CHAPTER 7

Shadow Tomography via Online Learning of Quantum States

The previous lecture covered an elegant framework for shadow tomography using only very simple randomized measurements. One drawback of this approach however is that in order for the sample complexity to be not too large, the observables being estimated have to be bounded in shadow norm, e.g., properties that are local or bounded in rank.

In this lecture, we present an alternative approach to shadow tomography which achieves efficient sample complexity for *general* observables, albeit at the cost of not being computationally efficient.

Throughout we assume that the Hermitian observables O_i are positive semidefinite and satisfy $\|O_i\|_{\mathsf{op}} \leq 1$; in other words, their eigenvalues all lie in the interval $[0, 1]$. The assumption of psd-ness is without loss of generality as we can write any observable as a difference of psd operators $O^+ - O^-$ and estimate O^+ and O^- separately.

1. Online Learning

We begin by considering the following abstract setting wherein a "student" trying to learn the observable values for an unknown d-dimensional quantum state ρ interacts with a "teacher." The teacher shows the student a sequence of observables $O_1, O_2, \ldots$ to the student, where $\|O_t\|_{\mathsf{op}} \leq 1$ for all i. Every time the student gets a new O_t, she needs to form a prediction $\hat{o}_t$ for $\mathrm{tr}(O_t\rho)$, at which point the teacher will declare either

- "Pass" if $|\hat{o}_t - \mathrm{tr}(O_t\rho)| \leq \frac{3}{4}\epsilon$, or
- "Fail" if $|\hat{o}_t - \mathrm{tr}(O_t\rho)| > \epsilon$.

If the student is in the "gray zone" where $|\hat{o}_t - \mathrm{tr}(O_t\rho)| \in (\frac{3}{4}\epsilon, \epsilon]$, the teacher's response can be either "pass" or "fail."

If however the teacher ever outputs "fail," then the teacher must tell the student whether $\hat{o}_t \geq \mathrm{tr}(O_t\rho)$ or $\hat{o}_t < \mathrm{tr}(O_t\rho)$.

This problem is called **online state learning**, first studied by [**ACH⁺18**]. Remarkably, as we will show in this section, there is an algorithm that the student can run to limit the number of "fails" they receive to a number independent of the number of rounds of interactions with the teacher!

Theorem 109. *In quantum state learning, there is an algorithm that the student can run which ensures that she only mistakes at most $O(\log(d)/\epsilon^2)$ mistakes in total, regardless of the length of the sequence of observables that the teacher provides, and regardless of the (potentially adversarial) strategy by which the teacher selects observables to give the student.*

Algorithm 5: MatrixMultiplicativeWeights($\{M_t\}_{t=1}^T, \eta$)

Input: Sequence of "cost" matrices $M_1, \ldots, M_T$, learning rate $\eta > 0$

Output: Sequence of responses $\rho_1, \ldots, \rho_T$

$H_1 \leftarrow 0$;

for $t = 1, \ldots, T$ **do**

  Receive cost matrix M_t;

  Respond with $\rho_t = \exp(-\eta H_t)/\operatorname{tr}\exp(-\eta H_t)$;

  $H_{t+1} \leftarrow H_t + M_t = \sum_{s=1}^t M_t$;

end

1.1. Matrix multiplicative weights

To prove Theorem 109, we consider an even more abstract setting. Suppose the student is given a sequence of matrices $M_1, M_2, \ldots, M_T$, again with $\|M_t\|_{\mathsf{op}} \le 1$ for all t, and every time the student receives some M_t, she must respond with a density matrix ρ_t. Here we will assume that for each t, either M_t or $-M_t$ is psd, so that the eigenvalues of M_t either all lie in $[0, 1]$ or all lies in $[-1, 0]$.

Over T rounds of interaction, the student's goal is to minimize the **regret**

$$\sum_{t=1}^T \operatorname{tr}(M_t \rho_t) - \min_\rho \sum_{t=1}^T \operatorname{tr}(M_t \rho)\,,$$

where the minimum is taken over all density matrices ρ. Intuitively, $\operatorname{tr}(M_t\rho_t)$ corresponds to a "cost" that the student incurs by predicting with ρ_t in round t, and $\sum_{t=1}^T \operatorname{tr}(M_t\rho)$ is the cost incurred by the "best-in-hindsight" strategy ρ.

This is a classic question within the classical literature on **online learning**, one of the crown jewels of which is the **matrix multiplicative weights** algorithm.

The specific update rule in Algorithm 5 may appear magical upon first glance, but it has a very simple interpretation in terms of the *maximum-entropy principle*.

Definition 110 (Entropy)**.** *Given a density matrix ρ, the* von Neumann entropy *of ρ is defined by $S(\rho) \triangleq \operatorname{tr}(-\rho \log \rho)$.*

Lemma 111. *Given a Hermitian operator H, the density matrix minimizing*

$$F(\rho) \triangleq \eta \operatorname{tr}(H\rho) - S(\rho)$$

is given by $\rho^* = \exp(-\eta H)/\operatorname{tr}\exp(-\eta H)$.

Proof. Define $Z = \operatorname{tr}\exp(-\eta H)$ and let $\rho' = \exp(-\eta H')/Z'$ for some Hermitian operator H', where $Z' \triangleq \operatorname{tr}\exp(-\eta H')$. Then the *relative entropy* $S(\rho'\|\rho^*) \triangleq \operatorname{tr}(\rho' \log \rho') - \operatorname{tr}(\rho' \log \rho^*)$ is given by

$$S(\rho'\|\rho^*) = -\langle \rho', \eta H' - \eta H + \log(Z'/Z)\cdot I\rangle = -\log(Z'/Z) + \eta\langle \rho', H - H'\rangle\,.$$

Note that

$$\begin{aligned} S(\rho') &= \langle \rho', \eta H' + (\log Z')\cdot I)\rangle \\ &= \eta \operatorname{tr}(H'\rho') + \log Z' \\ &= \eta \operatorname{tr}(H\rho') + \log Z' - \langle \rho', H - H'\rangle\,, \end{aligned}$$

so $-\log Z' + \langle \rho', H - H' \rangle = F(\rho')$. We conclude that $S(\rho' \| \rho^*) = F(\rho') - F(\rho^*)$. By (the quantum version of) Gibbs' inequality, $S(\rho' \| \rho^*) \geq 0$, so $F(\rho^*) \leq F(\rho')$ for all ρ' as claimed. $\square$

Thus, the state ρ_t in matrix multiplicative weights is simply the state which optimally balances between minimizing correlation with all of the cost matrices seen so far, and maximizing entropy. This tradeoff is modulated by the learning rate η, which can be interpreted physically as an inverse temperature parameter. Larger η corresponds to overfitting more to previous observations and maintaining less entropy, which may make it more difficult to adapt to future cost matrices. Conversely, smaller η corresponds to adapting less to previous observations less and maintaining large entropy, which may be disadvantageous if similar cost matrices as those seen in the past also appear in the future. The following main guarantee for matrix multiplicative weights ensures that there is a way to balance between these two extremes and achieve *sublinear regret*, that is, regret which is of lower order than the length of the time horizon T.

Theorem 112. *Given a sequence of cost matrices $M_1, \ldots, M_T$, possibly adaptively chosen,* MatrixMultiplicativeWeights*$(\{O_t\}, \eta)$ produces a sequence of responses $\rho_1, \ldots, \rho_T$ such that the regret is bounded by*

$$\sum_{t=1}^{T} \operatorname{tr}(M_t \rho_t) - \min_{\rho} \sum_{t=1}^{T} \operatorname{tr}(M_t \rho) \leq \eta T + \frac{\log d}{\eta} .$$

In particular, taking $\eta = \sqrt{\log(d)/T}$ results in a regret bound of $2\sqrt{T \log d}$.

The proof is based on an elegant potential function argument.

Proof. Let $Z_t \triangleq \operatorname{tr}\exp(-H_t)$. First note that

$$Z_t \geq \exp(-\eta \sigma_{\min}(H_t)) \tag{29}$$

for all t. Let $\sum_{s=1;\succeq 0}^{t}$ (resp. $\sum_{s=1;\preceq 0}^{t}$) denote the sum over all $1 \leq s \leq t$ for which M_s (resp. $-M_s$) is psd. We will show that

$$Z_t \leq d \exp\Big(-\eta_1 \sum_{s=1;\succeq 0}^{t-1} \operatorname{tr}(M_s \rho_s) - \eta_2 \sum_{s=1;\preceq 0}^{t-1} \operatorname{tr}(M_s \rho_s)\Big) , \tag{30}$$

for $\eta_1 \triangleq 1 - e^{-\eta}$ and $\eta_2 \triangleq e^{\eta} - 1$. Combining Eqs. (29) and (30) for $t = T+1$ and taking logs on both sides and dividing by η, we conclude that

$$\begin{aligned}
\min_{\rho} \operatorname{tr}(H_T \rho) = \sigma_{\min}(H_T) & \\
&\geq \frac{\log d}{\eta} + \frac{\eta_1}{\eta} \sum_{t=1;\succeq 0}^{T} \operatorname{tr}(M_t \rho_t) - \frac{\eta_2}{\eta} \sum_{t=1;\preceq 0}^{T} \operatorname{tr}(M_t \rho_t) \\
&\geq \frac{\log d}{\eta} + (1-\eta) \sum_{t=1;\succeq 0}^{T} \operatorname{tr}(M_t \rho_t) - (1+\eta) \sum_{t=1;\preceq 0}^{T} \operatorname{tr}(M_t \rho_t) \\
&\geq \frac{\log d}{\eta} + \eta T + \sum_{t=1}^{T} \operatorname{tr}(M_t \rho_t) ,
\end{aligned}$$

where the third line follows by $\eta_1 \geq \eta(1-\eta)$ and $\eta_2 \leq \eta(1+\eta)$, and the fourth line follows by the fact that $|\operatorname{tr}(M_t \rho_t)| \leq 1$ for all t.

It remains to show Eq. (30). We proceed by induction on t. For the inductive step, we will use the following nice fact that says that products of matrix exponentials behave like products of scalar exponentials upon taking trace:

Fact 113 (Golden-Thompson inequality)**.** *For all Hermitian operators A and B, $\operatorname{tr}\exp(A+B) \le \operatorname{tr}(\exp(A)\exp(B))$.*

By this, we have

$$\begin{aligned} Z_{t+1} &= \operatorname{tr}\exp(-\eta\sum_{s=1}^{t} M_s) \\ &= \operatorname{tr}(\exp(-\eta\sum_{s=1}^{t-1} M_s)\cdot\exp(-\eta M_t)) \end{aligned}$$

If M_t is psd, then $\exp(-\eta M_t) \preceq \mathrm{Id} - \eta_1 M_t$, using the scalar inequality $e^{-zx} \le 1-(1-e^{-z})x$ for all $z \ge 0$ and $x \in [0,1]$, and the above is thus at most

$$\operatorname{tr}(\exp(-\eta\sum_{s=1}^{t-1} M_s)\cdot(\mathrm{Id}-\eta_1 M_t)) = Z_t(1-\eta_1\operatorname{tr}(M_t\rho_t))\,.$$

Similarly, if $-M_t$ is psd, then $\exp(-\eta M_t) \preceq \mathrm{Id} - \eta_2 M_t$, using the scalar inequality $e^{-zx} \le 1-(e^{z}-1)x$ for all $z \ge 0$ and $x \in [-1,0]$, and the above is thus instead at most

$$\operatorname{tr}(\exp(-\eta\sum_{s=1}^{t-1} M_s)\cdot(\mathrm{Id}-\eta_2 M_t)) = Z_{t-1}(1-\eta_2\operatorname{tr}(M_t\rho_t))\,.$$

As $Z_1 = \operatorname{tr}\exp(-\eta H_1) = \operatorname{tr}(\mathrm{Id}) = d$, Eq. (30) follows by induction. □

1.2. Proof of Theorem 109

Theorem 109 is now almost immediate from the regret bound for matrix multiplicative weights.

Proof. The student will run matrix multiplicative weights, maintaining an estimate of the state ρ_t, but she will only perform an update after rounds in which she "fails." In such a round t, let

$$M_t = \begin{cases} O_t & \text{if } \operatorname{tr}(O_t\rho_t) > \operatorname{tr}(O_t\rho) \\ -O_t & \text{otherwise} \end{cases}.$$

(Note that the student can determine whether $M_t = O_t$ or $M_t = -O_t$ using the teacher's feedback.) By design, $\operatorname{tr}(M_t\rho_t) - \operatorname{tr}(M_t\rho) = |\operatorname{tr}(M_t\rho_t) - \operatorname{tr}(M_t\rho)|$. Let $\hat{o}_t \triangleq \operatorname{tr}(M_t\rho_t)$.

Suppose she ends up getting "fail" at least T times. Then the total cost she incurs in those T rounds is the sum of $\operatorname{tr}(M_t\rho) + |\hat{o}_t - \operatorname{tr}(M_t\rho)|$ from those rounds, whereas the total cost she would have incurred if she had chosen ρ in every round instead of ρ_t is $\operatorname{tr}(M_t\rho)$. Her regret is at least $\frac{3}{4}\epsilon T$ because $|\hat{o}_t - \operatorname{tr}(M_t\rho)| > \frac{3}{4}\epsilon$ in each round where she fails, but by Theorem 112 her regret is also at most $2\sqrt{T\log d}$. From

$$\frac{3}{4}\epsilon T < 2\sqrt{T\log d}$$

we conclude the desired bound on the number of mistakes T. □