Universidade de São Paulo

Biblioteca Digital da Produção Intelectual - BDPI

Departamento de Ciência da Computação - IME/MAC | Artigos e Materiais de Revistas Científicas - IME/MAC

2012

TWO NEW WEAK CONSTRAINT QUALIFICATIONS AND APPLICATIONS

Two new weak constraint qualifications and applications *

Roberto Andreani[†]   Gabriel Haeser[‡]

María Laura Schuverdt[§]   Paulo J. S. Silva[¶]

February 13, 2012

Abstract

We present two new constraint qualifications (CQ) that are weaker than the recently introduced Relaxed Constant Positive Linear Dependence (RCPLD) constraint qualification. RCPLD is based on the assumption that many subsets of the gradients of the active constraints preserve positive linear dependence locally. A major open question was to identify the exact set of gradients whose properties had to be preserved locally and that would still work as a CQ. This is done in the first new constraint qualification, that we call Constant Rank of the Subspace Component (CRSC) CQ. This new CQ also preserves many of the good properties of RCPLD, like local stability and the validity of an error bound. We also introduce an even weaker CQ, called Constant Positive Generator (CPG), that can replace RCPLD in the analysis of the global convergence of algorithms. We close this work extending convergence results of algorithms belonging to all the main classes of nonlinear optimization methods: SQP, augmented Lagrangians, interior point algorithms, and inexact restoration.

[†]Department of Applied Mathematics, Institute of Mathematics, Statistics and Scientific Computing, University of Campinas, Campinas, SP, Brazil. Email: `andreani@ime.unicamp.br`. Phone: +55-19-3521 5960.

[‡]Institute of Science and Technology, Federal University of São Paulo, São José dos Campos, SP, Brazil. Email: `gabriel.haeser@unifesp.br`. Phone: +55-12-3309 9585.

[§]CONICET, Department of Mathematics, FCE, University of La Plata, CP 172, 1900 La Plata Bs. As., Argentina. Email: `schuverd@mate.unlp.edu.ar`. Phone: +54-221-422 9850.

[¶]Institute of Mathematics and Statistics, University of São Paulo, São Paulo, SP, Brazil. Email: `pjssilva@ime.usp.br`. Phone: +55-11-3091 5178.

1

1 Introduction

Let us consider a nonlinear optimization problem in the form

$$
\begin{aligned}
\min \quad & f_0(x) \\
\text{s.t.} \quad & f_i(x) = 0, \ i = 1, \ldots, m, \\
& f_j(x) \le 0, \ j = m+1, \ldots, m+p,
\end{aligned}
\tag{NOP}
$$

where the functions $f_i : \mathbb{R}^n \to \mathbb{R}$, $i = 0, \ldots, m+p$ are continuously differentiable. We denote its feasible set by F. The constraints that hold as equalities in a point x are said *active* at x. If x is a feasible point, the active constraints contain all the equality constraints together with a, possibly empty, subset of inequalities. We will denote by $\mathcal{A}(x)$ the index set of the active inequality constraints $\mathcal{A}(x) \stackrel{def}{=} \{i \mid f_i(x) = 0, \ i = m+1, \ldots, m+p\}$.

One of the main subjects in the theory of nonlinear optimization is the characterization of optimality, which is often achieved through conditions that use the derivatives of the constraints at a prospective optimum. Among such conditions arguably the most important is the Karush-Kuhn-Tucker (KKT) condition which is extensively used in the development of algorithms to solve (NOP) [8, 35].

In order to ensure that the KKT conditions are necessary for optimality a constraint qualification (CQ) is needed. Constraint qualifications are properties of the algebraic description of the feasible set that allow its local geometry at a feasible point x to be recovered from the gradients of the active constraints at x. In order to make this sentence clear we need to recall some definitions.

Definition 1.1. *Let x be a feasible point of* (NOP)*, that is, $x \in F$. The* tangent cone *of F at x is defined as*

$$
\mathcal{T}(x) \stackrel{def}{=} \left\{ y \in \mathbb{R}^n \ \middle| \ \begin{array}{c} \exists x^k \in F, \ x^k \to x \\ \frac{x^k - x}{\|x^k - x\|} \to \frac{y}{\|y\|} \end{array} \right\} \cup \{0\}.
$$

This cone is composed by the limit of directions that move inward the feasible set. It is inherently a geometric object, as it captures the local "shape" of the set around x. Using it, we can easily present a geometric necessary optimality condition for local optimality at x

$$
-\nabla f_0(x) \in \mathcal{T}(x)^\circ,
\tag{1}
$$

where $\mathcal{T}(x)^\circ$ is the polar of $\mathcal{T}(x)$ [8].

However, the tangent cone is not an algebraic object and hence it cannot be directly used in algorithms. Constraint qualifications are conditions that ensure that $\mathcal{T}(x)^\circ$ can be recast using the algebraic information of the gradients. More specifically, we may try to approximate the tangent cone using the *linearized cone of F at x* which uses only information of the gradients and is given by

$$
\mathcal{F}(x) \stackrel{def}{=} \{y \mid \nabla f_i(x)'y = 0, \ i \in 1, \ldots, m, \ \nabla f_j(x)'y \le 0, \ j \in \mathcal{A}(x)\},
\tag{2}
$$

Note that this cone always contains the tangent cone, that is $\mathcal{T}(x) \subset \mathcal{F}(x)$.

The polar of $\mathcal{F}(x)$ can be easily computed and is given by

$$\mathcal{F}(x)^\circ = \left\{ y \mid y = \sum_{i=1}^{m} \lambda_i \nabla f_i(x) + \sum_{j \in \mathcal{A}(x)} \mu_j \nabla f_j(x), \ \mu_j \geq 0 \right\}. \qquad (3)$$

If $\mathcal{F}(x)^\circ = \mathcal{T}(x)^\circ$ the optimality condition (1) can be rewritten as

$$-\nabla f_0(x) \in \mathcal{F}(x)^\circ,$$

which is exactly the KKT condition. The condition $\mathcal{F}(x)^\circ = \mathcal{T}(x)^\circ$, was introduced by Guignard [15] and the discussion above suggests that it is the most general constraint qualification possible. In fact, Gould and Tolle proved in [13] that it is equivalent to the necessity of the KKT condition for all possible objective functions.

Another possibility is to require directly that $\mathcal{F}(x) = \mathcal{T}(x)$. Even though this condition is more stringent than Guignard's CQ, it is in some cases easier to work with since it does not involve the polar operation. Such constraint qualification was introduced by Abadie in [1] and it is widely used in the optimization theory [41, 7, 9, 26].

Clearly both Guignard's and Abadie's CQ are enforcing the equality between geometric objects that capture the local structure of the feasible set around x, namely $\mathcal{T}(x)$ and its polar, with objects that use gradient information at the point x. The gradients have local information of the respective constraint functions, but they cannot always express the interrelationship among all functions while defining the feasible set. In this sense, we can say that a CQ is a condition that tries to restrict how the gradients, and hence the constraints themselves, vary *together* in a neighborhood of x. Such variation should be well behaved enough to assert Guignard's condition.

The most simple CQ, called *linear independence constraint qualification* (LICQ), asks for linear independence of the gradients of the active constraints at the point of interest x. This condition is still important today and is required in many special cases, specially when connected to convergence results for numerical algorithms [8, 35]. When the problem has inequality constraints it is usually better to consider the Mangasarian-Fromovitz's CQ (MFCQ) that asks that the gradients of the active constraints must be positively linearly independent[1], relaxing LICQ [27, 40]. Even though these two conditions appear pointwise conditions, they actually constraint how the gradients may vary together in a neighborhood of x, as linear independence and positive linear independence are conditions that are preserved locally.

The LICQ was relaxed by Janin in [22] while studying the directional derivative of the marginal function associated to the right-hand side of (NOP). In particular, Janin showed that if the ranks of all subsets of the gradients of the

---

[1]For a precise definition of what is positive linear independence in this context, see Section 2.

active constraints remain constant in a neighborhood of x, then the KKT conditions are necessary for optimality. This condition is known as the constant rank constraint qualification (CRCQ). Clearly, LICQ is a particular case of CRCQ.

The CRCQ was further relaxed by Qi and Wei [38], in the context of studying sequential quadratic programming algorithms. The authors introduced the constant positive linear dependence (CPLD) condition, that was shown to be a constraint qualification by Andreani, Martínez and Schuverdt [6]. In [2, 3] Andreani et al. showed that this constraint qualification is enough to ensure the convergence of an augmented Lagrangian method to a KKT point. The CPLD condition asks that the positive linear dependence of any subset of the active gradients must be preserved locally.

More recently, Minchenko and Stakhovski showed that the constant rank constraint qualification can be relaxed to consider only the full set of the equality constraints [32]. More precisely they showed that the following condition is a constraint qualification.

Definition 1.2. *We say that the* relaxed constant rank *constraint qualification (RCRCQ) holds at a feasible point x if there is a neighborhood $N(x)$ of x where for all subsets $\mathcal{J} \subset \mathcal{A}(x)$ and all $y \in N(x)$ the set of gradients $\{\nabla f_i(y) \mid i \in \{1, \ldots, m\} \cup \mathcal{J}\}$ has constant rank.*

Interesting relations between this condition and the original constant rank condition were unveiled in [24]. The relaxed constant rank condition was further extended to take into account positive linear independence in the place of the rank in [5], where a relaxed version of the CPLD, called RCPLD, is introduced. This work also shows that RCPLD is enough to ensure the validity of an error bound and the global convergence of an augmented Lagrangian method.

These last developments are interesting as they do not take into account all the subsets of the gradients of the equality constraints. Only the full set of gradients $\{\nabla f_1(x), \ldots, \nabla f_m(x)\}$ is important. So, if the problem only has equality constraints, these conditions basically require that the linearized cone of F must have constant dimension locally, only tilting to support the feasible set at each point. This is a strong geometric condition that is easy to understand and visualize.

However, if the problem has inequalities, the results described above still require local conditions on all subsets of the gradients of the active inequalities. The simplicity of considering only one set of gradients whose properties must be stable is lost. The main purpose of this paper is to fill this gap, showing that only a single subset of the inequality constraints needs to be considered.

When the feasible set is described with inequalities, the rank preservation of the gradients is not the right concept to describe its structure. For example, consider the constraints $y \geq 0$, $y - x^2 \geq 0$. They conform to MFCQ at 0 but their rank increases locally. The rank is a tool that is better suited to deal with the gradients of the equality constraints as they generate a subspace contained in $\mathcal{F}(x)^\circ$ where the notion of dimension can be applied.

For inequality constraints the idea of constant positive linear dependence looks like the best choice. On the other hand, in some cases, inequality con-
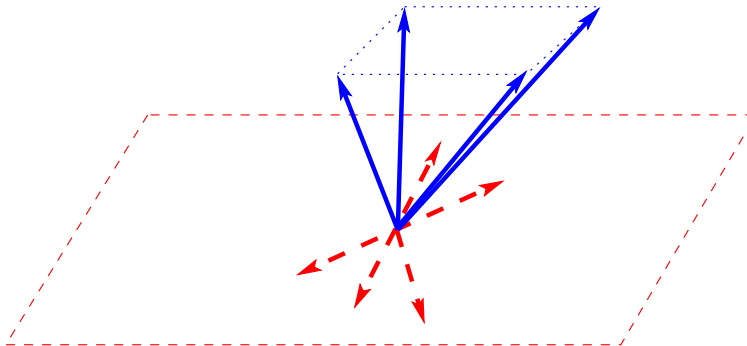
Figure 1: Linear space and pointed cone components of $\mathcal{F}(x)^\circ$. The subspace is generated by the gradients of the equality constraints together with the gradients of constraints with indexes in $\mathcal{J}_-$. The pointed cone is generated by the gradients of the active inequality constraints that are not in $\mathcal{J}_-$.

straints may behave like, or even be, equality constraints in disguise. For example, $x \geq 0$ and $x \leq 0$ that together mean $x = 0$. In this case, rank preservation is the right concept.

How to reconcile these two possibilities? One way is to try to identify which inequalities actually behave like equalities in the description of the polar of the linearized cone. With this objective in mind, let us consider the maximal subspace contained in $\mathcal{F}(x)^\circ$ that we call its *subspace component*. The description given in (3) seems to suggest that this subspace is generated by the gradients of the equalities. The other term in the sum, associated to the gradients of the inequalities, is expected to be a pointed cone. Most of the problems arise when this division is not clear, that is, when gradients of inequality constraints fall into the subspace component of the polar of the linearized cone. See Figure 1. Formally this happens whenever the set

$$\mathcal{J}_- \stackrel{def}{=} \{j \in \mathcal{A}(x) \mid -\nabla f_j(x) \in \mathcal{F}(x)^\circ\} \tag{4}$$

is nonempty. This index set appears implicitly in the Mangasarian-Fromovitz's CQ which is equivalent to require that $\mathcal{J}_-$ must be empty, while the gradients of the equality constraints, that generate the linear space component of the polar of the linearized cone, must be linearly independent thus preserving its dimension locally.

In order to generalize the constraint qualifications described above we need to generalize the notion of a basis of a subspace to deal with cones spanned by linear combinations using signed coefficients. We then require that such special spanning sets must be preserved locally. The precise definition of this new CQ is given in Section 3. In particular, we show that many of the constraint qualifications discussed above imply that the subspace component of the polar of the linearized cone has the same dimension locally, which in turn implies the new CQ.

The preservation of the dimension of the subspace component is an intermediate CQ that plays a fundamental role in the applications. Let us formalize it below.

Definition 1.3. *Let x be a feasible point of* (NOP) *and define the index set $\mathcal{J}_-$ as in* (4). *We say that the* constant rank of the subspace component *condition (CRSC) holds at x if there is a neighborhood $N(x)$ of x such that the rank of $\{\nabla f_l(y) \mid l \in \{1, \ldots, m\} \cup \mathcal{J}_-\}$ remains constant for $y \in N(x)$.*

Note that the fact that constant positive linear dependence CQs, in particular RCPLD, imply CRSC as proved in Theorem 4.2, is somewhat surprising. In particular this fact reconciles constant rank and constant positive linear dependence constraint qualifications: both are actually ensuring that the subspace spanned by the gradients of the equality constraints and the gradients of the inequality constraints with indexes in $\mathcal{J}_-$ has constant dimension locally. The fact that the dimension of the linear space component is locally constant has deep geometrical consequences: *it basically says that the polar of the linearized cone has the same shape locally, it can only tilt preserving its structure.* Moreover, this condition is clearly more general than RCPLD, as the simple feasible set $\{x \mid x \leq 0, -x \leq 0, x^2 \leq 0\}$ conforms to CRSC at its only point, the origin, while RCPLD fails.

The rest of this paper is organized as follows. Section 2 introduces the notion of positively linearly independent spanning pairs that replaces the idea of a basis for cones. Section 3 uses this idea to introduce a new constraint qualification that we call constant positive generator condition (CPG) and that generalizes CRSC and many of the CQs described above. Section 4 shows the relation between RCPLD, CRSC, and CPG. It shows that CPG implies Abadie's constraint qualification. Finally, Section 5 shows some important applications of CRSC and CPG. It discusses when an error bound holds and also show that many algorithms converge under the weak CPG condition.

2 Positively Linearly Independent Spanning Pairs

One of the main objects in the study of constraint qualification is $\mathcal{F}(x)^\circ$, the polar of the linearized cone of the feasible set at a feasible point x, see (3). This cone is spanned by the gradients of the active constraints at x with some sign conditions on the combination coefficients. This notion of spanning cones using vectors and coefficients with sign conditions is fundamental in our development. Let us formalize it in the next definition.

Definition 2.1. *Let $V = (v_1, v_2, \ldots, v_K)$ be a tuple[2] of vectors in $\mathbb{R}^n$ and $\mathcal{I}, \mathcal{J} \subset \{1, 2, \ldots, K\}$ a pair of index sets. We call a* positive combination *of*

---

[2]We use a tuple instead of a regular set to allow for vectors to appear more than once. It is natural to consider this possibility in our discussion as the gradients of different constraints may be equal in a given point.

elements of V associated to the (ordered) pair $(\mathcal{I}, \mathcal{J})$ a vector in the form

$$\sum_{i \in \mathcal{I}} \lambda_i v_i + \sum_{j \in \mathcal{J}} \mu_j v_j, \ \mu_j \geq 0, \forall j \in \mathcal{J}.$$

The set of all such positive combinations is called the positive span of V *associated to $(\mathcal{I}, \mathcal{J})$ and it is denoted by $\mathrm{span}_+(\mathcal{I}, \mathcal{J}; V)$. It is clearly a cone. If the tuple V is clear from the context we may omit it and use* positive combinations *of $(\mathcal{I}, \mathcal{J})$,* positive span of $(\mathcal{I}, \mathcal{J})$, *and write $\mathrm{span}_+(\mathcal{I}, \mathcal{J})$. On the other hand, if the set $\mathcal{I} = \emptyset$, that is if all coefficients are supposed to be non-negative, we may talk about* positive combinations of V *and* positive span of V.

The vectors $v_\ell, \ell \in \mathcal{I} \cup \mathcal{J}$, or the pair $(\mathcal{I}, \mathcal{J})$ when V is clear from the context, are said to be positively linearly independent *if the only way to write the zero vector using positive combinations is to use trivial coefficients. Otherwise we say that the vectors, or the pair, are* positively linearly dependent.

Let $\mathcal{I}', \mathcal{J}' \subset \{1, 2, \ldots, K\}$ be another pair of indexes, we say that $(\mathcal{I}', \mathcal{J}')$ spans positively $\mathrm{span}_+(\mathcal{I}, \mathcal{J}; V)$ *if $\mathrm{span}_+(\mathcal{I}', \mathcal{J}'; V) = \mathrm{span}_+(\mathcal{I}, \mathcal{J}; V)$. We may also say that $(\mathcal{I}', \mathcal{J}')$ is a* positive spanning pair *for such cone.*

Now, let us recall the definition of the polar of the linearized cone $\mathcal{F}(x)^\circ$ given in (3). If we set $\mathcal{I}$ as the indexes of the equality constraints $\{1, 2, \ldots, m\}$, $\mathcal{J}$ as the indexes of the inequality constraints that are active at x, that is $\mathcal{A}(x)$, and V as the tuple of gradients with indexes in $\mathcal{I} \cup \mathcal{J}$, then $\mathcal{F}(x)^\circ$ is the positive span of V associated to the pair $(\mathcal{I}, \mathcal{J})$.

Next, let us try to generalize the idea of a basis from linear spaces to positive spanned cones in the form $\mathrm{span}_+(\mathcal{I}, \mathcal{J}; V)$. In other words, we want to define a "minimal" spanning pair for such a cone. A first attempt is to look for a positively linearly independent spanning pair for it, however, the usual technique to find such pair may not apply. For example, for $V = \{v_1 = -1, v_2 = 1\} \subset \mathbb{R}$, $\mathcal{I} = \emptyset$ and $\mathcal{J} = \{1, 2\}$ it is not possible to obtain such a pair simply removing vectors from $\mathcal{I}$ and $\mathcal{J}$, as it is possible in the linear case. In order to find such spanning pair we need to *remove vectors from $\mathcal{J}$ and put them in $\mathcal{I}$*. In fact, $\mathcal{I}' = \{1\}$ and $\mathcal{J}' = \emptyset$ form a positively linearly independent spanning pair for the same cone. We make this procedure clear in the next result.

Theorem 2.1. *Let $V = (v_1, v_2, \ldots, v_K)$ be a tuple of vectors in $\mathbb{R}^n$ and $\mathcal{I}, \mathcal{J} \subset \{1, 2, \ldots, K\}$ such that the pair $(\mathcal{I}, \mathcal{J})$ is positively linearly dependent. Then the pair $(\mathcal{I}', \mathcal{J}')$ defined below spans positively $\mathrm{span}_+(\mathcal{I}, \mathcal{J}; V)$.*

1. *If $\mathcal{I}$ is associated to linearly dependent vectors, define $\mathcal{I}'$ as a proper subset of $\mathcal{I}$ such that $\mathrm{span}\{v_i \mid i \in \mathcal{I}'\} = \mathrm{span}\{v_i \mid i \in \mathcal{I}\}$ and set $\mathcal{J}' = \mathcal{J}$.*

2. *Otherwise, $\mathcal{I}$ is associated to linearly independent vectors and there is a $j' \in \mathcal{J}$ such that $-v_j \in \mathrm{span}_+(\mathcal{I}, \mathcal{J})$. Define $\mathcal{I}' = \mathcal{I} \cup \{j'\}$ and $\mathcal{J}' = \mathcal{J} \setminus \{j'\}$, a proper subset of $\mathcal{J}$.*

Proof. In the first case it is trivial to see that the cones coincide.

In the case 2, as $(\mathcal{I}, \mathcal{J})$ is positively linearly dependent, there must be coefficients $\bar{\lambda}_i$, for $i \in \mathcal{I}$, and non-negative $\bar{\mu}_j$, for $j \in \mathcal{J}$, such that

$$\sum_{i \in \mathcal{I}} \bar{\lambda}_i v_i + \sum_{j \in \mathcal{J}} \bar{\mu}_j v_j = 0. \tag{5}$$

Note that not all $\bar{\mu}_j$, $j \in \mathcal{J}$ are zero, otherwise v_i, $i \in \mathcal{I}$, would not be linearly independent. Then, there is at least one $j' \in \mathcal{J}$ such that $\bar{\mu}_{j'} > 0$. Dividing the equation above by $\bar{\mu}_{j'}$ we get

$$\sum_{i \in \mathcal{I}} \frac{\bar{\lambda}_i}{\bar{\mu}_{j'}} v_i + \sum_{j \in \mathcal{J} \setminus \{j'\}} \frac{\bar{\mu}_j}{\bar{\mu}_{j'}} v_j = -v_{j'}.$$

Now define the index sets $\mathcal{I}' \stackrel{def}{=} \mathcal{I} \cup \{j'\}$ and $\mathcal{J}' \stackrel{def}{=} \mathcal{J} \setminus \{j'\}$. Clearly, $\mathrm{span}_+(\mathcal{I}', \mathcal{J}') \supset \mathrm{span}_+(\mathcal{I}, \mathcal{J})$. On the other hand, let

$$\sum_{i \in \mathcal{I}'} \lambda_i v_i + \sum_{j \in \mathcal{J}'} \mu_j v_j$$

be an element of $\mathrm{span}_+(\mathcal{I}', \mathcal{J}')$. Then, it clearly belongs to $\mathrm{span}_+(\mathcal{I}, \mathcal{J})$ if the coefficient of $v_{j'}$ is non-negative. Otherwise,

$$\sum_{i \in \mathcal{I}'} \lambda_i v_i + \sum_{j \in \mathcal{J}'} \mu_j v_j = \sum_{i \in \mathcal{I}} \lambda_i v_i + \lambda_{j'} v_{j'} + \sum_{j \in \mathcal{J}'} \mu_j v_j$$

$$= \sum_{i \in \mathcal{I}} \lambda_i v_i + |\lambda_{j'}| \left(\sum_{i \in \mathcal{I}} \frac{\bar{\lambda}_i}{\bar{\mu}_{j'}} v_i + \sum_{j \in \mathcal{J} \setminus \{j'\}} \frac{\bar{\mu}_j}{\bar{\mu}_{j'}} v_j \right)$$

$$+ \sum_{j \in \mathcal{J}'} \mu_j v_j,$$

and we see that it is actually in $\mathrm{span}_+(\mathcal{I}, \mathcal{J})$. $\square$

We can then, easily construct positively linearly independent spanning pairs.

Corollary 2.2. *Let $V = (v_1, v_2, \ldots, v_K)$ be a tuple of vectors in $\mathbb{R}^n$ and $\mathcal{I}, \mathcal{J} \subset \{1, 2, \ldots, K\}$ a pair of index sets. Then, there exist $\mathcal{I}', \mathcal{J}' \subset \{1, 2, \ldots, K\}$ such that $(\mathcal{I}', \mathcal{J}')$ is positively linearly independent and $\mathrm{span}_+(\mathcal{I}', \mathcal{J}'; V) = \mathrm{span}(\mathcal{I}, \mathcal{J}; V)$. We call such pairs* positively linearly independent spanning pairs *of $\mathrm{span}_+(\mathcal{I}, \mathcal{J}; V)$.*

Proof. Start with $(\mathcal{I}, \mathcal{J})$ and apply the construction given in Theorem 2.1 while possible. Clearly this can only be done a finite number of times and the resulting pair $(\mathcal{I}', \mathcal{J}')$ is positively linearly independent. $\square$

The case 2 in Theorem 2.1 simply states that if both v_j and $-v_j$ belong to $\mathrm{span}_+(\mathcal{I}, \mathcal{J})$ for some index $j \in \mathcal{J}$, then this index may have been misplaced and should be moved to $\mathcal{I}$. If we recall the natural definitions $\mathcal{I}$, $\mathcal{J}$, and V

when considering $\mathcal{F}(x)^\circ$, moving an index from $\mathcal{J}$ to $\mathcal{I}$ is associated to stating that an inequality constraint should be viewed as an equality, something which is not usual in optimization.

To see why this is acceptable, let us recall that $\mathcal{F}(x)^\circ$ is the polar to the linearized cone. The fact that an inequality constraint f_j has both $\nabla f_j(x)$ and $-\nabla f_j(x)$ in $\mathcal{F}(x)^\circ$ implies that $\mathcal{F}(x)$, and hence $\mathcal{T}(x)$, lies in the subspace orthogonal to $\nabla f_j(x)$. That is, if we consider the feasible set F, f_j is interacting with the other constraints that define it and acting more closely to an equality constraint than to an inequality constraint.

We end this section with an alternative characterization of the positively linearly independent spanning pairs given above. We start with a definition, already suggested in the introduction.

Definition 2.2. *Let $V = (v_1, v_2, \ldots, v_K)$ be a tuple of vectors in $\mathbb{R}^n$ and $\mathcal{I}, \mathcal{J} \subset \{1, 2, \ldots, K\}$ a pair of index sets. Define*

$$\mathcal{J}_- \stackrel{def}{=} \{j \in \mathcal{J} \mid -v_j \in \text{span}_+(\mathcal{I}, \mathcal{J}; V)\} \quad and \quad \mathcal{J}_+ \stackrel{def}{=} \mathcal{J} \setminus \mathcal{J}_-.$$

Lemma 2.3. *Let $V = (v_1, v_2, \ldots, v_K)$ be a tuple of vectors in $\mathbb{R}^n$ and let $\mathcal{I}, \mathcal{J} \subset \{1, 2, \ldots, K\}$ be a pair of index sets. If $(\mathcal{I}', \mathcal{J}')$ is a positively linearly independent spanning pair for $\text{span}_+(\mathcal{I}, \mathcal{J}; V)$. Then,*

1. *$\mathcal{J}' \subset \mathcal{J}_+$.*

2. *$(\mathcal{I}', \mathcal{J}_+)$ is also a positively linearly independent spanning pair for $\text{span}_+(\mathcal{I}, \mathcal{J}; V)$.*

3. *$\mathcal{I}' \subset \mathcal{I} \cup \mathcal{J}_-$ and it is composed of indexes of a basis of the subspace spanned by $\{v_\ell \mid \ell \in \mathcal{I} \cup \mathcal{J}_-\}$.*

Proof. 1. Let $\ell \in \mathcal{J}'$. Suppose, by contradiction, that $\ell \notin \mathcal{J}_+$, in other words $-v_\ell \in \text{span}_+(\mathcal{I}, \mathcal{J}; V) = \text{span}_+(\mathcal{I}', \mathcal{J}'; V)$. In this case,

$$-v_\ell = \sum_{i \in \mathcal{I}'} \lambda_i v_i + \sum_{j \in \mathcal{J}' \setminus \{\ell\}} \mu_j v_j + \mu_\ell v_\ell, \ \mu_j \geq 0, \forall j \in \mathcal{J}'.$$

Which implies

$$0 = \sum_{i \in \mathcal{I}'} \lambda_i v_i + \sum_{j \in \mathcal{J}' \setminus \{\ell\}} \mu_j v_j + (\mu_\ell + 1) v_\ell, \ \mu_j \geq 0, \forall j \in \mathcal{J}'.$$

As $(\mu_\ell + 1) > 0$ this is a contradiction to the assumption that $(\mathcal{I}', \mathcal{J}')$ is positively linearly independent.

2. First, observe that as $\mathcal{J}' \subset \mathcal{J}_+$, then $\text{span}_+(\mathcal{I}, \mathcal{J}; V) = \text{span}_+(\mathcal{I}', \mathcal{J}'; V) \subset \text{span}_+(\mathcal{I}, \mathcal{J}_+; V) \subset \text{span}_+(\mathcal{I}, \mathcal{J}; V)$. Hence, $(\mathcal{I}', \mathcal{J}_+)$ is also a spanning pair. Now, suppose by contradiction that it is positively linearly dependent, that is, there are coefficients λ_i, for $i \in \mathcal{I}'$ and $\mu_j \geq 0$, for $j \in \mathcal{J}_+$, not all zero, such that

$$\sum_{i \in \mathcal{I}'} \lambda_i v_i + \sum_{j \in \mathcal{J}_+} \mu_j v_j = 0.$$

Since $(\mathcal{I}', \mathcal{J}')$ is positively linearly independent, the vectors with indexes in $\mathcal{I}'$ are linearly independent. Hence, at least one of the coefficients $\mu_{j'}, j' \in \mathcal{J}_+$ is strictly positive. We can then rearrange the above equality to solve for $-v_{j'}$ and get a contradiction to the definition of $\mathcal{J}_+$.

3. If $j \in \mathcal{I}'$, $-v_j \in \text{span}_+(\mathcal{I}, \mathcal{J}; V)$. Hence, j must belong to either $\mathcal{I}$ or $\mathcal{J}_-$, by definition of such index sets. Now, clearly, the vectors with indexes in $\mathcal{I}'$ are linearly independent, as $(\mathcal{I}', \mathcal{J}')$ is positively linearly independent. We only need to show that any v_ℓ, $\ell \in \mathcal{I} \cup \mathcal{J}_-$ is a linear combination of the vectors with indexes in $\mathcal{I}'$. Now, as both $v_\ell, -v_\ell \in \text{span}_+(\mathcal{I}', \mathcal{J}')$, there must be coefficients λ_i^+, λ_i^-, $i \in \mathcal{I}'$, and non-negative μ_j^+, μ_j^-, $j \in \mathcal{J}'$, such that

$$v_\ell = \sum_{i \in \mathcal{I}'} \lambda_i^+ v_i + \sum_{j \in \mathcal{J}'} \mu_j^+ v_j, \tag{6}$$
$$-v_\ell = \sum_{i \in \mathcal{I}'} \lambda_i^- v_i + \sum_{j \in \mathcal{J}'} \mu_j^- v_j.$$

Summing up these two inequalities we get,

$$0 = \sum_{i \in \mathcal{I}'} (\lambda_i^+ + \lambda_i^-) v_i + \sum_{j \in \mathcal{J}'} (\mu_j^+ + \mu_j^-) v_j.$$

As $(\mathcal{I}', \mathcal{J}')$ is positively linearly independent, we know that all coefficients in the summation above are zero. Since $\forall j \in \mathcal{J}', \mu_j^+, \mu_j^- \geq 0$ we conclude that $\forall j \in \mathcal{J}', \mu_j^+ = \mu_j^- = 0$. It follows from (6) that v_ℓ is spanned by the vectors in $\mathcal{I}'$.

$\square$

Corollary 2.4. *The positively linearly independent spanning pairs given by Corollary 2.2 have the form*

$$\mathcal{I}' \subset \mathcal{I} \cup \mathcal{J}_-, \quad \mathcal{J}' = \mathcal{J}_+,$$

where $\mathcal{I}'$ is composed by indexes of a basis of the space spanned by $\{v_\ell \mid \ell \in \mathcal{I} \cup \mathcal{J}_-\}$.

Proof. This is an immediate consequence of the lemma above and the fact that the procedure described in Corollary 2.2 never moves vectors from $\mathcal{J}_+$ to the set $\mathcal{I}'$. $\square$

Corollary 2.5. *The set $\text{span}_+(\mathcal{I}, \mathcal{J}_-)$ is a subspace.*

Proof. By definition, $v_j \in \mathcal{J}_-$ if, and only if, $-v_j$ is a positive linear combination of the other vectors in $\mathcal{I} \cup \mathcal{J}$. But in this positive combination the vectors in $\mathcal{J}_+$ can only appear with zero coefficients, otherwise they would belong to $\mathcal{J}_-$. $\square$

3 Constant positive generators

Now we are ready to introduce a new constraint qualification:

Definition 3.1. *Consider the nonlinear optimization problem* (NOP). *For* $y \in \mathbb{R}^n$ *define the tuple* $Gf(y) \overset{def}{=} (\nabla f_1(y), \nabla f_2(y), \ldots, \nabla f_{m+p}(y))$. *Let* x *be a feasible point and define the index sets* $\mathcal{I} \overset{def}{=} \{1, 2, \ldots, m\}$ *and* $\mathcal{J} \overset{def}{=} \mathcal{A}(x)$, *the set of active inequality constraints.*

We say that the constant positive generator *(CPG) condition holds at* x *if there is a positively linearly independent spanning pair* $(\mathcal{I}', \mathcal{J}_+)$ *of* $\mathrm{span}_+(\mathcal{I}, \mathcal{J}; Gf(x))$ *such that*

$$\mathrm{span}_+(\mathcal{I}', \mathcal{J}_+; Gf(y)) \supset \mathrm{span}_+(\mathcal{I}, \mathcal{J}; Gf(y)), \tag{7}$$

for all y *in a neighborhood of* x.

Note that we used implicitly Lemma 2.3 in this definition. Actually, if $(\mathcal{I}', \mathcal{J}')$ is a positively linearly independent spanning pair for $\mathrm{span}_+(\mathcal{I}, \mathcal{J}; Gf(x))$, the lemma says that $(\mathcal{I}', \mathcal{J}_+)$ is also a spanning pair. As $\mathcal{J}_+ \supset \mathcal{J}'$, it may be easier to show that the inclusion (7) holds using $\mathcal{J}_+$ in the place of a smaller $\mathcal{J}'$. Hence, we decided to state the definition already using the larger index set $\mathcal{J}_+$.

Another remark is that one may think that if the inclusion required in CPG holds, it must hold as an equality. This is not always true. For example consider the following feasible set

$$F = \{(x_1, x_2) \in \mathbb{R}^2 \mid x_1^3 - x_2 \le 0, \; x_1^3 + x_2 \le 0, \; x_1 \le 0\}$$

at the origin. In this point CPG holds with the inclusion holding in the proper sense. See Figure 2.

Finally, an extension of this example can also be used to show that it is possible to have the inclusion (7) holding only for a specific choice for $\mathcal{I}'$. In order to see this, let us add a constraint to the feasible set above and consider

$$F = \{(x_1, x_2) \in \mathbb{R}^2 \mid x_1^3 - x_2 \le 0, \; x_1^3 + x_2 \le 0, \; x_1 \le 0, x_2^3 \le 0\}$$

at the origin. Here, the constraints associated to $\mathcal{J}_-$ are the first, second and fourth, that is $\mathcal{J}_- = \{1, 2, 4\}$, while $\mathcal{J}_+ = \{3\}$ and $\mathcal{I} = \emptyset$. There are two possible choices for $\mathcal{I}'$ that are associated to positively linearly independent spanning pairs at the origin. Either $\mathcal{I}' = \{1\}$ which shows that CPG holds, or $\mathcal{I}' = \{2\}$, and the inclusion in the CPG definition is not valid. See Figure 3.

Now, we move to prove that CPG is actually a constraint qualification. First let us recall the definition of approximate-KKT points [4].

Definition 3.2. *We say that a feasible point* x *of* (NOP) *conforms to the Approximate-KKT condition (AKKT condition) if there exist sequences* $x^k \to x$, $\epsilon^k \to 0$, *and* $\{\lambda^k\} \subset \mathbb{R}^m$, $\{\mu^k\} \subset \mathbb{R}^p$, $\mu^k \ge 0$ *such that*

$$\nabla f_0(x^k) + \sum_{i \in \mathcal{I}} \lambda_i^k \nabla f_i(x^k) + \sum_{j \in \mathcal{A}(x)} \mu_j^k \nabla f_{m+j}(x^k) = \epsilon^k.$$

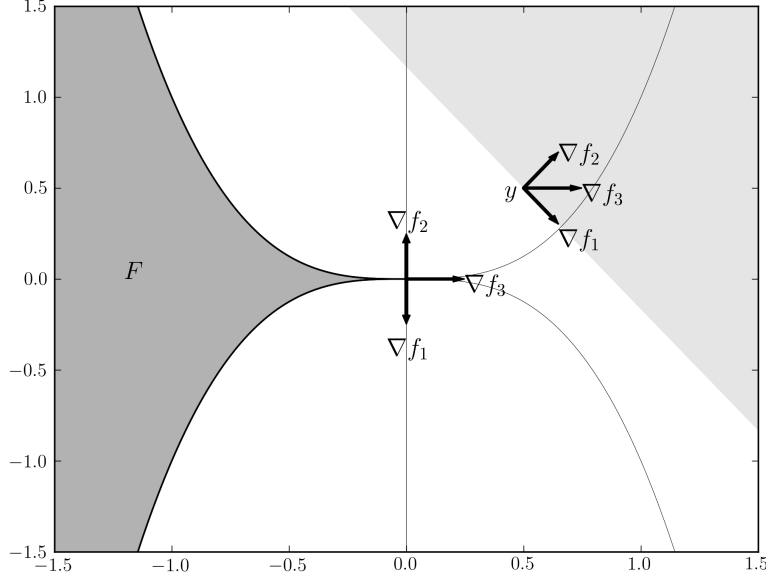*In this case, we may also say that* x *is an AKKT point.*

Figure 2: Consider $F = \{(x_1, x_2) \in \mathbb{R}^2 \mid f_1(x_1, x_2) = x_1^3 - x_2 \leq 0, \ f_2(x_1, x_2) = x_1^3 + x_2 \leq 0, \ f_3(x_1, x_2) = x_1 \leq 0\}$ at the origin. Then, we can take $\mathcal{I}' = \{1\}$ and $\mathcal{J}_+ = \{3\}$ in the definition of CPG. Then, for all $y \neq 0$, $\mathrm{span}_+(\{1\}, \{3\}; Gf(y))$ is a semispace, pictured in light gray above, that properly contains the pointed cone $\mathrm{span}_+(\emptyset, \{1, 2, 3\}; Gf(y))$, positively generated by the gradients.

It is well-known from [4] that if x is a local minimum then it must be an AKKT point. Therefore, to prove that CPG is a constraint qualification all we need to show is that if CPG holds at an approximate-KKT point then it has to be a KKT point. Another important property is that many methods for nonlinear optimization are guaranteed to converge to AKKT points. Hence, it will be a corollary of Theorem 3.1 below that if one of such algorithms generates a sequence converging to a point where CPG holds, such point has to be a KKT point. This will be the main tool used in Section 5.2 where we describe applications of CPG to the convergence analysis of nonlinear optimization methods.

Theorem 3.1. *Let x be a feasible point of* (NOP) *that satisfies the AKKT condition. If x also satisfies CPG, then x is a KKT point.*

Proof. Let x^k, ϵ^k, λ^k, and μ^k be the sequences given by the AKKT condition. Let $(\mathcal{I}', \mathcal{J}_+)$ be the positively linearly independent spanning pair given by CPG. Then, for each sufficiently large k there must be $\bar{\lambda}_i^k, i \in \mathcal{I}'$ and $\bar{\mu}_j^k \geq 0, j \in \mathcal{J}_+$
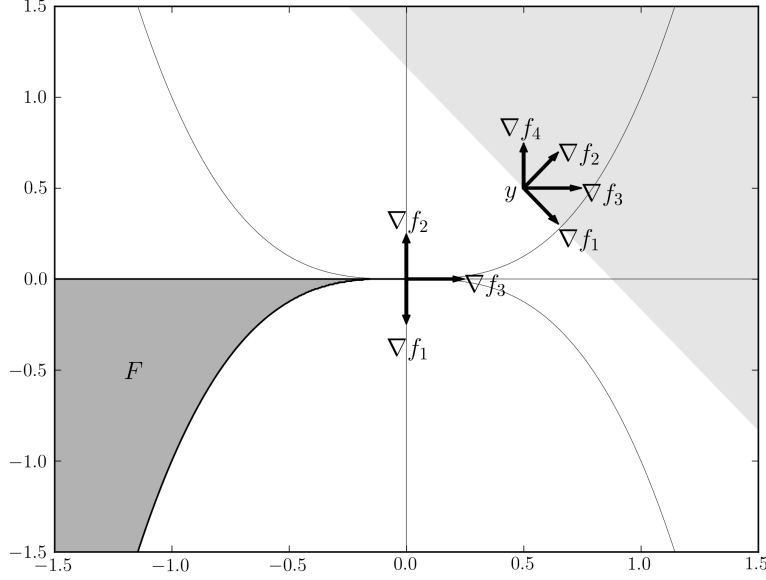
Figure 3: Consider $F = \{(x_1, x_2) \in \mathbb{R}^2 \mid f_1(x_1, x_2) = x_1^3 - x_2 \leq 0, \ f_2(x_1, x_2) = x_1^3 + x_2 \leq 0, \ f_3(x_1, x_2) = x_1 \leq 0, f_4(x_1, x_2) = x_2^3 \leq 0\}$ at the origin. Then, $\mathrm{span}_+(\{1\}, \{3\}; Gf(y))$ is the light gray semispace and contains all the gradients. On the other hand, $\nabla f_4(y) \notin \mathrm{span}_+(\{2\}, \{3\}; Gf(y))$ whenever $y \neq 0$.

such that

$$\nabla f_0(x^k) + \sum_{i \in \mathcal{I}'} \bar{\lambda}_i^k \nabla f_i(x^k) + \sum_{j \in \mathcal{J}_+} \bar{\mu}_j^k \nabla f_j(x^k) = \epsilon^k. \tag{8}$$

Define $M_k = \max\{|\bar{\lambda}_i^k|, i \in \mathcal{I}'; \ \bar{\mu}_j^k, j \in \mathcal{J}_+\}$. There are two possibilities:

1. If M_k has a bounded subsequence, we can assume, by possibly extracting a further subsequence, that for all $i \in \mathcal{I}'$ and $j \in \mathcal{J}_+$ the subsequences of $\bar{\lambda}_i^k$ and $\bar{\mu}_j^k$ have limits $\bar{\lambda}_i^*$ and $\bar{\mu}_j^* \geq 0$ respectively. Then, taking the limit at (8) we arrive at

$$\nabla f_0(x) + \sum_{i \in \mathcal{I}'} \bar{\lambda}_i^* \nabla f_i(x) + \sum_{j \in \mathcal{J}_+} \bar{\mu}_j^* \nabla f_j(x) = 0.$$

As

$$\sum_{i \in \mathcal{I}'} \bar{\lambda}_i^* \nabla f_i(x) + \sum_{j \in \mathcal{J}_+} \bar{\mu}_j^* \nabla f_j(x^k) \in \mathrm{span}_+(\mathcal{I}, \mathcal{J}; Gf(x)),$$

we see that x is KKT.

13

2. If $M_k \to \infty$, we can divide (8) by M_k for k large enough and get

$$\frac{1}{M_k}\nabla f_0(x^k) + \sum_{i \in \mathcal{I}'} \frac{\bar{\lambda}_i^k}{M_k}\nabla f_i(x^k) + \sum_{j \in \mathcal{J}_+} \frac{\bar{\mu}_j^k}{M_k}\nabla f_j(x^k) = \frac{\epsilon^k}{M_k}. \qquad (9)$$

We can then take the limit in the equation above and derive a contradiction to the fact that $(\mathcal{I}', \mathcal{J}_+)$ is positively linearly independent.

$\square$

Corollary 3.2. *The CPG condition is a constraint qualification.*

4 Relation with other constraint qualifications

Now that we know that CPG is a constraint qualification, it is natural to ask what is its relation with respect to other constraint qualifications in the literature. Let us start with its relation to RCRCQ which is naturally connected to CRSC defined in the introduction.

Theorem 4.1. *The constant rank of the subspace component condition (CRSC) implies CPG.*

Proof. Let $(\mathcal{I}', \mathcal{J}_+)$ be a positively linearly independent spanning pair of $\mathrm{span}_+(\mathcal{I}, \mathcal{J}; Gf(x))$. It suffices to show that in a neighborhood of x, $\nabla f_\ell(y) \in \mathrm{span}\{\nabla f_i(y) \mid i \in \mathcal{I}'\}, \forall \ell \in \mathcal{I} \cup \mathcal{J}_-$. We know from Lemma 2.3 that $\mathcal{I}'$ is the set of indexes of a basis for $\mathrm{span}\{\nabla f_i(x) \mid i \in \mathcal{I} \cup \mathcal{J}_-\}$. As the rank of $\{\nabla f_i(y) \mid i \in \mathcal{I} \cup \mathcal{J}_-\}$ remains constant in a neighborhood of x, this basis has to remain a basis in a (possibly smaller) neighborhood of x. $\square$

Note that, in particular, the theorem above shows that RCRCQ implies CPG, as RCRCQ implies CRSC. Moreover, CRSC successfully eliminates the need to test all subsets involving the gradients of active inequality constraints. CRSC simplified RCRCQ as RCRCQ simplified Janin's constraint qualification for feasible sets with only equality constraints.

Another constraint qualification in the same family is RCPLD, which is related to RCRCQ as CPLD is related to the original constant rank. That is, RCPLD trades the constant rank assumption in RCRCQ by the local preservation of positive linear dependence, a weaker condition:

Definition 4.1. *Let x be a feasible point of* (NOP). *Let $\tilde{\mathcal{I}}$ be the set of indexes of a basis of $\mathrm{span}\{\nabla f_i(x) \mid i \in \mathcal{I}\}$. We say that x satisfies RCPLD if there is a neighborhood $N(x)$ of x where*

1. *For all $y \in N(x)$, $\{\nabla f_i(y) \mid i \in \mathcal{I}\}$ has constant rank.*

2. *For all subsets of indexes of active inequality constraints $\tilde{\mathcal{J}} \subset \mathcal{A}(x)$, if $(\tilde{\mathcal{I}}, \tilde{\mathcal{J}})$ is positively linearly dependent at x then it remains positively linearly dependent (or equivalently, linearly dependent) in $N(x)$.*

We prove below that the RCPLD, just as RCRCQ, also preserves locally the rank of $\{\nabla f_i(y) \mid i \in \mathcal{I} \cup \mathcal{J}_-\}$, that is, it also implies CRSC.

Theorem 4.2. *RCPLD implies CRSC.*

Proof. From Corollary 2.5 we know that if $j \in \mathcal{J}_-$, $-\nabla f_j(x)$ can be positively spanned by the other vectors in the pair $(\mathcal{I}, \mathcal{J}_-)$. By the definition of *RCPLD* this fact remains true in $N(x)$ and hence, $\mathrm{span}_+(\mathcal{I}, \mathcal{J}_-; Gf(y))$ is actually a subspace for all $y \in N(x)$. What we want to show is that these subspaces have the same dimension as the subspace $\mathrm{span}_+(\mathcal{I}, \mathcal{J}_-; Gf(x))$ in a smaller neighborhood of x.

Let $\tilde{N}(x)$ a neighborhood of x contained in $N(x)$ such that the dimension of $\mathrm{span}_+(\mathcal{I}, \mathcal{J}_-; Gf(y))$ is greater or equal to the dimension of $\mathrm{span}_+(\mathcal{I}, \mathcal{J}_-; Gf(x))$, which exists as linear independence is preserved locally. We need then to show that the dimension cannot increase, remaining constant.

We start by noting that if $\tilde{\mathcal{I}}$ is as in the definition of RCPLD, $\forall y \in \tilde{N}(x)$, $\mathrm{span}_+(\mathcal{I}, \mathcal{J}_-; Gf(y)) = \mathrm{span}_+(\tilde{\mathcal{I}}, \mathcal{J}_-; Gf(y))$. Let $\tilde{m} = \#\tilde{\mathcal{I}}$, the cardinality of $\tilde{\mathcal{I}}$, $n_- = \#\mathcal{J}_-$, $\tilde{\mathcal{I}} = \{i_1, i_2, \ldots, i_{\tilde{m}}\}$ and $\mathcal{J}_- = \{j_1, j_2, \ldots, j_{n_-}\}$. Define

$$v_l(y) \overset{def}{=} \nabla f_{i_l}(y), \ l = 1, \ldots, \tilde{m},$$

$$v_{\tilde{m}+l}(y) \overset{def}{=} -\nabla f_{i_l}(y), \ l = 1, \ldots, \tilde{m},$$

$$v_{2\tilde{m}+l}(y) \overset{def}{=} \nabla f_{j_l}(y), \ l = 1, \ldots, n_-.$$

and define the set $\mathcal{A} \overset{def}{=} \{1, 2, \ldots, 2\tilde{m} + n_-\}$.

It is clear that the subspace $\mathrm{span}_+(\tilde{\mathcal{I}}, \mathcal{J}_-; Gf(y))$ is the cone positively spanned by $A(y) \overset{def}{=} \{v_l(y) \mid l \in \mathcal{A}\}$, in particular, it is linearly spanned by $A(y)$. Moreover, if a subset of vectors in $A(x)$ is linearly dependent using only non-negative weights, RCPLD asserts that, for $y \in \tilde{N}(y)$, the respective vectors in $A(y)$ remain linearly dependent using only non-negative weights.

Now let $v_{l'}(x)$ be a vector in $A(x)$ that can be positively spanned by the other vectors in $A(x)$, then $A(x) \setminus \{v_{l'}(x)\}$ still positively spans the same space. Moreover, as $A(x)$ spans the space positively, we know that $-v_{l'}(x)$ can be written as a positive combination of the remaining vectors in $A(x)$, that is, there are $\alpha_l \geq 0$ such that

$$-v_{l'}(x) = \sum_{l \in \mathcal{A} \setminus \{l'\}} \alpha_l v_l(x).$$

Using Carathéodory's Lemma [8, Exercise B.1.7] we can reduce this sum to a subset $\mathcal{A}' \subset \mathcal{A} \setminus \{l'\}$ such that the respective $\alpha_l > 0$ and the vectors $v_l(x)$, $l \in \mathcal{A}'$, are positively linearly independent. As RCPLD holds, this fact remains true in $\tilde{N}(x)$ and hence the vector $v_{l'}(y)$ is not necessary to describe the subspace linearly spanned by $A(y)$.

Hence, if we iteratively delete from $A(x)$ vectors that can be positively spanned by the other vectors in the set, delete from $\mathcal{A}$ the respective index, and call $\tilde{\mathcal{A}}$ the final index set, we can see that:

1. The subspace $\mathrm{span}_+(\mathcal{I}, \mathcal{J}_-; Gf(x))$ is the cone positively generated by the vectors in $\tilde{A}(x) \overset{def}{=} \{v_l(x) \mid l \in \tilde{\mathcal{A}}\}$ and $\tilde{A}(x)$ is a positive basis for this subspace [39].

2. for all $y \in \tilde{N}(x)$, the subspace $\mathrm{span}_+(\mathcal{I}, \mathcal{J}_-; Gf(y))$ is the subspace linearly spanned by $\tilde{A}(y) \overset{def}{=} \{v_l(y) \mid l \in \tilde{\mathcal{A}}\}$.

We can then apply Lemma 6 from [39] to $\tilde{A}(x)$ to see that there is a partition of the index set $\tilde{\mathcal{A}}$ into p pairwise disjoint subsets $\tilde{\mathcal{A}}_1 \cup \ldots \cup \tilde{\mathcal{A}}_p$ such that the positive cone generated by $\{v_l(x) \mid l \in \tilde{\mathcal{A}}_1 \cup \ldots \cup \tilde{\mathcal{A}}_{p'}\}$ is a linear subspace of dimension $(\sum_{k=1}^{p'} \#\tilde{A}_k) - p'$ for each $p' = 1, 2, \ldots, p$. In particular the dimension of the space positively spanned by $\tilde{A}(x)$ is $\#\tilde{\mathcal{A}} - p$.

Take $p' = 1$, the partition properties ensure that if we delete a vector $v_{l_1}(x)$ from $\{v_l(x) \mid l \in \tilde{\mathcal{A}}_1\}$ the remaining ones are linearly independent. Moreover, $v_{l_1}(x)$ is not only linearly dependent with the remaining ones, it is positively linearly dependent as its negative has to be positively spanned by the others. This positive linear dependence is preserved by RCPLD and hence the space linearly spanned by $\{v_l(y) \mid l \in \tilde{\mathcal{A}}_1\}$ is the same as the space linearly spanned by $\{v_l(y) \mid l \in \tilde{\mathcal{A}}_1, \ l \neq l_1\}$ for $y \in \tilde{N}(x)$.

Now, take $p' = 2$. There is vector $v_{l_2}(x) \in \{v_l(x) \mid l \in \tilde{\mathcal{A}}_2\}$ such that $\{v_l(x) \mid l \in \tilde{\mathcal{A}}_1 \cup \tilde{\mathcal{A}}_2, \ l \notin \{l_1, l_2\}\}$ is a basis of the subspace positively spanned by $\{v_l(x) \mid l \in \tilde{\mathcal{A}}_1 \cup \tilde{\mathcal{A}}_2\}$. As this space is positively spanned, we can see that there must be non-negative coefficients α_l such that

$$-v_{l_2}(x) = \sum_{l \in \tilde{\mathcal{A}}_1 \cup \tilde{\mathcal{A}}_2 \setminus \{l_2\}} \alpha_l v_l(x).$$

Again using Carathéodory's Lemma we can see that RCPLD ensures that for $y \in \tilde{N}(y)$ the vector $v_{l_2}(y)$ is not necessary to describe the subspace linearly spanned by $\{v_l(y) \mid l \in \tilde{\mathcal{A}}_1 \cup \tilde{\mathcal{A}}_2\}$. That is, for $y \in \tilde{N}(x)$ the subspace linearly spanned by $\{v_l(y) \mid l \in \tilde{\mathcal{A}}_1 \cup \tilde{\mathcal{A}}_2\}$ is the same as the one linearly spanned by $\{v_l(y) \mid l \in \tilde{\mathcal{A}}_1 \cup \tilde{\mathcal{A}}_2, \ l \neq l_2\}$ which in turn is the same as the one linearly spanned by $\{v_l(y) \mid l \in \tilde{\mathcal{A}}_1 \cup \tilde{\mathcal{A}}_2, \ l \notin \{l_1, , l_2\}\}$.

This process can be carried on p times and at the end we conclude that for all $y \in \tilde{N}(x)$ there are p vectors in $\tilde{A}(y)$ that are not necessary to describe its linearly spanned set, which in turn is $\mathrm{span}_+(\mathcal{I}, \mathcal{J}_-; Gf(y))$. Hence, the dimension of $\mathrm{span}_+(\mathcal{I}, \mathcal{J}_-; Gf(y))$ is less or equal to $\#\tilde{A}(y) - p = \#\tilde{\mathcal{A}} - p$. This last value is the dimension of the space linearly spanned by $\tilde{A}(x)$, namely $\mathrm{span}_+(\mathcal{I}, \mathcal{J}_-; Gf(x))$. $\qquad\square$

Note that CRSC condition is not equivalent to the CPG condition. Actually, consider once again the feasible set pictured in Figure 2

$$\{(x_1, x_2) \in \mathbb{R}^2 \mid x_1^3 - x_2 \leq 0, \ x_1^3 + x_2 \leq 0, \ x_1 \leq 0\}.$$

Then, at the origin $\mathcal{J}_- = \{1, 2\}$ and rank of $\{\nabla f_1(0), \nabla f_2(0)\}$ is 1. On the other hand, for any $y \neq 0$, the rank increases while CPG holds. In particular, CPG is a proper generalization of RCPLD.

Finally, let us show that CPG implies Abadie's constraint qualification. In order to achieve this we start with a result that can be directly deduced from the proof of Theorem 4.3.1 in [7]:

Lemma 4.3. *Let x be a feasible point of* (NOP) *that conforms to Mangasarian-Fromovitz's constraint qualification, i.e., the set* $\{\nabla f_i(x) \mid i \in \mathcal{I}\}$ *is linearly independent and there is a direction* $0 \neq d \in \mathbb{R}^n$ *such that*

$$\nabla f_i(x)'d = 0, \ i \in \mathcal{I}, \qquad \nabla f_i(x)'d < 0, \ i \in \mathcal{J}.$$

Then, there is a scalar $T > 0$ *and a continuously differentiable arc* $\alpha : [0, T] \to \mathbb{R}^n$ *such that*

$$\alpha(0) = x, \tag{10}$$
$$\dot{\alpha}(0) = d, \tag{11}$$
$$f_i(\alpha(t)) = 0, \ \forall t \in [0, T], \ i \in \mathcal{I}, \tag{12}$$
$$\nabla f_i(\alpha(t))'\dot{\alpha}(t) = 0, \ \forall t \in [0, T], \ i \in \mathcal{I}, \tag{13}$$
$$f_j(\alpha(t)) < 0, \ \forall t \in (0, T], \ j \in \mathcal{J}, \tag{14}$$
$$\nabla f_j(\alpha(t))'\dot{\alpha}(t) < 0, \ \forall t \in [0, T], \ j \in \mathcal{J}. \tag{15}$$

Now we use the lemma above to find special differentiable arcs that move inward the feasible set under CPG.

Lemma 4.4. *Let x be a feasible point for* (NOP) *where CPG holds and* $(\mathcal{I}', \mathcal{J}_+)$ *be the associated positively linearly independent spanning pair. Then there exists* $0 \neq d \in \mathbb{R}^n$ *such that*

$$\nabla f_i(\bar{x})'d = 0, \ i \in \mathcal{I}', \qquad \nabla f_j(\bar{x})'d < 0, \ j \in \mathcal{J}_+.$$

Moreover, for any such d, there is a scalar $T > 0$ *and a continuously differentiable arc* $\alpha : [0, T] \to \mathbb{R}^n$ *such that*

$$\alpha(0) = \bar{x}, \tag{16}$$
$$\dot{\alpha}(0) = d, \tag{17}$$
$$f_i(\alpha(t)) = 0, \ \forall t \in [0, T], \ i \in \mathcal{I}, \tag{18}$$
$$f_j(\alpha(t)) \leq 0, \ \forall t \in (0, T], \ j \in \mathcal{J}. \tag{19}$$

Proof. As $(\mathcal{I}', \mathcal{J}_+)$ is positively linearly independent, the feasible set described by

$$\{x \mid f_i(x) = 0, \ i \in \mathcal{I}', \ f_j(x) \leq 0, \ j \in \mathcal{J}_+\}$$

conforms to Mangasarian-Fromovitz's constraint qualification at x, therefore the desired direction d exists.

Let $\alpha : [0, T] \mapsto \mathbb{R}^n$ be the curve given by Lemma 4.3 and take $0 < T' \leq T$ to ensure that for all $t \in [0, T']$, $\alpha(t) \in N(x)$, where $N(x)$ is the neighborhood of x given by CPG. We already know that (18)-(19) hold for constraints with indexes in $\mathcal{I}' \cup \mathcal{J}_+$, hence all we need to show is that they also hold for $\ell \in (\mathcal{I} \cup \mathcal{J}_-) \setminus \mathcal{I}'$.

Fix such an index ℓ. We know that $\nabla f_\ell(y)$ belongs to $\mathrm{span}_+(\mathcal{I}', \mathcal{J}_+; Gf(y))$ for all $y \in N(x)$. That is, there are scalars $\lambda_i(y),\ i \in \mathcal{I}'$ and $\mu_j(y) \geq 0,\ j \in \mathcal{J}_+$ such that

$$\nabla f_\ell(y) = \sum_{i \in \mathcal{I}'} \lambda_i(y) \nabla f_i(y) + \sum_{j \in \mathcal{J}_+} \mu_j(y) \nabla f_j(y).$$

Define $\varphi_\ell(t) = f_\ell(\alpha(t))$, it follows that

$$\begin{aligned}
\varphi_\ell'(t) &= \nabla f_\ell(\alpha(t))' \dot{\alpha}(t) \\
&= \sum_{i \in \mathcal{I}'} \lambda_i(\alpha(t)) \nabla f_i(x)' \dot{\alpha}(t) + \sum_{j \in \mathcal{J}_+} \mu_j(\alpha(t)) \nabla f_j(\alpha(t))' \dot{\alpha}(t) \\
&\leq 0.
\end{aligned}$$

The last inequality follows from the sign structure given in Lemma 4.3. Hence, if ℓ is associated to an inequality constraint, (19) is proved. On the other hand, if ℓ is associated to an equality constraint, we know that $-\nabla f_\ell(x)$ also belongs to $\mathrm{span}_+(\mathcal{I}', \mathcal{J}_+; Gf(y))$ for $y \in N(x)$. We can then proceed as above to see that

$$-\varphi_\ell'(t) \leq 0.$$

And hence we conclude that (18) holds. $\qquad \square$

We are ready to show that CPG implies Abadie's constraint qualification.

Theorem 4.5. *CPG constraint qualification at x implies Abadie's constraint qualification at x.*

Proof. Let d be the direction given in Lemma 4.4 and $\epsilon > 0$ arbitrary. Given

$$\bar{d} \in \{h \mid \nabla f_i(\bar{x})' h = 0,\ i \in \mathcal{I},\ \nabla f_j(\bar{x})' h \leq 0,\ j \in \mathcal{J}\},$$

we need to show that $\bar{d}$ belongs to the tangent cone of the feasible set at x (Definition 1.1).

Clearly $\bar{d} + \epsilon d$ inherits from d the properties required to apply Lemma 4.4. Hence there is a $T > 0$ and a feasible continuously differentiable arc $\alpha : [0, T] \to \mathbb{R}^n$ such that

$$\alpha(0) = x, \qquad \dot{\alpha}(0) = \bar{d} + \epsilon d.$$

If follows that $\bar{d} + \epsilon d$ belongs to the tangent cone of the feasible set at x. Moreover, as this cone is closed, $\bar{d}$ also belongs to it. $\qquad \square$

In Figure 4, we show a complete diagram picturing the relations of CRSC and CPG with other constraint qualifications, including pseudo and quasi-normality whose definition can be found in [8]. To obtain all the relations, we used the results presented here together with the examples and results from [5].
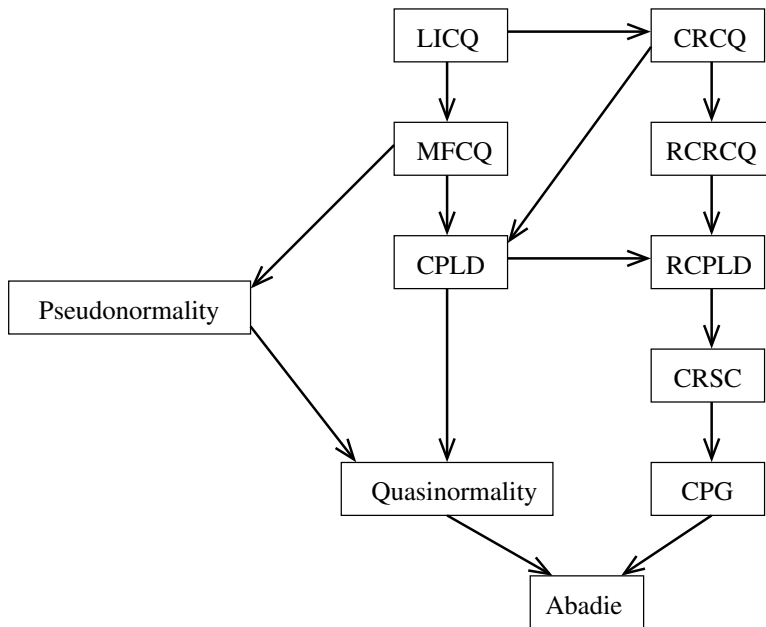
Figure 4: Complete diagram showing the relations of CRSC and CPG with other well-known constraint qualifications. An arrow between two CQs means that one is strictly stronger than the other, while conditions that are not connected by a directed path are independent from each other. Note that pseudonormality does not imply CPG as the Example 3 in [5] shows.

5 Applications of CRSC and CPG

5.1 Error bound

One interesting question about a constraint qualification is whether it implies an error bound. That is, if close to a feasible point x it is possible to estimate the distance to the feasible set F using a natural measure of infeasibility.

Definition 5.1. *[41] We say that an* error bound *holds in a neighborhood $N(x)$ of a feasible point $x \in F$ if there exists $\alpha > 0$ such that for every $y \in N(x)$*

$$\min_{z \in F} \|z - y\| \leq \alpha \max\{|f_i(y)|, \ i = 1, \dots, m;$$

$$\max\{0, f_j(y)\}, \ j = m + 1, \dots, m + p\}.$$

This property is valid for many constraint qualifications, and in particular, to weak ones such as RCPLD [5] and quasi-normality [34]. It has important theoretical and algorithmic implications, see for example [36, 41].

Unfortunately, such property does not hold for CPG, as the example in Figure 3 shows. In this case, there is no error bound around the origin. To see this, consider the sequence $x^k = (-\sqrt[3]{1/k}, 1/k)$. The distance of x^k to the feasible set is exactly $1/k$, while the infeasibility measure is $1/k^3$. Note that, by increasing the exponent that appears in the definition of the violated constraint f_4 and adapting the sequence accordingly, it is possible to make the infeasibility converge to zero as fast as $1/k^{2p+1}$, for any positive integer p, while the distance to the feasible set remains $1/k$. On the other hand, we will now show that the constant rank of the subspace component CQ (CRSC), is enough to ensure the validity of an error bound.

Throughout this subsection, we use x to denote a fixed feasible point that verifies CRSC and we denote by $B \subset \mathcal{I}$ an index set such that $\{\nabla f_i(x)\}_{i \in B}$ is a basis of $\mathrm{span}\{\nabla f_i(x)\}_{i \in \mathcal{I}}$. We will also need to compute the sets $\mathcal{J}$, $\mathcal{J}_-$, and $\mathcal{J}_+$ that appear in the definition of CRSC and CPG in points that are not x. Hence, given a feasible point y, we will use the following definitions:

$$\mathcal{J}(y) \stackrel{def}{=} \mathcal{A}(y),$$

$$\mathcal{J}_-(y) \stackrel{def}{=} \{j \in \mathcal{J}(x) \mid -\nabla f_j(y) \in \mathcal{F}(x)^\circ\},$$

$$\mathcal{J}_+(y) \stackrel{def}{=} \mathcal{J}(y) \setminus \mathcal{J}_-(y).$$

Using this notation, CRSC ensures that the rank of the vectors $\{\nabla f_i(y) \mid i \in B \cup \mathcal{J}_-(x)\}$ is constant in a neighborhood of x. Moreover, if K is an index set, let us denote by f_K the function whose components are the f_i such that i belongs to K.

We start the analysis of CRSC with a technical result.

Lemma 5.1. *Let x be a feasible point that verifies CRSC. Then, there exist scalars λ_i, $i \in B$, μ_j with $\mu_j > 0$ for all $j \in \mathcal{J}_-(x)$ such that*

$$\sum_{i \in B} \lambda_i \nabla f_i(x) + \sum_{j \in \mathcal{J}_-(x)} \mu_j \nabla f_j(x) = 0. \tag{20}$$

Proof. We know that for any index $l \in \mathcal{J}_-(x)$ there exist scalars $\lambda_i^l, i \in B, \mu_j^l$ with $\mu_j^l \geq 0$ such that

$$-\nabla f_l(x) = \sum_{i \in B} \lambda_i^l \nabla f_i(x) + \sum_{j \in \mathcal{J}_-(x)} \mu_j^l \nabla f_j(x).$$

Thus, adding for all $l \in \mathcal{J}_-(x)$ both sides of the above equality and rearranging the resulting terms, we get

$$\sum_{i \in B} \gamma_i \nabla f_i(x) + \sum_{j \in \mathcal{J}_-(x)} \theta_j \nabla f_j(x) = 0,$$

where $\gamma_i = \sum_{i \in B} \lambda_i^l$ and $\theta_j = 1 + \sum_{j \in \mathcal{J}_-(x)} \mu_j^l \geq 1 > 0$. $\qquad\square$

The next lemma extends an important result from Lu for CRCQ [23] to CRSC. Namely, it shows that the constraints f_j with $j \in \mathcal{J}_-(x)$ are actually equality constraints under the disguise of inequalities.

Lemma 5.2. *Let x be a feasible point that verifies CRSC. Then, there exists a neighborhood $N(x)$ of x such that, for every $i \in \mathcal{J}_-(x)$, $f_i(y) = 0$ for all feasible points $y \in N(x)$.*

Proof. From the previous lemma there exist scalars $\lambda_i, i \in B$ and $\mu_j > 0$ for all $j \in \mathcal{J}_-(x)$ such that (20) holds.

Since the rank of the vectors $\{\nabla f_i(y) \mid i \in B \cup \mathcal{J}_-(x)\}$ is constant for y in a neighborhood of x we can use [23, Proposition 1], defining the index sets K and J_0 in [23] as the sets $\mathcal{J}_-(x)$ and B respectively to complete the proof. □

Observe that, even though the hypothesis considered in [23, Proposition 1] is the constant rank constraint qualification, the proof is obtained applying the respective Lemma 1 where only the constant rank of the gradients in $K = \mathcal{J}_-(x)$ and $J_0 = B$ is used. Actually such lemma can be viewed as a variation of the Constant Rank Theorem [25] where only the rank of all gradients has to remain constant.

Now we are ready to show that the CRSC condition is preserved locally. That is, if it holds at a feasible point x, it must hold at all feasible points in a neighborhood of x. We start doing this by showing that the index set $\mathcal{J}_-$ is stable locally.

Lemma 5.3. *Let x be a feasible point that verifies CRSC. Then, there exists a neighborhood $N(x)$ of x such that $\mathcal{J}_-(y) = \mathcal{J}_-(x)$ for all feasible points $y \in N(x)$.*

Proof. From Lemma 5.1 we know that there exist scalars $\lambda_i, i \in B, \mu_j$ with $\mu_j > 0$ for all $j \in \mathcal{J}_-(x)$ such that (20) holds.

Let us take a subset $\widehat{\mathcal{J}} \subset \mathcal{J}_-(x)$ such that the set of gradients $\{\nabla f_i(x)\}_{i \in B \cup \widehat{\mathcal{J}}}$ is a basis of $\mathrm{span}\{\nabla f_i(x)\}_{i \in \mathcal{I} \cup \mathcal{J}_-(x)}$. Clearly the set of gradients

$$\{\nabla f_i(x)\}_{i \in B \cup \widehat{\mathcal{J}}} \tag{21}$$

is linearly independent.

Define the function

$$h(y) = - \sum_{j \in \mathcal{J}_-(x) \setminus \widehat{\mathcal{J}}} \mu_j f_j(y)$$

and let us consider a new feasible set F^h adding to the original feasible set F the equality constraint $h(y) = 0$, which is locally redundant by Lemma 5.2. Let us define $\mathcal{J}_-^h(\cdot)$ analogously for F^h as we define $\mathcal{J}_-(\cdot)$ for the original feasible set F. Thus, we have:

1. $h(y) = 0$, for all $y \in F \cap N(x)$.

2. $\nabla h(x) \in \mathcal{J}_{-}^{h}(x)$.

3. the set of gradients

$$\{\nabla h(y), \nabla f_i(y) \mid i \in B \cup \widehat{\mathcal{J}}\} \tag{22}$$

   has constant rank in a neighborhood of x, as ∇h is a combination of ∇f_i, $i \notin B \cup \widehat{\mathcal{J}}$, and each of the later gradients are generated by ∇f_i, $i \in B \cup \widehat{\mathcal{J}}$, by CRSC.

Recalling (20) we get

$$\nabla h(x) = - \sum_{j \in \mathcal{J}_{-}(x) \setminus \widehat{\mathcal{J}}} \mu_j \nabla f_j(x) = \sum_{i \in B} \lambda_i \nabla f_i(x) + \sum_{j \in \widehat{\mathcal{J}}} \mu_j \nabla f_j(x) \tag{23}$$

and therefore, using conditions (21)-(22) we can apply [23, Corollary 1] to obtain neighborhoods $N(x)$ of x, Z of $(f_B(x), f_{\widehat{\mathcal{J}}}(x))$, with Z being convex, and a continuously differentiable function $g : Z \to \mathbb{R}$ such that

$$h(x) = g(f_B(x), f_{\widehat{\mathcal{J}}}(x)) \tag{24}$$

and, for every $z \in Z$

$$\mathrm{sgn}\left(\frac{\partial g}{\partial z_i}(z)\right) = \mathrm{sgn}(\lambda_i), \ \forall i \in B, \tag{25}$$

$$\mathrm{sgn}\left(\frac{\partial g}{\partial z_i}(z)\right) = \mathrm{sgn}(\mu_i), \ \forall i \in \widehat{\mathcal{J}}. \tag{26}$$

Thus, by the definition of h and (24), it follows that for all $y \in F$ in a neighborhood of x

$$\nabla h(y) = - \sum_{j \in \mathcal{J}_{-}(x) \setminus \widehat{\mathcal{J}}} \mu_j \nabla f_j(y) \tag{27}$$

$$= \sum_{i \in B} \frac{\partial g}{\partial z_i}(f_B(y), f_{\widehat{\mathcal{J}}}(y)) \nabla f_i(y) + \sum_{j \in \widehat{\mathcal{J}}} \frac{\partial g}{\partial z_j}(f_B(y), f_{\widehat{\mathcal{J}}}(y)) \nabla f_j(y). \tag{28}$$

and, using (25)-(26) and (28), there are scalars $\gamma_i(y) = \frac{\partial g}{\partial z_i}(f_B(y), f_{\widehat{\mathcal{J}}}(y))$ and $\theta_j(y) = \frac{\partial g}{\partial z_j}(f_B(y), f_{\widehat{\mathcal{J}}}(y)) > 0$ such that

$$\sum_{j \in \mathcal{J}_{-}(x) \setminus \widehat{\mathcal{J}}} \mu_j \nabla f_j(y) + \sum_{i \in B} \gamma_i(y) \nabla f_i(y) + \sum_{j \in \widehat{\mathcal{J}}} \theta_j(y) \nabla f_j(y) = 0.$$

From the last expression, Lemma 5.2, and the definition of $\mathcal{J}_{-}(y)$ we obtain that $\mathcal{J}_{-}(y) = \mathcal{J}_{-}(x)$. $\qquad\square$

This fact shows that the constraint qualification CRSC is preserved locally, as the set $\mathcal{J}_{-}(x)$ is constant in a neighborhood of a feasible point where CRSC holds. We are ready to show that CRSC implies an error bound.

Theorem 5.4. *If $x \in F$ verifies CRSC and the functions f_i, $i = 1, \ldots, m+p$ defining F admit second derivatives in a neighborhood of x, then an error bound holds in a neighborhood of x.*

Proof. First, let us recall that Lemma 5.2 states that the constraints in $\mathcal{J}_-(x)$ are actually equality constraints in a neighborhood of x. Hence, it is natural to consider the feasible set F^E

$$F^E = \{y \in \mathbb{R}^n \mid f_i(y) = 0, \ \forall i \in \mathcal{I} \cup \mathcal{J}_-(x), \ f_j(y) \leq 0, \ \forall j \in \mathcal{J}_+(x)\},$$

which is equivalent to the original feasible set F close to x. It is trivial to see that the CRSC point (with respect to F) x verifies RCPLD as a feasible point of the set F^E. Now, using [5, Theorem 7] it follows that there exist $\alpha > 0$ and a neighborhood $N(x)$ of x such that for every $y \in N(x)$

$$\min_{z \in F} \|z - y\| = \min_{z \in F^E} \|z - y\| \leq \alpha r^E(y), \tag{29}$$

with

$$r^E(y) = \max\{\|f_{\mathcal{I} \cup \mathcal{J}_-(x)}(y)\|_\infty, \ \|\max\{0, f_{\mathcal{J}_+(x)}(y)\}\|_\infty\}. \tag{30}$$

Now, from Lemma 5.1 we know that there are scalars λ_i, $i \in B$ and μ_j, with $\mu_j > 0$ for all $j \in \mathcal{J}_-(x)$, such that (20) holds. Let $\widehat{\mathcal{J}}$ be as in the proof of Lemma 5.3, that is, $\widehat{\mathcal{J}}$ is a subset of $\mathcal{J}_-(x)$ such that the set of gradients $\{\nabla f_i(x)\}_{i \in B \cup \widehat{\mathcal{J}}}$ is a basis for $\mathrm{span}\{\nabla f_i(x)\}_{i \in \mathcal{I} \cup \mathcal{J}_-(x)}$. Let us consider also the function

$$h(y) = -\sum_{j \in \mathcal{J}_-(x) \setminus \widehat{\mathcal{J}}} \mu_j f_j(y).$$

Following the proof of Lemma 5.3, there are a neighborhood $N(x)$ of x, a neighborhood Z of $(f_B(x), f_{\widehat{\mathcal{J}}}(x))$, with Z being convex, and a continuously differentiable function $g : Z \to \mathbb{R}$ such that (24)-(26) holds. By shrinking $N(x)$ if necessary, we can assume that the partial derivatives of g will preserve the signs at $(f_B(x), f_{\widehat{\mathcal{J}}}(x))$. That is, we may assume the existence of constants $0 < \mu_m \leq \mu_M$ and λ_M such that $\mu_m \leq \frac{\partial g}{\partial z_j}(z) \leq \mu_M$, $\forall j \in \mathcal{J}_-(x)$ and $|\frac{\partial g}{\partial z_i}(z)| \leq \lambda_M$, $\forall i \in B$ and $\forall z \in Z$.

Thus, from the convexity of Z and the differentiability of g we can apply the Mean Value Theorem to see that, for each $y \in N(x)$ there exist $\xi_y \in Z$ between $(0,0) = (f_B(x), f_{\widehat{\mathcal{J}}}(x))$ and $(f_B(y), f_{\widehat{\mathcal{J}}}(y))$ such that

$$g(f_B(y), f_{\widehat{\mathcal{J}}}(y)) = \sum_{i \in B \cup \widehat{\mathcal{J}}} \frac{\partial g}{\partial z_i}(\xi_y) f_i(y).$$

This implies that

$$-\sum_{j \in \mathcal{J}_-(x) \setminus \widehat{\mathcal{J}}} \mu_j f_j(y) = \sum_{i \in B \cup \widehat{\mathcal{J}}} \frac{\partial g}{\partial z_i}(\xi_y) f_i(y) \tag{31}$$

and, for every $l \in \mathcal{J}_-(x) \setminus \widehat{\mathcal{J}}$ we can write

$$-\mu_l f_l(y) = \sum_{i \in B \cup \widehat{\mathcal{J}}} \frac{\partial g}{\partial z_i}(\xi_y) f_i(y) + \sum_{j \in (\mathcal{J}_-(x) \setminus \widehat{\mathcal{J}}) \setminus \{l\}} \mu_j f_j(y).$$

Since $\mu_l > 0$, it follows that

$$
\begin{aligned}
|f_l(y)| &\leq \frac{1}{\mu_l} \left(\sum_{i \in B \cup \widehat{\mathcal{J}}} \left| \frac{\partial g}{\partial z_i}(\xi_y) \right| |f_i(y)| + \sum_{j \in (\mathcal{J}_-(x) \setminus \widehat{\mathcal{J}}) \setminus \{l\}} |\mu_j| \max\{0, f_j(y)\} \right) \\
&\leq \frac{\max\{\mu_M; \ |\mu_j|, j \in \mathcal{J}_-(x) \setminus \widehat{\mathcal{J}}\}}{\mu_m} \left(\sum_{i \in B \cup \widehat{\mathcal{J}}} |f_i(y)| \right. \\
&\qquad\qquad\qquad\qquad\qquad\qquad \left. + \sum_{j \in \mathcal{J}_-(x) \setminus \widehat{\mathcal{J}}} \max\{0, f_j(y)\} \right).
\end{aligned}
$$

Thus, $\forall l \in \mathcal{J}_-(x) \setminus \widehat{\mathcal{J}}$, there is a $K > 0$ large enough such that

$$|f_l(y)| \leq K \max\{|f_i(y)|, \ i \in \mathcal{I}; \ \max\{0, f_j(y)\}, \ j \in \mathcal{J}\}. \tag{32}$$

If $l \in \widehat{\mathcal{J}}$, from (31), we obtain a similar bound

$$|f_l(y)| \leq \widetilde{K} \max\{|f_i(y)|, \ i \in \mathcal{I}; \ \max\{0, f_j(y)\}, \ j \in \mathcal{J}\}, \tag{33}$$

for some $\widetilde{K} > 0$.

Using (32)-(33) and (29)-(30) we obtain the desired result. $\qquad\square$

5.2 Algorithmic Applications of CPG

In this section, we show how the CPG condition can be used in the analysis of many algorithms for nonlinear optimization. The objective is to show that CPG can replace other, more stringent, constraint qualifications in the assumptions that ensure global convergence. We will show specific results for the main classes of algorithms for optimization, namely sequential quadratic programming (SQP), interior point methods, augmented Lagrangians, and inexact restoration.

5.2.1 Sequential quadratic programming

We start by extending the global convergence result of the general sequential quadratic programming (SQP) method studied by Qi and Wei [38]. In their work, Qi and Wei introduced the CPLD CQ and extended convergence results for SQP methods that previously were based on the MFCQ. In order to do so, their main tool was the notion of approximate KKT sequences.

Definition 5.2. *We say that $\{x^k\}$ is an approximate-KKT sequence, or simply an AKKT sequence, of* (NOP) *if there is a sequence* $\{(\lambda^k, \mu^k, \epsilon^k, \delta^k, \gamma_k)\} \in \mathbb{R}^m \times \mathbb{R}^p \times \mathbb{R}^n \times \mathbb{R}^p \times \mathbb{R}$ *such that*

$$
\begin{cases}
\nabla f_0(x^k) + \sum_{i=1}^m \lambda_i \nabla f_i(x^k) + \sum_{j=1}^p \mu_j \nabla f_{m+j}(x^k) = \epsilon^k \\
f_{m+j}(x^k) \leq \delta^k, \ j = 1, \ldots, p \\
\mu^k \geq 0 \\
\mu_j^k (f_{m+j}(x^k) - \delta_j^k) = 0, \ j = 1, \ldots, p \\
|f_i(x^k)| \leq \gamma_k, i = 1, \ldots, m,
\end{cases}
$$

where $\{(\epsilon^k, \delta^k, \gamma_k)\}$ *converges to zero.*

It is easy to see that approximate KKT sequences are closely related to approximate KKT feasible points from Definition 3.2. Actually, AKKT (feasible) points are exactly the limit points of AKKT sequences. Hence we can easily recast the results from [38] in terms of AKKT points.

In particular, Theorem 2.7 from [38], which ensures that limits of AKKT sequences are actually KKT, is just a particular case of Theorem 3.1 above, requiring CPLD, a more stringent constraint qualification, in the place of CPG. Hence, we may use Theorem 3.1 to generalize some convergence results from [38] replacing CPLD by CPG.

In order to do so, let us recall the general SQP method from [38]:

Algorithm 5.1. (General SQP) *Let $C > 0$, $x^0 \in F$, $H_0 \in \mathbb{R}^{n \times n}$ a symmetric positive definite matrix.*

1. *(Initialization) Set $k = 0$*

2. *(Computation of a search direction) Compute d^k solving the quadratic programming problem*

$$
\begin{aligned}
\min \quad & \frac{1}{2} d' H_k d + \nabla f(x^k)' d \\
s.t. \quad & f_i(x^k) + \nabla f_i(x^k)' d = 0, \ i = 1, \ldots, m, \qquad \text{(QP)} \\
& f_i(x^k) + \nabla f_i(x^k)' d \leq 0, \ i = m+1, \ldots, m+p.
\end{aligned}
$$

   if $d^k = 0$, stop.

3. *(Line search and additional correction) Determine the steplength $\alpha_k \in (0,1)$ and a correction direction $\bar{d}^k$ such that*

$$
\|\bar{d}^k\| \leq C\|d^k\|.
$$

4. *(Updates) Compute a new symmetric positive definite Hessian approximation H_{k+1}. Set $x^{k+1} = x^k + \alpha_k d^k + \bar{d}^k$ and $k = k+1$. Go to Step 1.*

As stated in [38], this algorithm is a general model for SQP methods, where specific choices for the Hessian approximations H_k, step length α_k, and correction steps $\bar{d}^k$ are defined. Moreover, if the algorithm stops at Step 2, then x^k is a KKT point for (NOP). Hence, when analyzing such method we only need to consider the case where it generates an infinite sequence. The result below is a direct generalization of Theorem 4.2 in [38] where we use CPG instead of CPLD.

Theorem 5.5. *Assume that the General SQP algorithm generates an infinite sequence $\{x^k\}$ and that this sequence has an accumulation point x^*. Let $\mathcal{L}$ be the index set associate to it, that is*

$$\lim_{k \in \mathcal{L}} x^k = x^*.$$

Suppose that CPG holds at x^* and that the Hessian estimates H_k are bounded. If

$$\liminf_{k \in \mathcal{L}} \|d_k\| = 0,$$

then x^* is a KKT point of (NOP).

Proof. We just follow the proof of Theorem 4.2 in [38] to see that it shows that under the assumptions above $\{x^k\}_{k \in \mathcal{L}}$ is an AKKT sequence. Hence, as discussed before, x^* is an AKKT point, which is KKT whenever CPG holds by Theorem 3.1. □

In order to present a concrete SQP algorithm that conforms to the assumptions of the theorem above, Qi and Wei recover the Panier-Tits SQP feasible algorithm for inequality constrained problems [37]. As pointed out by Qi and Wei, this method can be seen as a special case of the General SQP algorithm.

The Panier-Tits method depends on the validity of MFCQ on the computed iterates to be well defined. However, as pointed out by Qi and Wei, MFCQ does not need to hold at the limit points, where CPLD suffices. Once again we can generalize this result using CPG.

Theorem 5.6. *Consider the Panier-Tits feasible SQP method described in [38, Algorithm B]. Let $\{x^k\}$ be an infinite sequence generated by this method and H_k the respective Hessian approximations. Suppose MFCQ holds at all feasible points that are not KKT, that the Hessian estimates H_k are bounded, and let x^* be an accumulation point of $\{x^k\}$ where CPG holds. Then, x^* is a KKT point of* (NOP).

Proof. Once again we only need to follow the proof from Theorem 5.3 in [38] and use Theorem 5.5 above instead of its particular case [38, Theorem 4.2]. □

Note that it is easy to build examples where MFCQ holds at all feasible points but one, where CPG holds and CPLD does not hold. See, Figure 2 above. Hence the theorem above is a real generalization of Qi and Wei's result.

5.2.2 Interior point methods

Let us now turn our attention to how CPG can be used in the analysis of interior point methods for nonlinear optimization. In this context the usual constraint qualification is Mangasarian-Fromovitz's CQ [10, 12, 14].

It is interesting to understand why the definition of CPLD did not result in the generalization of the convergence conditions for such methods. To this effect, let us focus on problems with inequality constraints only. In this case, it is natural to assume that the optimization problem satisfies a sufficient interior property, that is, that every local minimizer can be arbitrarily approximated by strictly feasible points. It is known from [16] that CPLD together with such sufficient interior property is equivalent to MFCQ. Hence, it is fruitless to use CPLD to generalize results based on MFCQ in the context of interior point methods. Moreover, it is possible to replace CPLD with CRSC in the previous discussion since Lemma 5.2 shows that $\mathcal{J}_-(x) = \emptyset$ whenever CRSC and the sufficient interior property hold at a feasible point x, that is, MFCQ holds.

On the other hand, the example in Figure 3 shows that CPG and the sufficient interior property can hold together even when other constraint qualifications fail, in particular, MFCQ. Moreover, it was proved in [4] that the classic barrier method generates sequences with AKKT limit points. Hence, Theorem 3.1 shows that such limit points satisfy the KKT condition if CPG holds. This fact opens the path to prove convergence of modern interior point methods under less restrictive constraint qualifications. In particular, we generalize below the convergence results for the quasi-feasible interior point method of Chen and Goldfarb [10].

This algorithm consists in applying a log-barrier strategy to solve the general optimization problem (NOP), yielding a sequence of subproblems (FP_{ζ_l}), where the barrier sequence $\{\zeta_l\}$ should be driven to 0.

$$
\begin{aligned}
\min \quad & f_0(x) - \zeta_l \sum_{i=m+1}^{m+p} \log(-f_i(x)) \\
s.t. \quad & f_i(x) = 0, \ i = 1, \ldots, m, \\
& f_j(x) < 0, \ j = m+1, \ldots, m+p.
\end{aligned}
\qquad (\text{FP}_{\zeta_l})
$$

Algorithm I in [10] uses an ℓ_2-norm penalization to deal with the equality constraints in (FP_{ζ_l}), and tries to solve it approximately employing a Newton-like approach. More formally, given a barrier parameter $\zeta_l > 0$ and an error tolerance $\varepsilon_l > 0$, Algorithm I tries to find $x^l \in \mathbb{R}^n$, $\lambda^l \in \mathbb{R}^m$, and $\mu^l \in \mathbb{R}^p$ such

that $f_j(x^l) < 0, \ j = m+1\ldots m+p$, and

$$\|\nabla f_0(x^l) + \sum_{i=1}^{m} \lambda_i^l \nabla f_i(x^l) + \sum_{j=1}^{p} \mu_j^l \nabla f_{m+j}(x^l)\| \leq \varepsilon_l, \tag{34}$$

$$\forall i = 1,\ldots,m, \quad |f_i(x^l)| \leq \varepsilon_l, \tag{35}$$

$$\forall j = 1,\ldots,p, \quad |f_{m+j}(x^l)\mu_j^l + \zeta_l| \leq \varepsilon_l, \tag{36}$$

$$\forall j = 1,\ldots,p, \quad \mu_j^l \geq -\varepsilon_l. \tag{37}$$

The conditions above are simply the termination criteria defining a successful run of Algorithm I as stated in [10, Eq. (3.13)]. Moreover, (34)–(37) is an approximate version of the KKT conditions for (FP_{ζ_l}).

Algorithm II is then defined in [10] as employing Algorithm I to approximately solve (FP_{ζ_l}) for a sequence of barrier parameters $\zeta_l > 0$ and error tolerance $\varepsilon_l > 0$ both converging to 0. We show below that it is possible to improve the convergence results of this method using CPG instead of MFCQ.

Theorem 5.7. *Assume that the standard assumptions A1–A2 of [10] hold, that is, there exists a point x_0 such that $f_i(x_0) < 0, i = m+1,\ldots,m+p$ and the functions $f_i, i = 0,\ldots,m+p$ are twice continuously differentiable. Consider Algorithm II with sequences $\zeta_l > 0$ and $\varepsilon_l > 0$ both converging to zero. There are two possibilities:*

1. *For each ζ_l and $\varepsilon_l > 0$, Algorithm I terminates satisfying conditions (34)–(37) and in particular Algorithm II generates a sequence $\{x^l\}$. If this sequence admits a limit point x^*, then it is feasible and, if CPG with respect to (NOP) holds at x^*, it is also a KKT point of (NOP).*

2. *For some barrier parameter ζ_l, the termination criteria of Algorithm I are never met. Let $\{x^{l,k}\}$ be the sequence computed by Algorithm I with penalty parameters associated to the equality constraints $\{r_{l,k}\}$. Suppose further that Assumptions A3–A4 of [10] hold, that is, the sequence $\{x^{l,k}\}$ and the modified Hessian sequence $\{\mathcal{H}^{l,k}\}$ used in Algorithm I is bounded. Let x^* be a limit point of $\{x^{l,k}\}$. If CPG with respect to the infeasibility problem*

$$\min \quad \sum_{i=1}^{m} f_i(x)^2$$
$$\text{s.t.} \quad f_i(x) \leq 0, \ i = m+1,\ldots,m+p$$

holds at x^* and $\{\nabla f_i(x^*)\}_{i=1}^{m}$ is linearly independent, then x^* is a KKT point of such infeasibility problem.

Proof. First let us consider the case where Algorithm I successfully terminates conforming to (34)-(37) for all barrier parameters ζ_l. Let x^* be an accumulation point of $\{x^l\}$ and $\mathcal{L}$ be the associated index set, that is, $x^l \to_{\mathcal{L}} x^*$.

To see that x^* is feasible, we start noting that (35) and $\varepsilon_l \to 0$ ensure that x^* respects all the equality constraints. Moreover, as a limit of points that obey (strictly) the inequalities, x^* also conforms to the inequality constraints.

Now, we show that x^* is AKKT. Let us start with the observation that for each $j = 1, \ldots, p$, inequality (37) implies that either $\mu_j^l \to_{\mathcal{L}} 0$, or there is a $\delta_j > 0$ and an infinite index set contained in $\mathcal{L}$ where $\mu_j^l > \delta_j$. Hence, repeating this procedure p times we can obtain a disjoint partition $I_1 \cup I_2 = \{1, \ldots, p\}$, an infinite index set $\mathcal{L}' \subset \mathcal{L}$, and a $\delta > 0$, such that $\forall j \in I_1, \mu_j^l \to_{\mathcal{L}'} 0$ and $\forall j \in I_2, \mu_j^l > \delta$. In particular, if $j \notin \mathcal{A}(x^*)$, inequality (36) together with $\zeta_l \to 0$ and $\varepsilon_l \to 0$ imply that $\mu_j^l \to_{\mathcal{L}} 0$. That is $j \in I_1$.

Next we recover (34) and see that for $l \in \mathcal{L}'$

$$\|\nabla f_0(x^l) + \sum_{i=1}^m \lambda_i^l \nabla f_i(x^l) + \sum_{j \in I_2} \mu_j^l \nabla f_{m+j}(x^l)\| \le \varepsilon'_{\zeta_l},$$

where ε'_{ζ_l} is defined as $\varepsilon_l + \|\sum_{j \in I_1} \mu_j^l \nabla f_{m+j}(x^l)\|$. Using the continuity of the gradients of the constraints, $\varepsilon_l \to 0$, and $\forall j \in I_1, \mu_j^l \to_{\mathcal{L}'} 0$, it follows that $\varepsilon'_{\zeta_l} \to_{\mathcal{L}'} 0$ and therefore x^* is AKKT.

Finally, we can use Theorem 3.1 to assert that the validity of CPG with respect to (NOP) holds at x^* is enough to ensure that x^* is a KKT point of (NOP).

Now consider the case where Algorithm I generates an infinite sequence for a fixed barrier parameter ζ_l. There are two possibilities:

1. The penalty parameters $r_{l,k}$ are driven to infinity. In this case we follow the proof of [10, Theorem 3.6]. Dividing [10, Equation (3.16)] by the previously defined $\alpha_{l,k} = \max\{r_{l,k}, \|\mu_{l,k}\|_\infty\}$, where $\mu_{l,k}$ is the current multiplier estimate for the inequalities, it follows easily that x^* is an AKKT point of the infeasibility problem above. Hence, it is a KKT point of such problem if it fulfills CPG.

2. If the infinite sequence generated by Algorithm I is such that $r_{l,k}$ is bounded, then we follow the proof of [10, Lemma 3.8] to arrive at a contradiction with respect to the linear independence of the gradients of equality constraints.

$\square$

Note that the assumption that CPG holds with linear independence of equality constraint gradients is a real weakening of MFCQ, as can be seen by the example in Figure 3.

5.2.3 Augmented Lagrangians and inexact restoration

Finally, let us look at augmented Lagrangians algorithms, in particular we consider the variant introduced in [3, 2] and show that it converges under CPG.

This algorithm can solve problems in the form

$$
\begin{aligned}
\min \quad & f_0(x) \\
s.t. \quad & f_i(x) = 0, \ i = 1, \ldots, m, \\
& f_j(x) \leq 0, \ j = m+1, \ldots, m+p, \\
& x \in X,
\end{aligned}
\tag{NOP-LA}
$$

where the set $X = \{x \mid \underline{f}_i(x) = 0, \ i = 1, \ldots, \underline{m}, \ \underline{f}_j(x) \leq 0, \ j = \underline{m}+1, \ldots, \underline{m}+\underline{p}\}$ is composed of easy constraints that can be enforced by a readily available solver.

In the original papers, the global convergence of the augmented Lagrangian algorithm was obtained assuming CPLD. Such results were recently extended to require only RCPLD [5]. In these works, the basic idea was to explore the fact that the algorithm can only converge to AKKT points and then use a special case of Theorem 3.1 above to show that the limit points are actually KKT points. The same line of reasoning can then be followed requiring only CPG and generalizing the convergence result.

Theorem 5.8. *Let x^* be a limit point of a sequence generated by the augmented Lagrangian algorithm described in [3, 2]. Then one of the four conditions below holds:*

1. *CPG with respect to the set X does not hold at x^*.*

2. *x^* is not feasible and it is a KKT point of the problem*

$$
\begin{aligned}
\min \quad & \sum_{i=1}^{m} f_i^2(x) + \sum_{j=m+1}^{m+p} \max\{0, f_j(x)\}^2 \\
s.t. \quad & x \in X.
\end{aligned}
$$

3. *x^* is feasible, but CPG fails at x^* when taking into account the full set of constraints.*

4. *x^* is KKT.*

We close this section mentioning that Theorem 3.1 also proves convergence of inexact restoration methods [28, 29, 30, 11] to KKT points under CPG, since limit points of sequences generated by these methods satisfy the LAGP optimality condition [4, 31], which implies AKKT [17].

6 Conclusion

We presented two new constraint qualifications that are weaker than the previous CQs based on constant rank and constant positive linear dependence.

The first CQ, that we called Constant Rank of the Subspace Component (CRSC), solves the open problem on identifying the specific set of gradients whose rank must be preserved locally and still ensure that the constraints are

qualified. We achieved this by defining the set of active inequality constraints that resemble equalities, the set $\mathcal{J}_-$. We proved that under CRSC those inequalities are actually equalities locally and showed that an error bound holds.

The second CQ is more general and was called the Constant Positive Generator (CPG) condition. It basically asks that a generalization of the notion of a basis for a cone must be preserved locally. This condition is very weak and can even hold in a point where Guignard's constraint qualification fails in a neighborhood. Despite its weakness, we showed that this condition is enough to ensure that AKKT points conform to the KKT optimality conditions and hence CPG can be used to extend global convergence results of many algorithms for nonlinear optimization.

The definition of these two new CQs opens the path for several new research directions. For example, it would be interesting to investigate if CRSC can be used to extend results on sensitivity and perturbation analysis that already exist for RCRCQ and CPLD [22, 23, 24, 33]. Other possibility would be to extend CRSC to the context of problems with complementarity or vanishing constraints [18, 21], as was done recently for CPLD in [19, 20]. Another interesting area of research is to search for alternative proofs or methods that allow to drop the CQs that are stronger than CPG and that are still required in the convergence analysis of SQP and interior point methods presented in Section 5.2.

References

[1] J. Abadie. *On the Kuhn-Tucker Theorem*, pages 21–36. John Wiley, New York, 1967.

[2] R. Andreani, E. G. Birgin, J. M. Martínez, and M. L. Schuverdt. On Augmented Lagrangian Methods with General Lower-Level Constraints. *SIAM Journal on Optimization*, 18(4):1286, November 2007.

[3] R. Andreani, E. G. Birgin, J. M. Martínez, and M. L. Schuverdt. Augmented Lagrangian methods under the constant positive linear dependence constraint qualification. *Mathematical Programming*, 111(1-2):5–32, December 2008.

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

[5] R. Andreani, G. Haeser, M. L. Schuverdt, and P. J. S. Silva. A relaxed constant positive linear dependence constraint qualification and applications. *Mathematical Programming*, to appear, May 2011. DOI: 10.1007/s10107-011-0456-0.

[6] R. Andreani, J. M. Martínez, and M. L. Schuverdt. On the Relation between Constant Positive Linear Dependence Condition and Quasinormality

Constraint Qualification. *Journal of Optimization Theory and Applications*, 125(2):473–483, May 2005.

[7] M. S. Bazaraa, H. D. Sherali, and C. M. Shetty. *Nonlinear programming: theory and algorithms.* John Wiley and Sons, 2006.

[8] D. P. Bertsekas. *Nonlinear programming.* Athena Scientific, Belmont Mass., 2nd ed. edition, 1999.

[9] D. P. Bertsekas. *Convex analysis and optimization.* Athena Scientific, 2003.

[10] L. Chen and D. Goldfarb. Interior-point ℓ_2-penalty methods for nonlinear programming with strong global convergence properties. *Mathematical Programming*, 108:1–26, 2006.

[11] A. Fischer and A. Friedlander. A new line search inexact restoration approach for nonlinear programming. *Computational Optimization and Applications*, 46(2):333–346, 2010.

[12] A. Forsgren, P. Gill, and M. Wright. Interior points for nonlinear optimization. *SIAM Reviews*, 44:525–597, 2002.

[13] F. J. Gould and J. W. Tolle. A Necessary and Sufficient Qualification for Constrained Optimization. *SIAM Journal on Applied Mathematics*, 20(2):164, July 1971.

[14] C. Grossmann, D. Klatte, and B. Kummer. Convergence of primal-dual solutions for the nonconvex log-barrier method without licq. *Kybernetika*, 20(5):571–584, 2004.

[15] M. Guignard. Generalized Kuhn-Tucker Conditions for Mathematical Programming Problems in a Banach Space. *SIAM Journal on Control*, 7(2):232, July 1969.

[16] G. Haeser. On the global convergence of interior–point nonlinear programming algorithms. *Comput. Appl. Math.*, 29(2):125–138, 2010.

[17] G. Haeser and M. L. Schuverdt. On approximate KKT condition and its extension to continuous variational inequalities. *Journal of Optimization Theory and Applications*, 149(3):125–138, 2011.

[18] T. Hoheisel and C. Kanzow. First- and second-order optimality conditions for mathematical programs with vanishing constraints. *Applications of Mathematics*, 52:495–514, 2007.

[19] T. Hoheisel, C. Kanzow, and A. Schwartz. Mathematical programs with vanishing constraints: A new regularization approach with strong convergence properties. *Optimization*, to appear.

[20] T. Hoheisel, C. Kanzow, and A. Schwartz. Theoretical and numerical comparison of relaxation methods for mathematical programs with complementarity constraints. *Mathematical Programming*, to appear.

[21] A. F. Izmailov and M. V. Solodov. Mathematical programs with vanishing constraints: optimality conditions, sensitivity and a relaxation method. *Journal of Optimization Theory and Applications*, 142:501–532, 2009.

[22] R. Janin. *Directional derivative of the marginal function in nonlinear programming*, pages 110–126. Springer Berlin Heidelberg, 1984.

[23] S. Lu. Implications of the constant rank constraint qualification. *Mathematical Programming*, 126(2):365–392, May 2009.

[24] S. Lu. Relation between the constant rank and the relaxed constant rank constraint qualifications. *Optimization*, 2010. DOI: 10.1080/02331934.2010.527972.

[25] P. Malliavin. *Géométrie différentielle intrinsèque*. Hermann, 1972.

[26] O. L. Mangasarian. *Nonlinear programming*. SIAM, 1994.

[27] 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(1):37–47, 1967.

[28] J. M. Martínez. Inexact restoration method with Lagrangian tangent decrease and new merit function for nonlinear programming. *Journal of Optimization Theory and Applications*, 111:39–58, 2001.

[29] J. M. Martínez and E. A. Pilotta. Inexact restoration algorithms for constrained optimization. *Journal of Optimization Theory and Applications*, 104:135–163, 2000.

[30] J. M. Martínez and E. A. Pilotta. Inexact restoration methods for nonlinear programming: advances and perspectives. In L. Qi, K. Teo, and X. Yang, editors, *Optimization and Control with applications*, pages 271–292. Springer, 2005.

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

[32] L. Minchenko and S. Stakhovski. On relaxed constant rank regularity condition in mathematical programming. *Optimization*, 60(4):429–440, 2011.

[33] L. Minchenko and S. Stakhovski. Parametric nonlinear programming problems under the relaxed constant rank condition. *SIAM Journal on Optimization*, 1(314–332), 2011.

[34] L. Minchenko and A. Tarakanov. On error bounds for quasinormal programs. *Journal of Optimization Theory and Applications*, 148:571–579, 2011.

[35] J. Nocedal and S. J. Wright. *Numerical optimization.* Springer, New York, 2nd ed. edition, 2006.

[36] J.-S. Pang. Error bounds in mathematical programming. *Mathematical Programming*, 79(1-3):299–332, October 1997.

[37] E. R. Panier and A. L. Tits. On combining feasibility, descent and superlinear convergence in inequality constrained optimization. *Mathematical Programming*, 59(1-3):261–276, March 1993.

[38] L. Qi and A. Wei. On the Constant Positive Linear Dependence Condition and Its Application to SQP Methods. *SIAM Journal on Optimization*, 10(4):963, July 2000.

[39] J. R. Reay. Unique Minimal Representations with Positive Bases. *The American Mathematical Monthly*, 73(3):253–261, 1966.

[40] R. T. Rockafellar. Lagrange Multipliers and Optimality. *SIAM Review*, 35(2):183–238, 1993.

[41] M. V. Solodov. *Wiley Encyclopedia of Operations Research and Management Science*, chapter Constraint Qualifications. John Wiley & Sons, Inc., Hoboken, NJ, USA, Feb 2011.