Helstrom measurement

10/24/2022

A common experimental task is to discriminate between two quantum states. For example, it arises when one uses a quantum system as a sensor. The system starts in a certain initial state and evolves under a Hamiltonian that depends on some external field. For simplicity, we assume that the field is either zero or has some known nonzero value. In each case, the final state can be computed; we just need to identify the physical state with one of the theoretical predictions so that we can tell whether the field was present or not.

More abstractly, suppose we have two density matrices, ρ_0 and ρ_1, and a physical state ρ that is described by one of them. The goal is to tell which one it is, meaning that we need to do some measurement on ρ and produce an answer, 0 or 1. In general, the measurement cannot not be absolutely reliable, that is, if the actual state is ρ_0, the answer will be 1 with some probability, and vice versa. The goal is to make it as reliable as possible.

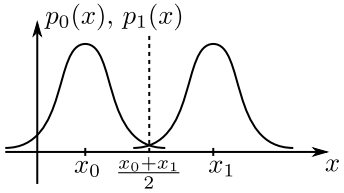
To completely specify the problem, we need to assign some *prior probabilities* q_0, q_1 to the two cases in question. If we think that the physical state is equally likely to be ρ_0 and ρ_1, it is reasonable to use $q_0 = q_1 = 1/2$. However, if the second possibility is less likely, we may want to assign q_1 a smaller value. This procedure is not very rigorous because we are trying to model our vague state of knowledge with probabilities.[1] But once q_0 and q_1 are fixed, the problem is well-defined. We just assume that somebody hands us a physical state that is ρ_0 with probability q_0 and ρ_1 with probability q_1.

Before elaborating on the quantum problem, let us discuss the classical case, where ρ_0 and ρ_1 are replaced with some probability distributions p_0 and p_1. First, a random number $j \in \{0, 1\}$ is generated according to the prior probabilities, but it is hidden from us. Then a random variable x is drawn from the probability distribution p_j. Thus, the pair (j, x) is realized with probability $q_j p_j(x)$. Conditioned on the variable x (known to us), the probability to have a given j is

$$\widetilde{q}_j = \frac{q_j p_j(x)}{p(x)}, \qquad \text{where} \quad p(x) = \sum_j q_j p_j(x). \tag{1}$$

The numbers $\widetilde{q}_0$, $\widetilde{q}_1$ are known as *posterior probabilities*. They describe our updated state of knowledge. Now, the optimal answer is almost obvious: if $\widetilde{q}_0 > \widetilde{q}_1$, then the hidden variable j is more likely equal to 0, and if $\widetilde{q}_0 < \widetilde{q}_1$, it is more likely equal to 1. For example, let p_0, p_1 be Gaussian distributions with different mean values but the same variance:

$$p_0(x) = \pi^{-1/2} e^{-(x-x_0)^2}, \quad p_1(x) = \pi^{-1/2} e^{-(x-x_1)^2} \tag{2}$$



[1] For example, if you see somebody walking a few blocks away, it's most likely a human, but there is a tiny chance it is a bear. Is it a 10^{-5} chance? Maybe. Similarly, when one looks for signs of a new particle in experimental data, it is assumed that the prior probability for it to exist is very low, say, 10^{-6}. The subjectivity of this assumption is compensated by repeated measurements, such that this number is not important in the end.

Assuming equal prior probabilities, $q_0 = q_1 = 1/2$, the more likely j is equal to 0 if $x < \frac{x_0+x_1}{2}$ and 1 if $x > \frac{x_0+x_1}{2}$.

In the quantum discrimination problem, the choice between 0 and 1 is determined by some POVM, which is given by a pair of Hermitian positive-semidefinite operators P_0, P_1 such that $P_0 + P_1 = I$. POVMs with just two possible outcomes are quite simple because in this case, P_0 and P_1 commute. One can also represent P_0 and P_1 by the operator $Y = I - 2P_0 = 2P_1 - I$. Thus,

$$P_0 = \frac{I - Y}{2}, \quad P_1 = \frac{I + Y}{2}, \qquad Y^\dagger = Y, \quad -1 \leqslant Y \leqslant 1. \tag{3}$$

The last inequality constrains the eigenvalues of Y and is equivalent to the condition $\|Y\| \leqslant 1$, modulo the fact that Y is Hermitian. Given an arbitrary state ρ, the discrimination procedure will output 0 with probability $\operatorname{Tr}(P_0\rho)$ and 1 with probability $\operatorname{Tr}(P_1\rho)$. By assumption, $\rho = \rho_j$ with probability q_j. Thus, the probability of a wrong answer (1 if $j = 0$ or 0 if $j = 1$) is given by the formula

$$\varepsilon = q_0 \operatorname{Tr}(P_1\rho_0) + q_1 \operatorname{Tr}(P_0\rho_1) = \frac{1}{2}\Big(1 - \operatorname{Tr}\big(Y(q_1\rho_1 - \rho_0\rho_0)\big)\Big). \tag{4}$$

Here we have used equation (3) together with $q_0 + q_1 = 1$ and $\operatorname{Tr}\rho_0 = \operatorname{Tr}\rho_1 = 1$.

To minimize ϵ, we should maximize $\operatorname{Tr}(YX)$ over Y, where

$$X = q_1\rho_1 - q_0\rho_0 \tag{5}$$

and Y satisfies the above constraints. Using the eigenvalue decomposition $X = \sum_j x_j|\xi_j\rangle\langle\xi_j|$, the optimal solution is

$$Y = \sum_j \operatorname{sgn}(x_j)|\xi_j\rangle\langle\xi_j|, \tag{6}$$

and we have $\operatorname{Tr}(YX) = \sum_j |x_j|\,|\xi_j\rangle\langle\xi_j| = \|X\|_1$. Thus,

$$\varepsilon_{\min} = \frac{1}{2}\big(1 - \|q_1\rho_1 - q_0\rho_0\|_1\big). \tag{7}$$

This construction is called *Helstrom measurement*. If the prior probabilities are $q_0 = q_1 = 1/2$, the equation is a bit simpler:

$$\boxed{\varepsilon_{\min} = \frac{1}{2}\left(1 - \frac{\|\rho_1 - \rho_0\|_1}{2}\right).} \tag{8}$$

Let us consider the special case of pure states, $\rho_j = |\psi_j\rangle\langle\psi_j|$. Since the relative position of two unit vectors $|\psi_0\rangle$, $|\psi_1\rangle$ is determined by their inner product, we may assume without loss of generality that $|\psi_0\rangle = |0\rangle$ and $|\psi_1\rangle = c_0|0\rangle + c_1|1\rangle$. Then

$$X = \rho_1 - \rho_0 = \begin{pmatrix} |c_0|^2 - 1 & c_0c_1^* \\ c_1c_0^* & |c_1|^2 \end{pmatrix}, \qquad x_0, x_1 = \pm|c_1|, \qquad \|X\|_1 = 2|c_1| = 2\sqrt{1 - |\langle\psi_0|\psi_1\rangle|^2}. \tag{9}$$

Thus,

$$\varepsilon_{\min} = \frac{1}{2}\Big(1 - \sqrt{1 - |\langle\psi_0|\psi_1\rangle|^2}\Big). \tag{10}$$

We now illustrate this result with a concrete example. Let $|\psi_0\rangle$ and $|\psi_1\rangle$ be Gaussian wavepackets, namely,

$$\psi_0(x) = \pi^{-1/4} e^{-(x-x_0)^2/2}, \qquad \psi_1(x) = \pi^{-1/4} e^{-(x-x_1)^2/2}. \tag{11}$$

In this case, $\langle\psi_0|\psi_1\rangle = e^{-(x_1-x_0)^2/4}$, and hence,

$$\varepsilon_{\min} = \frac{1}{2}\Big(1 - \sqrt{1 - e^{-(x_1-x_0)^2/2}}\Big) \approx \frac{1}{4}\, e^{-(x_1-x_0)^2/2}. \tag{12}$$

The last expression is for the limit of almost perfect separation, $|x_1 - x_0| \gg 1$.

Unfortunately, the optimal measurement is hard to implement. In practice, one typically measures x and does the classical discrimination of the corresponding probability distributions (2). This procedure has error probability

$$\begin{aligned} \varepsilon &= \frac{1}{2}\int_{-\infty}^{0} |\psi_1(x)|^2\, dx + \frac{1}{2}\int_{0}^{+\infty} |\psi_0(x)|^2\, dx = \pi^{-1/2} \int_{(x_0-x_1)/2}^{+\infty} e^{-x^2}\, dx \\ &= \frac{1}{2}\Big(1 - \operatorname{erf}\big((x_0 - x_1)/2\big)\Big) \approx \frac{e^{-(x_0-x_1)^2/4}}{\sqrt{\pi}\,(x_0 - x_1)}. \end{aligned} \tag{13}$$

Notice the factor of 2 difference in the exponent, indicating that the optimal quantum measurement is quadratically more efficient than the classical measurement.