TTIC 31290: Machine Learning for Algorithm Design (Fall 2025)
Avrim Blum and Dravyansh Sharma

Lecture 13: 11/13/25 Lecturer: Dravyansh Sharma

Outline for today

- Online Algorithm Design
- Dispersion and Regret Bounds

1 Online Algorithm Design

In the last lecture we saw some fundamental concepts related to online learning. We will now see how to model algorithm design as an online learning problem. In the statistical learning setting, input instances are i.i.d. and given to us all at once. Here we will view different algorithms as "experts". Given an online sequence of inputs, our goal will be to be competitive with the best algorithm in hindsight, on average for these inputs. The learner can keep changing the algorithm for the new inputs based on all the inputs it has seen so far.

Formal setup. We consider a sequence of rounds $t = 1, \ldots, T$. In round t, the learner selects an algorithm $A_t \in \mathcal{A}$ (say by selecting a parameter $\rho_t \in \mathcal{P} \subseteq \mathbb{R}^d$), and applies it to a new problem instance $x_t \in \Pi$. The performance of the algorithm on the instance x_t is given by a utility function $u_{x_t}(\rho_t)$[1] In the *full information* setting, the learner now observes $u_{x_t}(\rho)$ for all parameters ρ. Intuitively, given sufficient time/compute the learner can potentially compute u_{x_t} for all algorithms in the family before the next input is presented. More challenging learning settings include bandit and semi-bandit settings where we get partial feedback (e.g. just $u_{x_t}(\rho_t)$, or $u_{x_t}(\rho)$ for a subset of algorithms which we know have a similar behavior as ρ_t on x_t). We want to maximize the total utility across all rounds, and ideally be close in performance to the best algorithm in the family. Formally, we will seek to minimize the regret

$$R_T := \mathbb{E}\left[\max_{\rho \in \mathcal{P}} \sum_t u_{x_t}(\rho) - u_{x_t}(\rho_t)\right]$$

where the expectation is over the randomness in the learner's choices or over the randomness in the utility functions (e.g. the algorithms may have randomization).

We would like to obtain expected regret that is sublinear in T, since then the average (per-round) performance of the algorithm approaches that of the best parameter in hindsight (commonly referred as achieving "no regret" in online learning).

But this goal may be unattainable for typical algorithm families that we have seen in this class. Here we will see a simple example.

[1]In our previous terminology, we will directly work with the dual utility functions here.

Recall the family of greedy algorithms for the knapsack problem. A problem instance x consists of n items given by values $v_1, \ldots, v_n \in \mathbb{R}_{\geq 0}$ and sizes $s_1, \ldots, s_n \in \mathbb{R}_+$, and an overall knapsack capacity $C \in \mathbb{R}_+$ and we want to find the subset of items with largest total value for which the total size is at most C. The family of greedy algorithms select items in decreasing order of the score v_i/s_i^ρ where ρ is the parameter. Let the parameter space $\mathcal{P} = [0, 1]$.

Lemma 1. *Fix any $\rho^* \in (0, 1)$ and choose $\Delta > 0$. There exists knapsack instances x_{ρ^*} and x'_{ρ^*} such that the utility functions are piecewise constant with exactly two pieces over $[0, 1]$ and the gap between the utility in the two pieces is at least Δ. Moreover, the instance x_{ρ^*} has higher utility for $\rho < \rho^*$ but the instance x'_{ρ^*} has higher utility for $\rho > \rho^*$.*

Proof. First, we define the instance x_{ρ^*}. Set the knapsack capacity $C = 1$. There will be two items. A first item with size $s_1 = 1$ and value $v_1 = 1 + \Delta$ and a second (smaller) item with size $s_2 = \epsilon \in (0, 1)$ (to be set later) and value $v_2 = 1$. The critical point in the dual utility function is given by $\frac{v_1}{s_1^{\rho^*}} = \frac{v_2}{s_2^{\rho^*}}$, or $s_2 = s_1 \left(\frac{v_2}{v_1}\right)^{1/\rho^*}$, giving us $\epsilon = (1 + \Delta)^{-1/\rho^*}$. For $\rho < \rho^*$, we add the larger item and get the total value $1 + \Delta$. For $\rho > \rho^*$, we choose the smaller item and get a total value 1. For x'_{ρ^*}, we will have three items and again set the capacity $C = 1$. We have a larger item with $v_1 = 3\Delta$, $s_1 = 3/4$, and two smaller items that have $s_2 = s_3 = 1/2$ and $v_2 = 3\Delta\,(2/3)^{\rho^*}$. In this case, for $\rho < \rho^*$ we pick the larger item as $\frac{v_1}{s_1^\rho} = \frac{v_2}{s_2^\rho} \cdot (3/2)^{\rho^* - \rho} > \frac{v_2}{s_2^\rho}$ to get a total value 3Δ. Else, we are able to add both the smaller items, and get a value of $6\Delta\,(2/3)^{\rho^*}$. The difference is $3\Delta(2\,(2/3)^{\rho^*} - 1) \geq 3\Delta(4/3 - 1) = \Delta$. □

We can now use Lemma 1 to show that sub-linear regret is an unattainable goal in general for worst-case sequences of problem instances. The adversary will set instances x_t in round t by selecting a value ρ_t^* and presenting either the instance $x_{\rho_t^*}$ or $x'_{\rho_t^*}$ with equal probability. The sequence ρ_t^* will be chosen to guarantee that the learner's choices ρ_t have a large regret with respect to some ρ^*.

The adversary sets $\rho_1^* = 1/2$. Now, if $x_{1/2}$ was presented, we will set $\rho_2^* = 1/4$, else if $x'_{1/2}$ was presented we set it to $\rho_2^* = 3/4$. In general, for $t > 1$, if $x_{\rho_t^*}$ was presented we set $\rho_{t+1}^* = \rho_t^* - \frac{1}{2^t}$, and else set $\rho_{t+1}^* = \rho_t^* + \frac{1}{2^t}$. Since the adversary sets the next instances after the random selection for instance x_t, it is able to ensure that the online learner suffers a regret at least $(t\Delta)/2$ with respect to some ρ^* after t rounds. In other words, the regret of any online learner is at least $\Omega(T)$.

This motivates the need for further assumptions to attain sub-linear regret. In this lecture we will look at one such sufficient condition.

2 Dispersion and Regret Bounds

One thing to note about the above lower bound construction is that location of the discontinuities get exponentially closer to each other over the rounds, which means that, for large T, most instances will have their discontinuities in a small region in the parameter space. Motivated by this observation, here we will define a condition which will explicitly prevent this from happening.

Definition 1 (Dispersion)**.** *The sequence of utility functions $u_1, \ldots, u_T$ is β-dispersed for the Lipschitz constant L if, for all T and for all $\epsilon \geq T^{-\beta}$, at most $\tilde{O}(\epsilon T)$ functions (the soft-O notation* *suppresses dependence on logarithmic terms and quantities beside ϵ, T and β) are not L-Lipschitz in any ball of size ϵ contained in $\mathcal{P}$. Further if the utility functions are obtained from some distribution, the random process generating them is said to be β-dispersed if the above holds in expectation, i.e., if for all T and for all $\epsilon \geq T^{-\beta}$,*

$$\mathbb{E}\left[\max_{\rho \in \mathcal{P}} \left|\{t \mid u_t \text{ not L-Lipschitz in } \mathcal{B}(\rho, \epsilon)\}\right|\right] \leq \tilde{O}(\epsilon T),$$

where $\mathcal{B}(\rho, \epsilon)$ is the ball of radius ϵ centered at ρ in $\mathcal{P}$ (assumed to be metric space).

Note that the above definition implies that the sequence of utility functions have some randomization (this may be due to randomization in the problem instances, which is less strong than the iid assumption, or in the algorithm itself), and may not now be fully adversarial. We think of β as a measure of "niceness" of the sequence of functions, that takes values between 0 and 1. Any function sequence is 0-dispersed as $\beta = 0$ holds trivially.

We will now see how to get sublinear regret that depends on the dispersion coefficient β for any $\beta < 1$. The algorithm is a generalization of the randomized weighted majority to the continuous setting (which can be recovered when the utility functions have a finite domain and binary range). Formally, in round t, we sample a random ρ_t according to the distribution

$$p_t(\rho) \propto \exp\left(\lambda\left(\sum_{s=1}^{t-1} u_s(\rho)\right)\right).$$

In other words, each "expert" ρ has a weight $\exp\left(\lambda\left(\sum_{s=1}^{t-1} u_s(\rho)\right)\right)$ that we update according to the observed utility for that "expert" so far, and we select a random expert in proportion to its weight. For now, we will ignore the computational aspects related to sampling from this distribution. For this algorithm, we have the following regret guarantee.

Theorem 1. *Let $u_{x_1}, \ldots, u_{x_T} : \mathcal{P} \to [0, 1]$ be a sequence of utility functions corresponding to problem instances $x_1, \ldots, x_T$. Assume that the sequence is β-dispersed for Lipschitz constant L and the domain $\mathcal{P} \subset \mathbb{R}^d$ is bounded and contained in a ball of radius R. The continuous weighted majority algorithm with some appropriate λ has expected regret bounded by $\tilde{O}\left(\sqrt{Td} + T^{1-\beta}(L+1)\right)$.*

Proof Sketch. The total weight of all experts at the end of round t is given by

$$W_t = \int_{\mathcal{P}} \exp\left(\lambda\left(\sum_{s=1}^{t} u_s(\rho)\right)\right) d\rho.$$

We will give upper and lower bounds on W_T in terms of the payoff of the algorithm and the optimal payoff for any ρ respectively.

The upper bound follows a straightforward generalization of the classical argument.

$$W_T \leq W_1 \exp\left((e^\lambda - 1)\sum_{t=1}^{T} u_{x_t}(\rho_t)\right).$$

The lower bound uses dispersion. For any point $\rho \in \mathcal{B}(\rho^*, \epsilon)$ with $\epsilon = T^{-\beta}$, we have

$$\sum_{t=1}^{T} u_{x_t}(\rho) \geq \sum_{t=1}^{T} u_{x_t}(\rho^*) - \tilde{O}(\epsilon T) - L\epsilon T.$$

Thus,

$$\begin{aligned} W_T = \int_{\mathcal{P}} \exp\left(\lambda\left(\sum_{s=1}^{t} u_s(\rho)\right)\right) d\rho &\geq \int_{\mathcal{B}(\rho^*,\epsilon)} \exp\left(\lambda\left(\sum_{s=1}^{t} u_s(\rho)\right)\right) d\rho \\ &\geq \mathrm{Vol}(\mathcal{B}(\rho^*, \epsilon))\left(\sum_{t=1}^{T} u_{x_t}(\rho^*) - \tilde{O}(\epsilon T) - L\epsilon T\right) \\ &\geq (\epsilon/R)^d W_1 \left(\sum_{t=1}^{T} u_{x_t}(\rho^*) - \tilde{O}(\epsilon T) - L\epsilon T\right), \end{aligned}$$

where we have used that $W_1 \leq \mathrm{Vol}(\mathcal{B}(\rho', R))$ for some $\rho' \in \mathcal{P}$ by our boundedness assumption. Putting together gives an upper bound on the regret, and choosing λ to minimize the upper bound on the regret gives the desired upper bound. □

Therefore, if we find conditions under which the utility functions are dispersed, we can guarantee no-regret using the above result.

For the knapsack problem, we will show $\frac{1}{2}$-dispersion by making a mild assumption on the values, namely they come from a distribution with bounded density. We note that this condition is strongly linked with the notion of *smoothed analysis* mentioned earlier in the class. A special case of this is that we start with arbitrary adversarial values for the items, and then add some random noise (Gaussian noise is an example of a noise with bounded density).

Theorem 2. *Let $x_1, \ldots, x_T$ be any sequence of knapsack instances with n items and capacity 1, where instance x_i has sizes $s_1^{(i)}, \ldots, s_n^{(i)} \in [1, C]$ and values $v_1^{(i)}, \ldots, v_n^{(i)} \in (0, 1]$. Assume that all the item values are come from a distribution with probability density bounded by b*[2]*. Then the utility functions $u_{x_1}, \ldots, u_{x_T}$ are $\frac{1}{2}$-dispersed.*

Proof Sketch. The overall strategy is to first show bound the expected number of discontinuities in any interval of width 2ϵ (ball of radius ϵ), which also gives a bound on the expected number of functions that have a discontinuity in any interval of width 2ϵ, and then show that the maximum number of discontinuities over any interval of width 2ϵ cannot be much larger than this expected number (with high probability, given sufficiently large T).

Let $c_{ij}^{(t)} = \log(v_i^{(t)}/v_j^{(t)})/\log(s_i^{(t)}/s_j^{(t)})$ be the critical parameter value such that at $\rho = c_{ij}^{(t)}$, items i and j swap their relative order in the t-th instance. Balcan et al. (FOCS, 2018) show that each critical value $c_{ij}^{(t)}$ is a random variable with a density function bounded by $b^2 \ln(C)/2$. Thus, for any interval $I = [\rho - \epsilon, \rho + \epsilon]$ of radius ϵ, the expected total number of critical values $c_{ij}^{(t)}$ summed

[2]For the Gaussian distribution with variance σ^2, $b = \frac{1}{\sqrt{2\pi\sigma^2}}$.

over all pairs of items and instances is at most $\epsilon T n^2 b^2 ln(C)$. This is also an upper bound on the expected number of utility functions in $u_{x_1}, \ldots, u_{x_T}$ that have a discontinuity on I.

To complete the second part of the proof, we use the following result due to Balcan et al. (UAI 2020).

Lemma 2. *Let $u_1, \ldots, u_T : \mathbb{R} \to \mathbb{R}$ be independent piecewise L-Lipschitz functions, each having at most K discontinuities. Let $D(T, \epsilon, \rho) = |\{1 \leq t \leq T : u_t$ is not L-Lipschitz on $[\rho - \epsilon, \rho + \epsilon]\}|$ be the number of functions that are not L-Lipschitz on the ball $[\rho - \epsilon, \rho + \epsilon]$. Then we have $\mathbb{E}[\max_{\rho \in \mathbb{R}} D(T, \epsilon, \rho)] \leq \max_{\rho \in \mathbb{R}} \mathbb{E}[D(T, \epsilon, \rho)] + O(\sqrt{T \log(TK)})$.*

Thus, the expected maximum number of non-Lipschitz functions in any ball of radius ϵ is at most $\epsilon T n^2 b^2 ln(C) + O(\sqrt{T \log(Tn^2)}) = \tilde{O}(\epsilon T + \sqrt{T})$. This implies $\frac{1}{2}$-dispersion. □

Additional Resources:

- Maria-Florina Balcan, Travis Dick, and Ellen Vitercik. "Dispersion for data-driven algorithm design, online learning, and private optimization." In 2018 IEEE 59th Annual Symposium on Foundations of Computer Science (FOCS), pp. 603-614. IEEE, 2018.
- Maria-Florina Balcan, Travis Dick, and Wesley Pegden. "Semi-bandit optimization in the dispersed setting." In Conference on Uncertainty in Artificial Intelligence (UAI), pp. 909-918. PMLR, 2020.
- Maria-Florina Balcan, Travis Dick, and Dravyansh Sharma. "Learning piecewise Lipschitz functions in changing environments." In International Conference on Artificial Intelligence and Statistics (AISTATS), pp. 3567-3577. PMLR, 2020.