A NOTE ON THE SQUARED SLACK VARIABLES
TECHNIQUE FOR NONLINEAR OPTIMIZATION

Ellen H. Fukuda   Masao Fukushima
Kyoto University   *Nanzan University*

Abstract   In constrained nonlinear optimization, the squared slack variables can be used to transform a problem with inequality constraints into a problem containing only equality constraints. This reformulation is usually not considered in the modern literature, mainly because of possible numerical instabilities. However, this argument only concerns the development of algorithms, and nothing stops us in using the strategy to understand the theory behind these optimization problems. In this note, we clarify the relation between the Karush-Kuhn-Tucker points of the original and the reformulated problems. In particular, we stress that the second-order sufficient condition is the key to establish their equivalence.

1.   Introduction

The technique for converting an optimization problem with inequality constraints into a problem containing only equality constraints using *squared slack variables* is well-known for decades. It had been used by many researchers, even before the emerging of modern studies of algorithms for *nonlinear programming* (NLP) in 1960's [6]. In the present days, it is still a useful tool for the optimization theory [2]. Although it had been also used in the development of algorithms [9–11], the approach is usually avoided for this purpose, especially in the optimization community. The increase of the dimension of the problem is one of the reasons to avoid it, but the computational capabilities nowadays makes it less problematic. In fact, the main reason lies certainly in the possible numerical instabilities caused by the reformulation [4, 8]. These difficulties were also shown recently in [1], where numerical experiments using the sequential quadratic programming method were performed.

If one considers the pros and cons of the squared slack variables, the argument against it may be more predominant. However, here we choose a similar path taken by Bertsekas in [2]. In this book, optimality conditions for problems containing only equality constraints are considered first. Then, the squared slack variables strategy is used to derive optimality conditions for problems with inequality constraints. Similarly, in this work, we consider the slack variables technique as a tool to understand the theory behind optimization problems. More specifically, our aim is to analyze the relation between the original problem containing inequality constraints and the reformulated problem with additional slack variables. It is well-known that these problems are equivalent in terms of global/local optimal solutions, but the relation between their stationary points, or Karush-Kuhn-Tucker (KKT) points, has been unclear until recently.

In fact, Fukuda and Fukushima [3] had established these relations in the context of *nonlinear second-order cone programming* (NSOCP). For such problems, the reformulation

using slack variables turns out to be an NLP problem with a particular structure. This can be viewed as an advantage, if one considers the second-order cone as an object that is difficult to deal with. Subsequently, Lourenço, Fukuda and Fukushima [5] extended this work for *nonlinear semidefinite programming* (NSDP) problems, although the main motivation in this case was in easily deriving their second-order conditions. In both NSOCP and NSDP cases, the equivalence between the original and the reformulated problems was established with the *second-order sufficient condition.*

Recalling that NLP problems are particular cases of NSOCP problems (which, in turn, are particular cases of NSDP problems), in this paper, we turn back to the former more primitive type of problems. There are two reasons for that. One is that such an analysis for NLP had apparently not been published in the literature. Another reason is to observe if there are some gap between the results obtained for NLP and the ones given for NSOCP and NSDP. As it can be seen in this work, it turns out that the results are similar. However, as expected, the analyses here are much easier to follow, because it does not involve complicated Jordan algebras or operations with matrices. Moreover, most researchers from optimization are familiar with NLP, but the same cannot be said for NSOCP and NSDP. This motivated us to write down this paper.

The following notations will be used here. The Euclidean inner product and norm are denoted by $\langle \cdot, \cdot \rangle$ and $\| \cdot \|$, respectively. The transpose of a vector $x \in \mathbb{R}^s$ is denoted by $x^\top$. For any vector $x := (x_1, \ldots, x_s)^\top \in \mathbb{R}^s$, we use $\mathrm{diag}(x)$ to represent the diagonal matrix with diagonal entries x_i, $i = 1, \ldots, s$. The gradient and the Hessian of a function $p \colon \mathbb{R}^s \to \mathbb{R}$ at $x \in \mathbb{R}^s$ are denoted by $\nabla p(x)$ and $\nabla^2 p(x)$, respectively. For a function $q \colon \mathbb{R}^{s+\ell} \to \mathbb{R}$, the gradient and the Hessian of q at $(x, y) \in \mathbb{R}^{s+\ell}$ with respect to x are denoted by $\nabla_x q(x, y)$ and $\nabla_x^2 q(x, y)$, respectively. For a function $r \colon \mathbb{R}^s \to \mathbb{R}^\ell$, the Jacobian of r at $x \in \mathbb{R}^s$ is given by $Jr(x) \in \mathbb{R}^{\ell \times s}$.

The paper is organized as follows. In Section 2, we introduce the definition of the problem, the KKT conditions, and other preliminary results. In Section 3, we show that the original problem is equivalent to the reformulated problem with squared slack variables in terms of KKT points, under the second-order sufficient conditions. Since KKT conditions are necessary for optimality under a constraint qualification, in Section 4, we also prove the equivalence between linear independence constraint qualification satisfied by KKT points of the original and the reformulated problems. We conclude with some final remarks in Section 5.

2. Preliminaries

Let us consider the following nonlinear programming (NLP) problem with inequality constraints:

$$\begin{aligned} \underset{x}{\text{minimize}} \quad & f(x) \\ \text{subject to} \quad & g(x) \geq 0, \end{aligned} \tag{P1}$$

where $f \colon \mathbb{R}^n \to \mathbb{R}$ and $g \colon \mathbb{R}^n \to \mathbb{R}^m$ are twice continuously differentiable functions. Also, let $g := (g_1, \ldots, g_m)^\top$ with $g_i \colon \mathbb{R}^n \to \mathbb{R}$, $i = 1, \ldots, m$. Introducing slack variables $y := (y_1, \ldots, y_m)^\top \in \mathbb{R}^m$, we obtain the following formulation:

$$\begin{aligned} \underset{x,y}{\text{minimize}} \quad & f(x) \\ \text{subject to} \quad & g_i(x) - y_i^2 = 0, \quad i = 1, \ldots, m. \end{aligned} \tag{P2}$$

The above problem is equivalent to (P1) in the following sense. If (x^*, y^*) is a global (local) optimal solution of (P2), then x^* is a global (local) optimal solution of (P1). Conversely,

if x^* is a global (local) optimal solution of (P1), then there exists y^* such that (x^*, y^*) is a global (local) optimal solution of (P2) [10, Proposition 3.1].

From the practical viewpoint, it is more important to examine the relation between stationary points, or KKT points, of the two problems, because we can only expect to compute such points in practice. However, the relation between stationary points is less clear than that between optimal solutions.

We say that $(x, \lambda) \in \mathbb{R}^{n+m}$ is a KKT pair of problem (P1) if the following KKT conditions hold:

$$\nabla f(x) - \sum_{i=1}^{m} \lambda_i \nabla g_i(x) = 0, \tag{P1.1}$$

$$\lambda_i \geq 0, \quad i = 1, \ldots, m, \tag{P1.2}$$

$$g_i(x) \geq 0, \quad i = 1, \ldots, m, \tag{P1.3}$$

$$\lambda_i g_i(x) = 0, \quad i = 1, \ldots, m. \tag{P1.4}$$

Also, we say that $(x, y, \lambda) \in \mathbb{R}^{n+2m}$ is a KKT triple of problem (P2) when

$$\nabla f(x) - \sum_{i=1}^{m} \lambda_i \nabla g_i(x) = 0, \tag{P2.1}$$

$$y_i \lambda_i = 0, \quad i = 1, \ldots, m, \tag{P2.2}$$

$$g_i(x) - y_i^2 = 0, \quad i = 1, \ldots, m. \tag{P2.3}$$

Notice that under a constraint qualification, the above conditions, for both problems, are necessary for optimality [2]. For a KKT pair (x, λ) of (P1), we define the following sets of indices:

$$\begin{aligned}
\mathcal{I}_0 &:= \big\{i \in \{1, \ldots, m\} \colon g_i(x) = 0\big\}, \\
\mathcal{I}_{00} &:= \big\{i \in \{1, \ldots, m\} \colon g_i(x) = 0, \lambda_i = 0\big\}, \\
\mathcal{I}_{0P} &:= \big\{i \in \{1, \ldots, m\} \colon g_i(x) = 0, \lambda_i > 0\big\}, \\
\mathcal{I}_{P0} &:= \big\{i \in \{1, \ldots, m\} \colon g_i(x) > 0, \lambda_i = 0\big\}.
\end{aligned} \tag{2.1}$$

Observe that these sets are also suitable for a KKT tripe (x, y, λ) of (P2). In the latter case, however, λ_i is not necessarily nonnegative. So, we also have to consider the following index set:

$$\mathcal{I}_{0N} := \big\{i \in \{1, \ldots, m\} \colon g_i(x) = 0, \lambda_i < 0\big\}. \tag{2.2}$$

Clearly, the sets $\mathcal{I}_{00}$, $\mathcal{I}_{0P}$ and $\mathcal{I}_{0N}$ constitute a partition of $\mathcal{I}_0$, and the sets $\mathcal{I}_0$ and $\mathcal{I}_{P0}$ constitute a partition of the whole set of indices $\{1, \ldots, m\}$. Moreover, from (P2.3), y_i is determined by the value of $g_i(x)$. In other words, $y_i = 0$ if and only if $i \in \mathcal{I}_0 = \mathcal{I}_{00} \cup \mathcal{I}_{0P} \cup \mathcal{I}_{0N}$, and $y_i \neq 0$ if and only if $i \in \mathcal{I}_{P0}$. We also point out that, for problem (P1), the well-known *strict complementarity* condition means that $\mathcal{I}_{00} = \emptyset$.

3. Equivalence Between KKT Points

Here, we will establish the equivalence between KKT points of problems (P1) and (P2). One of the implications is simple, as shown in the next proposition.

Proposition 3.1. *Let $(x, \lambda) \in \mathbb{R}^{n+m}$ be a KKT pair of* (P1)*. Then, there exists $y \in \mathbb{R}^m$ such that (x, y, λ) is a KKT triple of* (P2)*.*

Proof. The condition (P2.1) holds trivially. Observe that (P1.3) implies the existence of $y_i \in \mathbb{R}$, $i = 1, \ldots, m$, such that (P2.3) holds. Moreover, from (P1.4) and (P2.3), we have

$$\left(y_i \lambda_i\right)^2 = g_i(x)\lambda_i^2 = 0$$

for all $i = 1, \ldots, m$. Then, (P2.2) also holds. □

The converse is not always true, that is, even if (x, y, λ) is a KKT triple of (P2), (x, λ) is not necessarily a KKT pair of (P1). In fact, the condition (P1.2), concerning the sign of the multiplier, may not hold. The following example illustrates this situation.

Example 3.2. *Let problem* (P1) *be defined as follows*

$$\begin{array}{cl} \underset{x}{minimize} & \frac{1}{2}x^2 \\ subject \ to & \sin x \geq 0, \end{array}$$

with $n = 1$. *Then,* $(x, y, \lambda) = (0, 0, 0)$ *and* $(x, y, \lambda) = (\pi, 0, -\pi)$ *are both KKT triples of* (P2). *However,* $(x, \lambda) = (0, 0)$ *is a KKT pair of* (P1), *and* $(x, \lambda) = (\pi, -\pi)$ *is not, since the condition* (P1.2) *fails to hold.*

We will show now that the converse is true when the second-order sufficient condition is assumed (see, for example, [2, Section 3.3] or [7, Section 12.5]). To this end, we define the Lagrangian functions $L \colon \mathbb{R}^{n+m} \to \mathbb{R}$ and $\mathcal{L} \colon \mathbb{R}^{n+2m} \to \mathbb{R}$ for problems (P1) and (P2), respectively, by

$$\begin{aligned} L(x, \lambda) &:= f(x) - \sum_{i=1}^{m} \lambda_i g_i(x), \\ \mathcal{L}(x, y, \lambda) &:= f(x) - \sum_{i=1}^{m} \lambda_i \left(g_i(x) - y_i^2\right). \end{aligned}$$

Thus, the second-order sufficient condition for (P1), which has only inequality constraints, is given below.

Definition 3.3. *Let* $(x, \lambda) \in \mathbb{R}^{n+m}$ *be a KKT pair of* (P1). *The* second-order sufficient condition (SOSC) *holds if*

$$\left\langle \nabla_x^2 L(x, \lambda)d, d \right\rangle > 0$$

for all nonzero $d \in \mathbb{R}^n$ *such that*

$$\left\langle \nabla g_i(x), d \right\rangle = 0, \quad i \in \mathcal{I}_{0P} \quad and \quad \left\langle \nabla g_i(x), d \right\rangle \geq 0, \quad i \in \mathcal{I}_{00},$$

where

$$\nabla_x^2 L(x, \lambda) = \nabla^2 f(x) - \sum_{i=1}^{m} \lambda_i \nabla^2 g_i(x).$$

The above definition shows that SOSC holds when the Hessian of the Lagrangian is positive definite on the critical cone. For problems with only equality constraints, the definition of SOSC is the same, but the critical cone is different. In fact, if an NLP problem has only $h_i(x) = 0$, $i = 1, \ldots, s$ as constraints, then a direction $d \in \mathbb{R}^n$ is in the critical cone if it satisfies $\langle \nabla h_i(x), d \rangle = 0$ for all $i = 1, \ldots, s$. This fact will be used to prove the next proposition, which shows that SOSC of the equality constrained problem (P2) can be written using the Lagrangian function associated to (P1).

Proposition 3.4. *Let* $(x, y, \lambda) \in \mathbb{R}^{n+2m}$ *be a KKT triple of* (P2). *The SOSC holds if*

$$\left\langle \nabla_x^2 L(x, \lambda) v, v \right\rangle + 2 \sum_{i=1}^m \lambda_i w_i^2 > 0 \tag{3.1}$$

for all nonzero $(v, w) \in \mathbb{R}^{n+m}$ *such that*

$$\langle \nabla g_i(x), v \rangle - 2 y_i w_i = 0, \quad i = 1, \ldots, m.$$

Proof. From the usual definition of SOSC in nonlinear programming, we observe that a KKT point (x, y, λ) satisfies SOSC when

$$\left\langle \nabla_{(x,y)}^2 \mathcal{L}(x, y, \lambda) d, d \right\rangle > 0$$

for all nonzero $d \in \mathbb{R}^{n+m}$ such that

$$\left(\nabla g_i(x)^\top, -2 y_i e_i^\top \right) d = 0, \quad i = 1, \ldots, m,$$

where e_i is the ith column of the identity matrix of dimension m and

$$\nabla_{(x,y)}^2 \mathcal{L}(x, y, \lambda) = \begin{bmatrix} \nabla_x^2 L(x, \lambda) & 0 \\ 0 & 2 \operatorname{diag}(\lambda) \end{bmatrix}.$$

The result follows by letting $d := (v, w)$ with $v \in \mathbb{R}^n$ and $w \in \mathbb{R}^m$. $\qquad \square$

Lemma 3.5. *Let* $(x, y, \lambda) \in \mathbb{R}^{n+2m}$ *be a KKT triple of* (P2) *and assume that it satisfies SOSC. Then, we have* $\mathcal{I}_{00} = \mathcal{I}_{0N} = \emptyset$.

Proof. Assume that there exists an index j such that $g_j(x) = y_j = 0$. Let us prove that in this case $\lambda_j > 0$. Taking $v = 0$ in (3.1), we have

$$\lambda_j w_j^2 + \sum_{i \neq j} \lambda_i w_i^2 > 0 \tag{3.2}$$

for all nonzero $w \in \mathbb{R}^m$ such that

$$y_i w_i = 0, \quad i = 1, \ldots, m.$$

In particular, the inequality (3.2) holds when $w_j \neq 0$ and $w_i = 0$ for all $i \neq j$. But this choice of w shows that $\lambda_j w_j^2 > 0$, which implies $\lambda_j > 0$. Therefore, we conclude that $\mathcal{I}_{00} = \mathcal{I}_{0N} = \emptyset$. $\qquad \square$

Proposition 3.6. *Let* $(x, y, \lambda) \in \mathbb{R}^{n+2m}$ *be a KKT triple of* (P2) *and assume that it satisfies SOSC. Then,* (x, λ) *is a KKT pair of* (P1).

Proof. Observe that (P1.1) trivially holds and that (P2.3) implies (P1.3). For each $i = 1, \ldots, m$, multiplying (P2.2) with y_i and recalling (P2.3), we obtain

$$y_i(y_i \lambda_i) = 0 \quad \Leftrightarrow \quad g_i(x)\lambda_i = 0,$$

and so (P1.4) is satisfied. Finally, (P1.2) holds because $\mathcal{I}_{0N} = \emptyset$ from Lemma 3.5. $\qquad \square$

The next proposition shows that the KKT pair (x, λ) of (P1) also satisfies SOSC. In addition, it also satisfies the strict complementarity.

Proposition 3.7. *Let $(x, y, \lambda) \in \mathbb{R}^{n+2m}$ be a KKT triple of* (P2) *that satisfies SOSC. Then,* (x, λ) *is a KKT pair of* (P1) *satisfying SOSC and the strict complementarity.*

Proof. Proposition 3.6 shows that (x, λ) is a KKT pair of (P1) and it also satisfies the strict complementarity ($\mathcal{I}_{00} = \emptyset$) from Lemma 3.5. Recalling that $\lambda_i = 0$ for all $i \in \mathcal{I}_{P0}$, we can rewrite the SOSC of (P2) as

$$\langle \nabla_x^2 L(x, \lambda)v, v \rangle + 2 \sum_{i \in \mathcal{I}_{0P}} \lambda_i w_i^2 > 0$$

for all nonzero $(v, w) \in \mathbb{R}^{n+m}$ such that

$$\begin{aligned} \langle \nabla g_i(x), v \rangle &= 0, \quad i \in \mathcal{I}_{0P}, \\ \langle \nabla g_i(x), v \rangle - 2y_i w_i &= 0, \quad i \in \mathcal{I}_{P0}. \end{aligned}$$

Since there is no restriction for w_i with $i \in \mathcal{I}_{0P}$, we can set $w_i = 0$ for all $i \in \mathcal{I}_{0P}$. Also, we observe that w_i, $i \in \mathcal{I}_{P0}$ are determined by the value of $v \in \mathbb{R}^n$. Indeed, if there exists a nonzero $v \in \mathbb{R}^n$ satisfying $\langle \nabla g_i(x), v \rangle = 0$ for all $i \in \mathcal{I}_{0P}$, then, for each $i \in \mathcal{I}_{P0}$, there exists $w_i \in \mathbb{R}$ such that $\langle \nabla g_i(x), v \rangle - 2y_i w_i = 0$, since $y_i \neq 0$. Thus, from the SOSC given above, we have

$$\langle \nabla_x^2 L(x, \lambda)v, v \rangle > 0$$

for all nonzero $v \in \mathbb{R}^n$ such that $\langle \nabla g_i(x), v \rangle = 0$, $i \in \mathcal{I}_{0P}$. This condition holds true vacuously, when there exists no $v \neq 0$ satisfying $\langle \nabla g_i(x), v \rangle = 0$ for all $i \in \mathcal{I}_{0P}$. Hence, recalling that $\mathcal{I}_{00} = \emptyset$, we conclude that (x, λ) satisfies the SOSC of (P1). $\square$

The above results show that if the SOSC of the reformulated problem (P2) is satisfied, then, in order to obtain a KKT point of the original problem (P1), it is sufficient to find a KKT point of the reformulated problem (P2). Moreover, such a KKT point also satisfies the SOSC of (P1) and the strict complementarity condition. However, in practice, whatever conditions we assume should be referred to the original problem (P1). So, we now show that the converse implication also holds. Observe that in this case, the strict complementarity condition is required.

Proposition 3.8. *Let $(x, \lambda) \in \mathbb{R}^{n+m}$ be a KKT pair of* (P1) *that satisfies SOSC and the strict complementarity. Then, there exists $y \in \mathbb{R}^m$ such that (x, y, λ) is a KKT triple of* (P2), *and any such triple satisfies SOSC.*

Proof. From Proposition 3.1, it is sufficient to show that (x, y, λ) satisfies SOSC of (P2). Note that (P1.2) implies $\mathcal{I}_{0N} = \emptyset$. This fact, together with the strict complementarity condition, shows that $\{1, \ldots, m\} = \mathcal{I}_{0P} \cup \mathcal{I}_{P0}$. Now, let $(v, w) \in \mathbb{R}^{n+m}$ be an arbitrary nonzero vector such that

$$\begin{aligned} \langle \nabla g_i(x), v \rangle &= 0, \quad i \in \mathcal{I}_{0P}, \\ \langle \nabla g_i(x), v \rangle - 2y_i w_i &= 0, \quad i \in \mathcal{I}_{P0}. \end{aligned} \tag{3.3}$$

From Proposition 3.4, we have to show that (3.1) holds.

First, let us consider the case where $v \neq 0$. From the SOSC of (P1), we clearly obtain $\langle \nabla_x^2 L(x, \lambda)v, v \rangle > 0$. Also, for any $w_i \in \mathbb{R}$, $\lambda_i w_i^2 = 0$ when $i \in \mathcal{I}_{P0}$, and $\lambda_i w_i^2 \geq 0$ when $i \in \mathcal{I}_{0P}$. Then, we conclude that

$$\langle \nabla_x^2 L(x, \lambda)v, v \rangle + 2 \sum_{i=1}^{m} \lambda_i w_i^2 > 0,$$

which means that the SOSC of (P2) is satisfied in this case.

Now, consider the case where $v = 0$ and $w \in \mathbb{R}^m$ is an arbitrary nonzero vector satisfying (3.3). Then, once again from Proposition 3.4, we have to prove that

$$\sum_{i=1}^{m} \lambda_i w_i^2 = \sum_{i \in \mathcal{I}_{0P} \cup \mathcal{I}_{P0}} \lambda_i w_i^2 > 0 \qquad (3.4)$$

for all nonzero $w \in \mathbb{R}^m$ such that $y_i w_i = 0$ for all $i \in \mathcal{I}_{P0}$. Since $y_i \neq 0$ in this case, we have to show that (3.4) holds for all nonzero $w \in \mathbb{R}^m$ such that

$$w_i = 0 \quad \text{for all} \quad i \in \mathcal{I}_{P0}. \qquad (3.5)$$

Note that if $\mathcal{I}_{0P} = \emptyset$ or, in other words, $\mathcal{I}_{P0} = \{1, \ldots, m\}$, then there exists no $w \neq 0$ satisfying (3.5). So, the condition (3.4) holds vacuously. Thus, let $\mathcal{I}_{0P} \neq \emptyset$, and choose an arbitrary $w \neq 0$ satisfying (3.5). For such a vector w, there exists an index j with $w_j \neq 0$. Since $j \in \mathcal{I}_{0P}$, we have $\lambda_j w_j^2 > 0$, which clearly implies (3.4). We then conclude that the SOSC of (P2) holds in this case. $\qquad \square$

4. Equivalence Between the Regularity Conditions

We now proceed with results concerning the regularity conditions. We recall that under the *linear independence constraint qualification* (LICQ), the KKT conditions are necessary for optimality. Moreover, we say that the LICQ condition of an NLP problem holds at a point if the gradients of the equality constraints and the gradients of active inequality constraints are linearly independent (see, for example, [2, Section 3.3] or [7, Section 12.1]).

Proposition 4.1. *Let $(x, y, \lambda) \in \mathbb{R}^{n+2m}$ be a KKT triple of (P2) and assume that it satisfies LICQ and SOSC. Then, (x, λ) is a KKT pair of (P1) that satisfies LICQ.*

Proof. From Proposition 3.6, (x, λ) is a KKT pair of (P1). We have to prove that (x, λ) satisfies LICQ of (P1), which means that the gradients of active constraints $\nabla g_i(x)$, $i \in \mathcal{I}_{0P} \cup \mathcal{I}_{00}$ are linearly independent. Since (x, y, λ) satisfies LICQ of (P2), the matrix $[Jg(x), -2\mathrm{diag}(y)]$ has linearly independent rows. Without loss of generality, we can write this matrix as

$$\begin{bmatrix} Jg_{\mathcal{I}_{0P} \cup \mathcal{I}_{00}}(x) & 0 & 0 \\ Jg_{\mathcal{I}_{P0}}(x) & 0 & -2\mathrm{diag}(y_i)_{i \in \mathcal{I}_{P0}} \end{bmatrix}$$

where $Jg_{\mathcal{I}_{0P} \cup \mathcal{I}_{00}}(x)$ and $Jg_{\mathcal{I}_{P0}}(x)$ denote the part of the Jacobian $Jg(x)$ with indices in $\mathcal{I}_{0P} \cup \mathcal{I}_{00}$ and $\mathcal{I}_{P0}$, respectively. Observe also that $\mathrm{diag}(y_i)_{i \in \mathcal{I}_{P0}}$ is nonsingular. Then, we conclude that the rows of $Jg_{\mathcal{I}_{0P} \cup \mathcal{I}_{00}}(x)$ are linearly independent, which is precisely the LICQ condition of (P1). $\qquad \square$

Proposition 4.2. *Let $(x, \lambda) \in \mathbb{R}^{n+m}$ be a KKT pair of (P1) and assume that it satisfies LICQ. Then, there exists $y \in \mathbb{R}^m$ such that (x, y, λ) is a KKT triple of (P2), and any such triple satisfies LICQ.*

Proof. From Proposition 3.1, it is sufficient to prove that (x, y, λ) satisfies LICQ of (P2). Assume, for the purpose of contradiction, that (x, y, λ) does not satisfy LICQ for (P2). Then, there exist α_i, $i = 1, \ldots, m$, not all zero such that

$$\sum_{i=1}^{m} \alpha_i \nabla g_i(x) = 0 \quad \text{and} \quad \alpha_i y_i = 0, \ i = 1, \ldots, m.$$

The latter equalities show that $\alpha_i = 0$ when $i \in \mathcal{I}_{P0}$. So, recalling that $\{1, \ldots, m\} = \mathcal{I}_{0P} \cup \mathcal{I}_{00} \cup \mathcal{I}_{P0}$, there exist α_i, $i \in \mathcal{I}_{0P} \cup \mathcal{I}_{00}$, not all zero such that

$$\sum_{i \in \mathcal{I}_{0P} \cup \mathcal{I}_{00}} \alpha_i \nabla g_i(x) = 0.$$

But this contradicts the LICQ condition of (P1), and so (x, y, λ) satisfies LICQ of (P2). $\qquad \square$

Summarizing the above discussions and the results of Section 3, we state the main result about the squared slack variables approach.

Theorem 4.3. *The following statements hold.*

(a) *Let $(x, \lambda) \in \mathbb{R}^{n+m}$ be a KKT pair of* (P1). *Assume that it satisfies LICQ, SOSC and the strict complementarity. Then, there exists $y \in \mathbb{R}^m$ such that (x, y, λ) is a KKT triple of* (P2), *and any such triple satisfies LICQ and SOSC.*

(b) *Let $(x, y, \lambda) \in \mathbb{R}^{n+2m}$ be a KKT triple of* (P2). *Assume that it satisfies LICQ and SOSC. Then, (x, λ) is a KKT pair of* (P1) *satisfying LICQ, SOSC and the strict complementarity.*

Proof. The item (a) follows from Propositions 3.8 and 4.2, and the item (b) follows from Propositions 3.7 and 4.1. $\qquad \square$

5. Final Remarks

We have analyzed the use of squared slack variables in the context of NLP. We have proved that, under the second-order sufficient conditions and the regularity conditions, KKT points of the original and the reformulated problems are essentially equivalent. A future research topic is to see if other conditions, that appear frequently in convergence analysis of optimization methods, can be considered instead of the second-order sufficient condition. In fact, from the proof of Proposition 3.6, we observe that in order to obtain the equivalence of the KKT points, it is sufficient to have $\mathcal{I}_{0N} = \emptyset$. From Lemma 3.5, it means that the SOSC assumption for (P2) is strong in the sense that it also gives $\mathcal{I}_{00} = \emptyset$. A similar question also arises in more general contexts, such as the nonlinear second-order cone programming and the nonlinear semidefinite programming problems, and should be a matter of investigation.

References

[1] P. Armand and D. Orban: The squared slacks transformation in nonlinear programming. *SQU Journal for Science*, **17** (2012), 22–29.

[2] D.P. Bertsekas: *Nonlinear Programming*, 2nd Edition (Athena Scientific, 1999).

[3] E.H. Fukuda and M. Fukushima: The use of squared slack variables in nonlinear second-order cone programming. *Journal of Optimization Theory and Applications*, **170** (2016), 394–418.

[4] H.Th. Jongen and O. Stein: On the complexity of equalizing inequalities. *Journal of Global Optimization*, **27** (2003), 367–374.

[5] B.F. Lourenço, E.H. Fukuda, and M. Fukushima: Optimality conditions for nonlinear semidefinite programming via squared slack variables. To appear in *Mathematical Programming* (First Online: June 25, 2016).

[6] S.G. Nash: SUMT (revisited). *Operations Research*, **46** (1998), 763–775.

[7] J. Nocedal and S.J. Wright: *Numerical Optimization*, 2nd Edition (Springer Verlag, New York, 2006).

[8] S.M. Robinson: Stability theory for systems of inequalities, part II: Differentiable nonlinear systems. *SIAM Journal on Numerical Analysis*, **13** (1976), 497–513.

[9] E. Spedicato: On a Newton-like method for constrained nonlinear minimization via slack variables. *Journal of Optimization Theory and Applications*, **36** (1982), 175–190.

[10] R.A. Tapia: A stable approach to Newton's method for general mathematical programming problems in $\mathbb{R}^n$. *Journal of Optimization Theory and Applications*, **14** (1974), 453–476.

[11] R.A. Tapia: On the role of slack variables in quasi-Newton methods for constrained optimization. In L.C.W. Dixon and G.P. Szegö (eds.): *Numerical Optimisation of Dynamic Systems* (North-Holland Publishing Company, 1980), 235–246.

Ellen Hidemi Fukuda
Graduate School of Informatics
Kyoto University
Sakyo-ku Yoshida Honmachi
Kyoto 606-8501, JAPAN
E-mail: `ellen@i.kyoto-u.ac.jp`