Corruption-Robust Contextual Search through Density Updates

Renato Paes Leme  
Google Research NYC  
RENATOPPL@GOOGLE.COM

Chara Podimata  
Harvard  
PODIMATA@G.HARVARD.EDU

Jon Schneider  
Google Research NYC  
JSCHNEI@GOOGLE.COM

Editors: Po-Ling Loh and Maxim Raginsky

Contextual Search is a fundamental primitive in online learning with binary feedback with applications to dynamic pricing (Kleinberg and Leighton, 2003) and personalized medicine (Bastani and Bayati, 2016). In each round, the learner chooses an action based on contextual information and observes only a single bit of feedback (e.g., “yes” or “no”). In the classic (realizable and noise-free) version, there exists a hidden vector \( \theta^* \in \mathbb{R}^d \) with \( \|\theta^*\| \leq 1 \) that the learner wishes to learn over time. Each round \( t \in [T] \) begins with the learner receiving a context \( u_t \in \mathbb{R}^d \) with \( \|u_t\| = 1 \). The learner then chooses an action \( y_t \in \mathbb{R} \), learns the sign \( \sigma_t = \text{sign}(\langle u_t, \theta^* \rangle - y_t) \in \{+1, -1\} \) and incurs loss \( \ell(y_t, \langle u_t, \theta^* \rangle) \). Importantly, the learner does not get to observe the loss they incur, only the feedback. A sequence of recent papers (Amin et al., 2014; Cohen et al., 2016; Lobel et al., 2017; Leme and Schneider, 2018; Liu et al., 2021) obtained the optimal regret bound for various loss functions, as highlighted on Table 1. The matching (up to \( \log d \)) upper and lower bounds in

<table>
<thead>
<tr>
<th>Loss</th>
<th>Lower Bound</th>
<th>Upper Bound</th>
</tr>
</thead>
<tbody>
<tr>
<td>( \varepsilon )-ball</td>
<td>( \Omega(d \log(1/\varepsilon)) )</td>
<td>( O(d \log(1/\varepsilon)) ) (Lobel et al., 2017)</td>
</tr>
<tr>
<td>absolute</td>
<td>( \Omega(d) )</td>
<td>( O(d \log d) ) (Liu et al., 2021)</td>
</tr>
<tr>
<td>pricing</td>
<td>( \Omega(d \log \log T) )</td>
<td>( O(d \log \log T + d \log d) ) (Liu et al., 2021)</td>
</tr>
</tbody>
</table>

Table 1: Optimal regret guarantees for realizable contextual search.

Table 1 indicate that the noiseless version of the problem is well understood. However, a lot of questions remain when the feedback is perturbed by some type of noise (as is often the case in practical settings). In the noisy model, the target value \( y_t^* = \langle u_t, \theta^* \rangle \) is perturbed to \( y_t^* + z_t \). Most of the literature thus far has focused on stochastic noise models (Javanmard and Nazerzadeh, 2016; Cohen et al., 2016; Javanmard, 2017; Shah et al., 2019; Liu et al., 2021; Xu and Wang, 2021, 2022).

A recent trend in machine learning is the study of adversarial noise models, often also called corrupted noise models. In this model, most of the data follows a learnable pattern but an adversary can corrupt a small fraction of it. The goal is to design learning algorithms whose performance is a function of how much corruption was added to the data. For the \( \varepsilon \)-ball loss, we give a tight regret bound of \( O(C + d \log(1/\varepsilon)) \) improving over the \( O(d^3 \log(1/\varepsilon)) \log^2(T) + C \log(T) \log(1/\varepsilon)) \) bound of Krishnamurthy et al. (2021). For the symmetric loss, we give an efficient algorithm with regret \( O(C + d \log T) \). Our techniques are a significant departure from prior approaches. Specifically, we keep track of carefully maintained density functions over the candidate vectors instead of a knowledge set consisting of the candidate vectors consistent with the feedback obtained.\(^2\)

1. We use the terms “regret” and “total loss” interchangeably.

References


Hamsa Bastani and Mohsen Bayati. Online decision-making with high-dimensional covariates. *Working paper, Stanford University*, 2016. 1


