Simultaneously Achieving Sublinear Regret and Constraint Violations for Online Convex Optimization with Time-varying Constraints

Qingsong Liu
liu-qs19@mails.tsinghua.edu.cn
IIIS, Tsinghua University, China

Wenfei Wu
wenfeiwu@pku.edu.cn
Peking University, China

Longbo Huang, Zhixuan Fang
{longbohuang, zfang}@tsinghua.edu.cn
IIIS, Tsinghua University, China

ABSTRACT

In this paper, we develop a novel virtual-queue-based online algorithm for online convex optimization (OCO) problems with long-term and time-varying constraints and conduct a performance analysis with respect to the dynamic regret and constraint violations. We design a new update rule of dual variables and a new way of incorporating time-varying constraint functions into the dual variables. To the best of our knowledge, our algorithm is the first parameter-free algorithm to simultaneously achieve sublinear dynamic regret and constraint violations. Our proposed algorithm also outperforms the state-of-the-art results in many aspects, e.g., our algorithm does not require the Slater condition. Furthermore, for a group of practical and widely-studied constrained OCO problems in which the variation of consecutive constraints is smooth enough across time, our algorithm achieves $O(1)$ constraint violations.

1. INTRODUCTION

Online Convex Optimization (OCO) with long-term constraints has become one of the most popular online learning frameworks in recent years due to its powerful modeling capability for various problems such as network routing [1], online display advertising [2], and resources management [3]. In the formulation of OCO with long-term constraints, the agent wants to minimize the accumulated loss while satisfying the constraints as much as possible in the long-term. Most existing works (e.g., [4, 5]) consider the scenarios where the constraints are time-invariant.

However, time-varying constraints arise in many practical applications in which the underlying time-varying system is dynamic and uncertain, e.g., smart grid with uncertain renewable energy supply [6] and data centers with dynamic user demands [7]. Thus, this paper considers this recently proposed OCO framework with long-term and time-varying constraints [8, 9, 10], which is more general and practical than the one with time-variant constraints setting.

Recent advances in OCO with long-term and time-varying constraints usually adopt dynamic regret and constraint violations as the metrics. In this setting, a crucial challenge is to achieve sublinear dynamic regret and constraint violation simultaneously. [8] is the first work to simultaneously achieve sublinear dynamic regret and constraint violations. But the performance bounds attained in [8] are only valid when the accumulated variations of the environment is known to the agent in advance, i.e., parameter-dependent. For parameter-free work, [3] analyzed the performance of a modified online saddle-point (MOSP) method and showed that sublinear dynamic regret and constraints violation may be achieved if the accumulated variations of the environment are sublinear. Later [11] improves upon it in terms of fewer assumptions but incurs a degradation of the performance. [9] proposed a variant of MOSP method and established the state-of-the-art performance upper bounds. However, all these parameter-free methods do not always guarantee the sublinear regret and constraint violations simultaneously, even given the accumulated variations of the environment is sublinear. Besides, most of them assume the Slater condition holds while it is not satisfied in many scenarios.

Thus, a challenging question remains that could we develop a parameter-free algorithm and achieve sublinear dynamic regret and constraint violations simultaneously without the Slater condition? The answer is yes. In our paper, we develop and analyze a novel parameter-free algorithm for OCO with long-term and time-varying constraints based on virtual queues. And we show that our algorithm can achieve sublinear dynamic regret and constraint violations simultaneously without Slater condition. The dynamic regret and constraint violations bounds of our developed algorithm outperform the state-of-the-art in many aspects. Furthermore, we show that when the variation of consecutive constraints is smooth enough across time, which holds in many practical applications [3], our algorithm can achieve $O(1)$ constraint violations.

2. PROBLEM SETUP

Let $\{f_t(x)\}_{t=1}^{\infty}$ and $\{\boldsymbol{g}_t(x)\}_{t=1}^{\infty}$ be time-varying continuous convex functions defined over a closed convex set $\chi \subseteq R^n$. In each round t, the agent incurs a loss function f_t and a constraint requirement $\boldsymbol{g}_t$. Specifically, we aim to solve the following online optimization problem:

$$\min_{\{x_t\}_{t=1}^{T}} \sum_{t=1}^{T} f_t(x_t), \text{ s.t. } \sum_{t=1}^{T} \boldsymbol{g}_t(x_t) \le \mathbf{0}. \tag{P1}$$

Where $\boldsymbol{g}_t(x)$ is defined as $[g_{t,1}(x), g_{t,2}(x), .., g_{t,K}(x)]^T$. In this paper, we consider the dynamic regret and constraint violations as the performance metrics, which are defined as follows, respectively,

$$\begin{aligned} \text{Regret} &= \sum_{t=1}^{T} f_t(x_t) - \sum_{t=1}^{T} f_t(x_t^*), \\ \text{Vio}_k &= \sum_{t=1}^{T} g_{t,k}(x_t),\ k \in \{1, 2, ..., K\}, \end{aligned} \tag{1}$$

Zhixuan Fang is the corresponding author.

Algorithm 1 VQB

1: Initialize: $\alpha_1, \gamma_0 > 0$, $\boldsymbol{g}_0 = \boldsymbol{\lambda}(0) = 0$, and $x_1 \in \chi$.
2: for round $t = 1...T-1$ do
3: Update the dual iterate $\boldsymbol{\lambda}(t)$:
4: $\boldsymbol{\lambda}(t) = \max\{\boldsymbol{\lambda}(t-1) + \gamma_{t-1}\boldsymbol{g}_{t-1}(x_t), -\gamma_{t-1}\boldsymbol{g}_{t-1}(x_t)\}$
5: Update the primal iterate that satisfies:
6: $x_{t+1} = \arg\min_{x\in\chi} \nabla f_t(x_t)^T(x - x_t) + [\boldsymbol{\lambda}(t) + \gamma_{t-1}\boldsymbol{g}_{t-1}(x_t)]^T(\gamma_t \boldsymbol{g}_t(x)) + \alpha_t ||x - x_t||_2^2$
7: Choose the action x_{t+1}
8: end for

where $\{x_t^*\}_{t=1}^T$ is the per-slot minimizers sequence which is defined as $x_t^* = \arg\min_{x\in\chi, \boldsymbol{g}_t(x)\leq 0} f_t(x)$. Our goal is to choose x_t in each round t such that both the dynamic regret and constraint violations grow sub-linearly with respect to the time horizon T. The performance bounds of any online algorithm should depend on how drastically $\{f_t\}$ and $\{\boldsymbol{g}_t\}$ vary across time. In this paper, we use the path-length and function variation to quantify the temporal variations of $\{f_t\}$ and $\{\boldsymbol{g}_t\}$, which are mainly-used regularities in the literature[3, 8, 9, 10, 11]. The path-length $V_x = \sum_{t=2}^T ||x_t^* - x_{t-1}^*||$ is the accumulated variation of per-slot minimizers $\{x_t^*\}$. The function variation $V_{\boldsymbol{g}}$ is the accumulated variation of consecutive constraints, i.e.,

$$V_{\boldsymbol{g}} = \sum_{t=2}^T \sup_{x\in\chi} ||\boldsymbol{g}_t(x) - \boldsymbol{g}_{t-1}(x)||.$$

We let $||\cdot||$ be the Euclidean norm throughout this paper. We assume the feasible set χ is closed, convex, and compact with diameter R, i.e., $\forall x, y \in \chi$, it holds that $||x - y|| \leq R$. We also assume the loss functions and constraint functions are convex, and bounded on χ, i.e., there exists a positive constant F such that $\max\{|f_t(x)|, ||\boldsymbol{g}_t(x)||\} \leq F, \forall x \in \chi, t$. Besides, the gradients of $g_{k,t}$ and f_t are upper-bounded by G over χ, i.e., $\max\{||\nabla f_t(x)||, ||\nabla g_{k,t}(x)||\} \leq G, \forall x \in \chi, k, t$.

Under these assumptions, we study problem (P1) in the full-information setting; that is, at round t, the agent can observe the complete loss and constraint functions after the decision x_t is submitted.

3. MAIN RESULTS

In this section, we first propose a novel virtual-queue-based algorithm, VQB, which is illustrated in Algorithm 1. Then we present the major theoretical results of our algorithm. VQB introduces a sequence of dual variables $\{\boldsymbol{\lambda}(t)\}$, which is also called virtual queue. The purpose of introducing the virtual queues is that we can characterize the regret and constraint violations through the drift-plus-penalty expression and then analyze the regret and constraint violations based on it. The upper bounds on the dynamic regret and constraint violations for VQB is given as follows.

Theorem 1. Set $\alpha_t = \sqrt{\frac{T}{R+\sum_{i\leq t}||x_i^*-x_{i-1}^*||}}$ and $\gamma_t^2 = \frac{1}{2\beta^2}\frac{1}{\sqrt{2R}}$ in VQB, then we have

$$\begin{aligned} &\text{Regret} \leq O(\max\{\sqrt{TV_x}, V_{\boldsymbol{g}}\}), \\ &\text{Vio}_k \leq O(\max\{\sqrt{T}, V_{\boldsymbol{g}}\}), \forall k = 1, 2, ..., K. \end{aligned} \tag{2}$$

Or set $\alpha_t = \sqrt{\frac{T}{R+\sum_{i\leq t}||x_i^*-x_{i-1}^*||}}$ and $\gamma_t^2 = \frac{1}{2\beta^2}\frac{1}{\sqrt{2R}}\frac{1}{\sqrt{t+1}}$ in VQB, then the performance guarantee of VQB becomes

$$\begin{aligned} &\text{Regret} \leq O(\sqrt{TV_x}), \\ &\text{Vio}_k \leq O(\max\{T^{\frac{3}{4}}, V_{\boldsymbol{g}}\}), \forall k = 1, 2, ..., K. \end{aligned} \tag{3}$$

Thus, VQB indeed simultaneously achieves sublinear regret and constraint violations without the Slater condition. Besides, our algorithm is parameter-free and the performance bounds of it outperform all existing works.

Extensions. Next, we consider a slightly stronger Slater condition: the slater constant ϵ satisfies

$$\epsilon > \max_t \max_{x\in\chi} ||\boldsymbol{g}_{t+1}(x) - \boldsymbol{g}_t(x)||,$$

which has been considered in [3]. We replace $\boldsymbol{g}_{t-1}(x_t)$ with $\boldsymbol{g}_t(x_t)$ in the update rule of $\boldsymbol{\lambda}(t)$ and x_{t+1} in VQB. Then this variant of VQB, named Algorithm 2, could guarantee the $O(1)$ constraint violations. We show this result in the following theorem.

Theorem 2. Set $\alpha_t = T^a$ and $\gamma_t^2 = \frac{1}{2\beta^2}T^a$ in Algorithm 2, then we have

$$\begin{aligned} &\text{Regret} \leq O(\max\{T^a V_x, T^a V_{\boldsymbol{g}}, T^{1-a}\}) \\ &\text{Vio}_k \leq O(1), \forall k = 1, 2, ..., K \end{aligned} \tag{4}$$

4. REFERENCES

[1] H. Yu, M. Neely, X. Wei, Online convex optimization with stochastic constraints, in: Advances in Neural Information Processing Systems, 2017, pp. 1428–1438.

[2] E. Hazan, Introduction to online convex optimization, arXiv preprint arXiv:1909.05207 (2019).

[3] T. Chen, Q. Ling, G. B. Giannakis, An online convex optimization approach to proactive network resource allocation, IEEE Transactions on Signal Processing 65 (24) (2017) 6350–6364.

[4] H. Yu, M. J. Neely, A low complexity algorithm with o($\sqrt{T}$) regret and o(1) constraint violations for online convex optimization with long term constraints, Journal of Machine Learning Research 21 (1) (2020) 1–24.

[5] S. Qiu, X. Wei, Beyond o($\sqrt{T}$) regret for constrained online optimization: Gradual variations and mirror prox, arXiv preprint arXiv:2006.12455 (2020).

[6] Y. Zhang, M. H. Hajiesmaili, S. Cai, M. Chen, Q. Zhu, Peak-aware online economic dispatching for microgrids, IEEE transactions on smart grid 9 (1) (2016) 323–335.

[7] Z. Liu, I. Liu, S. Low, A. Wierman, Pricing data center demand response, ACM SIGMETRICS Performance Evaluation Review 42 (1) (2014) 111–123.

[8] X. Cao, K. R. Liu, Online convex optimization with time-varying constraints and bandit feedback, IEEE Transactions on Automatic Control 64 (7) (2018) 2665–2680.

[9] T. Chen, G. B. Giannakis, Bandit convex optimization for scalable and dynamic iot management, IEEE INTERNET OF THINGS JOURNAL 6 (1) (2019).

[10] X. Yi, X. Li, T. Yang, L. Xie, T. Chai, K. H. Johansson, Distributed bandit online convex optimization with time-varying coupled inequality constraints, IEEE Transactions on Automatic Control (2020).

[11] T. Chen, Q. Ling, Y. Shen, G. B. Giannakis, Heterogeneous online learning for "thing-adaptive" fog computing in iot, IEEE Internet of Things Journal 5 (6) (2018) 4328–4341.