Chapter 2

Theory and algorithms

In this chapter the ideas behind rigorous verification will be presented. For a start, we give a detailed introduction into semidefinite programming. Notations and techniques of interval arithmetic are provided later. This allows for a more general approach to the rigorous verification (uncertainties in the input data can thus be involved in the computation). The actual theorems and algorithms for rigorously bounding the optimal value are presented in the subsequent sections. Finally, an important case of infeasible problems is examined in detail.

2.1 Semidefinite programming

Let us first define a *semidefinite program* in its primal form

$$\begin{aligned} p^* := \min\langle C, X\rangle \quad \text{s.t.} \quad \langle A_i, X\rangle = b_i \quad \text{for } i = 1, \ldots, m, \\ X \succeq 0, \end{aligned} \tag{2.1}$$

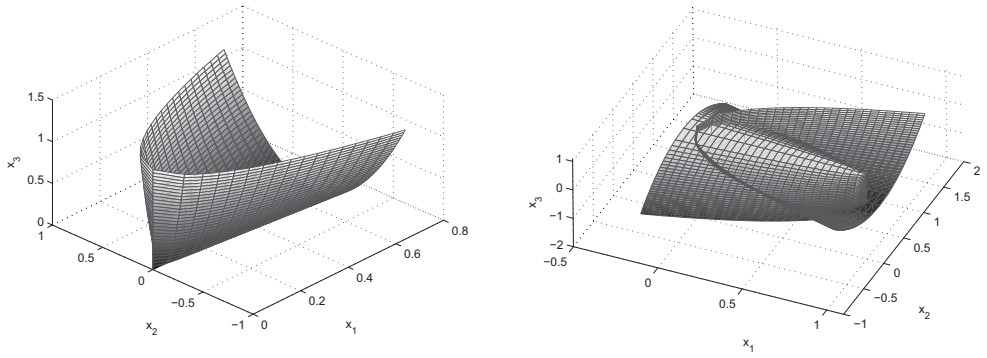
where $C \in \mathbb{S}^s$, $A_i \in \mathbb{S}^s$ and $b \in \mathbb{R}^m$ are given problem parameters, and $X \in \mathbb{S}^s$ is the optimization variable. Here, $\mathbb{S}^s$ denotes the space of real symmetric matrices of order s.

$$\langle C, X\rangle = \text{ trace } (C^T X) \tag{2.2}$$

in its turn denotes the *inner product* over $\mathbb{S}^s$. Moreover, $\succeq$ is the *Löwner partial order*, that is $X \succeq Y$ iff $X - Y$ is positive semidefinite.

The *Lagrangian dual* of (2.1) is

$$d^* := \max b^T y \quad \text{s.t.} \quad \sum_{i=1}^{m} y_i A_i \preceq C, \tag{2.3}$$



(a) (b)

Figure 2.1: Semidefinite cones for $s = 2$: $X \succeq 0$ (left) and $0 \preceq X \preceq I$ (right).

where $y \in \mathbb{R}^m$. The constraints $\sum_{i=1}^m y_i A_i \preceq C$ are called *linear matrix inequalities (LMI)*. We use the convention that $p^* = -\infty$ if (2.1) is unbounded and $p^* = \infty$ if (2.1) is infeasible. The analogous convention is used for (2.3).

Since we have formulated semidefinite programming in its standard form, it is easy to see, that the optimization domain is the intersection of the cone of positive semidefinite matrices with an affine space. The objective function is linear. The introduced problem can thus be seen as a subclass of cone programming and also as a generalization of linear programming. Indeed, if we demand all symmetric matrices to be diagonal, (2.1) will define a standard linear programming problem.

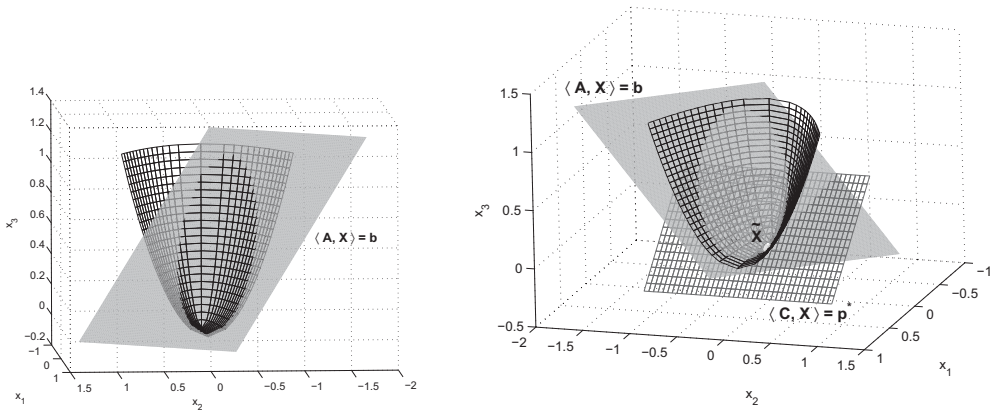
Example 2.1. *To get basic ideas about the geometry of the problem, let us consider the simplest case of* $s = 2$ *and a single equality constraint (*$m = 1$*) in Figures 2.1 and 2.2. The positive semidefiniteness condition*

$$X = \begin{pmatrix} x_1 & x_2 \\ x_2 & x_3 \end{pmatrix} \succeq 0 \tag{2.4}$$

is fulfilled for all the points in the interior and on the boundary of the cone in Figure 2.1a. The intersection of two semidefinite cones, shown in Figure 2.1b, is also a typical configuration. We have this situation, for example, in Chapter 3, where the condition on one-particle reduced density matrix is exactly $0 \preceq$ *1RDM* $\preceq I$*.*

For the semidefinite optimization example in Figure 2.2 we use the following data:

$$C = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}, \quad A = \begin{pmatrix} 2 & 0.5 \\ 0.5 & 1 \end{pmatrix} \quad \textit{and} \quad b = 1. \tag{2.5}$$



(a) (b)

Figure 2.2: Intersection of the semidefinite cone from Figure 2.1a with the linear constraint plane (left) and the objective function plane passing through the optimal solution $\tilde{X}$ (right).

The solution of the obtained semidefinite optimization problem is

$$\tilde{X} = \begin{pmatrix} 0.3867 & 0.1602 \\ 0.1602 & 0.0664 \end{pmatrix} \tag{2.6}$$

and the objective value $p^* = \langle C, \tilde{X} \rangle = 0.4531$. *When trying to graphically solve the problem, the idea is again similar to that in LP. We shift the plane* $\langle C, X \rangle = const$ *downwards until we reach the edge of the feasibility region. In our example we have to stop at* $const = 0.4531$. $\tilde{X}$ *is then the only intersection point of the cone* $X \succeq 0$, *the affine constraint* $\langle A, X \rangle = b$ *and the objective function plane* $\langle C, X \rangle = 0.4531$.

As a standard solution algorithm, interior point methods proved to be an approach of choice. They have their roots in Karmarkar's work [40], where he introduced an algorithm to solve an LP with polynomial iteration complexity[1]. As applied to SDPs, the idea found its development in the works of Nesterov and Nemirovski (see for example [51] and [52]) and Alizadeh [2]. The concept, similarly to any other barrier function method, would be to substitute the initial problem with an optimization problem (more precisely a sequence of them) without the semidefiniteness constraint

$$p^* := \min \langle C, X \rangle + \mu\phi(X) \quad \text{s.t.} \quad \langle A_i, X \rangle = b_i \quad \text{for } i = 1, \ldots, m, \tag{2.7}$$

[1] However the development of the idea can be tracked back to the works of Frisch [16] on logarithmic barrier functions and Huard [30] on the method of centers.

where $\phi(X)$ is the *barrier function* and $\mu > 0$ is the *barrier parameter*. Standard barrier function would be a logarithmic function of the type

$$\begin{aligned} &\phi(X) = -\ln\det X = \ln(\det X)^{-1} \quad \text{if} \quad X \succ 0, \\ &\phi(X) = +\infty \quad \text{otherwise.} \end{aligned} \tag{2.8}$$

Later on by sequentially decreasing μ towards 0, we solve the initial problem.

The duality theory of semidefinite programming is a bit more subtle compared to linear programming. The programs satisfy the *weak duality* condition

$$d^* \leq p^*, \tag{2.9}$$

but strong duality requires in contrast to linear programming additional conditions (see for example Nemirovski [50], Ramana, Tunçel, and Wolkowicz [63] and Vandenberghe and Boyd [74]).

Theorem 2.1 (Strong Duality Theorem)**.**

a) If (2.1) is strictly feasible (i.e. there exists a feasible positive definite matrix X) and p^* is finite, then $p^* = d^*$ and the dual supremum is attained.

b) If (2.3) is strictly feasible (i.e. there exists some $y \in \mathbb{R}^m$ such that $C - \sum_{i=1}^m y_i A_i$ is positive definite) and d^* is finite, then $p^* = d^*$, and the primal infimum is attained.

In general, one of the problems (2.1) and (2.3) may have optimal solutions while its dual is infeasible, or the duality gap may be positive at optimality. The strict feasibility assumptions in Theorem 2.1 are called *Slater's constraint qualifications*.

As a matter of convenience, in the rest of the thesis semidefinite programs will be considered in more general block diagonal form reflecting the sparsity of the problem. The primal problem (2.1) becomes then

$$\begin{aligned} p^* := \min \sum_{j=1}^n \langle C_j, X_j \rangle \quad \text{s.t.} \quad & \sum_{j=1}^n \langle A_{ij}, X_j \rangle = b_i \quad \text{for } i = 1, \ldots, m, \\ & X_j \succeq 0 \quad \text{for } j = 1, \ldots, n, \end{aligned} \tag{2.10}$$

where $C_j \in \mathbb{S}^{s_j}$, $A_{ij} \in \mathbb{S}^{s_j}$ and $X_j \in \mathbb{S}^{s_j}$. Finally, instead of the dual problem (2.3) we now have

$$d^* := \max b^T y \quad \text{s.t.} \quad \sum_{i=1}^m y_i A_{ij} \preceq C_j \quad \text{for } j = 1, \ldots, n. \tag{2.11}$$

2.2 Notation

Throughout this thesis we use the following notation. $\mathbb{R}$, $\mathbb{R}^n$, $\mathbb{R}^n_+$, and $\mathbb{R}^{m\times n}$ denote the sets of real numbers, real vectors, real nonnegative vectors, and real $m \times n$ matrices, respectively. $\mathbb{S}^n$, in its turn, stands for the set of real symmetric matrices. Comparisons $\leq$, absolute value $|\cdot|$, $\min$, $\max$, $\inf$ and $\sup$ are used entrywise for vectors and matrices. The identity matrix is denoted by I.

For a symmetric matrix A the eigenvalues are sorted non-increasingly, $\lambda_{\max}(A) = \lambda_1(A) \geq \lambda_2(A) \geq \ldots \geq \lambda_{\min}(A)$.

For $\mu \in \mathbb{R}$ the operator

$$\text{svec}(A, \mu) := (A_{11}, \mu A_{21}, \ldots, \mu A_{n1}, A_{22}, \mu A_{32}, \ldots, \mu A_{n\,n-1}, A_{nn})^T, \qquad (2.12)$$

transforms symmetric $n \times n$ matrices into $((n+1)n/2)$-dimensional vectors with the property that the inner product of two symmetric matrices A, B is

$$\langle A, B\rangle = \text{svec}(A, 2)^T \text{svec}(B, 1) = \text{svec}(A, \sqrt{2})^T \text{svec}(B, \sqrt{2}), \qquad (2.13)$$

and $\text{svec}(A, \sqrt{2})$ is the customary svec operator. We prefer the first representation of the inner product, since this avoids conversion errors of the input data of semidefinite programs in its vector representation form. The inverse operator of svec is denoted by $\text{smat}(a, \mu)$, where a is the vector representation (2.12).

For block matrices with blocks A_j for $j = 1, \ldots, n$ we define the concatenated vector

$$\text{svec}((A_j), \mu) := (\text{svec}(A_1, \mu); \ldots; \text{svec}(A_n, \mu)). \qquad (2.14)$$

A block diagonal matrix with blocks $B_1, \ldots, B_n$ will be written as

$$\text{Diag}(B_1, \ldots, B_n). \qquad (2.15)$$

Other necessary notation concerning, for example, interval arithmetic, will be introduced in the corresponding sections.

2.3 Interval arithmetic

Rigorous verification requires to consider rounding errors of the floating point arithmetic. One needs tools to control machine rounding and to estimate error propagation. Interval arithmetic provides us with such tools. Besides that, in real life applications

many values or model parameters are measurement results. Since no devices possess infinite precision, such values have to be considered with measurement errors. To cope with this, we allow interval input in all problems discussed in the thesis.

We require only some elementary facts about interval calculations, which are described here. There are a number of textbooks on interval arithmetic and self-validating methods that can be highly recommended to readers. These include Alefeld and Herzberger [1], Moore [49], and Neumaier [54], [55].

If $\mathbb{V}$ is one of the spaces $\mathbb{R}$, $\mathbb{R}^n$, $\mathbb{R}^{m\times n}$, and $\underline{v}, \overline{v} \in \mathbb{V}$, then the box

$$\mathbf{v} := [\underline{v}, \overline{v}] := \{v \in \mathbb{V} : \ \underline{v} \leq v \leq \overline{v}\} \tag{2.16}$$

is called an *interval quantity* in $\mathbb{IV}$ with *lower bound* $\underline{v}$ and *upper bound* $\overline{v}$. In particular, $\mathbb{IR}$, $\mathbb{IR}^n$, and $\mathbb{IR}^{m\times n}$ denote the set of real intervals $\mathbf{a} = [\underline{a}, \overline{a}]$, the set of real interval vectors $\mathbf{x} = [\underline{x}, \overline{x}]$, and the set of real interval matrices $\mathbf{A} = [\underline{A}, \overline{A}]$, respectively. The real operations $A \circ B$ with $\circ \in \{+, -, \cdot, /\}$ between real numbers, real vectors and real matrices can be generalized to *interval operations*. The result $\mathbf{A} \circ \mathbf{B}$ of an interval operation is defined as the interval hull of all possible real results, that is

$$\mathbf{A} \circ \mathbf{B} := \cap\{\mathbf{C} \in \mathbb{IV} : \ A \circ B \in \mathbf{C} \quad \text{for all} \quad A \in \mathbf{A}, B \in \mathbf{B}\}. \tag{2.17}$$

All interval operations can be easily executed by working appropriately with the lower and upper bounds of the interval quantities. In the simple cases of addition and subtraction, we obtain

$$\begin{array}{rcl} \mathbf{A} + \mathbf{B} & = & [\underline{A} + \underline{B}, \overline{A} + \overline{B}], \\ \mathbf{A} - \mathbf{B} & = & [\underline{A} - \overline{B}, \overline{A} - \underline{B}]. \end{array} \tag{2.18}$$

Interval multiplications and divisions require a distinction of cases. Let $\mathbf{a} = [\underline{a}, \overline{a}] \in \mathbb{IR}$ and $\mathbf{b} = [\underline{b}, \overline{b}] \in \mathbb{IR}$, then

$$\begin{array}{rcl} \mathbf{a} \cdot \mathbf{b} & := & [\min\{\underline{a}\underline{b}, \underline{a}\overline{b}, \overline{a}\underline{b}, \overline{a}\overline{b}\}, \max\{\underline{a}\underline{b}, \underline{a}\overline{b}, \overline{a}\underline{b}, \overline{a}\overline{b}\}], \\ \mathbf{a}/\mathbf{b} & := & [\underline{a}, \overline{a}] \cdot \left[\frac{1}{\overline{b}}, \frac{1}{\underline{b}}\right], \quad \text{assuming} \quad 0 \notin \mathbf{b}. \end{array} \tag{2.19}$$

The rules of commutativity and associativity remain valid also for operations on $\mathbb{IR}$. The sub-distributivity rule

$$\mathbf{a}(\mathbf{b} + \mathbf{c}) \subseteq \mathbf{a}\mathbf{b} + \mathbf{a}\mathbf{c} \tag{2.20}$$

with $\mathbf{c} \in \mathbb{IR}$ substitutes the distributivity from $\mathbb{R}$. $\mathbf{x} = [0, 0]$ and $\mathbf{y} = [1, 1]$ are the unique neutral elements with respect to addition and multiplication. A fundamental

property of interval arithmetic is inclusion monotonicity:

$$\mathbf{a} \subseteq \mathbf{a}',\ \mathbf{b} \subseteq \mathbf{b}' \quad \Rightarrow \quad \mathbf{a} \circ \mathbf{b} \subseteq \mathbf{a}' \circ \mathbf{b}', \quad \circ \in \{+, -, \cdot, /\}. \tag{2.21}$$

This property follows directly from the set-theoretical definitions of the interval arithmetic operations (2.17). Thus rational interval functions are inclusion monotonic, as are natural interval extensions of all the standard functions used in computing. With proper rounding procedures, rounded interval arithmetic operations are also inclusion monotonic (Moore [49]).

Similarly all operations (2.17) between interval vectors and interval matrices can be executed by replacing every real operation by the corresponding interval operation. For example the i, j component of the product of two interval matrices $\mathbf{C}, \mathbf{X} \in \mathbb{IR}^{n \times n}$ is

$$(\mathbf{CX})_{ij} := \sum_{k=1}^{n} \mathbf{C}_{ik}\mathbf{X}_{kj}, \tag{2.22}$$

and the inner product

$$\langle \mathbf{C}, \mathbf{X} \rangle = \text{ trace } (\mathbf{C}^T\mathbf{X}) = \sum_{i,j=1}^{n} \mathbf{C}_{ij}\mathbf{X}_{ij}. \tag{2.23}$$

For interval quantities $\mathbf{A}, \mathbf{B} \in \mathbb{IV}$ we define

$$\begin{aligned}
\text{mid}\mathbf{A} &:= (\underline{A} + \overline{A})/2 \quad \text{as the } \textit{midpoint}, && (2.24)\\
\text{rad}\mathbf{A} &:= (\overline{A} - \underline{A})/2 \quad \text{as the } \textit{radius,} && (2.25)\\
|\mathbf{A}| &:= \sup\{|A| :\ A \in \mathbf{A}\} \quad \text{as the } \textit{absolute value}, && (2.26)\\
\mathbf{A}^+ &:= \max\{0, \overline{A}\}, && (2.27)\\
\mathbf{A}^- &:= \min\{0, \underline{A}\}. && (2.28)
\end{aligned}$$

Moreover, the comparison in $\mathbb{IV}$ is defined by

$$\mathbf{A} \leq \mathbf{B} \quad \text{iff} \quad \overline{A} \leq \underline{B},$$

and other relations are defined analogously. Real quantities v are embedded in the interval quantities by identifying $v = \mathbf{v} = [v, v]$.

We call $\mathbf{A} \in \mathbb{IR}^{n \times n}$ *symmetric*, if $\mathbf{A}_{ij} = \mathbf{A}_{ji}$ for all i, j, and $\mathbf{A}$ is called positive semidefinite if all $A \in \mathbf{A}$ have this property.