Chance-Constrained Semidefinite Programming

K. A. Ariyawansa and Yuntao Zhu

February 2006
Chance-Constrained Semidefinite Programming

K. A. Ariyawansa* and Yuntao Zhu†

Abstract

Semidefinite programs are a class of optimization problems that have been the focus of intense research during the past fifteen years. Semidefinite programs extend linear programs, and both are defined using deterministic data. However, uncertainty is naturally present in applications leading to optimization problems. Stochastic linear programs with recourse have been studied since the fifties as a way to deal with uncertainty in data defining linear programs. Recently, the authors have defined an analogous extension of semidefinite programs termed stochastic semidefinite programs with recourse to deal with uncertainty in data defining semidefinite programs. A prominent alternative for handling uncertainty in data defining linear programs is chance-constrained linear programming. In this paper we introduce an analogous extension of semidefinite programs termed chance-constrained semidefinite programs for handling uncertainty in data defining semidefinite programs.

AMS Subject Classifications: 90C15, 90C22, 90C05, 90C51

Keywords: Linear programming, stochastic programming, semidefinite programming, chance-constraints

1 Introduction

Semidefinite programs [1, 16, 18, 14] are a class of optimization problems that have been studied extensively during the past fifteen years. Semidefinite programs extend linear programs, and since both are defined using deterministic data we shall refer to them as deterministic semidefinite programs (DSDP’s)1 and deterministic linear programs (DLP’s) respectively.

Uncertainty is naturally present in applications leading to optimization problems. (Two-stage) stochastic linear programs (with recourse) (SLP’s) [6, 17, 7, 3, 8, 11] have been studied since the fifties as a way to deal with uncertainty in data defining DLP’s. Indeed, the incorporation of uncertainty present in applications into models is so important that stochastic programming is currently one of the most active research subfields of optimization.

*Department of Mathematics, Washington State University, Pullman, WA 99164-3113, USA. (ari@wsu.edu). The work of this author was supported in part by the U.S. Army Research Office under Grant DAAD 19-00-1-0465.
†Department of Mathematics, Washington State University, Pullman, WA 99164-3113, USA. (zhuyt@wsu.edu). The material in this note is part of the doctoral dissertation of this author in preparation at Washington State University.

1As is customary in mathematical programming and optimization literature, we use the term deterministic semidefinite program to mean the generic form of a problem, and the term deterministic semidefinite programming to mean the field of activities based on that problem. Both will be denoted by the acronym DSDP, while the acronym DSDP’s will denote the plural of the first usage. Acronyms DLP, SLP, SSDP, CCLP and CCSDP are defined and used in the same sense.
In a recent paper [2] (see also [9]), the authors have defined a class of optimization problems termed (two-stage) stochastic semidefinite programs (with recourse) (SSDP’s) to deal with uncertainty in data defining DSDP’s. SSDP’s are related to DSDP’s in the same way that SLP’s are related to DLP’s.

Chance-constrained linear programming (CCLP) [4, 5, 10, 11, 12] is a prominent alternative to SLP for handling uncertainty in data defining DLP’s. In this paper, we introduce a class of apparently new optimization problems that we refer to as chance-constrained semidefinite programs (CCSDP’s). CCSDP’s are related to DSDP’s in the same way that CCLP’s are related to DLP’s, and CCSDP is an alternative to SSDP for handling uncertainty in data defining DSDP’s.

The rest of this paper is structured as follows. In §2, we indicate our notation and some preliminaries setting the stage for our definition of a CCSDP given in §3. The main reason that semidefinite programming and stochastic programming are two of the most vibrant research fields of optimization is their applicability. In order to demonstrate the applicability of SSDP, in [2], the authors indicated how an SSDP may be formulated for dealing with uncertainty in data defining the minimum-volume covering ellipsoid problem [16, 13] which is often cited as an example DSDP. In §4, we show how CCSDP as defined in §3 of this paper provides an alternative way to handle uncertainty in data defining the minimum-volume covering ellipsoid problem. We conclude the paper briefly in §5 commenting on how the definitions in this paper and [2] indicate the usefulness of possible joint activities by researchers in semidefinite programming and stochastic programming.

2 Preliminaries

Our notation on DSDP’s follows that of Todd [14]. Let \( \mathbb{R}^{m \times n} \) and \( \mathbb{R}^{n \times n} \) denote the vector spaces of real \( m \times n \) matrices and real symmetric \( n \times n \) matrices respectively. For \( U, V \in \mathbb{R}^{n \times n} \), we write \( U \succeq 0 \) (\( U > 0 \)) to mean that \( U \) is positive semidefinite (positive definite), and \( U \succeq V \) or \( V \preceq U \) to mean that \( U - V \succeq 0 \). For \( U, V \in \mathbb{R}^{m \times n} \), we write \( U \bullet V := \text{trace}(U^T V) \) to denote the Frobenius inner product between \( U \) and \( V \). Given \( U_i \in \mathbb{R}^{n_i \times n_i} \) for \( i = 1, 2, \ldots, n \), we use \( \text{diag}(U_1, U_2, \ldots, U_n) \) to denote the matrix in \( \mathbb{R}^{(\sum_{i=1}^n n_i) \times (\sum_{i=1}^n n_i)} \) with \( U_1, U_2, \ldots, U_n \) on the diagonal and zeros elsewhere.

A DSDP in primal standard form is

\[
\begin{align*}
\text{minimize} & \quad C \bullet X \\
\text{subject to} & \quad A_i \bullet X = b_i, \quad i = 1, 2, \ldots, m \\
& \quad X \succeq 0,
\end{align*}
\]  

where \( A_i \in \mathbb{R}^{n \times n} \) for \( i = 1, 2, \ldots, m \), \( b \in \mathbb{R}^m \) and \( C \in \mathbb{R}^{n \times n} \) constitute given data, and \( X \in \mathbb{R}^{n \times n} \) is the variable. A DSDP in dual standard form is

\[
\begin{align*}
\text{maximize} & \quad b^T y \\
\text{subject to} & \quad \sum_{i=1}^m y_i A_i \preceq C
\end{align*}
\]

where \( A_i \in \mathbb{R}^{n \times n} \) for \( i = 1, 2, \ldots, m \), \( b \in \mathbb{R}^m \) and \( C \in \mathbb{R}^{n \times n} \) constitute given data, and \( y \in \mathbb{R}^m \) is the variable.

It is possible to convert a problem in the form of (2) to an equivalent problem in the form of (1) and vice versa (see [18, 16]). Also, it is appropriate to refer to (1) and (2) as the primal form and the dual form respectively due to the duality theory that exists between them (see [14, §4]).

A DLP in primal standard form

\[
\begin{align*}
\text{minimize} & \quad c^T x \\
\text{subject to} & \quad Ax = b \\
& \quad x \geq 0
\end{align*}
\]
and its dual

\[
\begin{align*}
\text{maximize} & \quad b^T y \\
\text{subject to} & \quad A^T y \leq c
\end{align*}
\]

where \( A \in \mathbb{R}^{m \times n} \), \( b \in \mathbb{R}^m \) and \( c \in \mathbb{R}^n \) constitute given data, \( x \in \mathbb{R}^n \) is the primal variable, and \( y \in \mathbb{R}^m \) is the dual variable are special cases of (1) and (2) respectively. This follows from the associations \( C := \text{diag}(c_1, c_2, \ldots, c_n) \), \( A_i := \text{diag}(A_{i1}, A_{i2}, \ldots, A_{in}) \) for \( i = 1, 2, \ldots, m \), and \( X := \text{diag}(x_1, x_2, \ldots, x_n) \).

Now consider the DLP

\[
\begin{align*}
\text{maximize} & \quad b^T y \\
\text{subject to} & \quad A^T y \leq c \\
& \quad T^T y \leq \xi
\end{align*}
\]

where \( A \in \mathbb{R}^{m \times n_1} \), \( T \in \mathbb{R}^{m \times n_2} \), \( b \in \mathbb{R}^m \), \( c \in \mathbb{R}^{n_1} \), and \( \xi \in \mathbb{R}^{n_2} \) are deterministic data, and \( y \in \mathbb{R}^m \) is the decision variable. Clearly, (4) is a DLP in dual standard form.

A CCLP based on (4) is defined as follows. Suppose that the vector \( \xi \) and the matrix \( T \) are both random depending on an underlying outcome \( \omega \) in an event space \( \Omega \) with a known probability function \( \mathbb{P} \). Let the symbol \( \mathbb{P} \) denote probability. Then it is meaningful to require that the probability of the constraints \( T^T(\omega)x \leq \xi(\omega) \) being satisfied is at least some prescribed value \( p \in (0, 1) \), rather than requiring that they hold for all outcomes \( \omega \in \Omega \). This leads to the problem

\[
\begin{align*}
\text{maximize} & \quad b^T y \\
\text{subject to} & \quad A^T y \leq c \\
& \quad \mathbb{P}(T^T(\omega)y \leq \xi(\omega)) \geq p
\end{align*}
\]

which is termed a CCLP [4, 5, 10, 11, 12].

Constraints in the form of the last constraint of (5) arise naturally in various applications and are called chance-constraints (or probabilistic-constraints). Such constraints can be viewed as a relaxation of the requirement that constraints are enforced for all possible values of uncertain data, which could be prohibitive or even impossible. In practice, \( p \in (0, 1) \) may be chosen close to 1.

3 Definition of a CCSDP

We define a CCSDP based on deterministic data \( A_i \in \mathbb{R}^{n_i \times n_1} \) for \( i = 1, 2, \ldots, m \), \( b \in \mathbb{R}^m \), and \( C \in \mathbb{R}^{n_1 \times n_1} \); and random data \( W_i \in \mathbb{R}^{n_2 \times n_2} \) for \( i = 1, 2, \ldots, m \), and \( D \in \mathbb{R}^{n_2 \times n_2} \) whose realizations depend on an underlying outcome \( \omega \) in an event space \( \Omega \) with a known probability function \( \mathbb{P} \). Given this data, we define a CCSDP as the problem

\[
\begin{align*}
\text{maximize} & \quad b^T y \\
\text{subject to} & \quad \sum_{i=1}^m y_i A_i \preceq C \\
& \quad \mathbb{P}(\sum_{i=1}^m y_i W_i(\omega) \preceq D(\omega)) \geq p
\end{align*}
\]

where \( y \in \mathbb{R}^m \) is the variable, and \( p \in (0, 1) \) is some prescribed value.

4 An application

In this section we describe an application and its formulation as a CCSDP. Our application is a stochastic version of the minimum-volume covering ellipsoid problem (see [16, 13]) often cited as
4.1 Preliminaries

Suppose that we are given ellipsoids $E_i \subset \mathbb{R}^n$, $i = 1, 2$ defined by

$$E_i := \{x \in \mathbb{R}^n : x^T H_i x + 2g_i^T x + v_i \leq 0 \},$$

where $H_i \in \mathbb{R}^{n \times n}$, $H_i > 0$, $g_i \in \mathbb{R}^n$ and $v_i \in \mathbb{R}$ for $i = 1, 2$. Then $E_1$ contains $E_2$ if and only if there is a $\tau \geq 0$ such that matrix inequality

$$\begin{bmatrix} H_1 & g_1 \\ g_1^T & v_1 \end{bmatrix} \preceq \tau \begin{bmatrix} H_2 & g_2 \\ g_2^T & v_2 \end{bmatrix}$$

holds [15]. Now consider the ball $B \subset \mathbb{R}^n$ represented by

$$B := \{x \in \mathbb{R}^n : x^T x - 2\bar{x}^T x + \gamma \leq 0 \}.$$

The center of $B$ is $\bar{x}$ and its radius is $\sqrt{\bar{x}^T \bar{x} - \gamma}$. The distance from the origin to the center of the ball is $\sqrt{x^T x}$.

The ball $B$ contains the ellipsoids $E_1$ and $E_2$ if and only if there exist $\tau_1 \geq 0$ and $\tau_2 \geq 0$ such that

$$\begin{bmatrix} I & -\bar{x} \\ -\bar{x}^T & \gamma \end{bmatrix} \preceq \tau_i \begin{bmatrix} H_i & g_i \\ g_i^T & v_i \end{bmatrix}, \quad i = 1, 2.$$

4.2 The Application

We describe the application in generic terms first. Suppose that we are given $n_f$ fixed ellipsoids $E_i := \{x \in \mathbb{R}^n : x^T H_i x + 2g_i^T x + v_i \leq 0 \} \subset \mathbb{R}^n$, $i = 1, 2, \ldots, n_f$. Here $H_i \in \mathbb{R}^{n \times n}$, $H_i > 0$, $g_i \in \mathbb{R}^n$ and $v_i \in \mathbb{R}$ for $i = 1, 2, \ldots, n_f$ are deterministic data. We are also given $n_r$ random ellipsoids $E_i(\omega) := \{x \in \mathbb{R}^n : x^T H_i(\omega) x + 2g_i(\omega)^T x + \nu_i(\omega) \leq 0 \}, \quad i = 1, 2, \ldots, n_r$. Here for $i = 1, 2, \ldots, n_r$, $H_i(\omega) \in \mathbb{R}^{n \times n}$, $H_i(\omega) > 0$, $g_i(\omega) \in \mathbb{R}^n$, $\nu_i(\omega) \in \mathbb{R}$ are random data whose realizations depend on an underlying outcome $\omega$ in an event space $\Omega$ with a known probability function $P$.

Suppose that we need to determine a ball subject to two types of constraints: the ball must contain all $n_f$ fixed ellipsoids; and it must contain the $n_r$ random ellipsoids with probability at least $p \in (0, 1)$. We assume that the cost of choosing the ball has two components: the cost of the center is proportional to the Euclidean distance to the center from the origin, and the cost of the radius is proportional to the square of the radius. The second type of constraint mentioned above can be viewed as a relaxation of the requirement that the ball contains all the realizations of the $n_r$ random ellipsoids, which could be prohibitive or even impossible. The center and the radius are to be determined so that the total cost is minimized.

Before proceeding to formulate a model for this generic application, we indicate a more concrete version of it. Let $n := 2$. The fixed ellipsoids contain targets that need to be destroyed, and the random ellipsoids contain targets that also need to be destroyed but are moving. Fighter aircrafts take off from the origin to destroy both types of targets with a planned disk of coverage having the following properties: the disk contains all the fixed ellipsoids; and it contains the realizations of the random ellipsoids with probability at least $p \in (0, 1)$. Our model determines the center and the radius of the disk of coverage so that the total cost is minimized.
4.3 Formulation of the Model

Our goal is to determine $\bar{x} \in \mathbb{R}^n$ and $\gamma \in \mathbb{R}$ such that the ball $\mathcal{B}$ defined by

$$\mathcal{B} := \{x \in \mathbb{R}^n : x^T x - 2\bar{x}^T x + \gamma \leq 0\}$$

contains the fixed ellipsoids $\mathcal{E}_i$ for $i = 1, 2, \ldots, n_f$, and the realizations of the random ellipsoids $\tilde{\mathcal{E}}_i(\omega)$ for $i = 1, 2, \ldots, n_r$ with probability at least $p \in (0, 1)$. Note that the center of $\mathcal{B}$ is $\bar{x}$ and that the square of the radius of $\mathcal{B}$ is $\bar{x}^T \bar{x}$. We introduce the two constraints

$$\begin{bmatrix} d_1 I & \bar{x} \\ \bar{x}^T & d_1 \end{bmatrix} \succeq 0 \quad (7)$$

and

$$\begin{bmatrix} I & \bar{x} \\ \bar{x}^T & d_2 + \gamma \end{bmatrix} \succeq 0. \quad (8)$$

By Schur Complements (7) is equivalent to $d_1 - \bar{x}^T (d_1 I)^{-1} \bar{x} \geq 0$ which in turn is equivalent to $d_1 \geq \sqrt{\bar{x}^T \bar{x}}$. Constraint (8) is similarly equivalent to $d_2 + \gamma - \bar{x}^T I^{-1} \bar{x} \geq 0$ and to the constraint $d_2 \geq \bar{x}^T \bar{x} - \gamma$. Thus $d_1$ is an upper bound on the distance between the center of the ball $\mathcal{B}$ and the origin, $\sqrt{\bar{x}^T \bar{x}}$. Similarly, $d_2$ is an upper bound on the square of the radius of the ball $\mathcal{B}$, $\bar{x}^T \bar{x} - \gamma$.

Let $\bar{c} \geq 0$ denote the cost per unit of the Euclidean distance between the center of the ball $\mathcal{B}$ and the origin, and $\alpha \geq 0$ be the cost per unit of the square of the radius of $\mathcal{B}$.

We define the decision variable $x \in \mathbb{R}^{(n_f + n_r + n + 3)}$ as

$$x := \begin{bmatrix} d_1, d_2, \bar{x}^T, \gamma, \tau^T, \delta^T \end{bmatrix}^T,$$

where except for the auxiliary variables $\tau \in \mathbb{R}^{n_f}$ and $\delta \in \mathbb{R}^{n_r}$, other variables are as specified above. We also introduce the unit cost vector $c \in \mathbb{R}^{(n_f + n_r + n + 3)}$ as

$$c := \begin{bmatrix} \bar{c}, \alpha, 0^T, 0, 0^T \end{bmatrix}^T.$$

Then we get the model

$$\begin{array}{ll}
\text{minimize} & c^T x \\
\text{subject to} & \begin{bmatrix} I & -\bar{x}^T \\ -\bar{x} & \gamma \end{bmatrix} \succeq \tau_i \begin{bmatrix} H_i & g_i \\ g_i^T & v_i \end{bmatrix}, \quad i = 1, 2, \ldots, n_f, \\
& 0 \leq \tau, \\
& 0 \leq \begin{bmatrix} d_1 I & \bar{x} \\ \bar{x}^T & d_1 \end{bmatrix}, \quad (9) \\
& 0 \leq \begin{bmatrix} I & \bar{x} \\ \bar{x}^T & d_2 + \gamma \end{bmatrix}, \\
& \mathbb{P}\left(\begin{bmatrix} I & -\bar{x}^T \\ -\bar{x} & \gamma \end{bmatrix} \succeq \delta_i \begin{bmatrix} \tilde{H}_i(\omega) & \tilde{g}_i(\omega) \\ \tilde{g}_i^T(\omega) & \tilde{v}_i(\omega) \end{bmatrix}, \quad i = 1, 2, \ldots, n_r\right) \geq p, \\
& 0 \leq \delta.
\end{array}$$
Problem (9) is a CCSDP as defined in §3. We now indicate the assignments that need to be made in (6) to obtain (9). First let \( n_1 := n_f + n_r + 3(n+1) \), \( m := n_f + n_r + n + 3 \), and \( n_2 := n + 1 \). Then let \( y := x \) and \( b := -c \). We use \( \bar{0}_n \) to denote the zero vector in \( \mathbb{R}^n \), and \( \bar{0}_n \) to denote the zero matrix in \( \mathbb{R}^{n \times n} \). Let \( e_i \) be the \( i \)-th elementary vector in \( \mathbb{R}^n \), and \( I_n \) be the identity matrix in \( \mathbb{R}^{n \times n} \). Then for \( i = 1, 2, \ldots, (n_f + n_r + n + 3) \), we define matrix \( A_i \in \mathbb{R}^{(n_f + n_r + 3(n+1)) \times (n_f + n_r + 3(n+1))} \) as follows:

\[
A_1 := \text{diag}(\bar{0}_{n+1}, \bar{0}_{n_f}, -I_{n+1}, \bar{0}_{n+1}, \bar{0}_{n_r});
\]

\[
A_2 := \text{diag}(\bar{0}_{n+1}, \bar{0}_{n_f}, \bar{0}_{(n+1)}, -\left[ \begin{array}{cc} \bar{0}_n & 0 \\ 0 & 1 \end{array} \right], \bar{0}_{n_r});
\]

for \( i = 3, 4, \ldots, (n+2) \)

\[
A_i := \text{diag}\left( \begin{bmatrix} \bar{0}_n & -e_{(i-2)} \\ -e_{(i-2)}^T & 0 \end{bmatrix} \right), \bar{0}_{n_f}, \begin{bmatrix} \bar{0}_n & -e_{(i-2)} \\ -e_{(i-2)}^T & 0 \end{bmatrix}, \bar{0}_{n_r};
\]

\[
A_{(n+3)} := \text{diag}\left( \begin{bmatrix} \bar{0}_n & 0 \\ 0 & 1 \end{bmatrix}, \bar{0}_{n_f}, \bar{0}_{(n+1)}, \begin{bmatrix} \bar{0}_n & 0 \\ 0 & -1 \end{bmatrix}, \bar{0}_{n_r};
\right)
\]

for \( i = (n+4), (n+5), \ldots, (n_f + n_r + 3) \)

\[
A_i := \text{diag}\left( -\left[ \begin{bmatrix} H_{(i-n-3)} \\ g_{(i-n-3)} \\ v_{(i-n-3)} \end{bmatrix} \right], \text{diag}(-e_{(i-n-3)}, \bar{0}_{(n+1)}, \bar{0}_{(n+1)}, \bar{0}_{n_r});
\]

and for \( i = (n_f + n + 4), (n_f + n + 5), \ldots, (n_f + n_r + n + 3) \)

\[
A_i := \text{diag}(\bar{0}_{(n+1)}, \bar{0}_{(n_f)}, \bar{0}_{(n+1)}, \bar{0}_{(n+1)}, \text{diag}(-e_{(i-n_f-n-3)})).
\]

Next we define \( C \in \mathbb{R}^{(n_f + n_r + 3(n+1)) \times (n_f + n_r + 3(n+1))} \) as

\[
C := \text{diag}\left( \begin{bmatrix} -I_n & 0 \\ 0 & 0 \end{bmatrix}, \bar{0}_{n_f}, \bar{0}_{(n+1)}, \begin{bmatrix} I_n & 0 \\ 0 & 0 \end{bmatrix}, \bar{0}_{n_r};
\right).
\]

Then for \( i = 1, 2, \ldots, (n_f + n_r + n + 3) \), we define matrix \( W_i(\omega) \in \mathbb{R}^{(n+1) \times (n+1)} \) as follows:

\[
W_1(\omega) := \bar{0}_{(n+1)};
\]

\[
W_2(\omega) := \bar{0}_{(n+1)};
\]

for \( i = 3, 4, \ldots, (n+2) \)

\[
W_i(\omega) := \begin{bmatrix} \bar{0}_n & -e_{(i-2)} \\ -e_{(i-2)}^T & 0 \end{bmatrix};
\]

\[
W_{(n+3)}(\omega) := \begin{bmatrix} \bar{0}_n & 0 \\ 0 & 1 \end{bmatrix};
\]

for \( i = (n+4), (n+5), \ldots, (n_f + n + 3) \)

\[
W_i(\omega) = \bar{0}_{(n+1)};
\]

and for \( i = (n_f + n + 4), (n_f + n + 5), \ldots, (n_f + n_r + n + 3) \)

\[
W_i(\omega) := -\begin{bmatrix} \bar{H}_{(i-n_f-n-3)}(\omega) & \bar{g}_{(i-n_f-n-3)}(\omega) \\ \bar{g}_{(i-n_f-n-3)}^T(\omega) & \bar{v}_{(i-n_f-n-3)}(\omega) \end{bmatrix}.
\]
Finally, we define $D(\omega) \in \mathbb{R}^{(n+1)\times(n+1)}$ as

$$D(\omega) := \begin{bmatrix}
-I_n & 0_n \\
0_n & 0
\end{bmatrix}.$$ 

With these assignments in (6) we get (9), and so (9) is a CCSDP as defined in §3.

5 Concluding Remarks

In this paper, we have defined an apparently new paradigm for stochastic optimization that we term chance-constrained semidefinite programming (CCSDP). CCSDP is an alternative to (two-stage) stochastic semidefinite programming (with recourse) (SSDP) defined in [2] for handling uncertainty in data defining semidefinite programs. Our developments of CCSDP in this paper, and of SSDP in [2] are such that they parallel the corresponding developments leading to chance-constrained linear programs (CCLP’s) and (two-stage) stochastic linear programs (with recourse) (SLP’s) as ways of handling uncertainty in data defining linear programs. As a consequence, research problems on applications, algorithms and theory pertinent to CCSDP and SSDP, and possible ways for their solution essentially suggest themselves. For example, it readily becomes apparent that an SSDP and a CCSDP can be combined into a single model similar to the way an SLP and a CCLP is combined in [17].

Semidefinite programming and stochastic programming are two of the most vibrant research subfields of optimization. This paper and [2] suggest exciting new possibilities for joint work by researchers in the two fields.

References


