Lecture 11 Trace distance and fidelity

Xiao Yuan

November 15, 2021

In this lecture, we study two quantities that measures the difference or similarity of two quantum states.

1 Trace distance

1.1 Classical case

Consider two probability distributions $\mathbf{p}$ and $\mathbf{q}$, their difference could be quantified by trace distance

$$D(\mathbf{p},\mathbf{q}) = \frac{1}{2}\sum_j |p_j - q_j|. \tag{1}$$

Two important properties of trace distance are

- Nonnegativity: $D(\mathbf{p},\mathbf{q}) \geq 0$ with equal sign achieved iff $\mathbf{p} = \mathbf{q}$.
- Triangle inequality: $D(\mathbf{p}_1,\mathbf{p}_2) + D(\mathbf{p}_2,\mathbf{p}_3) \geq D(\mathbf{p}_1,\mathbf{p}_3)$.

The operational meaning of trace distance is in probability distinguishability. Suppose we aim to distinguish between an event P with probability $\mathbf{p}$ and another one Q with probability $\mathbf{q}$. A most general strategy is that when j happens, we judge it as P and Q with probability $f(j)$ and $1-f(j)$, with $f(j) \in [0,1]$. Then the success probabilities are $\sum_j p(j)f(j)$ and $\sum_j q(j)(1-f(j))$ when the event is P and Q, respectively. Now, suppose P and Q are uniformly produced, then the probability that we can distinguish P and Q is

$$P_{\text{succ}} = \frac{1}{2}\sum_j p(j)f(j) + q(j)(1-f(j)) = \frac{1+\sum_j f(j)(p(j)-q(j))}{2}. \tag{2}$$

We can easily maximize P_{succ} by choosing $f(j)$ to be 1 if $p(j)-q(j) \geq 0$ and 0 otherwise. Then we have

$$P_{\text{succ}} = \frac{1+\sum_{j:p(j)\geq q(j)} |p(j)-q(j)|}{2}. \tag{3}$$

Note that $0 = \sum_j p(j) - q(j) = \sum_{j:p(j)\geq q(j)} |p(j)-q(j)| - \sum_{j:p(j)<q(j)} |p(j)-q(j)|$, thus $\sum_{j:p(j)\geq q(j)} |p(j)-q(j)| = D(\mathbf{p},\mathbf{q})$ and

$$P_{\text{succ}} = \frac{1+D(\mathbf{p},\mathbf{q})}{2}. \tag{4}$$

When $\mathbf{p}$ and $\mathbf{q}$ are maximally different, we have $D(\mathbf{p},\mathbf{q}) = 1$ and hence they are maximally distinguishable with $P_{\text{succ}} = 1$. Otherwise, we have $D(\mathbf{p},\mathbf{q}) = 0$ and we cannot do a better job than blind guess.

1.2 Quantum case

A natural question is to generalize the definition of trace distance to quantum states. Consider two quantum states ρ and σ, suppose $\rho = \sum_j p_j |j\rangle\langle j|$ and $\sigma = \sum_j q_j |j\rangle\langle j|$, we can similarly define their distance as Eq. (1). However, when they have a different basis (under spectral decomposition), we need to change the definition.

1.2.1 Trance norm

The key idea is to use the *trace norm* of matrices. For matrix A with a singular value decomposition $A = \sum_j \lambda_j |s_j\rangle \langle v_j|$, the trace norm of A is

$$\|A\|_1 = \mathrm{Tr}[|A|] = \mathrm{Tr}[\sqrt{AA^\dagger}] = \sum_j |\lambda_j|. \tag{5}$$

Several important properties of the trace norm are

- Non-negativity: $\|A\|_1 \geq 0$ with equal sign achieved iff $A = 0$.
- Triangle inequality: $\|A\|_1 + \|B\|_1 \geq \|A + B\|_1$.
- Isometric invariance: $\|UAV\|_1 = \|A\|_1$ for isometry U and V[1].
- Convexity: $\|\lambda A + (1-\lambda)B\|_1 \leq \lambda \|A\|_1 + (1-\lambda)\|B\|_1$.
- Variational characterization 1: $\|A\|_1 = \max_{U:\text{unitary}} \mathrm{Tr}[UA]$.
- Variational characterization 2: $\|A\|_1 = \max_{-\mathbb{I} \leq P \leq \mathbb{I}} \mathrm{Tr}[PA]$ for hermitian A.

1.2.2 Trance distance

The trace distance between two quantum states is defined as

$$D(\rho, \sigma) = \frac{1}{2}\|\rho - \sigma\|_1. \tag{6}$$

Trace distance satisfy the following properties.

- Symmetric: $D(\rho, \sigma) = D(\sigma, \rho)$.
- Nonnegativity: $D(\rho, \sigma) \geq 0$ with equal sign achieved iff $\rho = \sigma$.
- Triangle inequality: $D(\rho, \sigma) + D(\sigma, \gamma) \geq D(\rho, \gamma)$.
- Unitary invariance: $D(\rho, \sigma) = D(U\rho U^\dagger, U\sigma U^\dagger)$.
- Tensor product property: $D(\rho \otimes \gamma, \sigma \otimes \gamma) = D(\rho, \sigma)$.
- Variational form: $D(\rho, \sigma) = \max_{0 \leq P \leq \mathbb{I}} \mathrm{Tr}[P(\rho - \sigma)]$.

Proof. Denote $A = \rho - \sigma$. Consider a spectral decomposition of A as $A = \sum_j \lambda_j |\psi_j\rangle \langle \psi_j|$. Denote

$$A_+ = \sum_{j:\lambda_j \geq 0} \lambda_j |\psi_j\rangle \langle \psi_j|, \; A_- = \sum_{j:\lambda_j \leq 0} \lambda_j |\psi_j\rangle \langle \psi_j|, \tag{7}$$

then $A = A_+ + A_-$ and $|A| = A_+ - A_-$. Note that $0 = \mathrm{Tr}[A] = \mathrm{Tr}[A_+] + \mathrm{Tr}[A_-]$, thus $D(\rho, \sigma) = \mathrm{Tr}[A_+] = -\mathrm{Tr}[A_-]$.

Then we have

$$\mathrm{Tr}[P(\rho - \sigma)] = \mathrm{Tr}[PA] = \mathrm{Tr}[PA_+] + \mathrm{Tr}[PA_-] \leq \mathrm{Tr}[PA_+] \leq \mathrm{Tr}[A_+] = D(\rho, \sigma). \tag{8}$$

We can further choose $P = \sum_{j:\lambda_j \geq 0} |\psi_j\rangle \langle \psi_j|$ to achieve the equal sign. □

[1] A matrix U is called an isometry when $U^\dagger U = \mathbb{I}$

- State distinguishability: trace distance $D(\rho,\sigma)$ measures the probability that we can distinguish between ρ and σ. Specifically, suppose we have a POVM $\{\Lambda, \mathbb{I}-\Lambda\}$, and the measurement corresponds to ρ if Λ happens and σ other wise. Then the probability that the strategy succeeds given uniformly random ρ and σ is

$$P_{\text{succ}} = \frac{1}{2}(\text{Tr}[\Lambda\rho] + \text{Tr}[(\mathbb{I}-\Lambda)\sigma]) = \frac{1}{2}(1 + \text{Tr}[\Lambda(\rho-\sigma)]). \tag{9}$$

Maximizing P_{succ} over all possible $\Lambda \in [0,\mathbb{I}]$ we have the maximal success probability $\max P_{\text{succ}} = \frac{1+D(\rho,\sigma)}{2}$.

- Monotonicity under quantum channels: $D(\rho,\sigma) \leq D(\mathcal{E}(\rho),\mathcal{E}(\sigma))$.

Proof. One can check Nielsen's book for a proof based on the above decomposition. Here we provide an alternative proof based on the definition of quantum channels. Note that any channel has the form $\mathcal{E}(\rho_A) = \text{Tr}_E[U(\rho_A \otimes |0\rangle\langle 0|_E)U^\dagger]$. We first note the monotonicity under partial trace, i.e.,

$$D(\rho_A,\sigma_A) \leq D(\rho_{AE},\sigma_{AE}), \tag{10}$$

with $\rho_A = \text{Tr}_E[\rho_{AE}]$ and $\sigma_A = \text{Tr}_E[\sigma_{AE}]$. We can prove it by using the variational form of trace distance as

$$\begin{aligned} D(\rho_{AE},\sigma_{AE}) &= \max_{0\leq P_{AE}\leq \mathbb{I}_{AE}} \text{Tr}[P_{AE}(\rho_{AE}-\sigma_{AE})], \\ &\geq \max_{0\leq P_A\leq \mathbb{I}_A} \text{Tr}[P_A\otimes \mathbb{I}_E(\rho_{AE}-\sigma_{AE})], \\ &= \max_{0\leq P_A\leq \mathbb{I}_A} \text{Tr}[P_A(\rho_A-\sigma_A)], \\ &= D(\rho_A,\sigma_A). \end{aligned} \tag{11}$$

Now we prove the monotonicity of trace distance under quantum channels

$$\begin{aligned} D(\rho_A,\sigma_A) &= D(\rho_A\otimes|0\rangle\langle 0|_E, \sigma\otimes|0\rangle\langle 0|_E), \\ &= D(U(\rho_A\otimes|0\rangle\langle 0|_E)U^\dagger, U(\sigma\otimes|0\rangle\langle 0|_E)U^\dagger), \\ &\geq D(\text{Tr}_E[U(\rho_A\otimes|0\rangle\langle 0|_E)U^\dagger], \text{Tr}_E[U(\sigma\otimes|0\rangle\langle 0|_E)U^\dagger]), \\ &= D(\mathcal{E}(\rho),\mathcal{E}(\sigma)). \end{aligned} \tag{12}$$

□

The meaning of monotonicity is that when we do the same operation on two quantum states, we cannot make them more different.

- Strong convexity: $D\big(\sum_j p_j\rho_j, \sum_j q_j\sigma_j\big) \leq D(\mathbf{p},\mathbf{q}) + \sum_j p_j D(\rho_j,\sigma_j)$.

Proof. We first apply the triangle inequality to have

$$\begin{aligned} D\big(\sum_j p_j\rho_j, \sum_j q_j\sigma_j\big) &\leq D\big(\sum_j p_j\sigma_j, \sum_j q_j\sigma_j\big) + D\big(\sum_j p_j\rho_j, \sum_j p_j\sigma_j\big), \\ &\leq \sum_j D\big(p_j\sigma_j, q_j\sigma_j\big) + \sum_j D\big(p_j\rho_j, p_j\sigma_j\big), \\ &\leq D(\mathbf{p},\mathbf{q}) + \sum_j p_j D(\rho_j,\sigma_j) \end{aligned} \tag{13}$$

Here the second lines uses the triangle inequality of the trace norm. □

- Reduction to classical trace distance. We can measure the quantum state to convert it to a classical distribution. Consider a POVM $\{E_j\}$, with $p_j = \mathrm{Tr}[\rho E_j]$ and $q_j = \mathrm{Tr}[\sigma E_j]$, then

$$D(\rho, \sigma) = \max_{\{E_j\}} D(\mathbf{p}, \mathbf{q}). \tag{14}$$

  We refer to Nielsen's book for one proof, here we give an alternative proof. Note that POVM could be understood as a quantum channel as

$$\mathcal{M}(\cdot) = \sum_j \mathrm{Tr}[E_j \cdot] |j\rangle \langle j| . \tag{15}$$

  Thus $D(\mathbf{p}, \mathbf{q}) = D(\mathcal{M}(\rho), \mathcal{N}(\sigma))$ and hence $D(\mathbf{p}, \mathbf{q}) \leq D(\rho, \sigma)$ due to the monotonicity property. We can achieve the equal sign with the eigenbasis of $\rho - \sigma$.

2 Fidelity

2.1 Classical case

Again consider two probability distributions $\mathbf{p}$ and $\mathbf{q}$, we can quantify their similarity as

$$F(\mathbf{p}, \mathbf{q}) = \sum_j \sqrt{p_j q_j}. \tag{16}$$

Fidelity F or infidelity $1 - F$ is not a distance measure, but it does quantify how close or different the two distributions are. In particular, we have $F = 1$ if $\mathbf{p} = \mathbf{q}$ and $F = 0$ if $\mathbf{p}$ and $\mathbf{q}$ are maximally distinguishable (prove it).

Trace distance and fidelity are also related. On the one hand,

$$\begin{aligned} D(\mathbf{p}, \mathbf{q}) &= \frac{1}{2} \sum_j |p_j - q_j| \\ &= \frac{1}{2} \sum_j |\sqrt{p_j} - \sqrt{q_j}||\sqrt{p_j} + \sqrt{q_j}| \\ &\geq \frac{1}{2} \sum_j |\sqrt{p_j} - \sqrt{q_j}|^2 \\ &= \frac{1}{2} \sum_j p_j + q_j - 2\sqrt{p_j q_j} \\ &= 1 - F(\mathbf{p}, \mathbf{q}). \end{aligned} \tag{17}$$

Therefore, whenever the fidelity $F(\mathbf{p}, \mathbf{q})$ is close to 1, it also means that their trace distance $D(\mathbf{p}, \mathbf{q})$ is close to 0.

There is another important relation

$$D(\mathbf{p}, \mathbf{q}) \leq \sqrt{1 - F(\mathbf{p}, \mathbf{q})^2}. \tag{18}$$

We leave its proof after we introduce the quantum fidelity.

2.2 Quantum case

Special pure states Consider two special pure states $|\psi\rangle = \sum_j \sqrt{p_j} |j\rangle$ and $|\phi\rangle = \sum_j \sqrt{q_j} |j\rangle$, their fidelity is defined as

$$F(\psi, \phi) = | \langle \psi | \psi \rangle | = \sum_j \sqrt{p_j q_j}. \tag{19}$$

Special mixed states Consider two special mixed states $\rho = \sum_j p_j |j\rangle\langle j|$ and $\sigma = \sum_j q_j |j\rangle\langle j|$, their fidelity is defined as

$$F(\rho,\sigma) = \mathrm{Tr}[\sqrt{\rho}\sqrt{\sigma}] = \sum_j \sqrt{p_j q_j}. \tag{20}$$

One may then suggest to define the fidelity of two general quantum states as $F(\rho,\sigma) = \mathrm{Tr}[\sqrt{\rho}\sqrt{\sigma}]$. However, this may fail since $\sqrt{\rho}\sqrt{\sigma}$ is no more hermitian and the trace may give complex numbers. In order to have a real valued fidelity, we thus define it as

$$F(\rho,\sigma) = \mathrm{Tr}[|\sqrt{\rho}\sqrt{\sigma}|] = \mathrm{Tr}\left[\sqrt{\sqrt{\rho}\sigma\sqrt{\rho}}\right]. \tag{21}$$

We examine the definition from the following aspects.

- The definition is consistent with the special pure and mixed state cases. Furthermore, when either ρ or σ is a pure state, say $\rho = \psi$, we have

$$F(\psi,\sigma) = \sqrt{\langle\psi|\sigma|\psi\rangle}. \tag{22}$$

  Therefore, the fidelity between ψ and σ could be understood as the expectation value of measuring observable σ of state ψ.

- The fidelity is also symmetric with $F(\rho,\sigma) = F(\sigma,\rho)$.

- $F(\rho,\sigma) \in [0,1]$ with $F(\rho,\sigma) = 1$ iff $\rho = \sigma$.

- Unitary invariance $F(\rho,\sigma) = F(U\rho U^\dagger, U\sigma U^\dagger)$.

- Multiplicativity: $F(\rho_1 \otimes \rho_2, \sigma_1 \otimes \sigma_2) = F(\rho_1,\sigma_1)F(\rho_2,\sigma_2)$.

- Triangle inequality. Although the quantum fidelity F violates the triangle inequality, we can define $A = \arccos F$ representing the angle between the two states. Then we have

$$A(\rho,\sigma) + A(\sigma,\gamma) \geq A(\rho,\gamma). \tag{23}$$

  The proof assumes uses the following Uhlmann's Theorem and the triangle inequality of angles (see Nielsen's book for the proof).

- *Uhlmann's Theorem*:

$$F(\rho,\sigma) = \max_{\psi,\phi} F(\psi,\phi), \tag{24}$$

  where ψ and ϕ are purifications (using the same ancillary system) of ρ and σ respectively.

  Proof. Denote the system of ρ and σ as A and the ancillary system as E. Consider the maximally entangled state $|\Phi^+\rangle = \sum_j |jj\rangle$, an explicit purification of ρ or σ is $|\psi\rangle_{AE} = \sqrt{\rho_A} \otimes \mathbb{I}_E |\Phi^+\rangle_{AE}$ or $|\phi\rangle_{AE} = \sqrt{\sigma_A} \otimes \mathbb{I}_E |\Phi^+\rangle_{AE}$. Then an arbitrary purification of ρ or σ is $U_E|\psi\rangle_{AE}$ or $V_E|\phi\rangle_{AE}$. Now we have

$$\begin{aligned} F(\psi,\phi) &= |\langle\phi_{AE}|V_E U_E|\psi_{AE}\rangle|, \\ &= |\langle\Phi^+|_{AE} \sqrt{\sigma_A} V_E U_E |\sqrt{\rho_A}|\Phi^+\rangle_{AE}|, \\ &= |\mathrm{Tr}[\sqrt{\sigma_A}\sqrt{\rho_A} \cdot V_E U_E]|. \end{aligned} \tag{25}$$

  Here the last line uses the identity $\mathrm{Tr}[A^\dagger B] = \langle\Phi^+|A \otimes B|\Phi^+\rangle$. Using the variational form of the trace norm, we have

$$\max_{\psi,\phi} F(\psi,\phi) = \max_{U_E V_E} |\mathrm{Tr}[\sqrt{\sigma_A}\sqrt{\rho_A} \cdot V_E U_E]| = |\mathrm{Tr}[|\sqrt{\sigma_A}\sqrt{\rho_A}|]| = F(\rho,\sigma). \tag{26}$$

  □

The Uhlmann's Theorem could be regarded as a variational form of the fidelity, : ich links complicated mixed fidelity with pure state fidelity.

- Monotonicity under quantum channels: $F(\rho,\sigma) \geq F(\mathcal{E}(\rho), \mathcal{E}(\sigma))$.

  Proof. Similar to the proof for trace distance, we only need to show the increase of fidelity under partial trace, which can be proved using Uhlmann's Theorem. □

- Strong concavity, $F(\sum_j p_j\rho_j, \sum_j q_j\sigma_j) \geq \sum_j \sqrt{p_jq_j}F(\rho_j,\sigma_j)$.

  Proof. Suppose we have the purifications ψ_j and ϕ_j of ρ_j and σ_j that achieves $F(\rho_j,\sigma_j) = F(\psi_j,\phi_j)$. Then $|\psi\rangle = \sum_j \sqrt{p_j}\,|\psi_j\rangle\,|j\rangle$ and $|\phi\rangle = \sum_j \sqrt{q_j}\,|\phi_j\rangle\,|j\rangle$ are respectively the purifications of $\sum_j p_j\rho_j$ and $\sum_j q_j\sigma_j$. Then we have

$$F(\sum_j p_j\rho_j, \sum_j q_j\sigma_j) \geq F(\psi,\phi) = \sum_j \sqrt{p_jq_j}|\,\langle\psi_j|\phi_j\rangle\,| = \sum_j \sqrt{p_jq_j}F(\rho_j,\sigma_j). \tag{27}$$

  Here the first inequality uses the the Uhlmann's Theorem. □

- Reduction to classical fidelity. Consider a POVM $\{E_j\}$ with $p_j = \mathrm{Tr}[\rho E_j]$ and $q_j = \mathrm{Tr}[\sigma E_j]$, then

$$F(\rho,\sigma) = \min_{\{E_j\}} F(\mathbf{p},\mathbf{q}). \tag{28}$$

  The proof of $F(\mathbf{p},\mathbf{q}) \geq F(\rho,\sigma)$ is similar to the one for trace distance. The POVM that achieves the the equal sign is the eigenbasis of the the Fuchs-Caves measurement $M = \rho^{-1/2}\sqrt{\sqrt{\rho}\sigma\sqrt{\rho}}\rho^{-1/2}$. Denote the spectral decomposition of $M = \sum_x \lambda_x\,|x\rangle\,\langle x|$ with non-negative eigenvalues (see Nielsen's book for the construction insights), and note $M\rho M = \sigma$. We have

$$\begin{aligned} F(\mathbf{p},\mathbf{q}) &= \sum_x \sqrt{\langle x|\,\rho\,|x\rangle \cdot \langle x|\,\sigma\,|x\rangle}, \\ &= \sum_x \sqrt{\langle x|\,\rho\,|x\rangle \cdot \langle x|\,M\rho M\,|x\rangle}, \\ &= \sum_x \sqrt{\langle x|\,\rho\,|x\rangle \cdot \langle x|\,\lambda_x\rho\lambda_x\,|x\rangle}, \\ &= \sum_x \lambda_x\,\langle x|\,\rho\,|x\rangle\,, \\ &= \mathrm{Tr}[M\rho], \\ &= \mathrm{Tr}\left[\rho^{-1/2}\sqrt{\sqrt{\rho}\sigma\sqrt{\rho}}\rho^{-1/2}\rho\right], \\ &= F(\rho,\sigma). \end{aligned} \tag{29}$$

  Note that when ρ is no invertible, we can define Π_ρ to be the projection onto the support of ρ and prove $F(\rho,\sigma) = F(\rho,\Pi_\rho\sigma\Pi_\rho)$.

3 Relation between trace distance and fidelity

Pure states For two pure states ψ and ϕ, trace distance and fidelity are actually equivalent

$$D(\psi,\phi) = \sqrt{1 - F(\phi,\psi)^2}. \tag{30}$$

General case In general, we have

$$1 - F(\rho,\sigma) \leq D(\rho,\sigma) \leq \sqrt{1 - F(\rho,\sigma)^2}. \tag{31}$$

The first inequality could be proven by using the reduction to classical trace distance and fidelity and their relation as we proved in Eq. (17). For the second inequality, we consider the purification of ρ and σ, i.e., ψ and ϕ, that achieves the fidelity according to Uhlmann's theorem. Then we have

$$D(\rho,\sigma) \leq D(\psi,\phi) = \sqrt{1 - F(\phi,\psi)^2} = \sqrt{1 - F(\rho,\sigma)^2}. \tag{32}$$