Application of the Mirror Descent Method to Minimize Average Losses Coming by a Poisson Flow*

Alexander Nazin1, Svetlana Anulova2, and Andrey Tremba1

Abstract—We treat a convex problem to minimize average loss function for a stochastic system operating in continuous time. The losses on time horizon $T$ arise at the jump times of a Poisson process with intensity being an unknown random process. The oracle gives randomly noised gradients of the loss function; the noises are additive, unbiased, with the bounded dual norm in average square sense. The goal consists in minimizing the average integral loss over a given convex compact set in the $N$-dimension space. We propose a mirror descent algorithm and prove an explicit upper bound for the average integral loss regret. The bound is of type “square root of $T$” with an explicit coefficient. Finally, we describe an example of optimization for a server processing a stream of incoming requests, and we discuss simulation results.

I. INTRODUCTION

The Mirror Descent Method (MDM) [5] represents a non-trivial generalization of a standard gradient scheme also treating as a primal-dual method for convex optimization [1], [7]. Essentially, it may give an advantage in hard situations with large dimension ensuring sub-optimal rate of convergence. It finds applications in different areas like tomography [2], classification [3], PageRank problem [4], [9], for instance.

MDM usually is implemented for a discrete-time models when the oracle gives observations sequentially as the time goes. In this paper, we treat a stochastic control system living in continuous time, and its current loss is observable at random discrete instants $\tau_i$ arising as a jump times of a Poisson process. The main result for the designed mirror descent algorithm which is proved in Section III is as follows. The Upper Bound for the average integral loss regret on any time horizon $T$ is of type $C\sqrt{T}$ with an explicit constant $C$ given. Numerical example of Section IV illustrates the theoretical result.

II. PROBLEM STATEMENT

A. Preliminary Consideration and Notation

Let us consider a stochastic control system which operates on a sequential random time intervals $[\tau_i, \tau_{i+1})$, such that the current system loss on $i$-th interval are observable, where $0 < \tau_1 < \tau_2 < \tau_3 \ldots$. The general idea is to minimize the average integral loss of the system at the given time horizon $T$ assuming that the current loss $\pi(i)$ depends on a control parameter $\theta(i)$ within a given convex compact set $\Theta \subset \mathbb{R}^N$, and the conditional expected loss $E\{\pi(i)|\theta(0), \ldots, \theta(i-1)\}$ represents a convex function $Q(\theta(i-1))$ of the control parameter. We assume that function $Q$ is unknown a priori, and the available current information corresponds to an oracle of a gradient type. The oracle output represents a stochastic subgradient for the current interval $[\tau_i, \tau_{i+1})$ and may be treated as noised subgradient of loss function $Q$.

To be more precise, introduce the following notation. The set indicator for set $A$ means $I_A(y)$ that equals 1 if $y \in A$ and 0 otherwise. The primal space $E = \mathbb{R}^N$ would be considered with a norm $\|x\|$ (as an example, it might be Euclidean norm $\|x\| = \sqrt{x_1^2 + \cdots + x_N^2}$); therefore, the dual space $E^* = \mathbb{R}^N$ should be equipped with norm $\|x\|_* = \sup_{\|x\| \leq 1} \langle x, y \rangle$, where $\langle \cdot, \cdot \rangle$ stands for the inner product.

Let $(\Omega, \mathcal{G}, \mathbb{P})$ be a probability space with two independent objects defined on it:

1) Poisson process $X_t, t \in [0, \infty)$, with intensity being a random process;

2) a noise sequence of i.i.d. unbiased square-integrable random variables $\xi(1), \xi(2), \ldots$ with values in $E^*$.

Denote $\tau_1, \tau_2, \tau_3, \ldots$ the sequence of jump times of the process $X$, and define $\tau_0 = 0$. Further, introduce a continuous-time process $\xi_t = \sum_{i=1}^{\infty} \xi(i) I_{\tau_i} (t), t \in [0, \infty)$ (in the indicator formula the point set $\tau_i$ is regarded as a subset of $[0, \infty), i = 1, 2, \ldots$).

B. Loss Function Properties and Control Problem Setup

Let the set of control parameters be a convex compact $\Theta \subset E$. The loss function is a continuous convex function $Q: \Theta \rightarrow \mathbb{R}_+$ with a subgradient $\partial Q: \Theta \rightarrow E^*$ (a bounded Borel vector-function) satisfying, therefore, the condition

$$\langle \partial Q(\theta), \theta' - \theta \rangle \leq Q(\theta') - Q(\theta), \quad \forall \theta, \theta' \in \Theta.$$ 

The loss function continuity implies the boundedness from below on a compact set, and there exists $\theta^* \in \Theta$ such that

$$Q^* = \min_{\theta \in \Theta} Q(\theta) = Q(\theta^*). \quad (1)$$

Now we proceed to statement of the control problem. The set of control strategies $\mathcal{U}$ consists of all possible piecewise constant random processes $\{\theta_t\}$ with values in $\Theta$ of the form

$$\theta_t = \sum_{i=1}^{\infty} \theta(i) I_{[\tau_i, \tau_{i+1})}(t),$$

defined as follows. Non-random $\theta_0 \in \Theta$ is chosen arbitrary; further define inductively

$$g(i) = \partial Q(\theta(i-1)) + \xi_{\tau_i}, \quad (2)$$

where $\theta(i)$ is an arbitrary measurable function of random variables $g(1), \ldots, g(i)$ with the values in $\Theta, i = 1, 2, \ldots$.  

---

*This work was supported by Russian Foundation for Basic Research under grant 12-08-01245.

1Alexander Nazin and Andrey Tremba are with Ya.Z. Tsypkin Laboratory of Adaptive and Robust Systems, V.A. Trapeznikov Institute of Control Sciences RAS nazine@ipu.ru, atremba@ipu.ru

2Svetlana Anulova is with the Laboratory of Statistical Information Treatment, V.A. Trapeznikov Institute of Control Sciences RAS anulovas@ipu.ru
Next specify the time horizon $T > 0$ and define the functional (control criterion) as follows:

$$
R_T(\{\theta_t\}) \triangleq E^0 \int_0^T Q(\theta_t) dX_t = E^0 \sum_{i=1}^{X_T} \eta(i). \tag{3}
$$

Recall that summation in RHS (3) reduces to 0 if $\{X_T = 0\}$.

The ideal optimization control problem is to minimize the functional over strategy set $\mathcal{U}$. Evidently, for any $u \in \mathcal{U}$ holds

$$
R_T(u) \geq Q^*EX_T.
$$

On the other hand, for a static control strategy $u = \{u_t = \theta, t \in [0, \infty)\}$ with a constant $\theta \in \Theta$ it holds

$$
R_T(u) = Q(\theta)EX_T;
$$

therefore, $\inf_{u \in \mathcal{U}} R_T(u)$ over all steady-state strategies $u$ equals $Q(\theta^*)EX_T$. Thus,

$$
\inf_{u \in \mathcal{U}} R_T(u) = Q^*EX_T. \tag{4}
$$

It means that the optimal steady-state strategy gives the minimum of control functional prescribed by value RHS (4). So, the goal is to design such implementable control strategy for which average integral loss at arbitrary horizon $T$ would be as close to $Q^*EX_T$ as possible; if such a control strategy is designed, it may be called the suboptimal control strategy.

### C. Further Assumptions

The following assumptions are stated for a given positive constants $L, \lambda, \sigma$:

A1. The loss function subgradient is bounded, 

$$
\sup_{\theta \in \Theta} \|\partial Q(\theta)\|_* \leq L; \tag{5}
$$

A2. The Poisson process intensity being a random non-anticipating process is bounded, that is $\lambda_x \in (0, \lambda]$ a.s.;

A3. $E\xi(i) = 0$ and $E\|\xi(i)\|^2 \leq \sigma^2$ for all $i = 1, 2, \ldots$.

In what follows, we describe a control strategy based on MDM. Then we will prove that it is suboptimal in a sense of the upper bound demonstrated in Theorem 1.

### III. MIRROR DESCENT METHOD (ALGORITHM)

Remind that the Mirror Descent Method is a primal-dual method [5], [7], [3], [6]. Primal and dual spaces are equipped with the corresponding norms as described in the previous section. The method contains a convex continuously differential function as a functional parameter $W_\beta : E^* \to \mathbb{R}$, incorporating a scalar parameter $\beta$. For this function Lipschitz condition should hold: $\forall \beta > 0$,

$$
\|\nabla W_\beta(z) - \nabla W_\beta(\tilde{z})\| \leq \frac{1}{\alpha \beta} \|z - \tilde{z}\|_* , \quad \forall z, \tilde{z} \in E^*, \tag{6}
$$

with some constant $\alpha > 0$ which does not depend on $\beta$.

### A. MD Continuous Time Algorithm

For the presented control problem, the Mirror Descent (MD) algorithm can be written in the continuous time form as follows.

1. Fix positive $\beta_0$ and sequence $\beta_i = \beta_0 \sqrt{i+1}, i = 0, 1, \ldots$ Fix pair of initial dual and primal variables as $\phi_0 = 0 \in E^*$ and $\theta_0 = -\nabla W_{\beta_0}(\phi_0)$, respectively.

2. For each $t > 0$ (as time $t$ increases continuously begining from zero), having parameter $\theta_t$, get oracle output $g_t = \sum_{i=1}^{\infty} g(i) I(t, \tau_{i, \tau_{i+1}}(t))$ using (2). Also form a continuous-time parameter

$$
\tilde{\beta}_t = \sum_{i=1}^{\infty} \beta_i I(t, \tau_{i, \tau_{i+1}}(t)).
$$

Form a trajectory of stochastic differential equation in $E^*$ alongside with algebraic equation which maps $E^*$ to $\Theta$ (see Proposition 1, item 2(ii) below) by

$$
d\phi_t = dg_t, \tag{7}
$$

$$
\theta_t = -\nabla W_{\tilde{\beta}_t}(\phi_t). \tag{8}
$$

3. At arbitrary horizon $T$ we get realization of loss functional (3) on strategy $\{\theta_t\}$ and, if loss function $Q$ is given, then integral loss is calculated

$$
\hat{R}_T(\{\theta_t\}) = \sum_{i=1}^{T} \eta(i),
$$

where $\tau_i$ are the jump times of process $X_t$.

**Remark 1:** Notice that the trajectory $\{\zeta_t\}$ of stochastic differential equation (7) may be formally written as follows:

$$
\zeta_t = \sum_{i=1}^{\infty} g(i) I(t, \tau_{i, \tau_{i+1}})(t), \quad t \in [0, T].
$$

### B. Proxy-Function Convexity Properties

For completeness let’s make some definitions and a proposition, explaining and expanding the essence of the MDM due to the convexity properties of proxy-function and its $\beta$-conjugate (cf. [3], section 3 and the references therein for details).

**Definition 1:** Let be $\alpha > 0$. Convex function $V : \Theta \to \mathbb{R}$ is called $\alpha$-strongly convex with respect to primal norm $\| \cdot \|$ , if

$$
V(sx + (1-s)y) \leq sV(x) + (1-s)V(y) - \frac{\alpha}{2} s(1-s)\|x-y\|^2
$$

for all $x, y \in \Theta$ and $s \in [0, 1]$.

**Proposition 1:** Let $V : \Theta \to \mathbb{R}$ be convex, and introduce a parameter $\beta > 0$. Then $\beta$-conjugate to $V$ is function

$$
W_\beta(z) = \sup_{\theta \in \Theta} \{ -(z, \theta) - \beta V(\theta) \}, \quad \forall z \in E^*, \tag{10}
$$

having the following properties:

1. Function $W_\beta : E^* \to \mathbb{R}$ is convex, and its conjugate function is $\beta V$, i.e.,

$$
\forall \theta \in \Theta, \quad \beta V(\theta) = \sup_{z \in E^*} \{ -(z, \theta) - W_\beta(z) \}. \tag{11}
$$
2. If function $V$ is $\alpha$-strongly convex with respect to primal norm $\| \cdot \|$, then
   (i) condition (6) holds,
   (ii) $\forall \beta \in \Theta$,
   \[ \min_{\beta \in \Theta} V(\beta) = V(\theta_\ast), \]
   (ii) condition (6) is valid for $\beta$-conjugate to $V$, that is for function $W_\beta$. □

**Definition 2:** A function $V : \Theta \to \mathbb{R}_+$ is called a proxy function, if it is convex and
   (i) there exists a point $\theta_\ast \in \Theta$ such that
   \[ \min_{\theta \in \Theta} V(\theta) = V(\theta_\ast), \]
   (ii) condition (6) is valid for $\beta$-conjugate to $V$, that is for function $W_\beta$.

Examples of proxy functions for minimization on a simplex are presented in [6] namely quadratic and entropy ones (see p. 1586).

**C. Main Result**

**Theorem 1:** Let the MD continuous-time algorithm described above with the proxy-function $V : \Theta \to \mathbb{R}_+$ having $\beta$-conjugate $W_\beta$ under condition (6) realizes the strategy \{\theta_j\} under assumptions A1–A3 of Section II-C. Then, for any horizon $T > 0$, the following inequality holds:

\[ \mathbb{E}\hat{R}_T(\{\theta_j\}) - \min_{\theta \in \Theta} R_T(\{\theta\}) \leq 2\sqrt{\frac{2(T\lambda + 1)\lambda(L^2 + \sigma^2)}{\alpha}}. \]

Parameter $\alpha$ is described in (9), and

\[ \beta_0 = \sqrt{\frac{2L^2 + \sigma^2}{\alpha \lambda}}, \quad \nabla \geq \max_{\theta \in \Theta} V(\theta). \]

□

**IV. EXAMPLE**

Consider a server processing incoming requests at Poisson-distributed instants with random intensity. A request contains tasks of $n$ different types, and each $j$-th type has its own random nonnegative processing cost $\eta_j$, $j = 1, \ldots, n$ (with unknown $\mathbb{E}\eta_j$, bounded second moments $\mathbb{E}(\eta_j^2) < \infty$ and independent to other task’s costs).

Coming at $i$-th time of instant $\tau_i$ request $\eta(i)$ can be characterized by vector of costs’ realizations $(\eta_1(i), \ldots, \eta_n(i))$. During requests each $j$-th type’s costs $\eta_j(i)$ form iid sequence.

The goal is to minimize total (integral) processing cost, counted over all task types and all events until horizon. For each request the server is allowed to redirect some tasks to other server, but in timeline average not more than small fraction $\delta$ of tasks is redirected. It is important that processing costs in a request are not known at the request’s arrival time and cannot be used for redirect decision, but become known afterwards and can be used for adjusting control.

Let’s apply following randomized redirection control strategy, described by vector of probabilities $\theta = (\theta_1, \ldots, \theta_n)$ of redirecting tasks with corresponding types. After $i$-th request comes each $j$-th type task is being redirected with probability $\theta_j(i - 1)$, then non-redirected tasks are processed (revealing costs $\eta_j(i)$) and $\theta$ is updated.

Formally the problem is to minimize

\[ R_T = \mathbb{E} \sum_{i: 0 < \tau_i \leq T} \sum_{j = 1}^n \eta_j(i), \]

where

\[ \eta_j(i) = \begin{cases} \eta_j(i), & \text{if } j\-th type task at } i\text{-th request was processed on the server}, \\ 0, & \text{if the task was redirected,} \end{cases} \]

under constraint

\[ \Theta = \{\theta : \sum_{j = 1}^n \theta_j \leq \delta n, \ 0 \leq \theta_j \leq 1, \ j = 1, \ldots, n\}. \]

Let’s also suppose that in average not more than one task is redirected, i.e. $\delta n < 1$. It means that sometimes no task redirected, thought it may happens that few tasks of one request are redirected.

It is evident that the optimal strategy would be to somehow determine the type with maximal expectation, and to redirect tasks of this type with probability $\delta n$ at each request. But in process of determining such type total processing cost may become large, while suggested MDM provide bound for total cost directly. Equation (14) has the same type as (3) with loss function

\[ Q(\theta) = \sum_{j = 1}^n (1 - \theta_j)\mathbb{E}\eta_j. \]

The gradient of the latter linear function is $(-\mathbb{E}\eta_1, \ldots, -\mathbb{E}\eta_n)$. It can be estimated only via realized costs of the just processed $i$-th request only, i.e. (2) expressed as

\[ g(i) = \left( \frac{-\eta_1(i)}{1 - \theta_1(i - 1)}, \ldots, \frac{-\eta_n(i)}{1 - \theta_n(i - 1)} \right)^T. \]

Decision and processing time intervals are considered insignificantly small, so that all data are processed until next request comes.

To apply MDM on non-unit simplex (15) we used proxy-function (cf. [3])

\[ V(\theta) = -\delta n \log \delta + \sum_{j = 1}^n \theta_j \log \theta_j, \]

with strong convexity parameter $\alpha = 1/(\delta n)$ and maximum value on the simplex $\nabla = \delta n \log n$. Primal norm is $\ell_1$-norm and dual one is $\ell_\infty$-norm.

**A. Simulation**

The numeric example has $n = 100$, $T = 1000$, $\lambda = 2$, $\delta = 0.004$, and each $j$-th type’s processing cost distribution was randomly chosen among

- uniform distribution on nonnegative interval $[a_j, b_j]$, or
- uniform discrete distribution on integer nonnegative interval $[c_j, d_j]$. 2196
After choosing distribution, its parameters were chosen randomly as well. Poisson process intensity is chosen uniformly on interval $[1, 8, 2]$ at each jump time being constant until next jump occurs.

On Figure 1 several realizations of differences between total processing costs $\hat{R}_t$ and optimal processing costs $\mathcal{R}_t^*$ = inf$_{u \in \mathcal{U}} R_t(u)$ alongside with its average and upper bound are plotted. Upper bound is calculated by (12) using parameters of the task costs distributions. Other experiments show that when increasing $\delta$ to $1/n$ the upper bound less tight to the algorithm’s runs, and more tight when $\delta$ is decreasing.

**APPENDIX**

Here goes short proof of Theorem 1. One may write

$\mathbb{E} \hat{R}_T(\{\theta_i\}) = \inf_{u \in \mathcal{U}} R_T(u)$

$= \mathbb{E} \left[l_{\{x_T \geq 1\}}(0) \sum_{i=1}^{X_T} Q(\theta_{x_i}) - X_T Q(\theta^*) \right]$

$= \mathbb{E} \sum_{n=1}^{\infty} l_{\{x_T = n\}}(\omega) \left\{ \sum_{i=1}^{n} (Q(\theta_{x_i}) - nQ(\theta^*)) \right\}.$

Conditioned on event $\{\omega : x_T = n\}$, the random variables $\xi_{x_1}, \xi_{x_2}, \ldots, \xi_{x_n}$ are distributed in just the same way, as $\xi(1), \xi(2), \ldots, \xi(n)$, and the algorithm works like MDM in the discrete time model of article [3]. Therefore the conditional expectation of variable $\sum_{i=1}^{n} Q(\theta_{x_i}) - nQ(\theta^*)$ depends only on $n$, and we define it $f(n)$. Remark that function $f$ is non-decreasing since

$\sum_{i=1}^{n} Q(\theta_{x_i}) - nQ(\theta^*) = \sum_{i=1}^{n} (Q(\theta_{x_i}) - Q(\theta^*))$

and every summand in RHS is a non-negative random variable due to (1). Thus

$\mathbb{E} \hat{R}_T(\{\theta_i\}) = \inf_{u \in \mathcal{U}} R_T(u)$

$= \mathbb{E} \sum_{n=1}^{\infty} l_{\{x_T = n\}}(\omega) f(n)$

$= \mathbb{E} f(X_T)$.

Now we remark that the expectation in the right hand part does not exceed the expectation for another Poisson process, namely with a constant intensity $\overline{\lambda}$. Argumentation (cf. [8] section 1.A, in particular, formula (1.A.7)): this process may be presented as a sum of the considered one and another one, with the intensity $\overline{\lambda} - \lambda_x$, therefore the distribution function of this process is not greater than that of the considered process; the function $f$ is non-decreasing. Thus we can use for the upper bound of the right hand side the upper bound for the Poisson process with the constant intensity equal to $\overline{\lambda}$. Now apply Propositions 2 and 3 from [3] estimating mathematical expectations of the approximation error with $\gamma_t \equiv 1$ and receive

$f(n) \leq \beta_0 \overline{V} + \sum_{i=0}^{n-1} \frac{L^2 + \sigma^2}{\alpha \beta_i}$

$\leq \sqrt{n+1} \left( \beta_0 \overline{V} + \frac{L^2 + \sigma^2}{\alpha \beta_0} \right).$

Hence, by Jensen inequality and by the “optimal” parameter $\beta_0$ defined in (13), we obtain

$\mathbb{E} f(X_T) \leq 2\beta_0 \overline{V} \mathbb{E} \sqrt{X_T} + 1 \leq 2\beta_0 \overline{V} \sqrt{X_T} + 1.$

We come to the desired inequality (12). The theorem is proved.

**ACKNOWLEDGMENT**

The authors thank anonymous referees for valuable notes.

**REFERENCES**


