Mirror Descent Algorithms in the Problems of Convex Optimization and Control

NAZIN A.V.

Institute of Control Sciences RAS, Moscow

nazine@ipu.ru, anazine@rambler.ru

June 20, 2012

IV Traditional Youth School "Control, Information, and Optimization"

Zvenigorod, 17–24 June 2012

Outline

1. Introduction to Mirror Descent Method (MDM).
2. General consideration and convex optimization problem.
3. Estimation of the Main Eigenvector of a Stochastic Matrix.
4. Multi-Armed Bandit Problem via MDM.
5. Conclusions.
6. Some References.

Introduction

MDA is a gradient-type recursive method for convex optimization, i.e. primal-dual method performing the descent in a dual space and mapping the resulted points to a primal space. See the following references:

1. Nemirovski and Yudin (1979/1983): [1]
2. Ben-Tal, Margalit, and Nemirovski (2001): [2]
3. Beck and Teboulle (2003): [3]
4. Nesterov (2005, 2007): [4], [5]
5. Juditsky, Nazin, Tsybakov, and Vayatis (2005): [6]
6. Juditsky, Lan, Nemirovski, and Shapiro (2007): [7]

1 Idea behind MDM (continuous time) [1]

Consider a primal-dual method, that is MDM:

$$\dot{\xi}(t) = -\nabla_x f(x(t)), \quad \xi(0) = \xi_0, \tag{1}$$

$$x(t) = \nabla_\xi W(\xi(t)), \quad t \geq 0. \tag{2}$$

Here:

- f is a convex function to be minimized in Banach space E,
- W is a uniform differentiable, convex function on dual space E^* .

For instance, “Euclidean” case of

$$W(\xi) = \frac{1}{2}\|\xi\|_2^2$$

gives a well-known standard gradient method

$$\dot{x}(t) = -\nabla_x f(x(t))\,.$$

Let us look at a simple analysis as follows. Assume

$$x^* = \arg\min f(x).$$

Then we have a candidate Lyapunov function

$$W_*(\xi) \triangleq W(\xi) - < \xi, x^* >,$$

since

$$\frac{dW_*(\xi(t))}{dt} = <\dot{\xi}(t), \nabla_\xi W(\xi(t)) - x^* > \qquad (3)$$

$$= - < \nabla_x f(x(t)), x(t) - x^* > \qquad (4)$$

$$\leq f(x^*) - f(x(t)) \qquad (5)$$

$$\leq 0\,, \qquad (6)$$

that is function $W_*(\xi)$ decreases along the trajectory $\{\xi(t)\}$. Furthermore, (3)–(5) lead to

$$f(x(t)) - f(x^*) \leq <\dot{\xi}(t), x^* > - \frac{dW(\xi(t))}{dt}\,, \qquad (7)$$

and, assuming that

$$\xi(0) = 0, \quad W(0) = 0,$$

and integrating by $t \in [0, T]$, we get

$$\int_0^T f(x(t))dt - Tf(x^*) \leq \quad < \xi(T), x^* > -W(\xi(T)) \quad (8)$$

$$\leq \quad V(x^*) \quad (9)$$

with the Legendre transformation

$$V(x) \triangleq \sup_{\xi}\{< \xi, x > -W(\xi)\}.$$

Now, introduce the average estimate

$$\widehat{x}(T) \triangleq \frac{1}{T}\int_0^T x(t)dt\,.$$

By Iensen lemma and convexity of $f(x)$, eqs (8)–(9) lead to

$$f(\widehat{x}(T)) - f(x^*) \quad \leq \quad \frac{1}{T}V(x^*). \tag{10}$$

Resume:

- Function $W : E^* \to \mathbb{R}$ is a parameter of MDM which ensures the Lyapunov function $W_* : E^* \to \mathbb{R}$; in particular, MDM reduces to standard gradient method; therefore, this additional degree of freedom may improve the accuracy algorithm, at least potentially.
- MDM leads to the average estimate $\widehat{x}(t)$, i.e. time-average to current estimates over the time interval $[0, t]$.

- Non-asymptotical upper bound on difference between current estimation function $f(\widehat{x}(t))$ and function minimum $f(x^*)$ is ensured; this upper bound is of type $O(T^{-1})$, and it is directly depending on $V(x^*)$; therefore, the given class function has to ensure the finite upper bound $\sup V(x)$. (Thus, further consideration is reduced to function minimization over a given compact convex set.)

- The previous consideration shows the role of Legendre transformation.

2 A Generalized View-Point

Proxy functions. Denote by $E = \ell_1^M$ the space $\mathbb{R}^M$ with the 1-norm

$$\|z\|_1 = \sum_{j=1}^{M} |z^{(j)}|$$

and by $E^* = \ell_\infty^M$ the dual space which is $\mathbb{R}^M$ equipped with the sup-norm

$$\|z\|_\infty = \max_{\|\theta\|_1=1} z^T\theta = \max_{1\le j\le M} |z^{(j)}|, \quad \forall\, z \in E^*.$$

Let Θ be a convex, closed set in E. For a given parameter $\beta > 0$ and a convex function $V : \Theta \to \mathbb{R}$, we call β-*conjugate* function of V the Legendre–Fenchel type transform of βV:

$$\forall\, z \in E^*, \quad W_\beta(z) = \sup_{\theta \in \Theta} \left\{ -z^T\theta - \beta V(\theta) \right\} . \qquad (11)$$

Assumption (L). *A convex function $V : \Theta \to \mathbb{R}$ is such that its β-conjugate W_β is continuously differentiable on E^* and its gradient ∇W_β satisfies*

$$\|\nabla W_\beta(z) - \nabla W_\beta(\tilde{z})\|_1 \le \frac{1}{\alpha\beta}\|z - \tilde{z}\|_\infty, \quad \forall\, z, \tilde{z} \in E^*,\ \beta > 0,$$

where $\alpha > 0$ is a constant independent of β.

Assumption (L) relates to the strong convexity w.r.t. 1-norm:

$$V(sx+(1-s)y) \leq sV(x)+(1-s)V(y)-\frac{\alpha}{2}s(1-s)\|x-y\|_1^2 \tag{12}$$

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

The following proposition sums up some properties of β-conjugates and, in particular, yields a sufficient condition for Assumption (L).

Proposition 1. *Let function* $V : \Theta \to \mathbb{R}$ *be convex and* $\beta > 0$. *Then, the* β*-conjugate* W_β *of* V *has the following properties.*

1. *The function* $W_\beta : E^* \to \mathbb{R}$ *is convex and has a conjugate* βV, *i.e.,*

$$\forall\, \theta \in \Theta, \quad \beta V(\theta) = \sup_{z \in E^*} \left\{ -z^T\theta - W_\beta(z) \right\} .$$

2. *If function* V *is* α*-strongly convex with respect to the* 1*-norm then*

   (i) *Assumption (L) holds true,*

   (ii) $\operatorname*{argmax}_{\theta \in \Theta} \left\{ -z^T\theta - \beta V(\theta) \right\} = -\nabla W_\beta(z) \in \Theta$.

Definition 1. *We call* $V : \Theta \to \mathbb{R}_+$ proxy function *if it is convex, and*

(i) there exists a point $\theta_* \in \Theta$ *such that* $\min_{\theta \in \Theta} V(\theta) = V(\theta_*)$,

(ii) Assumption (L) holds true.

Example: Consider a standard simplex $\Theta = \Theta_M$ and an entropy-type proxy function

$$V(\theta) = \ln(M) + \sum_{j=1}^{M} \theta^{(j)} \ln \theta^{(j)} \tag{13}$$

(where $0 \ln 0 \triangleq 0$) which has a single minimizer $\theta_* = (1/M, \ldots, 1/M)^T$ with $V(\theta_*) = 0$. It is directly checked that this function is α-strongly convex w.r.t. the 1-norm, with the parameter $\alpha = 1$. This leads to a β-conjugate function $W_\beta(z)$ of (14) and to a Mirror Descent Algorithm:

Indeed, for $\beta > 0$, one get function

$$W_\beta(z) = \beta \ln \left(\frac{1}{M} \sum_{k=1}^{M} e^{-z^{(k)}/\beta} \right), \quad z \in \mathbb{R}^M, \tag{14}$$

with partial derivatives relating to a Gibbs distribution on the coordinates of vector $z = (z^{(1)}, \ldots, z^{(M)})^T$, with β being a "temperature" parameter:

$$-\frac{\partial W_\beta(z)}{\partial z^{(j)}} = e^{-z^{(j)}/\beta} \left(\sum_{k=1}^{M} e^{-z^{(k)}/\beta} \right)^{-1}, \ j = 1, \ldots, M. \tag{15}$$

$\square$

Convex Stochastic Optimization Problem

$$A(\theta) \triangleq \mathbb{E}\, Q(\theta, Z) \;\rightarrow \min_{\theta \in \Theta}$$

with loss function $Q : \Theta \times \mathcal{Z} \rightarrow \mathbb{R}_+$ being such that the random function $Q(\cdot\,, Z) : \Theta \rightarrow \mathbb{R}_+$ is convex a.s., on a convex closed set $\Theta \subset \mathbb{R}^M$.

Let a learning sample be given in the form of an i.i.d. sequence $(Z_1, \ldots, Z_{t-1})$, where each Z_i has the same distribution as Z.

Denote stochastic subgradients

$$u_i(\theta) = \nabla_\theta Q(\theta, Z_i)\,, \quad i = 1, 2, \ldots, \tag{16}$$

which are measurable functions on $\Theta \times \mathcal{Z}$ such that, for any $\theta \in \Theta$, the expectation $\mathbb{E}\, u_i(\theta)$ belongs to the subdifferential of the function $A(\theta)$.

Mirror Descent Algorithm (MDA)

The algorithm is defined as follows:

- Fix the initial value $\zeta_0 = 0 \in \mathbb{R}^M$.
- For $i = 1, \ldots, t-1$, do the recursive update

$$\begin{aligned} \zeta_i &= \zeta_{i-1} + \gamma_i u_i(\theta_{i-1}) \,, \\ \theta_i &= -\nabla W_{\beta_i}(\zeta_i) \,. \end{aligned} \tag{17}$$

- Output at iteration t the following convex combination:

$$\widehat{\theta}_t = \sum_{i=1}^{t} \gamma_i \theta_{i-1} \left(\sum_{i=1}^{t} \gamma_i \right)^{-1} \,. \tag{18}$$

A Particular Case of the Algorithm

Let

$$\gamma_i \equiv 1\,, \quad \beta_i = \beta_0\sqrt{i+1} \quad (i \geq 1), \quad \beta_0 > 0\,. \tag{19}$$

Then the algorithm becomes simpler and can be implemented in the following recursive form:

$$\zeta_i = \zeta_{i-1} + u_i(\theta_{i-1})\ , \tag{20}$$

$$\theta_i = -\nabla W_{\beta_i}(\zeta_i)\ , \tag{21}$$

$$\widehat{\theta}_i = \widehat{\theta}_{i-1} - \frac{1}{i}\left(\widehat{\theta}_{i-1} - \theta_{i-1}\right), \quad i = 1, 2, \ldots, \tag{22}$$

with initial value $\zeta_0 = 0\,.$

Theorem 0. *Assume that*

$$\sup_{\theta\in\Theta} \mathbb{E} \, \|\nabla_\theta \, Q(\theta, Z)\|_\infty^2 \leq L_{\Theta,Q}^2 \ , \tag{23}$$

where $L_{\Theta,Q} \in (0, +\infty)$. *Let* V *be a proxy function on* Θ *satisfying Assumption (L) with a parameter* $\alpha > 0$, *and assume that there exists* $\theta_A^* \in \operatorname*{Argmin}_{\theta\in\Theta} A(\theta)$. *Furthermore, let* $V(\theta_A^*) \leq \overline{V} < +\infty$, *and we set* $\beta_0 = L_{\Theta,Q} \, (\alpha \, \overline{V}\,)^{-1/2}$.

Then, with sequences $(\gamma_i)_{i\geq 1}$ and $(\beta_i)_{i\geq 1}$ from (19), for any integer $t \geq 1$, the estimate $\widehat{\theta}_t$ being defined in (17)–(18) with stochastic subgradients (16) satisfies inequality

$$\mathbb{E}\, A(\widehat{\theta}_t) - \min_{\theta\in\Theta} A(\theta) \leq 2\, L_{\Theta,Q} \left(\alpha^{-1}\overline{V}\right)^{1/2} \frac{\sqrt{t+1}}{t}\,. \qquad (24)$$

In particular, if Θ is a convex compact set, we can take $\overline{V} = \max_{\theta\in\Theta} V(\theta)$. □

Proof: see [6].

3 Main Eigenvalue Estimation to a Stochastic Matrix.

Let $A = \|a_{ij}\|_{N\times N}$ be a given left stochastic matrix, N be *a large number*. Denote $\Theta_N \subset \mathbb{R}^N$, the standard simplex.

Our goal: We are to approximate a positive solution to a linear system equations

$$Ax = x\,, \quad x \in \Theta_N\,. \tag{25}$$

Motivation: Calculations for PageRank problem.

Notations

- Let $\mathcal{A}_N$ be a set of all left stochastic $N \times N$-matrices A.
- $A^{(j)}$ and $A_{(i)}$ mean j-column and i-row in the matrix A.
- Given a matrix $A \in \mathcal{A}_N$, define set

$$X_* \triangleq \operatorname*{Arg\,min}_{x \in \Theta_N} \|Ax - x\|_2 = \{x \in \Theta_N \ : \ Ax = x\} \quad (26)$$

being convex compact of all solutions $x_* \in X_*$.

- Define risk functions

$$\mathcal{R}_A(x) \triangleq \frac{1}{2}\|Ax - x\|_2^2 , \quad x \in \mathbb{R}^N , \quad (27)$$

$$\mathcal{Q}_A(x) \triangleq \|Ax - x\|_2 , \quad x \in \mathbb{R}^N . \quad (28)$$

Related Optimization Problem

Minimize risk function $\mathcal{R}_A(x)$ on simplex Θ_N. This gives

$$\nabla_x \mathcal{R}_A(x) = A^T A x - A^T x - A x + x \,. \tag{29}$$

Stochastic Gradients by Randomization

On iteration $k \geq 1$, one can prove

$$\mathbb{E}\left(\zeta_k \mid x_1, \ldots, x_k\right) = \left.\nabla_x \mathcal{R}_A(x)\right|_{x=x_k} \,, \tag{30}$$

where x_t means the result of t^{th} iteration, $t = 1, \ldots, k$,

$$\zeta_k \triangleq \left(A_{(\xi_k)}\right)^T - \left(A_{(\eta_k)}\right)^T - A^{(\eta_k)} + x_k \,; \tag{31}$$

having the two random indexes $\xi_k\,, \eta_k \in \{1, \ldots, N\}$ with

$$\mathbb{P}\left(\eta_k = j \mid x_1, \ldots, x_k\right) = x_k^{(j)}\,, \quad j = 1, \ldots, N\,, \tag{32}$$

and

$$\mathbb{P}\left(\xi_k = i \mid x_1, \eta_1, \ldots, x_k, \eta_k\right) = a_{i\eta_k}\,, \quad i = 1, \ldots, N\,. \tag{33}$$

Important bounds hold

$$\begin{aligned} \|\zeta_k\|_\infty &\leq \|\left(A_{(\xi_k)}\right)^T - \left(A_{(\eta_k)}\right)^T\|_\infty + \|x_k - A^{(\eta_k)}\|_\infty \quad (34) \\ &\leq 2\,. \quad (35) \end{aligned}$$

Optimization MD Algorithms

- Fix $x_0 \in \Theta_N$ and $\psi_0 = 0 \in \mathbb{R}^N$. Fix positive $(\gamma_k)_{k\geq 1}$, $(\beta_k)_{k\geq 1}$, and horizon $n > 1$.
- For $k = 0, \ldots, n-1$ generate η_k and ξ_k by (32) and (33); then calculate stochastic gradient ζ_k (31), and iterate

$$\begin{aligned} \psi_k &= \psi_{k-1} + \gamma_k \zeta_k \,, \\ x_k &= -\nabla W_{\beta_k}(\psi_k) \,. \end{aligned} \tag{36}$$

- Output n^{th} iteration of convex combination

$$\widehat{x}_n = \frac{\sum_{k=1}^n \gamma_k x_{k-1}}{\sum_{k=1}^n \gamma_k} \,. \tag{37}$$

▲

Here function $W_\beta(z)$ and its gradient $\nabla W_\beta(\cdot)$ being Gibbs potential are as follows: $\forall\, z \in \mathbb{R}^N$,

$$W_\beta(z) = \beta \ln \left(\frac{1}{N} \sum_{k=1}^{N} e^{-z^{(k)}/\beta} \right), \tag{38}$$

$$\frac{\partial W_\beta(z)}{\partial z^{(j)}} = -e^{-z^{(j)}/\beta} \left(\sum_{k=1}^{N} e^{-z^{(k)}/\beta} \right)^{-1}, \; j = 1, \ldots, N \tag{39}$$

Remark: Another conjugate function $W_\beta(z) = \frac{\beta}{2}\|z\|_2^2$ would give Euclidean potential $\nabla W_\beta(z) = \beta z$ which leads to an ordinary stochastic gradient algorithm (projected SA) with time averaging. Cf Polyak–Juditsky SA with averaging.

Main Results I: Uniform Upper Bounds

Theorem 1. *Let $N \geq 2$, and let estimation $\widehat{x}_n$ be defined by randomized algorithm (36)–(39) with stochastic gradient ζ_k (31) and the parameters*

$$\gamma_k \equiv 1\,, \quad \beta_k = \beta_0\sqrt{k+1}\,, \quad \beta_0 = 2(\ln N)^{-1/2}. \tag{40}$$

Then, under arbitrary iteration number $n \geq 1$, one holds

$$\mathbb{E}\,\|A\widehat{x}_n - \widehat{x}_n\|_2^2 \leq 8\,(\ln N)^{1/2}\,\frac{\sqrt{n+1}}{n}\,. \tag{41}$$

Remark: The projected SA (i.e., the MDA with Euclidean potential) would give Upper Bound like $O(N/n)$, instead of $O(\sqrt{\ln N/n})$ (41). For instance, condition $\sqrt{(\ln N)/n} \geq N/n$ implies $n \geq N^2/(\ln N)$.

4 Multi-Armed Bandit Problem.

Presented at the 17th IFAC World Congress:

1. Juditsky, A., A.V. Nazin, A.B. Tsybakov, N. Vayatis. Gap-free Bounds for Stochastic Multi-Armed Bandit. *Proc. 17th IFAC World Congress, Seoul, Korea, 6–11 July 2008, pp.11560–11563*.

Let $X = \{x(1), \ldots, x(N)\}$ be a set of N available actions. At each time $t = 1, 2, \ldots$, we have to choose sequentially an action $x_t \in X$. We denote by η_t the observable (instantaneous) loss for the choice of x_t, and introduce the average loss up to horizon T which is to be minimized:

$$\Phi_T = \frac{1}{T} \sum_{t=1}^{T} \eta_t \,. \tag{42}$$

A strategy $\mathcal{U}$ is a sequence of rules for the choice x_t at times $t = 1, \ldots, T$. In the stochastic setup that we consider here, the sequence of losses $(\eta_t)_{t \geq 1}$ is a stochastic process and x_t is a measurable function (random, in general) depending only on the vector of past decisions and losses $(x_1, \ldots, x_{t-1}; \eta_1, \ldots, \eta_{t-1})$.

Any strategy $\mathcal{U}$ generates a flow of σ-algebras $\mathcal{F}_t = \sigma\{x_1, \ldots, x_t; \eta_1, \ldots, \eta_t\}$, $t \geq 1$ (for brevity we do not indicate the dependence of $\mathcal{F}_t$ on $\mathcal{U}$). Throughout the paper we denote by $z^{(j)}$ the jth component of vector $z \in \mathbb{R}^N$.

Two basic assumptions:

A1. With probability 1, the conditional expectations satisfy

$$\mathbb{E}\{\eta_t \,|\, \mathcal{F}_{t-1}\,,\, x_t = x(k)\} = a_k,\;\; k = 1, \ldots, N, \qquad (43)$$

where $a_k \in \mathbb{R}$ are unknown deterministic values.

The value a_k characterizes the expected loss for deciding to take the action $x_t = x(k)$ at time t. Assumption A1 says that this loss should not depend on t.

A2. The second conditional moment of the loss η_t is a.s. bounded by a constant:

$$\mathbb{E}\{\eta_t^2 \,|\, \mathcal{F}_{t-1}\,,\, x_t\} \leq \sigma^2 < \infty\,. \qquad (44)$$

It is easy to prove (see, e.g., [12]) that under these assumptions all the limiting points of the average loss sequence $(\Phi_t)_{t\geq 1}$ cannot be almost surely (a.s.) less than

$$a_{\min} \triangleq \min_{k=1,\dots,N} a_k \, .$$

Thus, the problem is to design a strategy $\mathcal{U}^*$ which has the asymptotically minimal average loss:

$$\Phi_T \to a_{\min} \quad \text{as} \quad T \to \infty \, , \tag{45}$$

in an appropriate probability sense.

We study here *convergence in mean*, trying to get the rate of convergence

$$\mathbb{E}(\Phi_T) \longrightarrow a_{\min}$$

as fast as possible.
In particular, we provide *non-asymptotic* upper bounds for the expected excess risk $\mathbb{E}(\Phi_T) - a_{\min}$ that are close, up to logarithmic factors, to the lower bound of the order $\sqrt{N/T}$ proved for arbitrary N by (see Theorem 6.11 in [14]).

We will suppose that the following assumption on the loss sequence $(\eta_t)_{t\geq 1}$ holds:

A3. The losses are nonnegative: $\eta_t \geq 0$ a.s.

Below we propose a randomized decision strategy in which, at each step $t+1$, the action x_{t+1} is drawn according to a distribution $p_t \triangleq \left(p_t^{(1)}, \ldots, p_t^{(N)}\right)^\top$ over X where:

$$p_t^{(k)} \triangleq \mathbb{P}(x_{t+1} = x(k) \,|\, \mathcal{F}_t)\,, \quad k = 1, \ldots, N\,. \qquad (46)$$

The update of the distribution p_t over time is given by the MDA.

Denote by Θ the simplex of all probability vectors over X:

$$\Theta \triangleq \left\{ p \in \mathbb{R}_+^N \;\middle|\; \sum\nolimits_{k=1}^{N} p^{(k)} = 1 \right\} . \qquad (47)$$

We then define the mean (over the set of actions) loss function A on Θ:

$$A(p) = \sum_{k=1}^{N} a_k p^{(k)} = a^\top p \,, \quad p \in \Theta \;, \qquad (48)$$

where $a = (a_1, \ldots, a_N)^\top$. Since p_t is a random vector, the quantity $A(p_t)$ is a random variable. The update rule for the probability distribution p_t uses a stochastic gradient of A.

The expected average loss equals to the average over time of the expectations $\mathbb{E}A(p_t)$, that is

$$\mathbb{E}(\Phi_T) = \frac{1}{T}\sum_{t=1}^{T}\mathbb{E}(\mathbb{E}(\eta_t \,|\, x_t\,, \mathcal{F}_{t-1})) = \frac{1}{T}\sum_{t=1}^{T}\mathbb{E}(A(p_{t-1}))\ . \tag{49}$$

Theorem. *Let assumptions A1-A2-A3 be satisfied and let the conditional distributions $(p_t)_{t\geq 0}$ be defined by the MDA. Then, for any horizon $T \geq 1$,*

$$\mathbb{E}\left(\Phi_T\right) - a_{\min} \leq 2\,\sigma\,\frac{\sqrt{(T+1)N\ln N}}{T}\,. \tag{50}$$

Remark: The known information lower bound (see [14], Theorem 6.11) differs from the upper bound (50) by logarithmic term $\sqrt{\ln N}$.

5 Conclusions.

See another problems, for instance:

- Robust PageRank [Polyak, Juditsky (2011)]:

$$f(x) \triangleq \|Ax\|_2 + \varepsilon\|x\|_2 \to \min_{x \in \Theta_N},$$

  where $\varepsilon > 0$ is given, Θ_N stands for standard simplex in $\mathbb{R}^N$, $A = P - I$, P is a given stochastic $N \times N$-matrix, I is the identical matrix.

- Classification (pattern recognition) [6]

- Control of finite Markov chains [16], [17].

References

[1] Nemirovskii, A.S. and Yudin, D.B., *Slozhnost' zadach i effektivnost' metodov optimizatsii*, Moscow: Nauka, 1979 (in Russian). Translated under the title *Problem Complexity and Method Efficiency in Optimization*, Chichester: Wiley, 1983.

[2] A. Ben-Tal, T. Margalit, A. Nemirovski. The ordered subsets mirror descent optimization method with applications to tomography. *SIOPT* 12(1), 79–108, 2001.

[3] A. Beck, M. Teboulle. Mirror descent and nonlinear projected subgradient methods for convex optimization. *Oper. Res. Lett.* 31(3), 167–175, 2003.

[4] Yu. Nesterov. Primal-dual subgradient methods for convex problems: Core discussion paper 2005/67. Louvain-la-Neuve, Belgium: Center for Operation Research and Econometrics, 2005.

[5] Yu. Nesterov. Primal-dual subgradient methods for convex problems // *Mathematical Programming*, 2007. DOI: 10.1007/s10107-007-0149-x.

[6] A.B. Juditsky, A.V. Nazin, A.B. Tsybakov, and N. Vayatis. Recursive aggregation of estimators by the mirror descent algorithm with averaging. *Problems of Information Transmission*, 41(4):368–384, 2005.

[7] A. Juditsky, G. Lan, A. Nemirovski, and A. Shapiro. Stochastic Approximation approach to Stochastic Programming. *Optimization Online*, 10/03/2007 http://www.optimization-online.org/DB_HTML/2007/09/1787.html

[8] A.V. Nazin. Solution to a Particular Deterministic Problem in Finite High Dimension by the Recursive Randomized Mirror Descent Algorithm. *8th Meeting on Mathematical Statistics*, CIRM, Luminy, France, 2008.

[9] A.V. Nazin and B.T. Polyak. The Randomized Algorithm for Finding an Eigenvector of the Stochastic Matrix with Application to PageRank. *Doklady Mathematics*, 2009, Vol. 79, No. 3, pp. 424–427.

[10] A.V. Nazin and B.T. Polyak. Adaptive Randomized Algorithm for Finding Eigenvector of Stochastic Matrix with Application to PageRank. *48th IEEE Conf. on Decision and Control*, Shanghai, China, December 2009. (Regular Paper.)

[11] A.V. Nazin. Estimating the Principal Eigenvector of a Stochastic Matrix: Mirror Descent Algorithms via Game Approach. *49th IEEE Conf. on Decision and Control*, December 2010. (Submitted as a Regular Paper.)

[12] A.V. Nazin and A.S. Poznyak. *Adaptive Choice of Variants*. Nauka, Moscow, 1986, (in Russian).

[13] K. Najim and A.S. Poznyak. *Learning automata: theory and applications*. Pergamon Press, Inc., Elmsford, NY, USA, 1994. ISBN 0-08-042024-9.

[14] N. Cesa-Bianchi and G. Lugosi. *Prediction, Learning, and Games*. Cambridge University Press, 2006.

[15] A.V. Nazin, B.T. Polyak.
A randomized algorithm for finding the eigenvector of a stochastic matrix with application to PageRank problem. *Automation and Remote Control*, 2011, vol. 72, no. 2, pp. 342–352.)

[16] Nazin, A.V., B.M. Miller. Mirror Descent Algorithm for Homogeneous Finite Controlled Markov Chains with Unknown Mean Losses. *Proc. 18th IFAC World Congress, Milano, Italy, August 28 – September 2, 2011*.

[17] Nazin, A.V., B.M. Miller. The Mirror Descent Control Algorithm for Weakly Regular Homogeneous Finite Markov Chains with Unknown Mean Losses. *The 50th IEEE Conference on Decision and Control and European Control Conference, December 12–15, 2011, Orlando, Florida USA.*

THANK YOU FOR YOUR ATTENTION !!!