Lecture 10

Instructor: Haipeng Luo

1 Dynamic Regret

In the previous lecture we discussed interval regret which only considers regret on a certain time interval, or in other words, it only considers the performance of the algorithm in a local region. In this lecture we will discuss how to measure the global performance of an algorithm over T rounds, still under some non-stationary environment where the best single fixed point in hindsight is not the right benchmark to compare to.

The most ambitious goal would be to compare with the best decision $w_t^\star = \operatorname{argmin}_{w \in \Omega} f_t(w)$ *for each round*. Equivalently, this is the same as asking for sublinear regret against all competitor sequence $u_1, \ldots, u_T \in \Omega$:

$$\mathcal{R}_T(u_1, \ldots, u_T) \stackrel{\text{def}}{=} \sum_{t=1}^{T} (f_t(w_t) - f_t(u_t)) = o(T).$$

This is the so-called *dynamic regret*. Perhaps not so surprisingly, sublinear dynamic regret is not achievable in general, especially in an adversarial setting. To see this, simply think about a 2-expert problem where the environment always assigns loss 0 to the expert that has the highest weight $p_t(i)$ from the algorithm and loss 1 to the other expert (recall that the environment sees p_t before assigning losses). Then on each round, clearly the best expert has loss 0, while the algorithm suffers loss at least $1/2$, leading to linear dynamic regret.

However, this does not exclude many interesting situations where sublinear regret is still possible. The first situation we consider is when the competitors stay the same most of time, that is, $\sum_{t=2}^{T} \mathbf{1}\{u_t \neq u_{t-1}\} \leq S - 1$ for some constant S. In other words, we allow the benchmark to divide the total T rounds into S disjoint intervals, and to select the best fixed point on each interval. This is sometimes called the switching regret or tracking regret. The regular regret is clearly a special case when $S = 1$.

Such benchmark can be significantly better than a single fixed point (think about the 2-expert example from last lecture). Such locally stationary environments also appear in practice naturally. For example, think about the problem of product recommendation. It is often the case that data from, say each month, stays stationary and thus comparing to a best fixed decision for each month is pretty reasonable.

How should we obtain small switching regret? In fact, we have already solved this problem implicitly. Indeed, if we have a strongly adaptive algorithm with regret $\mathcal{O}(\sqrt{|\mathcal{I}| \ln T})$ for any interval $\mathcal{I}$, then just by running the exact same algorithm we have by dividing the T rounds into S intervals $\mathcal{I}_1, \ldots, \mathcal{I}_S$ so that the competitor stays the same on each interval,

$$\begin{aligned}
\mathcal{R}_T(u_1, \ldots, u_T) &= \sum_{m=1}^{S} \sum_{t \in \mathcal{I}_m} (f_t(w_t) - f_t(u_t)) = \mathcal{O}\left(\sum_{m=1}^{S} \sqrt{|\mathcal{I}_m| \ln T}\right) \\
&\leq \mathcal{O}\left(\sqrt{S \sum_{m=1}^{S} |\mathcal{I}_m| \ln T}\right) = \mathcal{O}(\sqrt{ST \ln T})
\end{aligned}$$

where the inequality is by Cauchy-Schwarz inequality. Therefore, as long as S is sublinear in T, the switching regret is also sublinear.

Next we consider an even more general situation where the competitors do not need to stay the same locally. Instead, the dynamic regret will depend on the variation of the loss functions $\ell_1, \ldots, \ell_T$, defined as

$$V_T = \sum_{t=2}^{T} \max_{w \in \Omega} |f_t(w) - f_{t-1}(w)|.$$

Small variation implies that the environment is slowly drifting over time and therefore learning is possible. This is very similar to the path length that we discussed before. Previously we derive adaptive regular regret in terms of path length with the hope of getting regret smaller than $\sqrt{T}$, and now we derive dynamic regret in terms of variation with the goal of avoiding linear regret.

Interestingly, this problem is again already solved by using a strongly adaptive algorithm.

Theorem 1 ([Zhang et al., 2017]). *A strongly adaptive algorithm with $\mathcal{R}_{\mathcal{I}} = \mathcal{O}(\sqrt{|\mathcal{I}| \ln T})$ for any interval $\mathcal{I}$ ensures*

$$\mathcal{R}_T(w_1^\star, \ldots, w_T^\star) = \mathcal{O}\left(T^{\frac{2}{3}} (V_T \ln T)^{\frac{1}{3}}\right)$$

where $w_t^\star = \operatorname{argmin}_{w \in \Omega} f_t(w)$.

Proof. Let $\mathcal{I}_1 = [s_1, e_1], \ldots, \mathcal{I}_M = [s_M, e_M]$ be any partition of the whole game $[1, T]$, and

$$V_{\mathcal{I}_m} = \sum_{t=s_m+1}^{e_m} \max_{w \in \Omega} |f_t(w) - f_{t-1}(w)|$$

be the variation of interval $\mathcal{I}_m$. For any $m \in [M]$, we have

$$\begin{aligned}
\sum_{t \in \mathcal{I}_m} (f_t(w_t) - f_t(w_t^\star)) &= \sum_{t \in \mathcal{I}_m} (f_t(w_t) - f_t(w_{s_m}^\star)) + \sum_{t \in \mathcal{I}_m} (f_t(w_{s_m}^\star) - f_t(w_t^\star)) \\
&\le \mathcal{O}(\sqrt{|\mathcal{I}_m| \ln T}) + 2|\mathcal{I}_m| V_{\mathcal{I}_m},
\end{aligned}$$

where the last step is by the guarantee of the strongly adaptive algorithm and the fact

$$\begin{aligned}
f_t(w_{s_m}^\star) - f_t(w_t^\star) &\le f_t(w_{s_m}^\star) - f_{s_m}(w_{s_m}^\star) + f_{s_m}(w_t^\star) - f_t(w_t^\star) \qquad \text{(by optimality of } w_{s_m}^\star) \\
&= \sum_{\tau=s_m+1}^{t} \left(f_\tau(w_{s_m}^\star) - f_{\tau-1}(w_{s_m}^\star)\right) + \sum_{\tau=s_m+1}^{t} (f_{\tau-1}(w_t^\star) - f_\tau(w_t^\star)) \\
&\le 2 \sum_{\tau=s_m+1}^{t} \max_{w \in \Omega} |f_\tau(w) - f_{\tau-1}(w)| \le 2V_{\mathcal{I}_m}.
\end{aligned}$$

Therefore, the dynamic regret is bounded by

$$\begin{aligned}
\mathcal{R}_T(w_1^\star, \ldots, w_T^\star) &\le \sum_{m=1}^{M} \mathcal{O}(\sqrt{|\mathcal{I}_m| \ln T}) + 2|\mathcal{I}_m| V_{\mathcal{I}_m} \\
&\le \mathcal{O}(\sqrt{MT \ln T}) + 2 \max_m |\mathcal{I}_m| \sum_{m=1}^{M} V_{\mathcal{I}_m} \\
&\le \mathcal{O}(\sqrt{MT \ln T}) + 2 \max_m |\mathcal{I}_m| V_T.
\end{aligned}$$

Finally we just need to balance M and $\max_m |\mathcal{I}_m|$. For a fixed M, it is clear that if we want to minimize $\max_m |\mathcal{I}_m|$, we should divide the game (almost) evenly into M intervals so that $\max_m |\mathcal{I}_m| = \mathcal{O}(T/M)$ and $\mathcal{R}_T(w_1^\star, \ldots, w_T^\star) = \mathcal{O}(\sqrt{MT \ln T} + TV_T/M)$. Setting M optimally to $\lfloor (T/\ln T)^{\frac{1}{3}} V_T^{\frac{2}{3}} \rfloor$ finishes the proof. □

According to the theorem, as long as V_T is sublinear, the dynamic regret is sublinear. Note that this does not contradict with the earlier impossibility result since V_T can be linear in T in the worst case.

The bound $\mathcal{O}(T^{\frac{2}{3}}(V \ln T)^{\frac{1}{3}})$ is worst-case optimal as shown in [Besbes et al., 2015], but it is not always tight. For example, suppose f_t stays the same most of the time except for $S-1$ rounds. Then assuming $f_t(w) \in [0,1]$ we have $V_T = S-1$ and thus $\mathcal{R}_T(w_1^\star, \ldots, w_T^\star) = \mathcal{O}(T^{\frac{2}{3}}(S \ln T)^{\frac{1}{3}})$. However, in this case the dynamic regret is clearly also the switching regret and we have showed $\mathcal{O}(\sqrt{ST \ln T})$ is possible (and better), and more importantly achieved by using the exact same strongly adaptive algorithm. Therefore, we can instead write the bound as

$$\mathcal{R}_T(w_1^\star, \ldots, w_T^\star) = \mathcal{O}\left(\min\left\{T^{\frac{2}{3}}(V_T \ln T)^{\frac{1}{3}}, \sqrt{\left(\sum_{t=2}^{T} \mathbf{1}\{w_t^\star \neq w_{t-1}^\star\}\right) T \ln T}\right\}\right).$$

2 Dynamic Regret for the Expert Problem

All the results discussed above apply to the special case of the expert problem of course, but in this section we will introduce a different type of dynamic regret that is more specific to the expert problem. First note that in this case the dynamic regret against a competitor sequence $u_1, \ldots, u_T \in \Delta(N)$ can be written as

$$\mathcal{R}_T(u_1, \ldots, u_T) = \sum_{t=1}^{T} \sum_{i=1}^{N} u_t(i) r_t(i)$$

where $r_t(i) = \langle p_t, \ell_t \rangle - \ell_t(i)$ is the instantaneous regret. It turns out that we can bound this regret with respect to another measure of non-stationarity:

$$A_T = \sum_{t=1}^{T} \sum_{i=1}^{N} [u_t(i) - u_{t-1}(i)]_+,$$

where $[x]_+ = \max\{x, 0\}$ and u_0 is defined as an all-zero vector for convenience. In other words, A_T is the sum of "one-sided ℓ_1 norms" between consecutive competitors, and one can see it as a generalized and soft version of the number of switches (indeed when there are $S-1$ switches in the sequence, we have $A_T \leq S$). Compared to V_T, which is the variation of the loss functions, A_T is the variation of the competitors. These two variations are related but in general not comparable (think about examples where $A_T = \mathcal{O}(1)$ while $V_T = \Omega(T)$, and vice versa.)

Somewhat surprisingly, all we need here is yet again a strongly adaptive algorithm:

Theorem 2 ([Luo and Schapire, 2015]). *For the expert problem, a strongly adaptive algorithm with $\mathcal{R}_\mathcal{I} = \mathcal{O}(\sqrt{|\mathcal{I}| \ln(NT)})$ for any interval $\mathcal{I}$ ensures*

$$\mathcal{R}_T(u_1, \ldots, u_T) = \mathcal{O}\left(\sqrt{TA_T \ln(NT)}\right)$$

for any competitor sequence $u_1, \ldots, u_T \in \Delta(N)$.

So again, whenever A_T is sublinear, the regret is sublinear. Rewriting the regret bound also leads to the following bound on the loss of the algorithm:

$$\sum_{t=1}^{T} \langle p_t, \ell_t \rangle \leq \min_{u_1, \ldots, u_T} \left(\sum_{t=1}^{T} \langle u_t, \ell_t \rangle + \mathcal{O}\left(\sqrt{TA_T \ln(NT)}\right)\right),$$

where one can pick $u_1, \ldots, u_T$ in different ways to balance the benchmark and the regret term.

Proof. The idea is to decompose the regret into weighted sum of several interval regrets. To this end, we fix an expert i and let $\alpha_m > 0$ and $\mathcal{I}_m \subset [1, T]$ ($m = 1, \ldots, M$) for some M be a set of weighted intervals such that $u_t(i) = \sum_{m=1}^{M} \mathbf{1}\{t \in \mathcal{I}_m\} \alpha_m$ for any t. In other words, each $u_t(i)$ is decomposed as the sum of weights of the intervals that cover t. There are many different ways to do this but later we will specify what the optimal way is. For now, note that the regret (against expert

i) can be decomposed as

$$\sum_{t=1}^{T} u_t(i) r_t(i) = \sum_{t=1}^{T}\sum_{m=1}^{M} \mathbf{1}\{t \in \mathcal{I}_m\}\alpha_m r_t(i) = \sum_{m=1}^{M}\alpha_m \sum_{t=1}^{T}\mathbf{1}\{t\in\mathcal{I}_m\} r_t(i) = \sum_{m=1}^{M}\alpha_m \mathcal{R}_{\mathcal{I}_m}(i)$$

$$= \mathcal{O}\left(\sum_{m=1}^{M}\alpha_m\sqrt{|\mathcal{I}_m|\ln(NT)}\right) \le \mathcal{O}\left(\sqrt{\sum_{m=1}^{M}\alpha_m}\sqrt{\sum_{m=1}^{M}\alpha_m|\mathcal{I}_m|\ln(NT)}\right) \quad \text{(Cauchy-Schwarz)}$$

$$= \mathcal{O}\left(\sqrt{\sum_{m=1}^{M}\alpha_m}\sqrt{\sum_{m=1}^{M}\alpha_m\sum_{t=1}^{T}\mathbf{1}\{t\in\mathcal{I}_m\}\ln(NT)}\right)$$

$$= \mathcal{O}\left(\sqrt{\sum_{m=1}^{M}\alpha_m}\sqrt{\sum_{t=1}^{T}u_t(i)\ln(NT)}\right).$$

Therefore, suppose we can pick α_m and $\mathcal{I}_m$ such that $\sum_{m=1}^{M}\alpha_m = \sum_{t=1}^{T}[u_t(i) - u_{t-1}(i)]_+$, then we prove the theorem since by applying Cauchy-Schwarz again

$$\mathcal{R}_T(u_1,\ldots,u_T) = \mathcal{O}\left(\sum_{i=1}^{N}\sqrt{\sum_{t=1}^{T}[u_t(i)-u_{t-1}(i)]_+}\sqrt{\sum_{t=1}^{T}u_t(i)\ln(NT)}\right)$$
$$\le \mathcal{O}\left(\sqrt{TA_T\ln(NT)}\right).$$

In the remainder of the proof, we show a recursive construction of α_m and $\mathcal{I}_m$ so that $\sum_{m=1}^{M}\alpha_m = \sum_{t=1}^{T}[u_t(i)-u_{t-1}(i)]_+$. For notation convenience we drop the index i. First we let $t^\star \in \operatorname{argmin}_t u_t$ and create an interval $\mathcal{I}_1 = [1, T]$ with weight $\alpha_1 = u_{t^\star}$. Then we recursively perform the same construction for the inputs $u_1 - u_{t^\star}, \ldots, u_{t^\star-1} - u_{t^\star}$ and $u_{t^\star+1} - u_{t^\star}, \ldots, u_T - u_{t^\star}$ respectively until there are no non-zero inputs left. Let $h(u_1,\ldots,u_T)$ denote the sum of the weights of the above construction. We use an induction (on the length of the input T) to prove $h(u_1,\ldots,u_T) = \sum_{t=1}^{T}[u_t - u_{t-1}]_+$. The base case $T=1$ holds trivially. Suppose the statement holds for any input length smaller than T. Then we have

$$h(u_1,\ldots,u_T) = u_{t^\star} + h(u_1 - u_{t^\star},\ldots,u_{t^\star-1}-u_{t^\star}) + h(u_{t^\star+1}-u_{t^\star},\ldots,u_T - u_{t^\star})$$
$$= u_{t^\star} + (u_1 - u_{t^\star}) + \sum_{t=2}^{t^\star-1}[u_t - u_{t-1}]_+ + (u_{t^\star+1}-u_{t^\star}) + \sum_{t=t^\star+2}^{T}[u_t - u_{t-1}]_+$$
$$= u_1 + \sum_{t=2}^{t^\star-1}[u_t - u_{t-1}]_+ + [u_{t^\star+1}-u_{t^\star}]_+ + \sum_{t=t^\star+2}^{T}[u_t-u_{t-1}]_+$$
$$= \sum_{t=1}^{T}[u_t - u_{t-1}]_+.$$

where the last step is by $[u_{t^\star} - u_{t^\star-1}]_+ = 0$. This finishes the proof. In fact, the above construction is *optimal* in minimizing $\sum_{m=1}^{M}\alpha_m$ and the proof can be found in [Luo and Schapire, 2015]. □

References

Omar Besbes, Yonatan Gur, and Assaf Zeevi. Non-stationary stochastic optimization. *Operations Research*, 63(5):1227–1244, 2015.

Haipeng Luo and Robert E. Schapire. Achieving All with No Parameters: AdaNormalHedge. In *28th Annual Conference on Learning Theory*, 2015.

Lijun Zhang, Tianbao Yang, Rong Jin, and Zhi-Hua Zhou. Strongly adaptive regret implies optimally dynamic regret. *arXiv preprint arXiv:1701.07570*, 2017.