ETH
Eidgenössische Technische Hochschule Zürich
Swiss Federal Institute of Technology Zurich

Quantum Information Theory
Solutions 1

HS 12
Prof. R. Renner

Exercise 1.1 Trace distance

The trace distance (or L_1-distance) between two probability distributions P_X and Q_X over a discrete alphabet $\mathcal{X}$ is defined as

$$\delta(P_X, Q_X) = \frac{1}{2} \sum_{x\in\mathcal{X}} |P_X(x) - Q_X(x)| \,. \tag{1}$$

The trace distance may also be written as

$$\delta(P_X, Q_X) = \max_{\mathcal{S}\subseteq\mathcal{X}} |P_X[\mathcal{S}] - Q_X[\mathcal{S}]| \,, \tag{2}$$

where we maximise over all events $S \subseteq \mathcal{X}$ and the probability of an event is given by $P_X[\mathcal{S}] = \sum_{x\in\mathcal{S}} P_X(x)$.

a) *Show that $\delta(\cdot,\cdot)$ is a good measure of distance by proving that $0 \le \delta(P_X, Q_X) \le 1$ and the triangle inequality $\delta(P_X, R_X) \le \delta(P_X, Q_X) + \delta(Q_X, R_X)$ for arbitrary probability distributions P_X, Q_X and R_X.*

The lower bound follows from the fact that each element of the sum (1) is non-negative. We get the upper bound from

$$\delta(P_X, Q_X) = \frac{1}{2} \sum_{x\in\mathcal{X}} |P_X(x) - Q_X(x)| \le \frac{1}{2} \sum_{x\in\mathcal{X}} P_X(x) + Q_X(x) = 1.$$

The triangle inequality can be written as

$$\frac{1}{2} \sum_{x\in\mathcal{X}} |P_X(x) - R_X(x)| \le \frac{1}{2} \sum_{x\in\mathcal{X}} |P_X(x) - Q_X(x)| + |Q_X(x) - R_X(x)| \,.$$

If the inequality is true for every $x \in \mathcal{X}$, it is also true for the above sum. It is thus sufficient to prove that $|P_X(x) - R_X(x)| \le |P_X(x) - Q_X(x)| + |Q_X(x) - R_X(x)|$ for all $x \in \mathcal{X}$. We know that $|\alpha + \beta| \le |\alpha| + |\beta|$ for $\alpha, \beta \in \mathbb{R}$. Hence the inequality follows with $\alpha = P_X(x) - Q_X(x)$ and $\beta = Q_X(x) - R_X(x)$.

b) *Show that definitions (2) and (1) are equivalent.*

To maximise $|P_X[\mathcal{S}] - Q_X[\mathcal{S}]| = \left|\sum_{x\in\mathcal{S}} P_X(x) - Q_X(x)\right|$ in (2), we choose

$$\mathcal{S} = \{x \in \mathcal{X} \;:\; P_X(x) \ge Q_X(x)\} \,.$$

Let $\bar{\mathcal{S}}$ be its complement, such that $\mathcal{S} \cup \bar{\mathcal{S}} = \mathcal{X}, \mathcal{S} \cap \bar{\mathcal{S}} = \emptyset$. We may now write

$$0 = \sum_{x\in\mathcal{X}} P_X(x) - Q_X(x) = \sum_{x\in S} |P_X(x) - Q_X(x)| - \sum_{x\in\bar{\mathcal{S}}} |P_X(x) - Q_X(x)| \,.$$

The terms $P_X(x) - Q_X(x)$ are positive in the first sum on the right-hand side and negative in the second sum. We can thus take the modulus after the sum in the first term and write

$$\left|\sum_{x\in\mathcal{S}} P_X(x) - Q_X(x)\right| = \sum_{x\in\bar{\mathcal{S}}} |P_X(x) - Q_X(x)| = \frac{1}{2} \sum_{x\in\mathcal{X}} |P_X(x) - Q_X(x)| \,.$$

This proves that the two definitions (1) and (2) are equivalent.

c) *Let us now find an operational meaning for the trace distance. Suppose that P_X and Q_X represent the probability distributions of the outcomes of two dice that look identical. You are only allowed to throw one of them once and then you have to guess which die it was. What is your best strategy? What is the probability that you guess correctly and how can you relate this probability to the trace distance $\delta(P_X, Q_X)$?*

Your best strategy is to say it was the die more likely to outcome the result you obtained, ie. if you define the event $\mathcal{S} = \{x \in \mathcal{X} : P_X(x) \geq Q_X(x)\}$ (the results that are more likely with die P), than you better say that you threw die P if you get an outcome $x \in \mathcal{S}$ and Q if $x \in \bar{\mathcal{S}}$.

The probability that your guess is right is

$$P_{\checkmark} = \frac{1}{2} P_X(\mathcal{S}) + \frac{1}{2} Q_X(\bar{\mathcal{S}}) = \frac{1}{2} \left(P_X(\mathcal{S}) + 1 - Q_X(\mathcal{S})\right) = \frac{1}{2} \left(1 + \delta(P_X, Q_X)\right),$$

by definition (2) of trace distance.

Exercise 1.2 Weak Law of Large Numbers

a) *Prove Markov's inequality*

$$P[A \geq \varepsilon] \leq \frac{\langle A \rangle}{\varepsilon}. \tag{3}$$

This is done by multiplying the summands by a fraction $a/\varepsilon \geq 1$:

$$P[A \geq \varepsilon] = \sum_{a \geq \varepsilon} P_A(a) \leq \sum_{a \geq \varepsilon} \frac{a P_a(a)}{\varepsilon} \leq \sum_{a} \frac{a P_a(a)}{\varepsilon} = \frac{\langle A \rangle}{\varepsilon}.$$

b) *Use Markov's inequality to prove the weak law of large numbers for i.i.d. X_i:*

$$\lim_{n \to \infty} P\left[\left(\frac{1}{n} \sum_i X_i - \mu\right)^2 \geq \varepsilon\right] = 0 \quad \textit{for any } \varepsilon > 0, \mu = \langle X_i \rangle. \tag{4}$$

First note that we can substitute $A \to (X - \mu)^2$ into Markov's inequality to get Chebyshev's inequality

$$P[(X - \mu)^2 \geq \varepsilon] \leq \frac{\sigma^2}{\varepsilon},$$

where σ is the standard deviation of X. If we now substitute $X \to \frac{1}{n} \sum X_i$ the expectation value remains the same, whereas the variance scales with $\frac{1}{n}$. We get

$$P\left[\left(\frac{1}{n} \sum_i X_i - \mu\right)^2 \geq \varepsilon\right] \leq \frac{\sigma^2}{n\varepsilon}.$$

and the weak law of large numbers follows with $n \to \infty$ for any fixed $\varepsilon > 0$.

Exercise 1.3 Conditional probabilities: how knowing more does not always help

You and your grandfather are trying to guess if it will rain tomorrow. All he knows is that it rains on 80% of the days. You know that and you also listen to the weather forecast and know that it is right 80% of the time and is always correct when it predicts rain.

Let us start by sorting the notation:

P_R - probability that it rains;
P_S - probability that it is sunny (no rain);
$P_{R|\hat{R}}$- probability that it rains *when* radio predicts rain;
$P_{\hat{R}}$ - probability that the radio predicts rain;
$P_{\hat{S}}$ - probability that the radio predicts sunshine;
$P_{R\hat{R}}$- probability that it rains *and* radio predicted rain.

Notice that $P_{R|\hat{R}}$ is a conditioned probability while $P_{R\hat{R}}$ is a joint probability:

$$P_{R\hat{R}} = P_{R|\hat{R}} P_{\hat{R}}. \tag{5}$$

a) *What is the optimal strategy for your grandfather? What is your optimal strategy?*

You were given the probabilities $P_R = 80\%, \quad P_{R\hat{R}} + P_{S\hat{S}} = 80\%, \quad P_{R|\hat{R}} = 100\%.$

The best thing your grandfather can do is to say it will rain every morning – this way he will win 80% of the time. As for you, if you use (5) you will compute the probabilities represented in Fig. 1 (note: this figure could be interpreted as a channel – check exercise sheet two for details).

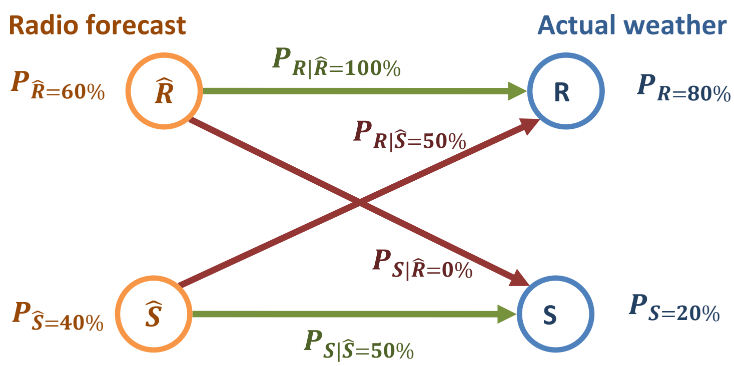

Figure 1: The radio forecast and the actual weather: marginal and conditional probabilities.

When the forecast is rain you should believe it. When the report predicts sun it fails with 50% chance, so any strategy in this case is equally good (or bad). You may for instance say it will always rain or follow the forecast.

b) *After some months who will have guessed correctly more often?*

"After some months" means "after so many days that you can apply the weak law of large numbers". Both you and your grandfather will be correct on approximately 80% of the days – this is easy to see since one of your optimal strategies is to copy your grandfather and say it will always rain. Formally, your success probability is $P_{\hat{R}} P_{\checkmark|\hat{R}} + P_{\hat{S}} P_{\checkmark|\hat{S}} = 0.6 \cdot 1 + 0.4 \cdot 0.5 = 0.8.$

c) *Can you design a method to convince your grandfather that the forecast is useful? Be precise.*

See exercise sheet two.