Lecture 06: Regret Matching and Blackwell Approachbility

Lecturer: John Lazarsfeld

October 02, 2025

Abstract

Regret Matching (RM) and Regret Matching+ (RM+) algorithms; Blackwell's Approchability theorem.

1 Regret Matching Algorithms

Overview and Intutition. In previous lectures we have seen several families of general algorithms (FTRL and OMD) for no-regret online learning. At a high-level, these algorithms obtain their regret guarantees by using a carefully-tuned learning rate to balance the exploration/exploitation tradeoff. In particular, recall that the FTRL algorithm essentially balances the *greedy* strategy of choosing the next action whose cumulative loss through the previous round is minimial with an additional regularization term that ensures stability in the iterates. Roughly speaking, this strategy is somewhat equivalent to selection actions proportionally to their (coordinate-wise) cumulative regret through the most recent round.

It turns out that implementing this strategy directly recovers a classical online learning algorithm known as Regret Matching (RM) that is of the most commonly used online algorithms in practice. The goal of this lecture is to introduce and analyze this algorithm, as well as to point out its connection to Blackwell approachability.[1]

Notation. For $u \in \mathbb{R}^n$ we write $[u]^+ \in \mathbb{R}^n$ to denote the vector with $[u]^+(i) = \max(0, u(i))$.

RM Algorithm. We now formally state the RM algorithm and its regret guarantee:

Algorithm 1 Regret-Matching Algorithm (RM) for Experts Setting

Input: Initial $x_1 \in \Delta_n$;
for $t = 1, \dots, T$ **do**:

1. Play action $x_t \in \Delta_n$, and incur cost $f_t(x_t) = \langle x_t, \ell_t \rangle$. Observe loss vector $\ell_t \in \mathbb{R}^n$.
2. Construct the instantaneous regret vector $r_t \in \mathbb{R}^n$ given by
$$r_t = \sum_{k=1}^{t} \langle \ell_k, x_k \rangle \mathbf{1} - \ell_k \ .$$
3. If $r_t = 0 \in \mathbb{R}^n$, set $x_{t+1} \in \Delta_n$ arbitrarily. Otherwise update:
$$x_{t+1} = \frac{[r_t]^+}{\|[r_t]^+\|_1} \ . \tag{1}$$

end for

We have the following regret guarantee for the RM algorithm in the experts setting:

Theorem 1. *Let $\{x_t\}$ be the iterates of RM (Algorithm 1) in the experts setting with loss vectors $\{\ell_t\}$. Suppose each $\ell_t \in [-1, 1]^n$. Then*
$$\mathrm{Reg}_{RM}(T) \le \sqrt{Tn} \ .$$

[1] The content of these notes roughly follows those of Farina (2021).

Remark 2. We make several remarks about the RM algorithm and its regret guarantee:

- **RM is a parameter-free algorithm**: In contrast to our previously-inroduced algorithms like FTRL and OMD, the RM algorithm has no parameter like a stepsize learning rate. Thus, this *parameter-free* algorithm is more easily implementable and requires no careful setting of a stepsize in order to obtain its regret guarantee.
- **Dimension dependence in regret guarantee**: On the other hand, observe that the bound on $\text{Reg}_{BM}(T)$ in the theorem has a $\sqrt{n}$ dependence, which is significantly larger (at least in theory) than the $\log n$ dependence for, e.g., the MWU algorithm. Thus, the price of being parameter-free is paid in the form of this worse dependence in the worst-case regret bound.

Proof of Theorem 1. Our goal will be to track the change in $\|[r_t]^+\|_2^2$ over time. For this, observe for each $t+1$ that

$$\begin{aligned}\Big\langle \langle \ell_{t+1}, x_{t+1}\rangle \mathbf{1} - \ell_{t+1}, x_{t+1} \Big\rangle &= \langle \ell_{t+1}, x_{t+1}\rangle \cdot \langle \mathbf{1}, x_{t+1}\rangle - \langle \ell_{t+1}, x_{t+1}\rangle \\ &= \langle \ell_{t+1}, x_{t+1}\rangle - \langle \ell_{t+1}, x_{t+1}\rangle = 0 \,, \end{aligned} \tag{2}$$

where the second equality follows from the fact that $\langle \mathbf{1}, x_{t+1}\rangle = 1$ since $x_{t+1} \in \Delta_n$.

Moreover, assuming $r_{t+1} \neq 0$, then we have by the update rule (1) of the RM algorithm that $x_{t+1} = \frac{[r_t]^+}{\|[r_t]^+\|_1}$. It follows from (2) that

$$\Big\langle \langle \ell_{t+1}, x_{t+1}\rangle \mathbf{1} - \ell_{t+1}, x_{t+1} \Big\rangle = \Big\langle \langle \ell_{t+1}, x_{t+1}\rangle \mathbf{1} - \ell_{t+1}, \frac{[r_t]^+}{\|[r_t]^+\|_1} \Big\rangle = 0 \,,$$

which further implies that

$$\Big\langle \langle \ell_{t+1}, x_{t+1}\rangle \mathbf{1} - \ell_{t+1}, [r_t]^+ \Big\rangle = 0 \,. \tag{3}$$

Now recall by definition of r_t that $r_{t+1} = r_t + (\langle x_{t+1}, \ell_{t+1}\rangle \mathbf{1} - \ell_{t+1})$. Then using the identity $\|[a+b]^+\|_2^2 \leq \|[a]^+ + b\|_2^2$, it follows that

$$\begin{aligned}\|[r_{t+1}]^+\|_2^2 &\leq \|[r_t]^+ + (\langle x_{t+1}, \ell_{t+1}\rangle \mathbf{1} - \ell_{t+1})\|_2^2 &(4)\\ &= \|[r_t]^+\|_2^2 + 2\Big\langle \langle x_{t+1}, \ell_{t+1}\rangle \mathbf{1} - \ell_{t+1}, [r_t]^+ \Big\rangle + \|\langle x_{t+1}, \ell_{t+1}\rangle \mathbf{1} - \ell_{t+1}\|_2^2 &(5)\\ &= \|[r_t]^+\|_2^2 + \|\langle x_{t+1}, \ell_{t+1}\rangle \mathbf{1} - \ell_{t+1}\|_2^2 &(6)\end{aligned}$$

where the final equality follows from (3). Iterating on this inequality, we find that

$$\|[r_T]^+\|_2^2 \leq \sum_{t=1}^{T} \|\langle x_t, \ell_t\rangle \mathbf{1} - \ell_t\|_2^2 \leq \sum_{t=1}^{T} \|\ell_t\|_2^2 \leq \sum_{t=1}^{T} n\|\ell_t\|_\infty^2 \leq Tn \,. \tag{7}$$

Finally, observe by definition of $\text{Reg}_{RM}(T)$ and r_T that

$$\text{Reg}_{RM}(T) = \max_{i\in[n]} r_T(i) \leq \max_{i\in[n]} [r_T(i)]^+ = \|[r_T]^+\|_\infty \leq \|[r_T]^+\|_2 \leq \sqrt{Tn} \,,$$

where in the final equality we apply the bound from (7). □

Regret Matching+ Algorithm. A more modern variant of the Regret Matching algorithm is Regret Matching+ (RM+), which adds an additional thresholding to the instantaneous regret term r_t. In particular, the RM+ algorithm instead defines

$$r_t = [r_{t-1} + \langle x_t, \ell_t\rangle \mathbf{1} - \ell_t]^+ \,,$$

where it is assumed $r_0 = 0 \in \mathbb{R}^n$. The algorithm then proceeds as in RM, and selects $x_{t+1} = \frac{[r_t]^+}{\|[r_t]^+\|_1}$. It can be shown that RM+ achieves a similar regret bound as RM in Theorem 1. See Farina (2021) for more details and discussion.

2 Blackwell Approachability

Overview. We now introduce the notion of *Blackwell Approachability* (Blackwell, 1956), which is among the earliest formalizations of external regret minimization. As we will see, Blackwell Approachability is intimately connected to the Regret Matching algorithm.

Blackwell's Approachability Game. A Blackwell Approachability game is specified by a tuple $(\mathcal{X}, \mathcal{Y}, u, S)$, where $\mathcal{X}$ and $\mathcal{Y}$ are closed, convex sets, and $u : \mathcal{X} \times \mathcal{Y} \to \mathbb{R}^n$ is a vector-valued biaffine function, and $S \subseteq \mathbb{R}^n$ is a closed, convex *target set*. The game proceeds between two players – a learner and an adversary – as follows:

1. The learner selects $x_t \in \mathcal{X}$.
2. The adversary selects $y_t \in \mathcal{Y}$, where y_t may depend on the entire history $\{x_t\}$.
3. The learner incurs loss $u(x_t, y_t) \in \mathbb{R}^n$.

The goal of the learner is to guarantee the average incurred loss converges to the target set S:

$$\min_{s \in S} \left\| s - \frac{1}{T} \sum_{t=1}^{T} u(x_t, y_t) \right\|_2 \to 0 \quad \text{as } T \to \infty. \tag{8}$$

Blackwell Approachability and Regret Minimization. Consider an instance of a Blackwell Approachability game $(\Delta_n, \mathbb{R}^n, u, \mathbb{R}^n_{\leq 0})$ where $u : \Delta_n \times \mathbb{R}^n \to \mathbb{R}^n$ is given by

$$u(x_t, \ell_t) = \langle \ell_t, x_t \rangle \mathbf{1} - \ell_t \ . \tag{9}$$

Moreover, define $R(T) = \min_{x \in \Delta_n} \sum_{t=1}^{T} \langle \ell_t, x_t - x \rangle$. Then the following holds:

$$\frac{R(T)}{T} \leq \min_{s \in \mathbb{R}^n_{\leq 0}} \left\| s - \frac{1}{T} \sum_{t=1}^{T} u(x_t, \ell_t) \right\|_2 . \tag{10}$$

In other words, a successful strategy for choosing the sequence $\{x_t\}$ and achieving the goal in (8) is also a successful strategy for achieving sublinear bounds on $R(T)$.

In particular, *Blackwell's Algorithm* presents a constructive way for achieving the goal in (8) for genearl Blackwell games. In the case of the instance $(\Delta_n, \mathbb{R}^n, u, \mathbb{R}^n_{\leq 0})$ described above, this algorithm reduces exactly to the Regret Matching algorithm. We defer details of this proof and relationship to Farina (2021) and Blackwell (1956).

References

David Blackwell. An analog of the minimax theorem for vector payoffs. 1956.

Gabriele Farina. Blackwell approachability and regret minimization on simplex domains, 2021. URL `https://www.mit.edu/~gfarina/2021/15888f21_L04_blackwell_rm/L04_blackwell_rm.pdf`.