Semidefinite Programming (SDP) in Quantum Information

"Semidefinite" : Positive semidefinite (PSD) constraints

"Programming" : Optimization problems

(e.g. linear programming)

⇒ SDP : Optimization problems with PSD constraints.

Why is SDP relevant in quantum info. ?

- Quantum states / measurements / maps are PSD operators.
  - states ↓ density op.
  - measurements ↓ POVM
  - maps ↓ Choi op.
- Many tasks in quantum info. are optimizations.

  e.g. state fidelity / discrimination.
- SDPs are computable (theoretically / numerically).
  - theoretically ↓ solve / bound
  - numerically ↓ efficient

More formal definition of SDP

From now on, we only consider finite dimensional, Hermitian operators.

An SDP can be written as

$$\max \quad \mathrm{Tr}\{AX\} \longrightarrow \text{Real linear function of } X$$

$$\text{s.t.} \quad \Phi_i(X) = B_i, \quad i = 1, 2, \cdots, m \longrightarrow \text{equality constraints}$$

$$\Gamma_j(X) \leq C_j, \quad j = 1, 2, \cdots, n \longrightarrow \text{inequality constraints}$$

$A \leq B \Leftrightarrow B - A$ is PSD.

X is an operator variable.

A, $\{B_i\}_i$, $\{C_j\}_j$ are operators defining the problem.

$\{\Phi_i\}_i$, $\{\Gamma_j\}_j$ are linear, hermiticity-preserving maps.

Nomenclature :

Feasible set $\mathcal{F} := \{X \mid \Phi_i(X) = B_i, \ \Gamma_j(X) \leq C_j, \ i = 1, \cdots, m, \ j = 1, \cdots, n\}$

Feasible operator: $X \in \mathcal{F}$

Optimal value : $V^* = \max\limits_{X \in \mathcal{F}} \mathrm{Tr}\{AX\}$

Optimal operator: $X^* = \underset{X \in \mathcal{F}}{\arg\max} \, \mathrm{Tr}\{AX\}$

key observation: $\mathcal{F}$ is convex, i.e. if $X_1, X_2 \in \mathcal{F}$,

$$p X_1 + (1-p) X_2 \in \mathcal{F} \quad \forall \, p \in [0,1]$$

$\Rightarrow$ There is no local maximum, i.e. all local maxima are global.

[Proof: Assume $\tilde{X}$ is a local maximum in some neighborhood of $\tilde{X}$, say $N(\tilde{X})$, s.t. $\exists \, X^* \in \mathcal{F}$, $\mathrm{Tr}\{AX^*\} > \mathrm{Tr}\{A\tilde{X}\}$. Then $p\tilde{X} + (1-p) X^* \in \mathcal{F}$, $\forall \, p \in [0,1]$

$$\begin{aligned} \mathrm{Tr}\{A[p\tilde{X} + (1-p)X^*]\} &= p\,\mathrm{Tr}\{A\tilde{X}\} + (1-p)\,\mathrm{Tr}\{AX^*\} \\ &> p\,\mathrm{Tr}\{A\tilde{X}\} + (1-p)\,\mathrm{Tr}\{A\tilde{X}\} \\ &= \mathrm{Tr}\{A\tilde{X}\}. \end{aligned}$$

However, as $p \to 1$, $p\tilde{X} + (1-p)X^* \in N(\hat{X})$. The inequality contradicts with the local optimality of $\tilde{X}$.]

This guarantees the efficiency of solving SDPs numerically.

$\Rightarrow$ X^* can always be an extreme point of $\mathcal{F}$.

[Proof: Let $X^* = \sum_i p_i \hat{X}_i$ where $\hat{X}_i$ is extremal, $\forall \, i$. Then

$$\mathrm{Tr}\{AX^*\} = \sum_i p_i \, \mathrm{Tr}\{A\hat{X}_i\} \le \max_i \mathrm{Tr}\{A\hat{X}_i\}$$

By optimality, $\mathrm{Tr}\{AX^*\} \ge \max_i \mathrm{Tr}\{A\hat{X}_i\}$

So X^* can be chosen as $\underset{i}{\arg\max} \, \mathrm{Tr}\{A\hat{X}_i\}$]

This shows why theoretically solving SDPs is possible.

Example: Given an operator H, what is its highest eigenvalue ?

$$\begin{aligned} \max \quad & \mathrm{Tr}\{H\rho\} \\ \text{s.t.} \quad & \mathrm{Tr}\{\rho\} = 1 \\ & \rho \ge 0 \end{aligned}$$

$\mathcal{F}$ = set of quantum states.

ρ^* = the eigenvector of H with the maximal eigenvalue.

pure / extremal.

"Standard" form of SDPs

$$\max \quad \mathrm{Tr}\{AX\}$$
$$\text{s.t.} \quad \Phi(X) = B$$
$$X \geq 0$$

We can always cast an SDP into this form. (See Watrous' notes on SDPs).

We use this form to introduce a very important concept about SDPs.

Duality

Primal SDP

$$\max \quad \mathrm{Tr}\{AX\}$$
$$\text{s.t.} \quad \Phi(X) = B \qquad \text{Lagrange multiplier } Y$$
$$X \geq 0 \qquad \text{Lagrange multiplier } Z$$

- Every SDP has a dual formulation, such that any feasible points of the dual problem provides an upper bound to the primal SDP.

- Under mild conditions, we have a tight bound, i.e.

$$V^*_{primal} = V^*_{dual}.$$

How to find the dual SDP?

Define the Lagrangian with the Lagrange multipliers Y, Z

$$L(X, Y, Z) := \mathrm{Tr}\{AX\} + \mathrm{Tr}\{Y[B - \Phi(X)]\} + \mathrm{Tr}\{ZX\} \qquad \mathrm{Tr}\{Y\Phi(X)\} = \mathrm{Tr}\{\Phi^\dagger(Y)X\}.$$

$$= \mathrm{Tr}\{[A + Z - \Phi^\dagger(Y)]X\} + \mathrm{Tr}\{BY\}$$

To get an upper bound, we require $Z \geq 0$, so that

$$L(X, Y, Z) \geq \mathrm{Tr}\{AX\}, \quad \forall X \in \mathcal{F}, Z \geq 0, Y$$

Thus, $g(Y, Z) := \max_X L(X, Y, Z) \geq V^*_{primal} \equiv \max_{X \in \mathcal{F}} \mathrm{Tr}\{AX\}, \quad \forall Z \geq 0, Y$

However, $g(Y, Z) \to \infty$ if $A + Z - \Phi^\dagger(Y) \neq 0$

The non-trivial upper bound is

$$g(Y, Z)\big|_{A + Z - \Phi^\dagger(Y) = 0}$$

The best upper bound is

$$\min \; g(Y,Z)\big|_{A+Z-\Phi^\dagger(Y)=0} \equiv \mathrm{Tr}\{BY\}$$
$$\text{s.t.} \quad \Phi^\dagger(Y) = A + Z$$
$$Z \geq 0$$

Primal SDP

$$\max \quad \mathrm{Tr}\{AX\}$$
$$\text{s.t.} \quad \Phi(X) = B$$
$$X \geq 0$$

Dual SDP

$$\min \quad \mathrm{Tr}\{BY\}$$
$$\text{s.t.} \quad \Phi^\dagger(Y) \geq A$$

Theorem (Weak duality)

For every SDP, it holds that

$$V^*_{\text{primal}} \leq V^*_{\text{dual}}.$$

[Proof. For every $X \in \mathcal{F}_{\text{primal}}$ and $Y \in \mathcal{F}_{\text{dual}}$,

$$\mathrm{Tr}\{AX\} \leq \mathrm{Tr}\{\Phi^\dagger(Y)X\} = \mathrm{Tr}\{Y\Phi(X)\} = \mathrm{Tr}\{BY\}.$$]

Theorem (Strong duality)

For every SDP with non-empty $\mathcal{F}_{\text{primal}}$ and $\mathcal{F}_{\text{dual}}$, if either of the following holds:

1. $\exists\, Y \in \mathcal{F}_{\text{dual}}$, s.t. $\Phi^\dagger(Y) > A$
2. $\exists\, X \in \mathcal{F}_{\text{primal}}$, s.t. $X > 0$.

then we have

$$V^*_{\text{primal}} = V^*_{\text{dual}}.$$

Proof. See Watrous' notes on SDPs.

Strong duality holds if the primal and dual SDPs are <u>strictly feasible</u>.

Note: The Lagrange multiplers method is general for obtaining the dual problem. You can also use it to go from the dual SDP to the primal SDP.

The SDP of trace distance

The trace norm of an operator A is defined as

$$\|A\|_1 := \mathrm{Tr}\{\sqrt{A^\dagger A}\} = \sum_i |\lambda_i(A)|,$$

where $\{|\lambda_i(A)|\}_i$ are eigenvalues of A.

The SDP of $\|A\|_1$:

$$\|A\|_1 = \min \ \mathrm{Tr}\{X\}$$

$$\text{s.t.} \quad -X \leq A \leq X$$

Proof: We show that $\mathrm{Tr}\{X\} \geq \|A\|_1 \ \forall X \in \mathcal{F}$ and the equality holds for an X^*.

Consider the eigendecomposition of A, $A = \sum_i \lambda_i |\psi_i\rangle\langle\psi_i|$

Define $P_+ \equiv \sum_{i|\lambda_i \geq 0} |\psi_i\rangle\langle\psi_i|$

$P_- \equiv \sum_{i|\lambda_i < 0} |\psi_i\rangle\langle\psi_i|$

so $P_+ + P_- = \mathbb{1}$ and $\mathrm{Tr}\{A(P_+ - P_-)\} = \|A\|_1$.

For $X \in \mathcal{F}$, $-X \leq A \leq X$,

$$\left.\begin{aligned} &\mathrm{Tr}\{(X-A)P_+\} \geq 0 \\ &\mathrm{Tr}\{(X+A)P_-\} \geq 0 \end{aligned}\right\} \Rightarrow \mathrm{Tr}\{X\} \geq \|A\|_1 \quad \forall X \in \mathcal{F}.$$

$X^* \equiv \sum_{i|\lambda_i \geq 0} \lambda_i |\psi_i\rangle\langle\psi_i| - \sum_{i|\lambda_i \leq 0} \lambda_i |\psi_i\rangle\langle\psi_i|$ s.t. $\mathrm{Tr}\{X^*\} = \|A\|_1$.

The trace distance between two states ρ and σ is

$$d_{\mathrm{Tr}}(\rho,\sigma) := \tfrac{1}{2}\|\rho - \sigma\|_1$$

$$= \min \ \tfrac{1}{2}\mathrm{Tr}\{X\}$$

$$\text{s.t.} \quad -X \leq \rho - \sigma \leq X$$

The dual SDP of $d_{\mathrm{Tr}}(\rho,\sigma)$ is

$$d_{\mathrm{Tr}}(\rho,\sigma) = \max \ \mathrm{Tr}\{(\rho-\sigma)Y\}$$

strong duality holds $\quad \text{s.t.} \quad 0 \leq Y \leq \mathbb{1}$

What is the maximal success probability to distinguish ρ and σ by a binary POVM $\{\underset{\text{"0"}}{Y}, \underset{\text{"1"}}{\mathbb{1} - Y}\}$, if ρ and σ are prepared with the equal probability $1/2$?

$$P^*_{succ} = \max_{0 \le Y \le \mathbb{1}} \frac{1}{2}\mathrm{Tr}\{Y\rho\} + \frac{1}{2}\mathrm{Tr}\{(\mathbb{1}-Y)\sigma\}$$
$$= \max_{0 \le Y \le \mathbb{1}} \frac{1}{2}\left(1 + \mathrm{Tr}\{(\rho-\sigma)Y\}\right)$$
$$= \frac{1}{2}\left(1 + d_{Tr}(\rho,\sigma)\right)$$

Directly prove it from SDP!

Data-processing inequality of d_{tr}

$$d_{Tr}(\rho,\sigma) \ge d_{Tr}(\mathcal{N}(\rho),\mathcal{N}(\sigma)) \qquad \forall\ \mathcal{N} \in \mathrm{CPTP}$$

Proof. Consider two sets $\mathcal{F} = \{Y \mid 0 \le Y \le \mathbb{1}\}$

$\mathcal{F}' = \{Y \mid 0 \le \mathcal{N}^\dagger(Y) \le \mathbb{1}\}$

Since $\mathcal{N}^\dagger$ is unital, $\forall\ Y \in \mathcal{F}$, $Y \in \mathcal{F}' \Rightarrow \mathcal{F} \subseteq \mathcal{F}'$

$$\max_{0 \le Y \le \mathbb{1}} \mathrm{Tr}\{(\mathcal{N}(\rho)-\mathcal{N}(\sigma))Y\} = \max_{0 \le Y \le \mathbb{1}} \mathrm{Tr}\{(\rho-\sigma)\mathcal{N}^\dagger(Y)\}$$
$$= \max_{Y \in \mathcal{F}} \mathrm{Tr}\{(\rho-\sigma)\mathcal{N}^\dagger(Y)\}$$
$$\le \max_{Y \in \mathcal{F}'} \mathrm{Tr}\{(\rho-\sigma)\mathcal{N}^\dagger(Y)\}$$
$$= \max_{0 \le \mathcal{N}^\dagger(Y) \le \mathbb{1}} \mathrm{Tr}\{(\rho-\sigma)\mathcal{N}^\dagger(Y)\}$$
$$= \max_{0 \le Z \le \mathbb{1}} \mathrm{Tr}\{(\rho-\sigma)Z\}$$

$\Rightarrow d_{Tr}(\mathcal{N}(\rho),\mathcal{N}(\sigma)) \le d_{Tr}(\rho,\sigma)$

Equality holds when $\mathcal{N}$ is reversible (unitary).

Also proved by using SDP!

You can also prove the triangle inequality for d_{Tr} by SDP.

Trace distance as an SDP constraint

Given a state σ, the ε-neighborhood of σ can be defined as a trace-norm ball:

$$B_\varepsilon(\sigma) := \{\rho \mid d_{Tr}(\rho,\sigma) \le \varepsilon,\ \rho \in \mathcal{S}\}.$$

Given an operator A, what is the maximal expectation value of A in the ε-neighborhood of a target state σ? (like a worst-case guarantee, if thinking A as Hamiltonian and σ as ground state).

$$\max_{\rho \in B_\varepsilon(\sigma)} \langle A \rangle_\rho = \max \operatorname{Tr}\{A\rho\}$$

$$\text{s.t.} \quad \operatorname{Tr}\{\rho\} = 1$$
$$\rho \geq 0$$
$$d_{Tr}(\rho, \sigma) \leq \varepsilon$$

How to make the maximisation be an SDP ?

Theorem (L. Vandenberghe & S. Boyd Semidefinite Programming , SIAM Review, 1996)

Given a matrix Z_{12}, there exists Z_{11} and Z_{22}, such that

$$\operatorname{Tr}\{Z_{11} + Z_{22}\} = 2\varepsilon$$

$$\begin{pmatrix} Z_{11} & Z_{12} \\ Z_{12}^{+} & Z_{22} \end{pmatrix} \geq 0$$

if and only if $\|Z_{12}\|_1 \leq \varepsilon$,

Proof. Let $Z_{12} = U \Sigma V^{+}$ be the singular value decomposition of Z_{12}, where U and V are d-dimensional unitaries and Σ is d-dimensional diagonal matrix.

If such Z_{11} and Z_{22} exist,

$$\operatorname{Tr}\left\{ \begin{pmatrix} UU^{+} & -UV^{+} \\ -VU^{+} & VV^{+} \end{pmatrix} \begin{pmatrix} Z_{11} & Z_{12} \\ Z_{12}^{+} & Z_{22} \end{pmatrix} \right\} \geq 0$$

since $\begin{pmatrix} UU^{+} & -UV^{+} \\ -VU^{+} & VV^{+} \end{pmatrix} = (U \ -V) \begin{pmatrix} U^{+} \\ -V^{+} \end{pmatrix} \geq 0$.

Consequently,

$$0 \leq -\operatorname{Tr}\{UV^{+}Z_{12}^{+}\} - \operatorname{Tr}\{VU^{+}Z_{12}\} + \operatorname{Tr}\{UU^{+}Z_{11}\} + \operatorname{Tr}\{VV^{+}Z_{22}\}$$

$$\leq -2\operatorname{Tr}\{\Sigma\} + \operatorname{Tr}\{Z_{11} + Z_{22}\}$$

$$= -2\|Z_{12}\|_1 + 2\varepsilon$$

$$\Rightarrow \|Z_{12}\|_1 \leq \varepsilon.$$

If $\|Z_{12}\|_1 \leq \varepsilon$, we can define

$$Z_{11} = U\Sigma U^{+} + \gamma \mathbb{1}, \qquad Z_{22} = V\Sigma V^{+} + \gamma \mathbb{1}$$

with $\gamma = (2\varepsilon - \|Z_{12}\|_1)/2d$, such that

$$\operatorname{Tr}\{Z_{11} + Z_{22}\} = 2\|Z_{12}\|_1 + 2d\gamma = 2\varepsilon.$$

$$\begin{pmatrix} Z_{11} & Z_{12} \\ Z_{12}^{+} & Z_{22} \end{pmatrix} = \begin{pmatrix} U\Sigma U^{+} + \gamma\mathbb{1} & U\Sigma V^{+} \\ V\Sigma U^{+} & V\Sigma V^{+} + \gamma\mathbb{1} \end{pmatrix}$$

$$= (U\Sigma^{1/2}, V\Sigma^{1/2}) \begin{pmatrix} \Sigma^{1/2}U^{+} \\ \Sigma^{1/2}V^{+} \end{pmatrix} + \gamma\mathbb{1} \geqslant 0.$$

So such Z_{11}, Z_{22} exist.

Thus, we can write

$$\max_{\rho \in B_\varepsilon(\sigma)} \langle A \rangle_\rho = \max \ \mathrm{Tr}\{A\rho\}$$

$$\begin{aligned} \text{s.t.} \quad & \mathrm{Tr}\{\rho\} = 1 \\ & \rho \geqslant 0 \\ & \mathrm{Tr}\{Z_1 + Z_2\} = 4\varepsilon \\ & \begin{pmatrix} Z_1 & \rho - \sigma \\ \rho - \sigma & Z_2 \end{pmatrix} \geqslant 0 \end{aligned}$$

References on SDPs

- P. Skrzypczyk and D. Cavalcanti, Semidefinite Programming in Quantum Information Science, IOP Publishing, 2023. https://iopscience.iop.org/book/mono/978-0-7503-3343-6
- J. Watrous, Lecture 7: Semidefinite programming, Theory of Quantum Information, 2011. https://johnwatrous.com/wp-content/uploads/TQI-notes.07.pdf
- L. Vandenberghe and S. Boyd, SIAM Review, 38(1): 49-95, March 1996. https://stanford.edu/~boyd/papers/sdp.html

Numerical implementation of SDPs

- CVXPY for Python (recommended): https://www.cvxpy.org
- CVX for MATLAB: https://cvxr.com/cvx/doc/quickstart.html