Math. Program., Ser. A (2012) 135:255–273
DOI 10.1007/s10107-011-0456-0

FULL LENGTH PAPER

A relaxed constant positive linear dependence constraint qualification and applications

Roberto Andreani · Gabriel Haeser · María Laura Schuverdt · Paulo J. S. Silva

Received: 16 June 2010 / Accepted: 14 April 2011 / Published online: 6 May 2011

Abstract In this work we introduce a relaxed version of the constant positive linear dependence constraint qualification (CPLD) that we call RCPLD. This development is inspired by a recent generalization of the constant rank constraint qualification by Minchenko and Stakhovski that was called RCRCQ. We show that RCPLD is enough to ensure the convergence of an augmented Lagrangian algorithm and that it asserts the validity of an error bound. We also provide proofs and counter-examples that show the relations of RCRCQ and RCPLD with other known constraint qualifications. In

This work was supported by PRONEX-Optimization (PRONEX-CNPq/FAPERJ E-26/171.510/2006-APQ1), Fapesp (Grants 2006/53768-0 and 2009/09414-7) and CNPq (Grants 300900/2009-0, 303030/2007-0 and 474138/2008-9).

The second author would like to acknowledge that he was a research assistant at the University of São Paulo while this work was carried out.

R. Andreani
Department of Applied Mathematics, Institute of Mathematics,
Statistics and Scientific Computing, University of Campinas, Campinas, SP, Brazil
e-mail: andreani@ime.unicamp.br

G. Haeser
Institute of Science and Technology, Federal University of São Paulo, São José dos Campos, SP, Brazil
e-mail: gabriel.haeser@unifesp.br

M. L. Schuverdt
CONICET, Department of Mathematics, FCE, University of La Plata,
CP 172, 1900 La Plata, Bs. As., Argentina
e-mail: schuverd@mate.unlp.edu.ar

P. J. S. Silva (✉)
Institute of Mathematics and Statistics, University of São Paulo, São Paulo, SP, Brazil
e-mail: pjssilva@ime.usp.br

particular, RCPLD is strictly weaker than CPLD and RCRCQ, while still stronger than Abadie's constraint qualification. We also verify that the second order necessary optimality condition holds under RCRCQ.

Keywords Nonlinear programming · Constraint qualifications · Augmented Lagrangian · Error bound property

Mathematics Subject Classification (2000) 90C30 · 49K99 · 65K05

1 Introduction

In this paper, we consider the nonlinear programming problem

$$\text{Minimize } f(x), \quad \text{subject to} \quad x \in \Omega, \tag{1}$$

where $\Omega = \{x \in \mathbb{R}^n \mid h(x) = 0, g(x) \leq 0\}$, $f : \mathbb{R}^n \to \mathbb{R}$, $h : \mathbb{R}^n \to \mathbb{R}^m$ and $g : \mathbb{R}^n \to \mathbb{R}^p$ are continuously differentiable functions. For each feasible point $x \in \Omega$, we define the set of active inequality constraints $A(x) = \{j \mid g_j(x) = 0, j = 1, \ldots, p\}$.

We say that a *constraint qualification* holds at a feasible point $x \in \Omega$ if whenever x is a local minimum of (1) for a given objective function f, then the KKT condition holds. That is, there exist Lagrange multipliers $\lambda \in \mathbb{R}^m$ and $\mu_i \geq 0$ for every $i \in A(x)$ such that

$$\nabla f(x) + \sum_{i=1}^{m} \lambda_i \nabla h_i(x) + \sum_{i \in A(x)} \mu_i \nabla g_i(x) = 0.$$

Even though the definition above may suggest that the validity of a constraint qualification depends on the objective functions, this is not the case. Actually, a constraint qualification is a property of an analytical description of the feasible set that ensures that the first order conic approximation of the feasible set at x captures its geometrical structure. The presence of a constraint qualification is then fundamental to derive (analytical) characterizations of the solutions to optimization and variational problems, as well as properties related to duality and sensitivity. It is widely used in the development and analysis of computational methods. See, for example, the discussion in [24].

The most common constraint qualification is the linear independence constraint qualification (LICQ). It requires the linear independence of the gradients $\left(\{\nabla h_i(x)\}_{i=1}^m, \{\nabla g_i(x)\}_{i \in A(x)}\right)$. A weaker condition is the Mangasarian–Fromovitz constraint qualification (MFCQ), which requires only positive-linear independence[1] of such gradients [17,22].

Next, we define the constant rank constraint qualification of Janin (CRCQ, [15]), which is also weaker than LICQ.

[1] The pair of families $\left(\{v_i\}_{i=1}^m, \{v_i\}_{i=m+1}^p\right)$ is said to be positive-linearly dependent if $\{v_i\}_{i=1}^p$ is linearly dependent with non-negative scalars associated to the second family of vectors. Otherwise we say that the pair of families is positive-linearly independent.

Definition 1 (CRCQ) We say that the *constant rank constraint qualification (CRCQ)* holds at a feasible point $x \in \Omega$ if there exists a neighborhood $N(x)$ of x such that for every $I \subset \{1, \ldots, m\}$ and every $J \subset A(x)$, the family of gradients $(\{\nabla h_i(y)\}_{i \in I}, \{\nabla g_i(y)\}_{i \in J})$ has the same rank for every $y \in N(x)$.

In [18], Minchenko and Stakhovski provide a relaxed form of the CRCQ, which they called *relaxed constant rank constraint qualification* (RCRCQ). Instead of requiring that the rank of every subset of equality and active inequality gradients remains constant in a neighborhood of a feasible point, they only require constant rank of subsets consisting of *all* equality gradients and any subset of active inequality gradients. The authors proved that this is still a constraint qualification, and they used it to prove an error bound property.

It is well known that CRCQ can be equivalently stated as a constant linear dependence condition, that is: for every $I \subset \{1, \ldots, m\}$ and every $J \subset A(x)$, whenever $(\{\nabla h_i(x)\}_{i \in I}, \{\nabla g_i(x)\}_{i \in J})$ is linearly dependent, we must have $(\{\nabla h_i(y)\}_{i \in I}, \{\nabla g_i(y)\}_{i \in J})$ linearly dependent for every $y \in N(x)$, for some neighborhood $N(x)$ of x.

This motivates the definition of the constant positive linear dependence constraint qualification (CPLD, [6,21]), which is weaker than MFCQ and CRCQ.

Definition 2 (CPLD) We say that the *constant positive linear dependence constraint qualification (CPLD)* holds at a feasible point $x \in \Omega$ if there exists a neighborhood $N(x)$ of x such that for every $I \subset \{1, \ldots, m\}$ and every $J \subset A(x)$, whenever $(\{\nabla h_i(x)\}_{i \in I}, \{\nabla g_i(x)\}_{i \in J})$ is positive-linearly dependent, then $(\{\nabla h_i(y)\}_{i \in I}, \{\nabla g_i(y)\}_{i \in J})$ is linearly dependent for every $y \in N(x)$.

It can be proved (see [23]) that the CPLD condition can be equivalently stated at a feasible point $x \in \Omega$ as: for every subset $I \subset \{1, \ldots, m\}$ and $J \subset A(x)$, whenever $(\{\nabla h_i(x)\}_{i \in I}, \{\nabla g_i(x)\}_{i \in J})$ is positive-linearly dependent, we have that $(\{\nabla h_i(y)\}_{i \in I}, \{\nabla g_i(y)\}_{i \in J})$ is positive-linearly dependent for every y in some neighborhood of x. That is, requiring that the gradients are positive-linearly dependent in a neighborhood instead of the apparently weaker requirement of linear dependence, is in fact the same thing. This result guarantees that CPLD is stable in the sense that if a feasible point $x \in \Omega$ satisfies CPLD, then every feasible point in some neighborhood of x will also satisfy CPLD.

In this work we will introduce a relaxed version of the CPLD, that we call RCPLD. This relaxation will keep many properties of the CPLD such as the convergence of an augmented Lagrangian method [2,3,9], the existence of an error bound [16,18], and stability. We provide proofs and counter-examples that give a complete picture of the relationship of RCRCQ and RCPLD with other well known constraint qualifications.

We will use the following notation:

- $\| \cdot \| = \| \cdot \|_2$,
- $|J|$ denotes the number of elements of the finite set J,
- $\operatorname{span}\{v_i\}_{i=1}^m$ denotes the subspace generated by the vectors $v_1, \ldots, v_m$.

2 Relaxed constant rank constraint qualification

We study the relaxed constant rank constraint qualification of Minchenko and Stakhovski (RCRCQ, [18]).

Definition 3 (RCRCQ) We say that the *relaxed constant rank constraint qualification (RCRCQ)* holds at a feasible point $x \in \Omega$ if there exists a neighborhood $N(x)$ of x such that for every $J \subset A(x)$, the family of gradients $\left(\{\nabla h_i(y)\}_{i=1}^m, \{\nabla g_i(y)\}_{i \in J}\right)$ has the same rank for every $y \in N(x)$.

Compared to original CRCQ this relaxation treats the set of equality constraints as a whole, without the need to impose restrictions on all their subsets. In [18], the authors proved that RCRCQ is still a constraint qualification, by showing that it implies Abadie's constraint qualification [1]. They also showed that RCRCQ is strictly weaker than CRCQ. In the case of only equality constraints, this condition was independently formulated in [4]. The RCRCQ has also been studied in the context of parametric problems in [16].

Since CRCQ is equivalent to the fact that for every subset of equality and active inequality gradients, linearly dependent vectors remain linearly dependent on some neighborhood, one could conjecture that a similar equivalence holds true for RCRCQ, considering only subsets that contain every equality gradient. But this is not the case. Consider the equality constraints $h_1(x_1, x_2) = x_1$, $h_2(x_1, x_2) = x_1$ and the inequality constraint $g_1(x_1, x_2) = x_2^2$ at the feasible point $x = (0, 0)$. RCRCQ does not hold since the gradients $\{\nabla h_1(y), \nabla h_2(y), \nabla g_1(y)\}$ have rank one at $y = x$ and rank two for y arbitrarily close to x with $y_2 \neq 0$, but subsets that contain both equality gradients are linearly dependent on every neighborhood of x.

We will provide a reformulation of RCRCQ in terms of constant linear dependence. We must keep the condition that the rank of the equality constraint gradients $\{\nabla h_i(y)\}_{i=1}^m$ is constant for every y in some neighborhood $N(x)$ of x. The key point is that in this situation, we may choose a subset $I \subset \{1, \ldots, m\}$ such that $\{\nabla h_i(x)\}_{i \in I}$ is a basis for $\operatorname{span}\{\nabla h_i(x)\}_{i=1}^m$, thus, since linearly independent vectors remain linearly independent in a neighborhood, and the rank is the same, we have that $\{\nabla h_i(y)\}_{i \in I}$ is a basis for $\operatorname{span}\{\nabla h_i(y)\}_{i=1}^m$ for every y in some neighborhood $N(x)$ of x. The reformulation requires that for every $J \subset A(x)$, linear dependence is maintained in a neighborhood of x whenever $(\{\nabla h_i(x)\}_{i \in I}, \{\nabla g_i(x)\}_{i \in J})$ is linearly dependent. Notice that when $(\{\nabla h_i(x)\}_{i \in I}, \{\nabla g_i(x)\}_{i \in J})$ is linearly dependent, there must exist an index $j \in J$ such that $\nabla g_j(x)$ is a linear combination of the remaining gradients, otherwise this would contradict the linear independence of $\{\nabla h_i(x)\}_{i \in I}$.

Theorem 1 *Let $I \subset \{1, \ldots, m\}$ be an index set such that $\{\nabla h_i(x)\}_{i \in I}$ is a basis for* $\operatorname{span}\{\nabla h_i(x)\}_{i=1}^m$. *A feasible point $x \in \Omega$ satisfies RCRCQ if, and only if, there exists a neighborhood $N(x)$ of x such that*

- $\{\nabla h_i(y)\}_{i=1}^m$ *has the same rank for every $y \in N(x)$,*
- *For every $J \subset A(x)$, if $(\{\nabla h_i(x)\}_{i \in I}, \{\nabla g_i(x)\}_{i \in J})$ is linearly dependent, then $(\{\nabla h_i(y)\}_{i \in I}, \{\nabla g_i(y)\}_{i \in J})$ is linearly dependent for every $y \in N(x)$.*

Proof Let $x \in \Omega$ satisfy RCRCQ. The first claim follows by taking $J = \emptyset$ in the definition of RCRCQ. Let $J \subset A(x)$ be such that $(\{\nabla h_i(x)\}_{i\in I}, \{\nabla g_i(x)\}_{i\in J})$ is linearly dependent. Since the gradients corresponding to the set I generate the remaining equality constraint gradients in a neighborhood, using RCRCQ we have that the rank of $(\{\nabla h_i(y)\}_{i\in I}, \{\nabla g_i(y)\}_{i\in J})$ is constant for every y in some neighborhood $N(x)$ of x, therefore, this set must be linearly dependent for $y \in N(x)$.

To prove the converse let $J \subset A(x)$. Choose $\hat{J} \subset J$ in such a way that $\left(\{\nabla h_i(x)\}_{i\in I}, \{\nabla g_i(x)\}_{i\in \hat{J}}\right)$ is a basis for $(\{\nabla h_i(x)\}_{i\in I}, \{\nabla g_i(x)\}_{i\in J})$. The case $\hat{J} = J$ is trivial. Now let $j \in J\backslash\hat{J}$. As $\left(\{\nabla h_i(x)\}_{i\in I}, \{\nabla g_i(x)\}_{i\in \hat{J}\cup\{j\}}\right)$ is linearly dependent, it must remain linearly dependent in $N(x)$. Hence the rank of $(\{\nabla h_i(y)\}_{i\in I}, \{\nabla g_i(y)\}_{i\in J})$ is not greater than $|I| + |\hat{J}|$. The result now follows from the fact that the rank cannot decrease in a neighborhood. □

Next we provide counter-examples to show where RCRCQ fits among other well known constraint qualifications. The following counter-example shows that MFCQ does not imply RCRCQ.

Counter-example 1 Consider the inequality constraints $g_1(x_1, x_2) = -x_2$ and $g_2(x_1, x_2) = x_1^2 - x_2$ at the feasible point $x = (0, 0)$. Clearly, MFCQ holds. RCRCQ does not hold since $\{\nabla g_1(y), \nabla g_2(y)\}$ has rank one at $y = x$ and rank two for y arbitrarily close to x with $y_1 \neq 0$.

We say that quasinormality (see [8,14]) holds at a feasible point $x \in \Omega$ if whenever $\sum_{i=1}^m \lambda_i \nabla h_i(x) + \sum_{i\in A(x)} \mu_i \nabla g_i(x) = 0$ for some $\lambda \in \mathbb{R}^m$ and $\mu_i \geq 0$ for every i, there is no sequence $y^k \to x$ such that for every k, $\lambda_i h_i(y^k) > 0$ when $\lambda_i \neq 0$ and $g_i(y^k) > 0$ when $\mu_i > 0$. The following counter-example shows that RCRCQ does not imply the quasinormality constraint qualification.

Counter-example 2 Consider the equality constraint $h_1(x_1, x_2) = -(x_1+1)^2 - x_2^2 + 1$ and the inequality constraints $g_1(x_1, x_2) = x_1^2 + (x_2 + 1)^2 - 1$, $g_2(x_1, x_2) = -x_2$, at the feasible point $x = (0, 0)$. Quasinormality does not hold, since we can write $\nabla g_1(x) + 2\nabla g_2(x) = 0$ and by taking $y^k = \left(\sqrt{1 - (1 - \frac{1}{k})^2 + \frac{1}{k}}, -\frac{1}{k}\right)$ we have $g_1(y^k) > 0$ and $g_2(y^k) > 0$ for every k. RCRCQ holds since there is a neighborhood $N(x)$ of x such that for every $y \in N(x)$, $\{\nabla h_1(y)\}$ has rank one and $\{\nabla h_1(y), \nabla g_1(y)\}$, $\{\nabla h_1(y), \nabla g_2(y)\}$, $\{\nabla h_1(y), \nabla g_1(y), \nabla g_2(y)\}$ have rank two.

In Fig. 1 we show relations of RCRCQ with other well known constraint qualifications, where pseudonormality appears in [8]. It holds at a feasible point $x \in \Omega$ if whenever $\sum_{i=1}^m \lambda_i \nabla h_i(x) + \sum_{i\in A(x)} \mu_i \nabla g_i(x) = 0$ for some $\lambda \in \mathbb{R}^m$ and $\mu_i \geq 0$ for every i, then there is no sequence $y^k \to x$ such that $\sum_{i=1}^m \lambda_i h_i(y^k) + \sum_{i\in A(x)} \mu_i g_i(y^k) > 0$ for every k. The proof that RCRCQ implies Abadie's constraint qualification [1] has been done in [18].

If the problem data is twice continuously differentiable there is the notion of second order necessary optimality conditions. Such conditions are usually associated to special constraint qualifications as their definitions assume the existence of special

Lagrange multipliers. As it was mentioned in [4], a second order optimality condition can be weak or strong depending on the tangent cone used to analyze the curvature of the Lagrangian function.

We say that the *second order necessary optimality condition* holds at a feasible point $x \in \Omega$ when there exist Lagrange multipliers $\lambda \in \mathbb{R}^m$, $\mu_i \geq 0 \quad \forall i \in A(x)$, satisfying the KKT condition for which

$$d^T\left(\nabla^2 f(x) + \sum_{i=1}^{m} \lambda_i \nabla^2 h_i(x) + \sum_{i=1}^{p} \mu_i \nabla^2 g_i(x)\right) d \geq 0, \tag{2}$$

for all directions $d \in \mathbb{R}^n$ in the critical cone:

$$\begin{aligned} V_1(x) = \{d \in \mathbb{R}^n : \nabla h_i(x)^T d = 0, i = 1, \ldots, m, \\ \nabla g_j(x)^T d = 0,\ j \in A^+(x), \\ \nabla g_j(x)^T d \leq 0,\ j \in A^0(x)\} \end{aligned}$$

where

$$A^+(x) = \{j \in A(x) : \mu_j > 0\},\ A^0(x) = \{j \in A(x) : \mu_j = 0\}.$$

Analogously, we say that the *weak second order necessary optimality condition* holds if there exists a Lagrange multiplier vector such that (2) holds for all directions $d \in \mathbb{R}^n$ in the following smaller cone:

$$V_2(x) = \{d \in \mathbb{R}^n : \nabla h_i(x)^T d = 0, i = 1, \ldots, m, \nabla g_j(x)^T d = 0,\ j \in A(x)\}.$$

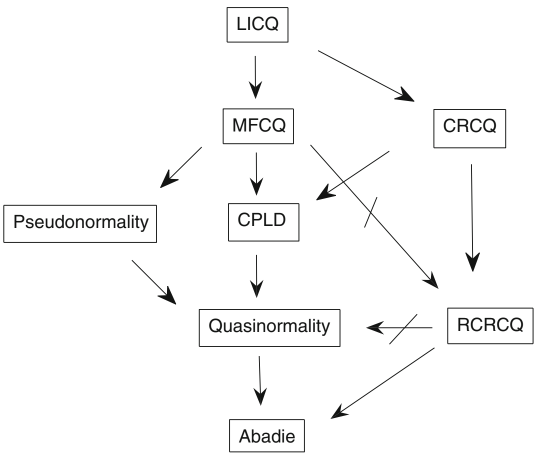

Fig. 1 Complete diagram showing relations of RCRCQ with other well known constraint qualifications, where an arrow between two constraint qualifications means that one is strictly stronger than the other

In [4], the authors proved that CRCQ is enough to ensure that any local solution not only verifies KKT but also verifies the second order necessary condition. Moreover, the counter-example defined in [7] shows that MFCQ is not even enough to ensure the weak second order necessary condition. Considering Fig. 1, it is still possible that RCRCQ can assert one of the second order necessary conditions above. Actually, it is easy to see, from the Remark 3.2 in [4], that the RCRCQ implies the second order necessary condition. In fact, it was shown in [4] that under RCRCQ, if x is a local solution of (1) then (2) holds for all $d \in V_1(x)$ and for every Lagrange multiplier vector.

3 Relaxed constant positive linear dependence constraint qualification

In [21], Qi and Wei proposed a relaxation of CRCQ, the constant positive linear dependence condition (CPLD), taking in consideration the positive sign of the multipliers associated to inequality constraints in the KKT condition. They used this condition to prove convergence of a sequential quadratic programming method.

In [6], it has been proved that CPLD is in fact a constraint qualification, and in [2,3], the authors proved convergence of an augmented Lagrangian method under CPLD. We now propose a relaxation of CPLD in a way similar to RCRCQ that we call *relaxed CPLD* (RCPLD). The definition is motivated by Theorem 1, considering only positive-linearly dependent gradients, as in the definition of CPLD.

Definition 4 (RCPLD) Let $I \subset \{1, \dots, m\}$ be such that $\{\nabla h_i(x)\}_{i \in I}$ is a basis for $\operatorname{span}\{\nabla h_i(x)\}_{i=1}^m$. We say that the *relaxed constant positive linear dependence constraint qualification (RCPLD)* holds at a feasible point $x \in \Omega$ if there exists a neighborhood $N(x)$ of x such that

- $\{\nabla h_i(y)\}_{i=1}^m$ has the same rank for every $y \in N(x)$.
- For every $J \subset A(x)$, if $(\{\nabla h_i(x)\}_{i \in I}, \{\nabla g_i(x)\}_{i \in J})$ is positive-linearly dependent, then $(\{\nabla h_i(y)\}_{i \in I}, \{\nabla g_i(y)\}_{i \in J})$ is linearly dependent for every $y \in N(x)$.

Observe that the definition of the RCPLD does not depend on the specific choice of the index set that selects the basis for $\operatorname{span}\{\nabla h_i(x)\}_{i=1}^m$. That is, if $I, I' \subset \{1, \dots, m\}$ are index sets of two bases for this space, then the assumption that $\{\nabla h_i(y)\}_{i=1}^m$ has the same rank for all $y \in N(x)$ implies that $(\{\nabla h_i(y)\}_{i \in I'}, \{\nabla g_i(y)\}_{i \in J})$ is (positive-) linearly dependent if, and only if, $(\{\nabla h_i(y)\}_{i \in I}, \{\nabla g_i(y)\}_{i \in J})$ has the same property.

Clearly, Theorem 1 shows that RCRCQ implies RCPLD. To prove that CPLD implies RCPLD we only need to show that the rank of the equality constraint gradients is constant in a neighborhood. This follows from the fact that constant linear dependence for every subset of gradients is equivalent to constant rank of all such sets.

An important tool to deal with positive-linearly dependent vectors (in particular, to deal with CPLD or RCPLD) is Carathéodory's Lemma [8, Exercise B.1.7]. We will state here a similar result that will be suitable to study the RCPLD. This result can be seen as a corollary of Carathéodory's Lemma, but we include a full proof for completeness.

Lemma 1 *If $x = \sum_{i=1}^{m+p} \alpha_i v_i$ with $v_i \in \mathbb{R}^n$ for every i, $\{v_i\}_{i=1}^m$ linearly independent and $\alpha_i \neq 0$ for every $i = m+1, \ldots, m+p$, then there exist $J \subset \{m+1, \ldots, m+p\}$ and scalars $\bar{\alpha}_i$ for every $i \in \{1, \ldots, m\} \cup J$ such that*

- $x = \sum_{i \in \{1,\ldots,m\} \cup J} \bar{\alpha}_i v_i$,
- $\alpha_i \bar{\alpha}_i > 0$ *for every* $i \in J$,
- $\{v_i\}_{i \in \{1,\ldots,m\} \cup J}$ *is linearly independent.*

Proof We assume that $\{v_i\}_{i=1}^{m+p}$ is linearly dependent, otherwise the result follows trivially. Then, there exists $\beta \in \mathbb{R}^{m+p}$, such that $\sum_{i=m+1}^{m+p} |\beta_i| > 0$ and $\sum_{i=1}^{m+p} \beta_i v_i = 0$. Thus, we may write $x = \sum_{i=1}^{m+p} (\alpha_i - \gamma \beta_i) v_i$, for every $\gamma \in \mathbb{R}$. Choosing $\gamma \neq 0$ as the number of smallest modulus such that $\alpha_i - \gamma \beta_i = 0$ for at least one index $i \in \{m+1, \ldots, m+p\}$, we are able to write the linear combination x with at least one less vector v_i, for some $i \in \{m+1, \ldots, m+p\}$. We may repeat this procedure until the vectors are linearly independent. □

We point out that we can obtain bounds $|\bar{\alpha}_i| \leq 2^{p-1} |\alpha_i|, \forall i = m+1, \ldots, m+p$ in the same way it is done in [12]. This may be useful, in particular, for applications to interior point methods.

We now prove that RCPLD is a constraint qualification. We will need a definition from [5]:

Definition 5 (AKKT) We say that $x \in \Omega$ satisfies the *Approximate-KKT condition (AKKT)* if there exist sequences $x^k \to x$, $\{\lambda^k\} \subset \mathbb{R}^m$, $\{\mu^k\} \subset \mathbb{R}^p$, $\mu^k \geq 0$ such that

$$\nabla f(x^k) + \sum_{i=1}^{m} \lambda_i^k \nabla h_i(x^k) + \sum_{i \in A(x)} \mu_i^k \nabla g_i(x^k) \to 0.$$

Observe that small, but equivalent, variations of this definition have appeared in the literature before, for example in [21] or even implicitly in [8] in the proof that each local minimizer of (1) is a Fritz-John point. However, we will use the version described above as it is better suited for our purposes.

Note that the definition of AKKT also depends on the objective function f, thus, it is a property of the optimization problem, rather than only of the constraint set. In Theorem 2.3 of [5] (with $I = \emptyset$), the authors proved that every local minimizer fulfills the AKKT condition (a simpler proof, specific for the case $I = \emptyset$, can be found in [13]). To prove that RCPLD is a constraint qualification, we need only to show that if RCPLD holds at a feasible point x such that AKKT also holds, then x is a KKT point.

Theorem 2 *Let $x \in \Omega$ be such that RCPLD and AKKT hold. Then x is a KKT point.*

Proof From the definition of AKKT, there exist sequences $\varepsilon_k \to 0$, $x^k \to x$, $\lambda^k \in \mathbb{R}^m$, $\mu_j^k \geq 0$, $\forall j \in A(x)$, such that

$$\nabla f(x^k) + \sum_{i=1}^{m} \lambda_i^k \nabla h_i(x^k) + \sum_{j \in A(x)} \mu_j^k \nabla g_j(x^k) = \varepsilon_k, \quad \text{for every } k.$$

Consider an index subset $I \subset \{1,\dots,m\}$ such that $\{\nabla h_i(x)\}_{i\in I}$ is a basis for $\text{span}\{\nabla h_i(x)\}_{i=1}^m$. Then $\{\nabla h_i(x^k)\}_{i\in I}$ is linearly independent for sufficiently large k. Since the rank of equality constraint gradients is constant, we have that $\{\nabla h_i(x^k)\}_{i\in I}$ is a basis for $\text{span}\{\nabla h_i(x^k)\}_{i=1}^m$ for sufficiently large k. Thus, there exists a sequence $\{\bar{\lambda}^k\} \subset \mathbb{R}^{|I|}$ such that $\sum_{i=1}^m \lambda_i^k \nabla h_i(x^k) = \sum_{i\in I} \bar{\lambda}_i^k \nabla h_i(x^k)$, and we may write

$$\nabla f(x^k) + \sum_{i\in I} \bar{\lambda}_i^k \nabla h_i(x^k) + \sum_{j\in A(x)} \mu_j^k \nabla g_j(x^k) = \varepsilon_k.$$

We apply Lemma 1 to obtain subsets $J_k \subset A(x)$ and multipliers $\tilde{\lambda}^k \in \mathbb{R}^{|I|}$ and $\tilde{\mu}_j^k \geq 0, \forall j \in J_k$ such that

$$\nabla f(x^k) + \sum_{i\in I} \tilde{\lambda}_i^k \nabla h_i(x^k) + \sum_{j\in J_k} \tilde{\mu}_j^k \nabla g_j(x^k) = \varepsilon_k,$$

and $\left(\{\nabla h_i(x^k)\}_{i\in I}, \{\nabla g_i(x^k)\}_{i\in J_k}\right)$ is linearly independent. We will consider a subsequence such that J_k is the same set J for every k (this can be done since there are finitely many possible sets J_k). Define $M_k = \max\{|\tilde{\lambda}_i^k|, \forall i \in I, \tilde{\mu}_j^k, \forall j \in J\}$. If there is a subsequence such that $M_k \to +\infty$, we may take a subsequence such that $\frac{\left(\tilde{\lambda}^k, \tilde{\mu}^k\right)}{M_k} \to (\lambda, \mu) \neq 0, \mu \geq 0$. Dividing by M_k and taking limits we have

$$\sum_{i\in I} \lambda_i \nabla h_i(x) + \sum_{j\in J} \mu_j \nabla g_j(x) = 0,$$

which contradicts RCPLD. Hence, we have that $\{M_k\}$ is a bounded sequence. Taking limits for a suitable subsequence such that $\lambda^k \to \lambda$ and $\mu^k \to \mu \geq 0$ we have

$$\nabla f(x) + \sum_{i=1}^m \lambda_i \nabla h_i(x) + \sum_{j\in J} \mu_j \nabla g_j(x) = 0,$$

which proves that x is a KKT point. □

Corollary 1 *RCPLD is a constraint qualification.*

Theorem 2 is also important to ensure the convergence of an augmented Lagrangian algorithm as we discuss in the next section.

Given a new constraint qualification, it is important to know its relation with other well known constraint qualifications. In particular, we would like to know if RCPLD can still guarantee that the tangent cone is polyhedral. In the following theorem we prove that this is the case by showing that RCPLD implies Abadie's constraint qualification.

Let us consider the feasible set Ω and $x \in \Omega$. We define the (upper) tangent cone of Ω at x as (see for example [8,11,14,24]):

$$T_\Omega(x) = \{0\} \cup \left\{ d \in \mathbb{R}^n : \exists \{x_k\} \subset \Omega, x_k \neq x, \ x_k \to x \text{ and } \frac{x_k - x}{\|x_k - x\|} \to \frac{d}{\|d\|} \right\}. \tag{3}$$

We define also the linearization cone at x as:

$$V_\Omega(x) = \{d \in \mathbb{R}^n : \nabla h_i(x)^T d = 0, i = 1, \ldots, m; \nabla g_j(x)^T d \leq 0, j \in A(x)\}. \tag{4}$$

We say that Abadie's constraint qualification [1] holds at a feasible point $x \in \Omega$ if $T_\Omega(x) = V_\Omega(x)$.

Theorem 3 *Let $x \in \Omega$ be such that RCPLD holds. Then x satisfies Abadie's constraint qualification.*

Proof The inclusion $T_\Omega(x) \subset V_\Omega(x)$ holds without any constraint qualification for every feasible point $x \in \Omega$.

The proof that $V_\Omega(x) \subset T_\Omega(x)$ analyzes first a simplified variation of the feasible set considered in [8,14]. Let us define the set of indexes

$$\hat{J} = \{i \in A(x) : \nabla g_i(x)^T d = 0, \quad \forall d \in V_\Omega(x)\}.$$

Now, define

$$\hat{X} = \{x \in \mathbb{R}^n : h_i(x) = 0, i = 1, \ldots, m; g_j(x) \leq 0, j \in \hat{J}\}.$$

In the degenerate case, where there are no equalities and the set $\hat{J}$ is empty, we have $\hat{X} = \mathbb{R}^n$ by convention. In this case, every point of $\hat{X}$ verifies RCPLD and Abadie's CQ.

By the definition of RCPLD, if a feasible point $x \in \Omega$ verifies the RCPLD then it verifies the RCPLD as a point in $\hat{X}$. Using this simplified set, let us prove first that x verifies Abadie's CQ as a point in $\hat{X}$.

We have that $T_{\hat{X}}(x) \subset V_{\hat{X}}(x)$ always holds. Let us take a direction $d \in V_{\hat{X}}(x)$. Let $\varepsilon > 0$, $k > 0$ and let $y(t,k)$ be the minimizer of the function

$$H(y,t,k) = \|y - x - td\|^2 + tk \left(\sum_{i=1}^{m} h_i(y)^2 + \sum_{i \in \hat{J}} \max\{0, g_i(y)\}^2 \right)$$

subject to $\|y - x\| \leq \varepsilon$.

We have that, for $t \geq 0$,

$$\|y(t,k) - x - td\|^2 \leq H(y(t,k),t,k) \leq H(x,t,k) = t^2\|d\|^2, \tag{5}$$

and analogously

$$0 \leq k \left(\sum_{i=1}^{m} h_i(y(t,k))^2 + \sum_{i \in \hat{J}} \max\{0, g_i(y(t,k))\}^2 \right) \leq t\|d\|^2. \tag{6}$$

By (5) we have

$$\|y(t,k) - x\| \leq 2t\|d\|. \tag{7}$$

Thus, for each $t > 0$, we have that the sequence $\{y(t,k)\}_k$ is a bounded sequence and there exists $y(t)$, $t > 0$ such that, taking a subsequence if necessary, we have

$$y(t,k) \to y(t). \tag{8}$$

Then, taking limits in (6) and by continuity:

$$0 \leq \left(\sum_{i=1}^{m} h_i(y(t))^2 + \sum_{i \in \hat{J}} \max\{0, g_i(y(t))\}^2 \right) \leq \lim_{k \to \infty} \frac{t}{k}\|d\|^2 = 0.$$

This implies that $y(t) \in \hat{X}$ for all $t > 0$.

From (7) we have that $y(t) \to x$ as $t \to 0$. Moreover, we can select a sequence of positive numbers t_r with $t_r \to 0$ such that the limit

$$d_0 = \lim_{r \to \infty} \frac{y(t_r) - x}{t_r}$$

exists. Since $y(t) \in \hat{X}$ for all $t > 0$ we obtain that $d_0 \in T_{\hat{X}}(x) \subset V_{\hat{X}}(x)$.

Let us consider r_0 large enough such that, by (7), $\|y(t_r,k) - x\| < \varepsilon, \forall r \geq r_0, \forall k$. By the definition of $y(t,k)$ we have that, for $r \geq r_0$, $\nabla_y H(y(t_r,k), t_r, k) = 0$, then

$$\frac{y(t_r,k) - x - t_r d}{t_r} + k \left(\sum_{i=1}^{m} h_i(y(t_r,k)) \nabla h_i(y(t_r,k)) + \sum_{i \in \hat{J}} \max\{0, g_i(y(t_r,k))\} \nabla g_i(y(t_r,k)) \right) = 0.$$

By the definition of RCPLD, we may take a subset $I \subset \{1, \ldots, m\}$ such that $\{\nabla h_i(y(t_r,k))\}_{i \in I}$ is a basis for $\mathrm{span}\{\nabla h_i(y(t_r,k))\}_{i=1}^m$ for sufficiently large k. Thus, there exists a sequence $\{\lambda^k(r)\} \subset \mathbb{R}^m$ such that $\sum_{i=1}^m k h_i(y(t_r,k)) \nabla h_i(y(t_r,k)) = \sum_{i \in I} \lambda_i^k(r) \nabla h_i(y(t_r,k))$. By applying Lemma 1, we have that there are subsets $J_k(r) \subset \hat{J}$ and multipliers $\bar{\lambda}_i^k(r), \forall i \in I$ and $\bar{\mu}_i^k(r) \geq 0, \forall i \in J_k(r)$ such that

$$\frac{y(t_r,k) - x - t_r d}{t_r} + \sum_{i \in I} \bar{\lambda}_i^k(r) \nabla h_i(y(t_r,k)) + \sum_{i \in J_k(r)} \bar{\mu}_i^k(r) \nabla g_i(y(t_r,k)) = 0 \tag{9}$$

and

$$\left(\{\nabla h_i(y(t_r,k))\}_{i\in I}, \{\nabla g_i(y(t_r,k))\}_{i\in J_k(r)}\right) \text{ is linearly independent.} \tag{10}$$

We will consider a subsequence such that $J_k(r)$ is the same set $J(r)$ (this can be done since there are finitely many possible sets $J_k(r)$).

Denote,

$$M_k(r) = \sqrt{1 + \sum_{i\in I}(\bar{\lambda}_i^k(r))^2 + \sum_{i\in J(r)}(\bar{\mu}_i^k(r))^2}.$$

Then, dividing (9) by $M_k(r)$ and taking limit when $k \to \infty$ for k in an appropriate subsequence we have that there are scalars $\mu_0(r), \lambda_i(r), i \in I, \mu_j(r), j \in J(r), \mu_j(r) \geq 0$ not all equal to zero such that

$$\mu_0(r)\left(\frac{y(t_r)-x}{t_r} - d\right) + \sum_{i\in I}\lambda_i(r)\nabla h_i(y(t_r)) + \sum_{i\in J(r)}\mu_i(r)\nabla g_i(y(t_r)) = 0, \quad \text{for} \quad t_r > 0. \tag{11}$$

Let us consider a subsequence such that $J(r)$ is the same set J. Using that $\|(\mu_0(r), \lambda_i(r), \mu_i(r))\| = 1$ and taking limit when $r \to \infty$ for r in an appropriate subsequence in (11) we have that there are scalars $\mu_0, \lambda_i, i \in I, \mu_j, j \in J, \mu_j \geq 0$ not all equal to zero such that

$$\mu_0(d_0 - d) + \sum_{i\in I}\lambda_i\nabla h_i(x) + \sum_{i\in J}\mu_i\nabla g_i(x) = 0. \tag{12}$$

If $\mu_0 = 0$, then we have that $(\{\nabla h_i(x)\}_{i\in I}, \{\nabla g_i(x)\}_{i\in J})$ is positive-linearly dependent, hence, since (10) holds, this contradicts RCPLD. Consequently, it must be $\mu_0 > 0$. Given $\hat{d} \in V_{\hat{X}}(x)$, let us prove that $\nabla g_j(x)^T\hat{d} = 0$ for every $j \in \hat{J}$. From the definition of $V_{\hat{X}}(x)$ we have that $\nabla g_j(x)^T\hat{d} \leq 0$ for every $j \in \hat{J}$, and from the definition of $\hat{J}$, for every $i \in A(x)\backslash\hat{J}$ there exists $d_i \in V_\Omega(x)$ such that $\nabla g_i(x)^T d_i < 0$. Defining $\bar{d} = \sum_{i\in A(x)\backslash\hat{J}} d_i$, we have that

$$\nabla g_i(x)^T\bar{d} < 0 \text{ for every } i \in A(x)\backslash\hat{J},$$

and

$$\nabla g_j(x)^T\bar{d} = 0 \text{ for every } j \in \hat{J}.$$

Thus, for sufficiently large $\alpha > 0$ we have

$$\nabla g_i(x)^T(\hat{d} + \alpha\bar{d}) = \nabla g_i(x)^T\hat{d} + \alpha\nabla g_i(x)^T\bar{d} < 0 \quad \text{for every } i \in A(x)\backslash\hat{J},$$

and

$$\nabla g_j(x)^T(\hat{d} + \alpha \bar{d}) = \nabla g_j(x)^T \hat{d} \le 0 \quad \text{for every } j \in \hat{J}.$$

Hence $\hat{d} + \alpha \bar{d} \in V_\Omega(x)$, which implies that for all $j \in \hat{J}$, $\nabla g_j(x)^T(\hat{d} + \alpha \bar{d}) = \nabla g_j(x)^T \hat{d} = 0$ as we wanted to prove. Thus, multiplying (12) by $d, d_0 \in V_{\hat{X}}(x)$, we obtain that

$$d_0^T(d - d_0) = 0 = d^T(d - d_0),$$

which implies that $d = d_0 \in T_{\hat{X}}(x)$.

We have proved that a feasible point x that verifies RCPLD as a point in $\hat{X}$ verifies Abadie's as a point in $\hat{X}$.

Now we have to prove that this implies that x verifies Abadie's CQ as a point in Ω. Let us define the set $\tilde{V}_\Omega(x) = \{d \in \mathbb{R}^n : \nabla h_i(x)^T d = 0, i = 1, \ldots, m; \nabla g_j(x)^T d = 0, j \in \hat{J}; \nabla g_j(x)^T d < 0, j \in A(x) \backslash \hat{J}\}$.

Since x verifies that $T_{\hat{X}}(x) = V_{\hat{X}}(x)$ and $\tilde{V}_\Omega(x) \subset V_{\hat{X}}(x)$ it is not hard to prove that $\tilde{V}_\Omega(x) \subset T_\Omega(x)$. In general, $T_\Omega(x)$ is a closed cone, thus, $V_\Omega(x) = \mathrm{cl}(\tilde{V}_\Omega(x)) \subset T_\Omega(x)$, where $\mathrm{cl}(\cdot)$ denotes the closure operator. Thus, we have that x verifies Abadie's CQ as a point in Ω as we wanted to prove. □

Note that Corollary 1 is also an immediate consequence of the above result. However, we derived it after Theorem 2 to emphasize the alternative technique to prove that a constraint qualification holds by showing that AKKT together with the candidate condition implies KKT. This strategy is specially appealing here as it is essential to derive the convergence of the augmented Lagrangian algorithm as we do in Sect. 4.

We will now provide some counter-examples to completely state the relation of RCPLD with respect to other known constraint qualifications. We observe that since MFCQ implies RCPLD, Counter-example 1 shows that RCPLD does not imply RCRCQ.

The following counter-example shows that pseudonormality does not imply RCPLD.

Counter-example 3 Consider the inequality constraints $g_1(x_1, x_2) = -x_1$ and $g_2(x_1, x_2) = x_1 - x_1^2 x_2^2$, at the feasible point $x = (0, 0)$. RCPLD does not hold since $(\emptyset, \{\nabla g_1(y), \nabla g_2(y)\})$ is positive-linearly dependent at $y = x$ but linearly independent for y arbitrarily close to x. Pseudonormality holds, since we can write $\mu \nabla g_1(x) + \mu \nabla g_2(x) = 0$ for every $\mu > 0$, but $\mu g_1(y_1, y_2) + \mu g_2(y_1, y_2) = -\mu y_1^2 y_2^2 \le 0$ for every y.

Since RCRCQ does not imply quasinormality and RCRCQ implies RCPLD, we have that RCPLD does not imply quasinormality. In Fig. 2 we show a complete diagram picturing the relations of RCPLD with other constraint qualifications.

Observe that since MFCQ implies RCPLD and the example from [7] shows that MFCQ does not imply the weak second order necessary condition, it follows immediately that RCPLD does not imply such condition either.

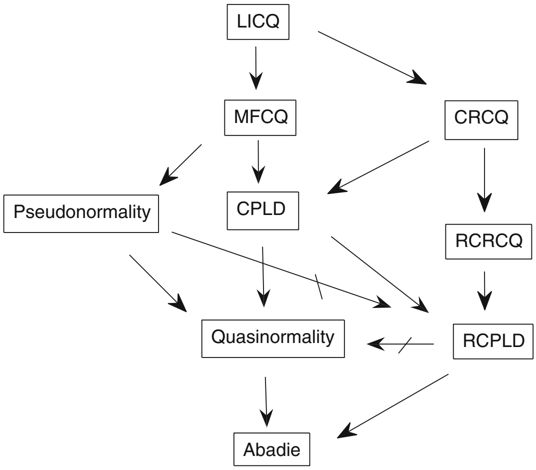

Fig. 2 Complete diagram showing relations of RCPLD with other well known constraint qualifications, where an arrow between two constraint qualifications means that one is strictly stronger than the other

We finish this section by proving that RCPLD can be equivalently stated requiring only positive-linear dependence in a neighborhood. This proves that RCPLD is stable in the sense that if it holds at a given feasible point, then it must hold at every feasible point of some neighborhood.

Theorem 4 *Let $I \subset \{1, \ldots, m\}$ be an index set such that $\{\nabla h_i(x)\}_{i \in I}$ is a basis for* $\text{span}\{\nabla h_i(x)\}_{i=1}^m$. *A feasible point $x \in \Omega$ satisfies RCPLD if, and only if, there exists a neighborhood $N(x)$ of x such that*

- $\{\nabla h_i(y)\}_{i=1}^m$ *has the same rank for every $y \in N(x)$,*
- *For every $J \subset A(x)$, if $(\{\nabla h_i(x)\}_{i \in I}, \{\nabla g_i(x)\}_{i \in J})$ is positive-linearly dependent, then $(\{\nabla h_i(y)\}_{i \in I}, \{\nabla g_i(y)\}_{i \in J})$ is positive-linearly dependent for every $y \in N(x)$.*

Proof Let us take a feasible point $x \in \Omega$ that satisfies RCPLD and $J \subset A(x)$ such that $(\{\nabla h_i(x)\}_{i \in I}, \{\nabla g_i(x)\}_{i \in J})$ is positive-linearly dependent. Thus, there are $\lambda_i \in \mathbb{R}, \forall i \in I, \mu_i \geq 0, \forall i \in J, \sum_{i \in J} \mu_i > 0$ such that $\sum_{i \in I} \lambda_i \nabla h_i(x) + \sum_{i \in J} \mu_i \nabla g_i(x) = 0$. We can assume that $\mu_i > 0$ for every $i \in J$. Since $J \neq \emptyset$, taking $j \in J$ we may write $\mu_j \nabla g_j(x) = \sum_{i \in I} -\lambda_i \nabla h_i(x) + \sum_{i \in J \setminus \{j\}} -\mu_i \nabla g_i(x)$. By Lemma 1, there exist $J' \subset J \setminus \{j\}$ and $\bar{\lambda}_i \in \mathbb{R}, \forall i \in I, \bar{\mu}_i > 0, \forall i \in J'$ such that

$$\mu_j \nabla g_j(x) = \sum_{i \in I} -\bar{\lambda}_i \nabla h_i(x) + \sum_{i \in J'} -\bar{\mu}_i \nabla g_i(x) \tag{13}$$

and

$$(\{\nabla h_i(x)\}_{i \in I}, \{\nabla g_i(x)\}_{i \in J'}) \tag{14}$$

is linearly independent.

Now, RCPLD ensures that Eq. (13) has a solution in $\bar{\lambda}$ and $\bar{\mu}$ when we change x for y in a neighborhood of x. As all the functions involved are continuous and (14) holds, it follows from the pseudo-inverse formula that $\bar{\lambda}$ and $\bar{\mu}$ will change continuously in a neighborhood of x, in particular preserving $\bar{\mu}_i > 0$ for every $i \in J'$. □

4 Applications of RCPLD

We now show how to apply the RCPLD constraint qualification to obtain a stronger convergence result for the general augmented Lagrangian method introduced in [2,3]. We define the method with some small changes in the penalty parameter update suggested in [9].

We consider the problem

$$\text{Minimize } f(x), \quad \text{subject to} \quad h(x) = 0, g(x) \le 0, \underline{h}(x) = 0, \underline{g}(x) \le 0, \tag{15}$$

where $f : \mathbb{R}^n \to \mathbb{R}, h : \mathbb{R}^n \to \mathbb{R}^m, g : \mathbb{R}^n \to \mathbb{R}^p, \underline{h} : \mathbb{R}^n \to \mathbb{R}^{\underline{m}}$ and $\underline{g} : \mathbb{R}^n \to \mathbb{R}^{\underline{p}}$ are continuously differentiable functions. The lower-level constraints $\underline{h}(x) = 0$ and $\underline{g}(x) \le 0$ are usually simple and we assume that a tailored optimization method can cope with them. For example, in the algorithm implemented in ALGENCAN[2] the lower-level constraints define a box and a method that can solve minimization problems inside a box is used. Given $\rho > 0, \lambda \in \mathbb{R}^m, \mu \in \mathbb{R}^p, \mu \ge 0, x \in \mathbb{R}^n$ we define the augmented Lagrangian function

$$\mathcal{L}_\rho(x, \lambda, \mu) = f(x) + \frac{\rho}{2}\left(\left\|h(x) + \frac{\lambda}{\rho}\right\|^2 + \left\|\max\left\{0, g(x) + \frac{\mu}{\rho}\right\}\right\|^2\right). \tag{16}$$

Algorithm Let $\varepsilon_k \ge 0, \varepsilon_k \to 0, \bar{\lambda}^k \in [\lambda_{min}, \lambda_{max}]^m, \bar{\mu}^k \in [0, \mu_{max}]^p$ for all k, $\rho_1 > 0, \tau \in (0, 1), \eta > 1$.

For all k, compute $x^k \in \mathbb{R}^n$ such that there exist $v^k \in \mathbb{R}^{\underline{m}}, w^k \in \mathbb{R}^{\underline{p}}, w^k \ge 0$ satisfying:

$$\left\|\nabla_x \mathcal{L}_{\rho_k}(x^k, \bar{\lambda}^k, \bar{\mu}^k) + \sum_{i=1}^{\underline{m}} v_i^k \nabla \underline{h}_i(x^k) + \sum_{i=1}^{\underline{p}} w_i^k \nabla \underline{g}_i(x^k)\right\| \le \varepsilon_k, \tag{17}$$

$$\|\underline{h}(x^k)\| \le \varepsilon_k, \|\max\{0, \underline{g}(x^k)\}\| \le \varepsilon_k, \tag{18}$$

and

$$w_i^k = 0 \text{ whenever } \underline{g}_i(x^k) < -\varepsilon_k. \tag{19}$$

[2] Freely available at http://www.ime.usp.br/~egbirgin/tango.

We define, for all $i = 1, \ldots, p$,

$$V_i^k = \max\left\{ g_i(x^k), \frac{-\bar{\mu}_i^k}{\rho_k} \right\}. \tag{20}$$

If $k = 1$ or

$$\max\{\|h(x^k)\|, \|V^k\|\} \leq \tau \max\{\|h(x^{k-1})\|, \|V^{k-1}\|\} \tag{21}$$

we define $\rho_{k+1} \geq \rho_k$. Else, we define $\rho_{k+1} \geq \eta\rho_k$.

Remark We can define the multiplier sequences $\{\bar{\lambda}^k\}$ and $\{\bar{\mu}^k\}$ using for example the first order update formula $\bar{\lambda}_i^{k+1} = P_{[\lambda_{min},\lambda_{max}]}(\bar{\lambda}_i^k + \rho_k h_i(x^k)), i = 1, \ldots, m$ and $\bar{\mu}_i^{k+1} = P_{[0,\mu_{max}]}(\bar{\mu}_i^k + \rho_k g_i(x^k)), i = 1, \ldots, p$, where $P_X(\cdot)$ denotes the euclidean projection in X.

The global convergence results for this variation of the augmented Lagrangian algorithms presented in [2,3,9] are based on the fact that they converge to AKKT points. This fact can be easily derived from the convergence proofs. Hence, any condition that ensures that AKKT points are actually KKT points can be used to ensure that the method converges to first order stationary points. In [2,3,9], the authors used the CPLD constraint qualification, but we have just showed in Theorem 2 that RCPLD can be used instead, weakening the required assumptions for convergence. In particular, the convergence analysis of [2,3,9] combined with Theorem 2 produces the following:

Theorem 5 *If x^* is a limit point of a sequence generated by the Algorithm, then it is feasible for the lower-level constraints, $\underline{h}(x) = 0$, $\underline{g}(x) \leq 0$, and one of the following holds:*

– *x^* is a feasible point of (15).*
– *x^* is a KKT point of the problem*

$$\textit{Minimize } \|h(x)\|^2 + \|\max\{0, g(x)\}\|^2, \quad \textit{subject to} \quad \underline{h}(x) = 0, \underline{g}(x) \leq 0. \tag{22}$$

– *RCPLD does not hold at x^* as a feasible point for the lower-level constraints.*

Theorem 6 *If x^* is a limit point of a sequence generated by the Algorithm such that x^* is feasible for (15), then one of the following holds:*

1. *x^* is a KKT point of problem (15).*
2. *RCPLD does not hold at x^* for the full set of constraints $h(x)=0$, $g(x) \leq 0$, $\underline{h}(x)= 0$, $\underline{g}(x) \leq 0$.*

Note that the above theorems are just a restatement of Theorem 3.1 from [9] with RCPLD playing the role of CPLD as the required constraint qualification.

Finally, note that the convergence analysis of the augmented Lagrangian method was recently extended to a version that does not require derivatives [10]. The convergence results are also based on CPLD and can also be generalized using RCPLD.

We now show that an error bound property holds under the RCPLD. This has been previously done for RCRCQ and CPLD in [18], and alternatively for RCRCQ in [16]. More recently such results were generalized by Minchenko and Tarakanov who showed that quasinormality also implies an error bound [19]. However, such result is not enough to show that RCPLD implies an error bound as RCPLD does not imply quasinormality, see Fig. 2.

As mentioned in [20,24], a main motivation to study error bounds arise in practical considerations in the analysis and implementation of iterative methods for solving optimization and equilibrium problems. An error bound is an estimate of the distance of a given point to the feasible set in terms of computable quantities measuring the violation of the constraints and can play a central role in the analysis of algorithms.

Definition 6 [24] We say that an *error bound* holds around a point $x \in \Omega$ if there exist $\alpha > 0$ and a neighborhood $N(x)$ of x such that for every $y \in N(x)$

$$\min_{z \in \Omega} \|z - y\| \le \alpha \max\{\|h(y)\|, \|\max\{0, g(y)\}\|\}.$$

Note that, as all norms in $\mathbb{R}^n$ are equivalent, the choice of the Euclidean norm in the above definition is not essential. Any norm can be used in its place. Finally, we point out that when the error bound property holds then the tangent cone coincides with the linearization cone, that is, Abadie's CQ holds [24]. In this sense, the Theorem below gives an alternative proof to Theorem 3 but requiring second derivatives of the constraints.

Theorem 7 *If $x \in \Omega$ satisfies RCPLD and the functions h and g defining Ω admit second derivatives in a neighborhood of x, then x satisfies an error bound.*

Proof If x is in the interior of Ω, then clearly the error bound property holds. We will assume that x lies in the frontier of Ω. For a fixed $y \in \mathbb{R}^n$, consider the problem

$$\text{Minimize } \|z - y\|, \quad \text{subject to} \quad h(z) = 0, g(z) \le 0. \tag{23}$$

In Theorem 2 of [18], the authors proved that if second derivatives are available, then the error bound property holds at $x \in \Omega$ if, and only if, there exists a neighborhood $N(x)$ of x such that there exist Lagrange multipliers for problem (23) that lie in a fixed compact set for all $y \in N(x)$, $y \notin \Omega$. Let us consider a sequence $y^k \to x$, $y^k \notin \Omega$ and let z^k be a solution to (23) for $y = y^k$. Since $\|z^k - y^k\| \le \|x - y^k\|$ we have also $z^k \to x$. It is a consequence of Theorem 4 that the RCPLD condition is preserved in a neighborhood, thus $z^k \in \Omega$ also satisfies RCPLD for sufficiently large k. Hence, there exist $\{\lambda^k\} \subset \mathbb{R}^m$ and $\{\mu^k\} \subset \mathbb{R}^p$, $\mu_i^k \ge 0$ such that

$$\frac{z^k - y^k}{\|z^k - y^k\|} + \sum_{i=1}^{m} \lambda_i^k \nabla h_i(z^k) + \sum_{i \in A(z^k)} \mu_i^k \nabla g_i(z^k) = 0,$$

for sufficiently large k, where $A(z^k) = \{i \in \{1, \dots, p\} \mid g_i(z^k) = 0\}$. From the definition of RCPLD we have that there exist $I \subset \{1, \dots, m\}$ and $\tilde{\lambda}_i^k$ for every $i \in I$ such that $\{\nabla h_i(z^k)\}_{i\in I}$ is linearly independent and $\sum_{i=1}^m \lambda_i^k \nabla h_i(z^k) = \sum_{i\in I} \tilde{\lambda}_i^k \nabla h_i(z^k)$ for sufficiently large k, hence, by Lemma 1, there exist $J_k \subset A(z^k)$, $\bar{\lambda}_i^k$ for every $i \in I$ and $\bar{\mu}_i^k \geq 0$ for every $i \in J_k$ such that:

$$\frac{z^k - y^k}{\|z^k - y^k\|} + \sum_{i\in I} \bar{\lambda}_i^k \nabla h_i(z^k) + \sum_{i\in J_k} \bar{\mu}_i^k \nabla g_i(z^k) = 0, \tag{24}$$

and

$$\left(\{\nabla h_i(z^k)\}_{i\in I}, \{\nabla g_i(z^k)\}_{i\in J_k}\right) \tag{25}$$

is linearly independent. Let us consider a subsequence such that J_k is the same set J for every k, where $J \subset A(z^k) \subset A(x)$. Define $M_k = \|(\bar{\lambda}^k, \bar{\mu}^k)\|_\infty$ and let us assume by contradiction that $\{M_k\}$ is unbounded. Taking a subsequence such that $\frac{(\bar{\lambda}^k, \bar{\mu}^k)}{M_k} \to (\lambda, \mu) \neq 0, \mu \geq 0$, we may divide (24) by M_k and take limits for this subsequence to obtain:

$$\sum_{i\in I} \bar{\lambda}_i \nabla h_i(x) + \sum_{i\in J} \bar{\mu}_i \nabla g_i(x) = 0.$$

Since RCPLD holds at x we have that $\left(\{\nabla h_i(z^k)\}_{i\in I}, \{\nabla g_i(z^k)\}_{i\in J}\right)$ must be linearly dependent for sufficiently large k, which contradicts (25). This concludes the proof. □

5 Final remarks

We introduced a generalization of the RCRCQ constraint qualification called RCPLD. We showed that this constraint qualification is strictly weaker than RCRCQ and CPLD. The RCPLD shares with CPLD many of its important properties. In particular, it is enough to ensure the convergence of an Augmented Lagrangian algorithm and the presence of an error bound.

An interesting question that was not touched in this paper is whether it is possible to extend the RCRCQ in a way that does not involve assumptions on the behavior of the gradients of *all subsets* of the active inequality constraints. Such extension would better fit the spirit of RCRCQ when the description of the feasible set does not have any inequalities. In this case, the assumption of constant rank has to be fulfilled only by the set of all gradients of the constraints.

Another question is whether RCPLD can still be weakened preserving the convergence of augmented Lagrangian algorithms. It may be also interesting to investigate its role in the convergence of other optimization methods, like sequential quadratic programming or inexact-restoration, as well as in the convergence of the extension of such methods to deal with variational inequalities. Finally, it may be valuable to search for an alternative proof that RCPLD implies the validity of an error bound that does not depend on the existence of second derivatives as required in Theorem 7.

Acknowledgments The authors would like to thank Prof. Alexey Izmailov from the Moscow State University for pointing out the work of Minchenko and Stakhovski [18] and also the fruitful suggestions of the anonymous referees and editor that greately improved the paper.

References

1. Abadie, J.: On the Kuhn-Tucker theorem. In: Abadie, J. (ed.) Nonlinear Programming, pp. 21–36. Wiley, New York (1967)
2. Andreani, R., Birgin, E.G., Martínez, J.M., Schuverdt, M.L.: On augmented Lagrangian methods with general lower-level constraints. SIAM J. Optim. **18**, 1286–1309 (2007)
3. Andreani, R., Birgin, E.G., Martínez, J.M., Schuverdt, M.L.: Augmented Lagrangian methods under the constant positive linear dependence constraint qualification. Math. Program. **112**, 5–32 (2008)
4. Andreani, R., Echagüe, C.E., Schuverdt, M.L.: Constant-rank condition and second-order constraint qualification. J. Optim. Theory Appl. **146**(2), 255–266 (2010)
5. Andreani, R., Haeser, G., Martínez, J.M.: On sequential optimality conditions for smooth constrained optimization. Optimization (2011). doi:10.1080/02331930903578700
6. Andreani, R., Martínez, J.M., Schuverdt, M.L.: On the relation between constant positive linear dependence condition and quasinormality constraint qualification. J. Optim. Theory Appl. **125**, 473–485 (2005)
7. Arutyunov, A.V.: Perturbations of extremum problems with constraints and necessary optimality conditions. J. Sov. Math. **54**, 1342–1400 (1991)
8. Bertsekas, D.P.: Nonlinear Programming. 2nd edn. Athena Scientific, Belmont (1999)
9. Birgin, E., Bueno, L.F., Krejić, N., Martınez, J.M.: Low Order-Value approach for solving VaR-constrained optimization problems. J. Global Optim. (2011). doi:10.1007/s10898-011-9656-7
10. Diniz-Ehrhardt, M.A., Martínez, J.M., Pedroso, L.G.: Derivative-free methods for nonlinear programming with general lower-level constraints. Comput. Appl. Math. **30**(1), 19–52 (2011)
11. Giorgi, G., Guerraggio, A.: On the notion of tangent cone in mathematical programming. Optimization **25**(1), 11–23 (1992)
12. Haeser, G.: On the global convergence of interior-point nonlinear programming algorithms. Comput. Appl. Math. **29**(2), 125–138 (2010)
13. Haeser, G., Schuverdt, M.L.: Approximate KKT condition for continuous variational inequality problems. J. Optim. Theory Appl. (2011). doi:10.1007/s10957-011-9802-x
14. Hestenes, M.R.: Optimization Theory: The Finite Dimensional Case. Wiley, New York (1975)
15. Janin, R.: Directional derivative of the marginal function in nonlinear programming. Math. Program. Stud. **21**, 110–126 (1984)
16. Lu, S.: Relation between the constant rank and the relaxed constant rank constraint qualifications. Optimization (2010). doi:10.1080/02331934.2010.527972
17. Mangasarian, O.L., Fromovitz, S.: The Fritz John optimality conditions in the presence of equality and inequality constraints. J. Math. Anal. Appl. **17**, 37–47 (1967)
18. Minchenko, L., Stakhovski, S.: On relaxed constant rank regularity condition in mathematical programming. Optimization (2011). doi:10.1080/02331930902971377
19. Minchenko, L., Tarakanov, A.: On error bounds for quasinormal programs. J. Optim. Theory Appl. **148**, 571–579 (2010)
20. Pang, J.-S.: Error bound in mathematical programming. Math. Program. **79**, 299–332 (1997)
21. Qi, L., Wei, Z.: On the constant positive linear dependence condition and its application to SQP methods. SIAM J. Optim. **10**, 963–981 (2000)
22. Rockafellar, R.T.: Lagrange multipliers and optimality. SIAM Rev. **35**(2), 183–238 (1993)
23. Schuverdt, M.L.: Métodos de Lagrangiano aumentado com convergência utilizando a condição de dependência linear positiva constante. PhD thesis, IMECC-UNICAMP, Departamento de Matemática Aplicada, Campinas-SP, Brazil (2006). http://www.libdigi.unicamp.br/document/?code=vtls000375801
24. Solodov, M.: Wiley Encyclopedia of Operations Research and Management Science, Chapter Constraint Qualifications. Wiley (2010). doi:10.1002/9780470400531