PRETTY-GOOD MEASUREMENT

ZAK WEBB

This will hopefully be a quick introduction to the pretty-good measurement, a procedure for producing a set of measurements for distinguishing a set of given states, that works "pretty good." It's not always an optimal measurement, but the fact that it is easily described allows it to be a first guess when an measurement is needed for some theoretical task. Further, in some cases it does provide an optimal distinguishing probability, which is useful theoretically.

1. Quantum hypothesis testing

The first thing we need to do when discussing the pretty-good measurement is to describe the problem it is used to solve. In particular, let us assume that we are given a quantum state on an N-dimensional system, where we know the ensemble from which the state is chosen. We then want to distinguish which state we received, and do so with a large probability of success.

More concretely, let us define the ensemble

$$\mathcal{E} = \big\{(p_i, \rho_i)\big\}_{i\in[M]}, \tag{1}$$

where each $\rho_i \in \mathcal{D}(\mathbb{C}^N)$, and the $(p_i)_{i\in[M]}$ form a probability distribution. Given a random σ distributed according to the distribution $\mathcal{E}$, we want to determine which $i \in [M]$ for which $\sigma = \rho_i$. The goal is then to construct a measurement scheme that maximizes the probability of correctly determining which state we are given.

This is a rather general problem in quantum information, and it depends greatly on the given distribution. For example, if $p_1 = 1$ the problem is trivial, while if each $\rho_i = \rho_j$, then the best solution is to simply guess whichever j has the largest probability of occurring.

This problem has a solution in the case for $M = 2$, as the fact that a given POVM's operators must sum to the identity means that we are really only trying to maximize over one positive semi-definite operator. In particular, we have that $p_1 = 1 - p_2 = p$, and $M_1 = \mathbb{I} - M_2 = M$, and the probability of successfully determining the correct $i \in [2]$ is given explicitly by

$$\Pr(\text{Success}) = p \operatorname{Tr}\big[M\rho_1\big] + (1-p)\operatorname{Tr}\big[(\mathbb{I} - M)\rho_2\big] \tag{2}$$

$$= 1 - p + \operatorname{Tr}\Big\{\big[p(\rho_1 + \rho_2) - \rho_2\big]M\Big\}. \tag{3}$$

Hence, in this special case, the success probability is maximized when M is a projector onto the positive eigenspace of $p(\rho_1 + \rho_2) - \rho_2$.

When we extend the problem to multiple outcomes, however, we no longer have as nice of a solution. In fact, as far as I am aware (and I could probably do a more thorough literature search), there isn't a nice closed form solution for this problem.

1.1. Holevo bound.

While we don't actually have a solution to this problem, we do have some upper bounds on the amount of information that can be extracted from the state σ. In particular, Holevo's theorem states that for a given ensemble $\mathcal{E}$ of quantum states, the amount of information that can be accessed about the initial $i \in [M]$ is bounded from above by

$$\xi(\mathcal{E}) = S\Big(\sum_{i\in[M]} p_i\rho_i\Big) - \sum_{i\in[M]} p_i S(\rho_i), \tag{4}$$

where $S(\rho)$ is the von Neumann entropy of the state ρ.

I might want to expound on this bound, but I'm not sure.

Date: September 16, 2016.

2. Intuition for the pretty-good measurement

The goal of the pretty good measure is to provide a measurement scheme for a given QHT that succeeds with a decent amount of probability, but is also somewhat intuitive. We will now proceed with two special cases of the QHT, for which an obvious measurement strategy works. The pretty-good measurement generalizes these special cases, with only a small amount of work.

2.1. **Pure and orthogonal states.** Let us first look at the case where each $\rho_i = |\psi_i\rangle\langle\psi_i|$, and where $\langle\psi_i|\psi_j\rangle = 0$ for $i \neq j$.

In this special case, we have several orthogonal states, each of which is pure. The natural idea is then to simply measure in the corresponding basis, where $M_i = \rho_i$ (possibly including an $M+1$'st measurement if the states $|\psi_i\rangle$ don't span the entire N-dimensional Hilbert space). This measurement scheme succeeds perfectly, as one might expect given that the states are all orthogonal.

2.2. **Mixed and orthogonal states.** Now let us look at the special case where each ρ_i can now be mixed, but the supports of the ρ_i are still disjoint.

In this case, if we again use the same $M_i = \rho_i$ as above, we are guaranteed that we never guess the incorrect answer, since

$$\mathrm{Tr}[M_i\rho_j] = \mathrm{Tr}[\rho_i\rho_j] = 0, \tag{5}$$

due to the fact that the density matrices have orthogonal support. However, in the correct case, we have that

$$\mathrm{Tr}[M_i\rho_i] = \mathrm{Tr}[\rho_i\rho_i] = \mathrm{Tr}[\rho_i^2] \leq 1, \tag{6}$$

since the states might no longer be pure.

What we want instead of the state itself is a projector onto the support of each ρ_i, as this will still keep the orthogonality condition, but will also always accept the state ρ_i. The way we do this is to use the Moore-Penrose pseudo-inverses, so that

$$M_i = \rho_i^{-1/2}\rho_i\rho_i^{-1/2}. \tag{7}$$

This is exactly what we wanted, as M_i is then the projector onto the support of ρ_i.

This collection of M_i is then guaranteed to succeed with probability 1 for the QHT with these special states.

2.3. **General collection of states.** Now let us not make any assumptions on the ensemble $\mathcal{E}$ other than it is an ensemble of states from $\mathcal{D}(\mathbb{C}^N)$. If we make the same definitions of the M_i as in the second case above, we run into a problem, in that the sum of the measurement operators have no relation with the identity! They can be bigger, smaller, or simply incomparable.

The problem in this case is simply that if the given states ρ_i have overlapping support, when we turn the measurement operators into projectors onto the support, we are essentially double-counting those subspaces. We somehow need to ensure that the sum of our measurement operators is always less than the identity.

The way that this is done for the pretty good measurement is to use the pseudo-inverse of a a slightly different operator, namely the density matrix of the actual state given to us:

$$S := \sum_{i\in[M]} p_i\rho_i, \tag{8}$$

and now we define the measurement operators as

$$M_i = p_i S^{-1/2}\rho_i S^{-1/2}. \tag{9}$$

Note that while both S and M_i contain the factors p_i, they cancel when the ρ_i have orthogonal support, and thus is a generalization of the previous constructions. Further, we have that

$$\sum_{i\in[M]} M_i = S^{-1/2}\left[\sum_{i\in[M]} p_i\rho_i\right]S^{-1/2} \leq \mathbb{I}, \tag{10}$$

where the sum is only less than the identity if the support of the ρ_i don't span the entire Hilbert space. As such, we have that these M_i form a valid POVM (after possibly including an additional operator).

Additionally, the inclusion of the p_i makes sense. If two of the ρ_i were equal, we would somehow want to ensure that one of them was chosen with a larger probability: this construction ensures that the state with a larger probability is chosen more often.

3. Bounds on the error

I'm not going to prove these, but there are some explicit bounds that we can place on the success probability of the pretty good measurement versus the optimal measurement.

$$1 - P_{PGM}(\mathcal{E} \leq 2 \sum_{i<j} \sqrt{p_i p_j} F(\sigma_i, \sigma_j), \tag{11}$$

where the fidelity is given by

$$F(\sigma, \rho) = \|\sqrt{\sigma}\sqrt{\rho}\|_1. \tag{12}$$

From this, we can then prove that

$$P_{PGM}(\mathcal{E}) \geq P_{\text{opt}}(\mathcal{E})^2, \tag{13}$$

so that the pretty good measurement is not that far from optimal.

Additionally, if the states ρ_i are all pure, there are some lower bounds we can make on the success probability, namely:

$$P_{PGM}(\mathcal{E}) \geq \sum_{i=\in[M]} \frac{p_i^2}{\sum_{j\in[M]} p_j |\langle\psi_j|\psi_i\rangle|^2}. \tag{14}$$

These bounds are also related to the Gram Matrix. Namely, let

$$G = \sum_{i,j\in[M]} \langle\psi_j|\psi_i\rangle |i\rangle\langle j|. \tag{15}$$

If the eigenvalues of G are given by $\{\lambda_i\}$, then we have

$$P_{PGM}\mathcal{E} \geq \frac{1}{n}\Big(\sum_{i=1}^{n} \sqrt{\lambda_i}\Big)^2. \tag{16}$$

If we again allow for mixed states, we have the slightly worse bounds of

$$P_{PGM}(\mathcal{E}) \geq \sum_{i\in[N]} \frac{p_i^2 \operatorname{Tr}(\rho_i^2)}{\sum_{j\in[M]} p_j F(\rho_i, \rho_j)}. \tag{17}$$

4. Application to dihedral HSP

References

[1] Howard Barnum and Emanuel Knill, *Reversing quantum dynamics with near-optimal quantum and classical fidelity*, Journal of Mathematical Physics **43** (2002), no. 5, 2097–2106, quant-ph/0004088.

[2] AS Holevo, *The capacity of quantum channel with general signal states*, IEEE Trans. Info. Theor. **44** (1996), 269–273, quant-ph/9611023.

[3] Ashley Montanaro, *On the distinguishability of random quantum states*, Communications in Mathematical Physics **273** (2007), no. 3, 619–636, quant-ph/0607011.

[4] John Watrous, *Theory of quantum information*, 2016.

[5] John Wright, *Lecture 20: Pretty good measurement.*