Strong Duality for a Trust-Region Type Relaxation of the Quadratic Assignment Problem

Kurt Anstreicher *   Xin Chen †   Henry Wolkowicz ‡   Ya-Xiang Yuan §

July 15, 1998

University of Waterloo

Department of Combinatorics and Optimization

Waterloo, Ontario N2L 3G1, Canada

Research Report CORR 98-31

AMS Subject Classifications: 49M40 52A41 90C20 90C27

Key words: Lagrangian relaxations, quadratic assignment problem, semidefinite programming, quadratically constrained quadratic programs.

Contents

---

*Department of Management Sciences, The University of Iowa. E-mail kurt-anstreicher@uiowa.edu

†Research supported by Chinese National Natural Science Foundation. E-mail chenx@lsec.cc.ac.cn

‡Research supported by NSERC. E-mail henry@orion.uwaterloo.ca.

§Research supported by Chinese National Natural Science Foundation. E-mail yyx@indigo9.cc.ac.cn

[0]This report is available by anonymous ftp at orion.uwaterloo.ca, in directory pub/henry/reports and also with URL http://orion.uwaterloo.ca/~hwolkowi/henry/reports/sdptrsqap.ps.gz

Abstract

Lagrangian duality underlies many efficient algorithms for convex minimization problems. A key ingredient is strong duality. Lagrangian relaxation also provides lower bounds for nonconvex problems, where the quality of the lower bound depends on the duality gap. Quadratically constrained quadratic programs (QQPs) provide important examples of nonconvex programs. For the simple case of one quadratic constraint (the trust region subproblem) strong duality holds. In addition, necessary and sufficient (strengthened) second order optimality conditions exist. However, these duality results already fail for the two trust region subproblem.

Surprisingly, there are classes of more complex, nonconvex QQPs where strong duality holds. One example is the special case of orthogonality constraints, which arise naturally in relaxations for the quadratic assignment problem (QAP). In this paper we show that strong duality also holds for a relaxation of QAP where the orthogonality constraint is replaced by a semidefinite inequality constraint. Using this strong duality result, and semidefinite duality, we develop new trust-region type necessary and sufficient optimality conditions for these problems. Our proof of strong duality introduces and uses a generalization of the Hoffman-Wielandt inequality.

1 Introduction

Quadratic programs with quadratic constraints (QQPs) are an important modelling tool for many optimization problems; almost as important as the linear programming model. Applications for QQP include e.g. hard combinatorial problems, e.g. [25], and SQP algorithms for nonlinear programming, e.g. [17]. These QQPs are often not convex and so are very hard to solve numerically. One approach is to use the Lagrangian relaxation of a QQP to obtain an approximate solution. The strength of such a relaxation depends on the duality gap, where a zero duality gap means that the relaxation is exact. In this paper we present a new technique for closing the duality gap for a class of nonconvex problems. This technique is to add certain redundant constraints before taking the Lagrangian relaxation.

The simplest of the nonconvex QQPs is the trust region subproblem, TRS, which consists of a quadratic objective with a single quadratic constraint. The constraint is usually the simple norm constraint (we normalize the right hand side to 1)

$$x^T x = 1 \quad (\text{or } \leq 1). \tag{1}$$

Surprisingly, see [29], the Lagrangian relaxation for this possibly nonconvex problem is exact. Moreover, there are (strengthened) second order necessary and sufficient optimality conditions for TRS, [19].

A visually similar problem to the equality-constrained TRS is the matrix quadratic problem with orthogonality constraints

$$XX^T = I. \tag{2}$$

Some such problems can be solved efficiently using eigenvalue techniques, such as the Hoffman-Wielandt inequality. However strong duality fails for the obvious Lagrangian dual based on relaxing the constraint (2).

In [3] it was shown that for a certain homogeneous QQP with the orthogonality constraints (2), strong duality *does* hold if the seemingly redundant constraint

$$X^T X = I$$

is added before the Lagrangian dual is formed. In this paper we extend this strong duality result to a problem where the orthogonality constraint (2) is replaced by the trust-region type semidefinite inequality

$$XX^T \preceq I, \tag{3}$$

where for two symmetric matrices, $S \preceq T$ denotes that $T - S$ is positive semidefinite. For this problem we also develop new strengthened second order necessary and sufficient optimality conditions that are similar to the conditions known to hold for TRS.

1.1 Background

1.1.1 General QQPs

Consider the quadratically constrained quadratic program

$$\begin{array}{lll} \text{QQP} & \min & q_0(x) \\ & \text{s.t.} & q_k(x) \leq 0 \text{ (or } = 0), \quad k = 1, \ldots m, \end{array}$$

where $q_i(x) := \frac{1}{2} x^T Q_i x + g_i^T x$ is a quadratic function. The Lagrangian function is

$$L(x, \lambda) := q_0(x) + \sum_{k=1}^{m} \lambda_k q_k(x),$$

where the multiplier λ_k is constrained to be nonnegative if the kth constraint is an inequality. It is unconstrained if it is an equality and it is a symmetric matrix $\Lambda \succeq 0$ in the case of the trust region type constraint (3). The Lagrangian dual or relaxation is then

$$\max_{\lambda} \min_{x} L(x, \lambda). \tag{4}$$

There has been a great deal of recent work on QQPs. The tractable case is the convex case, i.e. the objective and constraint functions are all convex (linear for equality constraints). In this case, the solution value is attained and there is a zero duality gap between QQP and its Lagrangian dual [18]. The bridge between the convex and the nonconvex case is the TRS problem discussed above. This problem is tractable, [30], and very efficient algorithms exist both for moderate dense problems, [19] and large sparse problems, [27, 28].

One view of the Lagrangian relaxation of QQPs is in terms of semi-infinite programming and valid inequalities. Let $\mathcal{F}$ denote the feasible set of the QQP. Then we trivially have

$$\lambda \geq 0 \quad \Rightarrow \quad \mathcal{F} \subset \mathcal{V}_\lambda := \{x : q_\lambda(x) := \sum_{k=1}^{m} \lambda_k q_k(x) \leq 0\}.$$

Thus q_λ provides a *valid inequality* for the feasible set. However, we now see that not all these valid inequalities are useful.

The outer maximization problem in the dual problem (4) has the hidden constraint that the Hessian

$$Q_0 + \sum_{k=1}^{m} \lambda_k Q_k \succeq 0,$$

since otherwise the inner minimization is unbounded below. Thus, for each vector of Lagrange multipliers $\lambda \geq 0$ such that the Hessian of the Lagrangian is positive semidefinite, we conclude that the useful valid inequalities for the feasible set of QQP are given by

$$\lambda \geq 0, \nabla^2_{xx} L(x, \lambda) \succeq 0 \quad \Rightarrow \quad \mathcal{F} \subset \mathcal{V}_\lambda.$$

(See [10, 16] for details for a linear objective function. The nonlinear case is being studied in [1].) Therefore, a zero duality gap means that we have enough of these useful valid inequalities. Otherwise, an obvious question is: *can we find additional quadratic constraints to close the duality gap.*

One of the highlights of the new results on QQPs is the result of Goemans-Williamson, e.g. [11], on the strength of the semidefinite programming, SDP, relaxation for the max-cut problem. This result essentially shows how well one can approximate the optimum of the QQP

$$\max \ x^T Q x \quad \text{s.t. } x_i^2 = 1, \ i = 1, \ldots, n,$$

where Q arises from the Laplacian matrix of the underlying (nonnegatively weighted) graph. This result has been extended in several ways: to allow for general Q [22]; to replace the constraints with interval constraints [31]; to allow for general homogeneous constraints [20, 9]; and other extensions [4, 21]. The above mentioned papers all characterize the quality of a tractable approximation to a nonconvex QQP, rather than finding special quadratic constraints to add in order to improve the approximation. The interpretation of the semidefinite relaxation in terms of valid quadratic inequalities is discussed in [10, 16].

1.1.2 Quadratic Assignment Problem and Relaxations

The *Quadratic Assignment Problem*, QAP, in the trace formulation is

$$\mu^* := \min_{X \in \Pi} \operatorname{tr}\left(AXBX^T + CX^T\right),$$

where Π denotes the set of permutation matrices, and A, B, C are $n \times n$ matrices. We assume throughout that A and B are real and symmetric. Applications of QAP include plant location problems, where the three matrices represent distances between sites, flows between plants, and location costs, respectively, and the permutation matrix X denotes which plant is located at which site. See for example [24, 6] for an extensive discussion of applications and algorithms for QAP.

The QAP is an NP-hard problem. In fact, this is one of the most difficult problems to solve in practice as there exist problems with dimension $n = 20$ still unsolved, [13, 24, 6]. For QAPs dimension $n = 25$ is considered "large scale." The problem consists of a, possibly nonconvex, quadratic objective function over the (discrete) set of permutation matrices. Since the set of permutation matrices is the intersection of the orthogonal matrices $\mathcal{O}$ with the doubly stochastic matrices $\mathcal{E}$ and the nonnegative matrices $\mathcal{N}$,

$$\Pi = \mathcal{O} \cap \mathcal{E} \cap \mathcal{N},$$

relaxations for the QAP often include quadratic constraints such as

$$XX^T = I;$$

or the trust region type of constraint

$$XX^T \preceq I.$$

As the objective in QAP is itself quadratic, these relaxations of QAP lead naturally to interesting classes of QQPs.

General nonlinear optimization over orthogonality constraints is considered in [7] while the partial order constraint $XX^T \preceq I$ is discussed in [23]. The relationship $Y = XX^T$ is used to model graph partitioning problems in [14, 2].

1.2 Outline

In this paper we study the trust region type relaxation for homogeneous ($C = 0$) QAP. We first find the explicit solution for the relaxation, and thus introduce an extension of the well-known Hoffman-Wielandt inequality. We then show that by adding the seemingly redundant constraint $X^T X^T \preceq I$ before forming the Lagrangian dual we can close the duality gap. Using this strong duality result, and semidefinite duality, we obtain new necessary and sufficient characterizations for optimality which are similar to the ordinary trust region subproblem result in nonlinear programming.

1.3 Notation

We now describe the notation used in the paper. Comprehensive up-to-date notation for SDP is available on the WWW with URL:
http://orion.uwaterloo.ca/~hwolkowi/henry/software/psd_tool.d/sdnotation.d/notation.ps.

Throughout this paper we work with real matrices. Let $\mathcal{S}_n$ denote the space of $n \times n$ symmetric matrices equipped with the trace inner product, $\langle A, B\rangle = \operatorname{tr} AB$. Let $A \succeq 0$ (resp. $A \succ 0$) denote positive semidefiniteness (resp. positive definiteness); $A \succeq B$ denotes $A - B \succeq 0$, i.e. $\mathcal{S}_n$ is equipped with the Löwner partial order. We let $\mathcal{P}$ denote the cone of symmetric positive semidefinite matrices; $\mathcal{M}_{m,n}$ denotes the space of general $m \times n$ matrices also equipped with the trace inner product, $\langle A, B\rangle = \operatorname{tr} A^T B$; while $\mathcal{M}_m$ denotes the space of general $m \times m$ matrices; $\mathcal{O}$ denotes the set of orthonormal (orthogonal) matrices; Π denotes the set of permutation matrices.

We let $\operatorname{Diag}(v)$ be the diagonal matrix formed from the vector v; its adjoint operator is $\operatorname{diag}(M)$ which is the vector formed from the diagonal of the matrix M. For $M \in \mathcal{M}_{m,n}$, the vector $m = \operatorname{vec}(M) \in \Re^{mn}$ is formed (columnwise) from M.

The Kronecker product of two matrices is denoted $A \otimes B$, and the Hadamard product is denoted $A \circ B$.

We use e to denote the vector of all ones, and $E = ee^T$ to denote the matrix of all ones. We use J to denote the matrix $J = (e_n, e_{n-1}, \cdots, e_1)$, where e_i is the ith unit vector.

2 Orthogonal Relaxation

One successful relaxation for the homogeneous ($C = 0$) QAP is the *eigenvalue relaxation* [8], i.e. one replaces Π with the set of orthogonal matrices

$$\mathcal{O} := \{X : XX^T = I\}.$$

We now consider strong duality results for this problem. The relaxed problem can be written

$$\mu^O := \min_{X \in \mathcal{O}} \operatorname{tr} AXBX^T. \tag{5}$$

The bound μ^O is often referred to as the *eigenvalue bound* for QAP. This bound is based on the following inequality, which can be viewed as a variant of the classical *Hoffman-Wielandt inequality*, see e.g. [8, 26, 5].

Theorem 1 *Let $V^TAV = \Sigma$, $U^TBU = \Lambda$, where $U, V \in \mathcal{O}$, $\Sigma = \mathrm{Diag}(\sigma)$, $\Lambda = \mathrm{Diag}(\lambda)$, $\sigma_1 \geq \sigma_2 \geq \cdots \geq \sigma_n$, $\lambda_1 \geq \lambda_2 \geq \cdots \geq \lambda_n$. Then for any $X \in \mathcal{O}$, we have*

$$\sum_{i=1}^n \lambda_i \sigma_{n-i+1} \leq \mathrm{tr}\, AXBX^T \leq \sum_{i=1}^n \lambda_i \sigma_i.$$

The upper bound is attained for $X = VU^T$, and the lower bound is attained for $X = VJU^T$, where $J = (e_n, e_{n-1}, \cdots, e_1)$ and e_i is the ith element unit vector. ■

It is clear that the eigenvalue bound is a tractable bound, i.e. it can be efficiently computed in polynomial time by computing the eigenvalues and ordering them appropriately. However, there can be a duality gap for the Lagrangian relaxation of (5) (and so also for the SDP relaxation, which is equivalent); see [32] for an example. Interestingly, we can close this duality gap by adding the seemingly redundant constraint $X^TX = I$ before forming the Lagrangian dual; see [3]. Define the primal problem

$$\begin{array}{llll} \text{QAPO} \qquad & \mu^O = & \min & \mathrm{tr}\, AXBX^T \\ & & \text{s.t.} & XX^T = I, \ X^TX = I. \end{array}$$

Using symmetric matrices S and T to relax the constraints $XX^T = I$ and $X^TX = I$, respectively, we arrive at a dual problem

$$\begin{array}{llll} \text{DQAPO} \qquad & \mu^O \geq \mu^{DO} := & \max & \mathrm{tr}\, S + \mathrm{tr}\, T \\ & & \text{s.t.} & (I \otimes S) + (T \otimes I) \preceq (B \otimes A) \\ & & & S = S^T, \ T = T^T. \end{array}$$

Theorem 2 *[3] Strong duality holds for* QAPO *and* DQAPO, *i.e. $\mu^{DO} = \mu^O$ and both primal and dual values are attained.* ■

3 Trust Region Relaxation

A further relaxation of the above orthogonal relaxation is the trust region relaxation studied in [15],

$$\begin{array}{rll} \mu^{\mathrm{Tr}} := & \min & \operatorname{tr} AXBX^T \\ & \text{s.t.} & XX^T \preceq I. \end{array}$$

The constraints $XX^T \preceq I$ are convex, and so it is hoped that solving this problem would be useful in obtaining bounds for QAP.

To begin, we will characterize the value μ^{Tr} by proving a generalization of Theorem 1. We require the following technical result.

Lemma 3 *Let B and X be $n \times n$ matrices, with B symmetric. Let $\lambda_1 \geq \lambda_2 \geq \ldots \geq \lambda_n$ be the eigenvalues of B, and $\lambda'_1 \geq \lambda'_2 \geq \cdots \geq \lambda'_n$ the eigenvalues of XBX^T. Let $X = P^T \Gamma Q$ be the singular value decomposition of X, where $P, Q \in \mathcal{O}$, $\Gamma = Diag(\gamma)$, $\gamma_1 \geq \gamma_2 \geq \cdots \geq \gamma_n \geq 0$. Then*

$$\begin{array}{ll} \gamma_n^2 \lambda_i \leq \lambda'_i \leq \gamma_1^2 \lambda_i, & \textit{for} \quad \lambda_i \geq 0, \\ \gamma_1^2 \lambda_i \leq \lambda'_i \leq \gamma_n^2 \lambda_i, & \textit{for} \quad \lambda_i < 0. \end{array}$$

Proof. Let $\mathcal{X}$ denote a subspace of $\Re^n$, and $|\mathcal{X}|$ denote the dimension of $\mathcal{X}$. First we assume that X is nonsingular. Because the eigenvalues of XBX^T are also those of $\Gamma Q B Q^T \Gamma$, by the Courant-Fisher theorem [12, Theorem 4.2.11] we have

$$\lambda'_i = \min_{|\mathcal{X}|=n-i+1} \max_{0 \neq x \in \mathcal{X}} \frac{x^T \Gamma Q B Q^T \Gamma x}{\| x \|^2}.$$

Then

$$\lambda'_i = \min_{|\mathcal{X}|=n-i+1} \max_{0 \neq \Gamma^{-1} Q y \in \mathcal{X}} \frac{y^T B y}{\| \Gamma^{-1} Q y \|^2}.$$

Let $\mathcal{Y} = Q^T \Gamma \mathcal{X}$. Due to the nonsingularity of Γ, $|\mathcal{Y}| = |\mathcal{X}|$, and in addition we clearly have

$$\frac{\|y\|^2}{\gamma_1^2} \leq \|\Gamma^{-1} Q y\|^2 \leq \frac{\|y\|^2}{\gamma_n^2}.$$

Moreover it is well know that the inertia of B is preserved under the transformation XBX^T [12, Theorem 4.5.8], and therefore the signs of λ_i and λ'_i coincide, for each i. It follows that for $\lambda'_i \geq 0$ we have

$$\begin{aligned} \lambda'_i &\leq \gamma_1^2 \min_{|\mathcal{Y}|=n-i+1} \max_{0\neq y\in\mathcal{Y}} \frac{y^T By}{\| y \|^2} = \gamma_1^2 \lambda_i, \\ \lambda'_i &\geq \gamma_n^2 \min_{|\mathcal{Y}|=n-i+1} \max_{0\neq y\in\mathcal{Y}} \frac{y^T By}{\| y \|^2} = \gamma_n^2 \lambda_i. \end{aligned}$$

While for $\lambda'_i < 0$ we have

$$\begin{aligned} \lambda'_i &\geq \gamma_1^2 \min_{|\mathcal{Y}|=n-i+1} \max_{0\neq y\in\mathcal{Y}} \frac{y^T By}{\| y \|^2} = \gamma_1^2 \lambda_i, \\ \lambda'_i &\leq \gamma_n^2 \min_{|\mathcal{Y}|=n-i+1} \max_{0\neq y\in\mathcal{Y}} \frac{y^T By}{\| y \|^2} = \gamma_n^2 \lambda_i. \end{aligned}$$

This completes the proof under the assumption that X is nonsingular. If X is singular, we can perturb the zero γ_i values and use the fact that the eigenvalues λ'_i are continuous functions of γ, to obtain the given bounds. ∎

Theorem 4 *Let $V^T AV = \Sigma$, $U^T BU = \Lambda$, where $U, V \in \mathcal{O}$, $\Sigma = \mathrm{Diag}(\sigma)$, $\Lambda = \mathrm{Diag}(\lambda)$, $\sigma_1 \geq \sigma_2 \geq \cdots \geq \sigma_n$, $\lambda_1 \geq \lambda_2 \geq \cdots \geq \lambda_n$. Then for any X with $XX^T \preceq I$ we have*

$$\sum_{i=1}^n \min\{0, \lambda_i \sigma_{n-i+1}\} \leq \mathrm{tr}\, AXBX^T \leq \sum_{i=1}^n \max\{0, \lambda_i \sigma_i\}.$$

The upper bound is attained for $X = V\,\mathrm{Diag}(\epsilon)U^T$, where $\epsilon_i = 1$ if $\sigma_i\lambda_i \geq 0$, and $\epsilon_i = 0$ otherwise. The lower bound is attained for $X = V\,\mathrm{Diag}(\epsilon)JU^T$, where $\epsilon_i = 1$ if $\sigma_i\lambda_{n+1-i} \leq 0$, and $\epsilon_i = 0$ otherwise, $J = (e_n, e_{n-1}, \cdots, e_1)$ and e_i is the ith element unit vector.

Proof: From Theorem 1 we have

$$\sum_{i=1}^n \sigma_i \lambda'_{n-i+1} \leq \mathrm{tr}\, AXBX^T \leq \sum_{i=1}^n \sigma_i \lambda'_i, \tag{6}$$

where $\lambda'_1 \geq \lambda'_2 \geq \ldots \geq \lambda'_n$ are the eigenvalues of XBX^T. In addition, the result of Lemma 3 (using $\gamma_1 \leq 1$, $\gamma_n \geq 0$) implies that for any i and j,

$$\sigma_i \lambda'_j \leq \begin{cases} \sigma_i\lambda_j & \text{if } \sigma_i\lambda_j \geq 0 \\ 0 & \text{otherwise} \end{cases}, \qquad \sigma_i \lambda'_j \geq \begin{cases} \sigma_i\lambda_j & \text{if } \sigma_i\lambda_j < 0 \\ 0 & \text{otherwise} \end{cases}. \tag{7}$$

The bounds of the theorem follow by combining (6) and (7). Attainment of the bounds may be verified by direct substitution of the indicated solutions into $\operatorname{tr} AXBX^T$. ■

For a scalar ξ, let $\xi^- := \min\{0, \xi\}$. From attainment of the lower bound in Theorem 4, we have $= \mu^{\mathrm{Tr}} \sum_{i=1}^n [\lambda_i \sigma_{n+1-i}]^-$. To establish a strong duality result for the trust region type relaxation, we will next prove that this same value is attained by the solution of a Lagrangian dual program. Note that since XX^T and X^TX have the same eigenvalues, the condition $XX^T \preceq I$ is equivalent to $X^TX \preceq I$. Explicitly using both sets of constraints, as in [3], we obtain the trust region type relaxation

$$\begin{array}{lrll} \text{QAPT} & \mu^{\mathrm{Tr}} = & \min & \operatorname{tr} AXBX^T \\ & & \text{s.t.} & XX^T \preceq I, \quad X^TX \preceq I. \end{array}$$

Next we apply Lagrangian relaxation to QAPT, using matrices $S \succeq 0$ and $T \succeq 0$ to relax the constraints $XX^T \preceq I$ and $X^TX \preceq I$, respectively. This results in the dual problem

$$\begin{array}{lrll} \text{DQAPT} & \mu^T \geq \mu^{DT} := & \max & -\operatorname{tr} S - \operatorname{tr} T \\ & & \text{s.t.} & (B \otimes A) + (I \otimes S) + (T \otimes I) \succeq 0 \\ & & & S \succeq 0, \ T \succeq 0. \end{array}$$

To prove that $\mu^{\mathrm{Tr}} = \mu^{DT}$ we will use the following simple result.

Lemma 5 *Let $\lambda \in \Re^n$, $\lambda_1 \leq \lambda_2 \leq \ldots \leq \lambda_n$. For $\sigma \in \Re^n$ consider the problem*

$$\min \quad z_\pi := \sum_{i=1}^n [\lambda_i \sigma_{\pi(i)}]^-,$$

where $\pi(\cdot)$ is a permutation of $\{1, \ldots, n\}$, Then the permutation that minimizes z_π satisfies $\sigma_{\pi(1)} \geq \sigma_{\pi(2)} \geq \ldots \sigma_{\pi(n)}$.

Proof. Assume that $\sigma_i < \sigma_{i+1}$ for some i. We will show that interchanging σ_i and σ_{i+1} cannot increase the value of $\sum_{i=1}^n [\lambda_i \sigma_i]^-$. The lemma then follows, since if $\bar{\pi}(\cdot)$ is a minimizing permutation we can go from $\bar{\pi}(\cdot)$ to $\pi(\cdot)$ with $\sigma_{\pi(1)} \geq \sigma_{\pi(2)} \geq \ldots \geq \sigma_{\pi(n)}$ by a sequence of pairwise interchanges.

Assume without loss of generality that $\sigma_1 < \sigma_2$. Our goal is to show that $v' \le v$, where

$$v := [\lambda_1\sigma_1]^- + [\lambda_2\sigma_2]^-, \quad v' := [\lambda_1\sigma_2]^- + [\lambda_2\sigma_1]^-.$$

We will demonstrate this via a case analysis, depending on the signs of λ_1, λ_2, σ_1, and σ_2. For convenience we number the cases as indicated in the following table.

	$0 \le \sigma_1 \le \sigma_2$	$\sigma_1 \le \sigma_2 < 0$	$\sigma_1 < 0 \le \sigma_2$
$0 \le \lambda_1 \le \lambda_2$	Case 1	Case 2′	Case 3′
$\lambda_1 \le \lambda_2 < 0$	Case 2	Case 1′	Case 4′
$\lambda_1 < 0 \le \lambda_2$	Case 3	Case 4	Case 1″

Case 1/1′/1″: In each of these cases $v = 0$, so $v' \le 0 \Rightarrow v' \le v$.

Case 2/2′: In these cases we need to show that $\lambda_1\sigma_2 + \lambda_2\sigma_1 \le \lambda_1\sigma_1 + \lambda_2\sigma_2$, which is equivalent to $(\lambda_2 - \lambda_1)(\sigma_2 - \sigma_1) \ge 0$, and this holds by assumption.

Case 3/3′: In Case 3 we need to show that $\lambda_1\sigma_2 \le \lambda_1\sigma_1$, which is equivalent to $\lambda_1(\sigma_2 - \sigma_1) \le 0$, and this holds by assumption. Case 3′ is similar.

Case 4/4′: In Case 4 we need to show that $\lambda_2\sigma_1 \le \lambda_2\sigma_2$, which is equivalent to $\lambda_2(\sigma_2 - \sigma_1) \ge 0$, and this holds by assumption. Case 4′ is similar. ■

Theorem 6 *Strong duality holds for QAPT and DQAPT, i.e. $\mu^D = \mu^{DT}$ and both primal and dual values are attained.*

Proof. Let $A = V\Sigma V^T$, $B = U\Lambda U^T$, where $V, U \in \mathcal{O}$, $\Lambda = \operatorname{Diag}(\lambda)$, $\Sigma = \operatorname{Diag}(\sigma)$. Then for any S and T,

$$(B \otimes A) + (I \otimes S) + (T \otimes I) = (U \otimes V)\left[(\Lambda \otimes \Sigma) + (I \otimes \bar{S}) + (\bar{T} \otimes I)\right](U^T \otimes V^T),$$

where $\bar{S} = V^T S V$, $\bar{T} = U^T T U$. Since $U \otimes V$ is nonsingular, $\operatorname{tr} S = \operatorname{tr} \bar{S}$ and $\operatorname{tr} T = \operatorname{tr} \bar{T}$, the dual problem DQAPT is equivalent to

$$\begin{aligned} \mu^{DT} = \quad \max \quad & -\operatorname{tr} S - \operatorname{tr} T \\ \text{s.t.} \quad & (\Lambda \otimes \Sigma) + (I \otimes S) + (T \otimes I) \succeq 0 \\ & S \succeq 0, \ T \succeq 0. \end{aligned} \tag{8}$$

However, since Λ and Σ are diagonal matrices, (8) is equivalent to the ordinary linear program:

$$\begin{array}{lll} \text{LD} & \max & -e^T s - e^T t \\ & \text{s.t.} & \lambda_i \sigma_j + s_j + t_i \geq 0, \quad i,j = 1,\ldots,n. \\ & & s \geq 0, \ t \geq 0. \end{array}$$

But LD is the dual of the linear "semi-assignment" problem:

$$\begin{array}{lll} \text{LP} & \min & \sum_{i,j} \lambda_i \sigma_j x_{ij} \\ & \text{s.t.} & \sum_{j=1}^{n} x_{ij} \leq 1, \quad i = 1,\ldots,n \\ & & \sum_{i=1}^{n} x_{ij} \leq 1, \quad j = 1,\ldots,n \\ & & x_{ij} \geq 0, \quad i,j = 1,\ldots,n. \end{array}$$

Then LP can be interpreted as the problem of finding a permutation $\pi(\cdot)$ of $\{1,\ldots,n\}$ so that $\sum_{i=1}^{n} [\lambda_i \sigma_{\pi(i)}]^-$ is minimized. Assume without loss of generality that $\lambda_1 \leq \lambda_2 \leq \ldots \leq \lambda_n$, and $\sigma_1 \geq \sigma_2 \geq \ldots \geq \sigma_n$. From Lemma 5 the optimal permutation is then $\pi(i) = i$, $i = 1,\ldots,n$, and from Theorem 4 the solution value μ^{DT} is exactly μ^D. ■

3.1 Necessary and Sufficient Optimality Conditions

In [15] the following sufficient conditions are conjectured to also be necessary for optimality in QAPT:

$$\begin{array}{rcl} XX^T & \preceq & I, \\ S \succeq 0, \ \operatorname{tr} S(XX^T - I) & = & 0, \\ AXB + SX & = & 0, \\ \operatorname{tr}(AhBh^T + Shh^T) & \geq & 0, \text{ if } Xh^T + hX^T \text{ is nsd on } \mathcal{N}(XX^T - I). \end{array} \tag{9}$$

These conditions are similar to the standard second order optimality conditions, and are in the spirit of results for the ordinary trust region problem, i.e. they contain strengthened second order conditions where the Hessian of the Lagrangian is positive semidefinite on a larger set than the standard tangent cone. (For the standard trust region problem, the Hessian of the Lagrangian is positive semidefinite on the whole space.)

Using the characterization of optimality in Theorem 4, we can show that for some special cases the conditions (9) are in fact necessary for optimality in QAPT.

Theorem 7 *Assume that $B = I$. Then the conditions* (9) *are necessary for X to be an optimal solution of QAPT.*

Proof: Let X be an optimal solution of QAPT. Then [15, Theorem 3.1] there exists S satisfying the first three conditions in (9). From the second condition it follows that $SXX^T = S$, and therefore, from the third, $AXX^T + S = 0$. Assume that

$$A = V \begin{pmatrix} \Sigma_1 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & \Sigma_3 \end{pmatrix} V^T, \qquad XX^T = V \begin{pmatrix} X_{11} & X_{21}^T & X_{31}^T \\ X_{21} & X_{22} & X_{32}^T \\ X_{31} & X_{32} & X_{33} \end{pmatrix} V^T$$

where $V \in \mathcal{O}$, $\Sigma_1 \prec 0$ and $\Sigma_3 \succ 0$ are diagonal matrices, and the blocks X_{11} and X_{33} have the same dimensions as Σ_1 and Σ_3, respectively. Then $\operatorname{tr} AXBX^T = \operatorname{tr}(\Sigma_1 X_{11} + \Sigma_3 X_{33}) \geq \operatorname{tr} \Sigma_1$, since $X_{33} \succeq 0$ and $X_{11} \preceq I$. Moreover from Theorem 4 the optimal solution value is $\mu^{\mathrm{Tr}} = \operatorname{tr} \Sigma_1$. It follows that we must have $X_{33} = 0$, and $X_{11} = I$. The facts that $XX^T \succeq 0$ and $X_{33} = 0$ together then imply that $X_{13} = 0$ and $X_{23} = 0$, while $XX^T \preceq I$ and $X_{11} = I$ together imply that $X_{21} = 0$. Therefore

$$S = -AXX^T = V \begin{pmatrix} -\Sigma_1 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix} V^T,$$

and $A + S \succeq 0$. Then $\operatorname{tr}(Ahh^T + Shh^T) \geq 0$ for *any* matrix h, so the conditions (9) hold. ■

In addition, if A and B are positive semidefinite, then the conjectured conditions (9) are necessary. However, as we next demonstrate, the conditions (9) may in fact fail to hold.

Example 8 *Let*

$$A = \begin{pmatrix} 2 & 0 \\ 0 & 1 \end{pmatrix}, \quad B = \begin{pmatrix} -3 & 0 \\ 0 & -1 \end{pmatrix}.$$

Using Theorem 4 one can show that $X = I$ *is the global optimum of QAPT, and therefore* $\mathcal{N}(XX^T - I) = \Re^2$. *The stationarity condition* $AXB + SX = 0$ *implies that*

$$S = \begin{pmatrix} 6 & 0 \\ 0 & 1 \end{pmatrix}, \quad \text{and } h = \begin{pmatrix} -1 & 0 \\ -1 & -1 \end{pmatrix}$$

satisfies $Xh^T + hX^T \preceq 0$. *However,* $\operatorname{tr} AhBh^T + Shh^T = -2$.

Thus the conditions (9) may fail to hold at an optimal solution X of QAPT. We will now use the strong duality result of Theorem 6, and the fact that DQAPT is a semidefinite program, to derive valid necessary and sufficient conditions for optimality in QAPT. These optimality conditions are exactly like the standard trust region optimality conditions, i.e. they contain strengthened second order conditions where the Hessian of the Lagrangian is positive semidefinite on the whole space.

For an $n^2 \times n^2$ matrix Y, we use $Y_{[ij]}$ to denote the $n \times n$ matrix which is the i,j block of Y, $i,j = 1,\ldots,n$. Define linear operators $\operatorname{bdiag}(\cdot)$ and $\operatorname{odiag}(\cdot)$, $\Re^{n^2 \times n^2} \to \Re^{n \times n}$, by

$$\begin{aligned} \operatorname{bdiag}(Y) &:= \sum_{i=1}^{n} Y_{[ii]}, \\ \operatorname{odiag}(Y)_{ij} &:= \operatorname{tr} Y_{[ij]}, \quad i,j = 1,\ldots,n. \end{aligned}$$

It is then easy to show that $\operatorname{bdiag}(\cdot)$ and $\operatorname{odiag}(\cdot)$ are the adjoints of the operators $S \to I \otimes S$, and $T \to T \otimes I$, respectively. (These adjoint operators arise in the derivation of an SDP relaxation for QAP in [32].) It follows that the semidefinite dual of the program DQAPT is the following semidefinite relaxation of QAPT:

$$\begin{array}{lll} \text{QAPSDP} & \min & \operatorname{tr}(B \otimes A)Y \\ & \text{s.t.} & \operatorname{bdiag}(Y) \preceq I \\ & & \operatorname{odiag}(Y) \preceq I \\ & & Y \succeq 0. \end{array}$$

Note that the objective of QAPT is $\operatorname{tr} AXBX^T = \operatorname{vec}(X)^T(B \otimes A)\operatorname{vec}(X) = \operatorname{tr}(B \otimes A)\operatorname{vec}(X)\operatorname{vec}(X)^T$. The problem QAPSDP can be derived directly from QAPT by relaxing $\operatorname{vec}(X)\operatorname{vec}(X)^T$ to an $n^2 \times n^2$ matrix $Y \succeq 0$. For $Y = \operatorname{vec}(X)\operatorname{vec}(X)^T$, note that $Y_{[ij]} = X_i X_j^T$, where X_i is the ith column of X. It follows that for such a Y,

$$\operatorname{bdiag}(Y) = XX^T, \qquad \operatorname{odiag}(Y) = X^T X, \tag{10}$$

so the constraints of QAPSDP are natural extensions of the conditions $X^T X \preceq I$ and $XX^T \preceq I$ to an arbitrary $Y \succeq 0$.

Since DQAPT and QAPSDP both have interior solutions strong duality must hold between these programs [2]. It follows that any optimal solutions Y and S, T satisfy the following optimality conditions:

$$\begin{aligned} Y \succeq 0,\ \operatorname{bdiag}(Y) \preceq I,\ \operatorname{odiag}(Y) &\preceq I, \\ S \succeq 0,\ \operatorname{tr} S(I - \operatorname{bdiag}(Y)) &= 0, \\ T \succeq 0,\ \operatorname{tr} T(I - \operatorname{odiag}(Y)) &= 0, \\ (B \otimes A) + (I \otimes S) + (T \otimes I) &\succeq 0, \\ \operatorname{tr} Y\Big((B \otimes A) + (I \otimes S) + (T \otimes I)\Big) &= 0. \end{aligned} \tag{11}$$

Theorem 9 *The matrix X is optimal for QAPT if and only if there exist symmetric matrices $S \succeq 0$, $T \succeq 0$ such that*

$$\begin{aligned} XX^T &\preceq I, && \textit{primal feasibility} \\ \operatorname{tr} S(I - XX^T) &= 0, && \textit{complementary slackness} \\ \operatorname{tr} T(I - X^T X) &= 0, && \textit{complementary slackness} \\ AXB + SX + XT &= 0, && \textit{stationarity} \\ (B \otimes A) + (I \otimes S) + (T \otimes I) &\succeq 0. && \textit{strengthened second order} \end{aligned}$$

Proof. From Theorem 6 there is an X with $XX^T \preceq I$ so that $Y = \operatorname{vec}(X)\operatorname{vec}(X)^T$ is optimal in QAPSDP. For such a Y, note that

$$\begin{aligned} Y\Big((B \otimes A) + (I \otimes S) + (T \otimes I)\Big) &= \operatorname{vec}(X)\operatorname{vec}(X)^T\Big((B \otimes A) + (I \otimes S) + (T \otimes I)\Big) \\ &= \operatorname{vec}(X)\left(\Big((B \otimes A) + (I \otimes S) + (T \otimes I)\Big)\operatorname{vec}(X)\right)^T \\ &= \operatorname{vec}(X)\operatorname{vec}(AXB + SX + XT)^T. \end{aligned} \tag{12}$$

But $Y \succeq 0$, $(B\otimes A)+(I\otimes S)+(T\otimes I) \succeq 0$, and $\operatorname{tr} Y\Big((B\otimes A)+(I\otimes S)+(T\otimes I)\Big) = 0$ together imply that $Y\Big((B \otimes A) + (I \otimes S) + (T \otimes I)\Big) = 0$, so (12) implies that $AXB + SX + XT = 0$. The remaining conditions follow from (11) and (10). ■

Notice that the conditions of Theorem 9 are equivalent to the usual second order necessary conditions for optimality, except for the fact that the Hessian of the Lagrangian is positive semidefinite everywhere rather than on just the tangent space at the optimum.

It is interesting to examine the optimality conditions of Theorem 9 in the case of Example 8, which provided a counterexample to the conjectured conditions (9). Since in this case A and B are diagonal it is easy to see that S and T may also be taken to be diagonal matrices $S = \operatorname{Diag}(s)$, $T = \operatorname{Diag}(t)$. The conditions $AXB + SX + XT = 0$ then become

$$\begin{aligned} -6 + s_1 + t_1 = 0, &\qquad t_1 = 6 - s_1 \geq 0, \\ -1 + s_2 + t_2 = 0, &\qquad t_2 = 1 - s_2 \geq 0, \end{aligned} \tag{13}$$

implying $0 \leq s_1 \leq 6$, $0 \leq s_2 \leq 1$. Since $X^T X = XX^T = I$, to satisfy the conditions of Theorem 9 it remains only to satisfy the strengthened second order condition, which can be written

$$\begin{aligned} -6 + s_1 + t_1 &\geq 0, \\ -3 + s_2 + t_1 &\geq 0, \\ -2 + s_1 + t_2 &\geq 0, \\ -1 + s_2 + t_2 &\geq 0. \end{aligned} \tag{14}$$

The first and fourth inequalities of (14) are satisfied with equality, from (13). Using (13) to eliminate t_1 and t_2, the second and third inequalities of (14) can be written

$$\begin{aligned} -3 + s_2 + (6 - s_1) = 3 + s_2 - s_1 &\geq 0, \\ -2 + s_1 + (1 - s_2) = -1 + s_1 - s_2 &\geq 0. \end{aligned}$$

Thus we require (s_1, s_2) having

$$0 \leq s_1 \leq 6,\ 0 \leq s_2 \leq 1,\ 1 \leq s_1 - s_2 \leq 3,$$

which is a feasible system of constraints; for example $s_1 = 4$, $s_2 = 1$, $t_1 = 2$, $t_2 = 0$ provide S and T such that the conditions of Theorem 9 are satisfied.

References

[1] A. ALFAKIH, S. KRUK, and H. WOLKOWICZ. A note on geometry of semidefinite relaxations. Technical Report in progress, University of Waterloo, Waterloo, Canada, 1996.

[2] F. ALIZADEH. Interior point methods in semidefinite programming with applications to combinatorial optimization. *SIAM Journal on Optimization*, 5:13–51, 1995.

[3] K.M. ANSTREICHER and H. WOLKOWICZ. On Lagrangian relaxation of quadratic matrix constraints. Research report, corr 98-24, University of Waterloo, Waterloo, Ontario, 1998.

[4] D. BERTSIMAS and Y. YE. Semidefinite relaxations, multivariate normal distributions, and order statistics. Technical report, Department of Management Sciences The University of Iowa, 1998.

[5] R. BHATIA. *Perturbation Bounds for Matrix Eigenvalues : Pitman Research Notes in Mathematics Series 162*. Longman, 1987.

[6] F. CELA. *The Quadratic Assignment Problem: Theory and Algorithms*. Kluwer, Massachessets, USA, 1998.

[7] ALAN EDELMAN, TOMAS ARIAN, and STEVEN T. SMITH. The geometry of algorithms with orthogonality constraints. *SIAM J. Matrix Anal. Appl.*, ??:???–???, 1998.

[8] G. FINKE, R.E. BURKARD, and F. RENDL. Quadratic assignment problems. *Annals of Discrete Mathematics*, 31:61–82, 1987.

[9] M. FU, Z. LUO, and Y. YE. Approximation algorithms for quadratic programming. *Journal of Combinatorial Optimization*, 2(1):29–50, 1998.

[10] T. FUJIE and M. KOJIMA. Semidefinite programming relaxation for nonconvex quadratic programs. *J. Global Optim.*, 10(4):367–380, 1997.

[11] M.X. GOEMANS and D.P. WILLIAMSON. Improved approximation algorithms for maximum cut and satisfiability problems using semidefinite programming. *Journal of Association for Computing Machinery*, 42(6):1115–1145, 1995.

[12] R.A. HORN and C.R. JOHNSON. *Matrix Analysis.* Cambridge University Press, New York, 1985.

[13] S. KARISCH. *Nonlinear Approaches for Quadratic Assignment and Graph Partition Problems.* PhD thesis, University of Graz, Graz, Austria, 1995.

[14] S.E. KARISCH and F. RENDL. Semidefinite programming and graph equipartition. In *Topics in Semidefinite and Interior-Point Methods*, volume 18 of *The Fields Institute for Research in Mathematical Sciences, Communications Series*, Providence, Rhode Island, 1998. American Mathematical Society.
URL http://www.diku.dk/%7Ekarisch/rep302.ps.

[15] S.E. KARISCH, F. RENDL, and H. WOLKOWICZ. Trust regions and the quadratic assignment problem. In *Proceedings of the DIMACS Workshop on Quadratic Assignment Problems*, volume 16 of *DIMACS Series in Discrete Mathematics and Theoretical Computer Science*, pages 199–219. American Mathematical Society, 1994.

[16] M. KOJIMA and L. TUNCEL. Cones of matrices and successive convex relaxations of nonconvex sets. Technical Report B-338, Dept. of Mathematical Sciences, Tokyo Institute of Technology, Tokyo, Japan, 1998.

[17] S. KRUK and H. WOLKOWICZ. SQ^2P, sequential quadratic constrained quadratic programming. Research report, corr 97-01, University of Waterloo, Waterloo, Ontario, 1997. accepted (subject to revision) for the Proceedings of Nonlinear Programming Conference in Beijing in honour of Professor M.J.D. Powell.

[18] Z-Q. LUO and S. ZHANG. On the extension of Frank-Wolfe theorem. Technical report, Erasmus University Rotterdam, The Netherlands, 1997.

[19] J.J. MORÉ and D.C. SORENSEN. Computing a trust region step. *SIAM J. Sci. Statist. Comput.*, 4:553–572, 1983.

[20] A. NEMIROVSKI, C. ROOS, and T. TERLAKY. On maximization of quadratic form over intersection of ellipsoinds with common center. Technical report, Delft University of Technology, Faculty of Technical Mathematics and Informatics, Delft University of Technology, Delft, The Netherlands, 1998.
URL http://ssor.twi.tudelft.nl/ ˜terlaky/files/ELLIP.DVI.

[21] Y. E. NESTEROV. Global quadratic optimization via conic relaxation. Technical report, CORE, Universite Catholique de Louvain, Belgium, 1998.

[22] Y. E. NESTEROV. Semidefinite relaxation and nonconvex quadratic optimization. *Optimization Methods and Software*, 9:141–160, 1998. Special Issue Celebrating the 60th Birthday of Professor Naum Shor.

[23] M.L. OVERTON and R.S. WOMERSLEY. Optimality conditions and duality theory for minimizing sums of the largest eigenvalues of symmetric matrics. *Mathematical Programming*, 62:321–357, 1993.

[24] P. PARDALOS, F. RENDL, and H. WOLKOWICZ. The quadratic assignment problem: A survey and recent developments. In *Proceedings of the DIMACS Workshop on Quadratic Assignment Problems*, volume 16 of *DIMACS Series in Discrete Mathematics and Theoretical Computer Science*, pages 1–41. American Mathematical Society, 1994.

[25] F. RENDL. Semidefinite programming and combinatorial optimization. Technical report, University of Graz, Graz, Austria, 1997.

[26] F. RENDL and H. WOLKOWICZ. Applications of parametric programming and eigenvalue maximization to the quadratic assignment problem. *Mathematical Programming*, 53:63–78, 1992.

[27] F. RENDL and H. WOLKOWICZ. A semidefinite framework for trust region subproblems with applications to large scale minimization. *Mathematical Programming*, 77(2):273–299, 1997.

[28] S.A. SANTOS and D.C. SORENSEN. A new matrix-free algorithm for the large-scale trust-region subproblem. Technical Report TR95-20, Rice University, Houston, TX, 1995.

[29] R. STERN and H. WOLKOWICZ. Indefinite trust region subproblems and nonsymmetric eigenvalue perturbations. *SIAM J. Optimization*, 5(2):286–313, 1995.

[30] Y. YE. A new complexity result on minimization of a quadratic function with a sphere constraint. In *Recent Advances in Global Optimization*, pages 19–31. Princeton University Press, 1992.

[31] Y. YE. Approximating quadratic programming with bound and quadratic constraints. Technical report, Department of Management Sciences The University of Iowa, Iowa city, IA 522542, 1998. To appear in Mathematical Programming.

[32] Q. ZHAO, S.E. KARISCH, F. RENDL, and H. WOLKOWICZ. Semidefinite programming relaxations for the quadratic assignment problem. *J. Combinatorial Optimization*, 2.1, 1998. CORR 95-27,
URL: ftp://orion.uwaterloo.ca/pub/henry/reports/qapsdp.ps.gz.