Conditions for Existence of Dual Certificates in Rank-One Semidefinite Problems

Paul Hand *

November 13, 2013

Abstract

We study the existence of dual certificates in convex minimization problems where a matrix $X_0$ is to be recovered under semidefinite and linear constraints. Dual certificates exist if and only if the problem satisfies strong duality. In the case that $X_0$ is rank 1, dual certificates are guaranteed to exist if there is nothing in the span of the linear measurement matrices that is positive and orthogonal to $X_0$. If there are such matrices in the span, then strong duality may fail and a dual certificate may not exist. We present a completeness condition on the measurement matrices and prove dual certificate existence if this completeness condition holds. If the condition fails, then the convex program can be supplemented with additional linearly independent measurements, resulting in an equivalent program that is guaranteed to have a dual certificate at the minimizer. If the set of linear measurements is not complete in the way described, we prove there is a convex program for which a dual certificate does not exist. This result informs the search space for the analytical construction of dual certificates in rank-one matrix completion problems.

1 Introduction

We consider the problem of showing that $X_0$ minimizes the semidefinite program

$$\min f(X) \text{ subject to } X \succeq 0, A(X) = b, \tag{1}$$

where $X \in S_n$ is a symmetric and real-valued matrix, $f$ is convex and continuous, $A$ is linear, and $A(X_0) = b \in \mathbb{R}^m$. Let $\langle X, Y \rangle = \text{tr}(Y^*X)$ be the Hilbert-Schmidt inner product. Matrix orthogonality is understood to be with respect to this inner product. The linear measurements $A(X) = b$ can be written as

$$A(X)_i = \langle X, A_i \rangle = b_i \text{ for } i = 1, \ldots, m,$$

for certain symmetric matrices $A_i$. Note that the adjoint of $A$ is $A^* \lambda = \sum_{i=1}^m \lambda_i A_i$.

A case of particular interest is when $X_0 = x_0x_0^*$ is rank one. For example, the phase retrieval problem motivated by X-ray crystallography can be written as a rank one matrix recovery problem through a process called lifting [4]. That matrix recovery problem is known as PhaseLift [7], and it is of form (1) with $f(X) = \text{tr}(X)$ and $A_i = zz_i^*$ for vectors $z_i$. Another example is the compressive phase retrieval problem [13, 15], where $X_0$ is additionally presumed to be sparse. In this case $f(X) = \|X\|_1 + c \text{ tr}(X)$, where the first term is the entry-wise $\ell^1$ norm of $X$. These programs are semidefinite instances of low-rank matrix recovery problems [6, 11].

*Massachusetts Institute of Technology, 77 Massachusetts Avenue E18-473, Cambridge, MA 02139, (hand@math.mit.edu).
1.1 Dual Certificates and Strong Duality

A common method for proving that a particular $X_0$ is a minimizer of a matrix recovery problem is certification [11, 7, 5, 9, 13, 1]. A certificate at $X_0$ defines a hyperplane that separates the feasible set of parameters from the set of points with smaller objectives. Under mild assumptions on $f$, a certificate necessarily exist at a minimizer. Finding a certificate guarantees that $X_0$ is a minimizer.

In many cases, the certificate can be represented in terms of the problem’s Lagrangian dual variables. Such certificates are known as dual certificates. The Lagrangian of (1) is given by

$$L(X, \lambda, Q) = f(X) + \langle \lambda, A(X) - b \rangle + \langle Q, X \rangle.$$  

where $Q \in S_n$ and $\lambda \in \mathbb{R}^m$. The Lagrangian dual problem for (1) is

$$\sup_{Q \preceq 0, \lambda} \inf_X L(X, \lambda, Q) \quad (2)$$

The dual variables $(\lambda, Q)$ are dual-feasible when $Q \preceq 0$. Let $p^*$ and $d^*$ be the optimal values of (1) and (2), respectively. The duality gap is $p^* - d^*$, which is always nonnegative. Problem (1) is said to satisfy strong duality when the duality gap is zero and the dual optimum is attained.

We call $(\lambda, Q)$ a dual certificate at $X_0$ if

$$A^*\lambda + Q \in -\partial f(X_0),$$  
$$Q \preceq 0, Q \perp X_0.$$  

That is, a dual certificate $(\lambda, Q)$ solves the KKT optimality conditions [3]. Observe that the optimality condition (3) ensures that 0 is in the subdifferential of $L$ with respect to $X$ at $X_0$. Conditions (4) enforce dual-feasibility and complementary slackness. Sometimes, we will refer to $Y = A^*\lambda + Q$ as the dual certificate.

By elementary arguments from convex optimization, a dual certificate at $X_0$ certifies that $X_0$ is a minimizer. Further, existence of a dual certificate is equivalent to strong duality.

**Theorem 1.** If $(\lambda, Q)$ is a dual certificate at $X_0$, then $X_0$ is a minimizer to (1).

**Theorem 2.** Let $X_0$ be a minimizer of (1). The following are equivalent:

(a) $(\lambda, Q)$ is a dual certificate at $X_0$,

(b) $(\lambda, Q)$ is dual optimal and strong duality holds.

1.2 Dual Certificates May Not Exist

For some semidefinite problems of form (1), a dual certificate fails to exist at a minimizer. We provide an example. Let $e_i$ be the $i$th standard basis vector, let $y \otimes q = yq^* + qy^*$ be the symmetric tensor product, let $\|X\|_F$ be the Frobenius matrix norm, and note that

$$X \succeq 0 \text{ and } \langle X, qq^* \rangle = 0 \text{ for } q \in \mathbb{R}^n \implies Xq = 0 \implies \langle X, y \otimes q \rangle = 0 \forall y \in \mathbb{R}^n. \quad (5)$$

**Example 1.** Let $X_0 = e_1e_1^*$. There is no dual certificate at $X_0$ for the problem

$$\min \frac{1}{2}\|X\|_F^2 \text{ subject to } X \succeq 0, \langle X, \begin{pmatrix} 0 & 0 \\ 0 & 1 \end{pmatrix} \rangle = 0, \langle X, \begin{pmatrix} 1 & 1 \\ 1 & 1 \end{pmatrix} \rangle = 1. \quad (6)$$
By (5), any feasible $X$ for Example 1 satisfies $\langle X, e_1 \otimes e_2 \rangle = 0$. Hence, the minimizer and only feasible point of (6) is $X_0$. The subdifferential of $f(X) = \frac{1}{2} \|X\|_F^2$ is $\partial f(X_0) = \{X_0\}$. We note that there is no dual certificate because there is no $(\lambda, Q)$ satisfying $Q \preceq 0$, $Q \perp X_0$, and

$$\begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix} = \lambda_1 \begin{pmatrix} 0 & 0 \\ 0 & 1 \end{pmatrix} + \lambda_2 \begin{pmatrix} 1 & 1 \\ 1 & 1 \end{pmatrix} + Q.$$

In this problem, one can show that there is no duality gap. Hence, the dual optimum is not attained.

Note that if (6) were supplemented with the constraint $\langle X, e_1 \otimes e_2 \rangle = 0$, there would exist a dual certificate at $X_0$.

1.3 Constraint Qualifications

The counterexample of the last section illustrates the well known fact that semidefinite programs may not satisfy strong duality [18, 14, 10]. A constraint qualification (CQ) is a condition such that strong duality is ensured. For example, the presence of a strictly feasible $X \succ 0$ such that $A(X) = b$ is a constraint qualification and is known as Slater’s condition [3]. Slater’s condition can fail for low-rank matrix recovery problems. For instance, if there is a measurement matrix that is positive semidefinite and orthogonal to $X_0$, then there are necessarily no strictly feasibility points. This is the case in Example 1.

The work in this paper will be based of the following constraint qualification. The Rockafellar-Pshenichnyi condition [14, 12, 20, 16] in the present context is that $X_0$ minimizes (1) if and only if there exists a $Y \in (-\partial f(X_0)) \cap \partial I_{X \succeq 0, A(X) = b}(X_0)$, where $I_{X \succeq 0, A(X) = b}$ is the indicator function of the feasible set. Let the cone of candidate dual certificates be

$$S := \left\{ \sum_i \lambda_i A_i + Q \mid Q \preceq 0, Q \perp X_0 \right\} = \partial I_{X \succeq 0}(X_0) + \partial I_{A(X) = b}(X_0).$$

A constraint qualification is thus that

$$\partial I_{X \succeq 0}(X_0) + \partial I_{A(X) = b}(X_0) = \partial I_{X \succeq 0, A(X) = b}(X_0).$$

This constraint qualification is known as a weakest constraint qualification because it is independent of the objective $f$ [14].

1.4 Main Results

In this paper, we interpret the Rockafellar-Pshenichnyi constraint qualification in the context of rank-one matrix recovery problems. We present conditions on the measurement matrices $A_i$ such that strong duality holds. We will also present a regularization procedure that adds implied measurements to (1) such that a dual certificate exists. The purpose of these results is primarily to aid matrix recovery proofs, as opposed to their computation. Knowing that strong duality holds ensures that the analytical construction of a dual certificate is not futile. Further, it allows negative results, where it is possible to prove that $X_0$ is not a minimizer of (1) by showing that no dual certificate can exist. See [13] for an example.

1.4.1 Sufficient conditions for dual certificate existence

The lack of a dual certificate in the example is because there is a linear combination of $A_i$ that is positive semidefinite and orthogonal to $X_0$. If this case is excluded, a dual certificate necessarily exists at the rank-one solution $X_0$. 

3
Theorem 3. If $X_0 = x_0x_0^*$ minimizes (1) and there does not exist a nonzero $A \in \text{span}\{A_i\}_{i=1}^m$ such that $A \succeq 0$ and $A \perp X_0$, then there exists a dual certificate at $X_0$.

This theorem can be proven by noting that the Theorem of the Alternative in Section 5.9.4 of [3] shows that Slater’s condition holds. For some measurement matrices $A_i$ of practical interest, Theorem 3 is applicable. For example, the following corollary implies a dual certificate exists with probability 1 if $X_0 = x_0x_0^*$ minimizes (1) with $A_i = z_iz_i^*$ for i.i.d. Gaussian $z_i$ and $m \leq n$. This is the case arising in the compressive phase retrieval problem [13].

Corollary 1. If $X_0 = x_0x_0^*$ minimizes (1) and $\{A_i,x_0\}_{i=1}^m$ are independent, then there exists a dual certificate at $X_0$.

If there is a positive semi-definite measurement matrix $A$ that is orthogonal to $X_0$, then [3] provides additional constraints on $X$ that may or may not be implied by the linear constraints $A(X) = b$ alone. For any $q \in \text{Range}(A)$, and for any $y$, all feasible $X$ satisfy $\langle X, y \otimes q \rangle = 0$. Hence $y \otimes q$ must be in $S$ in order for strong duality to hold. We say that $S$ is complete at $X_0$ if the following condition holds:

If $A = A^*\lambda \succeq 0, A \perp X_0$, then $y \otimes q \in S \forall y$ and $\forall q \in \text{Range}(A)$.  \hspace{1cm} (9)

The main theorem is as follows.

Theorem 4. Let $X_0 = x_0x_0^*$ minimize (1). If $S$ satisfies the completeness condition (9) then strong duality holds and a dual certificate exists at $X_0$.

Roughly, the theorem states that if all of the measurements implied by $X \succeq 0$ and $A(X) = b$ are included, then strong duality holds and a dual certificate exists. We note that completeness of $S$ does not imply that $y \otimes q \in \text{span}\{A_i\}$. Instead, completeness requires that $y \otimes q$ differs from $\text{span}\{A_i\}$ by something negative.

1.4.2 Regularization process

If the problem (1) does not satisfy the completeness condition (9), an optimality certificate for (1) can be expressed as a dual certificate for the problem augmented with linear constraints implied by $X \succeq 0$, $A(X) = b$. This completed problem is equivalent to (1) and its existence is a corollary of Theorem 4.

Corollary 2. If $X_0 = x_0x_0^*$ minimizes (1), then there exists an equivalent equivalent problem

$$\min f(X) \text{ such that } X \succeq 0, \tilde{A}(X) = \tilde{b}$$

such that there exists a dual certificate at $X_0$. This problem is equivalent to (1) in the sense that $X \succeq 0$ and $A(X) = b \iff X \succeq 0$ and $A(X) = b$.

The following procedure outlines a theoretical process to complete the set of measurement matrices $\{A_i\}$ in order to obtain an equivalent problem satisfying the completeness condition (9):

1. Consider all $A \succeq 0, A \in \text{span}\{A_i\}, \langle A, X_0 \rangle = 0$.
2. Write each $A = \sum_k c_k q_kq_k^*$ with $c_k > 0$.
3. For every $j$, if $e_j \otimes q_k \notin \text{span}\{A_i\}$, append $\langle X, e_j \otimes q_k \rangle = 0$ to $A(X) = b$. 
4. Repeat until $\mathcal{A}$ remains unchanged.

This process will produce a set $S$ satisfying (9), and it will terminate after finitely many repetitions because $\dim(\text{span}\{\mathcal{A}_i\})$ increases each time. Because the resulting $\hat{\mathcal{A}}$ will satisfy (9), we apply Theorem 4 and have thus proven Corollary 2. The semidefinite feasibility problem implicit to step 1 is of unknown computational complexity [17]. Hence, this procedure is of limited computational use. See [8] for computational preprocessing and regularization of semidefinite programs that fail Slater’s condition.

1.4.3 Weak necessity of the completeness condition

If the measurement matrices fail to satisfy the completeness condition (9), a particular problem of form (1) may or may not have a dual certificate at a minimizer $X_0$. Nonetheless, there is a problem for which a dual certificate does not exist.

**Theorem 5.** Fix $X_0 = x_0 x_0^*$ and the matrices $\{\mathcal{A}_i\}_{i=1}^m$. If $S$ does not satisfy the completeness condition (9), there exists a convex problem (1) such that $X_0$ is a minimizer and a dual certificate does not exist at $X_0$.

This weak form of necessity of the completeness condition arises in part because of an equivalence between completeness and the additivity of subgradients of indicator functions over the constraints, as proven in Lemma 1.

1.5 Discussion

The important message of this paper is that orthogonal measurements in semidefinite programs can give rise to situations where strong duality does not hold and a dual certificate does not exist. If a semidefinite program involves such measurements, then there are additional implied measurements that should be included when building a dual certificate. For example, if $\mathcal{A}(X) = b$ includes the measurement $\langle X, qq^* \rangle = 0$, then $\langle X, q e_i^t + e_i q^t \rangle = 0$ should be appended for all $i$, unless they are already implied by $\mathcal{A}(X) = b$.

We remark that the completeness condition (9) is only a sufficient condition for existence of dual certificates. For any particular problem, it may not be necessary. It is, however, necessary for some particular problem, as per Theorem 5.

We caution the reader of a subtlety of semidefinite programs of form (1). Consider a semidefinite program of form (1) that satisfies strong duality. Appending generic measurements to $\mathcal{A}(X) = b$ results in a program that may or may not satisfy strong duality. This is because these additional measurements may cause the completeness condition (9) to be unsatisfied. In this case, Theorem 4 is not applicable, and the program may or may not satisfy strong duality as written. To guarantee strong duality, a regularization process like above is needed to ensure (9) holds.

We now place the regularization procedure from Section 1.4.2 in context. Regularization is the modification of a semidefinite program or its dual in order to ensure strong duality. One approach for this is a minimal cone regularization [2, 18], where the conic constraint is modified. Another approach is the extended Lagrange-Slater Dual (ELSD), which is an alternative to the Lagrangian dual [17, 18]. It can be constructed in polynomially many additional variables. The regularization procedure in the present paper is different because it attains strong duality without changing the structure of the program. The conic constraint and overall form remain the same as only additional measurements are added. The procedure can not be written down mechanically. Hence it is not suitable for numerical computations. Instead, its simplicity in form makes it more useful for analytical constructions of dual certificates.
As stated, Theorem 4 is proven when the minimizer $X_0$ has rank one. It is an interesting problem to see if a corresponding result holds in the case of low rank $X_0$. Because this paper is motivated by vector recovery problems that are lifted to rank-one matrix recovery problems, this extension is left for future work.

1.6 Organization of this paper

In Section 1.7 we present the notation used throughout the paper. In Section 2 we prove Theorems 1 and 2 which are elementary results from convex optimization. In Section 3, we prove Theorems 3 and 5 and Corollary 1. Corollary 2 was proven in Section 1.4.2. The proofs of Theorems 3 and 4 rely on technical lemmas concerning the additivity of subdifferentials of indicator functions, and on the closedness of $S$. These lemmas are proven in Section 4.

1.7 Notation

Let $S_n$ be the space of symmetric, real-valued matrices of size $n$. Matrices will be denoted with boldface capital letters, and vectors will be denoted with boldface lowercase letters. Let $\langle \cdot , \cdot \rangle$ be the usual inner product for vectors and the Hilbert-Schmidt inner product for matrices. Let $X \preceq Y$.

Proof of Theorem 1. Consider a feasible $X$. Because $A^*X + Q \in - \partial f(X_0)$,

$$f(X) - f(X_0) \geq - \langle A^*X + Q, X - X_0 \rangle = - \langle Q, X \rangle \geq 0,$$

where the last equality uses $A(X) = A(X_0)$ and $Q \perp X_0$, and the last equality uses $Q \preceq 0$ and $X \succeq 0$. $\square$

Proof of Theorem 2. To prove (a) $\Rightarrow$ (b), we observe that (3) implies $0 \in \partial_X \mathcal{L}(X_0, \lambda, Q)$. Hence, $X_0$ minimizes $\mathcal{L}(X, \lambda, Q)$ over $X$. Hence $g(\lambda, Q) = \mathcal{L}(X_0, \lambda, Q)$. By slackness and feasibility of $X_0$, $\mathcal{L}(X_0, \lambda, Q) = f(X_0) = p^*$. Hence $(\lambda, Q)$ is dual optimal and strong duality holds.

To prove (b) $\Rightarrow$ (a), we observe that strong duality and dual optimality of $(\lambda, Q)$ imply

$$f(X_0) = \inf_{X} f(X) + \langle \lambda, A(X) - b \rangle + \langle Q, X \rangle$$

(11)

In particular, $f(X_0) \leq f(X_0) + \langle \lambda, A(X_0) - b \rangle + \langle Q, X_0 \rangle$, which implies $(Q, X_0) \geq 0$ by feasibility of $X_0$. By dual feasibility, $Q \preceq 0$ and hence $\langle Q, X_0 \rangle \leq 0$. We thus have $Q \perp X_0$. The infimum in (11) is achieved by $X_0$. Hence, $0 \in \partial_X \mathcal{L}(X_0, \lambda, Q)$, and we conclude $A^*X + Q \in - \partial f(X_0)$. $\square$
3 Proofs of Main Results

In this section, we present the proofs of the main theorems and Corollary 1.

3.1 Proof of Theorems 3 and 4 and Corollary 1

Under the assumptions of Theorem 3, the set $S$ trivially satisfies the completeness condition (9). The theorem is thus a special case of Theorem 4, and we will prove them together.

The strategy of proof involves rewriting (1) in an unconstrained form. Existence of a dual certificate is guaranteed when the subdifferentials of the sum of two indicator functions is the sum of their respective subdifferentials. In Lemma 1, we use a separating hyperplane argument to prove additivity of these subdifferentials under the condition (9).

Proof of Theorems 3 and 4. We first rewrite the problem (1) without constraints.

$$
\min f(\mathbf{X}) + I_{\mathbf{X}_{\geq 0}, A(\mathbf{X}) = \mathbf{b}}(\mathbf{X}),
$$

which, by convexity, happens if and only if

$$
0 \in \partial (f + I_{\mathbf{X}_{\geq 0}, A(\mathbf{X}) = \mathbf{b}})(\mathbf{X}_0).
$$

By assumption, $f$ is continuous everywhere. Hence, the Moreau-Rockefellar Theorem [19] guarantees that

$$
\partial (f + I_{\mathbf{X}_{\geq 0}, A(\mathbf{X}) = \mathbf{b}})(\mathbf{X}_0) = \partial f(\mathbf{X}_0) + \partial I_{\mathbf{X}_{\geq 0}, A(\mathbf{X}) = \mathbf{b}}(\mathbf{X}_0).
$$

By Lemma 1, the completeness condition (9) is equivalent to

$$
\partial I_{\mathbf{X}_{\geq 0}, A(\mathbf{X}) = \mathbf{b}}(\mathbf{X}_0) = \partial I_{\mathbf{X}_{\geq 0}}(\mathbf{X}_0) + \partial I_{A(\mathbf{X}) = \mathbf{b}}(\mathbf{X}_0).
$$

We note that

$$
\partial I_{\mathbf{X}_{\geq 0}}(\mathbf{X}_0) = \{Q \mid Q \preceq 0, Q \perp \mathbf{X}_0\},
$$

$$
\partial I_{A(\mathbf{X}) = \mathbf{b}}(\mathbf{X}_0) = \{A^* \lambda\},
$$

$$
\partial I_{\mathbf{X}_{\geq 0}}(\mathbf{X}_0) + \partial I_{A(\mathbf{X}) = \mathbf{b}}(\mathbf{X}_0) = \{A^* \lambda + Q \mid Q \preceq 0, Q \perp \mathbf{X}_0\} = S.
$$

We conclude there exists a dual certificate $(\lambda, Q)$ by combining (13), (14), (15), and (18).

The corollary follows from Theorem 3 because the independence assumption implies that there are no nontrivial linear combinations of measurement matrices that are positive and orthogonal to $\mathbf{X}_0$.

Proof of Corollary 1. Consider $A \succeq 0$, $A = \sum_i \lambda_i A_i$, $A \perp \mathbf{x}_0 \mathbf{x}_0^T$. By (5), $A\mathbf{x}_0 = 0$. Hence, $\sum_i \lambda_i A_i \mathbf{x}_0 = 0$. By independence of $\{A_i \mathbf{x}_0\}$, $\lambda_i = 0$ for all $i$. Hence the conditions of Theorem 3 are met and there exists a dual certificate at $\mathbf{X}_0$. 

3.2 Proof of Theorem 5

Theorem 5 provides a weak form of necessity for the completeness condition (9). If \(-\partial f(X_0)\) only contains matrices that are not of form \(S\), there will be no dual certificate. When \(S\) does not satisfy (9), there is a matrix orthogonal to all feasible points, and we choose \(f\) to have a gradient in the opposite direction. This argument also plays an important role in the primary technical lemma establishing equivalence between the completeness condition and additivity of subgradients.

**Proof of Theorem 5** If \(S\) does not satisfy the completeness condition (9), then there is a \(q \perp x_0\) and a \(y\) such that \(qq^* \in S\) and \(y \otimes q \notin S\). Consider the problem

\[
\min (-y \otimes q, X) \text{ subject to } X \succeq 0, A(X) = A(X_0).
\]

Because \(qq^* \in S\), \(qq^* = A^*\lambda + Q\) for some \(Q \succeq 0, Q \perp X_0\). Hence,

\[
\langle X, qq^* \rangle = \langle X - X_0, qq^* \rangle = \langle X - X_0, A^*\lambda + Q \rangle = \langle X - X_0, Q \rangle = \langle X, Q \rangle \leq 0,
\]

where the third equality uses \(A(X - X_0) = 0\). Because feasibility of \(X\) implies \(\langle X, qq^* \rangle \geq 0\), we conclude all feasible \(X \perp qq^*\). Hence, \(\langle -y \otimes q, X \rangle = 0\) for any feasible \(X\). Hence \(X_0\) is a minimizer. There is no dual certificate because \(-\partial f(X_0)\) contains the single element \(y \otimes q \notin S\). \(\square\)

4 Technical Lemmas

The main technical lemma establishes additivity of subgradients of a class of indicator functions. One direction follows from the same argument as the proof of Theorem 5. The other direction follows by showing \(Y \notin S \Rightarrow Y \notin \partial I_{S \geq 0, \Lambda(A(X)) = b}(X_0)\). To show \(Y\) is not such a subgradient, we explicitly build a feasible \(X\) using a separating hyperplane argument. That argument requires \(S\) be closed, as proven in Lemma 2. It also hinges on Lemma 4 which classifies when perturbations from \(X_0\) remain positive semidefinite.

**Lemma 1.** Let \(X_0 = x_0x_0^*\) and \(A(X_0) = b\). \(S = \{A^*\lambda + Q \mid Q \succeq 0, Q \perp X_0\}\) satisfies the completeness condition (9) if and only if

\[
\partial I_{S \geq 0, A(X) = b}(X_0) = \partial I_{S \geq 0}(X_0) + \partial I_{A(X) = b}(X_0).
\]

**Proof of Lemma 1** First, we show \(-9 \Rightarrow -20\). There exists \(q \perp x_0\) such that \(qq^* \in S\) but \(y \otimes q \notin S\) for some \(y\). Following the calculation in the proof of Theorem 5, all feasible \(X\) are orthogonal to \(y \otimes q\). Hence, \(y \otimes q \in \partial I_{S \geq 0, A(X) = b}(X_0)\), but \(y \otimes q \notin S = \partial I_{S \geq 0}(X_0) + \partial I_{A(X) = b}(X_0)\).

Next, we show \(9 \Rightarrow 20\). One inclusion in (20) is automatic:

\[
\partial I_{S \geq 0, A(X) = b}(X_0) = \partial I_{S \geq 0}(X_0) + I_{A(X) = b}(X_0) \supset \partial I_{S \geq 0}(X_0) + I_{A(X) = b}(X_0).
\]

To prove the other inclusion, we let \(Y \notin S = \partial I_{S \geq 0}(X_0) + \partial I_{A(X) = b}(X_0)\). We will show that \(Y \notin \partial I_{S \geq 0, A(X) = b}(X_0)\) by exhibiting a feasible \(X\) such that \(\langle Y, X - X_0 \rangle > 0\).

As we will prove in Lemma 2 (9) implies that \(S\) is closed. By the separating hyperplane theorem, for any \(Z \notin S\), there exists a \(A_Z\) such that

\[
\begin{align*}
A(A_Z) &= 0, \quad (22) \\
\langle A_Z, Q \rangle &\leq 0 \text{ for all } Q \succeq 0, Q \perp X_0, \quad (23) \\
\langle A_Z, M \rangle &= 0 \text{ if } \pm M \in S, \quad (24) \\
\langle A_Z, Z \rangle &> 0. \quad (25)
\end{align*}
\]
We observe that (23) implies $\mathcal{P}_{x_0^\perp} \Lambda Z \succeq 0$.

Let $B = \{qq^* \mid qq^* \perp X_0, qq^* \notin S\}$. We will build $\tilde{A}$ satisfying the following properties:

$$
\langle \tilde{A}, Q \rangle \leq 0 \text{ for all } Q \preceq 0, Q \perp X_0,
$$

$$
\langle \tilde{A}, M \rangle = 0 \text{ if } \pm M \in S,
$$

$$
\langle \tilde{A}, qq^* \rangle > 0 \text{ for all } qq^* \in B.
$$

We build $\tilde{A}$ through the following process. Choose a $q_1q_1^* \in B$ and find a corresponding $\Lambda_{q_1,q_1^*}$. Restrict $B$ to a set $\tilde{B}$ containing only the elements that are orthogonal to $\Lambda_{q_1,q_1^*}$. All elements in $B \setminus \tilde{B}$ have a positive inner product with $\Lambda_{q_1,q_1^*}$. Choose $q_2q_2^* \in \tilde{B}$ and find $\Lambda_{q_2,q_2^*}$. Further restrict $\tilde{B}$ to only the elements that are orthogonal to $\Lambda_{q_2,q_2^*}$. Now, all elements in $B \setminus \tilde{B}$ have a positive inner product with $\Lambda_{q_1,q_1^*}$ or $\Lambda_{q_2,q_2^*}$. Repeat this process until $B$ is empty. The process will complete after a finite number of repetitions because the set $\tilde{B}$ is restricted to a space of strictly decreasing dimension at each step. Let $\tilde{A} = \sum_i \Lambda_{q_i,q_i^*}$. We observe (26) hold due to (22)–(24). Every element of $B$ has a positive inner product with $\Lambda_{q_i,q_i^*}$ for some $i$. Hence, we have (29).

Let $\Lambda = \Lambda Y + \varepsilon \tilde{A}$, where $\varepsilon$ is small enough that $\langle \Lambda, Y \rangle > 0$. By Lemma 3 if (a) $\mathcal{P}_{x_0^\perp} \Lambda \succeq 0$ and (b) $\Lambda \perp qq^*$ and $qq^* \perp X_0 \Rightarrow \Lambda \perp x_0 \otimes q$, then there exists $\delta > 0$ such that $X_0 + \delta \Lambda \succeq 0$. By (23) and (27), (a) holds. To show (b) holds, we consider a $qq^* \perp \Lambda, qq^* \perp X_0$. By (29) and the definition of $\Lambda, qq^*$ must be in $S$. By (19), $\pm x_0 \otimes q \in S$. Hence, by (24) and (28), $\Lambda \perp x_0 \otimes q$, and (b) holds.

As given by Lemma 4, let $X = X_0 + \delta \Lambda$. Because $X \succeq 0$ and $\mathcal{A}(\Lambda) = 0$, $X$ is feasible. Additionally, $\langle Y, X - X_0 \rangle > 0$ because $\langle \Lambda, Y \rangle > 0$. Hence, $Y \notin \partial I_{X \succeq 0, \mathcal{A}(X) = \mathcal{B}(X)}$.

The hyperplane separation argument above requires that $S$ be closed. The following lemma reduces the closedness of $S \subset \mathcal{S}_n$ to an $n - 1 \times n - 1$ case without the orthogonality constraint, which is proved in Lemma 3.

**Lemma 2.** If $S = \{\sum_i \lambda_i A_i + Q \mid Q \preceq 0, Q \perp X_0\}$ satisfies the completeness condition (9) then $S$ is closed.

**Proof of Lemma 2.** Without loss of generality let $X_0 = e_1 e_1^*$. This can be seen by letting $V$ be an orthogonal matrix with $x_0/\|x_0\|$ in the first column, and by considering the set $V^* SV$. If necessary, linearly recombine the $A_i$ such that the first columns of $A_1, \ldots, A_\ell$ are independent and the first columns of the remaining $A_{\ell+1}, \ldots, A_m$ are zero.

Consider a Cauchy sequence $A^{(k)} + Q^{(k)} \to X$, where $A^{(k)} = \sum_{\ell=1}^m \lambda_i^{(k)} A_i$. We will establish that $X \in S$. Because $Q^{(k)} \preceq 0$ and $Q^{(k)} \perp e_1 e_1^*$, it is zero in the first row and column. Hence the first column of $\sum_{\ell=1}^m \lambda_i^{(k)} A_i$ converges to the first column of $X$. By independence, we obtain that $\lambda_i^{(k)}$ converges to some $\lambda_i^{(\infty)}$ for each $1 \leq i \leq \ell$. As a result,

$$
\sum_{i=\ell+1}^m \lambda_i^{(k)} A_i + Q^{(k)} \to \overline{X},
$$

where $\overline{X} = X - \sum_{i=1}^{\ell} \lambda_i^{(\infty)} A_i$, and $\overline{X}$ is zero in the first row and column.

The problem has now been reduced to one of size $n - 1 \times n - 1$ without an orthogonality constraint, and Lemma 3 completes the proof. Let $\tilde{A}_i$ be the lower-right $n - 1 \times n - 1$ sub matrix
of $A_i$. Let $\tilde{S} = \{\sum_{i=\ell+1}^m \lambda_i \tilde{A}_i + \tilde{Q} \mid \tilde{Q} \preceq 0\} \in S_{n-1}$. If $\tilde{q}\tilde{q}^* \in \tilde{S}$ then $\begin{pmatrix} 0 \\ \tilde{q} \end{pmatrix} = \begin{pmatrix} 0 \\ q \end{pmatrix}^* \in S$. By (9),
$$\begin{pmatrix} 0 \\ \tilde{y} \end{pmatrix} \otimes \begin{pmatrix} 0 \\ q \end{pmatrix} \in S$$ $\forall y \in \mathbb{R}^{n-1}$. By independence of the first columns of $A_1, \ldots, A_\ell$, $\tilde{y} \otimes \tilde{q} \in \tilde{S}$. The conditions of Lemma 3 are met. Hence, $\tilde{X} = \sum_{i=\ell+1}^m \lambda_i(\infty) A_i + Q(\infty)$ with $Q(\infty) \preceq 0, Q(\infty) \perp e_1e_1^*$. We conclude $X \in S$ and $S$ is closed.

\[\Box\]

The closedness of $S$ above relies on the closedness of a lower dimensional $\tilde{S}$ without the orthogonality constraint. For any Cauchy sequence in $\tilde{S}$, we show the limit belongs to $\tilde{S}$. If there is no intersection between the range of $A^*$ and the space of negative semidefinite matrices, then these sets have a fixed angle between them. Hence, dual variables corresponding to a Cauchy sequence $P$ cannot diverge separately yet converge when combined. If there is an intersection between the range of $A^*$ and the space of negative semidefinite matrices, the variant of the completeness condition allows us to project away from those directions.

**Lemma 3.** The set $\tilde{S} = \{\sum_{i} \lambda_i A_i + Q \mid Q \preceq 0\} \subset S_n$ is closed if
$$qq^* \in \tilde{S} \Rightarrow y \otimes q \in \tilde{S} \forall y. \quad (30)$$

**Proof of Lemma 3.** Consider a Cauchy sequence $A^{(k)} + Q^{(k)} \rightarrow X$, where $A^{(k)} = \sum_{i} \lambda_i^{(k)} A_i$. Let $V = \text{span}\{q \mid qq^* \in \tilde{S}\}$. For each $q \in V$, (30) gives that $y \otimes q \in \tilde{S} \forall y$. Because $P_{V^\perp}$ is the projection of matrices onto matrices with row and column spaces living in $V^\perp$,
$$\pm(X - P_{V^\perp}X) \in \tilde{S} \text{ for any } X. \quad (31)$$

The Cauchy sequence satisfies
$$P_{V^\perp}A^{(k)} + P_{V^\perp}Q^{(k)} \rightarrow P_{V^\perp}X. \quad (32)$$

If $\|P_{V^\perp}A^{(k)}\|_F \rightarrow \infty$, then $\frac{\|P_{V^\perp}A^{(k)}\|_F}{\|P_{V^\perp}Q^{(k)}\|_F} \rightarrow 1$ and $\left(\frac{P_{V^\perp}A^{(k)}}{\|P_{V^\perp}A^{(k)}\|_F}, \frac{P_{V^\perp}Q^{(k)}}{\|P_{V^\perp}Q^{(k)}\|_F}\right) \rightarrow -1$ as $k \rightarrow \infty$. The sets $\{A \in P_{V^\perp}\text{span } A_i\} \cap \{\|A\|_F = 1\}$ and $\{Q \preceq 0\} \cap \{\|Q\|_F = 1\}$ are compact. Hence $(A, Q)$ achieves its minimum. That minimum value must be $-1$, which implies that there exists a nonzero, positive semidefinite matrix $-Q \in P_{V^\perp}\text{span } A_i$. This is impossible by the construction of $V$. Suppose $P_{V^\perp}A^*\lambda \succeq 0$. By (31), we see $P_{V^\perp}A^*\lambda \in \tilde{S}$. Hence every rank-1 component $qq^*$ of $P_{V^\perp}A^*\lambda \succeq 0$ belongs to $\tilde{S}$. We reach a contradiction because $q$ would belong to $V$ and can not be in the column space of $P_{V^\perp}A^*\lambda$.

Hence, $P_{V^\perp}A^{(k)}$ has a bounded subsequence. Thus, there is a further subsequence that converges and $P_{V^\perp}X$ is of the form $P_{V^\perp}(\sum_{i} \lambda_i^{(\infty)} A_i + Q^{(\infty)})$. By (31), we conclude $X = \sum_{i=1}^m \lambda_i^{(\infty)} A_i + Q^{(\infty)}$ with $Q^{(\infty)} \preceq 0$.

\[\Box\]

The following lemma establishes a necessary and sufficient condition for when a symmetric perturbation from a positive rank 1 matrix remains positive.

**Lemma 4.** Let $X_0 = x_0x_0^* \in S_n$. $X_0 + \delta A \succeq 0$ for some $\delta > 0$ if and only if (a) $P_{x_0^\perp}A \succeq 0$ and (b) $A \perp qq^*$ and $q \perp x_0 \Rightarrow A \perp x_0 \otimes q$. 

10
Proof. Without loss of generality, assume $X_0 = e_1 e_1^*$, in which case $P_{X_0^\perp}$ is the restriction to the lower-right $n - 1 \times n - 1$ block. Let $\Lambda_{X_0^\perp} \in S_{n-1}$ be that lower-right block of $\Lambda$. Write the block form

$$
\Lambda = \begin{pmatrix}
\Lambda_{11} & \rho^*

\rho & \Lambda_{x_0^\perp}
\end{pmatrix}.
$$

First we prove $X_0 + \delta \Lambda \succeq 0 \Rightarrow (a)$ and (b). We immediately have (a) because $X_0$ is zero on the lower-right subblock. Using a Schur complement,

$$
\text{if } 1 + \delta \Lambda_{11} > 0, \text{ then } X_0 + \delta \Lambda \succeq 0 \iff \Lambda_{x_0^\perp} - \frac{\delta}{1 + \delta \Lambda_{11}} \rho \rho^* \succeq 0.
$$

(33)

If necessary, $\delta$ can be reduced to enforce $1 + \delta \Lambda_{11} > 0$. If (b) does not hold, then there is $\xi \in \mathbb{R}^{n-1}$ such that $\Lambda_{x_0^\perp} \perp \xi \xi^*$ and $\rho \not\perp \xi$. By testing against $\xi$, we see $\Lambda_{x_0^\perp} - \frac{\delta}{1 + \delta \Lambda_{11}} \rho \rho^* \not\succeq 0$.

Second, we prove (a) and (b) $\Rightarrow X_0 + \delta \Lambda$ for some $\delta > 0$. Assume (a) and (b) hold. Using the property (33) about Schur complements, it suffices to show

$$
\Lambda_{x_0^\perp} - \frac{\delta}{1 + \delta \Lambda_{11}} \rho \rho^* \succeq 0.
$$

(34)

Let $V = \text{span} \{ q \mid \Lambda_{x_0^\perp} \perp qq^* \} \subset S_{n-1}$. There is some $\epsilon$ such that $\Lambda_{x_0^\perp} \succeq \epsilon I_V$. If not, there would be a sequence of $x^{(e)} \in V^\perp$ such that $\|x^{(e)}\| = 1$ and $0 < x^{(e)} \Lambda_{x_0^\perp} x^{(e)*} < \epsilon$. Such $x^{(e)}$ would have a convergent subsequence to some $x^{(0)} \in V^\perp$ such that $x^{(0)} \Lambda_{x_0^\perp} x^{(0)*} = 0$, which is impossible by choice of $V$.

We note that for any $q \in V$, (b) guarantees $\rho \perp q$. Hence $\rho \in V^\perp$ and there is a sufficiently small $\delta$ such that $\frac{\delta}{1 + \delta \Lambda_{11}} \rho \rho^* \preceq \epsilon I_{V^\perp}$. We conclude that (33) holds, and hence $\exists \delta > 0$ such that $X_0 + \delta \Lambda \succeq 0$.

Acknowledgements

P.H. is partially supported by an NSF Mathematical Sciences Postdoctoral Research Fellowship and thanks Laurent Demanet, Radu Balan, Henry Wolkowicz and Yuen-Lam Cheung for useful discussions.

References

[1] A. Ahmed, B. Recht, J. Romberg. Blind Deconvolution using Convex Programming. arXiv preprint 1211.5608, 2012

[2] J. Borwein, H. Wolkowicz. Regularizing the Abstract Convex Program J. Math. Anal. Appl. 83, 495–530, 1981.

[3] S. Boyd, L. Vandenberghe. Convex Optimization. Cambridge University Press, 2004.

[4] E. J. Candes, Y. Eldar, T. Strohmer, V. Voroninski. Phase retrieval via matrix completion. SIAM J. on Imaging Sciences 6(1), 2011. 199–225.

[5] E. J. Candes, X. Li. Solving Quadratic Equations via PhaseLift when There Are About As Many Equations as Unknowns. arXiv 1208.6247, 2012
[6] E. J. Candes, B. Recht. Exact matrix completion via convex optimization. *Found. of Comput. Math.*, 9, 2008, 717–772.

[7] E. J. Candes, T. Strohmer, V. Voroninski. PhaseLift: Exact and Stable Signal Recovery from Magnitude Measurements via Convex Programming To appear in *Comm. Pure Appl. Math.*, 2012

[8] Y-L. Cheung, S. Schurr, and H. Wolkowicz. Preprocessing and regularization for degenerate semidefinite programs. In D.H. Bailey, H.H. Bauschke, P. Borwein, F. Garvan, M. Thera, J. Vanderwerf?, and H. Wolkowicz, editors, *Computational and Analytical Mathematics, In Honor of Jonathan Borweins 60th Birthday, volume 50 of Springer Proceedings in Mathematics and Statistics*. Springer, 2013.

[9] L. Demanet, P. Hand. Stable optimizationless recovery from phaseless linear measurements. arXiv preprint 1208.1803, 2012

[10] R. Rm. Freund. *Complexity of an Algorithm for Finding an Approximate Solution of a Semidefinite Program with No Regularity Assumption*, Technical report OR 302-94, MIT, Cambridge, MA, 1994.

[11] D. Gross. Recovery Low-Rank Matrices From Few Coefficients In Any Basis. *IEEE Trans. Inf. Theory*, 57(3), Mar 2011

[12] M. Guignard. Generalized Kuhn-Tucker conditions for mathematical programming problems in a Banach space. SIAM J. Control, 7(2):232-241, 1969.

[13] X. Li, V. Voroninski. Sparse Signal Recovery from Quadratic Measurements via Convex Programming. arXiv preprint 1209.4785, 2012

[14] L. Tuncel, H. Wolkowicz. *Strong duality and minimal representations for cone optimization*. Comput. Optim. Appl 53, 619-648, 2012.

[15] H. Ohlsson, A. Y. Yang, R. Dong, S. S. Sastry. Compressive phase retrieval from squared output measurements via semidefinite programming. *arXiv preprint 1111.6323*, 2012.

[16] B.N. Pshenichniyi. *Necessary conditions for an extremum*. Pure and Applied Mathematics, vol. 4. Dekker, New York (1971).

[17] M. V. Ramana, *An Exact Duality Theory for Semidefinite Programming and its Complexity Implications*, DIMACS Technical report 95-02R, RUTCOR, Rutgers University, New Brunswick, NJ, 1995.

[18] M. V. Ramana, L. Tuncel, H. Wolkowicz. Strong Duality for Semidefinite Programming SIAM J. Optim. 7(3), 641–662, 1997.

[19] R. Rockafellar. *Convex Analysis*. Princeton University Press, 1970

[20] H. Wolkowicz. *Geometry of optimality conditions and constraint qualifications: the convex case*. Math. Programming, 19(1):32-60, 1980.