

ON CONSTRAINT QUALIFICATIONS: MOTIVATION, DESIGN AND INTER-RELATIONS

ZITENG WANG AND SHU-CHERNG FANG

Edward P. Fitts Department of Industrial and Systems Engineering
North Carolina State University
Raleigh, NC, 27695, USA

WENXUN XING

Department of Mathematical Sciences
Tsinghua University
Beijing, 100084, China

(Communicated by Kok Lay Teo)

ABSTRACT. Constraint qualification (CQ) is an important concept in nonlinear programming. This paper investigates the motivation of introducing constraint qualifications in developing KKT conditions for solving nonlinear programs and provides a geometric meaning of constraint qualifications. A unified framework of designing constraint qualifications by imposing conditions to equate the so-called “locally constrained directions” to certain subsets of “tangent directions” is proposed. Based on the inclusion relations of the cones of tangent directions, attainable directions, feasible directions and interior constrained directions, constraint qualifications are categorized into four levels by their relative strengths. This paper reviews most, if not all, of the commonly seen constraint qualifications in the literature, identifies the categories they belong to, and summarizes the inter-relationship among them. The proposed framework also helps design new constraint qualifications of readers’ specific interests.

1. Introduction. Consider the following general nonlinear programming problem:

$$\begin{aligned} \min \quad & f(x) \\ \text{s.t.} \quad & g_i(x) \leq 0, \quad i \in I, \\ & h_j(x) = 0, \quad j \in J, \\ & x \in \mathbb{R}^n, \end{aligned} \tag{1}$$

where I and J are index sets, $f(x), g_i(x), h_j(x)$ are real valued functions on $\mathbb{R}^n$ for each $i \in I$ and $j \in J$. In this paper, we assume that $f(x), g_i(x)$ and $h_j(x)$ are all continuously differentiable in $\mathbb{R}^n$ and there exists a finite optimal solution. We use $\mathcal{F}$ to denote the feasible region and $\bar{x}$ to denote an optimal solution, globally or locally. The cardinality of an index set A is denoted by $|A|$.

2010 *Mathematics Subject Classification.* Primary: 90C30; Secondary: 90C46.

Key words and phrases. Constraint qualification, nonlinear programming, optimality conditions, optimization, KKT conditions.

Ziteng Wang’s research was supported by North Carolina State University Edward P. Fitts Fellowship. Shu-Cherng Fang’s research was supported by US NSF Grant # DMI-0553310. Wenxun Xing’s research was supported by National Natural Science Foundation of China, Grant # 11171177.

Establishing optimality conditions of problem (1) is one of the fundamentals of nonlinear programming both in theory and computation. In general, an optimality condition is an explicit characterization of the fact that $f(\bar{x}) \leq f(x)$, globally or locally. Since $\mathcal{F}$ is represented by the constraints $g_i(x) \leq 0$ and $h_j(x) = 0$, optimality conditions are usually of the form that reveals the relations among $f(x)$, $g_i(x)$ and $h_j(x)$, particularly at $\bar{x}$. Although an optimality condition that is both necessary and sufficient is preferred, such kind of conditions can only be valid under some ideal assumptions of the problem, for example, convexity. Therefore, most research has focused on determining necessary optimality conditions of (1). Related research started with the inequality constrained problems, i.e., $J = \emptyset$ in (1). One of the earliest necessary optimality conditions was proposed by Fritz John [24]. It says that if $\bar{x}$ is an optimal solution to (1) with $J = \emptyset$, then there exists a vector $\mu = (\mu_0, \mu_1, \dots, \mu_{|I|}) \in \mathbb{R}^{|I|+1}$ such that

$$\begin{cases} \mu_0 \nabla f(\bar{x}) + \sum_{i \in I} \mu_i \nabla g_i(\bar{x}) = 0, \\ \mu_i g_i(\bar{x}) = 0, & i \in I, \\ \mu_i \geq 0, & i \in I, \\ \mu \neq 0. \end{cases} \quad (2)$$

Fritz John conditions (2) require no additional conditions on $\{g_i(x), i \in I\}$ other than being smooth at $\bar{x}$.

Kuhn and Tucker [27] later introduced a condition called constraint qualification (CQ). When this constraint qualification is satisfied by the constraint functions $\{g_i(x), i \in I\}$, μ_0 in (2) becomes nonzero. Then (2) turns into the best-known KKT optimality conditions: If $\bar{x}$ is an optimal solution of (1) with $J = \emptyset$, and the constraint qualification is satisfied, then there exists a vector $\mu = (\mu_1, \dots, \mu_{|I|}) \in \mathbb{R}^{|I|}$ such that

$$\begin{cases} \nabla f(\bar{x}) + \sum_{i \in I} \mu_i \nabla g_i(\bar{x}) = 0, \\ \mu_i g_i(\bar{x}) = 0, & i \in I, \\ \mu_i \geq 0, & i \in I. \end{cases} \quad (3)$$

A solution x is called a KKT solution if it satisfies (3). From a practical point of view, the introduction of constraint qualification eliminates the possibility that Fritz John conditions hold trivially with $\mu_0 = 0$ at some non-optimal solutions for certain representations $\{g_i(x) \leq 0\}$ of $\mathcal{F}$. The set of candidates for being optimal solutions is then narrowed [16].

Since the original work of Kuhn and Tucker [27], there have been numerous papers on designing other types of constraint qualifications and discussing the inter-relations among them. (See for example Hurwicz [22], Slater [44], Karlin [26], Arrow et al. [4], Cottle [12], Canon et al. [11], Abadie [1], Varaiya [47], Guignard [19], Zangwill [48], Mangasarian [33] and Ritter [40].) There are several review papers in this area. For example, Bazaraa et al. [5] classified constraint qualifications into two groups in terms of their relative strengths and highlighted the inter-relations among them. Peterson [38] reviewed most of the known constraint qualifications by 1973 and provided a guidance for generating new constraint qualifications. Both reviews focused on the inter-relations of constraint qualifications rather than the motivations of derivation and geometrical meanings. Besides, they only discussed the case with $J = \emptyset$ in (1).

One approach of designing constraint qualifications for problem (1) with $J \neq \emptyset$ is to replace each equality constraint equivalently by two inequality constraints [16]. (See for example Mangasarian [33], Bazaraa et al. [6].) Certain additional

assumptions, for example the linear independence of $\{\nabla h_j(\bar{x}), j \in J\}$, are added in some constraint qualifications to make them complete. There are other approaches. For example, Mangasarian and Fromovitz [34] derived a constraint qualification in a direct way by generalizing Fritz John conditions (2) for (1). Gould and Tolle [16] pointed out a weakest constraint qualification for (1). This area still enjoys new development in recent years. Also, there are constraint qualifications designed in the sense of quasinormality of $\bar{x}$ [21]. (See for example, Janin [23], Qi and Wei [39], Minchenko and Stakhovski [36] and Andreani et al. [2].) They require certain properties of the constraint functions not only at $\bar{x}$ but also in a neighborhood of $\bar{x}$. Solodov [45] provides a brief review on this type of constraint qualifications.

In this paper, we study the motivation of introducing constraint qualifications by investigating the development of KKT conditions for problem (1) with $J = \emptyset$ and provide a geometric meaning of constraint qualifications. We use the cone of tangent directions [1] as a geometric representation of the local approximation of $\mathcal{F}$ at $\bar{x}$. A tangent direction is the limit direction of a sequence of feasible solutions that approach $\bar{x}$. A general condition for $\bar{x}$ being optimal is that $\nabla f(\bar{x})$ lies in the dual cone of tangent directions. To derive KKT conditions from this condition, we need an analytic representation of the local approximation of $\mathcal{F}$ at $\bar{x}$ to replace the cone of tangent directions and to link $\nabla f(\bar{x})$ with $\{\nabla g_i(\bar{x}), i \in I\}$. The cone of locally constrained directions [4] is considered for this purpose. A locally constrained direction has a nonpositive projection on the gradients of active constraint functions. In general, the cone of tangent directions is contained in the cone of locally constrained directions. KKT conditions will hold at $\bar{x}$ when the dual cones of tangent directions and locally constrained directions become equal. A condition on the constraint functions that makes these two cones equal then forms a constraint qualification. Therefore, the essence of constraint qualifications is requiring the analytic and geometric representations of the local approximation of the feasible region $\mathcal{F}$ at the optimal solution $\bar{x}$ to coincide with each other.

Based on the above mentioned, we propose a unified framework of designing constraint qualifications by imposing conditions such that the cone of locally constrained directions equals to certain subsets of the cone of tangent directions. These subsets include the cones of attainable directions [4], feasible directions [48] and interior constrained directions [5]. The cone of attainable directions contains the limit directions of continuous paths along which $\bar{x}$ can be approached by the solutions in $\mathcal{F}$ while the paths are restricted to straight lines for the cone of feasible directions. The cone of interior constrained directions is the interior of the cone of locally constrained directions and it is contained in the cone of feasible directions. Based on the inclusion relations of these cones, we categorize all constraint qualifications into four levels by their relative strengths. For example, as we will see, the cone of feasible directions is contained in the cone of attainable directions. Then a constraint qualification that makes the cone of locally constrained directions equal to the cone of feasible directions is stronger than a constraint qualification that makes the cone of locally constrained directions equal to the cone of attainable directions. Therefore, these two constraint qualifications will be categorized into two different levels. We use the proposed framework and the four-level categorization scheme to review most, if not all, of the commonly seen constraint qualifications in the literature. The inter-relationship among these constraint qualifications is then revealed.

To design constraint qualifications of problem (1) with $J \neq \emptyset$, we replace each equality constraint in (1) by two inequality constraints and reformulate (1) into

a problem with only inequality constraints. For the reformulated problem to be further discussed in Section 5, however, the cone of interior constrained directions will always be an empty set. Therefore, we consider the relative interior of the cone of locally constrained directions instead and call it the cone of relative interior constrained directions. To replace the role of interior constrained directions with relative interior constrained directions in the inclusion relations of subsets of the cone of tangent directions, additional conditions are introduced. These additional conditions include that $\{h_j(x), j \in J\}$ are affine and that $\{\nabla h_j(\bar{x}), j \in J\}$ are linearly independent. The constraint qualifications for problem (1) with $J = \emptyset$ are then generalized to problem (1) with $J \neq \emptyset$. We will review the constraint qualifications designed in the sense of quasinormality of $\bar{x}$ [21] too. These constraint qualifications are not explicitly specified using the subsets of the cone of tangent directions, but they are closely related to the constraint qualifications that can be designed in the framework we proposed. We try to identify the appropriate levels of these constraint qualifications in the four-level categorization scheme.

In this paper, we focus on constraint qualifications in finite dimensional smooth optimization. For constraint qualifications in infinite dimensional spaces, see for example [18], [10], [41], [19], [20], [22], [32]. For constraint qualifications in nonsmooth optimization, see for example [35], [46], [29], [25], [28].

The remaining of the paper is organized as follows. Section 2 provides definitions and preliminary results in convex analysis and nonlinear programming. The definitions and inter-relations of different cones of directions are particularly studied. Section 3 investigates the motivation of introducing a constraint qualification in developing KKT conditions for problem (1) with $J = \emptyset$ and its geometrical meaning. A unified framework of designing constraint qualifications is proposed. A four-level categorization scheme of constraint qualifications is also introduced in this section. Section 4 reviews most, if not all, of the commonly seen constraint qualifications, places them in appropriate categories and shows their inter-relations. Section 5 generalizes the constraint qualifications for problem (1) with $J \neq \emptyset$. Constraint qualifications designed in the sense of quasinormality are also reviewed in this section. Section 6 concludes the paper.

2. Preliminaries. In this section, we present some definitions and results that will be used later. For any set $\mathcal{S} \subseteq \mathbb{R}^n$, we denote the closure of $\mathcal{S}$ by $\text{cl}(\mathcal{S})$ and denote the set of interior points of $\mathcal{S}$ by $\text{int}(\mathcal{S})$. $\mathbb{Z}_+$ is used to denote the set of positive integer numbers. The index set of active inequality constraints at $\bar{x}$ is denoted by $I_{ac}(\bar{x})$, i.e., $I_{ac}(\bar{x}) = \{i \in I | g_i(\bar{x}) = 0\}$.

2.1. Convexity.

Definition 2.1 (convex set). A set $\mathcal{S} \subseteq \mathbb{R}^n$ is convex if $\lambda x + (1 - \lambda)\tilde{x} \in \mathcal{S}$ holds for any $x \in \mathcal{S}$, $\tilde{x} \in \mathcal{S}$ and $\lambda \in [0, 1]$.

Definition 2.2 (convex hull). The convex hull of a set $\mathcal{S} \subseteq \mathbb{R}^n$ is

$$\text{conv}(\mathcal{S}) = \{d \in \mathbb{R}^n | d = \sum_{k=1}^r \lambda^k x^k \text{ for some } r \in \mathbb{Z}_+, x^k \in \mathcal{S}, \lambda^k \geq 0, \text{ and } \sum_{k=1}^r \lambda^k = 1\}.$$

Note that $\text{conv}(\mathcal{S})$ is the smallest convex set that contains $\mathcal{S}$.

Definition 2.3 (convex function). Let $\mathcal{S}$ be a convex set in $\mathbb{R}^n$. A function $f(x) : \mathcal{S} \rightarrow \mathbb{R}$ is convex if $f(\lambda x + (1 - \lambda)\tilde{x}) \leq \lambda f(x) + (1 - \lambda)f(\tilde{x})$ holds for any $x \in \mathcal{S}$, $\tilde{x} \in \mathcal{S}$ and $\lambda \in [0, 1]$.

Definition 2.4 (pseudoconvex function). Let $\mathcal{S}$ be a nonempty convex open set in $\mathbb{R}^n$ and $f(x) : \mathcal{S} \rightarrow \mathbb{R}$ be a differentiable function on $\mathcal{S}$. f is pseudoconvex if $f(\tilde{x}) < f(x)$ implies $\nabla f(x)^T(\tilde{x} - x) < 0$ for any $x \in \mathcal{S}$ and $\tilde{x} \in \mathcal{S}$.

Definition 2.5 (quasiconvex function). Let $\mathcal{S}$ be a convex set in $\mathbb{R}^n$. A function $f(x) : \mathcal{S} \rightarrow \mathbb{R}$ is quasiconvex if $f(\lambda x + (1 - \lambda)\tilde{x}) \leq \max\{f(x), f(\tilde{x})\}$ for any $x \in \mathcal{S}$, $\tilde{x} \in \mathcal{S}$ and $\lambda \in [0, 1]$.

A function f is concave (pseudoconcave or quasiconcave) if $-f$ is convex (pseudoconvex or quasiconvex). A convex (concave) function must be pseudoconvex (pseudoconcave), but not vice versa. A pseudoconvex (pseudoconcave) function must be quasiconvex (quasiconcave), but not vice versa.

2.2. Cone and dual cone.

Definition 2.6 (cone). A set $\mathcal{K} \subseteq \mathbb{R}^n$ is a cone if $\lambda x \in \mathcal{K}$ holds for any $x \in \mathcal{K}$ and $\lambda > 0$.

Definition 2.7 (polyhedral cone). A cone $\mathcal{K} \subseteq \mathbb{R}^n$ is a polyhedral cone if it can be represented as

$$\mathcal{K} = \{x \in \mathbb{R}^n | \mathbf{M}x \leq 0\},$$

for some $m \times n$ matrix $\mathbf{M}$.

Definition 2.8 (conic hull). The conic hull of a set $\mathcal{S} \subseteq \mathbb{R}^n$ is

$$\text{cone}(\mathcal{S}) = \{d \in \mathbb{R}^n | d = \sum_{k=1}^r \lambda^k x^k \text{ for some } r \in \mathbb{Z}_+, x^k \in \mathcal{S}, \lambda^k \geq 0\}.$$

Note that $\text{cone}(\mathcal{S})$ is the smallest convex cone that contains $\mathcal{S}$.

Definition 2.9 (dual cone). Let $\mathcal{K} \subseteq \mathbb{R}^n$ be a cone. The dual cone of $\mathcal{K}$ is

$$\mathcal{K}^* = \{d \in \mathbb{R}^n | d^T x \geq 0, \forall x \in \mathcal{K}\}.$$

Geometrically, $d \in \mathcal{K}^*$ means that d has a nonnegative projection on any $x \in \mathcal{K}$. Basic properties of dual cones are summarized in the following lemma:

Lemma 2.10 (See for example [13], [14]). Let $\mathcal{K}$, $\mathcal{K}_1$, $\mathcal{K}_2$ be cones in $\mathbb{R}^n$. The following properties hold:

- (i) If $\mathcal{K}_1 \subseteq \mathcal{K}_2$, then $\mathcal{K}_2^* \subseteq \mathcal{K}_1^*$.
- (ii) $\mathcal{K}^*$ is a closed convex cone.
- (iii) The dual cone of $\mathcal{K}^*$ is the closed convex hull of $\mathcal{K}$, i.e., $(\mathcal{K}^*)^* = \text{cl}(\text{conv}(\mathcal{K}))$.

2.3. Cones of directions at $\bar{x}$. We review some cones of directions at the optimal solution $\bar{x} \in \mathcal{F}$. They all can be viewed as a local approximation of the feasible region $\mathcal{F}$ at $\bar{x}$.

Definition 2.11 (feasible direction, [48]). A vector $d \in \mathbb{R}^n$ is a feasible direction at $\bar{x}$ if there exists $\delta_0 > 0$ such that $\bar{x} + \delta d \in \mathcal{F}, \forall \delta \in [0, \delta_0]$.

Note that the set of feasible directions is a cone. We denote the cone of feasible directions at $\bar{x}$ by $\mathcal{D}(\bar{x})$.

Geometrically, if d is a feasible direction, then we can move $\bar{x}$ along a straight line d by a certain range of step to $\bar{x} + \delta d \in \mathcal{F}$ such that $\bar{x} + \delta d$ remains feasible. Note that $\mathcal{D}(\bar{x})$ is not necessarily closed [48].

Definition 2.12 (attainable direction, [4]). A vector $d \in \mathbb{R}^n$ is an attainable direction at $\bar{x}$ if there exists $\delta_0 > 0$ and a continuous path $\alpha : \mathbb{R} \rightarrow \mathbb{R}^n$ such that $\alpha(\delta) \in \mathcal{F}$ for $\delta \in [0, \delta_0]$, $\alpha(0) = \bar{x}$, and $d = \lim_{\delta \rightarrow 0^+} \frac{\alpha(\delta) - \alpha(0)}{\delta}$.

Note that the set of attainable directions is a cone. We denote the cone of attainable directions at $\bar{x}$ by $\mathcal{A}(\bar{x})$.

Definition 2.13 (tangent direction, [1]). A vector $d \in \mathbb{R}^n$ is a tangent direction at $\bar{x}$ if there exists a sequence $\{\theta^k \geq 0\}_{k=1}^\infty \subseteq \mathbb{R}$ and a sequence $\{x^k\}_{k=1}^\infty \subseteq \mathcal{F}$ such that $\bar{x} = \lim_{k \rightarrow \infty} x^k$ and $d = \lim_{k \rightarrow \infty} \theta^k(x^k - \bar{x})$.

Note that the set of tangent directions is a cone. We denote the cone of tangent directions at $\bar{x}$ by $\mathcal{T}(\bar{x})$.

The geometric meanings of attainable and tangent directions are that, besides a straight line, $\bar{x}$ can be approached by a continuous path and a sequence of feasible solutions in $\mathcal{F}$, respectively. Other definitions of the tangent directions can be seen, for example, in [17], [6], [38] and [43]. A comparison of these definitions can be seen in [15]. Moreover, $\mathcal{A}(\bar{x})$ and $\mathcal{T}(\bar{x})$ are both closed cones [38], [1].

If $\mathcal{F}$ is convex, then $\mathcal{D}(\bar{x})$, $\mathcal{A}(\bar{x})$ and $\mathcal{T}(\bar{x})$ are all convex [38]. In general, none of $\mathcal{D}(\bar{x})$, $\mathcal{A}(\bar{x})$ or $\mathcal{T}(\bar{x})$ is necessarily to be convex. An example that shows $\mathcal{T}(\bar{x})$ is not convex is given below.

Example 2.1.

$$\begin{aligned} \min \quad & -x_1 - x_2 \\ \text{s.t.} \quad & x_1 x_2 \leq 0, \\ & x_1 \leq 0, \\ & x_2 \leq 0, \\ & x \in \mathbb{R}^2. \end{aligned}$$

In this example, $\mathcal{F} = \{x \in \mathbb{R}^2 | x_1 \leq 0 \text{ and } x_2 = 0, \text{ or } x_1 = 0 \text{ and } x_2 \leq 0\}$. The optimal solution is $\bar{x} = (0, 0)^T$. $\mathcal{T}(\bar{x})$ coincides with $\mathcal{F}$ and is not convex.

$\mathcal{D}(\bar{x})$, $\mathcal{A}(\bar{x})$ and $\mathcal{T}(\bar{x})$ can be viewed as local approximations of $\mathcal{F}$ at $\bar{x}$. They are all defined by the geometric properties rather than any specific representations of $\mathcal{F}$. We are interested in describing them by the constraint functions so that we can have analytic representations which are more convenient for use. The cone of locally constrained directions and interior constrained directions are introduced for this purpose. In this section, we review these two cones for inequality constrained problems. For problems with both inequality and equality constraints, the discussion is a little more complicated and is arranged in Section 5.

Definition 2.14 (locally constrained direction, [4]). For problem (1) with $J = \emptyset$, a vector d is a locally constrained direction at $\bar{x}$ if $\nabla g_i(\bar{x})^T d \leq 0$ holds for any $i \in I_{ac}(\bar{x})$.

Note that the set of locally constrained directions is a cone. We denote the cone of locally constrained directions at $\bar{x}$ by $\mathcal{G}(\bar{x})$.

Definition 2.15 (interior constrained direction). For problem (1) with $J = \emptyset$, a vector d is an interior constrained direction at $\bar{x}$ if $\nabla g_i(\bar{x})^T d < 0$ holds for any $i \in I_{ac}(\bar{x})$.

Note that the set of interior constrained directions is a cone. We denote the cone of interior constrained directions at $\bar{x}$ by $\mathcal{G}^0(\bar{x})$. $\mathcal{G}^0(\bar{x})$ is also referred to as the strictly inward cone [5].

Obviously, $\mathcal{G}(\bar{x}) \neq \emptyset$ because $0 \in \mathcal{G}(\bar{x})$. Actually, $\mathcal{G}(\bar{x})$ is a polyhedral cone and thus closed and convex. It is easy to see that $\mathcal{G}^0(\bar{x}) = \text{int}(\mathcal{G}(\bar{x}))$. Therefore, $\text{cl}(\mathcal{G}^0(\bar{x})) = \mathcal{G}(\bar{x})$ if and only if $\mathcal{G}^0(\bar{x}) \neq \emptyset$. Geometrically, each vector $d \in \mathcal{G}^0(\bar{x})$ has a negative projection on the gradients of active inequality constraint functions. This implies that moving from $\bar{x}$ along d within a certain range of steps will not violate the constraints. Hence d is a feasible direction. $\mathcal{G}(\bar{x})$ and $\mathcal{G}^0(\bar{x})$ provide analytic representations of the local approximation of $\mathcal{F}$ at $\bar{x}$. The following theorem reveals the inclusion relations of the cones of directions at $\bar{x}$ defined above.

Theorem 2.16 (See also [5]). *For problem (1) with $J = \emptyset$, $\mathcal{G}^0(\bar{x}) \subseteq \mathcal{D}(\bar{x}) \subseteq \mathcal{A}(\bar{x}) \subseteq \mathcal{T}(\bar{x}) \subseteq \mathcal{G}(\bar{x})$.*

Proof. First, we show that $\mathcal{G}^0(\bar{x}) \subseteq \mathcal{D}(\bar{x})$. Suppose that $d \in \mathcal{G}^0(\bar{x})$, we need to show there exists $\delta_0 > 0$ such that $\bar{x} + \delta d \in \mathcal{F}$, $\forall \delta \in [0, \delta_0]$. Actually, when $i \notin I_{ac}(\bar{x})$, we know $g_i(\bar{x} + \delta d) < 0$ when δ is sufficiently small because $g_i(x)$ is continuous at $\bar{x}$. When $i \in I_{ac}(\bar{x})$, consider the Taylor series expansion of $g_i(\bar{x} + \delta d)$ at $\bar{x}$:

$$g_i(\bar{x} + \delta d) = g_i(\bar{x}) + \delta \nabla g_i(\bar{x})^T d + o(\delta) \|d\|, \text{ as } \delta \rightarrow 0.$$

Consequently, $g_i(\bar{x} + \delta d) < g_i(\bar{x}) = 0$ when δ is sufficiently small. Therefore, $d \in \mathcal{D}(\bar{x})$ and $\mathcal{G}^0(\bar{x}) \subseteq \mathcal{D}(\bar{x})$.

From the definitions of $\mathcal{D}(\bar{x})$, $\mathcal{A}(\bar{x})$ and $\mathcal{T}(\bar{x})$, it is easy to check that $\mathcal{D}(\bar{x}) \subseteq \mathcal{A}(\bar{x}) \subseteq \mathcal{T}(\bar{x})$.

Finally, we prove $\mathcal{T}(\bar{x}) \subseteq \mathcal{G}(\bar{x})$. Suppose that $d \in \mathcal{T}(\bar{x})$ and there exists $d \in \mathcal{T}(\bar{x})$ and $i \in I_{ac}(\bar{x})$ such that $\nabla g_i(\bar{x})^T d > 0$. Then, for $\{x^k\}_{k=1}^\infty$ associated with d in the definition of tangent directions, there exists k_0 such that $\nabla g_i(\bar{x})^T (x^k - \bar{x}) > 0$ when $k > k_0$. When $k \rightarrow \infty$, we have

$$\begin{aligned} g_i(x^k) &= g_i(\bar{x}) + \nabla g_i(\bar{x})^T (x^k - \bar{x}) + o(\|x^k - \bar{x}\|) \\ &= \nabla g_i(\bar{x})^T (x^k - \bar{x}) + o(\|x^k - \bar{x}\|) \\ &> 0. \end{aligned}$$

This contradicts to the feasibility of x^k . Hence $d \in \mathcal{G}(\bar{x})$ and $\mathcal{T}(\bar{x}) \subseteq \mathcal{G}(\bar{x})$. $\square$

Theorem 2.16 indicates that $\mathcal{G}(\bar{x})$ can be slightly larger than (if not equal to) $\mathcal{D}(\bar{x})$, $\mathcal{A}(\bar{x})$ and $\mathcal{T}(\bar{x})$. This means the analytic and geometric representations of the local approximation of $\mathcal{F}$ at $\bar{x}$ are close to each other. By Lemma 2.10, the inclusion relations of the dual cones of the directions at $\bar{x}$ become $\mathcal{G}(\bar{x})^* \subseteq \mathcal{T}(\bar{x})^* \subseteq \mathcal{A}(\bar{x})^* \subseteq \mathcal{D}(\bar{x})^* \subseteq \mathcal{G}^0(\bar{x})^*$.

We finish this section by reviewing the cone of gradients, which is the cone generated by the negative gradients of active inequality constraint functions.

Definition 2.17 (cone of gradients, [17]). The cone of gradients at $\bar{x}$ is

$$\mathcal{C}(\bar{x}) = \{d \in \mathbb{R}^n \mid d = - \sum_{i \in I_{ac}(\bar{x})} \mu_i \nabla g_i(\bar{x}) \text{ for some } \mu_i \geq 0, i \in I_{ac}(\bar{x})\}.$$

It is easy to prove $\mathcal{C}(\bar{x})$ and $\mathcal{G}(\bar{x})$ are dual cones of each other, i.e., $\mathcal{C}(\bar{x}) = \mathcal{G}(\bar{x})^*$ and $\mathcal{G}(\bar{x}) = \mathcal{C}(\bar{x})^*$. (See for example [16], [17].)

3. Motivation and design of constraint qualifications. In this section, we investigate the motivation of introducing a constraint qualification and its geometric meaning in developing KKT conditions for inequality constrained nonlinear programs. This leads to a unified framework of designing constraint qualifications and

a four-level categorization scheme of the constraint qualifications. Constraint qualifications for problems with both inequality and equality constraints will be discussed in Section 5.

Consider the problem (1) with $J = \emptyset$. Recall that we assume that $\bar{x} \in \mathbb{R}^n$ is an optimal solution, globally or locally. Geometrically, being optimal means that we cannot decrease the objective function value by moving $\bar{x}$ to other solutions in $\mathcal{F}$ along any feasible directions. As we know, if the steepest descent direction $-\nabla f(\bar{x})$ has a positive projection on some vector d , then d is a descent direction of $f(x)$ at $\bar{x}$. Therefore, for each feasible direction d at $\bar{x}$, $-\nabla f(\bar{x})$ should have a nonpositive projection on it. Actually, this holds for all tangent directions. We have the following lemma:

Lemma 3.1 ([47], [19], [7]). *If $\bar{x}$ is an optimal solution to problem (1) with $J = \emptyset$, then the gradient of $f(x)$ at $\bar{x}$ lies in the dual cone of tangent directions at $\bar{x}$, i.e., $\nabla f(\bar{x}) \in \mathcal{T}(\bar{x})^*$.*

Lemma 3.1 is generally difficult to check because $\mathcal{T}(\bar{x})$ and $\mathcal{T}(\bar{x})^*$ are specified by the geometric structure of the feasible region $\mathcal{F}$. Instead, analytically represented optimal conditions are preferred. Theorem 2.16 suggests that $\mathcal{G}(\bar{x})$ and its dual cone $\mathcal{G}(\bar{x})^*$ can be used to replace $\mathcal{T}(\bar{x})$ and $\mathcal{T}(\bar{x})^*$, respectively. Recall that $\mathcal{G}(\bar{x})^* \subseteq \mathcal{T}(\bar{x})^*$ and $\mathcal{G}(\bar{x})^* = \mathcal{C}(\bar{x})$. If $\mathcal{G}(\bar{x})^* = \mathcal{T}(\bar{x})^*$, then we have $\nabla f(\bar{x}) \in \mathcal{C}(\bar{x})$. Consequently, there exists $\mu_i \geq 0$ for each $i \in I_{ac}(\bar{x})$ such that

$$-\nabla f(\bar{x}) = \sum_{i \in I_{ac}(\bar{x})} \mu_i \nabla g_i(\bar{x}). \quad (4)$$

This means that the steepest descent direction (i.e., negative gradient) of the objective function, $-\nabla f(\bar{x})$, is a conic combination of the gradients of the active constraint functions. Geometrically, this means that in general the active constraints could be violated if we move $\bar{x}$ along a descent direction of $f(x)$ at $\bar{x}$. If we include the gradients of inactive constraints and set $\mu_i = 0$ for all $i \notin I_{ac}(\bar{x})$, then the equation (4) becomes equivalent to the KKT optimality conditions (3).

In summary, we have obtained the KKT conditions under the condition of $\mathcal{G}(\bar{x})^* = \mathcal{T}(\bar{x})^*$. In other words, any condition on the constraint functions $\{g_i(x), i \in I\}$ such that $\mathcal{G}(\bar{x})^* = \mathcal{T}(\bar{x})^*$ forms a constraint qualification. Geometrically, a constraint qualification requires that, for the local approximation of the feasible region $\mathcal{F}$ at $\bar{x}$, the analytic representation (the set of locally constrained directions) and the geometric representation (the set of tangent directions) be equivalent.

We now provide a framework of designing constraint qualifications. To make $\mathcal{G}(\bar{x})^* = \mathcal{T}(\bar{x})^*$ hold, we can impose conditions requiring that $\mathcal{G}(\bar{x})^*$ be equal to the dual cone of certain subsets of $\mathcal{T}(\bar{x})$. Based on the cones of directions $\mathcal{T}(\bar{x})$, $\mathcal{A}(\bar{x})$, $\mathcal{D}(\bar{x})$ and $\mathcal{G}^0(\bar{x})$ defined in Section 2, there can be four major types of constraint qualifications listed below.

1. $\mathcal{G}(\bar{x})^* = \mathcal{T}(\bar{x})^*$,
2. $\mathcal{G}(\bar{x})^* = \mathcal{A}(\bar{x})^*$,
3. $\mathcal{G}(\bar{x})^* = \mathcal{D}(\bar{x})^*$,
4. $\mathcal{G}(\bar{x})^* = \mathcal{G}^0(\bar{x})^*$.

Furthermore, we can design more constraint qualifications by adding additional assumptions, such as convexity and closeness, on the cones of directions. For example, $\mathcal{G}(\bar{x}) = \mathcal{A}(\bar{x})$ can be a constraint qualification because it implies $\mathcal{G}(\bar{x})^* = \mathcal{A}(\bar{x})^*$,

yet $\mathcal{G}(\bar{x}) = \mathcal{A}(\bar{x})$ requires the convexity of $\mathcal{A}(\bar{x})$. In Section 4, we will provide a detailed discussion about the constraint qualifications designed in this framework.

Based on the inclusion relations of $\mathcal{T}(\bar{x})$, $\mathcal{A}(\bar{x})$, $\mathcal{D}(\bar{x})$ and $\mathcal{G}^0(\bar{x})$, we categorize constraint qualifications into four levels in terms of their relative strengths. Each level corresponds to one of the four major types listed above. For example, Level 1 contains the constraint qualifications that impose conditions on $\mathcal{G}(\bar{x})$ and $\mathcal{T}(\bar{x})$ such that $\mathcal{G}(\bar{x})^* = \mathcal{T}(\bar{x})^*$ holds. The condition of $\mathcal{G}(\bar{x}) = \text{conv}(\mathcal{T}(\bar{x}))$ then belongs to this level. Obviously, Level 1 ($\mathcal{G}(\bar{x})^* = \mathcal{T}(\bar{x})^*$) is of the weakest strength and Level 4 ($\mathcal{G}(\bar{x})^* = \mathcal{G}^0(\bar{x})^*$) the strongest. In general, a weaker constraint qualification can be satisfied by more kinds of representations while a stronger constraint qualification may be easier to check. In Section 4, we will review most of the commonly seen constraint qualifications in the literature and identify the levels they belong to.

Note that, if $\mathcal{G}(\bar{x})^* = \mathcal{T}(\bar{x})^*$ is not satisfied at $\bar{x}$, we cannot make conclusion about whether KKT conditions hold at $\bar{x}$ or not. It may depend on the objective function $f(x)$. We take the following two examples to show both cases: (See also [27].)

Example 3.1.

$$\begin{aligned} \min \quad & f(x) = (x_1 - 2)^2 + (x_2 + 1)^2 \\ \text{s.t.} \quad & g_1(x) = (x_1 - 1)^3 + x_2 \leq 0, \\ & g_2(x) = -x_1 \leq 0, \\ & g_3(x) = -x_2 \leq 0, \\ & x \in \mathbb{R}^2. \end{aligned}$$

The optimal solution is $\bar{x} = (1, 0)^T$. At $\bar{x}$, we have $I_{ac}(\bar{x}) = \{1, 3\}$, $\nabla f(\bar{x}) = (-2, 2)^T$, $\nabla g_1(\bar{x}) = (0, 1)^T$ and $\nabla g_3(\bar{x}) = (0, -1)^T$. The KKT conditions fail to hold at $\bar{x}$. Moreover, $\mathcal{T}(\bar{x}) = \{d \in \mathbb{R}^2 | d_1 \leq 0, d_2 = 0\}$ and $\mathcal{G}(\bar{x}) = \{d \in \mathbb{R}^2 | d_2 = 0\}$. Therefore, $\mathcal{T}(\bar{x})^* = \{d \in \mathbb{R}^2 | d_1 \leq 0\}$ and $\mathcal{G}(\bar{x})^* = \{d \in \mathbb{R}^2 | d_1 = 0\}$. The constraint qualification $\mathcal{G}(\bar{x})^* = \mathcal{T}(\bar{x})^*$ is not satisfied.

Example 3.2.

$$\begin{aligned} \min \quad & f(x) = (x_1 - 1)^2 + (x_2 + 1)^2 \\ \text{s.t.} \quad & g_1(x) = (x_1 - 1)^3 + x_2 \leq 0, \\ & g_2(x) = -x_1 \leq 0, \\ & g_3(x) = -x_2 \leq 0, \\ & x \in \mathbb{R}^2. \end{aligned}$$

The optimal solution is $\bar{x} = (1, 0)^T$. The constraint qualification $\mathcal{G}(\bar{x})^* = \mathcal{T}(\bar{x})^*$ is not satisfied at $\bar{x}$. However, the KKT conditions hold at $\bar{x}$ because $\nabla f(\bar{x}) + 0 \cdot \nabla g_1(\bar{x}) + 2 \cdot \nabla g_3(\bar{x}) = 0$.

4. Commonly seen constraint qualifications. In this section, we continue the discussion in Section 3 about the constraint qualifications that can be designed in our framework. We review most, if not all, of the commonly seen constraint qualifications in the literatures, identify the categories they belong to and show the inter-relations among them.

4.1. Level 1: $\mathcal{G}(\bar{x})^* = \mathcal{T}(\bar{x})^*$. Constraint qualifications on the conditions of $\mathcal{G}(\bar{x})$ and $\mathcal{T}(\bar{x})$ such that $\mathcal{G}(\bar{x})^* = \mathcal{T}(\bar{x})^*$ belong to this level. Examples are $\mathcal{G}(\bar{x}) = \text{cl}(\text{conv}(\mathcal{T}(\bar{x})))$, $\mathcal{G}(\bar{x}) = \text{conv}(\text{cl}(\mathcal{T}(\bar{x})))$, $\mathcal{G}(\bar{x}) = \text{conv}(\mathcal{T}(\bar{x}))$, $\mathcal{G}(\bar{x}) = \text{cl}(\mathcal{T}(\bar{x}))$ and $\mathcal{G}(\bar{x}) = \mathcal{T}(\bar{x})$. It is known that $\text{cl}(\text{conv}(\mathcal{T}(\bar{x}))) = \text{conv}(\text{cl}(\mathcal{T}(\bar{x}))) = \text{conv}(\mathcal{T}(\bar{x}))$ [38] and easy to see $\text{cl}(\mathcal{T}(\bar{x})) = \mathcal{T}(\bar{x})$. Therefore, $\mathcal{G}(\bar{x}) = \text{cl}(\text{conv}(\mathcal{T}(\bar{x})))$ and $\mathcal{G}(\bar{x}) = \mathcal{T}(\bar{x})$ are two different types of constraint qualifications in this level. They are

referred to as Guignard's and Abadie's constraint qualifications, respectively, in the literature.

4.1.1. Guignard's constraint qualification. Guignard's constraint qualification [19] requires that $\mathcal{G}(\bar{x}) = \text{cl}(\text{conv}(\mathcal{T}(\bar{x})))$.

Gould and Tolle [16] proved that it is the necessary and sufficient condition that KKT conditions hold true for any objective function of which $\bar{x}$ is an optimal solution in $\mathcal{F} = \{x \in \mathbb{R}^n | g_i(x) \leq 0, \forall i \in I\}$. In this sense, Guignard's CQ is the weakest possible constraint qualification.

4.1.2. Abadie's constraint qualification. Abadie's constraint qualification [1] is $\mathcal{G}(\bar{x}) = \mathcal{T}(\bar{x})$.

Abadie's CQ requires that $\mathcal{T}(\bar{x})$ be a convex cone. Hence this condition is stronger than Guignard's CQ. Abadie's CQ is equivalent to a specialization of Hestenes' constraint qualification [20] in a finite-dimensional space [38].

4.2. Level 2: $\mathcal{G}(\bar{x})^* = \mathcal{A}(\bar{x})^*$. Similar to Section 4.1, we can design constraint qualifications in this level. Note that $\text{cl}(\text{conv}(\mathcal{A}(\bar{x}))) = \text{conv}(\text{cl}(\mathcal{A}(\bar{x}))) = \text{conv}(\mathcal{A}(\bar{x}))$ [38] and $\text{cl}(\mathcal{A}(\bar{x})) = \mathcal{A}(\bar{x})$. Then $\mathcal{G}(\bar{x}) = \text{cl}(\text{conv}(\mathcal{A}(\bar{x})))$ and $\mathcal{G}(\bar{x}) = \mathcal{A}(\bar{x})$ are two different types of constraint qualifications in this level. They are seen in the literature as AHU's constraint qualification and Kuhn-Tucker's constraint qualification, respectively.

4.2.1. AHU's constraint qualification. Arrow-Hurwicz-Uzawa's constraint qualification (AHU's CQ) [4] is $\mathcal{G}(\bar{x}) = \text{cl}(\text{conv}(\mathcal{A}(\bar{x})))$.

4.2.2. Kuhn-Tucker's constraint qualification. Kuhn-Tucker's constraint qualification [4] is $\mathcal{G}(\bar{x}) = \mathcal{A}(\bar{x})$.

Kuhn-Tucker's CQ is equivalent to Hurwicz's constraint qualification [22] specialized in a finite-dimensional space [4]. Note that the definition of $\mathcal{A}(\bar{x})$ in Kuhn and Tucker's original paper [27] was slightly different. In the definition of attainable directions in [27], $\alpha(\delta)$ was required to be differentiable everywhere in $[0, 1]$. This requirement was relaxed in [4] which is used here.

4.3. Level 3: $\mathcal{G}(\bar{x})^* = \mathcal{D}(\bar{x})^*$. The constraint qualifications in this level can be designed in the same way as in Sections 4.1 and 4.2. Note that $\text{cl}(\text{conv}(\mathcal{D}(\bar{x}))) = \text{conv}(\text{cl}(\mathcal{D}(\bar{x})))$ [38]. Then, $\mathcal{G}(\bar{x}) = \text{cl}(\text{conv}(\mathcal{D}(\bar{x})))$, $\mathcal{G}(\bar{x}) = \text{cl}(\mathcal{D}(\bar{x}))$, $\mathcal{G}(\bar{x}) = \text{conv}(\mathcal{D}(\bar{x}))$ and $\mathcal{G}(\bar{x}) = \mathcal{D}(\bar{x})$ are four different types of constraint qualifications in this level. Especially, $\mathcal{G}(\bar{x}) = \text{cl}(\mathcal{D}(\bar{x}))$ is seen in the literature as Zangwill's constraint qualification. The other two types may be viewed as new constraint qualifications.

4.3.1. Zangwill's constraint qualification. Zangwill's constraint qualification [48] is $\mathcal{G}(\bar{x}) = \text{cl}(\mathcal{D}(\bar{x}))$.

4.4. Level 4: $\mathcal{G}(\bar{x})^* = \mathcal{G}^0(\bar{x})^*$. Note that, if $\mathcal{G}^0(\bar{x}) = \emptyset$, then $\text{cl}(\text{conv}(\mathcal{G}^0(\bar{x}))) = \emptyset$. If $\mathcal{G}^0(\bar{x}) \neq \emptyset$, then $\mathcal{G}(\bar{x}) = \text{cl}(\text{conv}(\mathcal{G}^0(\bar{x}))) = \text{conv}(\text{cl}(\mathcal{G}^0(\bar{x}))) = \text{cl}(\mathcal{G}^0(\bar{x}))$ and $\text{conv}(\mathcal{G}^0(\bar{x})) = \mathcal{G}^0(\bar{x}) \neq \mathcal{G}(\bar{x})$. Therefore, $\mathcal{G}(\bar{x}) = \text{cl}(\mathcal{G}^0(\bar{x}))$ (or equivalently $\mathcal{G}^0(\bar{x}) \neq \emptyset$) is a constraint qualification in this level. Besides, a constraint qualification that is a sufficient condition for $\text{cl}(\mathcal{G}^0(\bar{x})) = \mathcal{G}(\bar{x})$ can also be categorized in this level. Such constraint qualifications include Slater's constraint qualification and the linear independence constraint qualification.

4.4.1. *Cottle's constraint qualification.* Cottle's constraint qualification [12] requires that $\mathcal{G}(\bar{x}) = \text{cl}(\mathcal{G}^0(\bar{x}))$.

Note that Cottle's CQ is equivalent to $\mathcal{G}^0(\bar{x}) \neq \emptyset$.

4.4.2. *Variants of $\mathcal{G}^0(\bar{x})$ and corresponding constraint qualifications.* We review two constraint qualifications that are closely related to Cottle's CQ. There exist other cones of directions in the literature that are between $\mathcal{G}^0(\bar{x})$ and $\mathcal{D}(\bar{x})$. The active constraint functions are classified by linear/nonlinear or pseudoconcave/non-pseudoconcave ones in this approach. They can be viewed as slightly relaxed variants of $\mathcal{G}^0(\bar{x})$. Define $\mathcal{G}^1(\bar{x})$ and $\mathcal{G}^2(\bar{x})$ as

$$\mathcal{G}^1(\bar{x}) = \left\{ d \in \mathbb{R}^n \mid \begin{array}{l} \nabla g_i(\bar{x})^T d < 0, \text{ if } g_i(x) \text{ is nonlinear at } \bar{x}, \\ \nabla g_i(\bar{x})^T d \leq 0, \text{ if } g_i(x) \text{ is linear at } \bar{x}, i \in I_{ac}(\bar{x}). \end{array} \right\}$$

$$\mathcal{G}^2(\bar{x}) = \left\{ d \in \mathbb{R}^n \mid \begin{array}{l} \nabla g_i(\bar{x})^T d < 0, \text{ if } g_i(x) \text{ is nonpseudoconcave at } \bar{x}, \\ \nabla g_i(\bar{x})^T d \leq 0, \text{ if } g_i(x) \text{ is pseudoconcave at } \bar{x}, i \in I_{ac}(\bar{x}). \end{array} \right\}$$

Abadie's second constraint qualification [1] is $\mathcal{G}^1(\bar{x}) \neq \emptyset$. AHU's second constraint qualification [4] is $\mathcal{G}^2(\bar{x}) \neq \emptyset$.

The following lemma and corollary reveal the inter-relations among Abadie's second CQ, AHU's second CQ and other constraint qualifications:

Lemma 4.1. *For problem (1) with $J = \emptyset$, let $\bar{x}$ be an optimal solution. Then $\mathcal{G}^0(\bar{x}) \subseteq \mathcal{G}^1(\bar{x}) \subseteq \mathcal{G}^2(\bar{x}) \subseteq \mathcal{D}(\bar{x})$.*

Proof. It's easy to see that $\mathcal{G}^0(\bar{x}) \subseteq \mathcal{G}^1(\bar{x}) \subseteq \mathcal{G}^2(\bar{x})$. Now we prove $\mathcal{G}^2(\bar{x}) \subseteq \mathcal{D}(\bar{x})$. If $d \in \mathcal{G}^2(\bar{x})$, we claim that $g_i(\bar{x} + \delta d) \leq 0$ with δ being sufficiently small for all $i \in I_{ac}(\bar{x})$ such that $g_i(x)$ is pseudoconcave at $\bar{x}$. Otherwise, $g_i(\bar{x} + \delta d) > 0 = g_i(\bar{x})$ and then $\delta \nabla g_i(\bar{x})^T d > 0$ by the pseudoconcavity. This becomes a contradiction. Hence $d \in \mathcal{D}(\bar{x})$ and $\mathcal{G}^2(\bar{x}) \subseteq \mathcal{D}(\bar{x})$. $\square$

Corollary 4.1. *For problem (1) with $J = \emptyset$, Cottle's CQ $\Rightarrow$ Abadie's second CQ $\Rightarrow$ AHU's second CQ $\Rightarrow$ Zangwill's CQ.*

4.4.3. *Slater's constraint qualification.* Slater's constraint qualification [44] requires that $g_i(x)$ be pseudoconvex at $\bar{x}$, $\forall i \in I_{ac}(\bar{x})$, and there exists $\tilde{x} \in \mathbb{R}^n$ such that $g_i(\tilde{x}) < 0, \forall i \in I_{ac}(\bar{x})$.

Lemma 4.2. *For problem (1) with $J = \emptyset$, Slater's CQ implies Cottle's CQ.*

Proof. Suppose that Slater's CQ is satisfied. For $i \in I_{ac}(\bar{x})$, we have $g_i(\tilde{x}) < g_i(\bar{x})$. Therefore, $\nabla g_i(\bar{x})^T (\tilde{x} - \bar{x}) < 0$ because $g_i(x)$ is pseudoconvex at $\bar{x}$. Let $d = \tilde{x} - \bar{x}$, then $d \in \mathcal{G}^0(\bar{x})$. This means $\mathcal{G}^0(\bar{x}) \neq \emptyset$ and Cottle's CQ holds. $\square$

4.4.4. *Linear independence constraint qualification.* Linearly independence constraint qualification (LICQ) is that the gradients $\{\nabla g_i(\bar{x}), i \in I_{ac}(\bar{x})\}$ are linearly independent.

Lemma 4.3. *For problem (1) with $J = \emptyset$, LICQ implies Cottle's CQ.*

Proof. If LICQ is satisfied at $\bar{x}$, we let

$$\mathbf{M} = \{\nabla g_i(\bar{x}), i \in I_{ac}(\bar{x})\}^T$$

be the matrix whose rows consist of the gradients of the active constraint functions and $b = (-1, -1, \dots, -1)^T$ as the vector that has $|I_{ac}(\bar{x})|$ number of -1 's. Then the linear equation $\mathbf{M}d = b$ has a solution d . Therefore $d \in \mathcal{G}^0(\bar{x})$ and Cottle's CQ is satisfied. $\square$

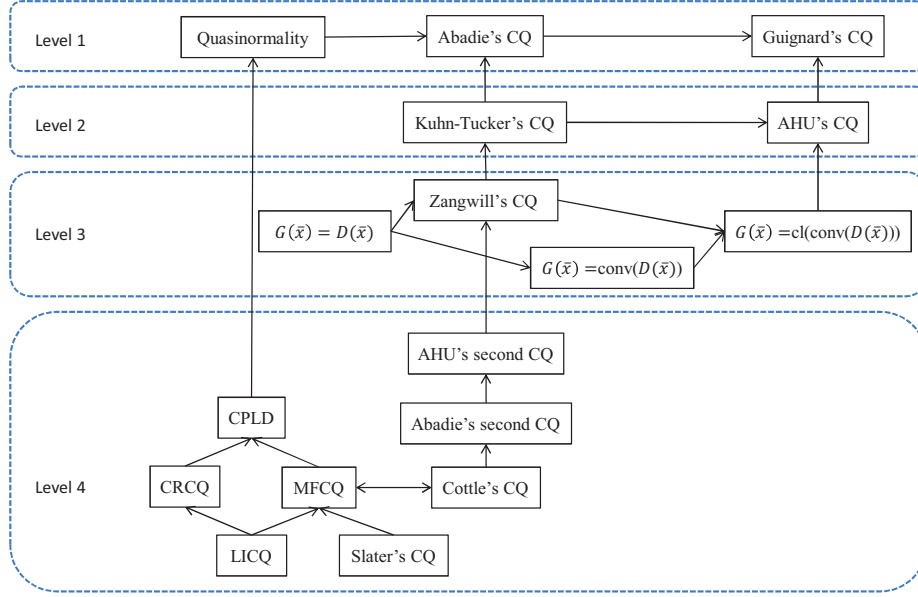


FIGURE 1. Inter-relations among constraint qualifications for problem (1) with $J = \emptyset$.

In Figure 1, we summarize the inter-relations among the constraint qualifications studied in this section using the proposed four-level categorization scheme.

5. Problem with inequality and equality constraints. In this section, we review the constraint qualifications for problem (1) with both inequality and equality constraints, i.e., problem (1) with $J \neq \emptyset$. For this problem, the KKT conditions (3) become

$$\begin{cases} \nabla f(\bar{x}) + \sum_{i \in I} \mu_i \nabla g_i(\bar{x}) + \sum_{j \in J} \nu_j \nabla h_j(\bar{x}) = 0, \\ \mu_i g_i(\bar{x}) = 0, \quad i \in I, \\ \mu_i \geq 0, \quad i \in I. \end{cases}$$

The approach we adopt to design constraint qualifications for problem (1) with $J \neq \emptyset$ is to replace each equality constraint $h_j(x) = 0$ by two inequality constraints $h_j(x) \leq 0$ and $-h_j(x) \leq 0$. The reformulation of problem (1) becomes

$$\begin{aligned} \min \quad & f(x) \\ \text{s.t.} \quad & g_i(x) \leq 0, \quad i \in I, \\ & h_j(x) \leq 0, \quad j \in J, \\ & -h_j(x) \leq 0, \quad j \in J, \\ & x \in \mathbb{R}^n. \end{aligned} \tag{5}$$

Consider the cones of directions $\mathcal{T}(\bar{x})$, $\mathcal{A}(\bar{x})$, $\mathcal{D}(\bar{x})$, $\mathcal{G}(\bar{x})$ and $\mathcal{G}^0(\bar{x})$ for problem (5). Note that the forms of $\mathcal{T}(\bar{x})$, $\mathcal{A}(\bar{x})$ and $\mathcal{D}(\bar{x})$ remain the same as defined in Section 2 because they are determined only by the geometric structure of $\mathcal{F}$. $\mathcal{D}(\bar{x})$ could be the set of zero vector $\{0\}$ if $h_j(x)$ is nonlinear for some $j \in J$. It is easy to see $\mathcal{G}(\bar{x})$ is of the form

$$\{d \in \mathbb{R}^n \mid \nabla g_i(\bar{x})^T d \leq 0, \forall i \in I_{ac}(\bar{x}), \nabla h_j(\bar{x})^T d = 0, \forall j \in J\}$$

and its dual cone $\mathcal{C}(\bar{x})$ is of the form

$$\{d \in \mathbb{R}^n | d = - \sum_{i \in I_{ac}(\bar{x})} \mu_i \nabla g_i(\bar{x}) + \sum_{j \in J} \nu_j \nabla h_j(\bar{x}) \text{ for some } \mu_i \geq 0, i \in I_{ac}(\bar{x}), \nu_j \in \mathbb{R}\}.$$

The major difference between the cones of directions at $\bar{x}$ for problem (5) and problem (1) with $J = \emptyset$ is that $\mathcal{G}^0(\bar{x})$ is always the empty set in (5) because $\mathcal{G}(\bar{x})$ has no interior points. Therefore, $\text{cl}(\mathcal{G}^0(\bar{x})) = \mathcal{G}(\bar{x})$ fails to hold. In this sense, we use the relative interior of $\mathcal{G}(\bar{x})$ to replace the role of $\mathcal{G}^0(\bar{x})$.

Definition 5.1 (relative interior constrained direction). For problem (1) with $J \neq \emptyset$, a vector d is a relative interior constrained direction at $\bar{x}$ if $\nabla g_i(\bar{x})^T d < 0$ holds, $\forall i \in I_{ac}(\bar{x})$, and $\nabla h_j(\bar{x})^T d = 0$ holds, $\forall j \in J$.

Note that the set of relative interior constrained directions at $\bar{x}$ is a cone. We denote the cone of relative interior constrained directions at $\bar{x}$ by $\mathcal{H}^0(\bar{x})$.

It is easy to check that $\mathcal{H}^0(\bar{x})$ is the relative interior of $\mathcal{G}(\bar{x})$. Hence $\text{cl}(\mathcal{H}^0(\bar{x})) = \mathcal{G}(\bar{x})$ if and only if $\mathcal{H}^0(\bar{x}) \neq \emptyset$. To replace the role of $\mathcal{G}^0(\bar{x})$ by $\mathcal{H}^0(\bar{x})$ in the inclusion relations of Theorem 2.16, we need additional conditions. For example, if the gradients of equality constraints $\{\nabla h_j(\bar{x}), j \in J\}$ are linearly independent, then $\mathcal{H}^0(\bar{x}) \subseteq \mathcal{A}(\bar{x}) \subseteq \mathcal{T}(\bar{x})$. Similarly, if $\{h_j(x), j \in J\}$ are affine and $\{\nabla h_j(\bar{x}), j \in J\}$ are linearly independent, then $\mathcal{H}^0(\bar{x}) \subseteq \mathcal{D}(\bar{x})$.

Remark 5.1 (on the condition of $\{\nabla h_j(\bar{x}), j \in J\}$ being linearly independent). Let $\mathcal{F}_1 = \{x \in \mathbb{R}^n | g_i(x) \leq 0, i \in I\}$ and $\mathcal{F}_2 = \{x \in \mathbb{R}^n | h_i(x) = 0, j \in J\}$. Denote the cones of tangent directions of $\mathcal{F}_1$ and $\mathcal{F}_2$ at $\bar{x}$ by $\mathcal{T}_{\mathcal{F}_1}(\bar{x})$ and $\mathcal{T}_{\mathcal{F}_2}(\bar{x})$, respectively, and denote the cones of locally constrained directions of $\mathcal{F}_1$ and $\mathcal{F}_2$ at $\bar{x}$ by $\mathcal{G}_{\mathcal{F}_1}(\bar{x})$ and $\mathcal{G}_{\mathcal{F}_2}(\bar{x})$, respectively. Moreover, denote the cones of interior constrained directions of $\mathcal{F}_1$ and $\mathcal{F}_2$ at $\bar{x}$ by $\mathcal{G}_{\mathcal{F}_1}^0(\bar{x})$ and $\mathcal{G}_{\mathcal{F}_2}^0(\bar{x})$, respectively. It is easy to verify that $\mathcal{F} = \mathcal{F}_1 \cap \mathcal{F}_2$, $\mathcal{T}(\bar{x}) = \mathcal{T}_{\mathcal{F}_1}(\bar{x}) \cap \mathcal{T}_{\mathcal{F}_2}(\bar{x})$, $\mathcal{G}(\bar{x}) = \mathcal{G}_{\mathcal{F}_1}(\bar{x}) \cap \mathcal{G}_{\mathcal{F}_2}(\bar{x})$ and $\mathcal{G}^0(\bar{x}) = \mathcal{G}_{\mathcal{F}_1}^0(\bar{x}) \cap \mathcal{G}_{\mathcal{F}_2}^0(\bar{x})$. Note that $\mathcal{G}^0(\bar{x}) = \emptyset$ because $\mathcal{G}_{\mathcal{F}_2}^0(\bar{x}) = \emptyset$. The relative interior constrained directions $\mathcal{H}^0(\bar{x}) = \mathcal{G}_{\mathcal{F}_1}^0(\bar{x}) \cap \mathcal{G}_{\mathcal{F}_2}(\bar{x})$. The necessary and sufficient condition of $\mathcal{H}^0(\bar{x}) \subseteq \mathcal{T}(\bar{x})$ becomes $\mathcal{G}_{\mathcal{F}_2}(\bar{x}) \subseteq \mathcal{T}_{\mathcal{F}_2}(\bar{x})$, or equivalently, $\mathcal{G}_{\mathcal{F}_2}(\bar{x}) = \mathcal{T}_{\mathcal{F}_2}(\bar{x})$. If the dimension of $\mathcal{T}_{\mathcal{F}_2}(\bar{x})$ is m and there exist $n - m$ linearly independent vectors in $\{\nabla h_j(\bar{x}), j \in J\}$, then $\mathcal{G}_{\mathcal{F}_2}(\bar{x}) = \mathcal{T}_{\mathcal{F}_2}(\bar{x})$ (See for example [31]). Note that $n - m \leq |J|$. Therefore, the linear independence of $\{\nabla h_j(\bar{x}), j \in J\}$ implies $\mathcal{G}_{\mathcal{F}_2}(\bar{x}) = \mathcal{T}_{\mathcal{F}_2}(\bar{x})$. Under this condition, we have $\mathcal{H}^0(\bar{x}) \subseteq \mathcal{T}(\bar{x})$. Similar discussion can be applied for $\mathcal{H}^0(\bar{x}) \subseteq \mathcal{A}(\bar{x})$.

Remark 5.2 (on the condition of $\{h_j(x), j \in J\}$ being affine). $\mathcal{D}(\bar{x})$ may contain no straight line segments if some of the equality constraint functions are nonlinear. In this case, $\mathcal{D}(\bar{x}) = \{0\}$. If $\mathcal{H}^0(\bar{x}) \neq \{0\}$, then $\mathcal{H}^0(\bar{x}) \subseteq \mathcal{D}(\bar{x})$ cannot hold. It is easy to verify that if $\{h_j(x), j \in J\}$ are affine and $\{\nabla h_j(\bar{x}), j \in J\}$ are linearly independent, then $\mathcal{H}^0(\bar{x}) \subseteq \mathcal{D}(\bar{x})$.

We summarize the discussion above in the following theorem:

Theorem 5.2. If $\bar{x}$ is an optimal solution to problem (1) with $J \neq \emptyset$, then

- (i) $\emptyset = \mathcal{G}^0(\bar{x}) \subseteq \mathcal{D}(\bar{x}) \subseteq \mathcal{A}(\bar{x}) \subseteq \mathcal{T}(\bar{x}) \subseteq \mathcal{G}(\bar{x})$;
- (ii) $\text{cl}(\mathcal{H}^0(\bar{x})) = \mathcal{G}(\bar{x})$ if and only if $\mathcal{H}^0(\bar{x}) \neq \emptyset$;
- (iii) If $\{\nabla h_j(\bar{x}), j \in J\}$ are linearly independent, then $\mathcal{H}^0(\bar{x}) \subseteq \mathcal{A}(\bar{x}) \subseteq \mathcal{T}(\bar{x}) \subseteq \mathcal{G}(\bar{x})$;
- (iv) If $\{h_j(x), j \in J\}$ are affine and $\{\nabla h_j(\bar{x}), j \in J\}$ are linearly independent, then $\mathcal{H}^0(\bar{x}) \subseteq \mathcal{D}(\bar{x}) \subseteq \mathcal{A}(\bar{x}) \subseteq \mathcal{T}(\bar{x}) \subseteq \mathcal{G}(\bar{x})$.

Next, we generalize the constraint qualifications in Sections 3 and 4 for problem (1) with $J \neq \emptyset$. The constraint qualifications designed in the sense of quasinormality of $\bar{x}$, including the quasinormality condition, the constant positive linear independence constraint qualification (CPLD) and the constant rank constraint qualification (CRCQ), are also reviewed. These constraint qualifications require properties of the constraints not only at $\bar{x}$ but also in a neighborhood of $\bar{x}$. Historically, they were not explicitly specified by the cones of directions at $\bar{x}$. We compare these constraint qualifications with the constraint qualifications designed in our framework to categorize them into appropriate levels by their relative strengths.

5.1. Level 1: $\mathcal{G}(\bar{x})^* = \mathcal{T}(\bar{x})^*$. In this level, Guignard's CQ of $\mathcal{G}(\bar{x}) = \text{cl}(\text{conv}(\mathcal{T}(\bar{x})))$ and Abadie's CQ of $\mathcal{G}(\bar{x}) = \mathcal{T}(\bar{x})$ can be generalized. The quasinormality condition is also categorized in this level because it is a relatively weak condition that implies Abadie's CQ. However, quasinormality does not admit an intuitive geometric interpretation [8].

5.1.1. Quasinormality condition. The quasinormality condition [21] requires that there exist no nonzero vector $(\mu, \nu) \in \mathbb{R}^{|I_{ac}(\bar{x})|+|J|}$ and no sequence $\{x^k\} \subseteq \mathbb{R}^n$ such that

- (i) $\sum_{i \in I_{ac}(\bar{x})} \mu_i \nabla g_i(\bar{x}) + \sum_{j \in J} \nu_j \nabla h_j(\bar{x}) = 0$;
- (ii) $\mu_i \geq 0, \forall i \in I_{ac}(\bar{x})$;
- (iii) $x^k \rightarrow \bar{x}$ as $k \rightarrow \infty$;
- (iv) $\mu_i g_i(x^k) > 0$ for all k and all i with $\mu_i \neq 0$, and $\nu_j h_j(x^k) > 0$ for all k and all j with $\nu_j \neq 0$.

The quasinormality condition implies Abadie's CQ [21]. Bertsekas and Ozdaglar [9], [37] extended the quasinormality condition and proposed the pseudonormality condition that is stronger than the quasinormality condition.

5.2. Level 2: $\mathcal{G}(\bar{x})^* = \mathcal{A}(\bar{x})^*$. Similar to Guignard's CQ and Abadie's CQ in Level 1, AHU's CQ of $\mathcal{G}(\bar{x}) = \text{cl}(\text{conv}(\mathcal{A}(\bar{x})))$ and Kuhn-Tucker's CQ of $\mathcal{G}(\bar{x}) = \mathcal{A}(\bar{x})$ can be directly generalized in this level.

5.3. Level 3: $\mathcal{G}(\bar{x})^* = \mathcal{D}(\bar{x})^*$. In this level, the four constraint qualifications $\mathcal{G}(\bar{x}) = \text{cl}(\text{conv}(\mathcal{D}(\bar{x})))$, $\mathcal{G}(\bar{x}) = \text{cl}(\mathcal{D}(\bar{x}))$, $\mathcal{G}(\bar{x}) = \text{conv}(\mathcal{D}(\bar{x}))$ and $\mathcal{G}(\bar{x}) = \mathcal{D}(\bar{x})$ can still be used for problem (1) with $J \neq \emptyset$. Zangwill's CQ is $\mathcal{G}(\bar{x}) = \text{cl}(\mathcal{D}(\bar{x}))$.

5.4. Level 4: $\mathcal{G}(\bar{x})^* = \mathcal{H}^0(\bar{x})^*$. When $\mathcal{H}^0(\bar{x}) \subseteq \mathcal{T}(\bar{x})$, we can replace $\mathcal{G}(\bar{x})^* = \mathcal{G}^0(\bar{x})^*$ by $\mathcal{G}(\bar{x})^* = \mathcal{H}^0(\bar{x})^*$ in Level 4. Recall that the linear independence of $\{\nabla h_j(\bar{x}), j \in J\}$ is a sufficient condition of $\mathcal{H}^0(\bar{x}) \subseteq \mathcal{T}(\bar{x})$. Under this condition, Cottle's CQ (equivalent to Mangasarian-Fromovitz's CQ in Section 5.4.2), Slater's CQ and LICQ can be generalized for problem (1). Note that Cottle's CQ implies Zangwill's CQ under the condition that $\{h_j(x), j \in J\}$ are affine. As will be seen, CPLD and CRCQ can be viewed as relaxations of MFCQ and LICQ. They are then categorized in this level.

5.4.1. Cottle's constraint qualification. Cottle's CQ requires that $\{\nabla h_j(\bar{x}), j \in J\}$ are linearly independent and $\mathcal{G}(\bar{x}) = \text{cl}(\mathcal{H}^0(\bar{x}))$.

5.4.2. *Mangasarian-Fromovitz's constraint qualification.* Mangasarian-Fromovitz's constraint qualification (MFCQ) [34] requires that $\{\nabla h_j(\bar{x}), j \in J\}$ are linearly independent and $\mathcal{H}^0(\bar{x}) \neq \emptyset$.

By Theorem 5.2, MFCQ is equivalent to Cottle's CQ. MFCQ is also equivalent to the positive-linear independence condition of the vectors in $\{\nabla g_i(\bar{x}), i \in I_{ac}(\bar{x})\} \cup \{\nabla h_j(\bar{x}), j \in J\}$. That is, the only solution of the linear system

$$\begin{aligned} \sum_{i \in I_{ac}(\bar{x})} \mu_i \nabla g_i(\bar{x}) + \sum_{j \in J} \nu_j \nabla h_j(\bar{x}) &= 0, \\ \mu_i &\geq 0, \quad \forall i \in I_{ac}(\bar{x}) \end{aligned}$$

is the zero vector. (See [34], [42]).

5.4.3. *Variants of $\mathcal{H}^0(\bar{x})$ and corresponding constraint qualifications.* Similar to Section 4.4.2, we can generalize Abadie's second CQ and AHU's second CQ by introducing some variants of $\mathcal{H}^0(\bar{x})$. Define $\mathcal{H}^1(\bar{x})$ and $\mathcal{H}^2(\bar{x})$ as

$$\begin{aligned} \mathcal{H}^1(\bar{x}) &= \left\{ d \in \mathbb{R}^n \left| \begin{array}{l} \nabla h_j(\bar{x})^T d = 0, j \in J, \\ \nabla g_i(\bar{x})^T d < 0, \text{ if } g_i(x) \text{ is nonlinear at } \bar{x}, \\ \nabla g_i(\bar{x})^T d \leq 0, \text{ if } g_i(x) \text{ is linear at } \bar{x}, i \in I_{ac}(\bar{x}). \end{array} \right. \right\} \\ \mathcal{H}^2(\bar{x}) &= \left\{ d \in \mathbb{R}^n \left| \begin{array}{l} \nabla h_j(\bar{x})^T d = 0, j \in J, \\ \nabla g_i(\bar{x})^T d < 0, \text{ if } g_i(x) \text{ is nonpseudoconcave at } \bar{x}, \\ \nabla g_i(\bar{x})^T d \leq 0, \text{ if } g_i(x) \text{ is pseudoconcave at } \bar{x}, i \in I_{ac}(\bar{x}). \end{array} \right. \right\} \end{aligned}$$

Abadie's second CQ becomes that $\{\nabla h_j(\bar{x}), j \in J\}$ are linearly independent and $\mathcal{H}^1(\bar{x}) \neq \emptyset$. AHU's second CQ becomes that $\{\nabla h_j(\bar{x}), j \in J\}$ are linearly independent and $\mathcal{H}^2(\bar{x}) \neq \emptyset$. Lemma 4.1 and Corollary 4.1 can also be generalized by considering Theorem 5.2.

5.4.4. *Constant positive linear dependence constraint qualification.* The constant positive linear dependence constraint qualification (CPLD) [39] requires that for any $\tilde{I} \subseteq I_{ac}(\bar{x})$ and any $\tilde{J} \subseteq J$, if there exists a nonzero solution to the linear system

$$\begin{aligned} \sum_{i \in \tilde{I}} \mu_i \nabla g_i(\bar{x}) + \sum_{j \in \tilde{J}} \nu_j \nabla h_j(\bar{x}) &= 0, \\ \mu &\geq 0, \quad \forall i \in \tilde{I}, \end{aligned}$$

then there exists a neighborhood $\mathcal{N}(\bar{x})$ of $\bar{x}$ such that $\{\nabla g_i(x), \nabla h_j(x), i \in \tilde{I}, j \in \tilde{J}\}$ are linear dependent for all $x \in \mathcal{N}(\bar{x})$.

CPLD provides a condition that the constraint functions may satisfy when the vectors in $\{\nabla g_i(\bar{x}), i \in I_{ac}(\bar{x})\} \cup \{\nabla h_j(\bar{x}), j \in J\}$ are positive-linear dependent. In this sense, CPLD can be viewed as a relaxation of MFCQ. Actually, MFCQ implies CPLD and CPLD implies the quasinormality condition [3].

Remark 5.3 (Relaxed CPLD). The relaxed constant positive linear dependence constraint qualification (RCPLD) requires that for any $\tilde{I} \subseteq I_{ac}(\bar{x})$ and some $\tilde{J} \subseteq J$ such that $\{\nabla h_j(\bar{x}), j \in \tilde{J}\}$ forms a basis of $\{\nabla h_j(\bar{x}), j \in J\}$, if there exists a nonzero solution to the linear system

$$\begin{aligned} \sum_{i \in \tilde{I}} \mu_i \nabla g_i(\bar{x}) + \sum_{j \in \tilde{J}} \nu_j \nabla h_j(\bar{x}) &= 0, \\ \mu &\geq 0, \quad \forall i \in \tilde{I}, \end{aligned}$$

then there exists a neighborhood $\mathcal{N}(\bar{x})$ of $\bar{x}$ such that $\{\nabla g_i(x), \nabla h_j(x), i \in \tilde{I}, j \in \tilde{J}\}$ are linear dependent for all $x \in \mathcal{N}(\bar{x})$ and $\{\nabla h_j(x), j \in J\}$ has the same rank for all $x \in \mathcal{N}(\bar{x})$.

RCPLD was proposed by Andreani et al. [2]. Note that CPLD imposes conditions on all subsets of J while RCPLD only needs a certain one of them. It was proved that CPLD implies RCPLD and RCPLD implies Abadie's CQ [2].

5.4.5. Slater's constraint qualification. Slater's CQ is that $g_i(x)$ is pseudoconvex at $\bar{x}$, $\forall i \in I_{ac}(\bar{x})$, $h_j(x)$ is both quasiconvex and quasiconcave at $\bar{x}$, $\forall j \in J$, $\{\nabla h_j(\bar{x}), j \in J\}$ are linearly independent, and there exists $\tilde{x} \in \mathbb{R}^n$ such that $g_i(\tilde{x}) < 0, \forall i \in I_{ac}(x^*)$, and $h_j(\tilde{x}) = 0, \forall j \in J$.

Lemma 5.3. *For problem (1) with $J \neq \emptyset$, Slater's CQ implies Cottle's CQ.*

Proof. Let $d = \tilde{x} - \bar{x}$. We only need to prove $\nabla h_j(\bar{x})^T d = 0$ for all $j \in J$. For each $h_j(x)$ and all $\lambda \in [0, 1]$, we have $h_j(\lambda\tilde{x} + (1-\lambda)\bar{x}) \leq \max\{h_j(\bar{x}), h_j(\tilde{x})\} = 0$ because $h_j(x)$ is quasiconvex at $\bar{x}$. Note that $h_j(x)$ is also quasiconcave at $\bar{x}$. Hence $h_j(\lambda\tilde{x} + (1-\lambda)\bar{x}) \geq \min\{h_j(\bar{x}), h_j(\tilde{x})\} = 0$. Consequently, $h_j(\lambda\tilde{x} + (1-\lambda)\bar{x}) = 0$. By the Taylor series expansions of $h_j(\lambda\tilde{x} + (1-\lambda)\bar{x})$ at $\bar{x}$, we have

$$\lambda \nabla h_j(\bar{x})^T d + o(\lambda) \|d\| = 0, \text{ as } \lambda \rightarrow 0.$$

Thus, $\nabla h_j(\bar{x})^T d = 0$. Similar to the proof of Lemma 4.2, we can prove that $\nabla g_i(\bar{x})^T d < 0, \forall i \in I_{ac}(\bar{x})$. Therefore, $d \in \mathcal{H}^0(\bar{x})$ and $\mathcal{H}^0(\bar{x}) \neq \emptyset$. Cottle's CQ is then satisfied. $\square$

5.4.6. Linear independence constraint qualification. Linearly independence constraint qualification (LICQ) requires that the gradients $\{\nabla g_i(\bar{x}), \nabla h_j(\bar{x}), i \in I_{ac}(\bar{x}), j \in J\}$ are linearly independent.

Obviously, LICQ implies MFCQ and Cottle's CQ.

5.4.7. Constant rank constraint qualification. The constant rank constraint qualification (CRCQ) [23] requires that there exists a neighborhood $\mathcal{N}(\bar{x})$ of $\bar{x}$ such that for each $\tilde{I} \subseteq I_{ac}(\bar{x})$ and each $\tilde{J} \subseteq J$, the set of gradients $\{\nabla g_i(x), \nabla h_j(x), i \in \tilde{I}, j \in \tilde{J}\}$ has the same rank for all $x \in \mathcal{N}(\bar{x})$.

CRCQ provides a condition that the constraint functions may satisfy when $\{\nabla g_i(\bar{x}), i \in I_{ac}(\bar{x}), \nabla h_j(\bar{x}), j \in J\}$ are linearly dependent. In this sense, CRCQ can be viewed as a relaxation of LICQ. For the relation between CRCQ and MFCQ, it was proved that CRCQ is neither weaker nor stronger than MFCQ [23]. Lu [30] proved that MFCQ and CRCQ are related in the following sense: if CRCQ holds, then MFCQ holds in an alternative representation of $\mathcal{F}$. Moreover, CRCQ implies CPLD [3].

Remark 5.4 (Relaxed CRCQ). The relaxed constant rank constraint qualification (RCRCQ) requires that there exists a neighborhood $\mathcal{N}(\bar{x})$ of $\bar{x}$ such that for each $\tilde{I} \subseteq I_{ac}(\bar{x})$, the set of gradients $\{\nabla g_i(x), \nabla h_j(x), i \in \tilde{I}, j \in J\}$ has the same rank for all $x \in \mathcal{N}(\bar{x})$.

RCRCQ was proposed by Minchenko and Stakhovski [36]. Note that CRCQ imposes conditions on all subsets of J while RCRCQ only needs J itself. Hence CRCQ implies RCRCQ. Moreover, RCRCQ implies RCPLD [2].

Note that the implication relation of AHU's second CQ $\Rightarrow$ Zangwill's CQ holds under the condition that $\{h_j(\bar{x}), j \in J\}$ are affine. In Figure 2, we summarize the inter-relations among the constraint qualifications studied in this section using the proposed four-level categorization scheme.

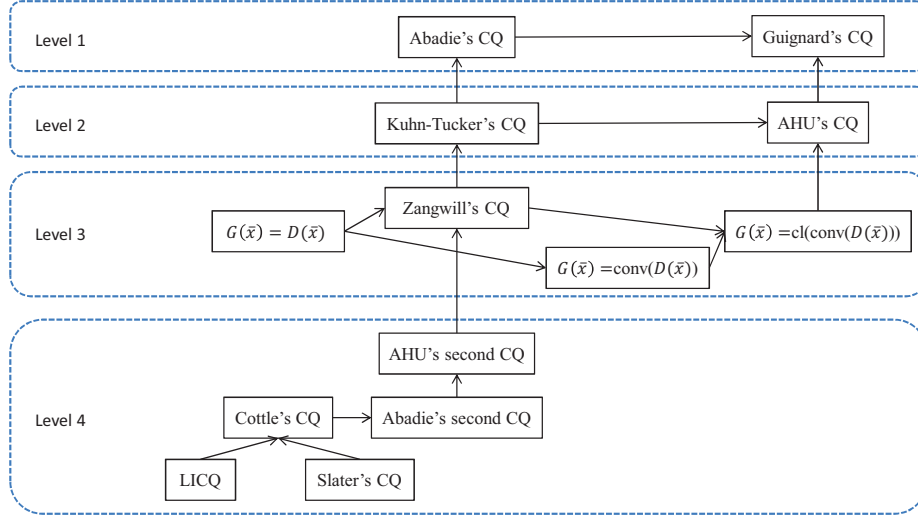


FIGURE 2. Inter-relations among constraint qualifications for problem (1) with $J \neq \emptyset$.

6. Conclusions. We have investigated the motivation of introducing constraint qualifications in developing KKT conditions for nonlinear programs and provided a geometric meaning of constraint qualifications. A unified framework of designing constraint qualifications has been proposed. This framework leads to a four-level categorization scheme of constraint qualifications by their relative strengths. By applying the scheme, we have reviewed most, if not all, of the commonly seen constraint qualifications in the literature and demonstrated the inter-relationship among them.

This paper reveals the essence of constraint qualifications and the important role of constraint qualifications in nonlinear programming. The proposed framework provides a systematic view of existing constraint qualifications in the literature and helps design new constraint qualifications of readers' specific interests. The four-level categorization scheme can be used to compare different constraint qualifications in terms of their relative strengths and mutual relations.

Constraint qualifications are essential in many aspects of nonlinear programming, such as duality theory, sensitivity analysis and convergence of computational methods. Studies on motivation and interpretation of constraint qualifications in these aspects are necessary and helpful. Extended research may include reviewing constraint qualifications for nonlinear programs in infinite-dimensional spaces and investigating abstract optimality conditions without constraint qualifications.

REFERENCES

- [1] J. Abadie, *On the Kuhn-Tucker theorem*, in "Nonlinear Programming" (ed. J. Abadie), North-Holland Pub. Co., (1967), 19–36.
- [2] R. Andreani, G. Haeser, M. L. Schuverdt and P. J. S. Silva, *A relaxed constant positive linear dependence constraint qualification and applications*, Mathematical Programming, **135** (2012), 255–273.

- [3] R. Andreani, J. M. Martinez and M. L. Schuverdt, *On the relation between constant positive linear dependence condition and quasinormality constraint qualification*, Journal of Optimization Theory and Applications, **125** (2005), 473–483.
- [4] K. J. Arrow, L. Hurwicz and H. Uzawa, *Constraint qualifications in maximization problems*, Naval Research Logistics Quarterly, **8** (1961), 175–191.
- [5] M. S. Bazaraa, J. J. Goode and C. M. Shetty, *Constraint qualifications revisited*, Management Science, **18** (1972), 567–573.
- [6] M. S. Bazaraa, H. D. Sherali and C. M. Shetty, “Nonlinear Programming: Theory and Algorithms,” 3rd edition, Wiley-Interscience, Hoboken, NJ, 2006.
- [7] D. P. Bertsekas, “Nonlinear Programming,” 2nd edition, Athena Scientific, 1999.
- [8] D. P. Bertsekas, A. Nedić and A. E. Ozdaglar, “Convex Analysis and Optimization,” Athena Scientific, Belmont, MA, 2003.
- [9] D. P. Bertsekas and A. E. Ozdaglar, *Pseudonormality and a Lagrange multiplier theory for constrained optimization*, Journal of Optimization Theory and Applications, **114** (2002), 287–343.
- [10] J. M. Borwein and H. Wolkowicz, *A simple constraint qualification in infinite dimensional programming*, Mathematical Programming, **35** (1986), 83–96.
- [11] M. Canon, C. Cullum and E. Polak, *Constrained minimization problems in finite-dimensional spaces*, SIAM Journal on Control, **4** (1966), 528–547.
- [12] R. W. Cottle, *A theorem of Fritz John in mathematical programming*, RAND Memorandum RM-3858-PR, RAND Corporation, 1963.
- [13] W. Fenchel, “Convex Cones, Sets and Functions,” Princeton University Press, Princeton, 1953.
- [14] D. Gale, “The Theory of Linear Economic Models,” McGraw-Hill Book Company, Inc., New York-Toronto-London, 1960.
- [15] G. Giorgi and A. Guerraggio, *On the notion of tangent cone in mathematical programming*, Optimization, **25** (1992), 11–23.
- [16] F. J. Gould and J. W. Tolle, *A necessary and sufficient qualification for constrained optimization*, SIAM Journal on Applied Mathematics, **20** (1971), 164–172.
- [17] F. J. Gould and J. W. Tolle, *Geometry of optimality conditions and constraint qualifications*, Mathematical Programming, **2** (1972), 1–18.
- [18] M. Gowda and M. Teboulle, *A comparison of constraint qualifications in infinite-dimensional convex programming*, SIAM Journal on Control and Optimization, **28** (1990), 925–935.
- [19] M. Guignard, *Generalized Kuhn-Tucker conditions for mathematical programming problems in a Banach space*, SIAM Journal on Control, **7** (1969), 232–241.
- [20] M. R. Hestenes, “Calculus of Variations and Optimal Control Theory,” John Wiley & Sons, Inc., New York-London-Sydney, 1966.
- [21] M. R. Hestenes, “Optimization Theory: The Finite Dimensional Case,” Pure and Applied Mathematics, Wiley-Interscience [John Wiley & Sons], New York-London-Sydney, 1975.
- [22] L. Hurwicz, *Programming in linear spaces*, in “Studies in Linear and Non-linear Programming” (eds. K. J. Arrow, L. Hurwicz and H. Uzawa), Stanford University Press, (1958), 38–102.
- [23] R. Janin, *Directional derivative of the marginal function in nonlinear programming. Sensitivity, stability and parametric analysis*, Mathematical Programming Studies, **21** (1984), 110–126.
- [24] F. John, *Extremum problems with inequalities as subsidiary conditions*, in “Studies and Essays Presented to R. Courant on his 60th Birthday, January 8, 1948” (eds. K. O. Friedrichs, O. E. Neugebauer and J. J. Stoker), Wiley-Interscience New York, (1948), 187–204.
- [25] A. Jourani, *Constraint qualifications and Lagrange multipliers in nondifferentiable programming problems*, Journal of Optimization Theory and Applications, **81** (1994), 533–548.
- [26] S. Karlin, “Mathematical Methods and Theory in Games, Programming, and Economics,” Addison-Wesley Publishing Co., Inc., Reading, Mass.-London, 1959.
- [27] H. W. Kuhn and A. W. Tucker, *Nonlinear programming*, in “Second Berkeley Symposium on Mathematical Statistics and Probability,” University of California Press, Berkeley and Los Angeles, (1951), 481–492.
- [28] L. Kuntz, *Constraint qualifications in quasidifferentiable optimization*, Mathematical Programming, **60** (1993), 339–347.
- [29] L. Kuntz and S. Scholtes, *A nonsmooth variant of the Mangasarian-Fromovitz constraint qualification*, Journal of Optimization Theory and Applications, **82** (1994), 59–75.

- [30] S. Lu, *Implications of the constant rank constraint qualification*, Mathematical Programming, **126** (2011), 365–392.
- [31] D. Luenberger and Y. Ye, “Linear and Nonlinear Programming,” 3rd edition, International Series in Operations Research & Management Science, **116**, Springer, 2008.
- [32] C. Ma, X. Li, K. F. C. Yiu, Y. Yang and L. Zhang, *On an exact penalty function method for semi-infinite programming problems*, Journal of Industrial and Management Optimization, **8** (2012), 705–726.
- [33] O. L. Mangasarian, “Nonlinear Programming,” McGraw-Hill Book Company, Inc., New York-London-Sydney, 1969.
- [34] O. L. Mangasarian and S. Fromovitz, *The Fritz John necessary optimality conditions in the presence of equality and inequality constraints*, Journal of Mathematical Analysis and Applications, **17** (1967), 37–47.
- [35] R. R. Merkovsky and D. E. Ward, *General constraint qualifications in nondifferentiable programming*, Mathematical Programming, **47** (1990), 389–405.
- [36] L. Minchenko and S. Stakhovski, *On relaxed constant rank regularity condition in mathematical programming*, Optimization, **60** (2011), 429–440.
- [37] A. E. Ozdaglar and D. P. Bertsekas, *The relation between pseudonormality and quasiregularity in constrained optimization*, Optimization Methods and Software, **19** (2004), 493–506.
- [38] D. W. Peterson, *A review of constraint qualifications in finite-dimensional spaces*, SIAM Review, **15** (1973), 639–654.
- [39] L. Qi and Z. Wei, *On the constant positive linear dependence condition and its application to SQP methods*, SIAM Journal on Optimization, **10** (2000), 963–981.
- [40] K. Ritter, *Optimization theory in linear spaces*, Mathematische Annalen, **184** (1970), 133–154.
- [41] R. T. Rockafellar, “Conjugate Duality and Optimization,” Lectures given at the Johns Hopkins University, Baltimore, Md., June, 1973, Conference Board of the Mathematical Sciences Regional Conference Series in Applied Mathematics, No. 16, Society for Industrial and Applied Mathematics, Philadelphia, Pa., 1974.
- [42] R. T. Rockafellar, *Lagrange multipliers and optimality*, SIAM Review, **35** (1993), 183–238.
- [43] R. T. Rockafellar and R. J. B. Wets, “Variational Analysis,” Grundlehren der Mathematischen Wissenschaften [Fundamental Principles of Mathematical Sciences], **317**, Springer-Verlag, Berlin, 1998.
- [44] M. Slater, *Lagrange multipliers revisited*, Cowles Foundation Discussion Paper No. 80, Cowles Foundation for Research in Economics, Yale University, 1950.
- [45] M. V. Solodov, *Constraint qualifications*, in “Wiley Encyclopedia of Operations Research and Management Science,” John Wiley & Sons, Inc., 2010.
- [46] O. Stein, *On constraint qualifications in nonsmooth optimization*, Journal of Optimization Theory and Applications, **121** (2004), 647–671.
- [47] P. Varaiya, *Nonlinear programming in Banach space*, SIAM Journal on Applied Mathematics, **15** (1967), 284–293.
- [48] W. I. Zangwill, “Nonlinear Programming: A Unified Approach,” Prentice-Hall International Series in Management, Prentice-Hall, Inc., Englewood Cliffs, NJ, 1969.

Received February 2013; revised March 2013.

E-mail address: zwang23@ncsu.edu

E-mail address: fang@ncsu.edu

E-mail address: wxing@math.tsinghua.edu.cn