

Article

Not peer-reviewed version

A Model-Based Stochastic Augmented Lagrangian Method for Online Stochastic Optimization

[Zewei Wang](#), [Dan Xue](#)^{*}, Yujia Zhai, Cong Li

Posted Date: 14 April 2026

doi: 10.20944/preprints202604.0934.v1

Keywords: online stochastic optimization; time-varying distribution; augmented Lagrangian method; machine learning; network resource allocation



Preprints.org is a free multidisciplinary platform providing preprint service that is dedicated to making early versions of research outputs permanently available and citable. Preprints posted at Preprints.org appear in Web of Science, Crossref, Google Scholar, Scilit, Europe PMC.

Copyright: This open access article is published under a [Creative Commons CC BY 4.0 license](#), which permit the free download, distribution, and reuse, provided that the author and preprint are cited in any reuse.

Disclaimer/Publisher's Note: The statements, opinions, and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions, or products referred to in the content.

Article

A Model-Based Stochastic Augmented Lagrangian Method for Online Stochastic Optimization

Zewei Wang, Dan Xue * and Yujia Zhai and Cong Li

School of Mathematics and Statistics, Qingdao University, Qingdao, 266071, China

* Correspondence: xuedan@qdu.edu.cn

Abstract

In this paper, we focus on the online stochastic optimization problems in which the random parameters follow time-varying distributions. At each round t , decision is obtained from solving current optimization problem. Then samples are drawn from distributions which are updated after obtaining decision. The objective and constraint are updated in this process, and the updated problem is used to obtain the next decision. For solving the online stochastic optimization problem, we propose a model-based stochastic augmented Lagrangian method, which is referred to as MSALM. At each round, we construct the model functions for the sample objective and constraint functions based on their properties, which reduced the computational complexity. The step size is designed in a dynamic form and decreases as t increases to accelerate convergence. Due to the setting of the online stochastic problem, we use stochastic dynamic regret and constraint violation to measure the performance of our algorithm. Under the assumptions, we prove that our algorithm's stochastic dynamic regret and constraint violation have a sublinear bound of total number of slots T . We design simulation experiments to verify the efficiency of our online algorithm. Its performance is evaluated on a range of information and system engineering problems, including adaptive filtering, online logistic regression, the time-varying smart grid energy dispatch, the online network resource allocation, and the path planning. In addition, in the context of the path planning problem, we integrate our algorithm with supervised learning to demonstrate its enhanced capabilities. The experimental results validate the performance of our new algorithm in practical applications.

Keywords: online stochastic optimization; time-varying distribution; augmented Lagrangian method; machine learning; network resource allocation

1. Introduction

In recent years, online optimization problems have garnered widespread attention because of their ability to make real-time decisions with partial information. In the online optimization process, decisions are made sequentially according to feedback from the environment. The general form of this problem can be written as follows:

$$\min \quad f_t(x), \quad t \in \{1, 2, \dots, T\}$$

where T is the total number of decision rounds. The meaning of online optimization is clearly defined in several classic review articles[1–3]. These articles define regret to measure the online optimization algorithm.

$$Regret_T := \sum_{t=1}^T f_t(x_t) - \min_{x \in X} \sum_{t=1}^T f_t(x)$$

Considering online optimization with constraint, many algorithms are proposed to solve constrained online optimization like in [4,5]. Constraint is added in the problem framework.

$$s.t. \quad g_t(x) \leq 0$$

Constraint violation is defined as:

$$Violation_T := \sum_{t=1}^T [g_t(x)]_+$$

In practice, each problem encountered during the online process is closer to being generated randomly. Therefore, studying online stochastic optimization problems is more suitable for meeting real-world demands, where objective and constraint contain random parameters. Taira et al. [6] study an online stochastic optimization algorithm, in which at each round the parameters are sampled independently from a fixed distribution. To more closely approximate the randomly generated nature of the problem, Cao et al. [7] study the following online stochastic optimization problem with time-varying distributions:

$$\begin{aligned} \min \quad & F_t(x) := E_{\theta \sim P_t} [f(x, \theta)] \\ \text{s.t.} \quad & G_t(x) := E_{\xi \sim Q_t} [g(x, \xi)] \leq 0, \quad t \in \{1, 2, \dots, T\} \end{aligned} \quad (1)$$

In their work, they employ the projected stochastic gradient method (PSGD), which is a deterministic offline version of online gradient descent (OGD) [1], to solve the problem. However, PSGD uses a fixed step size and lacks an adaptive mechanism, which may lead to slower convergence in online settings. Research on this kind of issue is still very limited at present. Thus, we aim to design a new algorithm to efficiently solve online stochastic optimization problem with time-varying distributions.

For stochastic optimization problems, existing stochastic gradient-based algorithms such as SGD [8], SVRG [9], and SPDAM [10] give us some insights into analyzing and handling stochastic problems. The Lagrangian method [11] is widely used in constrained optimization problems. Several works use this method in online optimization, for example, [12,13]. Liu et al. proposed a model-based augmented Lagrangian method to solve online constrained optimization [14]. This method performs well for specific structured problems. It simplifies the computation by approximating the objective and constraint functions at each round. Together with the corresponding primal-dual update formula, this method guarantees a sublinear upper bound. Especially for online optimization, [15] shows that the augmented Lagrangian method converges super-linearly asymptotically.

The regret of online algorithms has a dynamic version, which is defined in [1]. Dynamic regret measures the ability of algorithm to track the optimal solution at each round. [7] defines dynamic regret and constraint violation to measure its performance and adapts to stochasticity:

$$\begin{aligned} \text{Regret}(T) &:= E \left[\sum_{t=1}^T F_t(x_t) \right] - \sum_{t=1}^T F_t(x_t^*) \\ \text{Violation}_i(T) &:= E \left[\sum_{t=1}^T G_{t,i}(x_t) \right] \end{aligned} \quad (2)$$

Recent studies on distributed online optimization with coupled constraints [16] and online composite optimization with time-varying regularizers [17] also provide valuable insights into handling dynamic and structured online optimization problems, further motivating our approach.

Inspired by these ideas, we solve the online stochastic optimization problem with time-varying distribution by using the model-based augmented Lagrangian method in [14]. At the same time, we incorporate time-varying distribution approximations and a dynamic step size. Stochastic dynamic regret and constraint violation are used to evaluate the performance of our algorithm [7].

The following are the main contributions of our work:

1. We propose a model-based stochastic augmented Lagrangian method (MSALM) for online stochastic optimization. In each round, we construct model functions to approximate the stochastic objective and constraint functions, which are sampled from time-varying distributions. This

construction reduces computational complexity. The step size is designed in a dynamic form and decreases as t increases to accelerate convergence.

2. We adopt dynamic regret and constraint violation as our performance metrics. These measures are particularly suited for online stochastic optimization with time-varying distributions. Under the assumptions, we prove that the algorithm's regret and constraint violation have sublinear bounds in terms of the total number of rounds T .
3. We demonstrate the practical efficacy of our proposed algorithm through a series of simulation experiments. In the contexts of adaptive filtering and online logistic regression, we compare our method with PSGD. The results show that MSALM attains lower regret and constraint violation bounds than PSGD, indicating that MSALM converges more rapidly toward the theoretical optimum while maintaining stricter adherence to constraints. In addition, results from the time-varying smart grid energy dispatch, online network resource allocation, and path planning problems collectively confirm that regret and constraint violation have bounds $O(\sqrt{T})$.

2. MSALM for Online Stochastic Optimization

This section presents the online stochastic optimization problem and the details of the model-based stochastic augmented Lagrangian method (MSALM). Then we describe the update strategies of our algorithm.

2.1. The Online Stochastic Optimization Problem

The online optimization problem is a process of making decisions sequentially with partial information. It generates a sequence of decisions through continuous interaction with the environment. The environment refers to the optimization objective (loss function) and its constraints at each round. If the generation process of the objectives and constraints is stochastic, then this problem becomes an online stochastic problem, as shown in 1. The random parameters θ and ξ represent the samples, which are drawn from the time-varying distributions defined in 1. The problem determines the online decision process. At round t , the decision x_{t-1} is selected based on previous information. Then the distributions P_{t-1} and Q_{t-1} are updated to P_t and Q_t . Parameters $\theta_t \sim P_t$ and $\xi_t \sim Q_t$ are drawn from the current distribution. After that, the f_t and g_t are obtained as the samples of F_t and G_t as:

$$f_t(x) = f(x, \theta_t), \quad g_t(x) = g(x, \xi_t)$$

where $x \in X \subset \mathbb{R}^n$. New decision x_t is selected by solving this optimization problem.

Based on the above setup, we make the following standard assumptions.

Assumption 1. X is a bounded set, and there exists a constant $R > 0$ for any $x, y \in X$ satisfying

$$\|x - y\| \leq R$$

Assumption 2. $f(x, \theta)$ and $g(x, \xi)$ are convex and differentiable for any $\theta \in \Theta$ and $\xi \in \Xi$.

Assumption 3 (Slater's condition). At each round t , there exists a decision $x_t \in X$ for all $i = 1, 2, \dots, m$ satisfying

$$g_t^{(i)}(x_t) \leq 0$$

2.2. MSALM Algorithm

To efficiently solve the online stochastic optimization problem with time-varying distributions defined in 1, we extend the model-based augmented Lagrangian method (MALM) proposed by Liu et al. [14] to the stochastic setting. In the MALM framework [14], model-based means that we conservatively approximate the objective and constraint based on the properties of functions. The

approximations $\hat{f}_t(x)$ and $\hat{g}_t(x)$ are the model functions to objective and constraint at x_t . Model functions satisfy the following conditions [14]:

Assumption 4. 1. For any $x \in X$,

$$\hat{f}_t(x) \leq f_t(x), \quad \hat{f}_t(x_t) = f_t(x_t)$$

and $\hat{g}_t(x) \leq g_t(x)$, with equality at $x = x_t$

2. For any $x \in X$ and any $i = 1, 2, \dots, m$,

$$\hat{g}_t^{(i)}(x) \leq g_t^{(i)}(x), \quad \hat{g}_t^{(i)}(x_t) = g_t^{(i)}(x_t)$$

3. $\hat{g}_t(\cdot) = [\hat{g}_t^{(1)}(\cdot), \hat{g}_t^{(2)}(\cdot), \dots, \hat{g}_t^{(m)}(\cdot)]$ is a bounded mapping on X , and there exists a constant $D > 0$ for any $x \in X$ satisfying

$$\|\hat{g}_t(x)\| \leq D$$

In many approximations, we need gradient information of $f_t(x)$ and $g_t(x)$. However, in our stochastic setting, the exact gradients $\nabla F_t(x)$ and $\nabla G_t(x)$ are not directly accessible. Instead, we use unbiased stochastic gradient estimates $\nabla f_t(x)$ and $\nabla g_t(x)$ based on the sampled functions.

$$\nabla f_t(x) = E_{\theta_t \sim P_t}[\nabla F_t(x)], \quad \nabla g_t(x) = E_{\xi_t \sim Q_t}[\nabla G_t(x)]$$

Efficient models that satisfy Assumption 4 can be designed. Depending on the properties of the $f_t(x)$ and $g_t(x)$, different models are selected for approximation. Several model functions are presented in the MALM algorithm [14]:

- Linearized model:

$$\begin{aligned} \hat{f}_t(x) &:= f_t(x) + \langle \nabla f_t(x), x - x_t \rangle \\ \hat{g}_t^{(i)}(x) &:= g_t^{(i)}(x) + \langle \nabla g_t^{(i)}(x), x - x_t \rangle, \quad i = 1, \dots, m \end{aligned}$$

- Quadratic model:

$$\hat{f}_t(x) := f_t(x) + \langle \nabla f_t(x), x - x_t \rangle + \frac{l}{2} \|x - x_t\|^2$$

- Truncated model:

$$\hat{f}_t(x) := [f_t(x) + \langle \nabla f_t(x), x - x_t \rangle]_+$$

- Plain model:

$$\begin{aligned} \hat{f}_t(x) &:= f_t(x) \\ \hat{g}_t^{(i)}(x) &:= g_t^{(i)}(x), \quad i = 1, \dots, m \end{aligned}$$

Then we define the model-based stochastic augmented Lagrangian of the problem as:

$$\hat{L}_{t,\sigma}(x, \lambda) := \hat{f}_t(x) + \frac{1}{2\sigma} [\|\lambda + \sigma \hat{g}_t(x)\|_+^2 - \|\lambda\|^2]$$

where the operator $[\cdot]_+$ means $\max\{\cdot, 0\}$.

In the MALM algorithm [14], the primal variable is updated by solving the following proximal augmented Lagrangian subproblem:

$$x_{t+1} = \arg \min_{x \in X} [\hat{L}_{t,\sigma}(x, \lambda_t) + \frac{\alpha}{2} \|x - x_t\|^2]$$

where $\alpha > 0$ is the parameter of the proximal term. The optimality condition of the problem can be transformed as:

$$x_{t+1} = x_t - \frac{1}{\alpha} \nabla_x \hat{L}_{t,\sigma}(x_{t+1}, \lambda_t)$$

The $\frac{1}{\alpha}$ plays the role of the step size in a gradient descent step.

We design the step size as $\alpha_t = \alpha_0 \sqrt{t}$ which decreases as t increases, to accelerate convergence. At the beginning of the iteration, the algorithm has a larger step size. With the iteration of the algorithm, the step size continues to decrease. This design meets the requirements of different stages of the iterative process. In the early stage, a larger step size can accelerate algorithm iteration. In the late stage, a smaller step size can control the update amplitude and pursue precision. The proposed update for MSALM is:

$$x_{t+1} = \arg \min_{x \in X} [\hat{L}_{t,\sigma}(x, \lambda_t) + \frac{\alpha_t}{2} \|x - x_t\|^2] \quad (3)$$

The multiplier λ is updated by,

$$\lambda_{t+1} = [\lambda_t + \sigma \hat{g}_t(x_{t+1})]_+ \quad (4)$$

The algorithm based on model-based augmented Lagrangian method is as follows,

Algorithm 1 MSALM

Require: Choose an initial point $x_0 \in X$ arbitrarily. Set parameters $\alpha_0 > 0$, $\sigma > 0$. Set the initial multiplier $\lambda_0 = 0$.

for $t=1, 2, \dots, T$ **do**

 Submit the decision x_t .

 Update distributions P_t and Q_t to determine F_t and G_t .

 Generate f_t and g_t by sampling $\theta_t \sim P_t$ and $\zeta_t \sim Q_t$.

 Approximate $f_t(x)$ and $g_t(x)$ as $\hat{f}_t(x)$ and $\hat{g}_t(x)$.

 Update x_{t+1} and λ_{t+1} by 3 and 4

end for

We measure the algorithm performance by stochastic dynamic regret and constraint violation 2. x_t^* is the theoretical optimal decision at each round.

3. Convergence Analysis

In this section, we analyze the performance of MSALM in online stochastic optimization by stochastic dynamic regret and constraint violation.

For proving stochastic dynamic regret, we adopt the definition of drift and the assumption on the drift from [7]:

$$\Delta(T) := \sum_{t=2}^T \|x_{t-1}^* - x_t^*\|$$

Assumption 5. Before the start of the algorithm, there is a $\bar{\Delta}(T)$ for any T satisfying

$$\Delta(T) \leq \bar{\Delta}(T)$$

The drift and assumption 5 ensure that although the problem with time-varying distributions retains the characteristic of sharing a common decision set, as in standard online optimization.

Assumption 6. The gradient of $f_t(x)$ is bounded, i.e., there is a constant $G_f > 0$ satisfying

$$\|\nabla f_t\| \leq G_f$$

Lemma 1. Under the Assumption 1-6, we have,

$$E\left[\sum_{t=1}^T \alpha_t (\|x_t^* - x_t\|^2 - \|x_t^* - x_{t+1}\|^2)\right] \leq \alpha_0 R \sqrt{T} (2\bar{\Delta}(T) + 3R)$$

Proof.

$$\begin{aligned} \|x_t^* - x_t\|^2 - \|x_t^* - x_{t+1}\|^2 &\leq \|x_t^* - x_t\|^2 - \|x_t^* - x_{t+1}^*\|^2 - \|x_{t+1}^* - x_{t+1}\|^2 \\ &\quad + 2\|x_t^* - x_{t+1}^*\| \|x_{t+1}^* - x_{t+1}\| \end{aligned}$$

Let $b_t = x_t^* - x_{t+1}^*$,

$$\|x_t^* - x_t\|^2 - \|x_t^* - x_{t+1}\|^2 \leq \|x_t^* - x_t\|^2 - \|b_t\|^2 - \|x_{t+1}^* - x_{t+1}\|^2 + 2\|b_t\| \cdot \|x_{t+1}^* - x_{t+1}\| \quad (5)$$

Let $A_t = E[\|x_t^* - x_t\|^2]$. By Assumption 1, we have $E[\|x_{t+1}^* - x_{t+1}\|] \leq R$ and $A_t \leq R^2$, then we take the expectation of 5,

$$E[\|x_t^* - x_t\|^2 - \|x_t^* - x_{t+1}\|^2] \leq A_t - A_{t+1} - \|b_t\|^2 + 2R\|b_t\|$$

Multiplying by α_t and summing,

$$E\left[\sum_{t=1}^T \alpha_t (\|x_t^* - x_t\|^2 - \|x_t^* - x_{t+1}\|^2)\right] \leq \sum_{t=1}^T \alpha_t (A_t - A_{t+1}) + \sum_{t=1}^T \alpha_t (2R\|b_t\| - \|b_t\|^2) \quad (6)$$

We examine the two terms on the right-hand side of the above inequality separately.

$$\begin{aligned} \sum_{t=1}^T \alpha_t (A_t - A_{t+1}) &= \alpha_1 A_1 - \alpha_T A_{T+1} + \sum_{t=2}^T (\alpha_t - \alpha_{t-1}) A_t \\ &\leq \alpha_0 R^2 + R^2 (\alpha_0 \sqrt{T} - \alpha_0) = \alpha_0 R^2 \sqrt{T} \\ \sum_{t=1}^T \alpha_t (2R\|b_t\| - \|b_t\|^2) &\leq 2\alpha_0 R \sqrt{T} \sum_{t=1}^T \|b_t\| \\ &\leq 2\alpha_0 R \sqrt{T} \Delta(T+1) \\ &\leq 2\alpha_0 R \sqrt{T} (\bar{\Delta}(T) + R) \end{aligned}$$

6 can be written as:

$$\begin{aligned} E\left[\sum_{t=1}^T \alpha_t (\|x_t^* - x_t\|^2 - \|x_t^* - x_{t+1}\|^2)\right] &\leq \alpha_0 R^2 \sqrt{T} + 2\alpha_0 R \sqrt{T} (\bar{\Delta}(T) + R) \\ &= \alpha_0 R \sqrt{T} (2\bar{\Delta}(T) + 3R) \end{aligned}$$

Theorem 1. Suppose Assumption 1-6 hold. The stochastic dynamic regret of MSALM has an sublinear upper bound, when the parameters are set $\alpha_0 = \frac{1}{\sqrt{\bar{\Delta}(T)}}$ and $\sigma = \frac{1}{\sqrt{T}}$.

$$\text{Regret}(T) \leq \frac{D^2 \sqrt{T}}{2} + (G_f^2 + R) \sqrt{T \bar{\Delta}(T)} + \frac{3R^2}{2} \sqrt{\frac{T}{\bar{\Delta}(T)}} = O(\sqrt{T \bar{\Delta}(T)})$$

Proof. We construct an auxiliary optimization problem:

$$\min_{x \in X} \hat{L}_{t,\sigma}(x, \lambda_t) + \frac{\alpha_t}{2} (\|x - x_t\|^2 - \|x - x_{t+1}\|^2)$$

The optimal solution of the auxiliary problem satisfies:

$$\nabla_x \hat{L}_{t,\sigma}(x, \lambda_t) + \alpha_t(x_{t+1} - x_t) = 0 \quad (7)$$

Comparing 7 with the optimality condition of ??, x_{t+1} is the optimal point of the auxiliary problem.

We have the inequality:

$$\hat{L}_{t,\sigma}(x_{t+1}, \lambda_t) + \frac{\alpha_t}{2} \|x_{t+1} - x_t\|^2 \leq \hat{L}_{t,\sigma}(x_t^*, \lambda_t) + \frac{\alpha_t}{2} (\|x_t^* - x_t\|^2 - \|x_t^* - x_{t+1}\|^2) \quad (8)$$

We analyze the two sides of the inequality separately. According to ??, the left-hand side becomes:

$$\hat{L}_{t,\sigma}(x_{t+1}, \lambda_t) = \hat{f}_t(x_{t+1}) + \frac{1}{2\sigma} [\|\lambda_{t+1}\|^2 - \|\lambda_t\|^2]$$

Combining $\hat{f}_t$ is convex and the assumption4,

$$\hat{f}_t(x_{t+1}) \geq \hat{f}_t(x_t) + \langle G_f, x_{t+1} - x_t \rangle \geq f_t(x_t) - G_f \|x_{t+1} - x_t\|$$

Th inequality 8 becomes:

$$\hat{L}_{t,\sigma}(x_{t+1}, \lambda_t) \geq f_t(x_t) - G_f \|x_{t+1} - x_t\| + \frac{1}{2\sigma} [\|\lambda_{t+1}\|^2 - \|\lambda_t\|^2]$$

Meanwhile, the right-hand side of the inequality 3 is bounded as:

$$\hat{L}_{t,\sigma}(x_t^*, \lambda_t) \leq f_t(x_t^*) + \langle \lambda_t, \hat{g}_t(x_t^*) \rangle + \frac{\sigma}{2} \|\hat{g}_t(x_t^*)\|^2$$

Since x_t^* is a feasible solution, we have $\langle \lambda_t, \hat{g}_t(x_t^*) \rangle \leq 0$. Considering the assumption4, we have:

$$\hat{L}_{t,\sigma}(x_t^*, \lambda_t) \leq f_t(x_t^*) + \frac{\sigma}{2} D^2$$

From the above two parts, the inequality8 becomes:

$$\begin{aligned} f_t(x_t) - G_f \|x_{t+1} - x_t\| + \frac{1}{2\sigma} [\|\lambda_{t+1}\|^2 - \|\lambda_t\|^2] + \frac{\alpha_t}{2} \|x_{t+1} - x_t\|^2 \\ \leq f_t(x_t^*) + \frac{\sigma}{2} D^2 + \frac{\alpha_t}{2} (\|x_t^* - x_t\|^2 - \|x_t^* - x_{t+1}\|^2) \end{aligned}$$

We rearrange and bound the above inequality as:

$$f_t(x_t) - f_t(x_t^*) \leq \frac{G_f^2}{2\alpha_t} - \frac{1}{2\sigma} [\|\lambda_{t+1}\|^2 - \|\lambda_t\|^2] + \frac{\sigma}{2} D^2 + \frac{\alpha_t}{2} (\|x_t^* - x_t\|^2 - \|x_t^* - x_{t+1}\|^2)$$

Summing from $t = 1$ to T , we have:

$$\begin{aligned} \sum_{t=1}^T (f_t(x_t) - f_t(x_t^*)) &\leq \sum_{t=1}^T \frac{G_f^2}{2\alpha_t} - \frac{1}{2\sigma} \sum_{t=1}^T [\|\lambda_{t+1}\|^2 - \|\lambda_t\|^2] + \frac{\sigma D^2 T}{2} \\ &\quad + \frac{1}{2} \sum_{t=1}^T \alpha_t (\|x_t^* - x_t\|^2 - \|x_t^* - x_{t+1}\|^2) \end{aligned} \quad (9)$$

Taking expectation of the inequality 9 and substituting $\lambda_1 = 0$:

$$E \left[\sum_{t=1}^T (f_t(x_t) - f_t(x_t^*)) \right] \leq \frac{1}{2} E \left[\sum_{t=1}^T \alpha_t (\|x_t^* - x_t\|^2 - \|x_t^* - x_{t+1}\|^2) \right] \\ + \sum_{t=1}^T \frac{G_f^2}{2\alpha_t} - \frac{1}{2\sigma} E[\|\lambda_{T+1}\|^2] + \frac{\sigma D^2 T}{2}$$

where for the term $\sum_{t=1}^T \frac{G_f^2}{2\alpha_t}$, we have:

$$\sum_{t=1}^T \frac{G_f^2}{2\alpha_t} \leq \frac{G_f^2 \sqrt{T}}{\alpha_0}$$

When the parameters are set $\alpha_0 = \frac{1}{\sqrt{\bar{\Delta}(T)}}$ and $\sigma = \frac{1}{\sqrt{T}}$, we substitute the result of Lemma1:

$$\text{Reg}(T) = E \left[\sum_{t=1}^T (f_t(x_t) - f_t(x_t^*)) \right] \leq \frac{D^2 \sqrt{T}}{2} + (G_f^2 + R) \sqrt{T \bar{\Delta}(T)} + \frac{3R^2}{2} \sqrt{\frac{T}{\bar{\Delta}(T)}} \\ = O(\sqrt{T \bar{\Delta}(T)})$$

Assumption 7. The gradient of $g_t(x)$ is bounded, i.e., there is a constant $G_g > 0$ satisfying

$$\|\nabla g_t\| \leq G_g$$

Lemma 2.

$$\frac{1}{2\sigma} E[\|\lambda_{t+1}\|^2 - \|\lambda_t\|^2] \leq 2G_f R + \frac{\sigma}{2} D^2 + \frac{\alpha_t}{2} E[\|x^s - x_t\|^2 - \|x^s - x_{t+1}\|^2] - \varepsilon_0 E[\|\lambda_t\|]$$

where x^s is the Slater's point.

Proof. According to inequality8, we consider a point x^s which satisfy the Slater's condition.

$$\hat{L}_{t,\sigma}(x_{t+1}, \lambda_t) + \frac{\alpha_t}{2} \|x_{t+1} - x_t\|^2 \leq \hat{L}_{t,\sigma}(x^s, \lambda_t) + \frac{\alpha_t}{2} (\|x^s - x_t\|^2 - \|x^s - x_{t+1}\|^2) \quad (10)$$

From the definition of the augmented Lagrangian function,

$$\hat{L}_{t,\sigma}(x^s, \lambda_t) = \hat{f}_t(x^s) + \frac{1}{2\sigma} [\|\lambda_t + \sigma \hat{g}_t(x^s)\|_+^2 - \|\lambda_t\|^2]$$

Using the non-expansiveness property of the projection operator,

$$\|[\lambda_t + \sigma \hat{g}_t(x^s)]_+\|^2 \leq \|\lambda_t\|^2 + 2\sigma \langle \lambda_t, \hat{g}_t(x^s) \rangle + \sigma^2 \|\hat{g}_t(x^s)\|^2$$

Therefore,

$$\hat{L}_{t,\sigma}(x^s, \lambda_t) \leq \hat{f}_t(x^s) + \langle \lambda_t, \hat{g}_t(x^s) \rangle + \frac{\sigma}{2} \|\hat{g}_t(x^s)\|^2$$

By assumption 3 and 4, we have

$$\hat{L}_{t,\sigma}(x^s, \lambda_t) \leq f_t(x^s) - \varepsilon_0 \|\lambda_t\| + \frac{\sigma}{2} D^2$$

The inequality 10 becomes:

$$\begin{aligned} f_t(x_t) - G_f \|x_{t+1} - x_t\| + \frac{1}{2\sigma} [\|\lambda_{t+1}\|^2 - \|\lambda_t\|^2] + \frac{\alpha_t}{2} \|x_{t+1} - x_t\|^2 \\ \leq f_t(x^s) - \varepsilon_0 \|\lambda_t\| + \frac{\sigma}{2} D^2 + \frac{\alpha_t}{2} (\|x^s - x_t\|^2 - \|x^s - x_{t+1}\|^2) \end{aligned}$$

We rearrange terms and take expectation,

$$\begin{aligned} \frac{1}{2\sigma} E[\|\lambda_{t+1}\|^2 - \|\lambda_t\|^2] \leq E[f_t(x^s) - f_t(x_t)] + G_f E[\|x_{t+1} - x_t\|] - \frac{\alpha_t}{2} E[\|x_{t+1} - x_t\|^2] \\ - \varepsilon_0 E[\|\lambda_t\|] + \frac{\sigma}{2} D^2 + \frac{\alpha_t}{2} E[\|x^s - x_t\|^2 - \|x^s - x_{t+1}\|^2] \end{aligned}$$

We bound each term.

$$\begin{aligned} E[f_t(x^s) - f_t(x_t)] &\leq G_f E[\|x^s - x_t\|] \leq G_f R \\ G_f E[\|x_{t+1} - x_t\|] - \frac{\alpha_t}{2} E[\|x_{t+1} - x_t\|^2] &\leq G_f R \end{aligned}$$

The inequality becomes:

$$\frac{1}{2\sigma} E[\|\lambda_{t+1}\|^2 - \|\lambda_t\|^2] \leq 2G_f R + \frac{\sigma}{2} D^2 + \frac{\alpha_t}{2} E[\|x^s - x_t\|^2 - \|x^s - x_{t+1}\|^2] - \varepsilon_0 E[\|\lambda_t\|] \quad (11)$$

Lemma 3. There exist constants $C_1, C_2, C_3, C_4 > 0$ for any $t \geq 0$ and positive integer s satisfying

$$E[\|\lambda_t\|] \leq \psi(\sigma, \alpha_0, s) := C_1 + C_2 \alpha_0 + C_3 \sigma + C_4 \sigma s$$

where $C_1 = \frac{4G_f R}{\varepsilon_0}$, $C_2 = \frac{\alpha_0 R^2 \sqrt{s}}{\varepsilon_0}$, $C_3 = \frac{D^2}{\varepsilon_0} - D$, $C_4 = 2D + \frac{\varepsilon_0}{2} + s \frac{8\sigma D^2}{\varepsilon_0} \log \frac{32D^2}{\varepsilon_0^2}$.

Proof. For any $t \geq 0$, the inequality 11 can be:

$$\frac{1}{2\sigma} E[\|\lambda_{t+1}\|^2 - \|\lambda_t\|^2] \leq 2G_f R + \frac{\sigma}{2} D^2 + \frac{\alpha_t}{2} E[\|x^s - x_t\|^2 - \|x^s - x_{t+1}\|^2] - \varepsilon_0 E[\|\lambda_t\|]$$

Summing from t to $t + s - 1$,

$$\begin{aligned} \frac{1}{2\sigma} \sum_{l=0}^{s-1} E[\|\lambda_{t+l+1}\|^2 - \|\lambda_{t+l}\|^2] &\leq \sum_{l=0}^{s-1} \frac{\alpha_{t+l}}{2} E[\|x^s - x_{t+l}\|^2 - \|x^s - x_{t+l+1}\|^2] \\ &\quad - \varepsilon_0 \sum_{l=0}^{s-1} E[\|\lambda_{t+l}\|] + s \left(2G_f R + \frac{\sigma}{2} D^2 \right) \end{aligned} \quad (12)$$

We have,

$$\begin{aligned} \sum_{l=0}^{s-1} \frac{\alpha_{t+l}}{2} E[\|x^s - x_{t+l}\|^2 - \|x^s - x_{t+l+1}\|^2] &\leq \frac{\alpha_0}{2} \sqrt{t+s-1} \sum_{l=0}^{s-1} E[\|x^s - x_{t+l}\|^2 - \|x^s - x_{t+l+1}\|^2] \\ &\leq \frac{1}{2} \alpha_0 \sqrt{t+s-1} R^2 \end{aligned}$$

Since the projection operator has the non-expansiveness:

$$\|\lambda_{t+1} - \lambda_t\| = \|[\lambda_t + \sigma \hat{g}_t(x_{t+1})]_+ - \lambda_t\| \leq \|\sigma \hat{g}_t(x_{t+1})\| \leq \sigma D$$

For any $l \geq 0$,

$$\|\lambda_{t+l}\| \geq \|\lambda_t\| - \sigma D l$$

Therefore,

$$\sum_{l=0}^{s-1} E[\|\lambda_{t+l}\|] \geq \sum_{l=0}^{s-1} (E[\|\lambda_t\|] - \sigma D l) = sE[\|\lambda_t\|] - \sigma D \frac{s(s-1)}{2}$$

Substituting the above results into the inequality 12,

$$\frac{1}{2\sigma} E[\|\lambda_{t+s}\|^2 - \|\lambda_t\|^2] \leq s \left(2G_f R + \frac{\sigma}{2} D^2 \right) + \frac{1}{2} \alpha_0 \sqrt{t+s-1} R^2 - \varepsilon_0 \left(sE[\|\lambda_t\|] - \sigma D \frac{s(s-1)}{2} \right)$$

Since $\|\lambda_{t+s}\|^2 \geq 0$, we have,

$$\varepsilon_0 s E[\|\lambda_t\|] \leq s \left(2G_f R + \frac{\sigma}{2} D^2 \right) + \frac{1}{2} \alpha_0 \sqrt{t+s-1} R^2 + \frac{1}{2\sigma} E[\|\lambda_t\|^2] + \varepsilon_0 \sigma D \frac{s(s-1)}{2}$$

From the update rule,

$$|\|\lambda_{t+1}\|^2 - \|\lambda_t\|^2| \leq 2\|\lambda_t\| \|\lambda_{t+1} - \lambda_t\| + \|\lambda_{t+1} - \lambda_t\|^2 \leq 2\sigma D \|\lambda_t\| + \sigma^2 D^2$$

We define:

$$\theta = C_1 + C_2 \alpha_0 + C_3 \sigma$$

where C_1, C_2, C_3 are appropriate constants.

Specifically, from the rearranged inequality,

$$\frac{1}{2\sigma} E[\|\lambda_{t+s}\|^2 - \|\lambda_t\|^2] \leq K_s - \varepsilon_0 s E[\|\lambda_t\|]$$

where K_s collects all the positive bounded terms.

From earlier, we have:

$$|\|\lambda_{t+1}\| - \|\lambda_t\|| \leq \sigma D$$

Now, from Lemma 2, we have:

$$\frac{1}{2\sigma} E[\|\lambda_{t+1}\|^2 - \|\lambda_t\|^2] \leq K_1 - \varepsilon_0 E[\|\lambda_t\|]$$

where $K_1 = 2G_f R + \frac{\sigma}{2} D^2 + \frac{\alpha_t}{2} E[\|x^s - x_t\|^2 - \|x^s - x_{t+1}\|^2]$.

Note that,

$$\|\lambda_{t+1}\|^2 - \|\lambda_t\|^2 = (\|\lambda_{t+1}\| - \|\lambda_t\|)(\|\lambda_{t+1}\| + \|\lambda_t\|)$$

When $\|\lambda_t\|$ is large, say $\|\lambda_t\| \geq \theta$, then,

$$\|\lambda_{t+1}\|^2 - \|\lambda_t\|^2 \leq -2\sigma \varepsilon_0 \theta + 2\sigma K_1$$

Based on this condition, we can obtain the inequality below by applying Lemma from [14],

$$\|\lambda_t\| \leq \theta + \sigma D + \frac{4\sigma^2 D^2}{\frac{\varepsilon_0}{2}} \log \left[\frac{8\sigma^2 D^2}{(\frac{\varepsilon_0}{2})^2} \right]$$

where $\theta = \frac{2K_1}{\varepsilon_0}$.

Taking expectation and substituting back,

$$\begin{aligned} E[\|\lambda_t\|] &\leq \frac{2G_f R}{\varepsilon_0} + \frac{\sigma D^2}{\varepsilon_0} + \frac{\alpha_0 R^2}{2\varepsilon_0} + \sigma D \\ &\quad + \sigma D \cdot \frac{8\sigma D}{\varepsilon_0} \log \left[\frac{32\sigma^2 D^2}{\varepsilon_0^2} \right] \end{aligned}$$

Simplifying and grouping terms, we obtain the desired form,

$$E[\|\lambda_t\|] \leq \psi(\sigma, \alpha_0, s) := C_1 + C_2\alpha_0 + C_3\sigma + C_4\sigma s$$

where,

$$\begin{aligned} C_1 &= \frac{2G_f R}{\varepsilon_0} \\ C_2 &= \frac{R^2}{2\varepsilon_0} \\ C_3 &= \frac{D^2}{\varepsilon_0} + D \\ C_4 &= \frac{8D^2}{\varepsilon_0} \log \left[\frac{32D^2}{\varepsilon_0^2} \right] \end{aligned}$$

Theorem 2. Suppose Assumption 1-7 hold. The constraint violation of MSALM has an sublinear upper bound, when the parameters are set $\alpha_0 = \frac{1}{\sqrt{\Delta(T)}}$ and $\sigma = \frac{1}{\sqrt{T}}$

$$\text{Violation}_i(T) := E \left[\sum_{t=1}^T G_{t,i}(x_t) \right] \leq \mathcal{O}(\sqrt{T})$$

Proof. According to the Lagrange multiplier update rule, assumption 4 and assumption 7, we have:

$$g_{t,i}(x_t) \leq \frac{1}{\sigma}(\lambda_{t+1,i} - \lambda_{t,i}) + G_g \|x_{t+1} - x_t\|$$

Summing from $t = 1$ to T ,

$$\sum_{t=1}^T g_{t,i}(x_t) \leq \frac{1}{\sigma} \sum_{t=1}^T (\lambda_{t+1,i} - \lambda_{t,i}) + G_g \sum_{t=1}^T \|x_{t+1} - x_t\| \leq \frac{1}{\sigma} \lambda_{T+1,i} + G_g \sum_{t=1}^T \|x_{t+1} - x_t\|$$

Taking expectation,

$$E \left[\sum_{t=1}^T g_{t,i}(x_t) \right] \leq \frac{1}{\sigma} E[\lambda_{T+1,i}] + G_g \sum_{t=1}^T E[\|x_{t+1} - x_t\|] \leq \frac{1}{\sigma} E[\|\lambda_{T+1}\|] + G_g \sum_{t=1}^T E[\|x_{t+1} - x_t\|]$$

According to the Lemma3, for any $t \geq 0$ and positive integer s ,

$$E[\|\lambda_t\|] \leq C_1 + C_2\alpha_0 + C_3\sigma + C_4\sigma s$$

Specifically, for $t = T + 1$, choosing $s = \sqrt{T}$,

$$E[\|\lambda_{T+1}\|] \leq C_1 + C_2\alpha_0 + C_3\sigma + C_4\sigma\sqrt{T}$$

Substituting the parameter choices $\alpha_0 = \frac{1}{\sqrt{\Delta(T)}}$ and $\sigma = \frac{1}{\sqrt{T}}$,

$$\begin{aligned} E[\|\lambda_{T+1}\|] &\leq C_1 + C_2 \frac{1}{\sqrt{\Delta(T)}} + C_3 \frac{1}{\sqrt{T}} + C_4 \frac{1}{\sqrt{T}} \sqrt{T} \\ &= C_1 + C_2 \frac{1}{\sqrt{\Delta(T)}} + C_3 \frac{1}{\sqrt{T}} + C_4 \end{aligned}$$

Since $\bar{\Delta}(T)$ is an upper bound of $\Delta(T)$, under reasonable assumptions $\bar{\Delta}(T)$ does not tend to zero, therefore:

$$E[\|\lambda_{T+1}\|] \leq \mathcal{O}(1)$$

Computing the gradient,

$$\nabla_x \hat{L}_{t,\sigma}(x, \lambda_t) = \nabla \hat{f}_t(x) + [\lambda_t + \sigma \hat{g}_t(x)]_+ \cdot \nabla \hat{g}_t(x)$$

The gradients are bounded:

$$\|\nabla \hat{f}_t(x)\| \leq G_f, \quad \|\nabla \hat{g}_t(x)\| \leq G_g$$

Meanwhile, according to Lemma 3, $\|\lambda_t\|$ is bounded, therefore,

$$\|\nabla_x \hat{L}_{t,\sigma}(x_{t+1}, \lambda_t)\| \leq G_f + G_g(\|\lambda_t\| + \sigma D)$$

Thus,

$$\|x_{t+1} - x_t\| \leq \frac{1}{\alpha_t}(G_f + G_g(\|\lambda_t\| + \sigma D))$$

Taking expectation,

$$E[\|x_{t+1} - x_t\|] \leq \frac{1}{\alpha_t}(G_f + G_g(E[\|\lambda_t\|] + \sigma D))$$

Substituting $\alpha_t = \alpha_0 \sqrt{t}$ and the bound from Lemma 3,

$$E[\|x_{t+1} - x_t\|] \leq \frac{1}{\alpha_0 \sqrt{t}}(G_f + G_g(C_1 + C_2 \alpha_0 + C_3 \sigma + C_4 \sigma s + \sigma D))$$

Summing from $t = 1$ to T ,

$$\sum_{t=1}^T E[\|x_{t+1} - x_t\|] \leq \frac{1}{\alpha_0} \sum_{t=1}^T \frac{1}{\sqrt{t}}(G_f + G_g(C_1 + C_2 \alpha_0 + C_3 \sigma + C_4 \sigma s + \sigma D))$$

Using the integral bound,

$$\sum_{t=1}^T \frac{1}{\sqrt{t}} \leq 1 + \int_1^T \frac{1}{\sqrt{t}} dt = 1 + 2(\sqrt{T} - 1) \leq 2\sqrt{T}$$

Therefore,

$$\sum_{t=1}^T E[\|x_{t+1} - x_t\|] \leq \frac{2\sqrt{T}}{\alpha_0}(G_f + G_g(C_1 + C_2 \alpha_0 + C_3 \sigma + C_4 \sigma s + \sigma D))$$

Substituting the parameter choices $\alpha_0 = \frac{1}{\sqrt{\bar{\Delta}(T)}}$ and $\sigma = \frac{1}{\sqrt{T}}$, and choosing $s = \sqrt{T}$,

$$\begin{aligned} \sum_{t=1}^T E[\|x_{t+1} - x_t\|] &\leq 2\sqrt{T\bar{\Delta}(T)}(G_f + G_g(C_1 + C_2 \frac{1}{\sqrt{\bar{\Delta}(T)}} \\ &\quad + C_3 \frac{1}{\sqrt{T}} + C_4 + D \frac{1}{\sqrt{T}})) \end{aligned}$$

Under reasonable assumptions (that $\bar{\Delta}(T)$ does not grow too fast), we have,

$$\sum_{t=1}^T E[\|x_{t+1} - x_t\|] \leq \mathcal{O}(\sqrt{T})$$

So we have,

$$\begin{aligned}
 \text{Vio}_i(T) &= E \left[\sum_{t=1}^T g_{t,i}(x_t) \right] \\
 &\leq \frac{1}{\sigma} E[\|\lambda_{T+1}\|] + G_g \sum_{t=1}^T E\|x_{t+1} - x_t\| \\
 &\leq \frac{1}{\sigma} \cdot \mathcal{O}(1) + G_g \cdot \mathcal{O}(\sqrt{T}) \\
 &= \sqrt{T} \cdot \mathcal{O}(1) + \mathcal{O}(\sqrt{T}) \\
 &= \mathcal{O}(\sqrt{T})
 \end{aligned}$$

4. Numerical Experiments

In this section, our algorithm is demonstrated to be capable of solving many real problems. Firstly, we explored the influence of some initial parameter values on our algorithm on different mathematical models. The best parameter combinations for each model are provided. Then we compare our algorithm with the PSGD algorithm in the same simulation environment. Secondly, we created a simulation experiment to test our algorithm. The performance of our algorithm when solving different real problems is observed. Then the results of the experiments are presented. Finally, we combined our algorithm with supervised learning for path planning. The results show that regret and constraint violation also have convergent bounds. This proves that our algorithm has the ability to solve online path planning problems.

In addition, in this paper, the model training was conducted on Python 3.12.6 and all the experiments were conducted on Matlab 2025b on a laptop with Windows 11 installed for fairness. The CPU of this laptop is AMD Ryzen AI 9 H 465 w/ Radeon 880M (2.00 GHz) and 32GB of RAM.

4.1. Comparative Experiment with the Existing Algorithm

We compared our algorithm with the PSGD algorithm using adaptive filtering and online logistic regression problems.

Adaptive filtering is a core recursive estimation technique in modern signal processing, system identification, and control. In many applications, the impulse response exhibits sparsity in the time domain. Meanwhile, we consider that physically realizable systems are all stable, so their impulse response energy must be finite. There are two constraints, based on the actual context of the problem. The mathematical model of adaptive filtering is as follows:

$$\begin{aligned}
 &\min_{x \in \mathbb{R}^n} a \cdot x + b \\
 &\text{s.t. } \|x\|_1 \leq \gamma_s, \quad \|x\|_2^2 \leq \gamma_e
 \end{aligned}$$

At each round, a and b are drawn from two independent normal distributions, whose means and standard deviations both vary smoothly over time in a sinusoidal or cosinusoidal manner.

Online logistic regression is a classic binary classification method in machine learning. It has significant application value in dynamic data stream environments. This problem is mathematically formulated as a constrained regularized empirical risk minimization problem. The objective function controls the model's complexity and prevents overfitting. The mathematical model of online logistic regression is as follows:

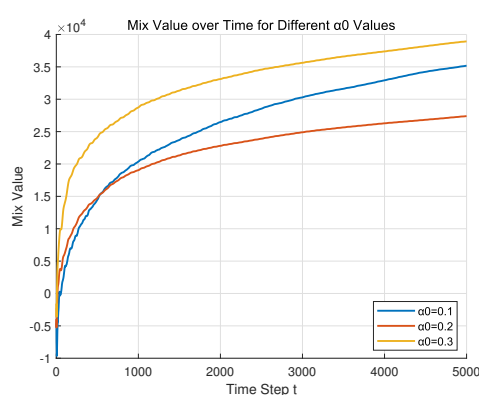
$$\begin{aligned}
 &\min_{\theta \in \mathbb{R}^n} \frac{1}{m} \sum_{i=1}^m [-y_i \log(\sigma(x_i^\top \theta)) - (1 - y_i) \log(1 - \sigma(x_i^\top \theta))] + \frac{\lambda}{2} \|\theta\|_2^2 \\
 &\text{s.t. } \|\theta\|_1 \leq \gamma_s, \quad \|\theta\|_2^2 \leq \gamma_e
 \end{aligned}$$

At each round, x_t is drawn from a normal distribution with fixed covariance, y_t is drawn from a Bernoulli distribution, and the true parameter vector $w_{\text{true}}(t)$ varies smoothly over time in a sinusoidal manner.

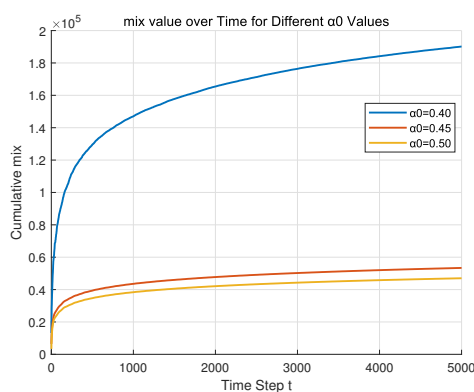
We explored the influence of initial parameter values on our algorithm. Different α_0 values in our algorithm were compared. There are two criteria for measuring the quality of an online algorithm, so a multi-objective planning approach was adopted to make a mixed measurement. We used AHP (Analytical Hierarchy Process) to determine the weights of the two quantities. The two measurements correspond to scale value 3 in the 1-9 scale method. So the weight of regret is 0.25 and the weight of constraint violation is 0.75. Our mixed measurement mix is defined as follows:

$$mix = 0.25 * regret + 0.75 * violation$$

In Figure 1, in adaptive filtering, we found that the algorithm performs best when $\alpha_0 = 0.2$. And in online logistic regression, the algorithm performs best when $\alpha_0 = 0.5$.



(a)



(b)

Figure 1. We plotted the images of the mix quantity under different α_0 to determine which parameter can achieve better performance of the algorithm. (a) The mixed value of adaptive filtering (b) The mixed value of online logistic regression.

Then we compared our algorithm with PSGD algorithm. The comparison results are as follows:

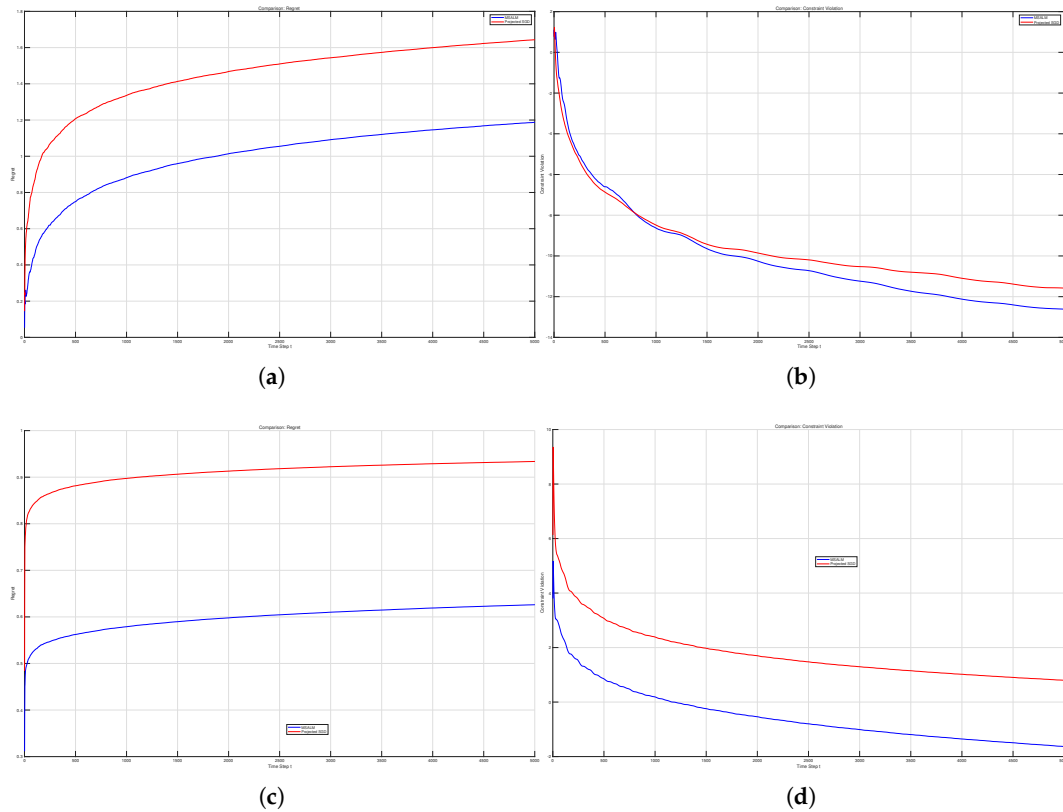


Figure 2. We compared our algorithm with PSGD algorithm. (a) The regret of adaptive filtering. (b) The constraint violation of adaptive filtering. (c) The regret of online logistic regression. (d) The constraint violation of online logistic regression.

we compared the performance of two algorithms in adaptive filtering and online logistic regression. The figures show that MSALM attains lower regret and constraint violation bounds than PSGD. This result indicates that our algorithm has the ability to converge more rapidly toward the theoretical optimum while adhering better to the constraints. According to the experimental results, the conclusion can be obtained. When choosing a suitable α_0 , MSALM is superior to the existing algorithm.

4.2. Experiments Under Existing Models

In order to test the practicality of our algorithm, we applied our algorithm in energy dispatch problem and network resource allocation.

Time-varying smart grid energy dispatch problem is one of the core tasks of modern smart grids. This problem is to economically and reliably coordinate multiple heterogeneous energy sources while meeting the changing electricity demand. The objective functions of this problem include power balance constraints and resource capacity constraints. The mathematical model of time-varying smart grid energy dispatch is as follows:

$$\begin{aligned} & \min_{x \in \mathbb{R}^n} c^\top x + \frac{1}{2} x^\top Q x \\ \text{s.t. } & \sum_{i=1}^{n_g} x_i - \sum_{j=1}^{n_s} x_{n_g+j} - \sum_{k=1}^{n_d} x_{n_g+n_s+k} = d \\ & 0 \leq x_i \leq u_i, \quad i = 1, \dots, n \end{aligned}$$

At each round, c , Q are drawn from independent normal distributions with fixed standard deviations. Their means evolve over time according to a random walk with a decaying step size.

Online network resource allocation is a key challenge in modern computing systems. The key task is to efficiently allocate limited resources to continuously arriving real-time tasks or user requests. The objective function consists of two parts: Quality of Service cost and resource usage cost. The mathematical model of online network resource allocation is as follows:

$$\min_{x \in \mathbb{R}^n} \sum_{i=1}^n (d_i - x_i)^2 + \alpha \sum_{i=1}^n x_i$$

$$\text{s.t. } x_i \leq c_i, \quad i = 1, \dots, n$$

At each round, the demand vector is drawn from a multivariate normal distribution with fixed standard deviation, and is then truncated to be nonnegative. Its mean vector varies smoothly over time following a sinusoidal pattern with a slow linear trend.

We conducted simulation experiments on two mathematical models, and the results are as follows,

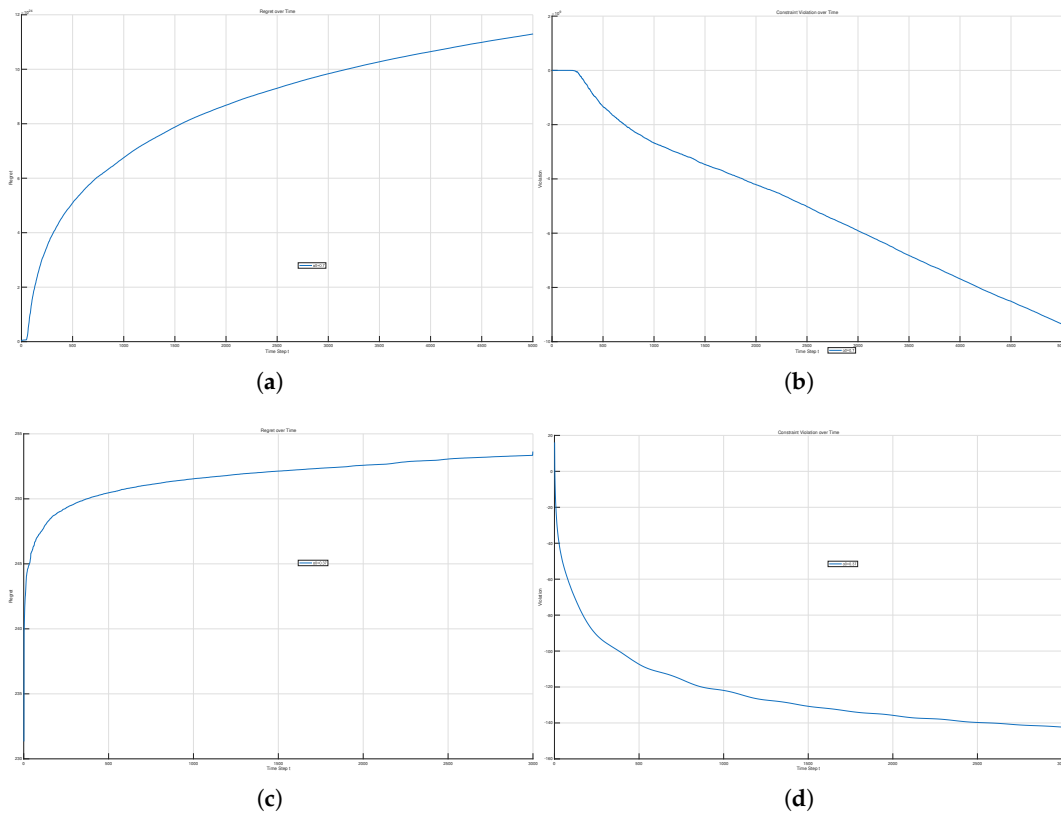


Figure 3. Our algorithm is applied to energy dispatch problem and online network resource allocation. (a) Regret of time-varying smart grid energy dispatch. (b) Constraint violation of time-varying smart grid energy dispatch. (c) Regret of online network resource allocation. (d) Constraint violation of online network resource allocation.

We took $\alpha_0 = 0.7$ and $\alpha_0 = 0.37$ for testing, respectively. The simulation experiments show that regret and constraint violation have a sublinear convergent bound of T , demonstrating that our algorithmic solution can gradually approach the theoretically optimal solution while adhering to the constraints. The results show that our algorithm can be applied to time-varying smart grid energy dispatch problem and the online network resource allocation problem.

4.3. Experiment Combining Our Algorithm with Supervised Learning

We combined MSALM with supervised learning to solve path planning problem. For obtaining an explicit function that is used in our algorithm, we applied parameter regression methods in supervised learning. First of all, 10 flight trajectories' data for the same flight at the same time slot on different dates were randomly selected. Afterwards, we utilized these data to train a mathematical model for

flight trajectories through parameter regression methods.B-spline interpolation was used to obtain the explicit function of the flight trajectory. In order to find the appropriate number of B-spline interpolation control points, we compared the errors between the fitted trajectory and 10 known trajectories under different numbers of control points. The results are as follows:

Table 1. Fitting error with different numbers of control points.

Number of Control Points	3D RMSE (m)	Mean Error (m)	Max Error (m)
6	16212.93	14630.05	42953.94
7	15278.20	13697.46	28088.96
8	10301.73	9577.57	20571.04
9	7767.29	7094.96	15264.24
10	6056.13	5466.36	12094.13
11	6474.88	5956.19	12599.98
12	6416.03	5793.27	12159.70

According to the above table, the accuracy reaches its highest when the trajectory has 10 control points. So we obtained the fitting function in this case. The coordinates of the feature points are as follows:

Table 2. Coordinates of the control points.

Control Point	X (m)	Y (m)	Z (m)
1	478288.923	4423806.307	1199.714
2	483007.653	4412762.620	1939.867
3	491176.780	4411013.545	1022.683
4	540450.109	4386975.935	6298.608
5	613313.197	4283290.653	6170.731
6	753322.541	4268797.429	6188.807
7	792318.111	4171779.364	6157.780
8	773451.251	4061749.319	3778.580
9	782481.198	4033694.042	2106.506
10	782988.940	4014086.869	1142.003

We set some random situations as constraints to simulate the real flight conditions. The constraints include 9 dynamic obstacle avoidance constraints and 4 airspace boundary constraints. The mathematical model of dynamic obstacle avoidance constraints has the following form:

$$g_{k,i} = d_{safe,k}^2(t) - ||p_i - c_k(t)||^2 \leq 0$$

where $c_k(t)$ is the obstacle center coordinates, $d_{safe,k}(t)$ is the safe distance, p_i is the control point with obstacle avoidance constraint. We set 3 obstacles and the 1st, 5th, and 10th control points with obstacle avoidance constraints, so $k = 1, 2, 3$ and $i \in \{1, 5, 10\}$.

The mathematical model of airspace boundary constraints has the following form:

$$x_{min} \leq x \leq x_{max}$$
$$y_{min} \leq y \leq y_{max}$$

Then we use our algorithm to solve this online optimization problem. The result of the experiment is as follows:

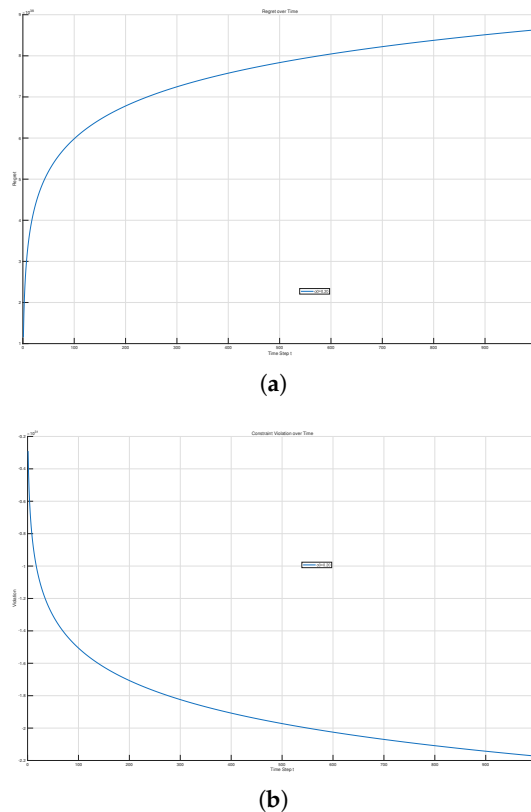


Figure 4. This is a figure in which our algorithm is applied to flight path planning. (a) The regret of flight path planning (b) The constraint violation of flight path planning.

In this figure, the results of the simulation experiment show that regret and constraint violation of our algorithm have a sublinear convergent bound of T , demonstrating progressive convergence to the theoretical optimum under compliance with the constraint. This indicates that the aircraft can make decisions that gradually approach the optimal choice during the flight and stay as safe as possible in the face of emergencies. This proves that our algorithm can combine with machine learning methods to be applied to similar path planning problems.

5. Conclusions

In this paper, we present a new online stochastic augmented Lagrangian method, the MSALM, for solving online stochastic optimization problems with time-varying distributions. At each round, we construct the model functions for the objective and constraint functions based on their properties, which reduces the computational complexity. The step size is designed in a dynamic form and decreases as t increases to accelerate convergence. Under standard assumptions, we proved that our algorithm achieves sublinear regret and constraint violation bounds of T . Simulation experiments demonstrate the performance and practical utility of the MSALM algorithm. Additionally, in the context of path planning, we combined our algorithm with supervised learning to further demonstrate its extensibility.

References

1. Zinkevich, M. Online convex programming and generalized infinitesimal gradient ascent. In *Proceedings of the 20th International Conference on Machine Learning*; 2003; pp. 928–935.
2. Shalev-Shwartz, S. Online learning and online convex optimization. *Found. Trends Mach. Learn.* **2012**, *4*, 107–194.
3. Hazan, E. Introduction to online convex optimization. *Found. Trends Optim.* **2016**, *2*, 157–325.
4. Chen, T.; Ling, Q.; Giannakis, G. B. An online convex optimization approach to proactive network resource allocation. *IEEE Trans. Signal Process.* **2017**, *65*, 6350–6364.

5. Zhang, Y.; Dall’Anese, E.; Hong, M. Online proximal-ADMM for time-varying constrained convex optimization. *IEEE Trans. Signal Inf. Process. Netw.* **2021**, *7*, 144–155.
6. Tsuchiya, T.; Ito, S. Fast rates in stochastic online convex optimization by exploiting the curvature of feasible sets. In *Advances in Neural Information Processing Systems 37 (NeurIPS 2024)*; 2024.
7. Cao, X.; Zhang, J.; Poor, H. V. Online stochastic optimization with time-varying distributions. *IEEE Trans. Autom. Control* **2020**, *66*, 1840–1847.
8. Kiefer, J.; Wolfowitz, J. Stochastic estimation of the maximum of a regression function. *The Annals of Mathematical Statistics* **1952**, *23*, 462–466.
9. Johnson, R.; Zhang, T. Accelerating stochastic gradient descent using predictive variance reduction. In *Advances in Neural Information Processing Systems* **2013**, *26*, 315–323.
10. Qi, R.; Xue, D.; Zhai, Y. A momentum-based adaptive primal–dual stochastic gradient method for non-convex programs with expectation constraints. *Mathematics* **2024**, *12*, 2393.
11. Arrow, K. J.; Hurwicz, L.; Uzawa, H. *Studies in Linear and Non-Linear Programming*; Stanford University Press: Stanford, CA, USA, 1958.
12. Yu, H.; Neely, M. J.; Wei, X. Online convex optimization with stochastic constraints. In *Advances in Neural Information Processing Systems*; 2017; pp. 1428–1438.
13. Lesage-Landry, A.; Wang, H.; Shames, I.; Mancarella, P.; Taylor, J. A. Online convex optimization of multi-energy building-to-grid ancillary services. *IEEE Trans. Control Syst. Technol.* **2020**, *28*, 2416–2431.
14. Liu, H.; Xiao, X.; Zhang, L. Augmented Lagrangian methods for time-varying constrained online convex optimization. *J. Oper. Res. Soc. China* **2025**, *13*, 364–392.
15. Rockafellar, R. T. Augmented Lagrangians and applications of the proximal point algorithm in convex programming. *Math. Oper. Res.* **1976**, *1*, 97–116.
16. Suo, W.; Li, W.; Zhang, B.; Liu, Y. Distributed online convex optimization with multiple coupled constraints: A double accelerated push-pull algorithm. *J. Franklin Inst.* **2024**, *361*, 106884.
17. Hou, R.; Li, X.; Shi, Y. Online composite optimization with time-varying regularizers. *J. Franklin Inst.* **2024**, *361*, 106884.

Disclaimer/Publisher’s Note: The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.