Complementi di Algoritmi e Strutture Dati

Ottimizzazione del prezzo base in aste al secondo prezzo

Docente: *Nicolò Cesa-Bianchi* versione del 14 giugno 2021

In online advertising, publishers sell their online ad space to advertisers through second-price auctions managed by ad exchanges. For each impression (ad display) created on the publisher's website, the ad exchange runs an auction on the fly. Empirical evidence shows that an informed choice of the seller's reserve price, disqualifying any bid below it, can indeed have a significant impact on the revenue of the seller. We assume the seller is also observing the highest bid together with the revenue. This richer feedback, which is key to proving the results in these notes, is made available by some ad exchanges such as AppNexus.

The seller's revenue in a second-price auction is computed as follows: if the reserve price r is not larger than the second-highest bid $b(2)$, then the item is sold to the highest bidder and the seller's revenue is $b(2)$. If r is between $b(2)$ and the highest bid $b(1)$, then the item is sold to the highest bidder and the seller's revenue is the reserve price. Finally, if r is bigger than $b(1)$, then the item is not sold and the seller's revenue is zero. Formally, the seller's revenue is

$$g\big(r, b(1), b(2)\big) = \max\big\{r, b(2)\big\}\mathbb{I}\{r \le b(1)\}$$

Note that the revenue only depends on the reserve price r and on the two highest bids $b(1) \ge b(2)$, which —by assumption— all belong to the unit interval $[0,1]$.

At the beginning of each auction $t = 1, 2, \ldots$, the seller computes a reserve price $r_t \in [0,1]$. Then, bids $b_t(1), b_t(2)$ are collected by the auctioneer, and the seller (which is not the same as the auctioneer) observes his revenue $g_t(r_t) = g\big(r_t, b_t(1), b_t(2)\big)$, together with the highest bid $b_t(1)$. Crucially, knowing $g_t(r_t)$ and $b_t(1)$ allows to compute $g_t(r)$ for all $r \ge r_t$. For technical reasons, we use losses $\ell_t(r_t) = 1 - g_t(r_t)$ instead of revenues, see Figure 1 for a pictorial representation.

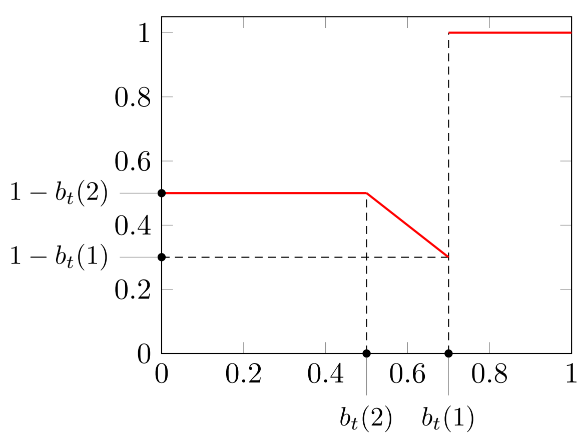

Figura 1: The loss function $\ell_t(r_t) = 1 - \max\{r_t, b_t(2)\}\mathbb{I}\{r_t \le b_t(1)\}$ when $b_t(1) = 0.7$ and $b_t(2) = 0.5$.

The loss functions $\ell_t : [0,1] \to [0,1]$ satisfy the semi-Lipschitz condition,

$$\ell_t(y + \delta) \ge \ell_t(y) - \delta \qquad \text{for all } 0 \le y \le y + \delta \le 1. \tag{1}$$

The learner's regret is defined by

$$R_T = \mathbb{E}\left[\sum_{t=1}^{T} \ell_t(r_t)\right] - \inf_{0 \le y \le 1} \sum_{t=1}^{T} \ell_t(y) \ ,$$

where the expectation is with respect to the randomness in the reserves r_t. We introduce the Exp3-RTB algorithm, a variant of Exp3 exploiting the richer feedback $\big\{\ell_t(r) \,:\, y \ge r_t\big\}$. The algorithm uses a discretization of the action space $[0,1]$ in $K = \lceil 1/\gamma \rceil$ actions $y_k := (k-1)\gamma$ for $k = 1, \dots, K$.

Algorithm 1 (Exp3-RTB)

Input: Exploration parameter $\gamma > 0$.

Set learning rate $\eta = \gamma/2$ and uniform distribution p_1 over $\{1, \dots, K\}$ where $K = \lceil 1/\gamma \rceil$
for $t = 1, 2, \dots$ **do**
  compute distribution $q_t(k) = (1-\gamma)p_t(k) + \gamma \mathbb{I}\{k = 1\}$ for $k = 1, \dots, K$;
  draw $I_t \sim q_t$ and choose $r_t = y_{I_t} = (I_t - 1)\gamma$;
  for each $k = 1, \dots, K$, compute the estimated loss

$$\widehat{\ell}_t(k) = \frac{\ell_t(y_k)}{\sum_{j=1}^{k} q_t(j)} \mathbb{I}\{I_t \le k\}$$

  for each $k = 1, \dots, K$, compute the new probability assignment

$$p_{t+1}(k) = \frac{\exp\big(-\eta \sum_{s=1}^{t} \widehat{\ell}_s(k)\big)}{\sum_{j=1}^{K} \exp\Big(-\eta \sum_{s=1}^{t} \widehat{\ell}_s(j)\Big)}$$

end for

Teorema 1 *The Exp3-RTB algorithm tuned with $\gamma > 0$ satisfies*

$$R_T \le \gamma T \left(2 + \frac{1}{4} \ln \frac{e}{\gamma}\right) + \frac{2 \ln \lceil 1/\gamma \rceil}{\gamma} \ .$$

In particular, $\gamma = T^{-1/2}$ gives $R_T = \widetilde{\mathcal{O}}\big(\sqrt{T}\big)$.

Dimostrazione. The proof follows the same lines as the regret analysis of Exp3. The key change is a tighter control of the variance term allowed by the richer feedback.

Pick any reserve price $y_k = (k-1)\gamma$. We first control the regret associated with actions drawn from p_t (the regret associated with q_t will be studied as a direct consequence). More precisely, since the estimated losses $\widehat{\ell}_t(j)$ are nonnegative, we can apply the standard analysis of Exp3 to get

$$\sum_{t=1}^{T} \sum_{i=1}^{K} p_t(i) \widehat{\ell}_t(i) - \sum_{t=1}^{T} \widehat{\ell}_t(k) \le \frac{\eta}{2} \sum_{t=1}^{T} \sum_{j=1}^{K} p_t(j) \widehat{\ell}_t(j)^2 + \frac{\ln K}{\eta} \tag{2}$$

Writing $\mathbb{E}_{t-1}[\cdot]$ for the expectation conditioned on $I_1, \dots, I_{t-1}$, we note that

$$\mathbb{E}_{t-1}\Big[\widehat{\ell}_t(j)\Big] = \ell_t(y_j) \quad \text{and} \quad \mathbb{E}_{t-1}\Big[p_t(j)\widehat{\ell}_t(j)^2\Big] = \frac{p_t(j)\ell_t(y_j)^2}{\sum_{i=1}^{j} q_t(i)} \le \frac{q_t(j)}{(1-\gamma)\sum_{i=1}^{j} q_t(i)}$$

where we used the definition of q_t and the fact that $|\ell_t(y_j)| \le 1$ by assumption. Therefore, taking expectation on both sides of (2) implies, again similarly to what is done in the analysis of Exp3,

$$\mathbb{E}\left[\sum_{t=1}^{T}\sum_{i=1}^{K} p_t(i)\ell_t(i)\right] - \sum_{t=1}^{T}\ell_t(y_k) \le \frac{\eta}{2(1-\gamma)}\sum_{t=1}^{T}\mathbb{E}\left[\sum_{j=1}^{K}\frac{q_t(j)}{\sum_{i=1}^{j} q_t(i)}\right] + \frac{\ln K}{\eta}$$

Setting $s_t(j) = \sum_{i=1}^{j} q_t(i)$, we can upper bound the sum with an integral,

$$\begin{aligned}\sum_{j=1}^{K}\frac{q_t(j)}{\sum_{i=1}^{j} q_t(i)} &= 1 + \sum_{j=2}^{K}\frac{s_t(j)-s_t(j-1)}{s_t(j)} = 1 + \sum_{j=2}^{K}\int_{s_t(j-1)}^{s_t(j)}\frac{\mathrm{d}x}{s_t(j)} \\ &\le 1 + \sum_{j=2}^{K}\int_{s_t(j-1)}^{s_t(j)}\frac{\mathrm{d}x}{x} = 1 + \int_{q_t(1)}^{1}\frac{\mathrm{d}x}{x} \le 1 - \ln q_t(1) \le 1 + \ln\frac{1}{\gamma}\end{aligned}$$

where we used $q_t(1) \ge \gamma$. Therefore, substituting into the previous bound, we get

$$\mathbb{E}\left[\sum_{t=1}^{T}\sum_{i=1}^{K} p_t(i)\ell_t(i)\right] - \sum_{t=1}^{T}\ell_t(y_k) \le \frac{\eta T\ln(e/\gamma)}{2(1-\gamma)} + \frac{\ln K}{\eta} \tag{3}$$

We now control the regret of the reserves $r_t = y_{I_t}$, where I_t is drawn from $q_t = (1-\gamma)p_t + \gamma\delta_1$. We have

$$\begin{aligned}\mathbb{E}\left[\sum_{t=1}^{T}\ell_t(r_t)\right] - \sum_{t=1}^{T}\ell_t(y_k) &= \mathbb{E}\left[\sum_{t=1}^{T}\left((1-\gamma)\sum_{i=1}^{K}p_t(i)\ell_t(i) + \gamma\ell_t(y_1)\right)\right] - \sum_{t=1}^{T}\ell_t(y_k) \\ &\le (1-\gamma)\,\mathbb{E}\left[\sum_{t=1}^{T}\sum_{i=1}^{K}p_t(i)\ell_t(i)\right] + \gamma T - \sum_{t=1}^{T}\ell_t(y_k) \\ &\le \frac{\eta T\ln(e/\gamma)}{2} + \frac{\ln K}{\eta} + \gamma T \end{aligned} \tag{4}$$

where the last inequality is by (3).

To conclude the proof, we upper bound the regret against any fixed $y \in [0,1]$. Since there exists $k \in \{1,\dots,K\}$ such that $y \in [y_k, y_k+\gamma]$, and since each ℓ_t satisfies the semi-Lipschitz condition (1), we have $\ell_t(y) \ge \ell_t(y_k) - \gamma$. This gives

$$\min_{k=1,\dots,K}\mathbb{E}\left[\sum_{t=1}^{T}\ell_t(y_k)\right] \le \min_{0\le y\le 1}\sum_{t=1}^{T}\ell_t(y) + \gamma T$$

Replacing the last inequality into (4), and recalling that $K = \lceil 1/\gamma \rceil$ and $\eta = \frac{\gamma}{2}$, finally yields

$$R_T \le \frac{\gamma T}{4}\ln\frac{e}{\gamma} + \frac{2\ln\lceil 1/\gamma\rceil}{\gamma} + 2\gamma T$$

Choosing $\gamma \approx T^{-1/2}$ concludes the proof. □