An extension of Yuan's Lemma and its applications in optimization

G. Haeser*

March 1, 2017. Reviewed April 7, 2017.

Abstract

We prove an extension of Yuan's Lemma to more than two matrices, as long as the set of matrices has rank at most 2. This is used to generalize the main result of [A. Baccari and A. Trad. *On the classical necessary second-order optimality conditions in the presence of equality and inequality constraints. SIAM J. Opt., 15(2):394–408, 2005]*, where the classical necessary second-order optimality condition is proved under the assumption that the set of Lagrange multipliers is a bounded line segment. We prove the result under the more general assumption that the Hessian of the Lagrangian evaluated at the vertices of the Lagrange multiplier set is a matrix set with at most rank 2. We apply the results to prove the classical second-order optimality condition to problems with quadratic constraints and without constant rank of the Jacobian matrix.

Keywords: Quadratic forms, second-order optimality conditions, global convergence.

1 Introduction

In a general nonlinear optimization problem, it is well known that when the set of Lagrange multipliers is bounded, a local minimizer satisfies a second-order optimality condition in the form of the maximum of many quadratic forms being non-negative in the critical cone, where each quadratic form is defined by the Hessian of the Lagrangian function evaluated at different Lagrange multipliers.

In this paper we are interested in second-order optimality conditions that can be verified with one single Lagrange multiplier. This is motivated by algorithmic considerations, since algorithms usually generate a primal-dual sequence and one is interested in proving that limit points of this sequence satisfy a first- or second-order optimality condition. Under this setting, one is usually restricted to considering a smaller subset of the true critical cone.

---

*Department of Applied Mathematics, University of São Paulo, São Paulo, SP, Brazil. Email: `ghaeser@ime.usp.br`. This research was conducted while holding a Visiting Scholar position at Department of Management Science and Engineering, Stanford University, Stanford CA, USA.

Yuan's Lemma [23] gives an important tool for this type of results, since it states that when the maximum of two quadratic forms is non-negative, some convex combination of the matrices is positive semidefinite. This implies that when the set of Lagrange multipliers is a bounded line segment, one can find a Lagrange multiplier such that the Hessian of the Lagrangian is positive semidefinite on a subset of the critical cone. This is the main result of [5]. However, this approach does not work when there are more than two quadratic forms, that is, when the Lagrange multiplier set is larger than a line segment.

In Section 2, we will extend Yuan's Lemma to more than two quadratic forms. This will be done under an assumption of redundancy on the matrices defining the quadratic forms, in the sense that it is a matrix set of rank at most 2. This gives, in Section 3, an optimality condition verifiable on one single Lagrange multiplier under an assumption of redundancy of the set of Hessians of the Lagrangian matrices that includes as a particular case the case of a Lagrange multiplier set equals to a line segment. In Section 4 we apply the results to a special form of quadratically-constrained problems, and we prove that our assumption holds when the rank of the Jacobian matrix increases at most by one in the neighborhood of a local solution. Section 5 gives some conclusions.

Notation: A set $K \subseteq \mathbb{R}^n$ is a first-order cone if K is the direct sum of a subspace and a ray, where a ray is a set of the form $\{td_0 \mid t \geq 0\}$, for some $d_0 \in \mathbb{R}^n$. Given a set $K \subseteq \mathbb{R}^n$ and a symmetric matrix $A \in \mathbb{R}^{n \times n}$, we say that A is positive semidefinite on K if $x^{\mathrm{T}} A x \geq 0$ for all $x \in K$. We denote by $\mathbb{R}^m_+$ the non-negative orthant of $\mathbb{R}^m$ and $\Lambda_m = \{t \in \mathbb{R}^m_+ \mid \sum_{i=1}^m t_i = 1\}$ the m-dimensional simplex. The rank of a set of matrices is the maximum number of linearly independent matrices viewed as vectors on the appropriate space.

2 Extension of Yuan's Lemma

An important result related to the trust-region subproblem is the following:

Lemma 1 (Yuan's Lemma [23]). *Let $A, B \in \mathbb{R}^{n \times n}$ be symmetric matrices and $K \subseteq \mathbb{R}^n$ be a first-order cone. Then the following conditions are equivalent:*

- $\max\{d^{\mathrm{T}} A d, d^{\mathrm{T}} B d\} \geq 0, \forall d \in K$,

- *there exists $t_1 \geq 0, t_2 \geq 0$ with $t_1 + t_2 = 1$ such that $t_1 A + t_2 B$ is positive semidefinite on K.*

Lemma 1 is known as Yuan's Lemma [23] and it has had many important generalizations [5, 13, 16, 10, 9]. In particular, our presentation of Yuan's Lemma corresponds to [5, Corollary 3.2]. Yuan's Lemma when $K = \mathbb{R}^n$ can be easily proved with a separation argument due to the convexity of the set $\{(x^{\mathrm{T}} A x, x^{\mathrm{T}} B x) \mid x \in \mathbb{R}^n\} \subseteq \mathbb{R}^2$ given in [11]. See also some related convexity results in [8, 20, 21, 19, 22].

In the next result we extend Yuan's Lemma to m matrices $A_1, \ldots, A_m \in \mathbb{R}^{n \times n}$ with $\mathrm{rank}(\{A_1, \ldots, A_m\}) \leq 2$. A similar assumption is used to obtain a

result of S-Lemma type in [19, Proposition 3.5], however, our proof is elementary and our result is more general in the sense that it holds for first-order cones.

Lemma 2. *Let $A_i \in \mathbb{R}^{n \times n}, i = 1, \ldots, m, m \geq 2$, be such that $A_i = \alpha_i A_1 + \beta_i A_2, i = 3, \ldots, m$ for some $(\alpha_i, \beta_i) \in \mathbb{R}^2, i = 3, \ldots, m$ and $K \subseteq \mathbb{R}^n$ a first-order cone. Then, the following are equivalent:*

$$\max_{i=1,\ldots,m} \{x^{\mathsf{T}} A_i x\} \geq 0, \forall x \in K, \tag{2.1}$$

$$\exists t \in \Lambda_m \text{ such that } \sum_{i=1}^{m} t_i A_i \text{ is positive semidefinite on } K. \tag{2.2}$$

Proof. The fact that (2.2) implies (2.1) can be easily seen by contradiction. The reciprocal implication is done by induction on m. If $m = 2$, the result follows from Lemma 1. Assume the assertion is true for $m - 1 \geq 2$. Assume (2.1). Since $A_m = \alpha_m A_1 + \beta_m A_2$, let us consider the following cases:

The case $(\alpha_m, \beta_m) = (0,0)$ is trivial. If $\alpha_m \geq 0$ and $\beta_m \geq 0, (\alpha_m, \beta_m) \neq (0,0)$, then $x^{\mathsf{T}} A_1 x < 0$ and $x^{\mathsf{T}} A_2 x < 0$ imply $x^{\mathsf{T}} A_m x < 0$. Thus, (2.1) implies $\max_{i=1,\ldots,m-1}\{x^{\mathsf{T}} A_i x\} \geq 0, \forall x \in K$. If $\alpha_m < 0$ and $\beta_m > 0$, we have $x^{\mathsf{T}} A_1 x < 0$ and $x^{\mathsf{T}} A_m x < 0$ imply $x^{\mathsf{T}} A_2 x < -\frac{\alpha_m}{\beta_m} x^{\mathsf{T}} A_1 x < 0$, hence, (2.1) implies $\max_{i=1,\ldots,m,i\neq 2}\{x^{\mathsf{T}} A_i x\} \geq 0, \forall x \in K$. It is clear that any matrix $A_i, i = 1, \ldots, m, i \neq 2$ can be written as the linear combination of two fixed matrices in this set. The case $\alpha_m < 0$ and $\beta_m = 0$ gives $\max\{x^{\mathsf{T}} A_1 x, x^{\mathsf{T}} A_m x\} \geq 0, \forall x \in K$. Also, the case $\alpha \geq 0, \beta < 0$ is analogous. In any of the above cases, the result follows from Lemma 1 or the inductive assumption. If $\alpha_m < 0$ and $\beta_m < 0$, we have $t_1 A_1 + t_2 A_2 + t_m A_m = (t_1 + \alpha_m t_m)A_1 + (t_2 + \beta_m t_m)A_2 = 0$, which is positive semidefinite on K, for $(t_1, t_2, t_m) = (\frac{-\alpha_m}{1-\alpha_m-\beta_m}, \frac{-\beta_m}{1-\alpha_m-\beta_m}, \frac{1}{1-\alpha_m-\beta_m}) \in \Lambda_3$. $\qquad \square$

Example 1. Let

$$A_1 = \begin{pmatrix} 1 & -1 \\ -1 & 1 \end{pmatrix}, A_2 = \begin{pmatrix} -2 & 1 \\ 1 & 1 \end{pmatrix}, A_3 = \begin{pmatrix} 4 & -3 \\ -3 & 1 \end{pmatrix}.$$

Note that $A_3 = 2A_1 - A_2$. A simple calculation shows that $\max_{i=1,2,3}\{x^{\mathsf{T}} A_i x\} \geq 0$ for all x, and, as predicted by Lemma 2, $t_1 A_1 + t_2 A_2 + t_3 A_3$ is positive semidefinite, for instance, for $(t_1, t_2, t_3) = (0, \frac{3}{5}, \frac{2}{5})$. As it is done in the proof, this is a consequence of the condition $\max\{x^{\mathsf{T}} A_2 x, x^{\mathsf{T}} A_3 x\} \geq 0$ for all x that necessarily holds.

Example 2. Let

$$A_1 = \begin{pmatrix} -1 & 0 \\ 0 & 1 \end{pmatrix}, A_2 = \begin{pmatrix} 1 & 2 \\ 2 & -2 \end{pmatrix}, A_3 = \begin{pmatrix} 0 & -2 \\ -2 & 1 \end{pmatrix}.$$

Note that $A_3 = -A_1 - A_2$. Therefore, $\frac{1}{3}A_1 + \frac{1}{3}A_2 + \frac{1}{3}A_3 = 0$ is positive semidefinite and clearly it can not exist $x \in \mathbb{R}^2$ with $x^{\mathsf{T}} A_1 x < 0, x^{\mathsf{T}} A_2 x < 0$ and $x^{\mathsf{T}} A_3 x < 0$, which implies the maximum of the three quadratic forms is non-negative. Note, however, that for all $i \neq j$, it is not the case that $\max\{x^{\mathsf{T}} A_i x, x^{\mathsf{T}} A_j x\} \geq 0$ for all x.

3

3 A Second-order Optimality Condition

Let us consider the nonlinear optimization problem

$$
\begin{array}{ll}
\text{Minimize} & f(x), \\
\text{subject to} & h(x) = 0, \\
& g(x) \leq 0,
\end{array}
\tag{3.1}
$$

where $f : \mathbb{R}^n \to \mathbb{R}, h : \mathbb{R}^n \to \mathbb{R}^{p_1}, g : \mathbb{R}^n \to \mathbb{R}^{p_2}$ are twice continuously differentiable functions. For a feasible x, we define $A(x) = \{i \in \{1, \ldots, p_2\} \mid g_i(x) = 0\}$, the index set of active inequality constraints. Given $(\lambda, \mu) \in \mathbb{R}^{p_1} \times \mathbb{R}_+^{p_2}$, we consider the Lagrangian function

$$
x \mapsto L(x, \lambda, \mu) = f(x) + \sum_{i=1}^{p_1} \lambda_i h_i(x) + \sum_{i=1}^{p_2} \mu_i g_i(x),
$$

with gradient $\nabla L(x, \lambda, \mu)$ and Hessian $\nabla^2 L(x, \lambda, \mu)$, where derivative is taken with respect to x. We will assume that at a local minimizer x^* of (3.1), the Mangasarian-Fromovitz constraint qualification holds:

Definition 1. A feasible point x^* of problem (3.1) satisfies the Mangasarian-Fromovitz constraint qualification [15] when $(\alpha, \beta) = (0, 0)$ is the only solution of

$$
\sum_{i=1}^{p_1} \alpha_i \nabla h_i(x^*) + \sum_{i \in A(x^*)} \beta_i \nabla g_i(x^*) = 0, \alpha_i \in \mathbb{R}, i = 1, \ldots, p_1; \beta_i \geq 0, i \in A(x^*).
$$

It is well known [12] that at a local minimizer x^*, the Mangasarian-Fromovitz constraint qualification is equivalent to the non-emptyness and boundedness of the Lagrange multiplier set

$$
\Lambda(x^*) = \{(\lambda, \mu) \in \mathbb{R}^{p_1} \times \mathbb{R}_+^{p_2} \mid \nabla L(x^*, \lambda, \mu) = 0, \mu^\mathsf{T} g(x^*) = 0\}.
$$

We observe also that $\Lambda(x^*)$ is always a closed polyhedron. The following optimality condition is well-known:

Theorem 1 ([7]). *Let x^* be a local minimizer of* (3.1) *satisfying the Mangasarian-Fromovitz constraint qualification. Then, for each $d \in C(x^*)$, there is a Lagrange multiplier $(\lambda, \mu) \in \Lambda(x^*)$ satisfying*

$$
d^\mathsf{T} \nabla^2 L(x^*, \lambda, \mu) d \geq 0,
\tag{3.2}
$$

where

$$
C(x^*) = \left\{ \begin{array}{ll} d \in \mathbb{R}^n : & \nabla h_i(x^*)^\mathsf{T} d = 0, i = 1, \ldots, p_1; \nabla g_i(x^*)^\mathsf{T} d \leq 0, i \in A(x^*); \\ & \nabla f(x^*)^\mathsf{T} d = 0. \end{array} \right\}
\tag{3.3}
$$

is the critical cone.

Theorem 1 is the classical second-order necessary optimality condition under Mangasarian-Fromovitz constraint qualification alone. A version of it without constraint qualifications and using Fritz-John multipliers is known as a no-gap optimality condition in terms of sufficiency, see details in [7].

Note that condition (3.2) holds with different Lagrange multipliers for different directions in the critical cone. In this paper we are interested in conditions that ensure the validity of (3.2) for the same Lagrange multiplier $(\lambda, \mu) \in \Lambda(x^*)$, even if for that we need to consider a smaller subset of the true critical cone $C(x^*)$. This is motivated by algorithmic considerations, since algorithms usually generate only one Lagrange multiplier approximation. See the discussion in [6].

In the next theorem, we will use Lemma 2 to formulate a new optimality condition of this type. It is a generalization of the main result of [5], where $\Lambda(x^*)$ was assumed to be a bounded line segment.

Theorem 2. *Let x^* be a local minimizer of* (3.1) *satisfying the Mangasarian-Fromovitz constraint qualification. Let $(\lambda^1, \mu^1), \ldots, (\lambda^v, \mu^v)$ be the vertices of the Lagrange multiplier set $\Lambda(x^*)$. If the rank of $\{\nabla^2 L(x^*, \lambda^i, \mu^i)\}_{i=1}^v$ is at most 2, then for every first-order cone $K \subseteq C(x^*)$, there exists $(\lambda, \mu) \in \Lambda(x^*)$ such that*

$$d^\mathrm{T} \nabla^2 L(x^*, \lambda, \mu) d \geq 0, \forall d \in K. \tag{3.4}$$

Proof. Let $K \subseteq C(x^*)$ be a first-order cone. From Theorem 1, we have

$$\max_{(\lambda,\mu) \in \Lambda(x^*)} \{d^\mathrm{T} \nabla^2 L(x^*, \lambda, \mu) d\} \geq 0, \forall d \in K. \tag{3.5}$$

Since Mangasarian-Fromovitz constraint qualification implies that $\Lambda(x^*)$ is a non-empy, compact, polyhedral set, let $(\lambda^1, \mu^1), \ldots, (\lambda^v, \mu^v)$ be its vertices. Since for each $d \in K$ the maximization problem (3.5) is a linear programming with non-empty and compact feasible region, its solution is attained at a vertex, which implies

$$\max_{i=1,\ldots,v} \{d^\mathrm{T} \nabla^2 L(x^*, \lambda^i, \mu^i) d\} \geq 0, \forall d \in K. \tag{3.6}$$

Lemma 2 gives $t \in \Lambda_v$ such that $\sum_{i=1}^v t_i \nabla^2 L(x^*, \lambda^i, \mu^i)$ is positive semidefinite on K, and the linearity of $\nabla^2 L(x^*, \cdot, \cdot)$ gives $\nabla^2 L(x^*, \lambda, \mu)$ positive semidefinite on K for $(\lambda, \mu) = \sum_{i=1}^v t_i (\lambda^i, \mu^i) \in \Lambda(x^*)$. $\qquad \square$

Note that when the critical cone is a first-order cone, Theorem 2 gives a full second-order optimality condition in terms of the true critical cone. One condition ensuring this, is the generalized scrict complementarity slackness, namely, that there is at most one index $i_0 \in A(x^*)$ such that $\mu_{i_0} = 0$ for all $(\lambda, \mu) \in \Lambda(x^*)$. See [5].

Let us revisit Examples 1 and 2. The fact that $\max_{i=1,2,3} \{x^\mathrm{T} A_i x\} \geq 0$ for all x, implies that $x^* = (0,0), z^* = 0$ is a global minimizer of the problem of minimizing z, subject to $\frac{1}{2} x^\mathrm{T} A_i x - z \leq 0, i = 1, 2, 3$. It is easy to see that the Mangasarian-Fromovitz constraint qualification holds, the set of Lagrange multipliers at (x^*, z^*) is the simplex Λ_3 and the critical cone is $\mathbb{R}^2 \times \{0\}$. Note

that the quadratic form defined by the Hessian of the Lagrangian at (x^*, z^*) evaluated at each of the three vertices of Λ_3 and restricted to the critical cone gives the quadratic forms defined by the three matrices A_1, A_2 and A_3. Hence, the optimality condition of Theorem 2 translates precisely to the fact that $t_1 A_1 + t_2 A_2 + t_3 A_3$ is positive semidefinite for some $(t_1, t_2, t_3) \in \Lambda_3$, which holds as observed in Examples 1 and 2. In the next session we will give a sufficient condition for the rank assumption to hold under this setting.

4 Application to a Quadratically-constrained Problem

It is well known that without the rank assumption, the optimality condition of Theorem 2 does not hold. Known counter-examples [3, 2, 4] are of the form

$$
\begin{aligned}
&\text{Minimize} \quad z, \\
&\text{subject to} \quad g_i(x, z) := \tfrac{1}{2} x^{\mathsf{T}} A_i x - z \leq 0, i = 1, \ldots, m,
\end{aligned}
\tag{4.1}
$$

where $A_i \in \mathbb{R}^{n \times n}, i = 1, \ldots, m$, are symmetric matrices and $(x^*, z^*) = (0, 0) \in \mathbb{R}^n \times \mathbb{R}$ is a solution. This is a special case of the well known quadratically-constrained quadratic programming problem, a well-studied difficult non-convex optimization problem [14, 18, 17]. Note that the linear objective can replace an objective function $f(x)$ by adding the constraint $f(x) - z \leq 0$. Note that the critical cone $C(x^*, z^*)$ is the subspace $\mathbb{R}^n \times \{0\}$. We will prove that for this type of problem, in order to fulfill our rank assumption, it is sufficient to assume that the rank of the Jacobian matrix (that is, the matrix of gradients of active constraints at the solution) increases at most by one in a neighborhood of the solution. It is conjectured in [1] that for the general problem (3.1) satisfying Mangasarian-Fromovitz constraint qualification at the local solution x^*, when the rank of the Jacobian matrix increases at most by one in a neighborhood of the solution, the optimality condition (3.4) holds with the cone K equal to the lineality space of $C(x^*)$. This is known to hold when the rank of the Jacobian matrix is constant [1], one less than the number of active constraints [5], one less than the dimension n [6], or when the Jacobian matrix has a smooth singular value decomposition [6]. We refer the reader to [6] for a discussion about this conjecture and related results. Our result proves this conjecture for the class of quadratically-constrained problems (4.1). Note that this is not a particular case of any known case where the conjecture holds. In particular, [6, Example 3.3] shows a problem of type (4.1) with $m = 3$ active constraints and $n = 3$ variables where the singular value decomposition is not smooth, and the rank of the Jacobian matrix increases from 1 to at most 2 in a neighborhood of the origin. Our theorem applies to this problem, whereas no other known result applies to obtain (3.4). To present our result, we start with the following lemma:

Lemma 3. *Given $a \neq 0$ and $A_i \in \mathbb{R}^{n \times n}$ symmetric matrices, $i = 1, \ldots, m$, define $v_i(x) = \begin{pmatrix} A_i x \\ a \end{pmatrix} \in \mathbb{R}^{n+1}, i = 1, \ldots, m$ and $J(x) \in \mathbb{R}^{(n+1) \times m}$ given*

column-wise by $J(x) = [v_1(x) \cdots v_m(x)]$. If the rank of $J(x)$ is at most 2 for all $x \in \mathbb{R}^n$, then the rank of $\{A_1, \ldots, A_m\}$ is at most 2.

Proof. Take any three distinct indexes $\{i_1, i_2, i_3\} \subseteq \{1, \ldots, m\}$ and let $A = A_{i_1}, B = A_{i_2}, C = A_{i_3}$. Let us prove that $\{A, B, C\}$ is linearly dependent. Since the submatrix of $J(x)$ consisting of columns i_1, i_2 and i_3 has rank at most 2, we have that for all $x \in \mathbb{R}^n$, there is $(\alpha_x, \beta_x, \gamma_x) \neq (0, 0, 0)$ such that $\alpha_x Ax + \beta_x Bx + \gamma_x Cx = 0$ and $\alpha_x + \beta_x + \gamma_x = 0$. That is,

$$\forall x \in \mathbb{R}^n, \alpha_x (A - C)x + \beta_x (B - C)x = 0, \text{ for some } (\alpha_x, \beta_x) \neq (0, 0). \quad (4.2)$$

Let $v_1, \ldots, v_n \in \mathbb{R}^n$ be an orthonormal basis of eigenvectors of the symmetric matrix $B - C$ with corresponding eigenvalues $\lambda_1, \ldots, \lambda_n \in \mathbb{R}$. If all eigenvalues are zero, then $B = C$ and the result follows. Suppose $\lambda_1 \neq 0$ and let us apply (4.2) with $x = v_1$:

$$\alpha_1 (A - C)v_1 + \beta_1 \lambda_1 v_1 = 0, \text{ for some } (\alpha_1, \beta_1) \neq (0, 0).$$

For $i = 1, \ldots, n$ let us take the inner product with v_i. We have:

$$\alpha_1 v_1^\mathsf{T}(A - C)v_1 + \beta_1 \lambda_1 = 0 \text{ and } \alpha_1 v_i^\mathsf{T}(A - C)v_1 = 0, i > 1.$$

Since $(\alpha_1, \beta_1) \neq (0, 0)$ and $\lambda_1 \neq 0$, we have $\alpha_1 \neq 0$. That is,

$$v_1^\mathsf{T}(A - C)v_1 = -\frac{\beta_1}{\alpha_1}\lambda_1 \text{ and } v_i^\mathsf{T}(A - C)v_1 = 0, i > 1. \quad (4.3)$$

Now, for each $i = 2, \ldots, n$, let us repeat the construction for $x = v_1 + v_i$. From (4.2), there is some $(\alpha_i, \beta_i) \neq (0, 0)$ such that:

$$\alpha_i (A - C)v_1 + \alpha_i (A - C)v_i + \beta_i(\lambda_1 v_1 + \lambda_i v_i) = 0.$$

Taking the inner product with v_1 and using (4.3) and the fact that $v_1^\mathsf{T}(A - C)v_i = v_i^\mathsf{T}(A - C)v_1$, we have:

$$-\alpha_i \frac{\beta_1}{\alpha_1}\lambda_1 + \beta_i \lambda_1 = 0,$$

which implies $\alpha_i \neq 0$ and $\frac{\beta_i}{\alpha_i} = \frac{\beta_1}{\alpha_1}, i > 1$. Thus, we conclude that for $u_1 = v_1, u_2 = v_1 + v_2, \ldots, u_n = v_1 + v_n$ it holds

$$(A - C)u_i + \delta(B - C)u_i = 0, \text{ for all } i = 1, \ldots, n,$$

where $\delta = \frac{\beta_1}{\alpha_1}$. Since $\{u_1, \ldots, u_n\}$ is a basis of $\mathbb{R}^n$, we have $A - C + \delta(B - C) = 0$, and the result follows. $\qquad \square$

Note that Lemma 3 does not hold without the assumption of symmetry of the matrices. A counter-example is $A_1 = \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix}$, $A_2 = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}$ and $A_3 = \begin{pmatrix} 1 & 0 \\ 1 & 0 \end{pmatrix}$ with $a = 1$. We give next our result for problem (4.1).

Theorem 3. *If* $(x^*, z^*) = (0, 0)$ *is a local minimizer of* (4.1) *and the rank of the Jacobian matrix* $J(x, z) = [\nabla_{x,z} g_1(x, z) \cdots \nabla_{x,z} g_m(x, z)]$ *increases at most by one in a neighborhood with respect to the rank at* $(x, z) = (x^*, z^*)$, *then there exists* $t \in \Lambda_m$ *such that* $\sum_{i=1}^m t_i A_i$ *is positive semidefinite.*

Proof. We have $\nabla_{x,z} g_i(x, z) = \begin{pmatrix} A_i x \\ -1 \end{pmatrix}, i = 1, \ldots, m$. Hence, Mangasarian-Fromovitz holds at (x^*, z^*), the critical cone $C(x^*, z^*)$ is the subspace $\mathbb{R}^n \times \{0\}$, and the set of Lagrange multipliers is the simplex Λ_m. Hence, its vertices are given by $\mu^i = e_i, i = 1, \ldots, m$ where $e_1, \ldots, e_m$ is the canonical basis of $\mathbb{R}^m$ and $\nabla_x^2 L(x^*, z^*, \mu^i) = A_i, i = 1, \ldots, m$. Since the rank of $J(x^*) = J(x^*, z^*)$ is one, the rank assumption implies that $J(x)$ has rank at most 2 in some neighborhood of x^*. This implies that $J(x)$ has rank at most 2 for all $x \in \mathbb{R}^n$, and the result follows from Lemma 3 and Theorem 2.

Note that the local minimality of $(0, 0)$ implies that $\max_{i=1,\ldots,m}\{x^{\mathsf{T}} A_i x\} \geq 0$ for all x in a small neighborhood of the origin, which implies it holds for all $x \in \mathbb{R}^n$, hence the result would also follow from Lemma 3 and Lemma 2. $\qquad\square$

5   Conclusions

Although no-gap necessary and sufficient second-order optimality conditions of Fritz-John type are well established in the optimization literature [7], these type of optimality condition has little relations to second-order global convergence results of algorithms. Hence the importance of developing weak conditions ensuring the existence of a Lagrange multiplier with a positive semidefinite Lagrangian Hessian on the critical cone, or a meaningful subset.

We developed a second-order optimality condition of this type where the Lagrange multiplier set can be larger than a bounded line segment, generalizing [5]. This was done by an extension of Yuan's Lemma [23] to a set of linearly dependent matrices. The results were applied to address a conjecture from [1] about a second-order optimality condition of this type under non-constant rank on a class of quadratically-constrained problems.

Acknowledgement

References

[1] R. Andreani, J. M. Martínez, and M. L. Schuverdt. On second-order optimality conditions for nonlinear programming. *Optimization*, 56:529–542, 2007.

[2] M. Anitescu. Degenerate nonlinear programming with a quadratic growth condition. *SIAM Journal on Optimization*, 10:4:1116–1135, 2000.

[3] A. V. Arutyunov. Second-order conditions in extremal problems. the abnormal points, 1998.

[4] A. Baccari. On the classical necessary second-order optimality conditions. *Journal of Optimization Theory and Applications*, 123(1):213–221, 2004.

[5] A. Baccari and A. Trad. On the classical necessary second-order optimality conditions in the presence of equality and inequality constraints. *SIAM Journal on Optimization*, 15(2):394–408, 2005.

[6] R. Behling, G. Haeser, A. Ramos, and D. S. Viana. On a conjecture in second-order optimality conditions. *Optimization-online*, 2016.

[7] J. F. Bonnans and A. Shapiro. *Pertubation Analysis of Optimization Problems*. Springer-Verlag, 2000.

[8] L. Brickman. On the field of values of a matrix. *Proceedings of the AMS*, (12):61–66, 1961.

[9] X. Chen and Y. Yuan. A note on quadratic forms. *Mathematical Programming*, (86):187–197, 1999.

[10] J. P. Crouzeix, J. E. Martinez-Legal, and A. Seeger. An alterntive theorem for quadratic forms and extensions. *Linear Algebra and its applications*, 215:121–134, 1995.

[11] L. L. Dines. On the mapping of quadratic forms. *Bulletin of the AMS*, (47):494–498, 1941.

[12] J. Gauvin. A necessary and sufficient regularity condition to have bounded multipliers in nonconvex programming. *Mathematical Programming*, (12):136–138, 1977.

[13] J.-B. Hiriart-Urruty and M. Torki. Permanently going back and forth between the quadratic world and the convexity world in optimization. *Applied Mathematics and Optimization*, (45):169–184, 2002.

[14] J. Linderoth. A simplicial branch-and-bound algorithm for solving quadratically constrained quadratic programs. *Mathematical Programming*, 103(2):251–282, 2005.

[15] O. L. Mangasarian and S. Fromovitz. The Fritz-John necessary conditions in presence of equality and inequality constraints. *Journal of Mathematical Analysis and applications*, 17:37–47, 1967.

[16] J. E. Martinez-Legaz and A. Seeger. Yuan's alternative theorem and the maximization of the minimum eigenvalue function. *Journal of Optimization Theory and Applications*, 82(1):159, 1994.

[17] J. J. Moré. Generalizations of the trust region problem. *Optimization Methods and Software*, 2(3-4):189–209, 1993.

[18] P. M. Pardalos and S. A. Vavasis. Quadratic programming with one negative eigenvalue is NP-hard. *Journal of Global Optimization*, 1(1):15–22, 1991.

[19] I. Pólik and T. Terlaky. A survey of the S-lemma. *SIAM Review*, 3(49):371–418, 2007.

[20] B.T. Polyak. Convexity of quadratic transformations and its use in control and optimization. *Journal of Optimization Theory and Applications*, (99):553–583, 1998.

[21] B.T. Polyak. Convexity of nonlinear image of a small ball with applications to optimization. *Set-Valued Analysis*, (9):159–168, 2001.

[22] Y. Xia. On local convexity of quadratic transformations. *arXiv:1405.6042*, 2014.

[23] Y. Yuan. On a subproblem of trust region algorithms for constrained optimization. *Mathematical Programming*, 47:53–63, 1990.