Necessary and Sufficient Conditions for S-Lemma and Nonconvex Quadratic Optimization*

V. Jeyakumar†, N. Q. Huy‡ and G. Y. Li§

Revised Version: December 5, 2008

Abstract

The celebrated S-lemma establishes a powerful equivalent condition for the nonnegativity of a quadratic function over a single quadratic inequality. However, this lemma fails without the technical condition, known as the Slater condition. In this paper, we first show that the Slater condition is indeed necessary for the S-lemma and then establishes a regularized form of the S-lemma in the absence of the Slater condition. Consequently, we present characterizations of global optimality and the Lagrangian duality for quadratic optimization problems with a single quadratic constraint. Our method of proof makes use of Brickman's theorem and conjugate analysis, exploiting the hidden link between the convexity and the S-lemma.

Key Words. Nonconvex quadratic optimization, S-lemma, regularized S-lemma, Slater's condition, necessary and sufficient global optimality conditions.

AMS subject classification. 41A65, 41A29, 90C30.

---

*Authors are grateful to the referee for his comments and suggestions which have contributed to the final preparation of the paper. This research was partially supported by the Australian Research Council.

†Department of Applied Mathematics, University of New South Wales, Sydney 2052, Australia. E-mail: jeya@maths.unsw.edu.au.

‡Department of Applied Mathematics, University of New South Wales, Sydney 2052, Australia. E-mail: nqhuy@maths.unsw.edu.au; Department of Mathematics, Hanoi Pedagogical University No.2, Phuc Yen, Vinh Phuc, Vietnam.

§Department of Applied Mathematics, University of New South Wales, Sydney 2052, Australia. E-mail: g.li@unsw.edu.au.

1 Introduction

The S-lemma provides an elegant and powerful equivalent condition for the nonnegativity of a quadratic function $f(x)$ over a single quadratic inequality $g(x) \leq 0$. In symbolic terms, it states that

$$[g(x) \leq 0 \Longrightarrow f(x) \geq 0] \Longleftrightarrow (\exists \mu \geq 0)(\forall x \in I\!R^n) \;\; f(x) + \mu g(x) \geq 0, \tag{1}$$

whenever $g(x_0) < 0$ for some x_0. This technical condition is known as the Slater condition. The S-lemma is a widely used tool in control theory, and in particular, in stability analysis of nonlinear systems [1, 2]. It also has important applications in quadratic optimization as well as in semidefinite optimization [5, 16, 17, 20]. For an excellent recent survey of the S-lemma, see [19]. In this paper, we show that the Slater condition is a necessary and sufficient condition for the S-lemma in the sense that (1) holds for each quadratic function f if and only if the Slater condition holds. As an application, we present a characterization of Lagrangian duality for quadratic optimization problems with a single quadratic inequality constraint in terms of the Slater condition.

On the other hand, the Slater condition limits the application of the S-lemma. For instance, the Slater condition is never satisfied for a system involving a homogeneous convex quadratic inequality, $g(x) \leq 0$, where $g(x) = x^T Bx$ and B is a symmetric positive semidefinite matrix. In the absence of the Slater condition, we establish that a regularized form of the S-lemma holds. Consequently, we derive a complete characterization of global optimality of a quadratic minimization problem over a single quadratic inequality constraint without the Slater condition.

Our method of proof combines the application of the Brickman's theorem [18, 3, 19] and conjugate analysis, exploiting the critical link between the convexity and the S-lemma. Our approach was motivated by the recent work on Farkas' lemma for convex inequality systems and its applications to convex programming problems (see [8, 7, 13, 11, 14, 12]).

The layout of the paper is as follows. Section 2 provides background material on convex conjugate analysis and on convexity properties of quadratic forms, used later in the paper. Section 3 presents a characterization of S-lemma in terms of the Slater condition and provides a characterization of the Lagrangian duality for quadratic programming problems with a single quadratic constraint. Section 4 establishes a regularized S-lemma and a complete characterization of global optimality of a quadratic minimization problem over a single quadratic inequality constraint without the Slater condition.

2 Preliminaries

We recall in this section notation and basic results which will be used later in the paper. The space of all $(n\times n)$ symmetric matrices is denoted by S^n. The $(n\times n)$ identity matrix is denoted by I_n. The notation $A \succeq B$ means that the matrix $A-B$ is positive semidefinite. Moreover, the notation $A \succ B$ means the matrix $A-B$ is positive definite. The positive semidefinite cone is defined by $S^n_+ := \{M \in S^n : M \succeq 0\}$. Let $A, B \in S^n$. Denote the (trace) inner product of A and B is defined by $A \cdot B = Tr[AB] = \sum_{i=1}^n \sum_{j=1}^n a_{ij} b_{ji}$ where a_{ij} is the (i,j) element of A and b_{ji} is the (j,i) element of B. Let K be a cone in S^n. The norm of $A \in S^n$ and the distance of A to the cone K is defined respectively by $\|A\| = (A \cdot A)^{1/2}$ and $d(A,K) = \inf_{B\in K} \|A-B\|$. The kernel of a matrix $A \in S^n$ is defined by $\mathrm{Ker}A := \{x \in \mathbb{R}^n : Ax = 0\}$. For a subset $D \subset I\!R^n$, the closure of D will be denoted by $\overline{D}$. A set $K \subset I\!R^n$ is called a cone if $\lambda K \subseteq K$ for any $\lambda \geq 0$. The (negative) polar of K is defined by $K^\circ := \{d : d^T x \leq 0\, \forall\, x \in K\}$.

For cones $K_1,\ K_2 \in I\!R^n$, the following polar formula holds (see [23]):

$$(K_1^\circ \cap K_2^\circ)^\circ = \overline{K_1 + K_2}.$$

In particular, letting $K_1 = -S^n_+$ and $K_2 = \bigcup_{t\geq 0} tH_2$, where H_2 is some matrix in S^n, we obtain that

$$\{X \in S^n_+ : H_2 \cdot X \leq 0\}^\circ = (K_1^\circ \cap K_2^\circ)^\circ = \overline{K_1 + K_2} = \overline{-S^n_+ + \bigcup_{t\geq 0} tH_2}. \tag{2}$$

Let $h : \mathbb{R}^n \to \mathbb{R} \cup \{-\infty, +\infty\}$. The conjugate function of h, $h^* : \mathbb{R}^n \to \mathbb{R} \cup \{+\infty\}$, is defined by

$$h^*(v) := \sup\{v(x) - h(x) \mid x \in \mathrm{dom}\ h\},$$

where $\mathrm{dom}\ h := \{x \in \mathbb{R}^n \mid h(x) < +\infty\}$ is the effective domain of h and $v(x) := v^T x$. The function h is said to be proper if h does not take on the value $-\infty$ and $\mathrm{dom}\ h \neq \emptyset$. The epigraph of h is defined by

$$\mathrm{epi}\ h := \{(x,r) \in \mathbb{R}^n \times \mathbb{R} \mid x \in \mathrm{dom}\ h, h(x) \leq r\}.$$

If $h : \mathbb{R}^n \to \mathbb{R}$ is a continuous convex function, the subdifferential of h at $x \in \mathbb{R}^n$ is defined by

$$\partial h(x) = \{a \in \mathbb{R}^n : a^T(y-x) \leq h(y) - h(x)\, \forall\, y \in \mathbb{R}^n\}. \tag{3}$$

For $\epsilon \geq 0$, the ϵ-subdifferential of h at $x \in \mathbb{R}^n$ is defined by

$$\partial_\epsilon h(x) = \{a \in \mathbb{R}^n : a^T(y-x) \leq h(y) - h(x) + \epsilon. \forall\, y \in \mathbb{R}^n\}. \tag{4}$$

It can be verified that $\partial h(x) \subseteq \partial_\epsilon h(x)$ for any $\epsilon > 0$ and $\partial h(x) = \{\nabla h(x)\}$ if h is a continuously differentiable convex function. For $h : \mathbb{R}^n \to \mathbb{R}$, $[h \leq 0] := \{x \in \mathbb{R}^n \mid h(x) \leq 0\}$. The normal cone of $[h \leq 0]$ at the point $x \in [h \leq 0]$ is $N_{[h\leq 0]}(x) := \{a \in \mathbb{R}^n : a^T(z-x) \leq 0, \forall\, z \in [h \leq 0]\}$.

If h is a continuously differentiable convex function on $\mathbb{R}^n$, then for each $x \in [h \leq 0]$,

$$N_{[h\leq 0]}(x) = \{a \in \mathbb{R}^n : (a, a^T x) \in \overline{\bigcup_{\lambda \geq 0} \mathrm{epi}(\lambda h)^*}\}, \tag{5}$$

and

$$\bigcup_{\lambda \geq 0, \lambda h(x) = 0} \{\lambda \nabla h(x)\} = \{a \in \mathbb{R}^n : (a, a^T x) \in \bigcup_{\lambda \geq 0} \mathrm{epi}(\lambda h)^*\}. \tag{6}$$

Thus, if $\bigcup_{\lambda \geq 0} \mathrm{epi}\ (\lambda h)^*$ is closed, then for each $x \in [h \leq 0]$,

$$N_{[h\leq 0]}(x) = \bigcup_{\lambda \geq 0, \lambda h(x) = 0} \{\lambda \nabla h(x)\}. \tag{7}$$

For details see [10].

Let f be a quadratic function defined by $f(x) = x^T A x + a^T x + \alpha$ where $A \in S^n$, $a \in \mathbb{R}^n$ and $\alpha \in \mathbb{R}$. We define H_f by

$$H_f = \begin{pmatrix} A & a/2 \\ a^T/2 & \alpha \end{pmatrix}. \tag{8}$$

A useful fact is that $f(x) \geq 0$ for all $x \in \mathbb{R}^n \;\Leftrightarrow\; H_f \in S^{n+1}_+$.

The following convexity results (cf [3, 18]) play a key role in deriving a regularized S-lemma later in the paper.

Lemma 2.1 (Brickman, [3]) *If $f,\ g : \mathbb{R}^n \to \mathbb{R}$ $(n \geq 3)$ are homogeneous quadratic functions, defined by $f(x) = x^T A x$ and $g(x) = x^T B x$, $A, B \in S^n$, then the set $\mathcal{V} := \{(f(x), g(x)) \mid \|x\| = 1\} \subset \mathbb{R}^2$ is a convex compact set.*

Lemma 2.2 (**Martínez-Legaz, [18, Lemma 3.1]**) *Let A, B be two 2×2 real symmetric matrices. Define $K = \{(x^T A x, x^T B x) : x \in \mathbb{R}^2\}$. Then, $K = \{(A \cdot X, B \cdot X) : X \in S^2_+\}$ and is convex.*

3 Characterizations of S-Lemma and Duality

In this section, we provide characterizations of S-Lemma and Lagrangian duality for a quadratic optimization over a single quadratic constraint in terms of the Slater condition.

Theorem 3.1 *(Characterization of S-Lemma) Let g be a quadratic function that is not identically zero. Suppose that $[g \leq 0] \neq \emptyset$. Then the following statements are equivalent:*

(i) *For each quadratic function $f : \mathbb{R}^n \to I\!R$,*
$$[g(x) \leq 0 \Longrightarrow f(x) \geq 0] \Longleftrightarrow (\exists \lambda \geq 0)(\forall x \in \mathbb{R}^n) \;\; f(x) + \lambda g(x) \geq 0.$$

(ii) $\bigcup_{\lambda \geq 0} \mathrm{epi}(\lambda g)^*$ *is a closed set in* $\mathbb{R}^{n+1}$.

(iii) *There exists $x_0 \in \mathbb{R}^n$ such that $g(x_0) < 0$.*

Proof. [(i) $\Rightarrow$ (ii)] Let $(u, \alpha) \in \mathrm{cl} \bigcup_{\lambda \geq 0} \mathrm{epi}\ (\lambda g)^*$. Then there exist sequences $\{u_k\} \subset \mathbb{R}^n$, $\{\alpha_k\}, \{\lambda_k\} \subset I\!R$ such that, for each k, $\lambda_k \in \mathbb{R}_+$, $(u_k, \alpha_k) \in \mathrm{epi}(\lambda_k g)^*$ with $\lim_{k\to\infty} u_k = u$ and $\lim_{k\to\infty} \alpha_k = \alpha$. So, $(\lambda_k g)^*(u_k) \leq \alpha_k$. Equivalently, for each $x \in \mathbb{R}^n$, $u_k^T x - \lambda_k g(x) \leq \alpha_k$. Now, for each $x \in \mathbb{R}^n$ with $g(x) \leq 0$, $\lambda_k g(x) \leq 0$, and so, $u_k^T x \leq \alpha_k$. Letting $k \to \infty$, $-u^T x + \alpha \geq 0$. Let $f(x) = -u^T x + \alpha \geq 0$. Then f is a quadratic function and so, by the assumption, there exists $\mu \geq 0$ such that $-u^T x + \alpha + \mu g(x) \geq 0$. Thus, $u^T x - \mu g(x) \leq \alpha$. This gives us that $(\mu g)^*(u) \leq \alpha$ which means that $(u, \alpha) \in \mathrm{epi}(\mu g)^* \subset \bigcup_{\lambda \geq 0} \mathrm{epi}\ (\lambda g)^*$, and hence the set $\bigcup_{\lambda \geq 0} \mathrm{epi}\ (\lambda g)^*$ is closed.

[(ii) $\Rightarrow$ (iii)] Let $g(x) = x^T B x + b^T x + \beta$, where $B \in S^n, b \in I\!R^n$ and $\beta \in I\!R$. On the contrary to (ii), suppose that $g(x) \geq 0$ for all $x \in \mathbb{R}^n$. Then g is convex. Otherwise B is not positive semidefinite. Then, there exists v_0 such that $v_0^T B v_0 < 0$. So, $g(\gamma v_0) \to -\infty$ as $\gamma \to \infty$, which is not possible as $g(x) \geq 0$. Hence, $[g \leq 0] = [g = 0] : \{x \in I\!R^n \ : g(x) = 0\} \neq \emptyset$. Since g is convex and it attains its minimum at each point of $[g = 0]$,

$$[g \leq 0] = \{x : \nabla g(x) = 0\} = \{x : 2Bx + b = 0\}. \tag{9}$$

Let $\overline{x} \in [g = 0]$. Then, $\nabla g(\overline{x}) = 0$ and $[g \leq 0] = \bar{x} + \mathrm{Ker} B$. Moreover, it follows from (ii) and (7) that

$$N_{[g\leq 0]}(\overline{x}) = \bigcup_{\lambda \geq 0, \lambda g(\overline{x}) = 0} \{\lambda \nabla g(\overline{x})\} = \bigcup_{\lambda \geq 0} \{\lambda \nabla g(\overline{x})\} = \{0\}. \tag{10}$$

Now, (10) and (9) give us that

$$\{0\} = N_{[g\leq 0]}(\overline{x}) = N_{\overline{x} + \mathrm{Ker} B}(\overline{x}) = (\mathrm{Ker} B)^\circ.$$

So, $\mathrm{Ker}B = I\!R^n$, which gives us that $B = 0$. Then, $g(x) = b^T x + \beta \geq 0$, but $[g \leq 0] \neq \emptyset$, so, we obtain that $b = 0$ and $\beta = 0$. Thus $g \equiv 0$. This is a contradiction.

[(iii) $\Rightarrow$ (i)] This always holds by the S-lemma. $\square$

As an application of Theorem 3.1 we derive the following characterization of Lagrangian duality for a quadratic minimization over a single quadratic constraint in terms of Slater's condition.

Theorem 3.2 *Let g be a quadratic function that is not identically zero. Then, the following statements are equivalent:*
(i) *For each quadratic function f,*

$$\inf\{f(x) : g(x) \leq 0\} = \max_{\lambda \geq 0} \min_{x \in \mathbb{R}^n} \{f(x) + \lambda g(x)\}.$$

(ii) *There exists $x_0 \in \mathbb{R}^n$ such that $g(x_0) < 0$.*

Proof. [(i) $\Rightarrow$ (ii)] Suppose that the Slater condition fails. Then, by Theorem 3.1, there exists a quadratic function $f_0 : \mathbb{R}^n \to I\!R$, such that $g(x) \leq 0 \Rightarrow f_0(x) \geq 0$, and there exists $x_0 \in \mathbb{R}^n$ such that $f_0(x_0) + \lambda g(x_0) < 0, \ \forall \lambda \geq 0$. This implies that

$$\inf\{f_0(x) : g(x) \leq 0\} \geq 0 \text{ and } \max_{\lambda \geq 0} \min_{x \in \mathbb{R}^n} \{f_0(x) + \lambda g(x)\} < 0.$$

This contradicts (i).

[(ii) $\Rightarrow$ (i)]. Firstly, note that the following weak duality inequality

$$\inf\{f(x) : g(x) \leq 0\} \geq \max_{\lambda \geq 0} \min_{x \in \mathbb{R}^n} \{f(x) + \lambda g(x)\}$$

always holds. If $r := \inf\{f(x) : g(x) \leq 0\} = -\infty$, then (i) holds trivially. Without loss of generality, we assume that $r = \inf\{f(x) : g(x) \leq 0\} > -\infty$. Then, $g(x) \leq 0 \Rightarrow f(x) - r \geq 0$. It follows from Theorem 3.1 that there exists $\bar{\lambda} \geq 0$ such that $f(x) + \bar{\lambda} g(x) \geq r, \ \forall x \in \mathbb{R}^n$. Thus, $\max_{\lambda \geq 0} \min_{x \in \mathbb{R}^n} \{f(x) + \lambda g(x)\} \geq \min_{x \in \mathbb{R}^n} \{f(x) + \bar{\lambda} g(x)\} \geq r$. This gives that

$$\inf\{f(x) : g(x) \leq 0\} = \max_{\lambda \geq 0} \min_{x \in \mathbb{R}^n} \{f(x) + \lambda g(x)\}.$$

$\square$

The following example illustrates that the S-lemma fails due to the absence of the Slater condition. However, some regularized forms of S-lemma still hold.

Example 3.1 Let $f(x) = -2x$ and $g(x) = x^2$. Then, clearly,

$$g(x) \leq 0 \Rightarrow f(x) \geq 0.$$

But, there exists no $\lambda \geq 0$ such that $f(x) + \lambda g(x) \geq 0$ for all $x \in \mathbb{R}$, as $f(x) + \lambda g(x) = -2x + \lambda x^2 = x(\lambda x - 2)$, for any $\lambda \geq 0$. So, the S-lemma fails. However, the following regularized versions of the S-lemma hold:

(1) $(\exists \{\lambda_k\} \subseteq \mathbb{R}_+)\ d(H_f + \lambda_k H_g, S_+^{n+1}) \to 0.$

(2) $(\forall \epsilon > 0)(\exists \lambda_\epsilon \geq 0)(\forall x \in \mathbb{R}^n)\ f(x) + \lambda_\epsilon g(x) + \epsilon(\|x\|^2 + 1) \geq 0.$

Indeed, for each $k \in \mathbb{N}$, let $\lambda_k = k$ and

$$A_k := \begin{pmatrix} 2k & -1 \\ -1 & \frac{2}{k} \end{pmatrix} \in S_+^2.$$

Then,

$$H_f + \lambda_k H_g = \begin{pmatrix} 2k & -1 \\ -1 & 0 \end{pmatrix} \text{ and } d(H_f + \lambda_k H_g, S_+^2) \leq \|(H_f + \lambda_k H_g) - A_k\| = \frac{2}{k} \to 0.$$

Therefore, (1) holds. To see (2), let $\epsilon > 0$, choose $\lambda_\epsilon = 2(\frac{1}{\epsilon} - \epsilon)$ if $\epsilon \in (0,1)$, and $\lambda_\epsilon = 1$ if $\epsilon \geq 1$. Then, for each $\epsilon \in (0,1)$

$$\begin{aligned} f(x) + \lambda_\epsilon g(x) + \epsilon(|x|^2 + 1) &= -2x + (\frac{2}{\epsilon} - \epsilon)x^2 + \epsilon \\ &= (\frac{2}{\epsilon} - \epsilon)(x - \frac{\epsilon}{2 - \epsilon^2})^2 + \frac{\epsilon(1 - \epsilon^2)}{2 - \epsilon^2} \geq 0, \end{aligned}$$

and for each $\epsilon \geq 1$,

$$\begin{aligned} f(x) + \lambda_\epsilon g(x) + \epsilon(|x|^2 + 1) &= -2x + (1 + \epsilon)x^2 + \epsilon \\ &= (1 + \epsilon)(x - \frac{1}{1 + \epsilon})^2 + (\epsilon - \frac{1}{1 + \epsilon}) \geq 0. \end{aligned}$$

This example prompts us to examine a regularized version of the S-lemma without the Slater condition.

4 Regularized S-Lemma without Slater's Condition

In this section, we establish a regularized version of the S-lemma allowing applications without the Slater condition. Consider two non-homogeneous quadratic functions f, g, where $f(x) = x^T Ax + a^T x + \alpha$ and $g(x) = x^T Bx + b^T x + \beta$, where $A, B \in S^n, a, b \in \mathbb{R}^n, \alpha, \beta \in \mathbb{R}$.

Lemma 4.1 *Let $f,\ g : \mathbb{R}^n \to \mathbb{R}$ be quadratic functions with $[g \leq 0] \neq \emptyset$. Then, the following two statements are equivalent:*
(i) $(\exists \{\lambda_k\} \subseteq \mathbb{R}_+)\ d(H_f + \lambda_k H_g, S_+^{n+1}) \to 0$.
(ii) $(\forall \epsilon > 0)\ (\exists \lambda_\epsilon \geq 0)\ (\forall x \in \mathbb{R}^n)\ f(x) + \lambda_\epsilon g(x) + \epsilon(\|x\|^2 + 1) \geq 0$.

Proof. [(i) $\Rightarrow$ (ii)] Suppose that (i) holds. Let $\epsilon > 0$. Then, there exists $\lambda_\epsilon \geq 0$ such that

$$d(H_f + \lambda_\epsilon H_g, S_+^{n+1}) \leq \epsilon.$$

Thus, there exists $A_\epsilon \in S_+^{n+1}$ such that $\|(H_f + \lambda_\epsilon H_g) - A_\epsilon\| \leq \epsilon$. This implies that

$$\begin{aligned} f(x) + \lambda_\epsilon g(x) &= \begin{pmatrix} x \\ 1 \end{pmatrix}^T \left(H_f + \lambda_\epsilon H_g\right) \begin{pmatrix} x \\ 1 \end{pmatrix} \\ &= \begin{pmatrix} x \\ 1 \end{pmatrix}^T \left((H_f + \lambda_\epsilon H_g) - A_\epsilon\right) \begin{pmatrix} x \\ 1 \end{pmatrix} + \begin{pmatrix} x \\ 1 \end{pmatrix}^T A_\epsilon \begin{pmatrix} x \\ 1 \end{pmatrix} \\ &\geq -\epsilon(\|x\|^2 + 1). \end{aligned}$$

Thus (ii) holds.

[(ii) $\Rightarrow$ (i)] Suppose that (ii) holds. Let $\epsilon_k = \frac{1}{k},\ k \in \mathbb{N}$. Then, there exists $\lambda_k \subseteq \mathbb{R}_+$ such that $f(x) + \lambda_k g(x) + \epsilon_k(\|x\|^2 + 1) \geq 0$ for each $x \in \mathbb{R}^n$. This implies that $H_f + \lambda_k H_g + \epsilon_k I_{n+1} \succeq 0$. Therefore, $d(H_f + \lambda_k H_g, S_+^{n+1}) \leq \|\epsilon_k I_{n+1}\|$. Hence, (i) holds by letting $k \to \infty$. $\square$

Theorem 4.1 *(Regularized S-lemma) Let $f,\ g : \mathbb{R}^n \to \mathbb{R}$ be quadratic functions with $[g \leq 0] \neq \emptyset$. Then, the following two statements are equivalent:*
(i) $g(x) \leq 0 \ \Rightarrow\ f(x) \geq 0$.
(ii) $(\forall \epsilon > 0)\ (\exists \lambda_\epsilon \geq 0)\ (\forall x \in \mathbb{R}^n)\ f(x) + \lambda_\epsilon g(x) + \epsilon(\|x\|^2 + 1) \geq 0$.

Proof. [(ii) $\Rightarrow$ (i)]. Let $x \in [g \leq 0]$. Let $\epsilon > 0$. Then, there exists $\lambda_\epsilon \geq 0$ such that

$$f(x) \geq \lambda_\epsilon g(x) - \epsilon(\|x\|^2 + 1) \geq -\epsilon(\|x\|^2 + 1).$$

Letting $\epsilon \to 0$, we obtain that $f(x) \geq 0$. Hence, (i) holds.

[(i) $\Rightarrow$ (ii)] We consider two cases:

Case 1: Let $f,\ g$ be homogeneous functions defined by $f(x) = x^T Ax$ and $g(x) = x^T Bx$ where $A, B \in S^n$.

If $n = 2$, then we have $(0,0) \notin \{(f(x), g(x)) : x \in \mathbb{R}^2\} + \mathrm{int}\mathbb{R}_+ \times \mathbb{R}_+$ (otherwise, there exists x_0 such that $f(x_0) < 0$ and $g(x_0) \leq 0$ which contradicts to (i)). Then, by Lemma 2.2, $(0,0) \notin \{(A \cdot X, B \cdot X) : X \in S^2_+\} + \mathrm{int}\mathbb{R}_+ \times \mathbb{R}_+$. In other words, $X \in S^2_+,\ B \cdot X \leq 0 \Rightarrow A \cdot X \geq 0$. This implies that $-A \in \{X \in S^2_+ : B \cdot X \leq 0\}^\circ$. Note from (2) that $\{X \in S^2_+ : B \cdot X \leq 0\}^\circ = \overline{-S^2_+ + \bigcup_{\lambda \geq 0} \lambda B}$. It now follows that, for each $\epsilon > 0$, there exists $\lambda_\epsilon \geq 0$ and $P_\epsilon \in S^2_+$ such that $\|(A + \lambda_\epsilon B) - P_\epsilon\| \leq \epsilon$. Thus for each $x \in \mathbb{R}^n$

$$f(x) + \lambda_\epsilon g(x) = x^T(A + \lambda_\epsilon B)x = x^T P_\epsilon x + x^T\big((A + \lambda_\epsilon B) - P_\epsilon\big)x \geq -\epsilon\|x\|^2.$$

Thus, (ii) holds in this case.

Let $n \neq 2$. Define a set $\mathcal{V} \subseteq \mathbb{R}^2$ by

$$\mathcal{V} := \{(f(x), g(x)) \mid \|x\| = 1\} + \mathbb{R}^2_+.$$

Then $\mathcal{V}$ is a closed convex set. Indeed, if $n = 1$ then $\{(f(x), g(x)) \mid \|x\| = 1\}$ is a singleton. Then $\mathcal{V}$ is clearly closed and convex. If $n \geq 3$. it follows from Brickman's theorem (Lemma 2.1) that $\mathcal{V}$ is also closed and convex, as the Minkowski sum of a compact convex set and a closed convex set is closed and convex.

Fix $\epsilon > 0$. Then $(-\epsilon, 0) \notin \mathcal{V}$. Otherwise, there exists $\|x_0\| = 1$ such that $f(x_0) \leq -\epsilon < 0$ and $g(x_0) \leq 0$, which contradicts (i). Then, by the strict separation theorem, there exist $r \in \mathbb{R}$, $(\lambda_1, \lambda_2) \in \mathbb{R}^2 \backslash \{(0,0)\}$ such that for all $(a_1, a_2) \in \mathcal{V}$

$$-\lambda_1 \epsilon < r < \lambda_1 a_1 + \lambda_2 a_2.$$

In particular, one has $(\lambda_1, \lambda_2) \in \mathbb{R}^2_+ \backslash \{(0,0)\}$ and for all $x \in \mathbb{R}^n$ with $\|x\| = 1$

$$-\lambda_1 \epsilon < r < \lambda_1 f(x) + \lambda_2 g(x). \tag{11}$$

Next, we show that there exists $\lambda_\epsilon \geq 0$ such that

$$A + \lambda_\epsilon B + \epsilon I_n \succeq 0. \tag{12}$$

Granting this we obtain that $f(x) + \lambda_\epsilon g(x) + \epsilon\|x\|^2 \geq 0$ for all $x \in \mathbb{R}^n$. Thus (ii) holds. To see the claim, we consider two subcases:

Subcase 1: If $\lambda_1 = 0$, then $\lambda_2 > 0$. It follows from (11) that for each $x \in \mathbb{R}^n$ with $\|x\| = 1$, $x^T Bx = g(x) > \frac{r}{\lambda_2} > 0$. Thus B is positive definite and hence there exists λ_0 (large enough) such that $A + \lambda_0 B \succeq 0$. Therefore, the claim holds by setting $\lambda_\epsilon = \lambda_0$.

Subcase 2: Suppose that $\lambda_1 > 0$. It follows from (11) that for each $x \in \mathbb{R}^n$ with $\|x\| = 1$

$$0 \leq \lambda_1 f(x) + \lambda_2 g(x) + \lambda_1 \epsilon = \lambda_1 f(x) + \lambda_2 g(x) + \lambda_1 \epsilon \|x\|^2.$$

Thus, for each $x \in \mathbb{R}^n$, $f(x) + \frac{\lambda_2}{\lambda_1} g(x) + \epsilon \|x\|^2 \geq 0$. Setting $\lambda_\epsilon = \frac{\lambda_2}{\lambda_1}$, it follows that $A + \lambda_\epsilon B + \epsilon I_n \succeq 0$ and the claim follows.

Case 2: Let $f(x) = x^T Ax + a^T x + \alpha$ and $g(x) = x^T Bx + b^T x + \beta$. Consider the following homogeneous functions over $\mathbb{R}^{n+1}$, generated by f and g respectively:

$$f(x, \rho) = x^T Ax + \rho a^T x + \rho^2 \alpha = \begin{pmatrix} x \\ \rho \end{pmatrix}^T H_f \begin{pmatrix} x \\ \rho \end{pmatrix}, \tag{13}$$

$$g(x, \rho) = x^T Bx + \rho b^T x + \rho^2 \beta = \begin{pmatrix} x \\ \rho \end{pmatrix}^T H_g \begin{pmatrix} x \\ \rho \end{pmatrix}. \tag{14}$$

Next, we claim that $g(x, \rho) \leq 0 \Rightarrow f(x, \rho) \geq 0$. Granting this, by Case 1, we obtain that $(\forall \epsilon > 0)$, $(\exists \lambda_\epsilon \geq 0)$ $(\forall (x, \rho) \in \mathbb{R}^{n+1})$ $f(x, \rho) + \lambda_\epsilon g(x, \rho) + \frac{\epsilon}{2}(\|(x, \rho)\|^2 + 1) \geq 0$. Letting $\rho = 1$, we get that $(\forall \epsilon > 0)$ $(\exists \lambda_\epsilon \geq 0)$ $(\forall x \in \mathbb{R}^n)$ $f(x) + \lambda_\epsilon g(x) \geq -\frac{\epsilon}{2}(\|x\|^2 + 2) \geq -\epsilon(\|x\|^2 + 1) \geq 0$. Thus (ii) holds.

To see our claim, we proceed by the method of contradiction. Suppose that there exists $(x_0, \rho_0) \in \mathbb{R}^n \times \mathbb{R}$ such that $f(x_0, \rho_0) < 0$ and $g(x_0, \rho_0) \leq 0$,. If $\rho \neq 0$, then $f(\frac{x_0}{\rho_0}) = \rho_0^{-2} f(x_0, \rho_0) < 0$ and $g(\frac{x_0}{\rho_0}) = \rho_0^{-2} g(x_0, \rho_0) \leq 0$ which contradicts to assumption (i). If $\rho = 0$, then $x_0^T Ax_0 < 0$ and $x_0^T Bx_0 \leq 0$. Now, fix an $\overline{x}$ with $g(\overline{x}) \leq 0$. It can be verify that, for all $t \in \mathbb{R}$,

$$f(tx_0 + \overline{x}) = t^2 x_0^T Ax_0 + 2t(a + A\overline{x})^T x_0 + f(\overline{x}) \text{ and } g(tx_0 + \overline{x}) = t^2 x_0^T Bx_0 + 2t(b + B\overline{x})^T x_0 + g(\overline{x}).$$

We now split the discussion into two subcases: Subcase 1: $(b + B\overline{x})^T x_0 \leq 0$; Subcase 2: $(b + B\overline{x})^T x_0 \geq 0$.

Suppose that subcase 1 holds. Since $x_0^T Ax_0 < 0$, for all $t > 0$ large enough one has $f(tx_0 + \overline{x}) < 0$. Moreover, since $x_0^T Bx_0 \leq 0$, $(b + B\overline{x})^T x_0 \leq 0$ and $g(\overline{x}) \leq 0$

$$g(tx_0 + \overline{x}) = t^2 x_0^T Bx_0 + 2t(b + B\overline{x})^T x_0 + g(\overline{x}) \leq 0.$$

This contradicts (i).

Suppose that the subcase 2 holds. Then, similarly, consider the point $-tx_0 + \overline{x}$ with $t > 0$ large enough. We obtain that $f(-tx_0 + \overline{x}) < 0$ and $g(-tx_0 + \overline{x}) \leq 0$. This also contradicts (i). $\square$

In the following corollary, we see that, in Theorem 4.1, if we further assume the Slater condition holds then the regularized S-lemma collapses to the standard S-lemma.

Corollary 4.1 *(S-lemma) Let $f,\ g : \mathbb{R}^n \to \mathbb{R}$ be quadratic functions, defined by $f(x) = x^T Ax + a^T x + \alpha$ and $g(x) = x^T Bx + b^T x + \beta$, $A, B \in S^n$, $a, b \in \mathbb{R}^n$ and $\alpha, \beta \in \mathbb{R}$. Let H_f, H_g be defined as in (8). Suppose that $[g \leq 0] \neq \emptyset$ and there exists $x_0 \in \mathbb{R}^n$ such that $g(x_0) < 0$. Then, the following statements are equivalent:*

(i) $g(x) \leq 0 \ \Rightarrow\ f(x) \geq 0$.

(ii) $(\exists\, \lambda \geq 0)\ f(x) + \lambda g(x) \geq 0.$

(iii) $(\exists\, \lambda \geq 0)\ H_f + \lambda H_g \succeq 0.$

(iv) $(\forall\, \epsilon > 0)\ (\exists\, \lambda_\epsilon \geq 0)\ (\forall x \in \mathbb{R}^n)\ f(x) + \lambda_\epsilon g(x) + \epsilon(\|x\|^2 + 1) \geq 0.$

Proof. Note that (ii) $\Leftrightarrow$ (iii) and (iii) $\Rightarrow$ (iv) always holds (by Lemma 4.1). Since (i) $\Leftrightarrow$ (iv) (by Theorem 4.1), it suffices to show that (iv) $\Rightarrow$ (iii).

[(iv) $\Rightarrow$ (iii)] Suppose that (iv) holds. By Lemma 4.1, $\exists\, \{\lambda_k\} \subseteq \mathbb{R}_+$, $d(H_f + \lambda_k H_g, S_+^{n+1}) \to 0$. This implies that

$$H_f \in \overline{S_+^{n+1} + \bigcup_{\lambda \geq 0} \lambda(-H_g)}.$$

Next, we show that the set $S_+^{n+1} + \bigcup_{\lambda \geq 0} \lambda(-H_g)$ is indeed closed. To see this, let $Z_k = P_k + \lambda_k(-H_g)$ with $Z_k \to Z$ where $P_k \in S_+^{n+1}$ and $\lambda_k \geq 0$. Let $X_0 = \begin{pmatrix} x_0 \\ 1 \end{pmatrix} \begin{pmatrix} x_0 \\ 1 \end{pmatrix}^T$, where $x_0 \in I\!R^n$. Then

$$0 \leq \begin{pmatrix} x_0 \\ 1 \end{pmatrix}^T P_k \begin{pmatrix} x_0 \\ 1 \end{pmatrix} = P_k \cdot X_0 = (Z_k + \lambda_k H_g) \cdot X_0 = Z_k \cdot X_0 + \lambda_k g(x_0).$$

Since $g(x_0) < 0$, it follows that $0 \leq \lambda_k \leq \frac{Z_k \cdot X_0}{-g(x_0)}$. Since $Z_k \to Z$ (and hence bounded), we see that λ_k is bounded. By passing to subsequence, we have $\lambda_k \to \lambda_0 \geq 0$. Thus $P_k = Z_k + \lambda_k H_g \to Z + \lambda_0 H_g \in S_+^{n+1}$, and so,

$$Z = (Z + \lambda_0 H_g) + \lambda_0(-H_g) \in S_+^{n+1} + \bigcup_{\lambda \geq 0} \lambda(-H_g).$$

Thus $S_+^{n+1} + \bigcup_{\lambda \geq 0} \lambda(-H_g)$ is closed. Therefore, one has

$$H_f \in S_+^{n+1} + \bigcup_{\lambda \geq 0} \lambda(-H_g).$$

Then, there exists $\lambda \geq 0$ such that $H_f + \lambda H_g \succeq 0$. $\square$

Consider the quadratic minimization problem:

$$(QP) \qquad \min_{x \in \mathbb{R}^n} \; f(x) \quad \text{s.t.} \quad g(x) \leq 0,$$

where f and g are quadratic functions defined by $f(x) = x^T Ax + a^T x + \alpha$ and $g(x) = x^T Bx + b^T x + \beta$, where $A, B \in S^n, a, b \in \mathbb{R}^n, \alpha, \beta \in \mathbb{R}$.

The following lemma on the relationship between subdifferentials and ϵ-subdifferentials will be used later to derive a sequential characterization of global optimality of (QP) without a constraint qualification.

Lemma 4.2 (Borwein [23, Theorem 3.1.1]) *Let f be a continuous convex function on $\mathbb{R}^n$. Let $\epsilon > 0$, $\beta \geq 0$ and $x_0^* \in \partial_\epsilon f(x_0)$. Then there exist $x_\epsilon, x_\epsilon^* \in \mathbb{R}^n$ such that*

$$x_\epsilon^* \in \partial f(x_\epsilon), \;\; \|x_\epsilon - x_0\| \leq \sqrt{\epsilon}$$

$$|f(x_\epsilon) - f(x_0)| \leq \sqrt{\epsilon}(\sqrt{\epsilon} + \beta^{-1}) \;\; \text{and} \;\; \|x_\epsilon^* - x_0^*\| \leq \sqrt{\epsilon}(1 + \beta\|x_0^*\|).$$

Theorem 4.2 *(Necessary and Sufficient Conditions for Global Optimality) For (QP), let $\overline{x} \in [g \leq 0]$. Then the following statements are equivalent:*

(i) *$\overline{x}$ is a global minimizer of (QP).*

(ii) *There exist $\{x_k\} \subset I\!R^n$ and $\{\lambda_k\} \subseteq \mathbb{R}_+$ such that*

$$x_k - \overline{x} \to 0, \; \lambda_k g(x_k) \to 0, \; \nabla(f + \lambda_k g)(x_k) \to 0 \;\; \text{and} \;\; d(A + \lambda_k B, S_+^n) \to 0.$$

Proof. [(i) $\Rightarrow$ (ii)]. Let $\overline{x}$ be a global minimizer of (QP). Then, $g(x) \leq 0 \;\Rightarrow\; f(x) - f(\overline{x}) \geq 0$. Let $\epsilon_k = \frac{1}{k}$, $k \in I\!N$. Then by Theorem 4.1 there exist $\lambda_k \subseteq \mathbb{R}_+$ such that for each $x \in \mathbb{R}^n$

$$h(x) := f(x) - f(\overline{x}) + \lambda_k g(x) + \epsilon_k(\|x\|^2 + 1) \geq 0. \tag{15}$$

In particular, one has

$$0 \geq \lambda_k g(\overline{x}) \geq -\epsilon_k \|\overline{x}\|^2 - \epsilon_k. \tag{16}$$

Note that as for each $x \in I\!R^n$, $h(x) \geq 0$,

$$\nabla^2 h = 2(A + \lambda_k B + \epsilon_k I_n) \succeq 0.$$

This implies that $d(A + \lambda_k B, S^n_+) \leq \|\epsilon_k I_n\| \to 0$, as $k \to \infty$.

On the other hand, from (16) and (15), for each $x \in \mathbb{R}^n$,

$$h(x) - h(\overline{x}) = \big(f(x) - f(\overline{x}) + \lambda_k g(x) + \epsilon_k \|x\|^2\big) - (\lambda_k g(\overline{x}) + \epsilon_k \|\overline{x}\|^2) \geq -\epsilon_k - \epsilon_k \|\overline{x}\|^2.$$

Define $\eta_k = \epsilon_k + \epsilon_k \|\overline{x}\|^2 > 0$. Then, $0 \in \partial_{\eta_k} h(\overline{x})$ and $\eta_k \to 0$ as $k \to \infty$. From Borwein's Theorem (letting $\beta = 1$, $\epsilon = \eta_k$ and $x^*_0 = 0$), there exist x_k such that $\|x_k - \overline{x}\| \leq \sqrt{\eta_k}$, $|h(\overline{x}) - h(x_k)| \leq \eta_k + \sqrt{\eta_k}$ and

$$0 \in \partial h(x_k) + \sqrt{\eta_k}\mathbb{B} = \nabla(f + \lambda_k g)(x_k) + 2\epsilon_k x_k + \sqrt{\eta_k}\mathbb{B},$$

where $\mathbb{B}$ denotes the closed unit ball in $\mathbb{R}^n$. Since $\epsilon_k \to 0$ and $\eta_k \to 0$, as $k \to \infty$, $x_k \to \overline{x}$, $\nabla(f + \lambda_k g)(x_k) \to 0$. To finish the proof, it suffices to show that $\lim_{k\to\infty} \lambda_k g(x_k) = 0$. To see this, since $|h(\overline{x}) - h(x_k)| \leq \eta_k + \sqrt{\eta_k}$, one has

$$|(\lambda_k g(\overline{x}) + \epsilon_k \|\overline{x}\|^2) - (f(x_k) - f(\overline{x}) + \lambda_k g(x_k) + \epsilon_k \|x_k\|^2)| \leq \eta_k + \sqrt{\eta_k}.$$

This together with $x_k \to \overline{x}$, $\epsilon_k \to 0$ and $\eta_k \to 0$ implies that $\lim_{k\to\infty} \lambda_k (g(\overline{x}) - g(x_k)) = 0$. It follows from (16) that

$$\lim_{k\to\infty} \lambda_k g(x_k) = \lim_{k\to\infty} \lambda_k g(\overline{x}) = 0.$$

[(ii) $\Rightarrow$ (i)] We proceed by the method of contradiction. Suppose that there exists x_0 such that $g(x_0) \leq 0$ and $f(x_0) < f(\overline{x})$. From (ii), there exist $x_k \to \overline{x}$, $\lambda_k \subseteq \mathbb{R}_+$ with $\lambda_k g(x_k) \to 0$ such that

$$\nabla(f + \lambda_k g)(x_k) \to 0 \text{ and } d(A + \lambda_k B, S^n_+) \to 0.$$

Note that

$$\begin{aligned}
f(x_0) - f(\overline{x}) &= \big(f(x_0) + \lambda_k g(x_0)\big) - \big(f(x_k) + \lambda_k g(x_k)\big) + (f(x_k) - f(\overline{x})) + \lambda_k g(x_k) - \lambda_k g(x_0) \\
&\geq \big(f(x_0) + \lambda_k g(x_0)\big) - \big(f(x_k) + \lambda_k g(x_k)\big) + (f(x_k) - f(\overline{x})) + \lambda_k g(x_k) \\
&= \big(\nabla(f + \lambda_k g)(x_k)\big)^T (x_0 - x_k) + \frac{1}{2}(x_0 - x_k)^T (A + \lambda_k B)(x_0 - x_k) \\
&\quad + (f(x_k) - f(\overline{x})) + \lambda_k g(x_k),
\end{aligned}$$

where the inequality holds since $\lambda_k \geq 0$ and $g(x_0) \leq 0$. Now, since $x_k \to \overline{x}$, $\nabla(f + \lambda_k g)(x_k) \to 0$, $\lambda_k g(x_k) \to 0$ and $f(x_0) < f(\overline{x})$, by passing to upper limit, we have

$$\limsup_{k\to\infty} \frac{1}{2}(x_0 - x_k)^T (A + \lambda_k B)(x_0 - x_k) \leq f(x_0) - f(\overline{x}) < 0. \tag{17}$$

On the other hand, since $d(A+\lambda_k B, S^n_+) \to 0$, there exist $P_k \in S^n_+$ such that $\|(A+\lambda_k B) - P_k\| \to 0$. Thus, we have

$$\begin{aligned}
&\limsup_{k\to\infty} \frac{1}{2}(x_0 - x_k)^T(A+\lambda_k B)(x_0 - x_k) \\
= \;& \limsup_{k\to\infty} \left(\frac{1}{2}(x_0 - x_k)^T(A+\lambda_k B - P_k)(x_0 - x_k) + \frac{1}{2}(x_0 - x_k)^T P_k (x_0 - x_k) \right) \\
\geq \;& \limsup_{k\to\infty} \frac{1}{2}(x_0 - x_k)^T(A+\lambda_k B - P_k)(x_0 - x_k) = 0.
\end{aligned}$$

This contradicts (17). □

References

[1] A. Ben-Tal and A. Nemirovski, *Lectures on Modern Convex Optimization: Analysis, Algorithms and Engineering Applications,* SIAM-MPS, Philadelphia, 2000.

[2] S. Boyd, L. El-Ghaoul, E. Feron and V. Balakhrishnan, *Linear Matrix Inequalities in Systems and Control Theory*, SIAM, Philadelphia, 1994.

[3] L. Brickman, *On the field of values of a matrix*, Proceedings of the American Mathematical Society, 12 (1961), 61-66.

[4] R. S. Burachik and V. Jeyakumar, *A new geometric condition for Fenchel duality in infinite dimensions*, Mathematical Programming, Ser B, **104** (2005), 229-233.

[5] K. Derinkuyu and M. C. Pinar, *On the S-procedure and some variants*, Mathematical Methods of Operations Research **64** (2006), 55–77.

[6] L. L. Dines, *On the mapping of quadratic forms*, Bullentin of the American Mathematical Society **47** (1941), 494–498.

[7] N. Dinh, V. Jeyakumar and G. M. Lee, *Sequential Lagrangian conditions for convex programs with applications to semidefinite programming*, Journal of Optimization Theory and Applications **125** (2005), 85–112.

[8] N. Dinh, M.A. Goberna, M.A López, *From linear to convex systems: Consistency, Farkas' Lemma and application*, Journal of Convex Analysis **13** (2006), 1–21.

[9] V. Jeyakumar, *Farkas Lemma: Generalizations*, Encyclopedia of Optimization, Vol. 2, Kluwer Academic Publishers, Boston, USA, 2001, 87-91.

[10] V. Jeyakumar and G. M. Lee, *Complete characterizations of stable Farkas lemma and cone-convex programming duality*, Mathematical Programming, Series A, **114**, (2008), 335-347.

[11] V. Jeyakumar, Z. Y. Wu, G. M. Lee and N. Dinh, *Liberating the subgradient optimality conditions from constraint qualifications,* Journal of Global Optimization, 36 (2006), 127–137.

[12] V. Jeyakumar, *The strong conical hull intersection property for convex programming*, Mathematical programming, Series A, **106**(2006), 81–92.

[13] V. Jeyakumar and B. M. Glover, *Nonlinear extensions of Farkas lemma with applications to global optimization and least squares*, Mathematics of Operations Research **20** (1995), 818–837.

[14] V. Jeyakumar, G. M. Lee and N. Dinh, *New sequential Lagrange multiplier conditions characterizing optimality without constraint qualifications*, SIAM Journal on Optimization, **14** (2003), 534-547.

[15] V. Jeyakumar, A. M. Rubinov, B. M. Glover and Y. Ishizuka, *Inequality systems and global optimization*, Journal of Mathematical Analysis and Applications, **202**(1996), 900-919.

[16] V. Jeyakumar, A. M. Rubinov and Z. Y. Wu, *Nonconvex quadratic minimization with quadratic constraints: Global optimality conditions*, Mathematical Programming, Ser. A, **110** (2007), 521–541

[17] Z.-Q. Luo, J. Sturm and S. S. Zhang, *Multivariate nonnegative quadratic mappings*, SIAM Journal on Optimization, **14** (2004), 1140–1162.

[18] J. E. Martínez-Legaz, *On Brickman's theorem*, Journal of Convex Analysis, 12 (2005), 139–143.

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

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

[21] J. F. Sturm and S. Z. Zhang, On cones of nonnegative quadratic functions, *Math. Oper. Res.,* **28**, (2003) , 246-267.

[22] V. A. Yakubovich, *The S-procedure in nonlinear control theory*, Vestnik Leningr. Univ., **4** (1977), 73-93.

[23] C. Zalinescu, *Convex Analysis in General Vector Space*, World Scientific Publishing Co. Pty. Ltd, Singapore, 2002.