Quantum state discrimination and Holevo–Helstrom theorem

Ryuhei Mori

Tokyo Institute of Technology

Discrimination of probability distributions

Discrimination of probability distributions:

- p_0, p_1: Known probability distributions on a finite set $\mathcal{X}$.
- $\lambda \in [0, 1]$: A probability of choosing the probability distributions.

- Input: $x \in \mathcal{X}$ drawn from p_0 with probability λ, and from p_1 with probability $1 - \lambda$.

- Output: $i \in \{0, 1\}$ that indicates the probability distribution p_i.

Maximum probability of success: Discrimination of probability distributions

Theorem

The maximum probability of success is equal to

$$\frac{1 + \|\lambda p_0 - (1-\lambda) p_1\|_1}{2}$$

where $\|a\|_1 := \sum_x |a(x)|$.

Proof.

Let $d \colon \mathcal{X} \to \{0,1\}$ be a discriminator. Assume we get 1 point if succeeds, and loose 1 point if fails. Then, expected point is bounded by

$$\begin{aligned}
&\lambda \sum_{x \in \mathcal{X}} p_0(x)(-1)^{d(x)} + (1-\lambda) \sum_{x \in \mathcal{X}} p_1(x)(-1)^{d(x)+1} \\
&= \sum_{x \in \mathcal{X}} \left(\lambda p_0(x) - (1-\lambda) p_1(x)\right)(-1)^{d(x)} \le \|\lambda p_0 - (1-\lambda) p_1\|_1 .
\end{aligned}$$

The equality is achieved by $d(x)$ that is 0 iff $\lambda p_0(x) \geq (1-\lambda) p_1(x)$. On the other hand, the expected point is $p_{\mathrm{succ}} - p_{\mathrm{fail}} = 2 p_{\mathrm{succ}} - 1$. □

Discrimination of quantum states

Discrimination of quantum states:

- ρ_0, ρ_1: Known quantum states.
- $\lambda \in [0, 1]$: A probability of choosing the quantum states.

- Input: A quantum state ρ_0 is given with probability λ, and ρ_1 is given with probability $1 - \lambda$.

- Output: $i \in \{0, 1\}$ that indicates the given state ρ_i.

Maximum probability of success: Discrimination of quantum states

Theorem (Holevo–Helstrom theorem)

The maximum probability of success is equal to

$$\frac{1 + \|\lambda\rho_0 - (1-\lambda)\rho_1\|_1}{2}$$

where $\|A\|_1 := \mathrm{Tr}(\sqrt{A^\dagger A})$, *which is a sum of the singular values of* A*.*

Proof.

Once a measurement $(P_y)_{y\in\mathcal{Y}}$ is fixed, we get a classical probability distribution $p_0(y) := \mathrm{Tr}(\rho_0 P_y)$ and $p_1(y) := \mathrm{Tr}(\rho_1 P_y)$. In this case, the maximum probability of success is given by

$$\frac{1 + \|\lambda p_0 - (1-\lambda)p_1\|_1}{2}$$

Hence, it's sufficient to show

$$\max_{(P_y)_{y\in\mathcal{Y}}} \|\lambda p_0 - (1-\lambda)p_1\|_1 = \|\lambda\rho_0 - (1-\lambda)\rho_1\|_1.$$

Holevo–Helstrom theorem

$$\max_{(P_y)_{y\in\mathcal{Y}}} \|\lambda p_0 - (1-\lambda)p_1\|_1 = \max_{(P_y)_{y\in\mathcal{Y}}} \sum_{y\in\mathcal{Y}} |\lambda p_0(y) - (1-\lambda)p_1(y)|$$

$$= \max_{(P_y)_{y\in\mathcal{Y}}} \sum_{y\in\mathcal{Y}} |\lambda \mathrm{Tr}(\rho_0 P_y) - (1-\lambda)\mathrm{Tr}(\rho_1 P_y)|$$

$$= \max_{(P_y)_{y\in\mathcal{Y}}} \sum_{y\in\mathcal{Y}} |\mathrm{Tr}\left((\lambda\rho_0 - (1-\lambda)\rho_1)P_y\right)|$$

Let

$$\lambda\rho_0 - (1-\lambda)\rho_1 = \sum_{x\in\mathcal{X}} \mu_x |\psi_x\rangle\langle\psi_x|$$

be a spectral decomposition. Then,

$$\max_{(P_y)_{y\in\mathcal{Y}}} \sum_{y\in\mathcal{Y}} |\mathrm{Tr}\left((\lambda\rho_0 - (1-\lambda)\rho_1)P_y\right)| = \max_{(P_y)_{y\in\mathcal{Y}}} \sum_{y\in\mathcal{Y}} \left|\sum_{x\in\mathcal{X}} \mu_x \langle\psi_x| P_y |\psi_x\rangle\right|$$

$$\leq \max_{(P_y)_{y\in\mathcal{Y}}} \sum_{y\in\mathcal{Y}} \sum_{x\in\mathcal{X}} |\mu_x| \langle\psi_x| P_y |\psi_x\rangle = \sum_{x\in\mathcal{X}} |\mu_x| = \|\lambda\rho_0 - (1-\lambda)\rho_1\|_1.$$

The maximum is achieved by $\mathcal{Y} = \mathcal{X}$ and $P_x = |\psi_x\rangle\langle\psi_x|$,
and $\mathcal{Y} = \{0,1\}$ and $P_0 = \sum_{x:\mu_x\geq 0} |\psi_x\rangle\langle\psi_x|$, $P_1 = \sum_{x:\mu_x<0} |\psi_x\rangle\langle\psi_x|$.

Discrimination of pure states

The maximum probability of success for discrimination of $|\psi\rangle\langle\psi|$ and $|\varphi\rangle\langle\varphi|$ is given by

$$\left\|\lambda |\psi\rangle\langle\psi| - (1-\lambda)|\varphi\rangle\langle\varphi|\right\|_1$$

Let $A := \lambda |\psi\rangle\langle\psi| - (1-\lambda)|\varphi\rangle\langle\varphi|$. Then, A has rank at most two. From

$$\begin{aligned}
\mathrm{Tr}(A) &= \mu_0 + \mu_1 = \lambda - (1-\lambda) \\
\mathrm{Tr}(A^2) &= \mu_0^2 + \mu_1^2 = \lambda^2 + (1-\lambda)^2 - 2\lambda(1-\lambda)|\langle\psi|\varphi\rangle|^2,
\end{aligned}$$

$$\begin{aligned}
2\mu_0\mu_1 &= (\mu_0+\mu_1)^2 - (\mu_0^2+\mu_1^2) = -2\lambda(1-\lambda)(1-|\langle\psi|\varphi\rangle|^2) \le 0. \\
(\mu_0-\mu_1)^2 &= (\mu_0^2+\mu_1^2) - 2\mu_0\mu_1 = (\lambda+(1-\lambda))^2 - 4\lambda(1-\lambda)|\langle\psi|\varphi\rangle|^2 \\
\|A\|_1 &= |\mu_0| + |\mu_1| = |\mu_0 - \mu_1| = \sqrt{1-4\lambda(1-\lambda)|\langle\psi|\varphi\rangle|^2}.
\end{aligned}$$

Trace norm

The trace norm is the sum of the singular values. Hence, it satisfies the unitary invariance

$$\|UAV\|_1 = \|A\|_1$$

for any unitary matrices U and V.
For

$$A = \sum_j \lambda_j |\psi_j\rangle \langle\varphi_j| \qquad \text{(singular decomposition)}$$

$$U = \sum_j |\tau_j\rangle \langle\psi_j| \qquad \text{(singular decomposition)}$$

$$\begin{aligned} |\mathrm{Tr}(UA)| &= \left| \mathrm{Tr}\left(\sum_j \lambda_j |\tau_j\rangle \langle\varphi_j| \right) \right| \\ &\le \sum_j \lambda_j \left|\mathrm{Tr}\left(|\tau_j\rangle \langle\varphi_j|\right)\right| = \sum_j \lambda_j |\langle\varphi_j|\tau_j\rangle| \le \sum_j \lambda_j = \|A\|_1 \end{aligned}$$

By setting $|\tau_j\rangle = |\varphi_j\rangle$, the equalities are satisfied. We obtain $\max_U |\mathrm{Tr}(UA)| = \|A\|_1$. Hence, the trace norm satisfies the triangle inequality.

$$\begin{aligned} \|A+B\|_1 &= \max_U |\mathrm{Tr}(U(A+B))| \le \max_U(|\mathrm{Tr}(UA)| + |\mathrm{Tr}(UB)|) \\ &\le \max_U |\mathrm{Tr}(UA)| + \max_U |\mathrm{Tr}(UB)| = \|A\|_1 + \|B\|_1 \end{aligned}$$

Unitary discrimination

Discrimination of unitary operators:

- Input: A unitary operator U_0 is given with probability λ, and U_1 is given with probability $1 - \lambda$ as an oracle $\mathcal{O}$ that can be used once.
- Output: $i \in \{0, 1\}$ that indicates the given unitary U_i.

Unitary discrimination

If algorithm call the oracle $\mathcal{O}$ on a state $|\psi\rangle$, we get either of $U_0 |\psi\rangle$ or $U_1 |\psi\rangle$.
Then, the maximum probability of success is given by

$$\max_{|\psi\rangle} \|\lambda U_0 |\psi\rangle \langle\psi| U_0^\dagger - (1-\lambda) U_1 |\psi\rangle \langle\psi| U_1^\dagger\|_1$$
$$= \max_{|\psi\rangle} \sqrt{1 - 4\lambda(1-\lambda)|\langle\psi| U_0^\dagger U_1 |\psi\rangle|^2}$$

For $V := U_0^\dagger U_1$, let $V = \sum_x \mu_x |\varphi_x\rangle \langle\varphi_x|$ be a spectral decomposition. Note that $|\mu_x| = 1$ for all x. Let $|\psi\rangle = \sum_x \alpha_x |\varphi_x\rangle$. Then, $\langle\psi| U_0^\dagger U_1 |\psi\rangle = \sum_x |\alpha_x|^2 \mu_x$, which is a convex combination of $(\mu_x)_x$. Let θ_{cover} be the smallest angle that covers all eigenvalues $(\mu_x)_x$. Then, $\min_{|\psi\rangle} |\langle\psi| U_0^\dagger U_1 |\psi\rangle| = \cos(\frac{\theta_{\text{cover}}}{2})$ if $\theta_{\text{cover}} \leq \pi$, and 0 otherwise.

The diamond distance

The diamond distance D is defined by

$$D(A, B) := \max_{|\psi\rangle} \|A|\psi\rangle\langle\psi|A^\dagger - B|\psi\rangle\langle\psi|B^\dagger\|_1.$$

Then, D is unitary invariant, i.e., $D(UAV, UBV) = D(A, B)$ for any unitary matrices U and V.

$$\begin{aligned}
D(A, B) &= \max_{|\psi\rangle} \|A|\psi\rangle\langle\psi|A^\dagger - B|\psi\rangle\langle\psi|B^\dagger\|_1 \\
&= \max_{|\psi\rangle} \|A|\psi\rangle\langle\psi|A^\dagger - C|\psi\rangle\langle\psi|C^\dagger + C|\psi\rangle\langle\psi|C^\dagger - B|\psi\rangle\langle\psi|B^\dagger\|_1 \\
&\le \max_{|\psi\rangle}(\|A|\psi\rangle\langle\psi|A^\dagger - C|\psi\rangle\langle\psi|C^\dagger\|_1 + \|C|\psi\rangle\langle\psi|C^\dagger - B|\psi\rangle\langle\psi|B^\dagger\|_1) \\
&\le \max_{|\psi\rangle} \|A|\psi\rangle\langle\psi|A^\dagger - C|\psi\rangle\langle\psi|C^\dagger\|_1 + \max_{|\psi\rangle} \|C|\psi\rangle\langle\psi|C^\dagger - B|\psi\rangle\langle\psi|B^\dagger\|_1 \\
&= D(A, C) + D(C, B)
\end{aligned}$$

The diamond distance D satisfies the triangle inequality.

Assignment

1. Show the maximum probability of success for discriminating $|0\rangle$ and $|+\rangle$ given with the uniform probability.

2. Show a binary optimal measurement for the discrimination of $|0\rangle$ and $|+\rangle$ given with the uniform probability.

3. Show the maximum probability of success for discriminating I and $R_Z(\theta)$ given with the uniform probability. Show the input state $|\psi\rangle$ for the oracle as well.

4. [Advanced] Show the maximum probability of success for discriminating $R_{\mathbf{v}}(\theta)$ and $R_{\mathbf{w}}(\eta)$ given with the uniform probability.