Journal of the Operations Research Society of Japan
Vol. 63, No. 4, October 2020, pp. 114–133

LIMIT OPERATION IN PROJECTIVE SPACE FOR CONSTRUCTING NECESSARY OPTIMALITY CONDITION OF POLYNOMIAL OPTIMIZATION PROBLEM

Tomoyuki Iori Toshiyuki Ohtsuka
Kyoto University

(Received November 26, 2019; Revised February 28, 2020)

Abstract This paper proposes a necessary optimality condition derived by a limit operation in projective space for optimization problems of polynomial functions with constraints given as polynomial equations. The proposed condition is more general than the Karush-Kuhn-Tucker (KKT) conditions in the sense that no constraint qualification is required, which means the condition can be viewed as a necessary optimality condition for every minimizer. First, a sequential optimality condition for every minimizer is introduced on the basis of the quadratic penalty function method. To perform a limit operation in the sequential optimality condition, we next introduce the concept of projective space, which can be regarded as a union of Euclidian space and its points at infinity. Through the projective space, the limit operation can be reduced to computing a point of the tangent cone at the origin. Mathematical tools from algebraic geometry were used to compute the set of equations satisfied by all points in the tangent cone, and thus by all minimizers. Examples are provided to clarify the methodology and to demonstrate cases where some local minimizers do not satisfy the KKT conditions.

Keywords: Nonlinear programming, optimality condition, Karush-Kuhn-Tucker conditions

1. Introduction

Optimality conditions are fundamental to nonlinear programming (NLP). They theoretically play an essential role in the analysis of optimization problems and practically provide important tools to solve optimization problems. A variety of optimality conditions have been investigated for optimization problems with both equality and inequality constraints [3, 5, 12, 17]. Among these, the Karush-Kuhn-Tucker (KKT) conditions are most popular and have become the basis of many numerical algorithms for solving NLP problems, such as Newton's method, the interior point method, and the augmented Lagrangian method [6, 19].

The KKT conditions cannot function as necessary optimality conditions on their own; they need an additional condition on minimizers called a constraint qualification (CQ). For example, the first-order necessary condition for optimality consists of the KKT conditions and the linear independence CQ (LICQ), where it is assumed that the gradients of all equality constraints and active inequality constraints are linearly independent at minimizers. There are various CQs, including the Slater CQ (SCQ), Mangasarian-Fromovitz CQ (MFCQ), Abadie's CQ (ACQ), and Guignard's CQ (GCQ) [12, 20]. GCQ is in a sense the weakest CQ that ensures the KKT conditions are necessary optimality conditions [13]. In other words, a local minimizer that violates the GCQ cannot be obtained by solving the KKT conditions.

Let us give a simple example in which the KKT conditions are no longer necessary conditions for optimality. Consider an NLP problem that minimizes a cost function $x_1 + x_2$ under equality constraints $(x_1-1)^2+x_2^2-1 = 0$ and $(10x_1-8)^2+(5x_2)^2-64 = 0$. The feasible set of this problem consists of three points, which are shown along with the contours of the cost function located at them in Figure 1. Obviously, this problem has a global minimizer

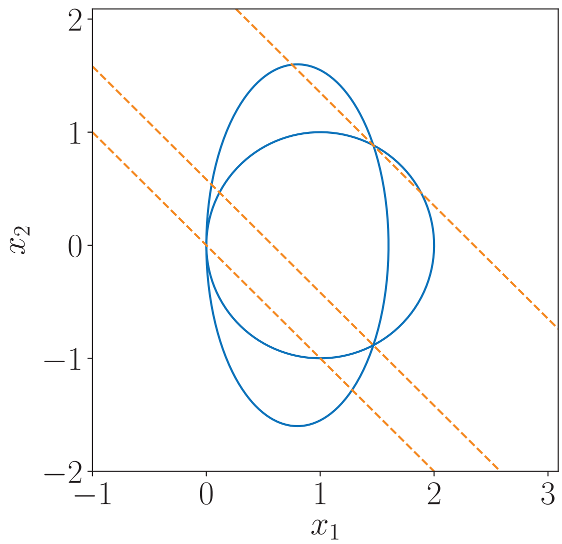


Figure 1: Feasible set (intersections of solid curves) and contours of cost function (dashed lines)

at $(x_1, x_2) = (0, 0)$. However, at this minimizer, the constraints violate the GCQ, and hence the minimizer does not satisfy the KKT conditions. For the following discussion, we say that a minimizer is *non-KKT type* if it does not satisfy the KKT conditions. Note that there are many optimization methods based on the KKT conditions (such as the augmented Lagrangian method) and that none of them can find non-KKT type minimizers.

As necessary optimality conditions for all local minimizers, including non-KKT type ones, Andreani et al. [2] proposed the approximate-KKT (AKKT) conditions. Roughly speaking, these conditions claim that each local minimizer is the convergence point of a sequence consisting of points that approximately satisfy the KKT conditions. Such optimality conditions described by the existence of a convergent sequence are called *sequential optimality conditions* [2] and are the subject of intensive investigation these days [1–3, 17]. Although AKKT conditions are especially useful for giving a termination criterion in numerical optimization algorithms [2, 14], they are difficult to verify for a given candidate. Indeed, to determine whether the given candidate can be a minimizer or not, we have to show *the existence of a sequence*, which is more difficult than, for example, showing whether or not the KKT conditions are satisfied at the candidate. From this viewpoint, it is important to develop *a necessary condition described by equations like the KKT conditions but satisfied even by non-KKT type minimizers.*

To find such a new necessary optimality condition, we first consider the quadratic penalty function method, which is a conventional method to relax a constrained NLP into an unconstrained one. With this method, if global minimizers of the penalty function converge to a point as a penalty parameter goes to infinity, the convergence point is a global minimizer of the original NLP. This convergence property does not assume any CQ, so the property is valid even for non-KKT type global minimizers. This is why we select the quadratic penalty function method as the starting point for finding a new necessary optimality condition. Indeed, it has previously been proven by using a certain penalty function method that the AKKT conditions are valid for all minimizers [2]. More precisely, that proof guarantees the existence of a sequence for each local minimizer that converges to the minimizer and that consists of stationary points of a certain penalty function. This is the basis of our proposed condition.

The stationary points of a penalty function can be viewed as functions of the penalty parameter defined by an implicit function representation consisting of the stationary conditions. This leads us to consider the limit operation of the penalty parameter to infinity in the stationary conditions. Note that just substituting infinity for the penalty parameter is meaningless because this ends up neglecting the terms from the cost function in the stationary conditions. Therefore, we utilize symbolic computation to perform the limit operation of the penalty parameter precisely. First, we extend the equations obtained from the stationary conditions defined on a Euclidian space to those defined on a projective space, which can be regarded as a union of a Euclidian space and its points at infinity. Next, we make some changes of variables through the projective space and transform the limit operation to infinity into the limit operation to zero. Finally, we use a tangent cone to precisely perform the limit operation to zero and to obtain the new necessary condition by using symbolic computation techniques from algebraic geometry.

Note that the Fritz-John (FJ) conditions [12] are also satisfied by all local minimizers and described by equations. However, the points that satisfy the FJ conditions often include an infinite number of points that are neither locally optimal nor even KKT. In fact, we can make all feasible points satisfy the FJ conditions by replacing an equality constraint with two inequality constraints or by adding a redundant inequality constraint [4]. In Section 6, we show a case where the new condition has finite solutions even though the FJ conditions have infinite ones.

In this paper, we consider nonlinear constrained optimization problems (COPs) defined as

$$\begin{aligned} &\min_{\boldsymbol{x}} f(\boldsymbol{x}) \\ &\text{s. t. } g(\boldsymbol{x}) = \mathbf{0}, \end{aligned} \tag{1.1}$$

where $\boldsymbol{x} \in \mathbf{R}^n$ is a vector of indeterminates, $f : \mathbf{R}^n \to \mathbf{R}$ is a cost function, and $g : \mathbf{R}^n \to \mathbf{R}^m$ is a vector-valued function describing the constraints. We assume that functions f and g consist of polynomials to utilize symbolic computation techniques from algebraic geometry. We also assume that the feasible set

$$\mathcal{X} := \{\boldsymbol{x} \in \mathbf{R}^n \mid g(\boldsymbol{x}) = \mathbf{0}\}$$

is nonempty throughout this paper. Note that although formulation (1.1) includes only equality constraints, it can also handle inequality constraints by introducing slack variables. Specifically, by introducing a slack variable $s \in \mathbf{R}$, an inequality constraint $\tilde{g}(\boldsymbol{x}) \leq 0$ is converted into an equality constraint $g(\boldsymbol{x}, s) := \tilde{g}(\boldsymbol{x}) + s^2 = 0$. Therefore, we can solve a COP with the equality constraint $g(\boldsymbol{x}, s) = 0$ and indeterminates $[\boldsymbol{x}^{\mathrm{T}}\ s]^{\mathrm{T}} \in \mathbf{R}^{n+1}$ instead of the original COP with inequality. It is well-known that the converted COP is equivalent to the original one in terms of minimizers, and this conversion technique has been used by many researchers from the 1960s [18] onward [5, 11]. For COPs (1.1), we provide a new necessary optimality condition with no CQs that is satisfied by all local minimizers.

This paper is organized as follows. In Section 2 of this paper, we introduce the various notations and concepts of the mathematical tools utilized in this work. Section 3 introduces the quadratic penalty function method and demonstrates the existence of a stationary point sequence that converges to each minimizer of the original COP. In Section 4, we show how the limit of a penalty parameter to infinity can be reduced to the limit of an additional parameter to the origin by utilizing a projective space. This transformation of a problem enables us to precisely compute the limit points of the trajectories that the stationary points of a penalty function move along. After that, in Section 5, we propose an algorithm to symbolically obtain the equations holding at the limit points by means of a tangent cone. In Section 6, examples are provided to illustrate the methodology. We conclude in Section 7 with a brief summary and mention of future work.

Notations

For the field of real numbers $\mathbf{R}$ and a vector $\boldsymbol{x} = [x_1 \cdots x_n]^{\mathrm{T}}$, $\mathbf{R}[\boldsymbol{x}]$ denotes the ring of polynomials in the components of $\boldsymbol{x}$ over $\mathbf{R}$. An ideal generated by polynomials $f_1, \ldots, f_s \in \mathbf{R}[\boldsymbol{x}]$ is defined as $I = \langle f_1, \ldots, f_s \rangle := \{a_1 f_1 + \cdots + a_s f_s \mid a_1, \ldots, a_s \in \mathbf{R}[\boldsymbol{x}]\}$, and the polynomials $f_1, \ldots, f_s$ are called generators of the ideal I. For an ideal $I \subset \mathbf{R}[\boldsymbol{x}]$, $\mathcal{V}(I) \subset \mathbf{R}^n$ is the set of elements in $\mathbf{R}^n$ where all polynomials in I vanish; it is called the algebraic set defined by I. Conversely, for an algebraic set $V \subset \mathbf{R}^n$, $\mathcal{I}(V) \subset \mathbf{R}[\boldsymbol{x}]$ is a set of all the polynomials that vanish everywhere in V; it is called the ideal of V. If an ideal I is generated by $\{f_1, \ldots, f_s\}$, $\mathcal{V}(I)$ equals the set of elements where all the generators $f_1, \ldots, f_s$ vanish [7]. In this case, $\mathcal{V}(I)$ is also denoted by $\mathcal{V}(f_1, \ldots, f_s)$. For a mapping $g(\boldsymbol{x}) = [g_1(\boldsymbol{x}) \ \cdots \ g_m(\boldsymbol{x})]^{\mathrm{T}} : \mathbf{R}^n \to \mathbf{R}^m$, $g_{\boldsymbol{x}}(\boldsymbol{x})$ denotes an $n \times m$ matrix whose (i,j) component consists of $\partial g_j(\boldsymbol{x})/\partial x_i$. In particular, when $m = 1$ holds, or equivalently when $g(\boldsymbol{x})$ is a function, $g_{\boldsymbol{x}}(\boldsymbol{x})$ denotes a column vector $[\partial g(\boldsymbol{x})/\partial x_1 \ \cdots \ \partial g(\boldsymbol{x})/\partial x_n]^{\mathrm{T}}$. $\mathrm{diag}[d_1, \ldots, d_n]$ denotes the diagonal matrix whose diagonal components are $d_1, \ldots, d_n$. Function $f(x, a)$ is also denoted by $f(x; a)$ when we emphasize that a is regarded as a parameter rather than a part of the variables of f.

2. Mathematical Preliminaries

This section introduces the mathematical concepts utilized in this paper. We refer to [7, 8] for most of the definitions and lemmas here.

2.1. Projective space

In this subsection, we introduce a few fundamental notations and concepts of projective geometry, which are useful for dealing with infinities. In projective geometry, we define the equivalence class

$$[L] := \{\text{all lines parallel to a given line } L \subset \mathbf{R}^n\}$$

as *a point at infinity* of $\mathbf{R}^n$. Note that this definition of points at infinity is a straightforward extension of the characterization of points in $\mathbf{R}^n$; namely, a point in $\mathbf{R}^n$ can be characterized by a set of all lines, any two of which cross each other at that point. This definition of points at infinity enables us to integrate $\mathbf{R}^n$ and points at infinity into a projective space $\mathbf{P}^n$, where it is no longer necessary to distinguish points of $\mathbf{R}^n$ from those at infinity.

Definition 2.1 (Projective Space)**.** *An n-dimensional projective space over $\mathbf{R}$ denoted by $\mathbf{P}^n$ is the set of equivalence classes of the equivalence relation $\sim$ on $\mathbf{R}^{n+1} \setminus \{\mathbf{0}\}$, where $\sim$ is defined for $\boldsymbol{x}, \boldsymbol{y} \in \mathbf{R}^{n+1} \setminus \{\mathbf{0}\}$ as*

$$\boldsymbol{x} \sim \boldsymbol{y} \overset{def}{\iff} \exists \lambda \in \mathbf{R} \setminus \{0\} \text{ s.t. } \boldsymbol{x} = \lambda \boldsymbol{y}. \tag{2.1}$$

An equivalence class $p \in \mathbf{P}^n$ is also denoted by $[X_0 : \cdots : X_n]$ with any point $[X_0 \ \cdots \ X_n]^{\mathrm{T}} \in p$, and this representation is called a homogeneous coordinate of p.

In this paper, we use uppercase characters to describe homogeneous coordinates. Note that a projective space is a topological space with the quotient topology induced from the natural topology of $\mathbf{R}^{n+1} \setminus \{\mathbf{0}\}$ by the quotient mapping $\mathbf{R}^{n+1} \setminus \{\mathbf{0}\} \ni [X_0 \ \cdots \ X_n]^{\mathrm{T}} \mapsto p = [X_0 : \cdots : X_n] \in \mathbf{P}^n$.

To show that a projective space $\mathbf{P}^n$ is the union of a Euclidian space $\mathbf{R}^n$ and a set of all points at infinity of $\mathbf{R}^n$, let us consider an open subset U_0 defined as

$$U_0 := \{[X_0 : \cdots : X_n] \in \mathbf{P}^n \mid X_0 \neq 0\} \subset \mathbf{P}^n. \tag{2.2}$$

This open subset can be identified with $\mathbf{R}^n$ because there is a homeomorphism $\phi_0 : \mathbf{R}^n \to U_0$ that maps $[x_1 \cdots x_n]^{\mathrm{T}} \in \mathbf{R}^n$ to $[1 : x_1 : \cdots : x_n] \in U_0$; the inverse mapping ϕ_0^{-1} can be defined as

$$\phi_0^{-1} : U_0 \ni [X_0 : \cdots : X_n] = [1 : X_1/X_0 : \cdots : X_n/X_0] \mapsto [X_1/X_0 \ \cdots \ X_n/X_0]^{\mathrm{T}} \in \mathbf{R}^n.$$

For the complement $H_0 := \mathbf{P}^n \setminus U_0$, each point $[0 : X_1 : \cdots : X_n] \in H_0$ uniquely determines a line L through the origin by $L = \{\boldsymbol{x} = [x_1 \ \cdots \ x_n]^{\mathrm{T}} \in \mathbf{R}^n \mid i \in \{1, \ldots, n\}, t \in \mathbf{R}, x_i = tX_i\}$. We can consider the equivalence class consisting of all lines parallel to L, which defines a point at infinity, and thus there is a bijective correspondence between all points of H_0 and those at infinity of $\mathbf{R}^n \cong U_0$. From this viewpoint, H_0 is called the *hyperplane at infinity* of U_0.

Note that there are many other pairs of subsets that have the same property as (U_0, H_0). Indeed, for each $i = 1, \ldots, n$, we can define subsets U_i and H_i as

$$U_i := \{[X_0 : \cdots : X_n] \in \mathbf{P}^n \mid X_i \neq 0\}, \tag{2.3}$$

$$H_i := \mathbf{P}^n \setminus U_i. \tag{2.4}$$

For each pair, we can identify U_i with $\mathbf{R}^n$ by a homeomorphism $\phi_i : \mathbf{R}^n \to U_i$ defined in the same way as U_0, and under this identification, H_i can be identified with the set of all points at infinity of U_i. The set of n pairs $\{(U_i, \phi_i)\}_{i=1}^n$ can be regarded as an *atlas* of the projective space $\mathbf{P}^n$ when we treat the space as a manifold, and each pair (U_i, ϕ_i) is called a *chart* of $\mathbf{P}^n$.

It can be easily shown that the intersection

$$U_i \cap H_j = \{[X_0 : \cdots : X_n] \in \mathbf{P}^n \mid X_i \neq 0, X_j = 0\}$$

is dense in H_j if and only if $i \neq j$ holds, which indicates each subset U_i contains almost all points at infinity of the other subsets U_j. This means that, by changing the chart from U_j to U_i $(j \neq i)$, we can reduce any computations in the hyperplane at infinity H_j to computations of finite values in $U_i \cong \mathbf{R}^n$.

2.2. Homogenization and dehomogenization

In this paper, we convert a COP into a set of equations and find its solutions where a penalty parameter goes to infinity. The projective space (mentioned in the previous subsection) enables us to treat such solutions as points in $\mathbf{R}^n$ rather than those at infinity. To do this, we extend a polynomial equation over $\mathbf{R}^n$ to that over $\mathbf{P}^n$, which leads to the *homogenizations* of polynomials.

Definition 2.2 (Homogenization of Polynomial). *For a real coefficient polynomial $f \in \mathbf{R}[x_1, \ldots, x_n]$, the homogenization of f is the homogeneous polynomial $f^{\mathrm{hom}} \in \mathbf{R}[X_0, \ldots, X_n]$ defined as*

$$f^{\mathrm{hom}}(X_0, \ldots, X_n) := X_0^d \cdot f\left(\frac{X_1}{X_0}, \ldots, \frac{X_n}{X_0}\right), \tag{2.5}$$

where d is the total degree of f.

Note that a homogenization f^{hom} of degree d is a homogeneous function of degree d because

$$f^{\mathrm{hom}}(\varepsilon X_0, \dots, \varepsilon X_n) = \varepsilon^d f^{\mathrm{hom}}(X_0, \dots, X_n)$$

holds. Therefore, if a homogeneous coordinate of $p \in \mathbf{P}^n$ satisfies $f^{\mathrm{hom}} = 0$, all homogeneous coordinates of p also satisfy the same equation, and thus it makes sense to consider the subset of $\mathbf{P}^n$ where f^{hom} vanishes. We also call this subset an algebraic set defined by f^{hom} and denote it by $\mathcal{V}(f^{\mathrm{hom}})$. The relationship between the algebraic set of the homogenization f^{hom} and that of the original polynomial f is

$$\mathcal{V}(f) = \mathcal{V}(f^{\mathrm{hom}}) \cap U_0,$$

which indicates $\mathcal{V}(f^{\mathrm{hom}}) \subset \mathbf{P}^n$ is an extension of $\mathcal{V}(f) \subset \mathbf{R}^n \cong U_0$.

Conversely, we can define a polynomial defined on U_0 corresponding to any homogeneous polynomial defined on $\mathbf{P}^n$, which is called *dehomogenization*.

Definition 2.3 (Dehomogenization of Homogeneous Polynomial)**.** *For a homogeneous polynomial $g(X_0, \dots, X_n) \in \mathbf{R}[X_0, \dots, X_n]$, the dehomogenization of g is a polynomial $g^{\mathrm{deh}} \in \mathbf{R}[x_1, \dots, x_n]$ defined as*

$$g^{\mathrm{deh}}(x_1, \dots, x_n) := g(1, x_1, \dots, x_n). \tag{2.6}$$

Note that we can consider the dehomogenization of g for any index $i = 0, \dots, n$ in the same way, although for simplicity we do not indicate i explicitly, and i is specified in accordance with the context.

It is readily known that there is also the relationship

$$\mathcal{V}(g^{\mathrm{deh}}) = \mathcal{V}(g) \cap U_0.$$

Note that, in general, a homogeneous polynomial f and the homogenization of its dehomogenization $(f^{\mathrm{deh}})^{\mathrm{hom}}$ may be different from each other. For example, a homogeneous polynomial

$$g(X_0, X_1, X_2) = X_0^3 X_1^2 + X_0^2 X_2^3$$

is different from

$$(g^{\mathrm{deh}})^{\mathrm{hom}}(X_0, X_1, X_2) = X_0 X_1^2 + X_2^3.$$

Remark 2.1. *Although $X_0 = 1$ is substituted into g in the definition* (2.6)*, it is consistent with the definition of a point in $\mathbf{P}^n$; for any homogeneous coordinate $[X_0 : \cdots : X_n]$ of a point $p \in \mathbf{P}^n$ (where X_0 is not necessarily equal to one), we can obtain the other homogeneous coordinate $[1 : X_1/X_0 : \cdots : X_n/X_0]$, which represents the same point p. From this viewpoint, dehomogenization corresponds to the changes of variables $x_i = X_i/X_0$ for $i = 1, \dots, n$. Indeed, applying changes of variables $X_i = x_i X_0$ to $g(X_0, \dots, X_n)$ of degree d, we obtain*

$$g(X_0, \dots, X_n) = X_0^d g^{\mathrm{deh}}(x_1, \dots, x_n),$$

which indicates that $g = 0$ is equivalent to $g^{\mathrm{deh}} = 0$ unless X_0 is equal to zero. This equation is the same as the definition of homogenization (2.5) *except that d is the total degree of g but not of g^{deh}.*

2.3. Tangent cone

In the penalty function method, we compute the limit solutions of infinite sequences of unconstrained optimization problems as the penalty parameter goes to infinity. In projective space, this divergence to infinity can be interpreted as a convergence to a point by changing the chart we focus on. Therefore, only the local information at the convergence point is needed to find the solutions, which means we can take advantage of a tangent cone at the convergence point. Throughout this subsection, an algebraic set V is assumed to include the origin $\mathbf{0}$.

The definition of a tangent cone is as follows [7].

Definition 2.4. *For an algebraic set $V = \mathcal{V}(f_1, \ldots, f_s) \subset \mathbf{R}^n$, its tangent cone at $\mathbf{0} \in V \subset \mathbf{R}^n$ is an algebraic set $C_{\mathbf{0}}(V) \subset \mathbf{R}^n$ defined as*

$$C_{\mathbf{0}}(V) := \mathcal{V}(\{f^{\min} \mid f \in \mathcal{I}(V)\}), \tag{2.7}$$

where $f^{\min}$ denotes the homogeneous component of the lowest degree in f, i.e., the sum of all the terms of f whose degrees are equal to the lowest degree of f.

Note that this definition indicates the set of polynomials defining a tangent cone consists of homogeneous polynomials. As can be seen in the above definition, a tangent cone is an algebraic set, which means there exists an ideal $J = \langle g_1, \ldots, g_t \rangle \subset \mathbf{R}[x_1, \ldots, x_n]$ that defines the tangent cone. There are various algorithms to compute generators of J [8, 9, 15, 16], and most of them use a Gröbner basis, which is a set of generators that has good properties for symbolic computations (see [7] for details). The following lemma gives us one of the methods to compute generators of J by using a Gröbner basis with respect to an elimination order [7] for X_0.

Lemma 2.1 (The Tangent Cone Algorithm [9, 15])**.** *Consider an ideal $I \subset \mathbf{R}[x_1, \ldots, x_n]$ generated by polynomials $f_1, \ldots, f_s$ and assume the algebraic set $\mathcal{V}(I) \subset \mathbf{R}^n$ includes the origin $\mathbf{0}$. Let $\{G_1, \ldots, G_t\}$ be a Gröbner basis of an ideal generated by the homogenizations $f_1^{\mathrm{hom}}, \ldots, f_s^{\mathrm{hom}}$ with respect to an elimination order for X_0. Then, for the dehomogenizations $G_1^{\mathrm{deh}}, \ldots, G_t^{\mathrm{deh}} \in \mathbf{R}[x_1, \ldots, x_n]$, the following equation holds:*

$$C_{\mathbf{0}}(\mathcal{V}(I)) = \mathcal{V}\left((G_1^{\mathrm{deh}})^{\min}, \ldots, (G_t^{\mathrm{deh}})^{\min}\right), \tag{2.8}$$

where $(G_i^{\mathrm{deh}})^{\min}$ denotes the homogeneous component of the lowest degree in G_i^{deh}.

Although Definition 2.4 consists of only algebraic statements, it is also useful to clarify the characterization of a tangent cone from the geometric viewpoint. Before that, let us introduce the following definition.

Definition 2.5. *We say that a sequence of lines $\{L_k\}_{k=1}^{\infty} \subset \mathbf{R}^n$ through $\mathbf{0}$ converges to a line L also through $\mathbf{0}$ if, for a given parametrization $L = \{t\boldsymbol{v} \in \mathbf{R}^n \mid \boldsymbol{v} \in \mathbf{R}^n, t \in \mathbf{R}\}$, there exist parametrizations of $L_k = \{t\boldsymbol{v}_k \in \mathbf{R}^n \mid \boldsymbol{v}_k \in \mathbf{R}^n, t \in \mathbf{R}\}$ such that $\lim_{k\to\infty} \boldsymbol{v}_k = \boldsymbol{v}$ holds.*

By using this definition, we can show the characterization of a tangent cone as Lemma 2.2.

Lemma 2.2. *A line $L \subset \mathbf{R}^n$ is called a* secant line *of an algebraic set $V \subset \mathbf{R}^n$ if the intersection of L and V consists of more than two points. Let $L_\infty \subset \mathbf{R}^n$ be a line through $\mathbf{0} \in V$. L_∞ is then a subset of the tangent cone $C_{\mathbf{0}}(V)$ if and only if there exists a sequence $\{\boldsymbol{q}_k\}_{k=1}^{\infty} \subset V \setminus \{\mathbf{0}\}$ such that $\lim_{k\to\infty} \boldsymbol{q}_k = \mathbf{0}$ holds and the sequence of secant lines $\{L_k \subset \mathbf{R}^n \mid \mathbf{0}, \boldsymbol{q}_k \in L_k \cap V\}_{k=1}^{\infty}$ converges to the line L_∞ as $k \to \infty$.*

3. Penalty Function Method and its Convergence Property

The quadratic penalty function $P(\boldsymbol{x}; r)$ for COP (1.1) is defined as

$$P(\boldsymbol{x}; r) := f(\boldsymbol{x}) + r g^{\mathrm{T}}(\boldsymbol{x}) g(\boldsymbol{x}), \tag{3.1}$$

where $r \in \mathbf{R}$ is a penalty parameter. It is well-known that the convergence points of the global minimizers of $P(\boldsymbol{x}; r)$, if any, are also the global minimizers of the original COP [5, 19]. As a more practical property of the quadratic penalty function method, such convergence property for local minimizers has been also established. Let X^* be an isolated compact set of local minimizers of COP (1.1) that corresponds to a certain local minimum value. It is proven that there exists a sequence of local minimizers of $P(\boldsymbol{x}; r)$ that converges to a *certain* point of X^* [10]. As a straightforward consequence of this convergence property, if X^* consists of only one minimizer, the existence of the sequence converging to the minimizer is guaranteed. However, if this is not the case, that is, if X^* consists of non-isolated minimizers, the existence of such a sequence is not guaranteed for *each* minimizer of X^*.

As necessary optimality conditions for all minimizers, the AKKT conditions were proposed in [2]. Roughly speaking, AKKT conditions state that, for any local minimizer, there exists a sequence of points converging to the local minimizer, where all points in the sequence approximately satisfy the KKT conditions. In other words, the existence of such a sequence is a necessary condition for all minimizers including even non-isolated ones. This is proved by using the convergence property of the Internal-External Penalty method in [10], and the proof gives a viewpoint on such a sequence not from the KKT conditions but from the penalty function method, which is the basis of our result.

Let us define a localized penalty function, which is localized to a specific local minimizer $\hat{\boldsymbol{x}}$:

$$P^{\mathrm{loc}}(\boldsymbol{x}; r, \hat{\boldsymbol{x}}) := f(\boldsymbol{x}) + \|\boldsymbol{x} - \hat{\boldsymbol{x}}\|^2 + r g^{\mathrm{T}}(\boldsymbol{x}) g(\boldsymbol{x}). \tag{3.2}$$

The reason we call it localized is as follows. For any local minimizer $\hat{\boldsymbol{x}}$, there exists a real number $\varepsilon > 0$ such that a minimization problem restricted on a ball $\mathcal{B}_\varepsilon(\hat{\boldsymbol{x}}) := \{\boldsymbol{x} \in \mathbf{R}^n \mid \|\boldsymbol{x} - \hat{\boldsymbol{x}}\| \leq \varepsilon\} \subset \mathbf{R}^n$, as

$$\begin{aligned} \min_{\boldsymbol{x} \in \mathcal{B}_\varepsilon(\hat{\boldsymbol{x}})} \ & f(\boldsymbol{x}) + \|\boldsymbol{x} - \hat{\boldsymbol{x}}\|^2 \\ \text{s. t. } & g(\boldsymbol{x}) = \mathbf{0}, \end{aligned} \tag{3.3}$$

has a *unique* local minimizer $\boldsymbol{x} = \hat{\boldsymbol{x}}$, and $P^{\mathrm{loc}}(\boldsymbol{x}; r, \hat{\boldsymbol{x}})$ is the penalty function for this minimization problem localized to $\hat{\boldsymbol{x}}$. By using this localized penalty function, a necessary optimality condition described by the existence of a stationary point sequence can be derived as follows.

Theorem 3.1. *Let $\{r_k\}_{k=1}^\infty$ be a sequence monotonically going to infinity as k does, $\hat{\boldsymbol{x}}$ be a minimizer of the original COP* (1.1)*, and $P^{\mathrm{loc}}(\boldsymbol{x}; r, \hat{\boldsymbol{x}})$ be a function defined as* (3.2) *with domain of definition $\boldsymbol{R}^n$. Then, there exists an integer $k^* > 0$ such that there exists a sequence of stationary points $\{\boldsymbol{x}_k\}_{k=k^*}^\infty$ satisfying*

$$P^{\mathrm{loc}}_{\boldsymbol{x}}(\boldsymbol{x}_k; r_k, \hat{\boldsymbol{x}}) = \mathbf{0} \tag{3.4}$$

and converging to $\hat{\boldsymbol{x}}$ as k goes to infinity. In other words, the existence of such a sequence $\{\boldsymbol{x}_k\}_{k=k^*}^\infty$ is a necessary condition for $\hat{\boldsymbol{x}}$ to be a local minimizer.

Proof. This proof is a part of the proof of Theorem 2.1 in [2] and is included here for the completeness of the paper. Let ε be such that the localized minimization problem (3.3) has a unique local minimizer and $\{r_k\}_{k=1}^\infty$ be a sequence of positive real numbers monotonically going to infinity as k does. Since $\hat{\boldsymbol{x}}$ is an interior point of $\mathcal{B}_\varepsilon$, there exists an open subset $U \subset \mathcal{B}_\varepsilon$ including $\hat{\boldsymbol{x}}$. According to the generally accepted result for the quadratic penalty function method [5], there exists a global minimizer sequence $\{\boldsymbol{x}_k\}_{k=1}^\infty \subset \mathcal{B}_\varepsilon$ of the penalty function (3.2) that converges to the global minimizer $\hat{\boldsymbol{x}} \in \mathcal{B}_\varepsilon$ of the localized COP (3.3). In particular, there exists a sufficiently large k^* such that $\{\boldsymbol{x}_k\}_{k=k^*}^{\infty} \subset U$ holds. In the problem settings of this paper, the penalty function (3.2) is continuous, and thus every $\boldsymbol{x}_k$ for all $k \geq k^*$ satisfies the stationary condition (3.4). □

Remark 3.1. *Since the sequence* $\{\boldsymbol{x}_k\}_{k=k^*}^{\infty}$ *in Theorem 3.1 is also an infinite sequence, if we replace its subscripts appropriately, we can assume* $k^* = 1$ *without loss of generality.*

Remark 3.2. *Note that although constraint* $\boldsymbol{x} \in \mathcal{B}_\varepsilon(\hat{\boldsymbol{x}})$ *is imposed in the localized minimization problem* (3.3)*, this is no longer assumed in Theorem 3.1. Due to the lacking this constraint, equation* (3.4) *can have a solution* $\tilde{\boldsymbol{x}}_k$ *that does not belong to* $\mathcal{B}_\varepsilon(\hat{\boldsymbol{x}})$ *and hence does not converge to* $\hat{\boldsymbol{x}}$*. However, this fact has nothing to do with the existence of the sequence* $\{\boldsymbol{x}_k\}_{k=k^*}^{\infty}$ *introduced in the proof, and Theorem 3.1 still gives a necessary optimality condition to be satisfied by every local minimizer* $\hat{\boldsymbol{x}}$*.*

The stationary condition $P_{\boldsymbol{x}}^{\mathrm{loc}}(\boldsymbol{x}; r, \hat{\boldsymbol{x}}) = \mathbf{0}$ is written as

$$P_{\boldsymbol{x}}^{\mathrm{loc}}(\boldsymbol{x}; r, \hat{\boldsymbol{x}}) = f_{\boldsymbol{x}}(\boldsymbol{x}) + 2(\boldsymbol{x} - \hat{\boldsymbol{x}}) + r g_{\boldsymbol{x}}(\boldsymbol{x}) g(\boldsymbol{x}) = \mathbf{0}. \tag{3.5}$$

When we try to use Theorem 3.1 to find a minimizer of COP (1.1), we have to solve equation (3.5) to find the sequence of stationary points $\{\boldsymbol{x}_k\}$. However, equation (3.5) is defined using the minimizer $\hat{\boldsymbol{x}}$, which we are seeking now. This deadlock can be avoided by regarding the minimizer $\hat{\boldsymbol{x}}$ as an additional variable $\boldsymbol{y} \in \boldsymbol{R}^n$, that is, regarding equation (3.5) as an equation defined on $\mathbf{R}^{2n+1}$:

$$P_{\boldsymbol{x}}^{\mathrm{loc}}(\boldsymbol{x}; r, \boldsymbol{y}) = f_{\boldsymbol{x}}(\boldsymbol{x}) + 2(\boldsymbol{x} - \boldsymbol{y}) + r g_{\boldsymbol{x}}(\boldsymbol{x}) g(\boldsymbol{x}) = \mathbf{0}, \tag{3.6}$$

where $\boldsymbol{x} = [x_1 \ \cdots \ x_n]^{\mathrm{T}}$ and $\boldsymbol{y} = [y_1 \ \cdots \ y_n]^{\mathrm{T}}$. By computing the convergence points of sequences $\{\boldsymbol{x}_k\}$ for all $\boldsymbol{y} \in \boldsymbol{R}^n$ and selecting pairs $(\boldsymbol{x}_\infty, \boldsymbol{y})$ such that $\boldsymbol{x}_\infty = \boldsymbol{y}$ holds, we can find candidates of minimizers that satisfy the necessary optimality condition stated as Theorem 3.1. For all $\boldsymbol{y} \in \mathbf{R}^n$ admitting the sequences $\{r_k\}_{k=1}^{\infty}$ and $\{\boldsymbol{x}_k\}_{k=1}^{\infty}$ that satisfy equation (3.6) for every k, symbolic computation enables us to compute a set of equations satisfied by every limit $\boldsymbol{x}_\infty$. The derived equations can be solved without iteratively increasing the penalty parameter, unlike other numerical algorithms for the penalty function method.

4. Limit Points in Projective Space

In this section, we utilize the concepts of projective space, homogenization, and dehomogenization to deal with the limit operation of the penalty parameter symbolically. When we regard equation (3.6) as an equation defined on $\mathbf{R}^{2n+1}$, the convergence points of its solutions as r goes to infinity lie in the hyperplane at infinity of $\mathbf{R}^{2n+1}$. In this section, we fix certain sequences ($\{r_k\}_{k=1}^{\infty}$ and $\{\boldsymbol{x}_k\}_{k=1}^{\infty}$) such that r_k monotonically goes to infinity as k does, $\boldsymbol{x}_k$ converges to a point $\boldsymbol{x}_\infty \in \mathbf{R}^n$, and all points $\left[\boldsymbol{x}_k^{\mathrm{T}} \ r_k\right]^{\mathrm{T}} \in \mathbf{R}^{n+1}$ satisfy equation (3.6) for a fixed $\boldsymbol{y} \in \mathbf{R}^n$. Moreover, for the following discussion, we denote the set consisting of all the components of $P_{\boldsymbol{x}}^{\mathrm{loc}}(\boldsymbol{x}; r, \boldsymbol{y})$ by F, that is,

$$F := \left\{ \left[P_{\boldsymbol{x}}^{\mathrm{loc}}(\boldsymbol{x}; r, \boldsymbol{y})\right]_1, \ldots, \left[P_{\boldsymbol{x}}^{\mathrm{loc}}(\boldsymbol{x}; r, \boldsymbol{y})\right]_n \right\} \subset \mathbf{R}[\boldsymbol{x}, \boldsymbol{y}, r], \tag{4.1}$$

where $\left[P_{\boldsymbol{x}}^{\mathrm{loc}}(\boldsymbol{x}; r, \boldsymbol{y})\right]_i$ is the i-th component of $P_{\boldsymbol{x}}^{\mathrm{loc}}(\boldsymbol{x}; r, \boldsymbol{y})$.

Let us consider a projective space $\mathbf{P}^{2n+1}$ and identify $\mathbf{R}^{2n+1}$ with an open subset,

$$U_0 := \{[X_0 : \cdots : X_{2n+1}] \in \mathbf{P}^{2n+1} \mid X_0 \neq 0\}, \tag{4.2}$$

by a homeomorphism $\phi_0 : \mathbf{R}^{2n+1} \to U_0$ that sends $[x_1 \ \cdots \ x_n \ y_1 \ \cdots \ y_n \ r]^{\mathrm{T}}$ to $[1 : x_1 : \cdots : x_n : y_1 : \cdots : y_n : r]$. As mentioned in subsection 2.1, the other open subset,

$$U_{2n+1} := \{[X_0 : \cdots : X_{2n+1}] \in \mathbf{P}^{2n+1} \mid X_{2n+1} \neq 0\}, \tag{4.3}$$

includes almost all points of the hyperplane at infinity H_0 of U_0; indeed, all the limit points we need. This is because the set difference

$$H_0 \setminus U_{2n+1} = \{[X_0 : \cdots : X_{2n+1}] \in \mathbf{P}^{2n+1} \mid X_0 = X_{2n+1} = 0\}$$

only includes the points at infinity of U_0 where the $(2n+1)$-coordinate, which corresponds to the penalty parameter r, is exactly zero.

There is also a homeomorphism $\phi_{2n+1} : \mathbf{R}^{2n+1} \to U_{2n+1}$ that sends a point $[\rho \ \xi_1 \ \cdots \ \xi_n \ \eta_1 \ \cdots \ \eta_n]^{\mathrm{T}}$ to the equivalent class $[\rho : \xi_1 : \cdots : \xi_n : \eta_1 : \cdots : \eta_n : 1]$. To construct a mapping from $\boldsymbol{x}$-$\boldsymbol{y}$-r space to $\boldsymbol{\xi}$-$\boldsymbol{\eta}$-ρ space, consider two open subsets:

$$D_r := \left\{ \left[\boldsymbol{x}^{\mathrm{T}} \ \boldsymbol{y}^{\mathrm{T}} \ r\right]^{\mathrm{T}} \in \mathbf{R}^{2n+1} \mid r \neq 0 \right\} \subset \mathbf{R}^{2n+1},$$
$$D_\rho := \left\{ \left[\rho \ \boldsymbol{\xi}^{\mathrm{T}} \ \boldsymbol{\eta}^{\mathrm{T}}\right]^{\mathrm{T}} \in \mathbf{R}^{2n+1} \mid \rho \neq 0 \right\} \subset \mathbf{R}^{2n+1},$$

and the restrictions of ϕ_0 and ϕ_{2n+1}^{-1}:

$$\phi_0|_{D_r} : D_r \ni \left[\boldsymbol{x}^{\mathrm{T}} \ \boldsymbol{y}^{\mathrm{T}} \ r\right]^{\mathrm{T}} \mapsto [1 : x_1 : \cdots : x_n : y_1 : \cdots : y_n : r] \in U_0 \cap U_{2n+1},$$
$$\phi_{2n+1}^{-1}|_{U_0 \cap U_{2n+1}} : U_0 \cap U_{2n+1} \ni [X_0 : \cdots : X_{2n+1}] \mapsto \left[\frac{X_0}{X_{2n+1}} \ \cdots \ \frac{X_{2n}}{X_{2n+1}}\right]^{\mathrm{T}} \in D_\rho.$$

Then, a composite mapping $\Phi := \left(\phi_{2n+1}^{-1}|_{U_0 \cap U_{2n+1}}\right) \circ \left(\phi_0|_{D_r}\right)$ is a homeomorphism between D_r and D_ρ defined by

$$\left[\boldsymbol{x}^{\mathrm{T}} \ \boldsymbol{y}^{\mathrm{T}} \ r\right]^{\mathrm{T}} \mapsto \left[\rho \ \boldsymbol{\xi}^{\mathrm{T}} \ \boldsymbol{\eta}^{\mathrm{T}}\right]^{\mathrm{T}} = \Phi\left(\left[\boldsymbol{x}^{\mathrm{T}} \ \boldsymbol{y}^{\mathrm{T}} \ r\right]^{\mathrm{T}}\right) = \frac{1}{r}\left[1 \ \boldsymbol{x}^{\mathrm{T}} \ \boldsymbol{y}^{\mathrm{T}}\right]^{\mathrm{T}}. \tag{4.4}$$

Note that both closures $\overline{D_r}$ and $\overline{D_\rho}$ are equal to $\mathbf{R}^{2n+1}$. Through the homeomorphism Φ, we obtain the triplet of sequences $\{\rho_k\}_{k=1}^{\infty}$, $\{\boldsymbol{\xi}_k\}_{k=1}^{\infty}$, and $\{\boldsymbol{\eta}_k\}_{k=1}^{\infty}$ as the image of sequences $\{r_k\}_{k=1}^{\infty}$ and $\{\boldsymbol{x}_k\}_{k=1}^{\infty}$ under Φ with a fixed $\boldsymbol{y}$.

The sequences $\{r_k\}_{k=1}^{\infty}$ and $\{\boldsymbol{x}_k\}_{k=1}^{\infty}$ satisfy equation (3.6) for a parameter $\boldsymbol{y}$, or in other words, every point $[\boldsymbol{x}_k^{\mathrm{T}} \ \boldsymbol{y}^{\mathrm{T}} \ r_k]^{\mathrm{T}} \in \mathbf{R}^{2n+1}$ belongs to $\mathcal{V}(F)$. This indicates there exists an equation satisfied by $\{\rho_k\}_{k=1}^{\infty}$, $\{\boldsymbol{\xi}_k\}_{k=1}^{\infty}$, and $\{\boldsymbol{\eta}_k\}_{k=1}^{\infty}$ for all $k = 1, \ldots, \infty$, and such equation can be obtained by homogenization and dehomogenization defined in subsection 2.2, as explained below. Let us define a set of homogeneous polynomials $F^{\mathrm{hom}} \subset \mathbf{R}[X_0, \ldots, X_{2n+1}]$ as

$$F^{\mathrm{hom}} := \left\{ f^{\mathrm{hom}} \in \mathbf{R}[X_0, \ldots, X_{2n+1}] \mid f \in F \right\}, \tag{4.5}$$

where $f^{\mathrm{hom}}(X_0, \ldots, X_{2n+1})$ is the homogenization of $f(x_1, \ldots, x_n, y_1, \ldots, y_n, r) \in F$ of total degree d, that is,

$$f^{\mathrm{hom}}(X_0, \ldots, X_{2n+1}) := X_0^d f\left(\frac{X_1}{X_0}, \ldots, \frac{X_{2n+1}}{X_0}\right). \tag{4.6}$$

Its dehomogenizations for index $2n+1$ yield the other polynomial set $\mathcal{F} \subset \mathbf{R}[\rho, \boldsymbol{\xi}, \boldsymbol{\eta}]$:

$$\mathcal{F} := F^{\mathrm{hom}}|_{X_0=\rho, X_1=\xi_1, \ldots, X_n=\xi_n, X_{n+1}=\eta_1, \ldots, X_{2n}=\eta_n, X_{2n+1}=1}. \tag{4.7}$$

The sequences $\{\rho_k\}_{k=1}^{\infty}$, $\{\boldsymbol{\xi}_k\}_{k=1}^{\infty}$, and $\{\boldsymbol{\eta}_k\}_{k=1}^{\infty}$ satisfy the algebraic equations

$$(f^{\mathrm{hom}})^{\mathrm{deh}}(\rho, \xi_1, \dots, \xi_n, \eta_1, \dots, \eta_n) = 0 \qquad \left(\forall (f^{\mathrm{hom}})^{\mathrm{deh}} \in \mathcal{F}\right)$$

because, for all $[\boldsymbol{x}^{\mathrm{T}}\ \boldsymbol{y}^{\mathrm{T}}\ r]^{\mathrm{T}} \in \mathcal{V}(F) \cap D_r$, we have

$$\begin{aligned}(f^{\mathrm{hom}})^{\mathrm{deh}} \circ \Phi\left(\left[\boldsymbol{x}^{\mathrm{T}}\ \boldsymbol{y}^{\mathrm{T}}\ r\right]^{\mathrm{T}}\right) &= (f^{\mathrm{hom}})^{\mathrm{deh}}(r^{-1}, r^{-1}x_1, \dots, r^{-1}x_n, r^{-1}y_1, \dots, r^{-1}y_n) \\ &= f^{\mathrm{hom}}(r^{-1}, r^{-1}x_1, \dots, r^{-1}x_n, r^{-1}y_1, \dots, r^{-1}y_n, 1) \\ &= r^{-d} f(x_1, \dots, x_n, y_1, \dots, y_n, r) \\ &= 0. \end{aligned} \tag{4.8}$$

Now, since the homeomorphism Φ and its inverse Φ^{-1} are both continuous, the following proposition is trivially obtained.

Proposition 4.1. *Let the pair of sequences $\{\rho_k\}_{k=1}^{\infty}$ and $\{\boldsymbol{\xi}_k\}_{k=1}^{\infty}$ be a part of the image of sequences $\{r_k\}_{k=1}^{\infty}$ and $\{\boldsymbol{x}_k\}_{k=1}^{\infty}$ under the homeomorphism Φ with a fixed $\boldsymbol{y}$. Then,*

$$\frac{\boldsymbol{\eta}_k}{\rho_k} = \boldsymbol{y} \tag{4.9}$$

holds for all k. Moreover, for the limit $\boldsymbol{x}_\infty$, sequences $\{r_k\}_{k=1}^{\infty}$ and $\{\boldsymbol{x}_k\}_{k=1}^{\infty}$ satisfy

$$\lim_{k \to \infty} r_k = \infty, \tag{4.10}$$

$$\lim_{k \to \infty} \boldsymbol{x}_k = \boldsymbol{x}_\infty \tag{4.11}$$

if and only if $\{\rho_k\}_{k=1}^{\infty}$ and $\{\boldsymbol{\xi}_k\}_{k=1}^{\infty}$ satisfy

$$\lim_{k \to \infty} \rho_k = 0, \tag{4.12}$$

$$\lim_{k \to \infty} \frac{\boldsymbol{\xi}_k}{\rho_k} = \boldsymbol{x}_\infty. \tag{4.13}$$

Proposition 4.1 shows that we can obtain the limit of the sequence $\lim_{k\to\infty} \boldsymbol{x}_k$ as the limit of fractions $\lim_{k\to\infty} \boldsymbol{\xi}_k/\rho_k$. Moreover, from

$$\lim_{k \to \infty} \boldsymbol{\xi}_k = \left(\lim_{k \to \infty} \rho_k\right)\left(\lim_{k \to \infty} \frac{\boldsymbol{\xi}_k}{\rho_k}\right) = \mathbf{0}, \quad \lim_{k \to \infty} \boldsymbol{\eta}_k = \left(\lim_{k \to \infty} \rho_k\right)\left(\lim_{k \to \infty} \frac{\boldsymbol{\eta}_k}{\rho_k}\right) = \mathbf{0},$$

the points $[\rho_k\ \boldsymbol{\xi}_k^{\mathrm{T}}\ \boldsymbol{\eta}_k^{\mathrm{T}}]^{\mathrm{T}} \in \mathbf{R}^{2n+1}$ converge to the origin as $k \to \infty$, which implies that $\mathcal{V}(\mathcal{F})$ contains the origin as a closed subset of the closure $\overline{D_\rho} = \mathbf{R}^{2n+1}$. This indicates that, to compute the limit points in $\boldsymbol{\xi}$-$\boldsymbol{\eta}$-ρ space, we need only the local information about this algebraic set at the origin.

Remark 4.1. *Note that Proposition 4.1 and the following discussion also indicate that $\mathcal{V}(\mathcal{F})$ includes the origin if and only if the convergence point $\boldsymbol{x}_\infty$ ($\|\boldsymbol{x}_\infty\| < \infty$) exists. Moreover, Theorem 3.1 guarantees the existence of such a convergence point if at least one local minimizer exists. Therefore, $\mathcal{V}(\mathcal{F})$ includes the origin and thus all the mathematical tools introduced in subsection 2.3 can be applied whenever COP* (1.1) *has at least one local minimizer.*

5. Computation of Limit Points and New Necessary Condition for Optimality

To focus on the origin in the algebraic set $\mathcal{V}(\mathcal{F}) \subset \overline{D_\rho} = \mathbf{R}^{2n+1}$, it is useful to consider a tangent cone $C_{\mathbf{0}}(\mathcal{V}(\mathcal{F})) \subset \mathbf{R}^{2n+1}$ at the origin. From Lemma 2.2, the tangent cone can be seen as the set of lines that approximates $\mathcal{V}(\mathcal{F})$ in a neighborhood of the origin. Note that the limit of fractions $[\boldsymbol{x}_\infty^{\mathrm{T}}\ \boldsymbol{y}^{\mathrm{T}}]^{\mathrm{T}} = \lim_{k\to\infty}[\boldsymbol{\xi}_k^{\mathrm{T}}\ \boldsymbol{\eta}_k^{\mathrm{T}}]^{\mathrm{T}}/\rho_k$ can be seen as the gradient of such a line with respect to ρ at the origin. This consideration leads us to relate the limit $\lim_{k\to\infty}[\boldsymbol{\xi}_k^{\mathrm{T}}\ \boldsymbol{\eta}_k^{\mathrm{T}}]^{\mathrm{T}}/\rho_k$ with the gradient of each line in $C_{\mathbf{0}}(\mathcal{V}(\mathcal{F}))$ at the origin, as stated in the following lemma.

Lemma 5.1. *Let $\{r_k\}_{k=1}^\infty$ be a sequence monotonically tending to infinity and $\{\boldsymbol{x}_k\}_{k=1}^\infty$ be a sequence such that each pair $(\boldsymbol{x}_k, r_k)$ satisfies equation* (3.6) *with a fixed $\boldsymbol{y}$. Suppose that $\{\boldsymbol{x}_k\}_{k=1}^\infty$ has a convergence point $\boldsymbol{x}_\infty$. Then, $[1\ \boldsymbol{x}_\infty^{\mathrm{T}}\ \boldsymbol{y}^{\mathrm{T}}]^{\mathrm{T}} \in C_{\mathbf{0}}(\mathcal{V}(\mathcal{F}))$ holds.*

Proof. Let us define sequences $\{\rho_k\}_{k=1}^\infty$, $\{\boldsymbol{\xi}_k\}_{k=1}^\infty$, and $\{\boldsymbol{\eta}_k\}_{k=1^\infty}$ as in Proposition 4.1. From the preceding discussion of the proposition, all polynomials in $\mathcal{F}$ vanish at each triplet $(\rho_k, \boldsymbol{\xi}_k, \boldsymbol{\eta}_k)$ because its preimage $(\boldsymbol{x}_k, \boldsymbol{y}, r_k)$ under Φ satisfies equation (3.6). Moreover, as mentioned in the proposition and the discussion following it,

$$\lim_{k\to\infty} \left[\rho_k\ \boldsymbol{\xi}_k^{\mathrm{T}}\ \boldsymbol{\eta}_k^{\mathrm{T}}\right]^{\mathrm{T}} = \mathbf{0} \in \mathbf{R}^{2n+1}$$

holds if $\lim_{k\to\infty} r_k = \infty$ and $\lim_{k\to\infty} \boldsymbol{x}_k = \boldsymbol{x}_\infty$ hold. Since algebraic sets are closed, this convergence implies $\mathbf{0}$ is an element of $\mathcal{V}(\mathcal{F}) \subset \mathbf{R}^{2n+1}$, and thus we can define the sequence of secant lines $\{L_k\}_{k=1}^\infty$ as those through $\mathbf{0}$ and the points $\boldsymbol{q}_k := [\rho_k\ \boldsymbol{\xi}_k^{\mathrm{T}}\ \boldsymbol{\eta}_k^{\mathrm{T}}]^{\mathrm{T}}$. Let us parametrize each line L_k by $\{t\boldsymbol{v}_k \mid \boldsymbol{v}_k = \boldsymbol{q}_k/\rho_k, t \in \mathbf{R}\}$; then, Proposition 4.1 readily shows that $\{L_k\}_{k=1}^\infty$ converges to the line:

$$L := \left\{t\boldsymbol{v} \mid \boldsymbol{v} = \left[1\ \boldsymbol{x}_\infty^{\mathrm{T}}\ \boldsymbol{y}^{\mathrm{T}}\right]^{\mathrm{T}}, t \in \mathbf{R}\right\}$$

as $k \to \infty$. Therefore, from Lemma 2.2, $[1\ \boldsymbol{x}_\infty^{\mathrm{T}}\ \boldsymbol{y}^{\mathrm{T}}]^{\mathrm{T}} \in L \subset C_{\mathbf{0}}(\mathcal{V}(\mathcal{F}))$ holds. □

This lemma shows that the intersection of the tangent cone and a hyperplane

$$C_{\mathbf{0}}(\mathcal{V}(\mathcal{F})) \cap \left\{\left[\rho\ \boldsymbol{\xi}^{\mathrm{T}}\ \boldsymbol{\eta}^{\mathrm{T}}\right]^{\mathrm{T}} \in \mathbf{R}^{2n+1}\ \middle|\ \rho = 1\right\}$$

includes all the points of the form $[1\ \boldsymbol{x}_\infty^{\mathrm{T}}\ \boldsymbol{y}^{\mathrm{T}}]^{\mathrm{T}}$ where $\boldsymbol{x}_\infty$ is a limit point, as r goes to infinity, of the sequences $\{\boldsymbol{x}_k\}_{k=1}^\infty$ whose elements satisfy equation (3.6) with a fixed $\boldsymbol{y}$. In other words, if we have a set of l polynomials $G = \{G_1, \ldots, G_l\} \subset \mathbf{R}[\rho, \boldsymbol{\xi}, \boldsymbol{\eta}]$ defining the tangent cone $C_{\mathbf{0}}(\mathcal{V}(\mathcal{F}))$, $G_i(1, \boldsymbol{x}_\infty, \boldsymbol{y}) = 0$ $(i = 1, \ldots, l)$ holds for every pair of $\boldsymbol{x}_\infty$ and $\boldsymbol{y}$ satisfying the assumptions of Lemma 5.1. Moreover, if $\boldsymbol{y} = \hat{\boldsymbol{x}}$ attains a minimum of COP (1.1), Theorem 3.1 guarantees the existence of a sequence $\{\boldsymbol{x}_k\}_{k=1}^\infty$ that converges to $\boldsymbol{x}_\infty = \hat{\boldsymbol{x}}$. This implies that every minimizer $\hat{\boldsymbol{x}}$ satisfies equation $G_i(1, \hat{\boldsymbol{x}}, \hat{\boldsymbol{x}}) = 0$ $(i = 1, \ldots, l)$. Finally, the whole process to obtain the polynomial set

$$\mathcal{G} := \{G_i(1, \boldsymbol{x}, \boldsymbol{x}) \in \mathbf{R}[\boldsymbol{x}] \mid G_i \in G\}$$

from COP (1.1) can be summarized as Algorithm 1, and by using $\mathcal{G}$, the new necessary condition for optimality can be stated as Theorem 5.1.

Theorem 5.1. *Let $\mathcal{G} = \{\mathcal{G}_1, \ldots, \mathcal{G}_l\} \subset \mathbf{R}[\boldsymbol{x}]$ be the set of equations obtained by Algorithm 1. Suppose that $\hat{\boldsymbol{x}}$ is a minimizer of COP* (1.1)*. Then, $\hat{\boldsymbol{x}}$ satisfies the equations*

$$\mathcal{G}_i(\hat{\boldsymbol{x}}) = 0 \quad \forall i = 1, \ldots, l. \tag{5.1}$$

Proof. Let $\hat{\boldsymbol{x}}$ be a minimizer of COP (1.1). For a sequence $\{r_k\}_{k=1}^{\infty}$ monotonically tending to infinity as k does, Theorem 3.1 guarantees that the existence of a sequence $\{\boldsymbol{x}_k\}_{k=1}^{\infty}$ converging to $\hat{\boldsymbol{x}}$ whose elements satisfy equation $P_{\boldsymbol{x}}^{\mathrm{loc}}(\boldsymbol{x}_k; r_k, \hat{\boldsymbol{x}}) = \boldsymbol{0}$ for every k. Then, Lemma 5.1 indicates that $[1\ \hat{\boldsymbol{x}}^{\mathrm{T}}\ \hat{\boldsymbol{x}}^{\mathrm{T}}]^{\mathrm{T}}$ is a point of $C_{\boldsymbol{0}}(\mathcal{V}(\mathcal{F}))$. Now, Let $G = \{G_1, \ldots, G_l\} \subset \mathbf{R}[\rho, \boldsymbol{\xi}, \boldsymbol{\eta}]$ be a set of generators of an ideal defining $C_{\boldsymbol{0}}(\mathcal{V}(\mathcal{F}))$, that is, $C_{\boldsymbol{0}}(\mathcal{V}(\mathcal{F})) = \mathcal{V}(G)$ holds. Since $[1\ \hat{\boldsymbol{x}}^{\mathrm{T}}\ \hat{\boldsymbol{x}}^{\mathrm{T}}]^{\mathrm{T}} \in C_{\boldsymbol{0}}(\mathcal{V}(\mathcal{F}))$,

$$\mathcal{G}_i(\hat{\boldsymbol{x}}) = G_i(1, \hat{\boldsymbol{x}}, \hat{\boldsymbol{x}}) = 0$$

holds for all $i = 1, \ldots, l$, which completes the proof. □

6. Numerical Examples

This section is devoted to two numerical examples. The first one demonstrates the proposed methodology and clarifies the relationships among the penalty function method, homogenization, and the tangent cone. The second one demonstrates how the proposed method can find non-KKT type global minimizers.

Example 1 (Illustrative example)

Let us consider the following COP:

$$\begin{aligned} \min_{x}\ & \frac{1}{2}(x_1^2 + x_2^2) \\ \text{s. t. }\ & (x_1 - 5)^2 - (x_2 - 4)^3 = 0, \end{aligned} \tag{6.1}$$

where $\boldsymbol{x} = [x_1\ x_2]^{\mathrm{T}} \in \mathbf{R}^2$ are indeterminates. Figure 2 shows the feasible set and contours of the cost function. For this problem, the penalty function of the localized COP (3.3) is obtained as

$$P^{\mathrm{loc}}(\boldsymbol{x}; r, \boldsymbol{y}) = \frac{1}{2}(x_1^2 + x_2^2) + (x_1 - y_1)^2 + (x_2 - y_2)^2 + r\left\{(x_1 - 5)^2 - (x_2 - 4)^3\right\}^2, \tag{6.2}$$

where $r \in \mathbf{R}$ is a penalty parameter and $\boldsymbol{y} = [y_1\ y_2]^{\mathrm{T}}$ are additional variables representing the coordinates of a minimizer. The stationary conditions $P_{\boldsymbol{x}}^{\mathrm{loc}}(\boldsymbol{x}; r, \boldsymbol{y}) = \boldsymbol{0}$ are the following polynomial equations:

$$\begin{aligned} x_1 + 2(x_1 - y_1) + 4r\left((x_1 - 5)^2 - (x_2 - 4)^3\right)(x_1 - 5) = 0, \\ x_2 + 2(x_2 - y_2) - 6r\left((x_1 - 5)^2 - (x_2 - 4)^3\right)(x_2 - 4)^2 = 0. \end{aligned} \tag{6.3}$$

Hence, the polynomial set $F \subset \mathbf{R}[x_1, x_2, y_1, y_2, r]$ in Section 4 consists of the left-hand sides of equations (6.3).

Algorithm 1 Symbolic Computation of $\mathcal{G}$

Input: COP (1.1)
Output: Set of polynomial $\mathcal{G} \subset \mathbf{R}[\boldsymbol{x}]$

1: Compute polynomial set $F \subset \mathbf{R}[\boldsymbol{x}, \boldsymbol{y}, r]$ in (4.1) by differentiating localized penalty function $P^{\mathrm{loc}}(\boldsymbol{x}; r, \boldsymbol{y}) = f(\boldsymbol{x}) + \|\boldsymbol{x} - \boldsymbol{y}\|^2 + rg^{\mathrm{T}}(\boldsymbol{x})g(\boldsymbol{x})$
2: Compute $\mathcal{F} \subset \mathbf{R}[\rho, \boldsymbol{\xi}, \boldsymbol{\eta}]$ by computing homogenization as in equation (4.5) and dehomogenization as in equation (4.7)
3: Compute set of generators $G \subset \mathbf{R}[\rho, \boldsymbol{\xi}, \boldsymbol{\eta}]$ of ideal defining tangent cone $C_{\boldsymbol{0}}(\mathcal{V}(\mathcal{F}))$ from polynomial set $\mathcal{F}$, for example, by using Gröbner basis described in Lemma 2.1
4: Define $\mathcal{G}$ as $G|_{\rho=1, \boldsymbol{\xi}=\boldsymbol{x}, \boldsymbol{\eta}=\boldsymbol{x}}$

By replacing x_1, x_2, y_1, y_2, and r with X_1/X_0, X_2/X_0, X_3/X_0, X_4/X_0, and X_5/X_0, respectively, the homogenizations of F are obtained as

$$X_0^4X_1 + 2X_0^4(X_1 - X_3) + 4X_5\left(X_0(X_1 - 5X_0)^2 - (X_2 - 4X_0)^3\right)(X_1 - 5X_0), \tag{6.4}$$

$$X_0^5X_2 + 2X_0^5(X_2 - X_4) - 6X_5\left(X_0(X_1 - 5X_0)^2 - (X_2 - 4X_0)^3\right)(X_2 - 4X_0)^2, \tag{6.5}$$

where $[X_0 : X_1 : X_2 : X_3 : X_4 : X_5] \in \mathbf{P}^5$ is a homogeneous coordinate. The dehomogenizations of polynomials (6.4) and (6.5), denoted as $\mathcal{F}$ in Section 5, are as follows:

$$\begin{aligned} &\rho^4\xi_1 + 2\rho^4(\xi_1 - \eta_1) + 4\left(\rho(\xi_1 - 5\rho)^2 - (\xi_2 - 4\rho)^3\right)(\xi_1 - 5\rho), \\ &\rho^5\xi_2 + 2\rho^5(\xi_2 - \eta_2) - 6\left(\rho(\xi_1 - 5\rho)^2 - (\xi_2 - 4\rho)^3\right)(\xi_2 - 4\rho)^2, \end{aligned} \tag{6.6}$$

where $\rho = X_0/X_5$, $\xi_1 = X_1/X_5$, $\xi_2 = X_2/X_5$, $\eta_1 = X_3/X_5$, and $\eta_2 = X_4/X_5$.

From polynomials (6.6), we obtain $G \subset \mathbf{R}[\rho, \xi_1, \xi_2, \eta_1, \eta_2]$, which define the tangent cone $C_{\mathbf{0}}(\mathcal{V}(\mathcal{F}))$, as a set of five polynomials with total degrees 5, 6, 7, 9, and 10. By substituting $\rho = 1$, $\boldsymbol{\xi} = \boldsymbol{\eta} = \boldsymbol{x}$ to the polynomial set G, we have the polynomial set $\mathcal{G}$ consisting of five polynomials:

$$3x_1{x_2}^2 - 22x_1x_2 + 48x_1 - 10x_2, \tag{6.7}$$

$$\begin{aligned} &4x_1{x_2}^3 - 4{x_1}^3 - 48x_1{x_2}^2 - 20{x_2}^3 + 60{x_1}^2 \\ &\qquad + 192x_1x_2 + 240{x_2}^2 - 556x_1 - 960x_2 + 1780, \end{aligned} \tag{6.8}$$

$$\begin{aligned} &6{x_2}^5 - 6{x_1}^2{x_2}^2 - 120{x_2}^4 + 48{x_1}^2x_2 + 60x_1{x_2}^2 + 960{x_2}^3 \\ &\qquad - 96{x_1}^2 - 480x_1x_2 - 3990{x_2}^2 + 960x_1 + 8880x_2 - 8544, \end{aligned} \tag{6.9}$$

$$\begin{aligned} &72x_1{x_2}^3 - 108{x_1}^3 - 864x_1{x_2}^2 - 540{x_2}^3 \\ &\qquad + 1620{x_1}^2 + 3376x_1x_2 + 6600{x_2}^2 - 12324x_1 - 26480x_2 + 48060, \end{aligned} \tag{6.10}$$

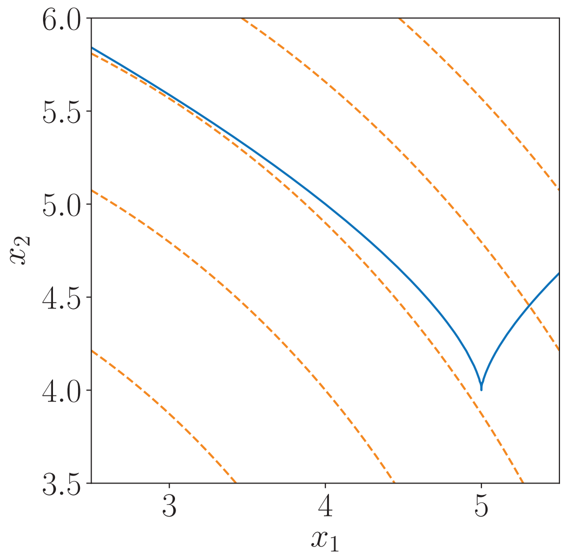
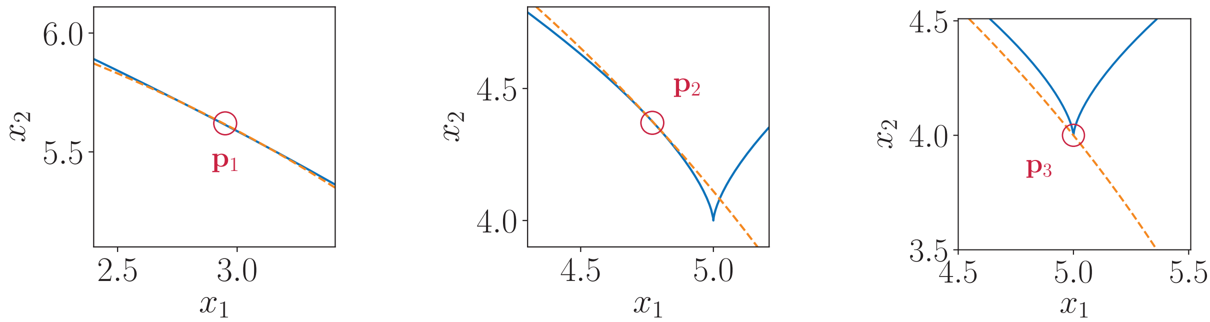


Figure 2: Feasible set (solid lines) and contours of cost function (dashed lines)

Table 1: Candidates derived from proposed necessary condition and corresponding values of cost function

	$\boldsymbol{p}_1$	$\boldsymbol{p}_2$	$\boldsymbol{p}_3$
x_1	2.95	4.77	5.00
x_2	5.62	4.37	4.00
$f(\boldsymbol{p}_i)$	20.1	20.9	20.5



(a) Candidate $\boldsymbol{p}_1$ (b) Candidate $\boldsymbol{p}_2$ (c) Candidate $\boldsymbol{p}_3$

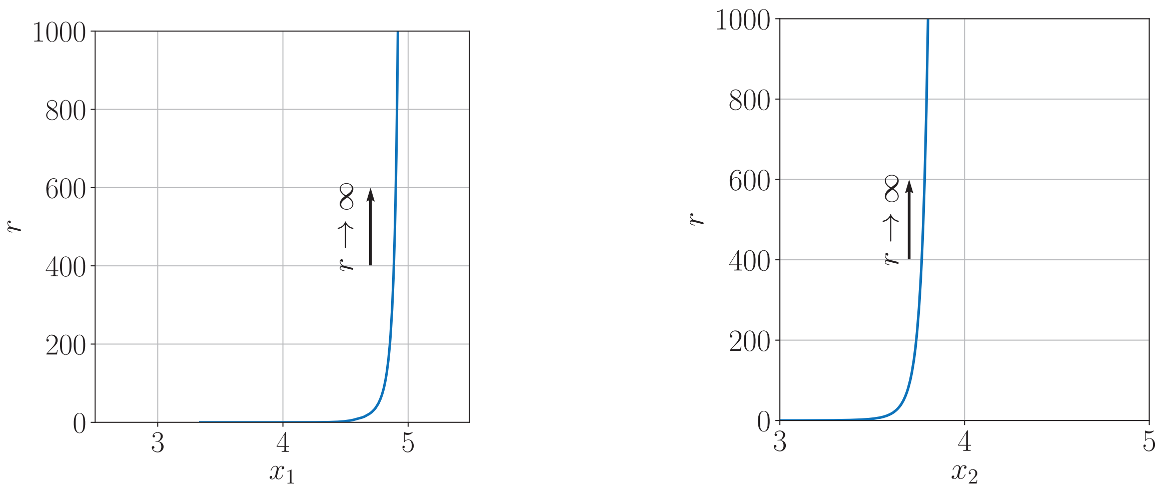
Figure 3: Candidates (circles) in feasible set (solid lines) and corresponding contours of cost function (dashed lines). Labels beside each point correspond to labels in Table 1

and

$$108x_1{}^4 - 48x_1{}^2x_2{}^2 - 1620x_1{}^3 + 592x_1{}^2x_2 + 6180x_1{}^2 - 2000x_1x_2 + 1800x_2{}^2 - 1980x_1 - 9600x_2. \quad (6.11)$$

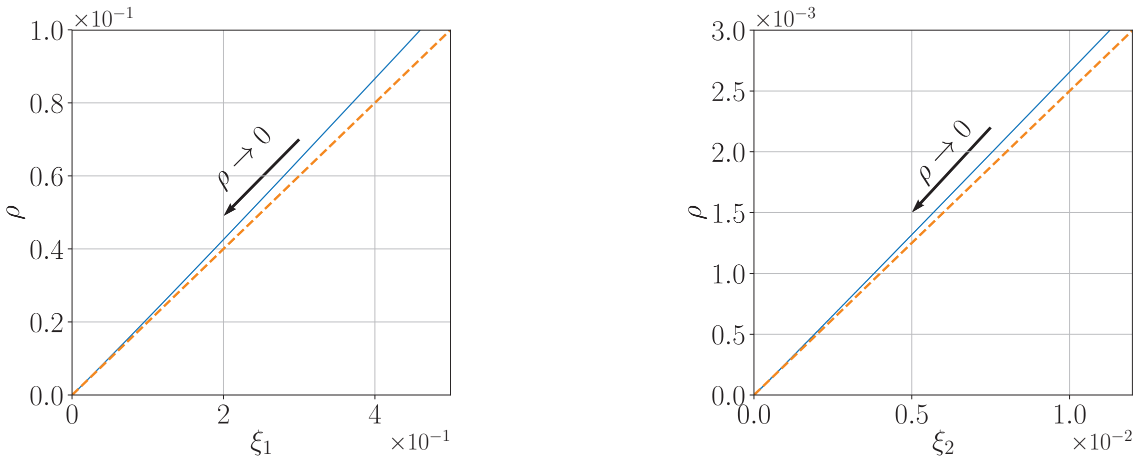
We can obtain three candidates of minimizers (listed in Table 1) by solving equations (6.7)–(6.11). Theorem 5.1 guarantees that these candidates include all minimizers and, consequently, global minimizers. Figure 3 shows the obtained candidates and contours of the cost function corresponding to those candidates. In Figure 3, the contours of the cost function are tangent to the feasible set at the candidate points $\boldsymbol{p}_1$ and $\boldsymbol{p}_2$, which are typical situations on the KKT points. In fact, $\boldsymbol{p}_1$ is a global minimizer, and $\boldsymbol{p}_2$ is a local maximizer. On the other hand, at $\boldsymbol{p}_3$, the contour of the cost function does not seem to be tangent to the feasible set; indeed, $\boldsymbol{p}_3$ has no Lagrange multipliers, so the KKT conditions do not hold. However, as shown in Figure 3(c), $\boldsymbol{p}_3$ is obviously a local minimizer because any feasible point in its neighborhood lies in the area where the value of the cost function is larger than $f(\boldsymbol{p}_3)$.

To illustrate Theorem 3.1 and Proposition 4.1, let us fix $\boldsymbol{y}$ to $[5\ 4]^{\mathrm{T}}$, a minimizer of COP (6.1). If we substitute $\boldsymbol{y} = [5\ 4]^{\mathrm{T}}$ into equations (6.3), Theorem 3.1 guarantees the existence of a stationary point sequence (or trajectory for continuously varying r) that converges to the minimizer, that is, $\boldsymbol{x}_\infty = \boldsymbol{y} = [5\ 4]^{\mathrm{T}}$. Figure 4 shows the solution trajectory of equations (6.3) with $\boldsymbol{y} = [5\ 4]^{\mathrm{T}}$ projected onto the x_1-r and x_2-r planes. We can see that the trajectory approaches the minimizer $[x_1\ x_2]^{\mathrm{T}} = [5\ 4]^{\mathrm{T}}$. However, x_2 still has a non-negligible error for $r = 1000$, which means the common solution method (where r is fixed to a number assumed to be sufficiently large) ends up with the wrong solution.



(a) Trajectory of x_1 with respect to r (b) Trajectory of x_2 with respect to r

Figure 4: Trajectory of $\boldsymbol{x}$ with respect to r satisfying equation (6.3) for $\boldsymbol{y} = [5\ 4]^{\mathrm{T}}$, which is projected onto x_1-r and x_2-r planes



(a) Trajectory of ξ_1 with respect to ρ (b) Trajectory of ξ_2 with respect to ρ

Figure 5: Trajectory satisfying equation (6.6) $= \mathbf{0}$ (solid lines) and its tangent cone at origin (dashed lines) for $\boldsymbol{y} = [5\ 4]^{\mathrm{T}}$, which is projected onto ξ_1-ρ and ξ_2-ρ planes

For Proposition 4.1, equation (4.9) shows that

$$\begin{bmatrix} \eta_1 \\ \eta_2 \end{bmatrix} = \rho \begin{bmatrix} 5 \\ 4 \end{bmatrix} \tag{6.12}$$

holds for $\boldsymbol{y} = [5\ 4]^{\mathrm{T}}$. Substituting equation (6.12) into polynomials (6.6), the algebraic set $\mathcal{V}(\mathcal{F})$ is obtained as the solid curves shown in Figure 5. We can regard the algebraic set $\mathcal{V}(\mathcal{F})$ as a trajectory of $\boldsymbol{\xi}$ with respect to ρ, and it is readily observed that the trajectory converges to the origin as $\rho \to 0$, as mentioned in the discussion following Proposition 4.1. Moreover, the tangent of the trajectory at the origin (dashed lines in Figure 5) has the gradient $[5\ 4]^{\mathrm{T}}$ at the origin, which is the consequence of Proposition 4.1 stated as equations (4.12) and (4.13).

Example 2 (Case with global minimizers violating KKT conditions)

Let us consider the following COP with three indeterminates and a constraint:

$$\begin{aligned} &\min_{\boldsymbol{x}} \frac{1}{2}\|\boldsymbol{x}\|^2 \\ &\text{s. t. } (x_2 - x_1^2 + 2)^2 - (x_3 - 1)^3 = 0, \end{aligned} \tag{6.13}$$

where $\boldsymbol{x} = [x_1\ x_2\ x_3]^{\mathrm{T}} \in \mathbf{R}^3$. The penalty function of the COP is obtained as

$$\frac{1}{2}x_1^2 + \frac{1}{2}x_2^2 + \frac{1}{2}x_3^2 + r\left\{(x_2 - x_1^2 + 2)^2 - (x_3 - 1)^3\right\}^2.$$

Figure 6 shows the feasible set $\mathcal{X}$ of the COP, where all points in the intersection $\mathcal{X} \cap \{\boldsymbol{x} \mid x_3 = 1\}$ are singular and form a parabola on a plane of $x_3 = 1$. As shown, the curve of singularities $\mathcal{X} \cap \{\boldsymbol{x} \mid x_3 = 1\}$ is like "the bottom of a ravine" and is defined by equations $x_3 = 1$ and $x_2 = x_1^2 - 1$. The bottom of the ravine lies above the origin, and thus the feasible point that is closest to the origin would lie on the bottom. In other words, the global minimizer would be included in the curve of singularities. If this is the case, these points cannot be KKT points because, for all points in the curve, the derivatives of the constraint function vanish, whereas those of the cost function do not vanish, which indicates the nonexistence of Lagrange multipliers. Therefore, the Lagrange multiplier method or other methods based on KKT conditions or assuming the existence of Lagrange multipliers cannot find the global minimizers.

For this problem, the proposed method yields a set of 14 polynomials of the highest degree seven as $\mathcal{G}$. These equations have five solutions $\boldsymbol{p}_1, \cdots, \boldsymbol{p}_5$, which are listed in Table 2. As shown in Figure 7, the proposed algorithm yields a set of candidates that includes all KKT points ($\boldsymbol{p}_1$ and $\boldsymbol{p}_2$) and some other non-KKT points on the singular curve ($\boldsymbol{p}_3$, $\boldsymbol{p}_4$, and $\boldsymbol{p}_5$). In Table 2, points $\boldsymbol{p}_4$ and $\boldsymbol{p}_5$ attain the minimum among the candidates, and thus they are the non-KKT type global minimizers.

Note again that there are no Lagrange multipliers corresponding to these global minimizers. Indeed, the derivative of the corresponding Lagrangian with respect to x_3 is

$$x_3 - 3\lambda(x_3 - 1)^2,$$

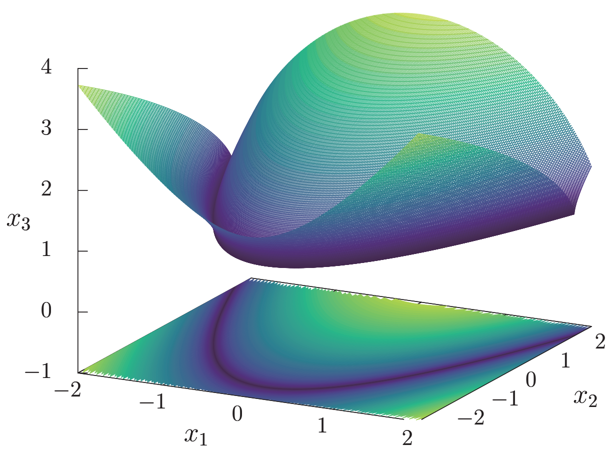


Figure 6: Feasible set of COP (6.13). Heat map on x_1-x_2 plane shows projection of feasible set, whose color corresponds to x_3-coordinate of projected points

Table 2: Candidates derived by proposed algorithm and corresponding values of cost function

	$\boldsymbol{p}_1$	$\boldsymbol{p}_2$	$\boldsymbol{p}_3$	$\boldsymbol{p}_4$	$\boldsymbol{p}_5$
x_1	0.00	0.00	0.00	1.22	−1.22
x_2	−1.35	−1.94	−2.00	−0.50	−0.50
x_3	1.75	1.16	1.00	1.00	1.00
$f(p_i)$	2.44	2.55	2.50	1.38	1.38

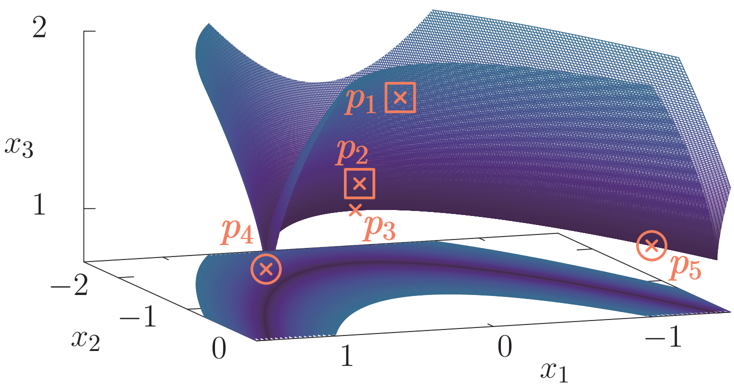


Figure 7: Candidates of minimizer obtained by proposed method (cross), KKT points (open square), and global minimizers (open circle). Labels beside each point correspond to labels in Table 2

where $\lambda \in \mathbf{R}$ is the Lagrange multiplier corresponding to the constraint of COP (6.13). It is obvious that this derivative cannot be zero if $x_3 = 1$ holds; in other words, when a point is included in the singular curve $\mathcal{X} \cap \{\boldsymbol{x} \mid x_3 = 1\}$.

Remark 6.1. *For the first example, the FJ conditions yield the same three points as in Table 1. However, for the second example, the FJ conditions are satisfied by all the points of $\mathcal{X} \cap \{\boldsymbol{x} \mid x_3 = 1\}$, which includes an infinite number of points neither locally optimal nor KKT. This indicates that the proposed condition is less conservative and can yield a significantly smaller number of candidates than the FJ conditions do.*

7. Conclusion and Future Work

We have proposed a new necessary optimality condition for polynomial optimization problems with polynomial constraints. The proposed necessary condition is satisfied by all minimizers and thus does not require any constraint qualifications. First, a sequential optimality condition based on the quadratic penalty function, which is described by the existence of a certain sequence converging to a minimizer as the penalty parameter tends to infinity, is introduced. By considering a projective space, the limit operation of the penalty parameter is symbolically performed as a computation of the tangent cone at the origin. The set of polynomials, which vanish at all the points satisfying the sequential optimality condition, is obtained. Two numerical examples are provided to illustrate the methodology and demonstrate that the proposed necessary condition can be satisfied even by non-KKT type minimizers.

One direction of further study is to generalize the problem settings. For instance, the algorithm should be readily applicable to parametric optimization problems. We will also extend the functions appearing in the problem to include more general functions than polynomials, as long as they can be homogenized in some sense.

It is worth mentioning that, in the proposed algorithm, the iterative computations for solving stationary conditions and updating the penalty parameter are reduced to solving the equations defining the tangent cone only once. This reduction can be applied to any equation with parameters if some convergence property of their solutions is guaranteed, for instance, as mentioned in Theorem 3.1. Therefore, the proposed method can be applied to a broader class of problems beyond optimization problems.

Acknowledgements

This work was partly supported by JSPS KAKENHI Grant Numbers JP18J22093 and JP15H02257.

References

[1] R. Andreani, N.S. Fazzio, M.L. Schuverdt, and L.D. Secchin: A sequential optimality condition related to the quasi-normality constraint qualification and its algorithmic consequences. *SIAM Journal on Optimization*, **29-1** (2019), 743–766.

[2] R. Andreani, G. Haeser, and J.M. Martínez: On sequential optimality conditions for smooth constrained optimization. *Optimization*, **60-5** (2011), 627–641.

[3] R. Andreani, J.M. Martínez, and B.F. Svaiter: A new sequential optimality condition for constrained optimization and algorithmic consequences. *SIAM Journal on Optimization*, **20-6** (2010), 3533–3554.

[4] N. Andreasson, A. Evgrafov, and M. Patriksson: *An Introduction to Continuous Optimization: Foundations and Fundamental Algorithms* (Dover Publications, New York, 2020).

[5] D. Bertsekas: *Nonlinear Programming* (Athena Scientific, NH, 2016).

[6] M. Bierlaire: *Optimization: Principles and Algorithms* (EPFL Press, Lausanne, 2015).

[7] D.A. Cox, J. Little, and D. O'Shea: *Ideals, Varieties, and Algorithms* (Springer International Publishing, Cham, 2015).

[8] D.A. Cox, J. Little, and D. O'Shea: *Using Algebraic Geometry* (Springer-Verlag, New York, 2005).

[9] D. Eisenbud: *Commutative Algebra* (Springer, NewYork, 2004).

[10] A.V. Fiacco and G.P. McCormick: *Nonlinear Programming: Sequential Unconstrained Minimization Techniques* (Wiley, New York, 1968).

[11] E.H. Fukuda and M. Fukushima: A note on the squared slack variables technique for nonlinear optimization. *Journal of the Operations Research Society of Japan*, **60-3** (2017), 262–270.

[12] M. Fukushima: *Fundamentals of Nonlinear Optimization* (Asakura Publishing, Tokyo, 2001) (in Japanese).

[13] F.J. Gould and J.W. Tolle: A necessary and sufficient qualification for constrained optimization. *SIAM Journal on Applied Mathematics*, **20-2** (1971), 164–172.

[14] C. Kanzow, D. Steck, and D. Wachsmuth: An augmented Lagrangian method for optimization problems in Banach spaces. *SIAM Journal on Control and Optimization*, **56-1** (2018), 272–291.

[15] M. Kreuzer and L. Robbiano: *Computational Commutative Algebra 2* (Springer-Verlag, Berlin, 2005).

[16] R. Krone: Numerical algorithms for dual bases of positive-dimensional ideals. *Journal of Algebra and Its Applications*, **12-6** (2013), 1–21.

[17] J.M. Martínez and B.F. Svaiter: A practical optimality condition without constraint qualifications for nonlinear programming. *Journal of Optimization Theory and Applications*, **118** (2003), 117–133.

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

[19] J. Nocedal and S.J. Wright: *Numerical Optimization* (Springer-Verlag, New York, 2006).

[20] D.W. Peterson: A review of constraint qualifications in finite-dimensional spaces. *SIAM Review*, **15-3** (1973), 639–654.

Tomoyuki Iori
Graduate School of Informatics
Kyoto University
Yoshida-honmachi, Sakyo-ku, Kyoto
606-8501, Japan
E-mail: iori.tomoyuki.25s@kyoto-u.jp