)^T, (0, 1)^T\} \cup conv\{(0, 0)^T, (1, 0)^T\}, & \text{if } a = b = 0, \end{cases}$$

where the convex hull of a set X , denoted by $conv(X)$, is the set of all convex combinations of elements of X .

The following result is very important for the proof of Theorem 4.1.

Lemma 3.3 *If $x^* \in X$ is MPVC-generalized pseudonormal, then all feasible points in a neighborhood of x^* are MPVC-generalized pseudonormal.*

Combining the proof techniques of Lemma 1 in [14], Lemma 2 in [15] and Lemma 3.1 in [16], we can give the proof of the above lemma. We omit the proof here.

Based on Lemma 4.1, Lemma 4.2 and Lemma 4.3, by using a different and elementary proof technique, we can prove the following exact result under the MPVC-generalized pseudonormality.

Theorem 3.1 *Let x^* be a local minimizer of (1.1). If the MPVC-generalized pseudonormality holds at x^* and f is locally Lipschitz continuous around x^* , then the penalty function $P_\alpha(x)$ is exact at x^* .*

Proof. Assume that the conclusion is not correct, i.e., there exists a positive number sequence $\{\alpha_k\} \rightarrow +\infty (k \rightarrow +\infty)$, such that x^* is not a local minimum of the function

$$P_{\alpha_k}(x) = f(x) + \alpha_k \left[\sum_{i=1}^p dist_{(-\infty, 0]}(g_i(x)) + \sum_{i=1}^l dist_{\{0\}}(h_i(x)) + \sum_{i=1}^m dist_C(G_i(x), H_i(x)) \right].$$

Let $\varepsilon > 0$ be sufficiently small such that

$$f(x^*) \leq f(x), \quad \forall x \in X, \quad \|x - x^*\| \leq \varepsilon. \tag{3.4}$$

Taking into account the continuity of $P_{\alpha_k}(x)$, we can assume that x^k minimizes $P_{\alpha_k}(x)$ over the set satisfying $\|x - x^*\| \leq \varepsilon$. Since x^* is not a local minimum of $P_{\alpha_k}(x)$, then we must have that $x^k \neq x^*$.

The following object is to show that x^k is infeasible for (1.1). Actually, assume that x^k is feasible for (1.1), then $P_{\alpha_k}(x^k) = f(x^k)$. Taking into account that $x^k \in X$ and $\|x^k - x^*\| \leq \varepsilon$, from (3.4), we can get $f(x^*) \leq f(x^k)$. In view of that $\|x^* - x^k\| \leq \varepsilon$, one gets that $P_{\alpha_k}(x^k) \leq P_{\alpha_k}(x^*)$. Noting that $P_{\alpha_k}(x^*) = f(x^*)$, we obtain that $P_{\alpha_k}(x^k) = P_{\alpha_k}(x^*)$. Obviously, this contradicts the fact that x^* is not a local minimum of $P_{\alpha_k}(x)$ over the set satisfying $\|x - x^*\| \leq \varepsilon$.

Hence, x^k is infeasible for (1.1), i.e.,

$$\sum_{i=1}^p dist_{(-\infty, 0]}(g_i(x^k)) + \sum_{i=1}^l dist_{\{0\}}(h_i(x^k)) + \sum_{i=1}^m dist_C(G_i(x^k), H_i(x^k)) > 0.$$

In view of that x^* is not a local minimum of $P_{\alpha_k}(x)$, we have

$$\begin{aligned} P_{\alpha_k}(x^k) &= f(x^k) + \alpha_k \left[\sum_{i=1}^p dist_{(-\infty, 0]}(g_i(x^k)) + \sum_{i=1}^l dist_{\{0\}}(h_i(x^k)) \right. \\ &\quad \left. + \sum_{i=1}^m dist_C(G_i(x^k), H_i(x^k)) \right] \leq f(x^*). \end{aligned} \tag{3.5}$$

Noting that $\{f(x)\}$ and $\{P_{\alpha_k}(x)\}$ are all bounded over the set of satisfying $\|x - x^*\| \leq \varepsilon$, (3.5) implies that

$$\lim_{k \rightarrow \infty} dist_{(-\infty, 0]}(g_i(x^k)) = 0, \forall i = 1, \dots, p, \quad \lim_{k \rightarrow \infty} dist_{\{0\}}(h_i(x^k)) = 0, \forall i = 1, \dots, l,$$

$$\lim_{k \rightarrow \infty} \text{dist}_C(G_i(x^k), H_i(x^k)) = 0, \forall i = 1, \dots, m.$$

Since $\{x^k\}$ is bounded, if \bar{x} is any of its accumulation points, we have that \bar{x} is feasible for (1.1) and $\|\bar{x} - x^*\| \leq \varepsilon$. Without loss of generality, we assume that $\{x^k\}$ converges to \bar{x} . Taking into account Proposition 2.4.3 in [17], the fact that x^k minimizes $P_{\alpha_k}(x)$ over the set satisfying $\|x - x^*\| \leq \varepsilon$ shows that for the sufficiently large σ ,

$$0 \in \partial[P_{\alpha_k}(x^k) + \sigma \text{dist}_{(-\infty, 0]}(\|x^k - x^*\| - \varepsilon)],$$

i.e., there are

$$\begin{aligned} \xi^k &\in \partial f(x^k), \quad \lambda_i^k \in \partial \text{dist}_{(-\infty, 0]}(g_i(x^k)), \quad \forall i = 1, 2, \dots, p, \\ \mu_i^k &\in \partial \text{dist}_{\{0\}}(h_i(x^k)), \quad \forall i = 1, 2, \dots, l, \quad \zeta^k \in \partial \text{dist}_{(-\infty, 0]}(\|x^k - x^*\| - \varepsilon), \\ (\nu_i^k, -\gamma_i^k) &\in \partial \text{dist}_C(G_i(x^k), H_i(x^k)), \quad \forall i = 1, 2, \dots, m \end{aligned}$$

such that

$$\xi^k + \alpha_k \left[\sum_{i=1}^p \lambda_i^k \nabla g_i(x^k) + \sum_{i=1}^l \mu_i^k \nabla h_i(x^k) - \sum_{i=1}^m \gamma_i^k \nabla H_i(x^k) + \sum_{i=1}^m \nu_i^k \nabla G_i(x^k) \right] + \sigma \zeta^k \frac{x^k - x^*}{\|x^k - x^*\|} = 0.$$

For the sufficiently large k , we let $\sigma = \sqrt{\alpha_k}$. From the above equation, we also can get

$$\frac{1}{\alpha_k} \xi^k + \sum_{i=1}^p \lambda_i^k \nabla g_i(x^k) + \sum_{i=1}^l \mu_i^k \nabla h_i(x^k) - \sum_{i=1}^m \gamma_i^k \nabla H_i(x^k) + \sum_{i=1}^m \nu_i^k \nabla G_i(x^k) + \frac{1}{\sqrt{\alpha_k}} \zeta^k \frac{x^k - x^*}{\|x^k - x^*\|} = 0.$$

Noting Lemma 4.2, it is to see that the sequence $\{(\xi^k, \lambda^k, \mu^k, \gamma^k, \nu^k, \zeta^k)\}$ is bounded. Thus, we can assume without loss of generality that the sequence $\{(\lambda^k, \mu^k, \gamma^k, \nu^k)\}$ converges to some limit $(\lambda, \mu, \gamma, \nu)$. Taking the limit as $k \rightarrow \infty$ in the above equation, this yields

$$\sum_{i=1}^p \lambda_i \nabla g_i(\bar{x}) + \sum_{i=1}^l \mu_i \nabla h_i(\bar{x}) - \sum_{i=1}^m \gamma_i \nabla H_i(\bar{x}) + \sum_{i=1}^m \nu_i \nabla G_i(\bar{x}) = 0$$

from the smoothness of g, h, G, H . Furthermore, we can obtain that from Lemma 4.2

$$\begin{aligned} \lambda_i &\geq 0, \quad \forall i \in I_g(\bar{x}), \quad \lambda_i = 0, \quad \forall i \notin I_g(\bar{x}), \\ \gamma_i &= 0, \quad \forall i \in I_+(\bar{x}), \quad \gamma_i \geq 0, \quad \forall i \in I_0(\bar{x}), \quad \gamma_i \text{ is free}, \quad \forall i \in I_{0+}(\bar{x}) \cup I_{00}(\bar{x}), \\ \nu_i &= 0, \quad \forall i \in I_{+-}(\bar{x}) \cup I_{0-}(\bar{x}) \cup I_{0+}(\bar{x}), \quad \nu_i \geq 0, \quad \forall i \in I_{+0}(\bar{x}) \cup I_{00}(\bar{x}), \\ \gamma_i \nu_i &= 0, \quad \forall i \in I_{00}(\bar{x}). \end{aligned}$$

On the other hand, it is easy to see that, for all k ,

$$\begin{aligned} \lambda_i g_i(x^k) &\geq 0, \quad \forall i = 1, 2, \dots, p, \quad \mu_i h_i(x^k) \geq 0, \quad \forall i = 1, 2, \dots, l, \\ -\gamma_i H_i(x^k) &\geq 0, \quad \forall i = 1, 2, \dots, m, \quad \nu_i G_i(x^k) \geq 0, \quad \forall i = 1, 2, \dots, m. \end{aligned}$$

Because x^k is infeasible for (1.1) for all k , at least one constraint for (1.1) has to be violated infinitely many times. If $g_i(x^k) \leq 0$ is violated infinitely many times, we have $\lambda_i g_i(x^k) > 0$ for those k . If $h_i(x^k) = 0$ is violated infinitely many times, we have $\mu_i h_i(x^k) > 0$ for those k . If $H_i(x^k) \geq 0$ and $G_i(x^k) H_i(x^k) \leq 0$ are violated infinitely many times, from Lemma 4.1 and Lemma 4.2, we have $-\gamma_i H_i(x^k) + \nu_i G_i(x^k) > 0$ for those k . This yields

$$\sum_{i=1}^p \lambda_i g_i(x^k) + \sum_{i=1}^l \mu_i h_i(x^k) - \sum_{i=1}^m \gamma_i H_i(x^k) + \sum_{i=1}^m \nu_i G_i(x^k) > 0$$

at least on a subsequence K . However, this implies that the MPVC-generalized pseudonormality does not hold at \bar{x} . This contradicts Lemma 4.3. The proof is complete. ■

In order to show the feasibility and effectiveness of proposed result, we can provide the following example.

Example 1. Consider the two-dimensional MPVC problem

$$\begin{aligned} \min \quad & f(x) = x_1^2 + x_2^2 \\ \text{s.t.} \quad & H_1(x) = x_1 + x_2 \geq 0; \\ & G_1(x)H_1(x) = -(x_1 + x_2)(x_1 + x_2) \leq 0, \end{aligned}$$

$x^* = (0, 0)^T$ is a minimizer of the above MPVC. Obviously, we can see that the MPVC-MFCQ at x^* does not hold. However, by Definition 2.1, we can conclude that the MPVC-generalized MFCQ at x^* holds. In view of Proposition 2.1, this implies that the MPVC-generalized pseudonormality at x^* holds. On the other hand, it is easy to see that x^* is a global minimizer of the following penalty function for any $\alpha > 0$:

$$P_\alpha(x) = x_1^2 + x_2^2 + \alpha \text{dist}_C(-(x_1 + x_2), x_1 + x_2) = x_1^2 + x_2^2.$$

This example shows that the penalty function P_α is exact at x^* under the MPVC-generalized pseudonormality which is strictly weaker than the MPVC-MFCQ.

Acknowledgments. The authors would like to express their thanks to the anonymous referees for their constructive comments on the paper, which have greatly improved its presentation. This work was supported in part by NNSF(No.11461015,11761014) of China and Guangxi Natural Science Foundation (No.2015GXNSFAA139010,2017GXNSFAA198243).

References

1. T.Hoheisel, C.Kanzow, J.V.Outrata. "Exact penalty results for mathematical programs with vanishing constraints," *Nonlinear Analysis: Theory, Methods and Applications*, vol.72, no.5, pp.2514-2526, 2010.
2. W.Achtziger, C.Kanzow. "Mathematical programs with vanishing constraints: Optimality conditions and constraints qualifications," *Mathematical Programming*, vol.114,no.1, pp.69-99, 2008.
3. W.Achtziger, T.Hoheisel, C.Kanzow. "A smoothing-regularization approach to mathematical programs with vanishing constraints," *Computational Optimization and Applications*, vol.55, no.1, pp.733-767, 2013.
4. W.Achtziger, T.Hoheisel, C.Kanzow. "On a relaxation method for mathematical programs with vanishing constraints," *GAMM-Mitt.*, vol.35, no.2, pp.110-130, 2012.
5. D.Dorsch, V.Shikhman, O.Stein. "Mathematical programs with vanishing constraints: critical point theory," *Journal of Global Optimization*, vol.52,no.3, pp.591-605, 2012.
6. T.Hoheisel, C.Kanzow. "On the Abadie and Guignard constraint qualification for mathematical programs with vanishing constraints," *Optimization*, vol.58, no.4, pp.431-448, 2009.
7. T.Hoheisel, C.Kanzow. "Stationary conditions for mathematical programs with vanishing constraints using weak constraint qualification," *Journal of Mathematical Analysis and Applications*, vol.337,no.1, pp.292-310, 2008.
8. T.Hoheisel, C.Kanzow. "First- and second-order optimality conditions for mathematical programs with vanishing constraints," *Applications of Mathematics*, vol.52,no.6, pp.495-514, 2007.
9. T.Hoheisel, C.Kanzow, A.Schwartz. "Convergence of a local regularization approach for mathematical programs with complementarity or vanishing constraints," *Optimization Method and Software*, vol.27, no.3, pp.483-512, 2012.
10. A.F.Izmailov, A.L. Pogosyan. "Optimality conditions and Newton-type methods for mathematical programs with vanishing constraints," *Computational Mathematics and Mathematical Physics*, vol.49, no.7, pp.1128-1140, 2009.
11. A.F.Izmailov, M. V. Solodov. "Mathematical Programs with Vanishing Constraints: Optimality Conditions, Sensitivity and a Relaxation Method," *Journal of Optimization Theory and Applications*, vol.142, no.3, pp.501-532, 2009.
12. D.P.Bertsekas, A.E.Ozdoglar. "Pseudonormality and a Lagrange multiplier theory for constrained optimization," *Journal of Optimization Theory and Applications*, vol.114, no.2, pp.287-343, 2002.
13. C.Kanzow, A.Schwartz. "Mathematical programs with equilibrium constraints: enhanced Fritz John-conditions, new constraint qualifications and improved exact penalty results," *SIAM Journal on Optimization*, vol.20, no.5, pp.2730-2753, 2010.

14. J.J.Ye, J.Zhang. "Enhanced Karush-Kuhn-Tucker condition and weaker constraint qualifications," *Mathematical Programming*, Ser.B. vol.139, no.1-2, pp.353-381, 2013.
15. D.P.Bertsekas, A.E.Ozdaglar. "The relation between pseudonormality and quasiregularity in constrained optimization," *Optimization Methods and Software*, vol.19, no.5, pp.493-506, 2004.
16. J.J.Ye, J.Zhang. "Enhanced Karush-Kuhn-Tucker condition for mathematical programs with equilibrium constraints," *Journal of Optimization Theory and Applications*, DOI 10.1007/s10957-013-0493-3, 2013.
17. F.H.Clarke. "Optimization and Nonsmooth Analysis," Wiley, New York, 1983.
18. C.Kirches, A.Potschka, H.G.Bock, S.Sager. "A parametric active set method for quadratic programs with vanishing constraints," *Pac. J. Optim.*, vol.9, no.2, pp.275-299, 2013.
19. R.A.Jabr. "Solution to economic dispatching with disjoint feasible regions via semidefinite programming," *IEEE Trans. Power Syst.*, vol.27, no.1, pp.572-573, 2012.
20. N.J.Michael, C.Kirches, S.Sager. "On perspective functions and vanishing constraints in mixed-integer nonlinear optimal control," *Facets of Combinatorial Optimization*, pp.387-417, 2013.