Presolving for Semidefinite Programs Without Constraint Qualifications

Gerald Gruber* Serge Kruk † Franz Rendl ‡ Henry Wolkowicz §

July 28, 1998

PROCEEDINGS of HPOPT97
Second workshop on High Performance Optimization Techniques
Rotterdam, Netherlands

University of Waterloo
Department of Combinatorics and Optimization
Waterloo, Ontario N2L 3G1, Canada

Research Report CORR 98-32

Key words: Semidefinite Programming, Ill-posed Problems, Presolving, Constraint Qualifications, Regularization.

Abstract

Presolving for linear programming is an essential ingredient in many commercial packages. This step eliminates redundant constraints and identically zero variables, and it identifies possible infeasibility and unboundedness. In semidefinite programming, identically zero variables corresponds to lack of a constraint qualification which can result in both theoretical and numerical difficulties. A nonzero duality gap can exist which nullifies the elegant and powerful duality theory. Small perturbations can result in infeasibility and/or large perturbations in solutions. Such problems fall into the class of ill-posed problems.

It is interesting to note that classes of problems where constraint qualifications fail arise from semidefinite programming relaxations of hard combinatorial problems. We look at several such classes and present two approaches to find regularized solutions. Some preliminary numerical results are included.

Contents

*Technische Universität Graz, Institut für Mathematik, Kopernikusgasse 24, A-8010 Graz, Austria, Email: gruber@opt.math.tu-graz.ac.at

†Research supported by Natural Sciences Engineering Research Council Canada. E-mail sgkruk@barrow.uwaterloo.ca

‡Technische Universität Graz, Institut für Mathematik, Kopernikusgasse 24, A-8010 Graz, Austria, Email: rendl@fmatbds01.tu-graz.ac.at

§Research supported by Natural Sciences Engineering Research Council Canada. E-mail hwolkowi@orion.uwaterloo.ca.

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

1 Introduction

> But in many applied problems we have to get along without a precise knowledge of the causes, and in others we are really trying to find *causes* that will produce a desired effect. We are then led to *ill-posed* problems. One might say that the majority of applied problems are, and always have been, ill-posed, particularly when they require numerical answers.
> ....
> Past experience suggests that the concepts and methods used in the discussion of ill-posed problems will in turn stimulate advances in *pure* mathematical analysis.
>
> Fritz John, Translation Editor's Preface in [65]

The presolve step has proven to be *essential* in simplex and interior-point methods for linear programming (LP). This step finds redundant constraints and fixed variables in order to reduce the complexity of the problem. It also identifies infeasibility and unboundedness. In semidefinite programming (SDP) there is an additional complication in that a duality gap can exist if no constraint qualification holds. Thus presolving can be essential in both regularizing ill-posed problems as well as reducing the complexity.

In this paper we present some of the ideas needed for presolving. We emphasize the problems where Slater's constraint qualification may fail.

1.1 Notation

The (primal) SDP we consider is an extension of LP where nonnegativity componentwise is replaced by positive semidefiniteness on matrix variables. In fact, SDP looks just like LP, e.g.

$$(\mathbf{SDP}) \quad p^* = \begin{array}{ccc} \min & b^t y & (\langle b, y\rangle) \\ \text{subject to} & \mathcal{A}^* y \succeq C & \text{(linear constraints)} \end{array}$$

and its dual

$$(\mathbf{DSDP}) \quad \begin{array}{cccc} d^* = & \max & \operatorname{trace} CX & (\langle C, X\rangle) \\ & \text{subject to} & \mathcal{A}X = b & \text{(linear constraints)} \\ & & X \succeq 0, \quad (X \in \mathcal{P}) & \text{(nonneg. constraints)} \end{array}$$

where A, B are in the space of real $n \times n$ symmetric matrices, denoted $\mathcal{S}_n$, $\preceq$ denotes the Löwner partial order, i.e. $A \preceq B$ if $B - A \succeq 0$, $B - A$ is symmetric positive semidefinite;

$$\mathcal{A} : \mathcal{S}_n \to \Re^m$$

is a linear operator defined by

$$(\mathcal{A}X)_i = \langle A_i, X \rangle = \operatorname{trace}(A_i X), \text{ for given } A_i \in \mathcal{S}_n;$$

and the adjoint operator is

$$\mathcal{A}^* y = \sum_i y_i A_i.$$

We let $\mathcal{P}$ denote the cone of positive semidefinite matrices, i.e. $\mathcal{P}$ in SDP replaces $\Re^n_+$, the nonnegative orthant, in LP.

Remark 1 *We have used the customary dual as our primal. This is to facilitate the notation for the coming discussions on convex programs. When discussing LP, we often denote $\mathcal{A}X$ for the customary linear constraint Ax. The dual constraint is equivalent to $\mathcal{A}^* y + Z = C$, where $Z \succeq 0$ is the (nonnegative) slack variable.*

By a cone K we mean a set closed under addition and nonnegative scalar multiplication, $K + K \subset K$, $\alpha K \subset K$, $\forall \alpha \geq 0$. The cone $T \subset K$ is a *face* of the cone K, denoted $T \lhd K$, if

$$x, y \in K, \ x + y \in T \Rightarrow x, y \in T.$$

Each face, $K \lhd \mathcal{P}$, is characterized by a subspace, $S \subset \Re^n$. Let W be $n \times k$ with columns spanning S. Then

$$K = \{X \in \mathcal{P} : \mathcal{N}(X) \supset S\} = \mathcal{P} \cap \{WW^t\}^\perp.$$

Moreover,

$$\operatorname{relint} K = \{X \in \mathcal{P} : \mathcal{N}(X) = S\}.$$

The complementary (or conjugate) face of K is $K^c = K^\perp \cap \mathcal{P}$

$$K^c = \{X \in \mathcal{P} : \mathcal{N}(X) \supset S^\perp\}.$$

Moreover,

$$\operatorname{relint} K^c = \{X \in \mathcal{P} : \mathcal{N}(X) = S^\perp\}.$$

The face K (respectively, K^c) is determined by the supporting hyperplane corresponding to any $X \in \operatorname{relint} K^c$ (respectively, $\operatorname{relint} K$); and

$$XY = 0, \forall X \in K, Y \in K^c.$$

1.2 Ill-Posed Problems

Ill-posed problems are roughly defined as those for which the optimum is not stable with respect to perturbations in the data, e.g. [65] and [23, 24] for the SDP case. Regularization [65] can be used to get a unique optimal solution which is continuous with respect to perturbations in the data. In other words, regularization allows one to solve a nearby problem.

We now look at two possible cases of ill-posedness. These two cases correspond to the conditions in the Mangasarian-Fromovitz [47] constraint qualification, M-F CQ, which is known to be equivalent to stability with respect to right hand side perturbations. (This constraint qualification for the primal problem corresponds to the Slater constraint qualification: the existence of a strictly feasible point. The constraint qualification for the dual problem states: there exists a strictly feasible point and the linear operator $\mathcal{A}$ is onto or full rank.) For LPs, one tries to identify which constraints are redundant and which variables are (nearly) identically 0 on the feasible set, thus, discarding both reduces the dimension and complexity of the LP problem and obtains M-F CQ. In fact, this type of preprocessing where one identifies redundant constraints and redundant variables (as well as infeasibility or unboundedness) is *essential* in many commercial LP codes, since it drastically reduces the size of the problem and avoids many instances of ill-posedness and *stalling and/or cycling*, see [46, 3] and the references therein. Moreover, the identification of identically zero variables is often done dynamically during the course of the solution algorithm.

In the case of SDPs, the presolve is doubly important. It does not only reduce the size of the problem but it can avoid a duality gap which prevents the proper solution of the problem.

1.2.1 Linear Equality Constraints

LP problems are considered to rarely have numerical instabilities. One contrary case is when the linear equality constraint in the dual has a matrix which is not (numerically) full rank or is close to being not full rank. In fact, most descriptions of the simplex method make this assumption from the start. In the case of linear dependence, one usually finds a submatrix by *discarding* redundant constraints, i.e. one makes a decision on the numerical rank of the problem. To identify if there is a redundancy one (implicitly) solves the system

$$\text{find } y \neq 0 \text{ such that } \langle y, \mathcal{A}X \rangle = 0, \forall X, \tag{1}$$

or equivalently, since $\mathcal{A}$ is linear,

$$\min e^t y \text{ subject to } \mathcal{A}^* y = 0. \tag{2}$$

An optimal value of 0 denotes full rank, while a value of $-\infty$ denotes linear dependence. One can normalize the problem by adding the linear constraints $||y||_1 \leq 1$ so that the optimal value will be finite in both cases. However, the exact determination of 0 on a computer is usually impossible without exploiting special structure.

We can modify the problem so that we can determine full rank in a numerically (backwards) stable way. To do this we can modify the following and solve it as a linear program

$$\min ||\mathcal{A}^* y||_1 \text{ subject to } ||y||_1 = 1. \tag{3}$$

In this case, the optimal value is positive if $\mathcal{A}$ is full rank and this value is numerically stable with respect to small perturbations in $\mathcal{A}$. This is equivalent to finding the smallest singular value of $\mathcal{A}$. One can then discard the row corresponding to the singular values that are numerically 0.

Note that the original linear system $Ax = b$ can be modified and solved as an LP without worrying about linear dependence, i.e. we can use

$$\begin{array}{ll} \min & e^t z \\ \text{s.t.} & \begin{bmatrix} A & -A & I \end{bmatrix} \begin{pmatrix} x' \\ x'' \\ z \end{pmatrix} = b \\ & x', x'', z \geq 0. \end{array} \tag{4}$$

However, the solution is not unique and so this problem is not well-posed. An optimal value being zero or positive denotes feasibility or infeasibility, respectively. However, this is not done because it does not avoid the ill-posedness of the problem and actually can cause instability when none existed [61, 57].

A similar identification and correction can be done for the linear equality constraints of SDP problems.

1.2.2 Identically Zero Variables

A second case of instability can occur if one of the variables in the dual LP is (numerically) identically 0 on the complete dual feasible set. In this case small perturbations in the data can result in infeasibility. Or we can look at the primal LP and look for a slack variable that is identically 0 on the primal feasible set. One can identify this latter problem by solving

$$\text{find } X \geq 0, X \neq 0 \text{ such that } \quad \langle X, \mathcal{A}^* y - C \rangle = 0, \ \forall y \text{ with } \mathcal{A}^* y \geq C. \tag{5}$$

Equivalently we can solve

$$\min \langle C, X \rangle + \alpha \text{ subject to } \mathcal{A}X = 0, \langle I, X \rangle + \alpha = 1, X \geq 0, \tag{6}$$

where I denotes a unit element, e.g. e the vector of ones. The second equality constraint guarantees that X is not 0. A positive optimal value guarantees the existence of a strictly feasible point for PSDP and this optimal value is (backwards) stable with respect to small perturbations in the data. The existence of a strictly feasible point can be seen easily from the dual program of (6).

$$(\mathbf{P}') \quad \begin{array}{ll} \max & w \\ \text{s.t.} & \mathcal{A}^* y + wI \preceq C \\ & w \leq 1 \\ & y \in \Re^m, w \in \Re. \end{array}$$

A similar identification can be done for SDP problems.

Unlike the correction for redundant linear constraints, a correction for the failure of the M-F CQ does not follow through in a straightforward way from LP to SDP. In fact, such SDPs can have a nonzero duality gap. One approach to handle this case for general cone programs is the regularization procedure presented in [11]. This approach projects the problem onto the minimal face of the cone of semidefinite matrices. Thus we get a lower dimensional problem with fewer constraints. This is currently ignored in the literature since there is no known efficient technique for identifying the minimal face of the semidefinite cone. Rather, what is currently done is to use an infeasible approach and hope for the best.

We note that one can add additional variables to ensure that there exists a strictly feasible solution, e.g. using the self-dual embedding [36, 37].

Instances where the CQ fails in LP abound. As mentioned above, this preprocessing is essential in many public domain codes. Instances where the CQ fails in SDP applications exist as well. In fact, there are several classes of problems where it is shown that the CQ fails for SDP relaxations of hard combinatorial problems, see [75, 74].

1.3 Background and History of SDP

SDP has become a very *hot* area of research. Why has SDP become so interesting?

In [15], George Dantzig describes his first presentation of the simplex method to a very distinguished audience at a meeting of the Econometric Society in Wisconsin, in 1948. One very well-known statistician, Hotelling, stood up and stated: "But we all know the world is nonlinear." Before Dantzig could compose a reply, Jon von Neumann requested to reply for him: "The speaker titled his talk 'linear programming' and carefully stated his axioms. If you have an application that satisfies the axioms, well use it. If it does not, then don't." But then Dantzig mentions in the article: "In the final analysis, of course, Hotelling was right. The world is highly nonlinear. Fortunately systems of linear inequalities (as opposed to equalities) permit us to approximate most of the kinds of nonlinear relations encountered in practical planning." To summarize, linear programming is so popular because we can efficiently solve LP problems.

But, more often than not, quadratic approximations are better than linear approximations. *But, quadratic approximations are too hard to solve!* However, we can solve relaxations of quadratic approximations efficiently using semidefinite programming.

1.3.1 History of SDP

Though SDP has recently drawn a great deal of attention, its history is almost as old as linear programming. The set of semidefinite matrices is a closed convex cone in the set of symmetric matrices. It is arguably the most important set of matrices and has applications in many diverse areas. The study of the geometry of the cone of semidefinite matrices predates LP, e.g. by Bohnenblust in 1948, [9], and further in [4]. In fact, SDP is a special case of cone programming, studied by e.g. Duffin in 1956, [17]. The more general topic of cone programming was a major area of study both in finite dimensional and abstract spaces e.g. [20, 8, 13, 14]. Observations about LP like duality using a Slater-like constraint qualification appear in 1963 by Bellman and Fan, [6]; while a strong duality without a constraint qualification is developed in [10, 72] with a regularization algorithm in [11]. Semidefinite programming itself appeared in the literature in many places, specifically statistics (covariance matrices, clustering) [51]; and control theory (Lyapunov stability, Ricatti equations), e.g. [51, 67, 66] where interior-point methods such as the ellipsoid algorithm appeared; other optimization applications of SDP occurred in [21, 22, 64]. A related area, positive semidefinite completion problems, was studied in [29] as well as many following papers; survey papers appear in [35, 43] as well as a paper in this volume. Before the seminal work of Nestorov and Nemirovsky [48, 49], algorithmic approaches for cone programs (and SDP) were (relatively) inefficient and were first based on discretization and semi-infinite programming, e.g. [26], and later, in the control theory area, LMIs were solved by the ellipsoid algorithm [67]. But the fact that the function $\log\det$ is a valid barrier for the cone of positive semidefinite matrices follows from its strict concavity which is proved and used in [29], though it was well-known earlier on, e.g. [5, 18]. The important framework for barrier methods for general cone programs, including SDP, only appeared in [48, 49]. That potential function approaches as well as other interior-point approaches from LP can be extended to SDP in a unified way was shown in [1], where it appears that the term *semidefinite programming* first originated. Also at this time there appeared several important theoretical and algorithmic results that dealt with applications of SDP to combinatorial optimization, see e.g. [44, 45, 30, 28] and the survey papers [27, 56, 2]. The combination of the strong numerical results, theoretical breakthroughs, and diverse applications has led to an explosion of papers in this area.

1.4 Outline

In this paper we present two approaches on handling ill-posed SDPs, i.e. we show how to handle SDPs for which the Slater CQ (strict feasibility) fails. One approach reduces the dimension of the problem to get a problem for which the CQ holds; or decides that the problem is infeasible. The other approach deals directly with the existing problem, but with a modified convergence criteria.

In Section 2 we start with ordinary convex programs, OCP, and show how lack of CQ can result in duality problems. We also introduce the regularization process. We then show how to move to cone constrained programs. We include some historical notes on Lagrange multipliers.

The special case of duality and regularization in SDP is reviewed in Section 3. The numerical tests appear in Section 4 We conclude with comments in Section 5.

2 Strong Duality in Convex Programming

We first consider the *ordinary convex programming problem*:

$$\text{OCP} \qquad \min\{f(x) : f^k(x) \le 0, k \in \mathcal{I} := \{1, \ldots, m\}\}$$

where the functions are all real valued and convex; and the constraints are also faithfully convex (see below). We follow closely the approaches in [7, 70, 71, 73], where missing details and references can be found.

2.1 Decomposition of Convex Functions

To understand what can cause duality gaps in OCP, we first look at what first order information is missing in a convex function. For simplicity, we restrict ourselves to the case that f is *faithfully convex*: f is affine on a line segment only if it is affine on the whole line containing that segment. For example, analytic convex and strictly convex functions are faithfully convex. In the case that f is faithfully convex, we have the decomposition

$$f(x) = h(Ax + b) + a^t x + c,$$

where h is strictly convex, A is $m \times n$ and $c \in \Re$.

To completely describe the local behaviour of f we need to define the *cones of directions at x of*: decrease; nonincrease; constancy; increase; nondecrease;

$$D_f^{\text{symbol}}(x) := \{d : f(x + \alpha d) \text{ symbol } f(x), \quad \forall 0 < \alpha < \bar{\alpha}, \quad \text{for some } \bar{\alpha}\},$$

where *symbol* represents $<, \le, =, >, \ge$, respectively.

Since f is convex, we see that

$$D_f^{<}(x) = \{d : \nabla f(x)^t d < 0\},$$

i.e. this is completely described using the first order information. However,

$$D_f^{>}(x) \supset \{d : \nabla f(x)^t d > 0\},$$

and

$$D_f^{\geq}(x) = \{d : \nabla f(x)^t d \ge 0\} \setminus D_f^{=}(x).$$

Thus the missing information is given by the cone of directions of constancy.

From our decomposition, we see that the cone $D_f^{=}$ is the null space of A intersect the null space of a, i.e.

$$D_f^{=} = \mathcal{N}(A) \cap \{a\}^{\perp}$$

which is independent of x. An algorithm for finding this cone is given in [69, 42].

2.2 Decomposition of Convex Programs

Now that we have seen how to decompose convex functions and get all the directional information, we continue to see how this helps regularize convex programs, see [73].

We let F denote the feasible set and

$$\mathcal{I}^{=} := \{k : x \in F \Rightarrow f^k(x) = 0\}$$

denote the implicit equality constraints. Then $\mathcal{I}^{<}$ denotes the complement of the equality constraints. The constancy cone

$$D^{\equiv}_{\mathcal{I}^=} = \cap_{k \in \mathcal{I}^=} D^{\equiv}_k$$

denotes the intersection of the cones of directions of constancy of the implicit equality constraints. This is independent of x since we are dealing with faithfully convex functions.

2.3 Characterizations of Optimality of OCP Without CQ

There are several different ways to characterize optimality for convex programs. Since the feasible set F is necessarily convex, one characterization is the Rockafellar-Pshnenichnyi condition.

Lemma 1 *The point $x \in F$ is optimal for OCP if and only if*

$$\partial f(x) \cap T^+(F, x) \neq \emptyset, \tag{7}$$

where $\partial f(x)$ denotes the subgradient of f at x, $T(F, x)$ denotes the tangent cone of F at x

$$T(F, x) = \mathrm{cl}\,(\mathrm{cone}\,(F - x)),$$

and $\cdot^+$ denotes the polar cone

$$S^+ = \{\phi : \langle \phi, s \rangle \geq 0, \quad \forall s \in S\}.$$

Proof. The proof follows from the definitions and the fact that x is optimal if and only if there is no feasible direction of decrease.

□

The above *geomatric* optimality characterization is complete but does not yield Lagrange multipliers directly. It uses the tangent cone to provide a local geometric description of the feasible set at x. To get Lagrange multipliers we need to do some arithmetic with polars. One can show using a hyperplane separation argument that if K, L are closed convex cones, then

$$(K \cap L)^+ = \mathrm{cl}\,(K^+ + L^+).$$

The closure is not needed if some constraint qualification holds, e.g. $\mathrm{int}\,(K) \cap L \neq \emptyset$.

The linearizing cone of an active set of constraints $\Omega \subset \mathcal{I}$ is the set of directions d with nonpositive directional derivative, i.e.

$$C_\Omega(x) := \{d : \phi^t d \leq 0, \quad \forall \phi \in \partial f_k(x), \quad \forall k \in \Omega\}.$$

This is a local *algebraic* description of the feasible set at x. Using a Farkas Lemma argument, one can show that

$$C^+_\Omega(x) = \mathrm{cl}\,(\mathrm{cone}\, B_\Omega(x)) := \mathrm{cl}\,\mathrm{cone}\,(\cup_{k \in \Omega} \partial f_k(x)),$$

i.e. the cone generated by the subgradients.

One now sees a connection between the geometric description of the feasible set given by the tangent cone and the algebraic description given by the linearizing cone. From the above, we see that the gap in this description is the cone of directions of constancy. In fact

$$T(S, x) = cl(\mathrm{conv}\; D^{\equiv}_{\mathcal{I}^=}(x)) \cap C_{active}(x).$$

Therefore the optimality characterization translates into the fact that a subgradient has to be in the set

$$T^+(S, x) = (D^{\equiv}_{\mathcal{I}^=})^+ - B_{active}(x), \tag{8}$$

where *active* denotes the active (tight) constraints. This is a Lagrange multiplier result, i.e. $x \in F$ is optimal if and only if

$$\begin{array}{c} \text{there exist } \phi \in \partial f(x), \quad \phi_k \in \partial f_k(x), \quad \lambda_k \geq 0 \\ \text{such that} \\ \phi + \sum_{k \text{ active}} \lambda_k \phi_k \in (D^{=}_{\mathcal{I}^=})^+. \end{array} \tag{9}$$

In addition, the above gives us a *geometric weakest constraint qualification*

$$T(S, x) = C_{active}(x).$$

The missing set was added on in order to complete the optimality conditions. Any appropriate set that satisfies (8) would do, i.e. we get in the differentiable case the more familiar Lagrange multiplier optimality characterization

$$\nabla f(x) + \sum_k \lambda_k \nabla f_k(x) \in K^+,$$

for the appropriate set K.

A primal approach to the regularization process consists in finding the implicit set of equality constraints and treating this set $\mathcal{I}^=$ and its complement $\mathcal{I}^<$ in different ways, i.e. if we denote

$$F^= := \{x : f_k(x) = 0, \quad \forall k \in \mathcal{I}^=\},$$

then the following program satisfies the generalized Slater's constraint qualification

$$\min\{f(x) : f_k(x) \leq 0, \ k \in \mathcal{I}^<, x \in F^=\}$$

We can apply the Pshnenichnyi-Rockafellar optimality conditions to get (9).

2.3.1 Regularization of OCP

From the above we see that the key to regularizing OCP is the identification of the implicit equality constraints, $\mathcal{I}^=$. Suppose that $\bar{x}$ is a feasible point, $g(\bar{x}) \leq 0$. If $\bar{x}$ is a Slater point, then we are done. Otherwise, if $\hat{x}$ is a Slater point, then the direction $d = \hat{x} - \bar{x}$ is a direction of decrease for each of the active constraints at $\bar{x}$. Therefore, such a Slater point does not exist if and only if

$$\begin{array}{c} \text{there exists } 0 \neq \lambda \geq 0 \text{ such that} \\ \lambda^t g(\bar{x}) = 0 \text{ (active set determination)} \\ \nabla g(\bar{x})\lambda = 0 \text{ (no common descent direction)}, \end{array} \tag{10}$$

where $\nabla g(\bar{x})$ is the Jacobian of g, i.e. the matrix of gradients of g_k. The constraints corresponding to $\lambda_k > 0$ are implicit equality constraints. We can then find the intersection of the cones of directions of constancy and project the problem onto a smaller dimensional space, say the range of the matrix P, i.e. we substitute $\bar{x} + Py$ and obtain a new problem in y. We then repeat the process. (See [73] for the details.)

To solve (10), we could solve a LP. We could then use the error analysis of LP to determine the accuracy of the subspace determined by P. If we had a bound on the optimal value of the original problem, we could now find bounds on the error in the optimal value.

2.4 Cone Programming and Constraint Qualifications

The OCP above can equivalently be considered as a cone constrained program with the constraint $F(x) \preceq_{\mathcal{R}^+} 0$, where $F(x)$ denotes the vector of constraints and $y \preceq_{\mathcal{R}^+} 0$ means nonpositive elementwise. Then the vector valued function F is convex with respect to the partial order induced by the cone R^+. The regularization is then equivalent to restricting the constraints to the minimal face, denoted $\mathcal{R}^f$, of the nonnegative orthant that contains the image of the feasible set and adding the set constraint

$$x \in F^{-1}(span\ (\mathcal{R}^f)) = F^{-1}(\mathcal{R}^f - \mathcal{R}^f) = F^{-1}(\mathcal{R}^f - \Re^m_+).$$

We now look at the optimality conditions for general cone programming. We let S be a closed convex cone. This induces a linear partial order

$$x \preceq_S y \text{ means } x - y \in -S.$$

The general cone program is then

$$\min\{f(x) : g(x) \preceq_S 0, \ x \in \Omega\}.$$

We restrict attention to the convex case, i.e. Ω is a convex set and g is S-convex or convex with respect to the induced partial order. We again denote the feasible set by F.

To regularize the problem, We follow the approach of using the minimal face of the cone S, see [11]. This is equivalent to solving the system (10) for OCP, i.e.

$$\begin{array}{c} \text{there exists } \Lambda \succeq_{S^+} 0 \text{ such that} \\ \langle \Lambda, g(\bar{x}) \rangle = 0 \text{ (active set determination)} \\ \langle \Lambda, \nabla g(\bar{x}) \rangle = 0 \text{ (no descent direction).} \end{array} \tag{11}$$

The conjugate face determined by Λ contains the image of the feasible set, $g(F)$. We can now project onto this face and then repeat the process to find the minimal face containing the image of the feasible set.

We can rewrite the program using the minimal face of S that contains the image of the feasible set $g(F) \subset -S^f$, i.e. we get

$$\min\{f(x) : g(x) \preceq_{S^f} 0, \ x \in \Omega^f\}$$

where

$$\Omega^f = \Omega \cap g^{-1}(S^f - S^f) = \Omega \cap g^{-1}(S^f - S).$$

The optimality conditions without any constraint qualification become: There exists a Lagrange multiplier vector $\Lambda \in (S^f)^+$ such that

$$f(x) + \Lambda g(x) \geq \mu^*, \ \forall x \in \Omega^f.$$

In the special case of linear cone programming, where K, L are closed convex cones, we get [72] the following dual pair for which strong duality holds.

The primal problem is:

$$\mu^* := \min\{cx : \mathcal{A}x \succeq_K b, \ x \succeq_L 0\}$$

The minimal cones are denoted K^f, L^f, i.e. $F \subset L^f$, and $\mathcal{A}(F) \subset K^f$. The dual program for which strong duality and attainment holds is

$$\mu^* = \nu^* := \max\{by : A^* y \preceq_{(L^f)^+} b, \ y \succeq_{(K^f)^+} 0\}$$

2.5 Some History on Lagrange Multipliers

The importance of Lagrange multipliers in optimization is well known. Its history has been discussed in several places and we include some details here.

In fact, Lagrange multipliers are actually due to Leonard Euler. But their importance is highlighted by their use by Euler's student, Lagrange, in his master's thesis. Thus the λ notation is in honour of both Lagrange and Leonard, see [50] where the basic multiplier rules are discussed and the details of the history are given.

Some interesting historical information is given in the 1976 article by Harold Kuhn in [40], where the following quote appears.

> The Lagrange multiplier method leads to several transformations which are important both theoretically and practically.
>
> By means of these transformations new problems equivalent to a given problem can be so formulated that stationary conditions occur simultaneously in equivalent problems. In this way we are led to transformations of the problems which are important because of their symmetric character. Moreover, for a given maximum problem with maximum M, we shall often be able to find an equivalent minimum problem with the same value M as minimum; this is a useful tool for bounding M from above and below. (Courant and Hilbert)

In addition, Kuhn credits the following as the first occurrence of duality in the literature.
Problem 1:

> Given three points in the plane, find a fourth point such that the sum of its distances to the three given points is minimum. (17th Century Fermat Problem)

Problem 2:

> In the three Sides of an equiangular Field stand three Trees, at the Distances of 10, 12, and 16 Chains from one another: To find the Content of the Field, it being the greatest the Data will admit of? (The Ladies Diary or Woman's Almanack (1755) problem posed by Mr Tho. Moss (pg 47))
>
> Given any triangle, circumscribe the largest possible equilateral triangle about it. (Annales de mathematiques Pures et Appliquees, edited by J.D. Gergonne, Vol I (1810-11), Problem posed on Page 384)

Solution:

> Thus the largest equilateral triangle circumscribing a given triangle has sides perpendicular to the lines joining the vertices of the given triangle to the point such that the sum of the distances to these vertices is a minimum.
> One can conclude that the altitude of the largest equilateral triangle that can be circumscribed about a given triangle is equal to the sum of distances from the vertices of the given triangle to the point at which the sum of distances is a minimum. (Rochat, Vecten, Fauguier, and Pilatte in vol II (1811-12) page 88-93)

In fact, Von Neumann had a duality for LP back in the 1940s, which was based on using game theory, [68]. Later Gale, Kuhn, Tucker [25] developed a duality theory for general problems. Duality theory for cone programs in abstract spaces followed as well, see e.g. [52, 32, 16, 17] and the references therein.

Optimality conditions for nonlinear programs date back to Fritz John [34]. Other related references and extensions to cone problems in abstract spaces include [41, 58]. The history of constraint qualifications is discussed in [40]. The essential equivalence between constraint qualifications and stability have been extensively studied, e.g. [59, 60, 58, 19]. The classical M-F CQ is introduced in [47].

Characterizations of optimality without constraint qualifications and discussions about zero duality gaps appear in [62, 63, 8, 38, 31, 39] and [7, 12, 11, 10, 72] for the convex and convex cone case.

3 Strong Duality in SDP

We now consider duality for the special cone programming problem SDP. We follow [10, 72, 55, 53]. For simplicity we consider linear objective and constraint functions; and, we consider the primal SDP in the following form.

$$
\textbf{(P)} \qquad \begin{array}{rcl} p^* := & \max & c^t x \\ & \text{s.t.} & \mathcal{A}x \preceq b \\ & & x \in \Re^m, \end{array}
$$

where $x \in \Re^m$, $b \in \mathcal{S}_n$. The Lagrangian dual is given as

$$
p^* \le d^* := \min_{U \succeq 0} \max_x L(x, U) = \langle U, b \rangle + x^t(c - \mathcal{A}^* U).
$$

The hidden constraint $c - \mathcal{A}^* U = 0$ yields the more familiar dual

$$
\textbf{(D)} \qquad \begin{array}{rcl} d^* = & \min & \operatorname{trace} bU \\ & \text{s.t.} & \mathcal{A}^* U = c \\ & & U \succeq 0. \end{array}
$$

If Slater's condition holds for (P),

$$
\exists \hat{x} \text{ such that } \mathcal{A}\hat{x} \prec b,
$$

then strong duality holds from the primal side, i.e. there is a zero duality gap and the dual is attained.

If Slater's condition fails for the primal SDP, then we would like to do something similar to what is done for OCP, i.e. find an equality set and project onto a smaller dimensional problem. For the OCP, we found the implicit equalities. If we consider the constraint $g(x) \leq 0$ in OCP as denoting containment in the nonpositive orthant

$$g(x) \in -\Re^m_+,$$

then the implicit equalities corresponds to finding a minimal face of -$\Re^m_+$ that contains $g(F)$, the image of the feasible set. We follow the approach in Section 2.4 and find the minimal face. This follows [11] and is outlined in [53] to show the comparison with the Ramana optimality conditions in [54].

The *minimal cone* of P is defined as

$$\mathcal{P}^f = \cap\{\text{faces of } \mathcal{P} \text{ containing } (b - \mathcal{A}(F))\}.$$

Therefore, an equivalent program is the *regularized P program*

$$\begin{array}{llll} & p^* = & \max & c^t x \\ (\mathbf{RP}) & & \text{s.t.} & \mathcal{A}x \preceq_{\mathcal{P}^f} b \\ & & & x \in \Re^m. \end{array}$$

And, there exists x such that $b - \mathcal{A}x \in \operatorname{relint} \mathcal{P}^f$, i.e. the generalized Slater's constraint qualification holds. Strong duality holds for the pair RP and

$$\begin{array}{llll} & p^* = & \min & \operatorname{trace} bU \\ (\mathbf{DRP}) & & \text{s.t.} & \mathcal{A}^* U = c \\ & & & U \succeq_{(\mathcal{P}^f)^+} 0. \end{array}$$

We note that the following duality theorem holds for the regularized primal-dual pair.

Theorem 1 *(***DUALITY THEOREM***)*

1. *If one of the problems is inconsistent, then the other is inconsistent or unbounded.*
2. **WEAK DUALITY**
Let the two problems be consistent, and let x^0 be a feasible solution for RP and U^0 be a feasible solution for RD. Then
$$c^t x^0 \leq \left\langle b, U^0 \right\rangle.$$
3. **STRONG DUALITY**
If both RP and RD are consistent, then their optimal values are equal and RD is attained.
4. **COMPLEMENTARY SLACKNESS**
Let x^0 and U^0 be feasible solutions of RP and RD, respectively. Then x^0 and U^0 are optimal if and only if
$$\left\langle U^0, (b - \mathcal{A}x^0) \right\rangle = 0.$$
if and only if
$$U^0(b - \mathcal{A}x^0) = 0.$$
5. **SADDLE POINT**
The vectors x^0, U^0 are optimal solutions of RP and RD, respectively, if and only if (x^0, U^0) is a saddle point of the Lagrangian $L(x, U)$ for all (x, U),
$$L(x, U^0) \leq L(x^0, U^0) \leq L(x^0, U)$$
and then
$$L(x^0, U^0) = c^t x^0 = \left\langle b, U^0 \right\rangle.$$

□

How do we find $\mathcal{P}^f$ and what are its properties? The following lemma presents a theorem of the alternative that allows us to find a proper face that contains the minimal cone.

Lemma 2 *Suppose $\mathcal{P}^f \lhd K \lhd \mathcal{P}$. Then the system*

$$\mathcal{A}^* U = 0, U \succeq_{K^+} 0, \operatorname{trace} Ub = 0$$

is consistent only if

$$\text{the minimal cone } \mathcal{P}^f \subset \{U\}^\perp \cap K.$$

Proof. Since $\operatorname{trace} U(\mathcal{A}x - b) = 0$, for all x, we get $\mathcal{A}(F) - b \subset U^\perp$, i.e. $\mathcal{P}^f \subset \{U\}^\perp$. □

The following property is rather surprising. This provides a geometric explanation of why the usual optimality conditions fail due to lack of closure; whereas, the regularized optimality conditions satisy the closure. (The proof is in [53].)

Lemma 3 *Suppose that $0 \neq K \lhd \mathcal{P}$ (proper face). Then*

$$\mathcal{P}^+ + K^\perp = cl(\mathcal{P}^+ + span\ K^c) \text{ is always closed.}$$

But

$$\mathcal{P}^+ + span\ K \text{ is never closed.}$$

□

We present the first step of the regularization procedure here. This applies Lemma 2. After the projection is done, this step can be repeated. (Recall that $\lhd$ denotes face and $\cdot^c$ denotes conjugate face.)

Step 1

Define $\mathcal{P}_0 := \mathcal{P}$ and

$$\mathcal{U}_1 := \{U \succeq 0 : \mathcal{A}^* U = 0, \operatorname{trace} Ub = 0\}.$$

Choose $\hat{U}_1 \in \operatorname{relint}(\mathcal{U}_1)$. (If $\hat{U}_1 = 0$, then Slater's condition holds for $(\mathbf{P})$ and we STOP. In fact, we do not have to have the relative interior restriction but use it to facilitate the comparison with the Ramana dual.) Further, let

$$\mathcal{P}_1 := (\mathcal{F}(\mathcal{U}_1))^c \; (= \{\hat{U}_1\}^\perp \cap \mathcal{P}_0 \lhd \mathcal{P}_0).$$

We can now define the following equivalent program to $(\mathbf{P})$ and its Lagrangian dual.

$$(\mathbf{RP_1}) \qquad \begin{array}{rll} p^* = & \max & c^t x \\ & \text{s.t.} & \mathcal{A}x \preceq_{\mathcal{P}_1} b \\ & & x \in \Re^m. \end{array}$$

$$(\mathbf{DRP_1}) \qquad \begin{array}{rll} d_1^* = & \min & \operatorname{trace} bU \\ & \text{s.t.} & \mathcal{A}^* U = c \\ & & U \succeq_{(\mathcal{P}_1)^+} 0. \end{array}$$

Note that $p^* \leq d_1^* \leq d^*$ and

$$(\mathcal{P}_1)^+ = (\mathcal{P} \cap \mathcal{P}_1)^+ = \mathcal{P} + (\mathcal{P}_1)^\perp.$$

Therefore

$$(\mathcal{P}_1)^+ = \mathcal{P} + ((\mathcal{F}(\mathcal{U}_1))^c)^\perp, \; (\mathcal{P}_1)^\perp = \mathcal{W}_1^S$$

and we get the following equivalent program to $(\mathbf{DRP_1})$.

$$(\mathbf{ELSD_1}) \qquad \begin{array}{rll} d_1^* = & \min & \operatorname{trace} b(U + (W + W^t)) \\ & \text{s.t.} & \mathcal{A}^*(U + (W + W^t)) = c \\ & & \mathcal{A}^* U_1 = 0, \operatorname{trace} U_1 b = 0 \\ & & U \succeq 0, \; \begin{bmatrix} I & W^t \\ W & U_1 \end{bmatrix} \succeq 0. \end{array}$$

This latter program is the first part of the Ramana ELSD dual program, [54, 55, 53]. The importance of the Ramana *dual* regularization process is that it is provable of polynomial size in the data.

If Slater's condition now holds for the first regularized step RP_1, then we can start to solve the problem. Otherwise we need to repeat the regularization step and project onto an even smaller problem.

Each step of the regularization can be expensive. Just as for LP one has to balance off the benefit of the regularization with its cost. However, one additional motivation is that the regularization allows one to solve a nearby problem; otherwise, one is not certain what is being solved, i.e. the optimal value is somewhere within the duality gap if one ignores regularization.

To apply Lemma 2 at the first step of the regularization, one can solve the (homogeneous) SDP

$$\begin{array}{ll} \min & \operatorname{trace} bU \\ \text{s.t.} & \mathcal{A}^* U = 0 \\ & \operatorname{trace} U = 1 \\ & U \succeq 0. \end{array} \tag{12}$$

If the optimal value is (numerically) 0, then we have found U that yields a proper face that contains the minimal face. However, have we just replaced one ill-posed problem with another? How do we ensure that this problem satisfies Slater's CQ? To do this we can preform several other preliminary presolving steps. For example, one can check if any $A_i \succeq 0$. For then we can regularize (12) using A_i, i.e. the feasible set of (12) is in the conjugate face of A_i. So, we can assume that no $A_i \succeq 0$ (or $\preceq 0$). In addition, we can check if b_i is linearly independent of the span of the A_i. If not, than any feasible point is optimal so we can change b to 0.

4 Computational Considerations

In this section we look at various ways of dealing with SDPs for which the Slater CQ fails. We consider two issues separately. The first, and easier one, is to tackle feasible SDPs for which there is no duality gap, but Slater's condition fails for at least one of PSDP and DSDP. A more subtle issue is to recognize the feasibility status of an SDP. The difficulty here lies in the fact, that an SDP can be infeasible, but there could exist asymptotically feasible sequences. We try to determine the feasibility status by solving certain auxiliary SDPs.

There exist various ways of dealing with the lack of CQ. For instance one could use the skew-symmetric embedding as mentioned in the introduction. In our present approach we propose an alternative, which is based on the so-called HKM [33] search direction. The key idea is to modify the HKM direction so as to slow down the convergence towards primal and dual feasibility.

Let **(P)** and **(D)** be as in Section 3. The search direction is computed by solving the following system of equations (KKT-System).

$$\begin{array}{rcl} \mathcal{A}^*(\Delta U) & = & -\delta(\mathcal{A}^*(U) - c) \\ \Delta W + \mathcal{A}^{\mathbf{t}}(\Delta y) & = & -\delta(W + \mathcal{A}(x) - b) \\ \Delta W U + W \Delta U & = & WU - \sigma\mu I \end{array} \tag{13}$$

where $0 < \delta \leq 1$. The damping factor $\delta < 1$ prevents U and W from approaching the boundary of $\mathcal{P}$ prematurely. Using this search direction we perform the line search and update in the usual way.

In the following sections we present some computational results. We denote by n the size of the problem, m is the number of constraints, and iter represents the number of iterations. (R_p) (R_d) stands for the primal (dual) residuum, respectively and gap is the duality gap. For vectors $||v||$, denotes the Euclidean norm of v. For matrices we write $||X||_F$ for the Frobenius norm of X, respectively.

4.1 Feasible SDPs without Constraint Qualification

First we consider feasible SDPs with a zero duality gap where *not* both **(P)** and **(D)** have interior feasible points. We consider two examples of test problems.

Example 1

$$(\mathbf{D}) \qquad \begin{array}{ll} \inf & \langle b, U\rangle \\ s.t. & \operatorname{diag}(U) = e \\ & \langle J, U\rangle = 0 \\ & U \succeq 0 \end{array}$$

where $e = (1, 1, \ldots, 1)$, J denotes the matrix of all ones, $\operatorname{diag}(U)$ is the vector of the diagonal entries of U and b is a random generated matrix in $\mathcal{S}_n$.

Obviously $(\mathbf{D})$ is feasible but the interior of the feasible set is empty. The corresponding $(\mathbf{P})$ has interior feasible points, and strong duality holds. Table 1 contains some computational results.

n	m	iter	$\|\|R_p\|\|_F$	$\|\|R_d\|\|$	gap	elapsed time
10	11	80	5.5230e-07	8.1368e-07	8.0641e-06	9.7
20	21	90	1.0844e-07	1.9315e-07	1.2090e-06	27.1
30	31	95	5.7089e-08	3.9180e-07	3.7956e-07	107.9

Table 1: Numerical Results for Example 1

Example 2

Next we generate a class of feasible problems having zero duality gap, where neither the primal nor the dual has an interior. The input data are generated as follows. We start with an orthonormal basis $(v_1, v_2, \ldots, v_n)$. Let $\mathcal{I}$ denote the index set $\{1, 2, \ldots, n\}$. Assume $\mathcal{J} \subseteq \mathcal{I}$, $|\mathcal{J}| = m$ and $b \in \Re^m$. We set $C = \sum_{i\in\mathcal{J}} \alpha_i A_i + \sum_{i\in\mathcal{I}\backslash\mathcal{J}} \beta_i A_i$, where $A_i = v_i v_i^T$, $i = 1, \ldots, n$ and $\alpha_i > 0$, $i \in \mathcal{J}$ and $\beta_i \geq 0$, $i \in \mathcal{I}\backslash\mathcal{J}$ are random numbers. We choose an arbitrary $k \in \{1, 2, \ldots, m\}$ and set $b_i = 1$, $\forall i \in \{1, 2, \ldots, m\}\backslash k$ and $b_k = 0$. For one arbitrary index $j \in \mathcal{I}\backslash\mathcal{J}$ we set $\beta_j = 0$ and $\beta_i > 0$ $\forall i \in \mathcal{I}\backslash\{\mathcal{J} \cup j\}$.

It is easily checked that $(\mathbf{P})$ and $(\mathbf{D})$ are feasible without strictly interior points and that the optimal solution is $\sum_{i\in\{1,2,\ldots,m\}\backslash k} \alpha_i$. Table 2 shows some computational results.

n	m	iter	$\|\|R_p\|\|_F$	$\|\|R_d\|\|$	gap	elapsed time
10	7	63	2.0431e-08	1.1849e-11	9.2849e-08	5.1
30	20	65	1.7941e-08	1.9862e-10	7.8882e-08	43.3
50	43	65	1.8039e-08	2.5823e-11	8.1138e-08	340.2

Table 2: Numerical Results for Example 2

Remark 2 *We achieve similar results if we set more of the b_i's and β_j's to zero.*

Discussion: Although the interior point assumption for the dual pair does not hold, such problems can be handled by using the damping strategy from (13). For both problem classes we are able to obtain feasible solutions in reasonable computation time. The duality gap is close to zero. We note however, that the number of Newton iterations is much larger than on typical problems where the central path exists.

4.2 Detecting Infeasibility

First we consider SDPs where $(\mathbf{P})$ is infeasible. In this case we introduce an auxiliary SDP called $\text{SDP}^{(\mathbf{P})}$. Its solution allows us to recognize whether or not the original $(\mathbf{P})$ is infeasible. $\text{SDP}^{(\mathbf{P})}$ reads as follows.

$$(\mathbf{P}') \qquad \begin{array}{ll} \sup & w \\ \text{s.t.} & \mathcal{A}(x) + wI \preceq b \\ & x \in \Re^m, w \in \Re \end{array}$$

$$(\mathbf{D}') \qquad \begin{array}{ll} \inf & \operatorname{trace} bU \\ \text{s.t.} & \mathcal{A}^*(U) = 0 \\ & \operatorname{trace}(U) = 1 \\ & U \succeq 0. \end{array}$$

Note:

- $(\mathbf{P})$ feasible $\Leftrightarrow w \geq 0$
- $(\mathbf{P})$ infeasible $\Leftrightarrow w < 0$
- $(\mathbf{P})$ satisfies Slaters condition $\Leftrightarrow w > 0$

Example 3

The problem data are as follows. We set $c = 0$, $b = \begin{bmatrix} 0 & E \\ E & 0 \end{bmatrix}$ where E is a block of ones of size k, k suitable, i.e., k depends on n and $\mathcal{A}^*(U) = \text{trace}\left[\begin{pmatrix} A & 0 \\ 0 & 0 \end{pmatrix} U\right] = \text{trace}\,(AU_{kk})$ where A denotes a matrix of order k in $\mathcal{P}$ and U_{kk} denotes the k-th leading principal matrix of U. It is easily verified that the optimal dual value is zero and that the feasible set of $(\mathbf{P})$ is empty. Table 3 shows the numerical results obtained by solving $\text{SDP}^{(\mathbf{P})}$ using our technique.

n	iter	$\|\|R_p\|\|_F$	$\|\|R_d\|\|$	gap	w	elapsed time
10	95	9.8282e-09	3.0648e-08	6.3821e-07	-3.7465e-04	3.1
20	97	9.7449e-09	3.9498e-08	8.6326e-06	-5.9502e-04	6.6
30	234	2.9578e-12	3.4950e-09	9.9411e-06	-1.0636e-05	26.6

Table 3: Numerical Results for Example 3

Now we propose an auxiliary SDP whose solution allows us to recognize whether or not the original $(\mathbf{D})$ is infeasible. We call it $\text{SDP}^{(\mathbf{D})}$. The test problem is example 1 with the only difference that we change the constraint $\langle J, U\rangle = 0$ to $\langle J, U\rangle = \varepsilon$ where $|\varepsilon|$ is small. We see that $(\mathbf{D})$ is feasible for $\varepsilon \geq 0$ and infeasible for $\varepsilon < 0$. $\text{SDP}^{(\mathbf{D})}$ reads as follows.

$$(\mathbf{P}') \quad \begin{array}{ll} \sup & c^t x \\ \text{s.t.} & -\mathcal{A}(x) \succeq 0 \\ & x^t \mathcal{A}^*(I) = -1 \\ & x \in \Re^m \end{array}$$

$$(\mathbf{D}') \quad \begin{array}{ll} \inf & w \\ \text{s.t.} & \mathcal{A}^*(U) - w\mathcal{A}^*(I) = c \\ & U \succeq 0. \end{array}$$

Note:

- $(\mathbf{D})$ feasible $\Leftrightarrow w \leq 0$
- $(\mathbf{D})$ infeasible $\Leftrightarrow w > 0$
- $(\mathbf{D})$ satisfies Slaters condition $\Leftrightarrow w < 0$

Table 4 shows the numerical results using this method.

iter	ε	$\|\|R_d\|\|$	$\|\|R_{p1}\|\|_F$	$\|R_{p2}\|$	gap	w	elapsed time
68	1e-02	3.9563e-15	2.1943e-17	2.2204e-16	5.7175e-15	-5.00e-04	19.3
68	1e-04	2.8474e-15	1.6997e-17	0	5.5043e-15	-5.00e-06	19.8
68	1e-08	2.0550e-15	1.9626e-17	0	5.6021e-15	-5.5e-10	20.0
69	0	1.7739e-15	9.8131e-18	0	3.5331e-15	1.7018e-16	19.6
69	-1e-08	2.2337e-15	9.8131e-18	0	3.2534e-15	5.00e-10	19.6
68	-1e-04	3.9490e-15	1.6997e-17	0	5.5256e-15	5.00e-06	19.3
68	-1e-02	8.0304e-16	2.6874e-17	2.2204e-16	5.4419e-15	5.00e-04	19.8

Table 4: Numerical Results for Example 1 with different ε, n=20

Discussion: We note that solving $\mathrm{SDP}^{(\mathbf{P})}$ ($\mathrm{SDP}^{(\mathbf{D})}$) is an efficient way to identify the feasibility status of the original **(P)** (**(P)**), respectively. We observe that w is in close connection with ε. For $\varepsilon \to 0$ we have $w \to 0$. In our tests we discovered that we attain higher accuracy for hard problems if we set $\delta < 0.5$. Such values of δ increase the number of iterations. If δ is chosen close to 1, than convergence is faster, but the algorithm is likely to get stuck on problems with no interior. Further investigations are clearly necessary, to get some theoretical convergence results for this type of problem.

5 Conclusion

In this paper we have begun a study of presolving for SDP. The presolving step is essential in most commercial packages for LP but has not been implemented efficiently for SDP problems as of yet. Our theoretical work has concentrated on the regularization step, i.e. solving problems for which a strictly feasible solution may not exist. (This corresponds to identifying variables which are fixed in LP.) The regularization allows us to solve a nearby problem in a backwards stable fashion.

The numerical tests are very preliminary and mainly concentrated on solving problems for which there was a zero duality gap. This is ongoing work. More details can be found in the forthcoming work by the authors and in particular in the theses of Gerald Gruber and Serge Kruk.

References

[1] F. ALIZADEH. *Combinatorial optimization with interior point methods and semidefinite matrices.* PhD thesis, University of Minnesota, 1991.

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

[3] E. D. ANDERSEN and K. D. ANDERSEN. Presolving in linear programming. *Mathematical Programming*, 71(2):221–245, 1995.

[4] G.P. BARKER and D. CARLSON. Cones of diagonally dominant matrices. *Pacific J. of Math.*, 57:15–32, 1975.

[5] R. BELLMAN. *Introduction to Matrix Analysis.* McGraw Hill, New York, 1957.

[6] R. BELLMAN and K. FAN. On systems of linear inequalities in Hermitian matrix variables. In *Proceedings of Symposia in Pure Mathematics, Vol 7, AMS*, 1963.

[7] A. BEN-ISRAEL, A. BEN-TAL, and S. ZLOBEC. *Optimality in nonlinear programming: a feasible directions approach.* John Wiley & Sons Inc., New York, 1981. A Wiley-Interscience Publication.

[8] A. BEN-ISRAEL, A. CHARNES, and K. KORTANEK. Duality and asymptotic solvability over cones. *Bulletin of American Mathematical Society*, 75(2):318–324, 1969.

[9] F. BOHNENBLUST. Joint positiveness of matrices. Technical report, UCLA, 1948. unpublished manuscript.

[10] J.M. BORWEIN and H. WOLKOWICZ. Characterizations of optimality for the abstract convex program with finite dimensional range. *J. Austra. Math. Soc. Series A*, 30:390–411, 1981.

[11] J.M. BORWEIN and H. WOLKOWICZ. Regularizing the abstract convex program. *Journal of Mathematical Analysis and its Applications*, 83:495–530, 1981.

[12] J.M. BORWEIN and H. WOLKOWICZ. Characterizations of optimality without constraint qualification for the abstract convex program. *Mathematical Programming Study*, 19:77–100, 1982.

[13] B. CRAVEN and J.J. KOLIHA. Generalizations of Farkas' thoerem. *SIAM J. Math. Anal.*, 8:983–997, 1977.

[14] B. CRAVEN and B. MOND. Linear programming with matrix variables. *Linear Algebra and Its Applications*, 38:73–80, 1981.

[15] G. DANTZIG. Linear programming. In *History of Mathematical Programming: A Collection of Personal Reminiscences*. CWI North-Holland, Amsterdam, 1991.

[16] R. J. DUFFIN, R. G. JEROSLOW, and L. A. KARLOVITZ. Duality in semi-infinite linear programming. In *Semi-infinite programming and applications (Austin, Tex., 1981)*, volume 215 of *Lecture Notes in Econom. and Math. Systems*, pages 50–62. Springer, Berlin, 1983.

[17] R.J. DUFFIN. Infinite programs. In A.W. Tucker, editor, *Linear Equalities and Related Systems*, pages 157–170. Princeton University Press, Princeton, NJ, 1956.

[18] K. FAN. On a theorem of weyl concerning eigenvalues of linear transformations ii. *Proc. Nat. Acad. Sci. U.S.A.*, 36:31–35, 1950.

[19] A.V. FIACCO. *Introduction to Sensitivity and Stability Analysis in Nonlinear Programming.* Academic Press, 1983.

[20] A.V. FIACCO and G.P. K.O. KORTANEK (eds.). *Semi-infinite Programming and Applications.* Springer-Verlag, Berlin, 1983.

[21] R. FLETCHER. A nonlinear programming problem in statistics (educational testing). *SIAM J. Sci. Stat. Comput.*, 2:257–267, 1981.

[22] R. FLETCHER. Semi-definite matrix constraints in optimization. *SIAM J. Control and Optimization*, 23:493–513, 1985.

[23] R.M. FREUND. Complexity of an algorithm for finding an approximate solution of a semi-definite program with no regularity assumption. Technical Report OR 302-94, MIT, Cambridge, MA, 1994.

[24] R.M. FREUND and J.R. VERA. Some characterizations and properties of the "distance to ill-posedness" and the condition measure of a conic linear system. Technical report, MIT, Cambridge, MA, 1997.

[25] D. GALE, H.W. KUHN, and A.W. TUCKER. Linear programming and the theory of games. In T.C. Koopmans, editor, *Activity Analysis of Production and allocation.* Wiley, New York, 1951.

[26] K. GLASHOFF and S. GUSTAFSON. *Linear Optimzation and Approximation*, volume 45 of *Applied Mathematical Sciences.* Springer-Verlag, Verlag Basel, 1978.

[27] M. X. GOEMANS. Semidefinite programming in combinatorial optimization. *Mathematical Programming*, 79:143–162, 1997.

[28] M.X. GOEMANS and D.P. WILLIAMSON. .878-approximation algorithms for MAX CUT and MAX 2SAT. In *ACM Symposium on Theory of Computing (STOC)*, 1994.

[29] B. GRONE, C.R. JOHNSON, E. MARQUES DE SA, and H. WOLKOWICZ. Positive definite completions of partial Hermitian matrices. *Linear Algebra and its Applications*, 58:109–124, 1984.

[30] M. GRÖTSCHEL, L. LOVÁSZ, and A. SCHRIJVER. *Geometric Algorithms and Combinatorial Optimization.* Springer Verlag, 1988.

[31] M. GUIGNARD. Generalized Kuhn-Tucker conditions for mathematical programming problems in a banach space. *SIAM J. of Control*, 7:232–241, 1969.

[32] T. D. GUO and S. Q. WU. A modified homogeneous and self-dual linear programming algorithm. *Systems Sci. Math. Sci.*, 8(3):270–277, 1995.

[33] C. HELMBERG, F. RENDL, R. J. VANDERBEI, and H. WOLKOWICZ. An interior point method for semidefinite programming. *SIAM Journal on Optimization*, pages 342–361, 1996. URL: ftp://orion.uwaterloo.ca/pub/henry/reports/sdp.ps.gz.

[34] F. JOHN. Extremum problems with inequalities as subsidiary conditions. In *Sudies and Essays, Courant Anniversary Volume*, pages 187–204. Interscience, New York, 1948.

[35] C.R. JOHNSON. Matrix completion problems: a survey. *Proceedings of Symposium in Applied Mathematics*, 40:171–198, 1990.

[36] E. DE KLERK. *Interior point methods for semidefinite programming*. PhD thesis, Delft University, 1997.

[37] E. DE KLERK, C. ROOS, and T. TERLAKY. Initialization in semidefinite programming via a self-dual skew-symmetric embedding. *Oper. Res. Lett.*, 20(5):213–221, 1997.

[38] K. KORTANEK. Constructing a perfect duality in infinite programming. *Applied Mathematics & Optimization*, 3(4):357–372, 1977.

[39] K.O. KORTANEK and Q. ZHANG. Prefect duality in semi-infinite and semidefinite programming. Technical report, Department of Management Sciences The University of Iowa, 1998.

[40] H.W. KUHN. Nonlinear programming: a historical view. In R.W. Cottle and C.E. Lemke, editors, *Nonlinear Programming*, pages 1–26, Providence, R.I., 1976. AMS.

[41] S. KURCYUSZ. On the existence and nonexistence of lagrange multipliers in banach spaces. *Journal of Optimization Theory and Applications*, 20:81–110, 1976.

[42] M. LAMOREAUX and H. WOLKOWICZ. Numerical decompositions of a convex function. *J. of Optimization Theory and Applications*, 47:51–56, 1985.

[43] M. LAURENT. Cuts, matrix completions and graph rigidity. *Mathematical Programming*, 79:255–284, 1997.

[44] L. LOVÁSZ. On the Shannon capacity of a graph. *IEEE Transactions on Information Theory*, 25:1–7, 1979.

[45] L. LOVÁSZ and A. SCHRIJVER. Cones of matrices and set-functions and 0-1 optimization. *SIAM Journal on Optimization*, 1(2):166–190, 1991.

[46] I. J. LUSTIG, R. E. MARSTEN, and D. F. SHANNO. Interior point methods for linear programming: Computational state of the art. *ORSA Journal on Computing*, 6:1–14, 1994.

[47] O. L. MANGASARIAN and S. FROMOVITZ. The Fritz John necessary optimality conditions in the presence of equality and inequality constraints. *J. Math. Anal. Appl.*, 17:37–47, 1967.

[48] Y. E. NESTEROV and A. S. NEMIROVSKI. Self–concordant functions and polynomial–time methods in convex programming. Book–Preprint, Central Economic and Mathematical Institute, USSR Academy of Science, Moscow, USSR, 1989. Published in Nesterov and Nemirovsky [49].

[49] Y. E. NESTEROV and A. S. NEMIROVSKI. *Interior Point Polynomial Algorithms in Convex Programming*. SIAM Publications. SIAM, Philadelphia, USA, 1994.

[50] B.H. POURCIAU. Modern multiplier methods. *American Mathematical Monthly*, 87(6):433–451, 1980.

[51] F. PUKELSHEIM. *Optimal Design of Experiments*. Wiley Series in Probability and Mathematical Statistics: Probability and Mathematical Statistics. John Wiley & Sons Inc., New York, 1993. A Wiley-Interscience Publication.

[52] M. C. PULLAN. A duality theory for separated continuous linear programs. *SIAM J. Control Optim.*, 34(3):931–965, 1996.

[53] M. RAMANA, L. TUNCEL, and H. WOLKOWICZ. Strong duality for semidefinite programming. *SIAM Journal on Optimization*, 7(3):641–662, 1997.
URL:ftp://orion.uwaterloo.ca/pub/henry/reports/strongdual.ps.gz.

[54] M.V. RAMANA. *An algorithmic analysis of multiquadratic and semidefinite programming problems.* PhD thesis, Johns Hopkins University, Baltimore, Md, 1993.

[55] M.V. RAMANA. An exact duality theory for semidefinite programming and its complexity implications. *Mathematical Programming*, 77:129–162, 1997.

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

[57] S. M. ROBINSON. Bounds for error in the solution set of a perturbed linear program. *Linear Algebra and Appl.*, 6:69–81, 1973.

[58] S. M. ROBINSON. First order conditions for general nonlinear optimization. *SIAM Journal on Applied Mathematics*, 30:597–607, 1976.

[59] S. M. ROBINSON. Regularity and stability for convex multivalued functions. *Mathematics of Operations Research*, 1:130–143, 1976.

[60] S. M. ROBINSON. Stability theorems for systems of inequalities, part ii: differentiable nonlinear systems. *SIAM J. Numerical Analysis*, 13:497–513, 1976.

[61] S. M. ROBINSON. A characterization of stability in linear programming. *Operations Res.*, 25(3):435–447, 1977.

[62] R.T. ROCKAFELLAR. Duality and stability in extremum problems involving convex functions. *Pacific J. Math.*, 21:167–187, 1967.

[63] R.T. ROCKAFELLAR. Ordinary convex programs without a duality gap. *JOTA*, 7:143–148, 1971.

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

[65] A.N. TIKHONOV and V.Y. ARSENIN. *Solutions of Ill-Posed Problems.* V.H. Winston & Sons, John Wiley & Sons, Washington D.C., 1977. Translation editor Fritz John.

[66] L. VANDENBERGHE and S. BOYD. Positive definite programming. In *Mathematical Programming: State of the Art, 1994*, pages 276–308. The University of Michigan, 1994.

[67] L. VANDENBERGHE and S. BOYD. Semidefinite programming. *SIAM Review*, 38:49–95, 1996.

[68] J. von NEUMANN. Discussion of a maximum problem. In A.H. Taub, editor, *Jon von Neumann, Collected Works*, pages 89–95. Pergamon, New York, 1963.

[69] H. WOLKOWICZ. Calculating the cone of directions of constancy. *Journal of Optimization Theory and Applications*, 25:451–457, 1978.

[70] H. WOLKOWICZ. *Constructive approaches to approximate solutions of operator equations and convex programming.* PhD thesis, McGill University, 1978.

[71] H. WOLKOWICZ. Geometry of optimality conditions and constraint qualifications: The convex case. *Mathematical Programming*, 19:32–60, 1980.

[72] H. WOLKOWICZ. Some applications of optimization in matrix theory. *Linear Algebra and its Applications*, 40:101–118, 1981.

[73] H. WOLKOWICZ. The method of reduction in convex programming. *Journal of Optimization Theory and Applications*, 40:349–378, 1983.

[74] H. WOLKOWICZ and Q. ZHAO. Semidefinite relaxations for the graph partitioning problem. *Discrete Applied Math.*, ??, to appear subject to revision.
URL: ftp://orion.uwaterloo.ca/pub/henry/reports/graphpart.ps.gz.

[75] 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.