DOI: https://doi.org/10.1070/RM9961

Online algorithm for aggregating experts' predictions with unbounded quadratic loss

A. A. Korotin, V. V. V'yugin, and E. V. Burnaev

We consider the problem of online aggregation of experts' predictions with a quadratic loss function. At the beginning of each round $t = 1, 2, \dots, T$, experts $n = 1, \dots, N$ provide predictions $\gamma_t^1, \dots, \gamma_t^N \in \mathbb{H}$ (where $\mathbb{H}$ is a Hilbert space). The player aggregates the predictions to a single prediction $\overline{\gamma_t} \in \mathbb{H}$. Then nature provides the true outcome $\omega \in \mathbb{H}$. The player and the experts $n = 1, \dots, N$ suffer the losses $h_t = \|\omega - \overline{\gamma_t}\|^2$ and $l_t^n = \|\omega - \gamma_t^n\|^2$, respectively, and the next round $t+1$ begins. The goal of the player is to minimize the regret, that is, the difference between the total loss of the player and the loss of the best expert: $R_T = \sum_{t=1}^T h_t - \min_{n=1,\dots,N} \sum_{t=1}^T l_t^n$.

Online regression is a widespread special case of this problem: the γ_t^n and ω_t are real numbers (the predictions and the outcome) and $\mathbb{H} = \mathbb{R}$ ([1], § 2.1). A more general case is the functional or probabilistic forecasting, for example, where the γ_t^n and ω_t are probability densities, that is, elements of $\mathbb{H} = \mathscr{L}^2(\mathbb{R}^D)$ (see [4]).

The problem of online prediction with experts' advice is considered in game theory, machine learning [1], and online optimization [3]. Existing aggregating algorithms provide strategies which guarantee a constant upper bound on the regret but assume that the losses are bounded. For example, if $l_t^n \leqslant B^2$ for all t and n, then the algorithm in § 2.1 of [1] guarantees a T-independent bound $R_T \leqslant O(B^2 \log N)$. However, the algorithm requires knowing B beforehand.

In this paper, we propose an algorithm for aggregating experts' predictions which does not require a prior knowledge of the upper bound on the losses. The algorithm is based on the exponential reweighing of experts' losses.

Parameters: game length T, number of experts N, Hilbert space $\mathbb{H}$.
$B_0^\dagger \leftarrow 0$; $L_0^n \leftarrow 0$ for $n = 1, \dots, N$;
for $t = 1, 2, \dots, T$ **do**
  Experts $n = 1, \dots, N$ provide predictions $\gamma_t^1, \dots, \gamma_t^N \in \mathbb{H}$;
  $B_t \leftarrow \max\left(B_{t-1}^\dagger, \max_{n,n'} \|\gamma_t^n - \gamma_t^{n'}\|\right)$; $\eta_t \leftarrow 1/[2(B_t)^2]$;
  $w_t^n \leftarrow \exp(-\eta_t L_{t-1}^n) / \sum_{n'=1}^N \exp(-\eta_t L_{t-1}^{n'})$;
  Player combines the predictions $\overline{\gamma_t} \leftarrow \sum_{n=1}^N w_t^n \cdot \gamma_t^n \in \mathbb{H}$;
  Nature reveals the true outcome $\omega_t \in \mathbb{H}$;
  Player and experts suffer losses $h_t = \|\omega_t - \overline{\gamma_t}\|^2$ and $l_t^n = \|\omega_t - \gamma_t^n\|^2$;
  $B_t^\dagger \leftarrow B_t$; $L_t^n \leftarrow L_{t-1}^n + l_t^n$ for $n = 1, \dots, N$;
  if $\max_n \sqrt{l_t^n} > B_t^\dagger$ **then**
    $B_t^\dagger \leftarrow \sqrt{2} \max_n \sqrt{l_t^n}$;

Algorithm 1: player's strategy when the bound on the losses is not known beforehand.

This work was partially supported by the Ministry of Science and Higher Education of the Russian Federation (grant no. 14.756.31.0001).

AMS 2020 *Mathematics Subject Classification*. Primary 68T05, 90C25.

The proposed algorithm assigns weights to the experts proportionally to the inverse exponential of their cumulative losses from the previous steps. The weights are used to perform linear (convex) aggregation of predictions. The learning rate (the parameter η of the algorithm) changes dynamically, enabling the algorithm to adapt to the maximum of the observed losses.

Theorem 1. *The regret admits the estimate* $R_T \leqslant O\big(\max_{t,n} l_t^n \cdot (\log N + 1)\big)$ *for Algorithm* 1.

Proof. Consider the step t. Let $\mathbb{H}_t = \operatorname{Span}\{\omega_t, \gamma_t^1, \dots, \gamma_t^N\}$, a ($\leqslant N+1$)-dimensional linear subspace of $\mathbb{H}$, and let

$$S_t = \bigcap_{n=1}^{N} \{\gamma \in \mathbb{H}_t \colon \|\gamma - \gamma_t^n\| \leqslant B_t\},$$

that is, S_t is the convex set of points $\gamma \in \mathbb{H}_t$ which are ($\leqslant B_t$)-close to all the predictions $\gamma_t^1, \dots, \gamma_t^N$. Let $\Gamma_t = \operatorname{ConvexHull}\{\gamma_t^1, \dots, \gamma_t^N\}$. Note that $\Gamma_t \subset S_t$. (This follows from the fact that $B_t \geqslant \max_{n,n'} \|\gamma_t^n - \gamma_t^{n'}\|$ by the definition of B_t in Algorithm 1.)

Let $\lfloor \omega_t \rfloor$ be the projection of ω_t on S_t. We define $f_t \colon \Gamma_t \to \mathbb{R}$ by $f_t(\gamma) = \|\gamma - \lfloor \omega_t \rfloor\|^2$. For any $\gamma \in \Gamma_t$ we have $\|\gamma - \lfloor \omega_t \rfloor\| \leqslant B_t$, so by Lemma 4.2 in [3], we conclude that f_t is an ($\eta_t = 1/[2(B_t)^2]$)-exponentially concave function. Thus,

$$\exp(-\eta_t f(\overline{\gamma_t})) \geqslant \sum_{n=1}^{N} w_t^n \exp(-\eta_t f(\gamma_t^n)),$$

where $\overline{\gamma_t}$ is the aggregated prediction of the player. We denote $f_t(\overline{\gamma_t}) = \|\overline{\gamma_t} - \lfloor \omega_t \rfloor\|^2$ by $\lfloor h_t \rfloor$ and $f(\gamma_t^n) = \|\gamma_t^n - \lfloor \omega_t \rfloor\|^2$ by $\lfloor l_t^n \rfloor$, and we obtain the inequality

$$\exp(-\eta_t \lfloor h_t \rfloor) \geqslant \sum_{n=1}^{N} w_t^n \exp(-\eta_t \lfloor l_t^n \rfloor).$$

For $\eta > 0$ we define

$$\lfloor m_t \rfloor(\eta) = -\eta^{-1} \log \sum_{n=1}^{N} w_t^n \exp(-\eta \lfloor l_t^n \rfloor).$$

To begin with, it follows from the previous paragraph that $\lfloor h_t \rfloor \leqslant \lfloor m_t \rfloor(\eta_t)$. Next, for $\eta > 0$ we define

$$m_t(\eta) = -\eta^{-1} \log \sum_{n=1}^{N} w_t^n \exp(-\eta l_t^n).$$

Since $\lfloor \omega_t \rfloor$ is the projection of ω_t on the convex set S_t, it follows that $\lfloor l_t^n \rfloor = \|\gamma_t^n - \lfloor \omega_t \rfloor\|^2 \leqslant \|\gamma_t^n - \omega_t\|^2 = l_t^n$. Thus, for any $\eta > 0$ we have $m_t(\eta) \geqslant \lfloor m_t \rfloor(\eta)$. In particular, for $\eta = \eta_t$ we have $\lfloor h_t \rfloor \leqslant \lfloor m_t \rfloor(\eta_t) \leqslant m_t(\eta_t)$.

Let us prove that $h_t \leqslant \lfloor h_t \rfloor + (B_t^\dagger)^2 - (B_t)^2$. If $\omega_t \in S_t$, then $\omega_t = \lfloor \omega_t \rfloor$, $h_t = \lfloor h_t \rfloor$, and $B_t^\dagger = B_t$, which results in the desired inequality. But if $\omega_t \notin S_t$, then $\max_n \sqrt{l_t^n} > B_t$ and $(B_t^\dagger)^2 = 2\max_n l_t^n > \max_n l_t^n + (B_t)^2$ by the definition in Algorithm 1. Thus, $\lfloor h_t \rfloor + (B_t^\dagger)^2 - (B_t)^2 > \lfloor h_t \rfloor + \max_n l_t^n \geqslant \max_n l_t^n \geqslant h_t$, where the last inequality follows from the convexity of the quadratic function f_t and the fact that $\overline{\gamma_t} \in \Gamma_t$.

We combine the derived inequalities and conclude that

$$h_t \leqslant m_t(\eta_t) + (B_t^\dagger)^2 - (B_t)^2.$$

Summing over $t = 1, 2, \dots, T$, we get that

$$\sum_{t=1}^{T} h_t \leqslant (B_T^\dagger)^2 + \sum_{t=1}^{T} m_t(\eta_t).$$

Note that $\eta_1 \geqslant \eta_2 \geqslant \cdots \geqslant \eta_T$ is a non-increasing dynamic learning rate. By Lemma 2 in [2],

$$\sum_{t=1}^{T} m_t(\eta_t) \leqslant -\eta_T^{-1} \log \sum_{t=1}^{T} N^{-1} \exp(-\eta_T L_T^n).$$

The latter quantity does not exceed

$$-\eta_T^{-1} \log \bigl(N^{-1} \exp(-\eta_T \min_n L_T^n)\bigr) = \eta_T^{-1} \log N + \min_n L_T^n.$$

From this we immediately obtain a regret bound for our algorithm:

$$\sum_{t=1}^{T} h_t - \min_n L_T^n \leqslant \eta_T^{-1} \log N + (B_T^\dagger)^2 = 2(B_t)^2 \log N + (B_t^\dagger)^2 \leqslant (2\log N + 1)(B_t^\dagger)^2.$$

Note that

$$B_t^\dagger \leqslant \max\bigl(\max_{t,n} \|\gamma_t^n - \gamma_t^{n'}\|, \sqrt{2} \max_{t,n} \sqrt{l_t^n}\bigr).$$

By the triangle inequality,

$$\|\gamma_t^n - \gamma_t^{n'}\| \leqslant \|\gamma_t^n - \omega_t\| + \|\gamma_t^{n'} - \omega_t\| = \sqrt{l_t^n} + \sqrt{l_t^{n'}} \leqslant 2\max_n \sqrt{l_t^n}$$

for any t, n, and n'. Thus, $B_T^\dagger \leqslant 2\max_{n,t} \sqrt{l_t^n}$, and the final regret bound is

$$\sum_{t=1}^{T} h_t - \min_n L_T^n \leqslant 4(2\log N + 1) \max_{t,n} l_t^n = O(\max_{t,n} l_t^n \cdot (\log N + 1)).$$

Bibliography

[1] N. Cesa-Bianchi and G. Lugosi, *Prediction, learning, and games*, Cambridge Univ. Press, Cambridge 2006, xii+394 pp.

[2] S. De Rooij, T. van Erven, P. D. Grünwald, and W. M. Koolen, *J. Mach. Learn. Res.* **15** (2014), 1281–1316.

[3] E. Hazan, *Found. Trends Optim.* **2**:3-4 (2016), 157–325.

[4] A. Korotin, V. V'yugin, and E. Burnaev, *Integral mixability: a tool for efficient online aggregation of functional and probabilistic forecasts*, 2020 (v1 – 2019), 22 pp., arXiv: 1912.07048.

Alexander A. Korotin
Skolkovo Institute of Science and Technology
E-mail: a.korotin@skoltech.ru

Vladimir V. V'yugin
Skolkovo Institute of Science and Technology
E-mail: v.vyugin@skoltech.ru

Evgenii V. Burnaev
Skolkovo Institute of Science and Technology
E-mail: e.burnaev@skoltech.ru

Presented by L. D. Beklemishev
Accepted 10/JUL/20
Translated by THE AUTHORS