Higher-order quantum transformations: From quantum circuit design to indefinite causal order

Lecture III. Semidefinite programming

Jessica Bavaresco

Contents

1 Introduction

Semidefinite programming is a class of convex optimization problems [1] widely applied in quantum information theory [2].

2 Semidefinite programming

An optimization problem is a semidefinite program (SDP) if it is of the following form.
Let

(i) $A \in \mathcal{L}(\mathbb{C}^d)$, $A = A^\dagger$ (Hermitian);

(ii) $B \in \mathcal{L}(\mathbb{C}^{d'})$, $B = B^\dagger$ (Hermitian); and

(iii) $\Lambda : \mathcal{L}(\mathbb{C}^d) \to \mathcal{L}(\mathbb{C}^{d'})$, such that $\forall x = x^\dagger$, $\Lambda(x) = \Lambda(x)^\dagger$ (Hermitian preserving).

Then

$$
\begin{aligned}
\text{given} \quad & A, B, \Lambda \\
\min_{x \in \mathcal{L}(\mathbb{C}^d)} \quad & \operatorname{tr}(Ax) && \text{(linear objective function)} \\
\text{subject to} \quad & \Lambda(x) = B && \text{(affine constraint)} \\
& x \geq 0 && \text{(positive semidefinite cone)}
\end{aligned}
$$

is an SDP.

This is usually referred to as the **canonical form**.

Linearity and the Riesz representation lemma

The objective function must be a linear, real-valued function. A function $f : \mathcal{L}(\mathbb{C}^d) \to \mathbb{R}$ is linear if for all $x, y \in \mathcal{L}(\mathbb{C}^d)$ and $\lambda, \lambda' \in \mathbb{R}$,

$$f(\lambda x + \lambda' y) = \lambda f(x) + \lambda' f(y). \tag{1}$$

All linear functions f can be written as an inner product. This is known as the Riesz representation lemma. For any linear function $f : \mathcal{L}(\mathbb{C}^d) \to \mathbb{R}$, there exists an operator $A \in \mathcal{L}(\mathbb{C}^d)$ such that

$$f(x) = \langle A, x\rangle = \mathrm{tr}(A^\dagger x). \tag{2}$$

Positive semidefinite cone

The optimization variable x of an SDP must obey affine constraints and be in the cone of positive semidefinite operators. The set of positive semidefinite operators $P = \{x \in \mathcal{L}(\mathcal{H}) \mid x \geq 0\}$ is a cone since

$$x \geq 0 \implies \lambda x \geq 0 \quad \forall\, \lambda \geq 0. \tag{3}$$

Feasible set

The feasible set of an SDP is the subset of operators $x \in \mathcal{L}(\mathbb{C}^d)$ that lies at the intersection of the positive semidefinite cone and the hyperplane defined by the affine constraint specified in the SDP:

$$F_P \coloneqq \{x \in \mathcal{L}(\mathbb{C}^d) \,|\, x \geq 0, \Lambda(x) = B\}. \tag{4}$$

Solution

The optimal values of x in the feasible set will be denoted x^*; that is,

$$x^* \coloneqq \arg \min_{x \in F_P} f(x). \tag{5}$$

This value is generally not unique. However, the optimal value for the objective function, denoted $f(x^*)$, is **always unique for an SDP**.

2.1 Equivalent forms

An SDP can be written in many different and equivalent forms, which may include several positivity and affine constraints. Let

(i) $f : \mathcal{L}(\mathbb{C}^d) \to \mathbb{R}$ be a linear function;

(ii) $\{B_i\}_i, \{C_j\}_j$, with $B_i, C_j \in \mathcal{L}(\mathbb{C}^{d'})$ be Hermitian;

(iii) $\{g_i\}_i, \{h_j\}_j$, with $g_i, h_j : \mathcal{L}(\mathbb{C}^d) \to \mathcal{L}(\mathbb{C}^{d'})$ be Hermitian preserving; and

(iv) $n, m \in \mathbb{N}$ be natural numbers.

Then

$$\begin{aligned} \text{given} \quad & f, \{g_i\}_i, \{h_j\}_j, \{B_i\}_i, \{C_j\}_j \\ \min \quad & f(x) \\ \text{s.t.} \quad & g_i(x) = B_i \quad \forall i \in \{1, \dots, n\} \\ & h_j(x) \geq C_j \quad \forall j \in \{1, \dots, m\} \end{aligned}$$

is an SDP.

Every SDP written in such a form can be equivalently expressed in the canonical form.

2.2 Feasibility vs. optimization problems

A feasibility problem is to determine whether a feasible set is empty, that is, whether there exists a non-empty set of values for the optimization variable that simultaneously satisfies all constraints:

A feasibility SDP can be expressed as:

$$\begin{aligned}
\text{given} \quad & B, \Lambda \\
\text{find} \quad & x \in \mathcal{L}(\mathbb{C}^d) \\
\text{subject to} \quad & \Lambda(x) = B \\
& x \geq 0.
\end{aligned}$$

This is equivalent to an optimization SDP where $A = 0$:

$$\begin{aligned}
\text{given} \quad & B, \Lambda \\
\min_{x \in \mathcal{L}(\mathbb{C}^d)} \quad & 0 \\
\text{s.t.} \quad & \Lambda(x) = B \\
& x \geq 0.
\end{aligned}$$

Every feasibility SDP can be turned into an optimization SDP. For example, by introducing an extra optimization variable and setting

$$\begin{aligned}
\text{given} \quad & B, \Lambda \\
\alpha^* := \max_{\alpha \in \mathbb{R},\, x \in \mathcal{L}(\mathbb{C}^d)} \quad & \alpha \\
\text{s.t.} \quad & \Lambda(x) = B \\
& x \geq \alpha \mathbb{1}.
\end{aligned}$$

If $\alpha^* \geq 0$, the SDP is feasible. If $\alpha^* < 0$, it is unfeasible.

Note that $\min f(x) = \max -f(x) = \min f'(x)$; both minimization and maximization problems are acceptable as an SDP.

Advantages of optimization over feasibility SDPs:

- Generally better behaved numerically.
- Generally easier to extract an infeasibility certificate.

2.3 Why is semidefinite programming relevant for us?

Because we know how to solve them:

1. SDPs can be solved efficiently numerically. Interior-point algorithms return a solution that is ϵ-close to optimal in polynomial time with respect to the dimension of the problem.
2. Duality properties of SDPs offer very useful techniques for analytical proofs.
3. Many relevant problems in quantum information can be solved or approximated with an SDP.

Example: State discrimination

Given a set of candidate states $\{\rho_i\}$ sampled with probability $\{p_i\}$, the maximal probability of successfully discriminating them is given by the SDP:

$$\begin{aligned} P^*_{\text{succ}} \coloneqq \max \quad & \sum_i p_i \text{tr}(\rho_i M_i) \\ \text{s.t.} \quad & M_i \geq 0 \quad \forall\, i \\ & \sum_i M_i = \mathbb{1}. \end{aligned}$$

The optimal solution for the SDP variables $\{M_i^*\}$ is the optimal quantum measurement for the discrimination of the state ensemble $\{p_i, \rho_i\}$.

3 Duality theory

Every SDP, referred from now on as the primal problem, has an associated SDP, referred to as the dual problem, with interesting properties. The main property is that it is constructed in such a way that it yields a lower bound to the primal problem.

Let us derive the dual problem of the SDP

$$\begin{aligned} \text{given} \quad & A, B, \Lambda & \\ \min_{x\in\mathcal{L}(\mathbb{C}^d)} \quad & \text{tr}(Ax) & \text{(linear objective function)} \\ \text{subject to} \quad & \Lambda(x) = B & \text{(affine constraint)} \\ & x \geq 0 & \text{(positive semidefinite cone)} \end{aligned}$$

We start by associating one dual variable (which corresponds to a Lagrange multiplier) to each constraint of the primal problem:

- $\Lambda(x) = B \longrightarrow y \in \mathcal{L}(\mathbb{C}^{d'})$
- $x \geq 0 \longrightarrow C \in \mathcal{L}(\mathbb{C}^d)$, $C \geq 0$

Then, we define the Lagrangian of the problem, which necessarily lower bounds the original primal problem. Loosely speaking, we define the real-valued Lagrangian function $L(x, y, C)$ of primal (x) and dual (y, C) variables according to

$$\begin{aligned} L(x,y,C) \coloneqq & \text{PRIMAL OBJ. FUNCTION} + \\ & -\sum \text{DUAL VAR.} \times \text{ASSOC. REAL-VALUED NULL CONSTR.} + \\ & -\sum \text{DUAL VAR.} \times \text{ASSOC. REAL-VALUED POS. CONSTR.} \end{aligned}$$

For the purpose of constructing the Lagrangian, let us first see how to transform the operator constraints in the primal problem into real-valued constraints to be used in the Lagrangian.

- **Operator equality constraint $\iff$ real-valued null constraint**

$$\Lambda(x) = B \iff \Lambda(x) - B = 0 \iff y(\Lambda(x) - B) = 0 \quad \forall\, y \iff \text{tr}\left[y(\Lambda(x) - B)\right] = 0 \quad \forall\, y$$

- **Operator inequality constraint $\iff$ real-valued positivity constraint**

$$x \geq 0 \iff xC \geq 0 \quad \forall\, C \geq 0 \iff \text{tr}\,(xC) \geq 0 \quad \forall\, C \geq 0$$

Hence,

$$\begin{aligned} L(x,y,C) &:= \operatorname{tr}(Ax) - \operatorname{tr}[y(\Lambda(x) - B)] - \operatorname{tr}(Cx) \\ &= \operatorname{tr}(yB) + \operatorname{tr}[x(A - \Lambda^\dagger(y) - C)]. \end{aligned} \tag{6}$$

This construction guarantees that the Lagrangian lower bounds the primal objective function for every value of the primal and dual variables, that is,

$$L(x,y,C) \leq f(x) \quad \forall\, x, y, C \geq 0. \tag{7}$$

When $A - \Lambda^\dagger(y) - C = 0$ is satisfied, the Lagrangian becomes independent of x:

$$L(x,y,C) = L(y,C) = \operatorname{tr}(yB) \leq f(x), \tag{8}$$

implying that

$$\begin{gathered} f(x) \geq \operatorname{tr}(yB) \quad \forall\, x \\ \text{when}\, C \geq 0 \quad \text{and} \quad A - \Lambda^\dagger(y) - C = 0 \end{gathered} \tag{9}$$

is a valid lower bound for $f(x)$. To find the best possible lower bound, we maximize this value, which yields the dual problem:

$$\begin{aligned} \max_{y,C} \quad & \operatorname{tr}(yB) \\ \text{s.t.} \quad & A - \Lambda^\dagger(y) - C = 0 \\ & C \geq 0 \end{aligned}$$

This can be further simplified into a dual problem with no redundant variables:

$$\begin{aligned} \text{given} \quad & A, B, \Lambda \\ \max_{y \in \mathcal{L}(\mathbb{C}^{d'})} \quad & \operatorname{tr}(yB) =: g(y) \\ \text{s.t.} \quad & A - \Lambda^\dagger(y) \geq 0 \end{aligned}$$

Note that:

- The dual problem of an SDP is always an SDP as well.
- The dual problem of the dual problem is the primal problem.

We can now define the feasible set of the dual SDP as

$$F_D := \{y \in \mathcal{L}(\mathbb{C}^{d'}) \,|\, A - \Lambda^\dagger(y) \geq 0\}, \tag{10}$$

and the optimal solution for the dual variable as

$$y^* := \arg\min_{y \in F_D} g(y). \tag{11}$$

Since the dual problem is an SDP, it also holds that this value is generally not unique. However, the optimal value for the objective function, denoted $g(y^*)$, is **always unique**. Furthermore, it holds that

$$g(y) \leq g(y^*) \leq f(x^*) \leq f(x) \quad \forall\, x \in F_P, y \in F_D. \tag{12}$$

3.1 Strong duality

Strong duality holds when the bound between the optimal solution of primal and dual problem is tight. That is, when

$$g(y^*) = f(x^*). \tag{13}$$

This can be checked via Slater's Condition (a condition that is sufficient but not necessary for strong duality): Strong duality holds if

- The primal problem is bounded and strictly feasible; and/or
- The dual problem is bounded and strictly feasible; and/or
- Primal and dual problems are strictly feasible.

Bounded means that

$$f(x^*) > -\infty \tag{14}$$
$$g(y^*) < \infty. \tag{15}$$

Strictly feasible means that

$$\exists\, x \in F_P \quad \text{s.t.} \quad x > 0 \tag{16}$$
$$\exists\, y \in F_D \quad \text{s.t.} \quad A - \Lambda^\dagger(y) > 0, \tag{17}$$

i.e., the inequality constraints can be strictly satisfied $\implies$ there exists a feasible point in the interior of the feasible set.

3.2 Why is duality relevant for us?

1. Useful method to find upper and lower bounds for a problem numerically.
2. Useful method to find upper and lower bounds, or to solve a problem analytically.
3. Either primal or dual SDP might be less costly to solve numerically (when not given in the canonical form)

4 Linear programming

Linear programming is a subclass of semidefinite programming that optimizes over real vectors instead of complex matrices $\longrightarrow$ equivalent to all matrices being diagonal.

References

[1] S. Boyd and L. Vandenberghe, Convex Optimization (Cambridge University Press, 2004).

[2] P. Skrzypczyk and D. Cavalcanti, Semidefinite Programming in Quantum Information Science (IOP Publishing, 2023).