SDP as a tool for quantum information

Lecture 1: What is an SDP?

Marco Túlio Quintino

QALYPSO 2026

Lecture 1: What is an SDP?

1. Warming up

Lecture 1: What is an SDP?

1. Warming up
2. What is an SDP?

Lecture 1: What is an SDP?

1. Warming up
2. What is an SDP?
3. Convex optimisation and the feasible set

Lecture 1: What is an SDP?

1. Warming up
2. What is an SDP?
3. Convex optimisation and the feasible set
4. State discrimination

Lecture 1: What is an SDP?

1. Warming up
2. What is an SDP?
3. Convex optimisation and the feasible set
4. State discrimination
5. Duality theory

1. Warming up

Linearity, matrices, and positive semidefinite

Linearity is at the core of quantum theory

Linearity

Linearity is at the core of quantum theory

Linearity

a linear map is a function between vector spaces

$$f : \mathcal{H} \longrightarrow \mathcal{H}'$$

Linearity is at the core of quantum theory

Linearity

a linear map is a function between vector spaces

$$f : \mathcal{H} \longrightarrow \mathcal{H}'$$

a function that respects

$$f(\alpha x + \beta y) = \alpha f(x) + \beta f(y)$$

for every $x, y \in \mathcal{H}$ and $\alpha, \beta \in \mathbb{C}$

Linearity is at the core of quantum theory

$$X \in \mathcal{L}(\mathbb{C}^d)$$

Linearity is at the core of quantum theory

$$X \in \mathcal{L}(\mathbb{C}^d)$$

X is a square $d \times d$ matrix

$$X = (X_{ij})_{i,j=1}^{d} = \begin{pmatrix} X_{11} & \cdots & X_{1d} \\ \vdots & \ddots & \vdots \\ X_{d1} & \cdots & X_{dd} \end{pmatrix}$$

Linearity is at the core of quantum theory

A map

$$\Lambda : \mathcal{L}(\mathbb{C}^d) \longrightarrow \mathcal{L}(\mathbb{C}^{d'})$$

is *linear* when

Linearity is at the core of quantum theory

A map

$$\Lambda : \mathcal{L}(\mathbb{C}^d) \longrightarrow \mathcal{L}(\mathbb{C}^{d'})$$

is *linear* when

$$\Lambda(\alpha X + \beta Y) = \alpha\Lambda(X) + \beta\Lambda(Y)$$

for every $X, Y \in \mathcal{L}(\mathbb{C}^d)$ and $\alpha, \beta \in \mathbb{C}$

Linearity is at the core of quantum theory

A map

$$\Lambda : \mathcal{L}(\mathbb{C}^d) \longrightarrow \mathcal{L}(\mathbb{C}^{d'})$$

is *linear* when

$$\Lambda(\alpha X + \beta Y) = \alpha\Lambda(X) + \beta\Lambda(Y)$$

for every $X, Y \in \mathcal{L}(\mathbb{C}^d)$ and $\alpha, \beta \in \mathbb{C}$

Quantum evolution

$$\rho \longmapsto \Lambda(\rho)$$

Linearity is at the core of quantum theory

A map

$$\Lambda : \mathcal{L}(\mathbb{C}^d) \longrightarrow \mathcal{L}(\mathbb{C}^{d'})$$

is *linear* when

$$\Lambda(\alpha X + \beta Y) = \alpha\Lambda(X) + \beta\Lambda(Y)$$

for every $X, Y \in \mathcal{L}(\mathbb{C}^d)$ and $\alpha, \beta \in \mathbb{C}$

Quantum evolution

$$\rho \longmapsto \Lambda(\rho)$$

Born rule

$$p(i \mid \rho, \{M_i\}_i) = \mathrm{tr}(M_i\rho)$$

PSD means non-negative in every direction

PSD = Positive SemiDefinite

$$X \succeq 0 \iff \begin{cases} X = X^\dagger \\ \langle\psi|X|\psi\rangle \geq 0 \quad \forall\ |\psi\rangle \in \mathbb{C}^d \end{cases}$$

PSD means non-negative in every direction

PSD = Positive SemiDefinite

$$X \succeq 0 \iff \begin{cases} X = X^\dagger \\ \langle\psi|X|\psi\rangle \geq 0 \quad \forall\ |\psi\rangle \in \mathbb{C}^d \end{cases}$$

$$\langle\psi|X|\psi\rangle \geq 0 \quad \forall\ |\psi\rangle \in \mathbb{C}^d \implies X = X^\dagger$$

PSD means non-negative in every direction

PSD = Positive SemiDefinite

$$X \succeq 0 \iff \begin{cases} X = X^\dagger \\ \langle\psi|X|\psi\rangle \geq 0 \quad \forall\ |\psi\rangle \in \mathbb{C}^d \end{cases}$$

$$\langle\psi|X|\psi\rangle \geq 0 \quad \forall\ |\psi\rangle \in \mathbb{C}^d \implies X = X^\dagger$$

$$\langle\psi|X|\psi\rangle \geq 0 \quad \forall\ |\psi\rangle \in \mathbb{R}^d \;\not\Rightarrow\; X = X^\mathsf{T}$$

PSD means non-negative in every direction

PSD = Positive SemiDefinite

$$X \succeq 0 \iff \begin{cases} X = X^\dagger \\ \langle\psi|X|\psi\rangle \geq 0 \quad \forall\, |\psi\rangle \in \mathbb{C}^d \end{cases}$$

$$X \succeq 0 \iff \begin{cases} X = X^\dagger \\ \text{eig}_i(X) \geq 0 \quad \forall\, i \end{cases}$$

PSD means non-negative in every direction

$$\textbf{PSD} = \text{Positive SemiDefinite}$$

$$X \succeq 0 \iff \begin{cases} X = X^\dagger \\ \langle\psi|X|\psi\rangle \geq 0 \quad \forall\, |\psi\rangle \in \mathbb{C}^d \end{cases}$$

$$X \succeq 0 \iff \begin{cases} X = X^\dagger \\ \text{eig}_i(X) \geq 0 \quad \forall\, i \end{cases}$$

Example:

$$A = \begin{pmatrix} 1 & -1 \\ -1 & 1 \end{pmatrix} \succeq 0$$

PSD means non-negative in every direction

PSD = Positive SemiDefinite

$$X \succeq 0 \iff \begin{cases} X = X^\dagger \\ \langle\psi|X|\psi\rangle \geq 0 \quad \forall\, |\psi\rangle \in \mathbb{C}^d \end{cases}$$

$$X \succeq 0 \iff \begin{cases} X = X^\dagger \\ \text{eig}_i(X) \geq 0 \quad \forall\, i \end{cases}$$

Example:

$$A = \begin{pmatrix} 1 & -1 \\ -1 & 1 \end{pmatrix} \succeq 0 \qquad\qquad A = A^\dagger$$

PSD means non-negative in every direction

PSD = Positive SemiDefinite

$$X \succeq 0 \quad \Longleftrightarrow \quad \begin{cases} X = X^\dagger \\ \langle\psi|X|\psi\rangle \geq 0 \quad \forall\ |\psi\rangle \in \mathbb{C}^d \end{cases}$$

$$X \succeq 0 \quad \Longleftrightarrow \quad \begin{cases} X = X^\dagger \\ \mathrm{eig}_i(X) \geq 0 \quad \forall\, i \end{cases}$$

Example:

$$A = \begin{pmatrix} 1 & -1 \\ -1 & 1 \end{pmatrix} \succeq 0$$

$$A = A^\dagger$$

A has eigenvalues 0 and 2

Positive and strictly positive

$$X \succeq 0 \qquad \Longleftrightarrow \qquad \langle\psi|X|\psi\rangle \geq 0 \quad \forall\ |\psi\rangle \in \mathbb{C}^d$$

positive semidefinite (PSD), also called *positive*

Positive and strictly positive

$$X \succeq 0 \qquad \Longleftrightarrow \qquad \langle\psi|X|\psi\rangle \geq 0 \quad \forall\ |\psi\rangle \in \mathbb{C}^d$$

positive semidefinite (PSD), also called *positive*

$$X \succ 0 \qquad \Longleftrightarrow \qquad \langle\psi|X|\psi\rangle > 0 \quad \forall\, 0 \neq |\psi\rangle \in \mathbb{C}^d$$

positive definite (PD), also called *strictly positive*

Positivity is everywhere in quantum theory

States

$$\rho \succeq 0, \qquad \mathrm{tr}(\rho) = 1$$

Positivity is everywhere in quantum theory

States

$$\rho \succeq 0, \qquad \mathrm{tr}(\rho) = 1$$

Measurements

$$M_i \succeq 0, \qquad \sum_i M_i = \mathbb{1}$$

Positivity is everywhere in quantum theory

States

$$\rho \succeq 0, \qquad \mathrm{tr}(\rho) = 1$$

Measurements

$$M_i \succeq 0, \qquad \sum_i M_i = \mathbb{1}$$

Channels

$$\Lambda : \mathcal{L}(\mathbb{C}^d) \longrightarrow \mathcal{L}(\mathbb{C}^{d'})$$

$\Lambda \in \mathrm{CPTP}$ (Completely Positive and Trace Preserving)

Positivity is everywhere in quantum theory

States

$$\rho \succeq 0, \qquad \mathrm{tr}(\rho) = 1$$

Measurements

$$M_i \succeq 0, \qquad \sum_i M_i = \mathbb{1}$$

Channels

$$\Lambda : \mathcal{L}(\mathbb{C}^d) \longrightarrow \mathcal{L}(\mathbb{C}^{d'})$$

$\Lambda \in \mathrm{CPTP}$ (Completely Positive and Trace Preserving)

$$\mathrm{CP} \iff \rho_\Lambda \succeq 0 \qquad \mathrm{TP} \iff \mathrm{tr}_{\mathrm{out}}(\rho_\Lambda) = \mathbb{1}_{\mathrm{in}}$$

$\rho_\Lambda \in \mathcal{L}(\mathbb{C}^d) \otimes \mathcal{L}(\mathbb{C}^{d'})$ is the Choi operator of Λ

2. What is an SDP?

Linear objective function, affine constraints, and positive semidefiniteness

An SDP has three ingredients

Linear objective function

An SDP has three ingredients

Linear objective function

$+$ affine constraints

An SDP has three ingredients

Linear objective function

$+$ affine constraints

$+$ positive-semidefinite constraint

Notation

$\mathbb{C}^d$	a d-dimensional complex vector space
$\mathcal{L}(\mathbb{C}^d)$	linear operators acting on $\mathbb{C}^d$ (d by d complex matrices)
$X = X^\dagger$	X is Hermitian (or, X is self-adjoint)

$$\Lambda : \mathcal{L}(\mathbb{C}^d) \longrightarrow \mathcal{L}(\mathbb{C}^{d'})$$

Notation

$\mathbb{C}^d$	a d-dimensional complex vector space
$\mathcal{L}(\mathbb{C}^d)$	linear operators acting on $\mathbb{C}^d$ (d by d complex matrices)
$X = X^\dagger$	X is Hermitian (or, X is self-adjoint)

$$\Lambda : \mathcal{L}(\mathbb{C}^d) \longrightarrow \mathcal{L}(\mathbb{C}^{d'})$$

Λ is Hermiticity-preserving when $X = X^\dagger \Rightarrow \Lambda(X) = \Lambda(X)^\dagger$

Semidefinite Programming

Given:

$$A \in \mathcal{L}(\mathbb{C}^d) \qquad B \in \mathcal{L}(\mathbb{C}^{d'})$$

$$\Lambda : \mathcal{L}(\mathbb{C}^d) \longrightarrow \mathcal{L}(\mathbb{C}^{d'})$$

where $A = A^\dagger$, $B = B^\dagger$, and Λ is a Hermiticity-preserving linear map

Semidefinite Programming

Given:

$$A \in \mathcal{L}(\mathbb{C}^d) \qquad B \in \mathcal{L}(\mathbb{C}^{d'})$$

$$\Lambda : \mathcal{L}(\mathbb{C}^d) \longrightarrow \mathcal{L}(\mathbb{C}^{d'})$$

where $A = A^\dagger$, $B = B^\dagger$, and Λ is a Hermiticity-preserving linear map

$$\begin{aligned} \text{maximise} \quad & \operatorname{tr}(A\rho) \\ \text{subject to} \quad & \Lambda(\rho) = B \\ & \rho \succeq 0 \end{aligned}$$

Semidefinite Programming

Given:

$$A \in \mathcal{L}(\mathbb{C}^d) \qquad B \in \mathcal{L}(\mathbb{C}^{d'})$$

$$\Lambda : \mathcal{L}(\mathbb{C}^d) \longrightarrow \mathcal{L}(\mathbb{C}^{d'})$$

where $A = A^\dagger$, $B = B^\dagger$, and Λ is a Hermiticity-preserving linear map

$$\begin{aligned} \text{maximise} \quad & \operatorname{tr}(A\rho) \\ \text{subject to} \quad & \Lambda(\rho) = B \\ & \rho \succeq 0 \end{aligned}$$

$$\alpha := \sup\{\operatorname{tr}(A\rho) \mid \Lambda(\rho) = B, \rho \succeq 0\}$$

Great! We know what an SDP is!

Great! We know what an SDP is!

1. Why is SDP important and useful?

Great! We know what an SDP is!

1. Why is SDP important and useful?

2. Examples: using SDP in quantum theory

Great! We know what an SDP is!

1. Why is SDP important and useful?

2. Examples: using SDP in quantum theory

Before getting there, how about a little digression?

Programming

Programming predates computers

program = plan

First known use of *programming*

1896

Merriam-Webster Dictionary: "program" | "programming"

In optimisation, programming means planning

Plan the best action under constraints

George B. Dantzig

"Programming in a Linear Structure"

Econometrica 17 (1949), pp. 73–74

The two uses developed independently

Computer programming

was not named after linear programming

nor vice versa

The two uses developed independently

Computer programming

was not named after linear programming

nor vice versa

1937

Bryce's IBM patent

"programing means"

1942

Mauchly's memorandum

"program device"

1948

Dantzig's report

linear programming

Both extend the older meaning: **an organised plan**

The ENIAC project established the computing usage

Sources: Bryce, US 2,244,241 (filed 1937); Mauchly (1942)
Haigh–Priestley (2016); The Craft of Coding (2022)

SDP has two closely related meanings

Semidefinite programming

the class of optimisation problems

SDP has two closely related meanings

Semidefinite programming

the class of optimisation problems

Semidefinite program

one particular optimisation problem

SDP has two closely related meanings

Semidefinite programming

the class of optimisation problems

Semidefinite program

one particular optimisation problem

The intended meaning is usually clear from context

3. Convex optimisation and the feasible set

Convex sets, convex cones, and geometrical interpretation

The feasible set

Given: $A \in \mathcal{L}(\mathbb{C}^d)$, $B \in \mathcal{L}(\mathbb{C}^{d'})$

$$\Lambda : \mathcal{L}(\mathbb{C}^d) \longrightarrow \mathcal{L}(\mathbb{C}^{d'})$$

$$\begin{aligned} \text{maximise} \quad & \operatorname{tr}(A\rho) \\ \text{subject to} \quad & \Lambda(\rho) = B \\ & \rho \succeq 0 \end{aligned}$$

The feasible set

$$\textbf{Given:} \quad A \in \mathcal{L}(\mathbb{C}^d), \quad B \in \mathcal{L}(\mathbb{C}^{d'})$$
$$\Lambda : \mathcal{L}(\mathbb{C}^d) \longrightarrow \mathcal{L}(\mathbb{C}^{d'})$$

$$\begin{aligned} \text{maximise} \quad & \operatorname{tr}(A\rho) \\ \text{subject to} \quad & \Lambda(\rho) = B \\ & \rho \succeq 0 \end{aligned}$$

$$\mathcal{F} = \left\{ \rho \in \mathcal{L}(\mathbb{C}^d) \mid \Lambda(\rho) = B,\ \rho \succeq 0 \right\}$$

The feasible set

Given: $A \in \mathcal{L}(\mathbb{C}^d)$, $B \in \mathcal{L}(\mathbb{C}^{d'})$

$$\Lambda : \mathcal{L}(\mathbb{C}^d) \longrightarrow \mathcal{L}(\mathbb{C}^{d'})$$

$$\begin{array}{rl} \text{maximise} & \operatorname{tr}(A\rho) \\ \text{subject to} & \Lambda(\rho) = B \\ & \rho \succeq 0 \end{array}$$

$$\mathcal{F} = \left\{ \rho \in \mathcal{L}(\mathbb{C}^d) \mid \Lambda(\rho) = B,\ \rho \succeq 0 \right\}$$

SDP: Maximise $\operatorname{tr}(A\rho)$ over $\mathcal{F}$

The feasible set

$$\mathcal{F} = \left\{\rho \in \mathcal{L}(\mathbb{C}^d) \mid \Lambda(\rho) = B,\ \ \rho \succeq 0\right\}$$

What can we say about the feasible set?

Geometrical interpretation

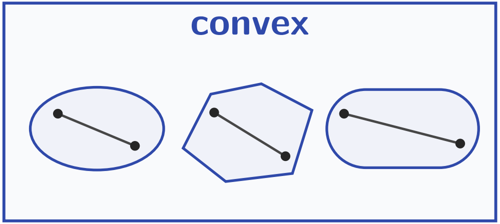
A set $\mathcal{C}$ is *convex* when

$$X, Y \in \mathcal{C}, \quad p \in [0,1] \quad \Longrightarrow \quad pX + (1-p)Y \in \mathcal{C}$$

Geometrical interpretation

A set $\mathcal{C}$ is *convex* when

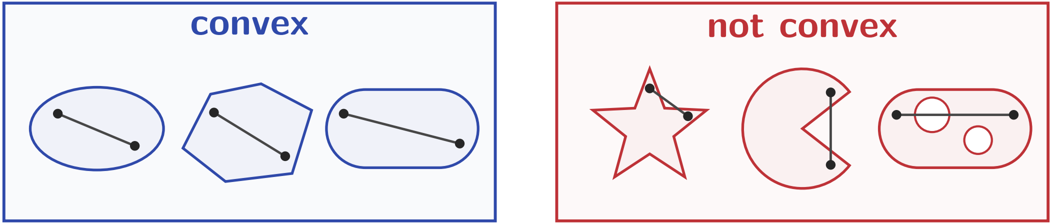
$$X, Y \in \mathcal{C}, \quad p \in [0,1] \quad \Longrightarrow \quad pX + (1-p)Y \in \mathcal{C}$$



Geometrical interpretation

A set $\mathcal{C}$ is *convex* when

$$X, Y \in \mathcal{C}, \quad p \in [0,1] \quad \Longrightarrow \quad pX + (1-p)Y \in \mathcal{C}$$



PSD matrices form a convex set

$$\mathcal{P}_d = \left\{\rho \in \mathcal{L}(\mathbb{C}^d) \mid \rho \succeq 0\right\}$$

PSD matrices form a convex set

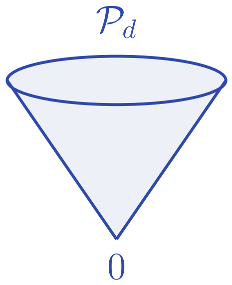
$$\mathcal{P}_d = \left\{\rho \in \mathcal{L}(\mathbb{C}^d) \mid \rho \succeq 0\right\}$$

$$\begin{aligned} &\rho_1, \rho_2 \in \mathcal{P}_d, \quad p \in [0,1] \\ \implies &\quad p\rho_1 + (1-p)\rho_2 \in \mathcal{P}_d \end{aligned}$$

PSD matrices form a convex cone

$$\mathcal{P}_d = \left\{\rho \in \mathcal{L}(\mathbb{C}^d) \mid \rho \succeq 0\right\}$$

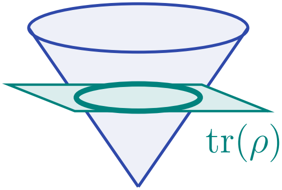
$$\rho \in \mathcal{P}_d, \quad t \geq 0 \quad \Longrightarrow \quad t\rho \in \mathcal{P}_d$$



PSD matrices form a convex cone

$$\mathcal{P}_d = \left\{\rho \in \mathcal{L}(\mathbb{C}^d) \mid \rho \succeq 0\right\}$$

$$\rho \in \mathcal{P}_d, \quad t \geq 0 \quad \implies \quad t\rho \in \mathcal{P}_d$$

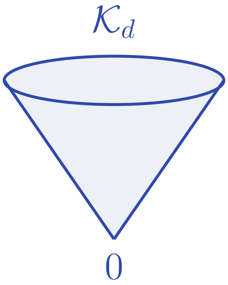


Quantum states

$$\mathrm{D}_d = \left\{\rho \in \mathcal{P}_d \mid \mathrm{tr}(\rho) = 1\right\} = \mathcal{P}_d \cap \left\{\rho \mid \mathrm{tr}(\rho) = 1\right\}$$

Every SDP feasible set is an affine slice

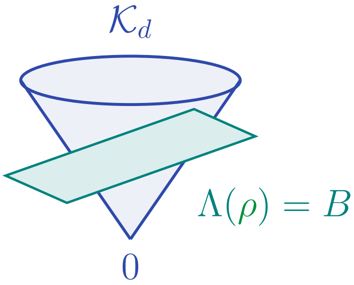
$$\mathcal{K}_d = \left\{\rho \mid \rho \succeq 0\right\}$$ PSD cone



Every SDP feasible set is an affine slice

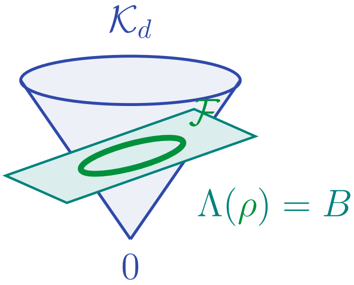
$$\mathcal{K}_d = \left\{ \rho \mid \rho \succeq 0 \right\}$$ PSD cone

$$\mathcal{A}_B = \left\{ \rho \mid \Lambda(\rho) = B \right\}$$ affine subspace



Every SDP feasible set is an affine slice

$\mathcal{K}_d = \left\{\rho \mid \rho \succeq 0\right\}$	PSD cone
$\mathcal{A}_B = \left\{\rho \mid \Lambda(\rho) = B\right\}$	affine subspace
$\mathcal{F} = \mathcal{K}_d \cap \mathcal{A}_B$	feasible set

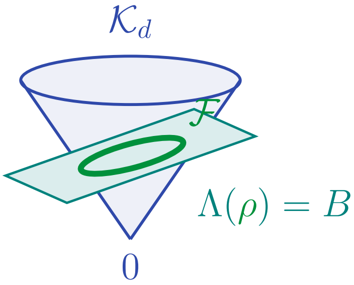


Every SDP feasible set is an affine slice

$\mathcal{K}_d = \{\rho \mid \rho \succeq 0\}$ PSD cone

$\mathcal{A}_B = \{\rho \mid \Lambda(\rho) = B\}$ affine subspace

$\mathcal{F} = \mathcal{K}_d \cap \mathcal{A}_B$ feasible set

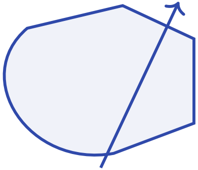


Problem: Does there exist ρ such that

$$\Lambda(\rho) = B, \qquad \rho \succeq 0$$

No objective function, but still an SDP (feasibility problem)

One value, possibly many maximisers

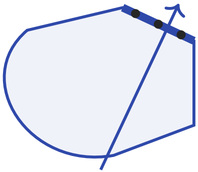


increasing $\mathrm{tr}(A\rho)$

The optimal value is one number

$$\alpha := \max_{\rho \in \mathcal{C}} \mathrm{tr}(A\rho)$$

One value, possibly many maximisers



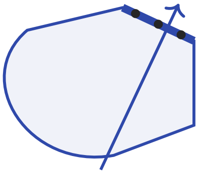
The optimal value is one number

$$\alpha := \max_{\rho \in \mathcal{C}} \mathrm{tr}(A\rho)$$

The set of maximisers may contain many points

$$\mathcal{O}^\star = \left\{\rho \in \mathcal{C} \mid \mathrm{tr}(A\rho) = \alpha\right\}$$

One value, possibly many maximisers



The optimal value is one number

$$\alpha := \max_{\rho \in \mathcal{C}} \mathrm{tr}(A\rho)$$

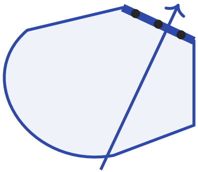
The set of maximisers may contain many points

$$\mathcal{O}^\star = \left\{\rho \in \mathcal{C} \mid \mathrm{tr}(A\rho) = \alpha\right\}$$

The highlighted edge is $\mathcal{O}^\star$

One optimal value, possibly many optimal points

One value, possibly many maximisers



increasing $\mathrm{tr}(A\rho)$

The optimal value is one number

$$\alpha := \max_{\rho \in \mathcal{C}} \mathrm{tr}(A\rho)$$

The set of maximisers may contain many points

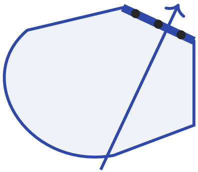
$$\mathcal{O}^\star = \left\{ \rho \in \mathcal{C} \mid \mathrm{tr}(A\rho) = \alpha \right\}$$

The highlighted edge is $\mathcal{O}^\star$

One optimal value, possibly many optimal points

The optimum is attainable at an extreme point

One value, possibly many maximisers



The optimal value is one number

$$\alpha := \max_{\rho \in \mathcal{C}} \mathrm{tr}(A\rho)$$

The set of maximisers may contain many points

$$\mathcal{O}^\star = \left\{\rho \in \mathcal{C} \mid \mathrm{tr}(A\rho) = \alpha\right\}$$

The highlighted edge is $\mathcal{O}^\star$

One optimal value, possibly many optimal points

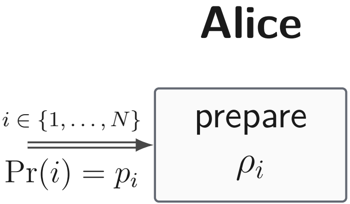
The optimum is attainable at an extreme point

Every local maximum is global — no spurious local optima

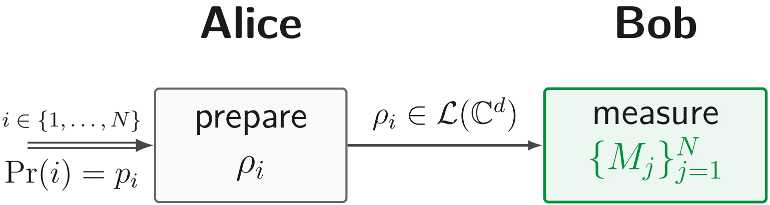
4. State discrimination

SDP in quantum theory

Can Bob identify the state Alice sent?

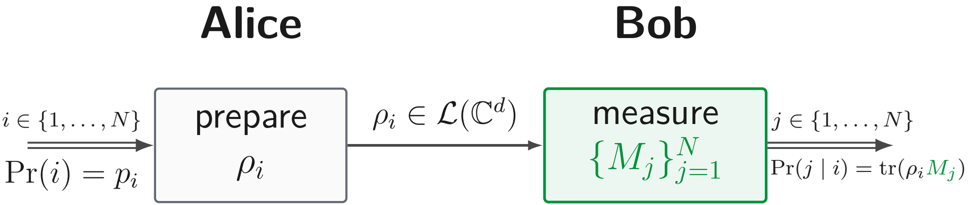


Can Bob identify the state Alice sent?



Bob knows the ensemble $\mathcal{E} = \{p_i, \rho_i\}_{i=1}^N$, but does not know i

Can Bob identify the state Alice sent?



Bob knows the ensemble $\mathcal{E} = \{p_i, \rho_i\}_{i=1}^N$, but does not know i

Bob's task is to guess i he wins if $j = i$

The probability of success

Given: $\{p_i, \rho_i\}_{i=1}^{N}$

The probability of success

Given: $\{p_i, \rho_i\}_{i=1}^{N}$

$$p_{\text{succ}}\left(\{M_j\}_{j=1}^{N}\right) = \sum_{i=1}^{N} p_i \sum_{j=1}^{N} \operatorname{tr}(\rho_i M_j)\, \delta_{ij}$$

$$= \sum_{i=1}^{N} p_i \operatorname{tr}(\rho_i M_i)$$

The probability of success

Given: $\{p_i, \rho_i\}_{i=1}^N$

$$p_{\text{succ}}\left(\{M_j\}_{j=1}^N\right) = \sum_{i=1}^N p_i \sum_{j=1}^N \operatorname{tr}(\rho_i M_j)\,\delta_{ij}$$

$$= \sum_{i=1}^N p_i \operatorname{tr}(\rho_i M_i)$$

$$p^\star_{\text{succ}} = \max_{\{M_i\}_{i=1}^N} \sum_{i=1}^N p_i \operatorname{tr}(\rho_i M_i)$$

The state-discrimination problem

Given: $\{p_i, \rho_i\}_{i=1}^{N}$

The state-discrimination problem

Given: $\{p_i, \rho_i\}_{i=1}^{N}$

maximise $\displaystyle\sum_{i=1}^{N} p_i \operatorname{tr}(\rho_i M_i)$

The state-discrimination problem

Given: $\{p_i, \rho_i\}_{i=1}^N$

$$\text{maximise} \quad \sum_{i=1}^N p_i \operatorname{tr}(\rho_i M_i)$$

$$\text{subject to} \quad \sum_{i=1}^N M_i = \mathbb{1}$$

$$M_i \succeq 0 \quad \forall i \in \{1, \ldots, N\}$$

Example: orthogonal states

Uniform prior $\quad p_1 = p_2 = \frac{1}{2}$

$$\rho_1 = |0\rangle\langle 0| \qquad \rho_2 = |1\rangle\langle 1| \qquad \langle 0 \mid 1 \rangle = 0$$

Example: orthogonal states

Uniform prior $\quad p_1 = p_2 = \frac{1}{2}$

$$\rho_1 = |0\rangle\langle 0| \qquad \rho_2 = |1\rangle\langle 1| \qquad \langle 0 \mid 1\rangle = 0$$

$$M_1 = |0\rangle\langle 0| \qquad M_2 = |1\rangle\langle 1|$$

Example: orthogonal states

Uniform prior $\qquad p_1 = p_2 = \frac{1}{2}$

$$\rho_1 = |0\rangle\langle 0| \qquad \rho_2 = |1\rangle\langle 1| \qquad \langle 0 \mid 1 \rangle = 0$$

$$M_1 = |0\rangle\langle 0| \qquad M_2 = |1\rangle\langle 1|$$

$$p_{\text{succ}} = \frac{1}{2}\operatorname{tr}(\rho_1 M_1) + \frac{1}{2}\operatorname{tr}(\rho_2 M_2) = 1$$

Perfect discrimination $\qquad p^{\star}_{\text{succ}} = 1$

Example: non-orthogonal states

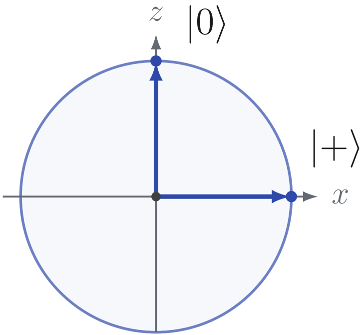
Uniform prior $\quad p_1 = p_2 = \frac{1}{2}$

$$\rho_1 = |0\rangle\langle 0| \qquad \rho_2 = |+\rangle\langle +|$$

$$|+\rangle = \frac{|0\rangle + |1\rangle}{\sqrt{2}}$$

Example: non-orthogonal states

$$\rho_1 = |0\rangle\langle 0| \qquad \rho_2 = |+\rangle\langle +|$$



Bloch circle in the x–z plane

Example: non-orthogonal states

Uniform prior $$p_1 = p_2 = \frac{1}{2}$$

$$\rho_1 = |0\rangle\langle 0| \qquad \rho_2 = |+\rangle\langle +|$$

$$|+\rangle = \frac{|0\rangle + |1\rangle}{\sqrt{2}}$$

$$|\langle 0 \,|\, +\rangle|^2 = \frac{1}{2}$$ not orthogonal

Perfect discrimination is impossible

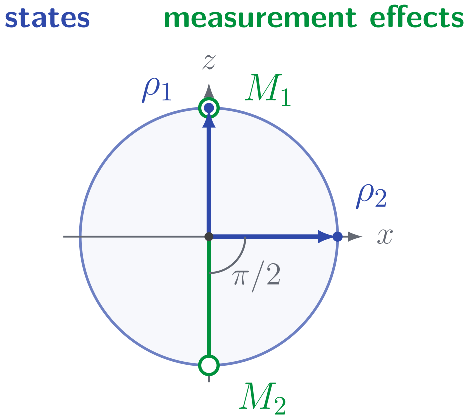
Example: non-orthogonal states

$$p_1 = p_2 = \frac{1}{2} \qquad \rho_1 = |0\rangle\langle 0| \qquad \rho_2 = |+\rangle\langle +|$$

Computational-basis measurement

$$M_1 = |0\rangle\langle 0| \qquad M_2 = |1\rangle\langle 1|$$

Example: non-orthogonal states



Example: non-orthogonal states

$$p_1 = p_2 = \frac{1}{2} \qquad \rho_1 = |0\rangle\langle 0| \qquad \rho_2 = |+\rangle\langle +|$$

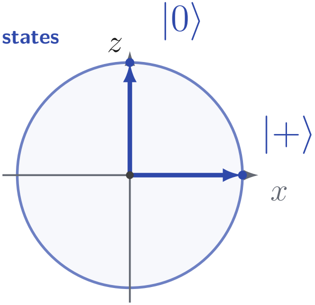
Computational-basis measurement

$$M_1 = |0\rangle\langle 0| \qquad M_2 = |1\rangle\langle 1|$$

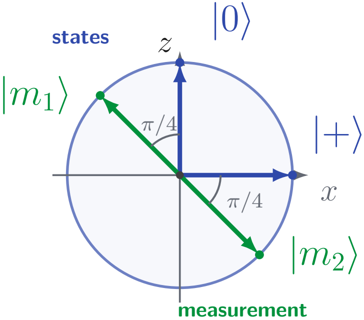
$$p_{\text{succ}} = \frac{1}{2}(1) + \frac{1}{2}\left(\frac{1}{2}\right) = \frac{3}{4} \qquad \Longrightarrow \qquad p^{\star}_{\text{succ}} \geq \frac{3}{4}$$

Can we do better?

How about an unbiased measurement?



How about an unbiased measurement?



$$|m_1\rangle := \cos\left(\frac{\pi}{8}\right)|0\rangle - \sin\left(\frac{\pi}{8}\right)|1\rangle = \frac{|0\rangle + |-\rangle}{\sqrt{2+\sqrt{2}}}$$

$$|m_2\rangle := \sin\left(\frac{\pi}{8}\right)|0\rangle + \cos\left(\frac{\pi}{8}\right)|1\rangle = \frac{|1\rangle + |+\rangle}{\sqrt{2+\sqrt{2}}}$$

How about an unbiased measurement?

$$M_1 := |m_1\rangle\langle m_1| \qquad M_2 := |m_2\rangle\langle m_2| \qquad M_1 + M_2 = \mathbb{1}$$

How about an unbiased measurement?

$$M_1 := |m_1\rangle\langle m_1| \qquad M_2 := |m_2\rangle\langle m_2| \qquad M_1 + M_2 = \mathbb{1}$$

$$p_{\text{succ}}(\{M_1, M_2\}) = \frac{1}{2}\,|\langle 0|m_1\rangle|^2 + \frac{1}{2}\,|\langle +|m_2\rangle|^2$$

$$= \cos^2\left(\frac{\pi}{8}\right) = \frac{2+\sqrt{2}}{4} \approx 0.8536 \quad \Longrightarrow \quad p^\star_{\text{succ}} \geq \frac{2+\sqrt{2}}{4} > \frac{3}{4}$$

How about an unbiased measurement?

$$M_1 := |m_1\rangle\langle m_1| \qquad M_2 := |m_2\rangle\langle m_2| \qquad M_1 + M_2 = \mathbb{1}$$

$$p_{\text{succ}}(\{M_1, M_2\}) = \frac{1}{2}\left|\langle 0|m_1\rangle\right|^2 + \frac{1}{2}\left|\langle +|m_2\rangle\right|^2$$

$$= \cos^2\left(\frac{\pi}{8}\right) = \frac{2+\sqrt{2}}{4} \approx 0.8536 \quad \Longrightarrow \quad p^{\star}_{\text{succ}} \geq \frac{2+\sqrt{2}}{4} > \frac{3}{4}$$

Can we do better?

How about an unbiased measurement?

$$M_1 := |m_1\rangle\langle m_1| \qquad M_2 := |m_2\rangle\langle m_2| \qquad M_1 + M_2 = \mathbb{1}$$

$$p_{\text{succ}}(\{M_1, M_2\}) = \frac{1}{2}\left|\langle 0|m_1\rangle\right|^2 + \frac{1}{2}\left|\langle +|m_2\rangle\right|^2$$

$$= \cos^2\left(\frac{\pi}{8}\right) = \frac{2+\sqrt{2}}{4} \approx 0.8536 \quad \Longrightarrow \quad p^\star_{\text{succ}} \geq \frac{2+\sqrt{2}}{4} > \frac{3}{4}$$

Can we do better? **How about using SDP?**

State discrimination as SDP

State discrimination

Given: $\{p_i, \rho_i\}_{i=1}^N$

$$p^\star_{\text{succ}} = \max_{\{M_i\}_{i=1}^N} \quad \sum_i p_i \operatorname{tr}(\rho_i M_i)$$

$$\text{subject to} \quad \sum_i M_i = \mathbb{1}$$

$$M_i \succeq 0 \quad \forall i$$

State discrimination as SDP

State discrimination

Given: $\{p_i, \rho_i\}_{i=1}^N$

$$
\begin{aligned}
p^\star_{\text{succ}} = \max_{\{M_i\}_{i=1}^N} \quad & \sum_i p_i \operatorname{tr}(\rho_i M_i) \\
\text{subject to} \quad & \sum_i M_i = \mathbb{1} \\
& M_i \succeq 0 \quad \forall i
\end{aligned}
$$

Standard form

Given: A, B, Λ

$$
\begin{aligned}
\alpha = \max_{\rho} \quad & \operatorname{tr}(A\rho) \\
\text{subject to} \quad & \Lambda(\rho) = B \\
& \rho \succeq 0
\end{aligned}
$$

State discrimination as SDP

State discrimination

Given: $\{p_i, \rho_i\}_{i=1}^N$

$$p^\star_{\text{succ}} = \max_{\{M_i\}_{i=1}^N} \quad \sum_i p_i \operatorname{tr}(\rho_i M_i)$$

$$\text{subject to} \quad \sum_i M_i = \mathbb{1}$$

$$M_i \succeq 0 \quad \forall i$$

Standard form

Given: A, B, Λ

$$\alpha = \max_{\rho} \quad \operatorname{tr}(A\rho)$$

$$\text{subject to} \quad \Lambda(\rho) = B$$

$$\rho \succeq 0$$

How to phrase state discrimination as SDP?

State discrimination as SDP

$$M \in \mathcal{L}(\mathbb{C}^d) \otimes \mathcal{L}(\mathbb{C}^N)$$

Work in the direct-sum matrix algebra

$$M := \sum_{i=1}^{N} M_i \otimes |i\rangle\langle i|_E = \bigoplus_{i=1}^{N} M_i$$

State discrimination as SDP

$$M \in \mathcal{L}(\mathbb{C}^d) \otimes \mathcal{L}(\mathbb{C}^N)$$

Work in the direct-sum matrix algebra

$$M := \sum_{i=1}^{N} M_i \otimes |i\rangle\langle i|_E = \bigoplus_{i=1}^{N} M_i$$

$$M \succeq 0 \quad \iff \quad M_i \succeq 0 \quad \forall i$$

State discrimination as SDP

Write $\{p_i, \rho_i\}_i$ in a single big matrix

$$P := \sum_{i=1}^{N} p_i \rho_i \otimes |i\rangle\langle i|_E = \bigoplus_{i=1}^{N} p_i \rho_i$$

State discrimination as SDP

$$P := \sum_{i=1}^{N} p_i \rho_i \otimes |i\rangle\langle i|_E = \bigoplus_{i=1}^{N} p_i \rho_i$$

$$M := \sum_{i=1}^{N} M_i \otimes |i\rangle\langle i|_E = \bigoplus_{i=1}^{N} M_i$$

State discrimination as SDP

$$P := \sum_{i=1}^{N} p_i \rho_i \otimes |i\rangle\langle i|_E = \bigoplus_{i=1}^{N} p_i \rho_i$$

$$M := \sum_{i=1}^{N} M_i \otimes |i\rangle\langle i|_E = \bigoplus_{i=1}^{N} M_i$$

Hence, $$\mathrm{tr}(PM) = \sum_{i=1}^{N} p_i \, \mathrm{tr}(\rho_i M_i)$$

State discrimination as SDP

$$\mathrm{tr}_E : \mathcal{L}(\mathbb{C}^d) \otimes \mathcal{L}(\mathbb{C}^N) \longrightarrow \mathcal{L}(\mathbb{C}^d)$$

State discrimination as SDP

$$\operatorname{tr}_E : \mathcal{L}(\mathbb{C}^d) \otimes \mathcal{L}(\mathbb{C}^N) \longrightarrow \mathcal{L}(\mathbb{C}^d)$$

$$\operatorname{tr}_E(X \otimes Y_E) := X \operatorname{tr}(Y)$$

State discrimination as SDP

$$\mathrm{tr}_E(M) = \mathrm{tr}_E\left(\sum_{i=1}^{N} M_i \otimes |i\rangle\langle i|_E\right)$$

State discrimination as SDP

$$\operatorname{tr}_E(M) = \operatorname{tr}_E\left(\sum_{i=1}^{N} M_i \otimes |i\rangle\langle i|_E\right)$$

$$= \sum_{i=1}^{N} M_i \operatorname{tr}(|i\rangle\langle i|_E)$$

State discrimination as SDP

$$\begin{aligned}
\mathrm{tr}_E(M) &= \mathrm{tr}_E\left(\sum_{i=1}^{N} M_i \otimes |i\rangle\langle i|_E\right) \\
&= \sum_{i=1}^{N} M_i \ \mathrm{tr}(|i\rangle\langle i|_E) \\
&= \sum_{i=1}^{N} M_i
\end{aligned}$$

State discrimination as SDP

Given: P, tr_E, $\mathbb{1}$

$$\begin{aligned} p^\star_{\mathrm{succ}} = \max_{M} \quad & \mathrm{tr}(PM) \\ \text{subject to} \quad & \mathrm{tr}_E(M) = \mathbb{1} \\ & M \succeq 0 \end{aligned}$$

State discrimination as SDP

State discrimination

Given: P, tr_E, $\mathbb{1}$

$$
\begin{aligned}
p^\star_{\mathrm{succ}} = \max_{M} \quad & \mathrm{tr}(PM) \\
\text{subject to} \quad & \mathrm{tr}_E(M) = \mathbb{1} \\
& M \succeq 0
\end{aligned}
$$

Standard form

Given: A, Λ, B

$$
\begin{aligned}
\alpha = \max_{\rho} \quad & \mathrm{tr}(A\rho) \\
\text{subject to} \quad & \Lambda(\rho) = B \\
& \rho \succeq 0
\end{aligned}
$$

State discrimination is an SDP

State discrimination as SDP

State discrimination

Given: P, tr_E, $\mathbb{1}$

$$\begin{aligned} p^\star_{\text{succ}} = \max_{M} \quad & \mathrm{tr}(PM) \\ \text{subject to} \quad & \mathrm{tr}_E(M) = \mathbb{1} \\ & M \succeq 0 \end{aligned}$$

Standard form

Given: A, Λ, B

$$\begin{aligned} \alpha = \max_{\rho} \quad & \mathrm{tr}(A\rho) \\ \text{subject to} \quad & \Lambda(\rho) = B \\ & \rho \succeq 0 \end{aligned}$$

State discrimination is an SDP

This technique is general.
The standard form covers any finite set of affine and PSD constraints

Linear objective functions

Riesz representation theorem

$$\ell : \mathcal{H} \longrightarrow \mathbb{F} \quad \Longrightarrow \quad \ell(x) = \langle a, x \rangle \qquad \mathbb{F} \in \{\mathbb{R}, \mathbb{C}\}$$

Every linear function on a Hilbert space is an inner product with a constant vector

Linear objective functions

Riesz representation theorem

$$\ell : \mathcal{H} \longrightarrow \mathbb{F} \quad \Longrightarrow \quad \ell(x) = \langle a, x \rangle \qquad \mathbb{F} \in \{\mathbb{R}, \mathbb{C}\}$$

Every linear function on a Hilbert space is an inner product with a constant vector

Matrices: Hilbert–Schmidt inner product

$$\langle A, B \rangle_{\mathrm{HS}} := \mathrm{tr}\big(A^\dagger B\big)$$

Linear objective functions

Riesz representation theorem

$$\ell : \mathcal{H} \longrightarrow \mathbb{F} \quad \Longrightarrow \quad \ell(x) = \langle a, x \rangle \qquad \mathbb{F} \in \{\mathbb{R}, \mathbb{C}\}$$

Every linear function on a Hilbert space is an inner product with a constant vector

Matrices: Hilbert–Schmidt inner product

$$\langle A, B \rangle_{\mathrm{HS}} := \operatorname{tr}\left(A^\dagger B\right)$$

$$A = A^\dagger, \quad \rho = \rho^\dagger \quad \Longrightarrow \quad \ell(\rho) = \langle A, \rho \rangle_{\mathrm{HS}} = \operatorname{tr}(A\rho) \in \mathbb{R}$$

$\operatorname{tr}(A\rho)$ represents every real linear objective on Hermitian primal variables

OK, state discrimination is an SDP,

but why does that help me?

5. Duality theory

Rigorous upper bounds

A feasible measurement gives a lower bound

PRIMAL – achievable

For every feasible POVM $\{M_i\}_{i=1}^{N}$

A feasible measurement gives a lower bound

PRIMAL – achievable

For every feasible POVM $\{M_i\}_{i=1}^{N}$

$$p_{\text{succ}}(\{M_i\}) := \sum_{i=1}^{N} p_i \operatorname{tr}(\rho_i M_i) \leq p^\star_{\text{succ}}$$

A feasible measurement gives a lower bound

PRIMAL – achievable

For every feasible POVM $\{M_i\}_{i=1}^N$

$$p_{\text{succ}}(\{M_i\}) := \sum_{i=1}^{N} p_i \operatorname{tr}(\rho_i M_i) \leq p_{\text{succ}}^\star$$

Feasible points correspond to attainable strategies

Achievability is only half of an optimality proof

$$\underbrace{p_{\text{succ}}(\{M_i\})}_{\text{achievable}} \quad \leq \quad p^\star_{\text{succ}} \quad \leq \quad \underbrace{\qquad\qquad}_{\text{universal upper bound?}}$$

How to obtain upper bounds?

Achievability is only half of an optimality proof

$$\underbrace{p_{\text{succ}}(\{M_i\})}_{\text{achievable}} \quad \leq \quad p^{\star}_{\text{succ}} \quad \leq \quad \underbrace{}_{\text{universal upper bound?}}$$

Duality theory!

Start again from the primal SDP

PRIMAL – achievable

Given: $A = A^\dagger \in \mathcal{L}(\mathbb{C}^d)$, $B = B^\dagger \in \mathcal{L}(\mathbb{C}^{d'})$

$\Lambda : \mathcal{L}(\mathbb{C}^d) \longrightarrow \mathcal{L}(\mathbb{C}^{d'})$ Hermiticity-preserving

Start again from the primal SDP

PRIMAL – achievable

Given: $A = A^\dagger \in \mathcal{L}(\mathbb{C}^d)$, $B = B^\dagger \in \mathcal{L}(\mathbb{C}^{d'})$

$\Lambda : \mathcal{L}(\mathbb{C}^d) \longrightarrow \mathcal{L}(\mathbb{C}^{d'})$ Hermiticity-preserving

$$\begin{aligned} \alpha = \sup_{\rho} \quad & \operatorname{tr}(A\rho) \\ \text{subject to} \quad & \Lambda(\rho) = B \\ & \rho \succeq 0 \end{aligned}$$

Each primal constraint gets a dual variable

PRIMAL – achievable

$$\alpha = \sup_{\rho} \quad \mathrm{tr}(A\rho)$$

$$\text{subject to} \quad \Lambda(\rho) = B$$

$$\rho \succeq 0$$

Each primal constraint gets a dual variable

PRIMAL – achievable

$$\alpha = \sup_{\rho} \quad \mathrm{tr}(A\rho)$$

$$\text{subject to} \quad \Lambda(\rho) = B \qquad \longleftrightarrow \qquad Y = Y^\dagger$$

$$\rho \succeq 0$$

Each primal constraint gets a dual variable

PRIMAL – achievable

$$
\begin{array}{rlcc}
\alpha = \sup\limits_{\rho} & \mathrm{tr}(A\rho) & & \\
\text{subject to} & \Lambda(\rho) = B & \longleftrightarrow & Y = Y^\dagger \\
 & \rho \succeq 0 & \longleftrightarrow & \sigma \succeq 0
\end{array}
$$

Each primal constraint gets a dual variable

PRIMAL – achievable

$$
\begin{array}{rlcl}
\alpha = \sup\limits_{\rho} & \mathrm{tr}(A\rho) & & \\
\text{subject to} & \Lambda(\rho) = B & \longleftrightarrow & Y = Y^\dagger \\
& \rho \succeq 0 & \longleftrightarrow & \sigma \succeq 0
\end{array}
$$

$$Y \in \mathcal{L}(\mathbb{C}^{d'}),\ \sigma \in \mathcal{L}(\mathbb{C}^{d})$$

Put everything into the Lagrangian

PRIMAL

$$\alpha = \sup_{\rho} \quad \mathrm{tr}(A\rho)$$

subject to $\quad \Lambda(\rho) = B$

$\quad \rho \succeq 0$

$$L(\rho, Y, \sigma) := \mathrm{tr}(A\rho)$$

objective

Put everything into the Lagrangian

PRIMAL

$$\alpha = \sup_{\rho} \quad \mathrm{tr}(A\rho)$$

subject to $\quad \Lambda(\rho) = B$

$\rho \succeq 0$

$$L(\rho, Y, \sigma) := \mathrm{tr}(A\rho) \qquad \text{objective}$$

$$+ \,\mathrm{tr}\Big[Y\Big(B - \Lambda(\rho)\Big)\Big] \qquad \text{equality constraint}$$

Put everything into the Lagrangian

PRIMAL

$$\alpha = \sup_{\rho} \quad \mathrm{tr}(A\rho)$$

$$\text{subject to} \quad \Lambda(\rho) = B$$

$$\rho \succeq 0$$

$$L(\rho, Y, \sigma) := \mathrm{tr}(A\rho) \qquad \text{objective}$$

$$+ \mathrm{tr}\Big[Y\Big(B - \Lambda(\rho)\Big)\Big] \qquad \text{equality constraint}$$

$$+ \mathrm{tr}(\sigma\rho) \qquad \text{positivity constraint}$$

Put everything into the Lagrangian

PRIMAL

$$\begin{aligned} \alpha = \sup_{\rho} \quad & \mathrm{tr}(A\rho) \\ \text{subject to} \quad & \Lambda(\rho) = B \\ & \rho \succeq 0 \end{aligned}$$

$$\begin{aligned} L(\rho, Y, \sigma) := \; & \mathrm{tr}(A\rho) && \text{objective} \\ & + \mathrm{tr}\Big[Y\Big(B - \Lambda(\rho)\Big)\Big] && \text{equality constraint} \\ & + \mathrm{tr}(\sigma\rho) && \text{positivity constraint} \end{aligned}$$

$$\rho \in \mathcal{F}, \quad \sigma \succeq 0 \quad \Longrightarrow \quad \mathrm{tr}(A\rho) \leq L(\rho, Y, \sigma)$$

Put everything into the Lagrangian

PRIMAL

$$\alpha = \sup_{\rho} \quad \mathrm{tr}(A\rho)$$

subject to $\Lambda(\rho) = B$

$\rho \succeq 0$

$$L(\rho, Y, \sigma) := \mathrm{tr}(A\rho)$$ objective

$$+\,\mathrm{tr}\Big[Y\Big(B - \Lambda(\rho)\Big)\Big]$$ equality constraint

$$+\,\mathrm{tr}(\sigma\rho)$$ positivity constraint

$$\rho \in \mathcal{F}, \quad \sigma \succeq 0 \quad \Longrightarrow \quad \mathrm{tr}(A\rho) \leq L(\rho, Y, \sigma)$$

$$\Longrightarrow \quad \alpha = \sup_{\rho \in \mathcal{F}} \mathrm{tr}(A\rho) \leq \sup_{\rho \in \mathcal{F}} L(\rho, Y, \sigma)$$

How about re-organising the Lagrangian?

The adjoint map $\Lambda^\dagger$ is defined by

$$\mathrm{tr}[Y\Lambda(\rho)] = \mathrm{tr}\left[\Lambda^\dagger(Y)\rho\right]$$

How about re-organising the Lagrangian?

The adjoint map $\Lambda^\dagger$ is defined by

$$\mathrm{tr}[Y\Lambda(\rho)] = \mathrm{tr}\left[\Lambda^\dagger(Y)\rho\right]$$

$$L(\rho, Y, \sigma) = \mathrm{tr}(BY) + \mathrm{tr}(A\rho) - \mathrm{tr}\left[\Lambda^\dagger(Y)\rho\right] + \mathrm{tr}(\sigma\rho)$$

How about re-organising the Lagrangian?

The adjoint map $\Lambda^\dagger$ is defined by

$$\mathrm{tr}[Y\Lambda(\rho)] = \mathrm{tr}\big[\Lambda^\dagger(Y)\rho\big]$$

$$\begin{aligned} L(\rho, Y, \sigma) &= \mathrm{tr}(BY) + \mathrm{tr}(A\rho) - \mathrm{tr}\big[\Lambda^\dagger(Y)\rho\big] + \mathrm{tr}(\sigma\rho) \\ &= \mathrm{tr}(BY) + \mathrm{tr}\Big[\Big(A - \Lambda^\dagger(Y) + \sigma\Big)\rho\Big] \end{aligned}$$

The dual problem

$$L(\rho, Y, \sigma) = \mathrm{tr}(BY) + \mathrm{tr}\big[\big(A - \Lambda^\dagger(Y) + \sigma\big)\rho\big]$$

The dual problem

$$L(\rho, Y, \sigma) = \mathrm{tr}(BY) + \mathrm{tr}\Big[\big(A - \Lambda^\dagger(Y) + \sigma\big)\rho\Big]$$

For L to be independent of ρ, set

$$A - \Lambda^\dagger(Y) + \sigma = 0$$

The dual problem

$$L(\rho, Y, \sigma) = \mathrm{tr}(BY) + \mathrm{tr}\big[\big(A - \Lambda^\dagger(Y) + \sigma\big)\rho\big]$$

For L to be independent of ρ, set

$$A - \Lambda^\dagger(Y) + \sigma = 0$$

For stronger arguments justifying this constraint, see the extra slides

The dual problem

$$L(\rho, Y, \sigma) = \mathrm{tr}(BY) + \mathrm{tr}\big[\big(A - \Lambda^\dagger(Y) + \sigma\big)\rho\big]$$

$$\begin{aligned} \text{best upper bound: minimise} \quad & L = \mathrm{tr}(BY) \\ \text{subject to} \quad & A - \Lambda^\dagger(Y) + \sigma = 0 \\ & \sigma \succeq 0 \end{aligned}$$

The dual problem

DUAL – upper bound

$$\begin{aligned} \beta = \inf \quad & \operatorname{tr}(BY) \\ \text{subject to} \quad & A - \Lambda^\dagger(Y) + \sigma = 0 \\ & \sigma \succeq 0 \\ & Y = Y^\dagger \end{aligned}$$

The slack variable σ is redundant

$$A - \Lambda^{\dagger}(Y) + \sigma = 0 \quad \Longleftrightarrow \quad \sigma = \Lambda^{\dagger}(Y) - A$$

The slack variable σ is redundant

$$A - \Lambda^\dagger(Y) + \sigma = 0 \quad \Longleftrightarrow \quad \sigma = \Lambda^\dagger(Y) - A$$

$$\sigma \succeq 0 \quad \Longleftrightarrow \quad \Lambda^\dagger(Y) \succeq A$$

The slack variable σ is redundant

DUAL – upper bound

$$\begin{aligned} \text{minimise} \quad & \operatorname{tr}(BY) \\ \text{subject to} \quad & A - \Lambda^\dagger(Y) + \sigma = 0 \\ & \sigma \succeq 0 \\ & Y = Y^\dagger \end{aligned} \quad \Longleftrightarrow \quad \begin{aligned} \text{minimise} \quad & \operatorname{tr}(BY) \\ \text{subject to} \quad & \Lambda^\dagger(Y) \succeq A \\ & Y = Y^\dagger \end{aligned}$$

The best upper bound is the dual SDP

PRIMAL – achievable

$$\begin{aligned} \alpha = \sup_{\rho} \quad & \operatorname{tr}(A\rho) \\ \text{subject to} \quad & \Lambda(\rho) = B \\ & \rho \succeq 0 \end{aligned}$$

DUAL – upper bound

The best upper bound is the dual SDP

PRIMAL – achievable

$$\begin{aligned} \alpha = \sup_{\rho} \quad & \mathrm{tr}(A\rho) \\ \text{subject to} \quad & \Lambda(\rho) = B \\ & \rho \succeq 0 \end{aligned}$$

DUAL – upper bound

$$\begin{aligned} \beta = \inf_{Y} \quad & \mathrm{tr}(BY) \\ \text{subject to} \quad & \Lambda^{\dagger}(Y) \succeq A \\ & Y = Y^{\dagger} \end{aligned}$$

The best upper bound is the dual SDP

PRIMAL – achievable

$$\begin{aligned} \alpha = \sup_{\rho} \quad & \mathrm{tr}(A\rho) \\ \text{subject to} \quad & \Lambda(\rho) = B \\ & \rho \succeq 0 \end{aligned}$$

DUAL – upper bound

$$\begin{aligned} \beta = \inf_{Y} \quad & \mathrm{tr}(BY) \\ \text{subject to} \quad & \Lambda^{\dagger}(Y) \succeq A \\ & Y = Y^{\dagger} \end{aligned}$$

$\alpha \leq \beta$ (weak duality)

Summary

Lecture 1 in brief

Quantum theory: **linearity and positivity**

Lecture 1 in brief

Quantum theory: **linearity and positivity**

SDP: linear objective + affine and PSD constraints

Lecture 1 in brief

Quantum theory: **linearity and positivity**

SDP: linear objective + affine and PSD constraints

State discrimination is an SDP

Lecture 1 in brief

Quantum theory: **linearity and positivity**

SDP: linear objective $+$ affine and PSD constraints

State discrimination is an SDP

Primal feasible points
attainable strategies

Dual feasible points
rigorous upper bounds

Lecture 1 in brief

Quantum theory: **linearity and positivity**

SDP: linear objective + affine and PSD constraints

State discrimination is an SDP

Primal feasible points
attainable strategies

Dual feasible points
rigorous upper bounds

see you in lecture 2!

Extra slides

Stronger arguments to derive the dual problem

A small recap

PRIMAL – achievable

$$\alpha := \sup_{\rho} \quad \mathrm{tr}(A\rho)$$

subject to $\Lambda(\rho) = B, \quad \rho \succeq 0$

DUAL – upper bound

$$\beta := \inf_{Y} \quad \mathrm{tr}(BY)$$

subject to $\Lambda^\dagger(Y) \succeq A, \quad Y = Y^\dagger$

A small recap

PRIMAL – achievable

$$\alpha := \sup_{\rho} \quad \mathrm{tr}(A\rho)$$
subject to $\Lambda(\rho) = B, \quad \rho \succeq 0$

DUAL – upper bound

$$\beta := \inf_{Y} \quad \mathrm{tr}(BY)$$
subject to $\Lambda^\dagger(Y) \succeq A, \quad Y = Y^\dagger$

Lagrangian

$$L(\rho, Y, \sigma) := \mathrm{tr}(A\rho) + \mathrm{tr}\Big[Y\Big(B - \Lambda(\rho)\Big)\Big] + \mathrm{tr}(\sigma\rho)$$

A small recap

PRIMAL – achievable

$$\alpha := \sup_{\rho} \quad \mathrm{tr}(A\rho)$$
subject to $\Lambda(\rho) = B, \quad \rho \succeq 0$

DUAL – upper bound

$$\beta := \inf_{Y} \quad \mathrm{tr}(BY)$$
subject to $\Lambda^\dagger(Y) \succeq A, \quad Y = Y^\dagger$

Lagrangian

$$L(\rho, Y, \sigma) := \mathrm{tr}(A\rho) + \mathrm{tr}\Big[Y\Big(B - \Lambda(\rho)\Big)\Big] + \mathrm{tr}(\sigma\rho)$$

$$L(\rho, Y, \sigma) = \mathrm{tr}(BY) + \mathrm{tr}\Big[\Big(A - \Lambda^\dagger(Y) + \sigma\Big)\rho\Big]$$

Now maximise over every Hermitian H

$$L(H, Y, \sigma) = \mathrm{tr}(BY) + \mathrm{tr}\Big[\Big(A - \Lambda^\dagger(Y) + \sigma\Big)H\Big]$$

Now maximise over every Hermitian H

$$L(H, Y, \sigma) = \operatorname{tr}(BY) + \operatorname{tr}\Big[\Big(A - \Lambda^\dagger(Y) + \sigma\Big)H\Big]$$

$$g(Y, \sigma) := \sup_{H = H^\dagger} L(H, Y, \sigma)$$

H ranges over all Hermitian matrices and need not be PSD

Now maximise over every Hermitian H

$$L(H, Y, \sigma) = \mathrm{tr}(BY) + \mathrm{tr}\big[\big(A - \Lambda^\dagger(Y) + \sigma\big)H\big]$$

$$g(Y, \sigma) := \sup_{H = H^\dagger} L(H, Y, \sigma)$$

H ranges over all Hermitian matrices and need not be PSD

$$\rho \in \mathcal{F}, \quad \sigma \succeq 0 \quad \Longrightarrow \quad \mathrm{tr}(A\rho) \leq L(\rho, Y, \sigma) \leq g(Y, \sigma)$$

Only the primal-feasible ρ must be PSD; the supremum may range over the larger Hermitian space

What do we want from the Lagrangian?

$$L(H, Y, \sigma) = \operatorname{tr}(BY) + \operatorname{tr}\big[\big(A - \Lambda^\dagger(Y) + \sigma\big)H\big]$$

What do we want from the Lagrangian?

$$L(H, Y, \sigma) = \mathrm{tr}(BY) + \mathrm{tr}\Big[\big(A - \Lambda^\dagger(Y) + \sigma\big)H\Big]$$

If ρ is primal feasible and $\sigma \succeq 0$, the Lagrangian is an upper bound

$$\mathrm{tr}(A\rho) \leq L(\rho, Y, \sigma) \leq g(Y, \sigma)$$

What do we want from the Lagrangian?

$$L(H, Y, \sigma) = \mathrm{tr}(BY) + \mathrm{tr}\big[\big(A - \Lambda^\dagger(Y) + \sigma\big)H\big]$$

If ρ is primal feasible and $\sigma \succeq 0$, the Lagrangian is an upper bound

$$\mathrm{tr}(A\rho) \leq L(\rho, Y, \sigma) \leq g(Y, \sigma)$$

We want the best upper bound: minimise the worst-case Lagrangian

$$g(Y, \sigma)$$

A finite bound cannot depend on H

$$C := A - \Lambda^\dagger(Y) + \sigma, \qquad C = C^\dagger$$

A finite bound cannot depend on H

$$C := A - \Lambda^\dagger(Y) + \sigma, \qquad C = C^\dagger$$

If $C \neq 0$, choose $H = tC$ with $t > 0$

This is allowed because H need only be Hermitian, not PSD

$$\operatorname{tr}(CH) = t\,\operatorname{tr}(C^2) \longrightarrow +\infty$$

A finite bound cannot depend on H

$$C := A - \Lambda^\dagger(Y) + \sigma, \qquad C = C^\dagger$$

If $C \neq 0$, choose $H = tC$ with $t > 0$

This is allowed because H need only be Hermitian, not PSD

$$\mathrm{tr}(CH) = t\,\mathrm{tr}(C^2) \longrightarrow +\infty$$

$g(Y, \sigma) < +\infty$ requires $A - \Lambda^\dagger(Y) + \sigma = 0$

A finite bound cannot depend on H

$$C := A - \Lambda^\dagger(Y) + \sigma, \qquad C = C^\dagger$$

If $C \neq 0$, choose $H = tC$ with $t > 0$

This is allowed because H need only be Hermitian, not PSD

$$\mathrm{tr}(CH) = t\,\mathrm{tr}(C^2) \longrightarrow +\infty$$

$g(Y, \sigma) < +\infty$ requires $A - \Lambda^\dagger(Y) + \sigma = 0$

Another argument: the dual problem cannot depend on primal variables