On the relationship between simulation logit dynamics in the population game theory and mirror descent method in the online optimization using the example of the shortest path problem

A. Lagunovskaya [1,2]

[1](Keldysh Institute of Applied Mathematics of RAS, Russia) [2]MIPT, Russia

"3 oracles",
IITP RAS
December, 23 2015

Outline

Goal and tasks

Goal of the research

Demonstrate the link between online optimization and game theory (the behavior of the users of the transport network)

Tasks

- Briefly describe several versions of the mirror descent method
- Formulate the problem that user of the transport network is to solve as the optimization problem
- Explain the behavior dynamics of the transport network users

General problem

The sequence $\{x^k\} \in Q$ is required to be chosen so that the pseudo regret is minimum:

$$\mathrm{Regret}_N\left(\{f_k(\cdot)\},\left\{x^k\right\}\right)=\frac{1}{N}\sum_{k=1}^{N} f_k\left(x^k\right)-\min_{x\in Q}\frac{1}{N}\sum_{k=1}^{N} f_k(x) \qquad (1)$$

based on the information known

$$\left\{\nabla_x f_1\left(x^1,\xi^1\right);\ldots;\nabla_x f_{k-1}\left(x^{k-1},\xi^{k-1}\right)\right\}$$

calculating the x^k.

General problem

The following conditions are to be satisfied:
$\max\limits_{x,y\in Q} \|x-y\| \leq \tilde{R}$.

1) for every $k = 1, ..., N$ (Ξ^{k-1} – sigma-algebra generated by $\xi^1, \ldots,$ ξ^{k-1})

$$E_{\xi_k}\left[\nabla_x f_k\left(x^k,\xi^k\right)\right] = \nabla f_k\left(x^k\right).$$

Regarding the class of functions from which we choose $\{f_k(\cdot)\}$, the following conditions are assumed *условия*:

2) $\{f_k(\cdot)\}$ $\quad -$ convex functions;
3) for every $k = 1, ..., N$, $x \in Q$

$$\|\nabla_x f_k(x,\xi)\|_*^2 \leq M^2.$$

Mirror descent method

$d(x) : d\left(x^1\right) = \nabla d\left(x^1\right) = 0.$
$x^1 = \arg\min_{x \in Q} d(x)$

$$V_x(y) = d(y) - d(x) - \langle \nabla d(x), y - x \rangle.$$

we believe that $d(x) = V_{x^1}(x) \leq R^2$ for all $x \in Q$.

$$\mathrm{Mirr}_{x^k}(g) = \arg\min_{y \in Q} \left\{ \left\langle g, y - x^k \right\rangle + V_{x^k}(y) \right\}.$$

Mirror descent method (MDM) for the pronlem described above will be:

$$x^{k+1} = \mathrm{Mirr}_{x^k}\left(\alpha \nabla_x f_k\left(x^k, \xi^k\right)\right), \quad k = 1, ..., N.$$

Then if the condition (2) is satisfied, for every $u \in Q$, $k = 0, ..., N-1$ the inequality holds:

$$\alpha \left\langle \nabla_x f_k\left(x^k, \xi^k\right), x_k - u \right\rangle \leq \frac{\alpha^2}{2} \left\| \nabla_x f_k\left(x^k, \xi^k\right) \right\|_*^2 + V_{x^k}(u) - V_{x^{k+1}}(u).$$

Example 1 (simplex)

$$Q = S_n(1) = \left\{ x \geq 0 : \quad \sum_{i=1}^{n} x_i = 1 \right\}.$$

Choose $\|\cdot\| = \|\cdot\|_1$, $\quad d(x) = \ln n + \sum\limits_{i=1}^{n} x_i \ln x_i$.

Then MDM will be: ($\alpha = M^{-1}\sqrt{2 \ln n / N}$): $x_i^1 = 1/n, \quad i = 1, ..., n$, $k = 1, ..., N$, $i = 1, ..., n$

$$x_i^{k+1} = \frac{\exp\left(-\sum\limits_{r=1}^{k} \alpha \frac{\partial f_r(x^r,\xi^r)}{\partial x_i}\right)}{\sum\limits_{l=1}^{n} \exp\left(-\sum\limits_{r=1}^{k} \alpha \frac{\partial f_r(x^r,\xi^r)}{\partial x_l}\right)} = \frac{x_i^k \exp\left(-\alpha \frac{\partial f_k\left(x^k,\xi^k\right)}{\partial x_i}\right)}{\sum\limits_{l=1}^{n} x_l^k \exp\left(-\alpha \frac{\partial f_k\left(x^k,\xi^k\right)}{\partial x_l}\right)}.$$

Estimation of the pseudo regret:

$$\mathrm{Regret}_N\left(\{f_k(\cdot)\}, \left\{x^k\right\}\right) \leq M\sqrt{\frac{2 \ln n}{N}}$$

Example 2 (the direct product of simplexes)

$$x = \left(z^1, \dots, z^m\right) \in Q = \prod_{j=1}^{m} S_{n_j}\left(d_j\right)$$

$$\|x\| = \sqrt{\sum_{j=1}^{m} \|z^j\|_1^2}, d\left(x\right) = \sum_{j=1}^{m} d_j d^j\left(z^j\right), d^j\left(z^j\right) = d_j \ln n_j + \sum_{i=1}^{n_j} z_i^j \ln\left(\frac{z_i^j}{d_j}\right).$$

Denote: $\alpha_j = M^{-1}\sqrt{2\ln n_j/N}$, MDM: $z_i^j = d_j/n_j, \quad i = 1, \dots, n_j$, if $k = 1, \dots, N,\ i = 1, \dots, n_j,\ j = 1, \dots, m$

$$z_i^{j,k+1} = d_j \frac{\exp\left(-\sum\limits_{r=1}^{k} \alpha_j \frac{\partial f_r(x^r)}{\partial z_i^j}\right)}{\sum\limits_{l=1}^{n} \exp\left(-\sum\limits_{r=1}^{k} \alpha_j \frac{\partial f_r(x^r)}{\partial z_l^j}\right)} = d_j \frac{x_i^k \exp\left(-\alpha_j \frac{\partial f_k\left(x^k\right)}{\partial z_i^j}\right)}{\sum\limits_{l=1}^{n} x_l^k \exp\left(-\alpha_j \frac{\partial f_k\left(x^k\right)}{\partial z_l^j}\right)}.$$

Estimation of the pseudo regret:

$$\mathrm{Regret}_N\left(\{f_k\left(\cdot\right)\}, \{x^k\}\right) \le \frac{M}{\sqrt{N}} \frac{\max\limits_{j=1,\dots,m}\{\ln n_j\}}{\sqrt{2\min\limits_{j=1,\dots,m}\{\ln n_j\}}} \left(\sum_{j=1}^{m} d_j^2 + 1\right)$$

Example 3(the choice between the vertices of the simplex

Additional conditions:

4) $f_k(x) = \langle l^k, x \rangle, \quad k = 1, ..., N;$

5) At each step the generation of the random value x^k according to the probability distribution p^k is performed independently. The choice of f_k is carried out without knowledge of x^k.

Given $p_i^1 = x_i^1 = 1/n$, $i = 1, ..., n$. If $k = 1, ..., N$, $i = 1, ..., n$ according to the probability distribution $(\alpha = M^{-1}\sqrt{2\ln n/N})$

$$p_i^{k+1} = \frac{\exp\left(-\sum\limits_{r=1}^{k} \alpha \frac{\partial f_r(x^r,\xi^r)}{\partial x_i}\right)}{\sum\limits_{l=1}^{n} \exp\left(-\sum\limits_{r=1}^{k} \alpha \frac{\partial f_r(x^r,\xi^r)}{\partial x_l}\right)} = \frac{p_i^k \exp\left(-\alpha \frac{\partial f_k(x^k,\xi^k)}{\partial x_i}\right)}{\sum\limits_{l=1}^{n} p_l^k \exp\left(-\alpha \frac{\partial f_k(x^k,\xi^k)}{\partial x_l}\right)},$$

generate random variable $i(k+1)$, and assume
$x_{i(k+1)}^{k+1} = 1, \;\; x_j^{k+1} = 0, \; j \neq i(k+1).$
Estimation of the pseudo regret:
$\text{Regret}_N\left(\{f_k(\cdot)\}, \{x^k\}\right) \leq M\sqrt{\frac{2}{N}}\left(\sqrt{\ln n} + 6\sqrt{\ln(\sigma^{-1})}\right)$ with $P \geq 1 - \sigma$

The Shortest path problem: problem set

$\langle V, E\rangle$ - the graph of the transportation network
$OD \subseteq V \otimes V$ $(|OD| = m)$ – a set of the pairs of source-drain;
d_w – correspondence, corresponding to the pair w;
x_p – a flow along the path p;
P_w – a set of the paths according to the correspondence w, $P = \bigcup\limits_{w\in OD} P_w$;
L – the maximum number of edges in the path from P.
The costs of passing the edge $e \in E$ are desctibed by the function $0 \le \tau_e(f_e) \le \tilde{M}$, where f_e – a flow along the edge e:

$$f_e(x) = \sum_{p\in P} \delta_{ep} x_p, \quad \delta_{ep} = \begin{cases} 1, & e \in p \\ 0, & e \notin p \end{cases}.$$

Given $M = \tilde{M}L$.
$G_p(x) = \sum\limits_{e\in E} \tau_e(f_e(x))\,\delta_{ep}$ – the costs on the path p.

The Shortest path problem: solution scheme

Similarly to the example 2 introduce:

$$X = \left\{ x \geq 0 : \sum_{p \in P_w} x_p = d_w, \; w \in OD \right\},$$

$$\Psi(x) = \sum_{e \in E} \int\limits_0^{f_e(x)} \tau_e(z)\, dz : \nabla \Psi(x) = G(x).$$

Given $d_w := d_w \cdot \bar{N}, \quad \bar{N} \gg 1, \quad w \in OD$, и $\tau_e(f_e) := \tau_e\left(f_e / \bar{N}\right)$.

The Shortest path problem: user behaviour

- Each individual user of the transport network solve independently its own problem as in the example 3.
- The user know at the step $k+1$ the exact information regarding the "cost history" $\left\{ l_p^r = \left\{ G_p \left(x^r \right) \right\}_{p \in P_w} \right\}_{r=1}^{k}$.
- We assume that $0 \leq \left\{ l_p^k \right\} \leq M$ can be chosen adversely to the user's intention.
- User tends to act optimally (with $i = p$, $n = |P_w|$).
- The number of the users is assumed to be $\bar{N} \gg 1$. All users behave as described above.
- So if we consider the entire transport system the estimation of the pseudo regret will be similar to the ones from the example 2.

The Shortest path problem: solution scheme

Thus, the optimization problem: $\Psi(x) \to \min\limits_{x \in X}$,

$$f_k(x) \equiv \Psi(x), \quad \bar{x}^N = \frac{1}{N}\sum_{k=1}^{N} x^k, \quad \Psi_* = \Psi(x_*),$$

$j = w$, $z^j = \{x_p\}_{p \in P_w}$, $n_j = |P_w|$,

$$\alpha_j = M^{-1}\sqrt{2\ln n_j/N}, \quad N\mathrm{Regret}_N \le \max_{j=1,\dots,m} \alpha_j^{-1}\left(R^2 + \frac{1}{2}M^2 N \max_{j=1,\dots,m} \alpha_j^2\right)$$

Pseudo regret estimation:

$$\Psi\left(\bar{x}^N\right) - \Psi_* \le \mathrm{Regret}_N \le \frac{M}{\sqrt{N}} \frac{\max\limits_{j=1,\dots,m}\{\ln n_j\}}{\sqrt{2\min\limits_{j=1,\dots,m}\{\ln n_j\}}}\left(\sum_{j=1}^{m} d_j^2 + 1\right)$$

Conclusion

The solution of the oprimization problem $\Psi(x) \to \min\limits_{x \in X}$ not only provides us with the equilibrium flow-distribution along the paths.
The individual stochastic dynamics (described in the example 3) arising from the solution the individual online optimization problem by each user of the transport network leads to the deterministic dynamics of the entire transport system when number of users tends increases with the estimations of the pseudo regret similar to the ones from example 2.

References

1 Lugosi G., Cesa-Bianchi N. Prediction, learning and games. New York: Cambridge University Press, 2006.

2 Patriksson M. The traffic assignment problem. Models and methods. Utrecht, Netherlands: VSP, 1994.

3 Sandholm W.H. Population games and Evolutionary dynamics. Economic Learning and Social Evolution. MIT Press; Cambridge, 2010.

4 Bubeck S. Introduction to online optimization. Princeton University: Lecture Notes, 2011.

5 Gasnikov A.V., Nesterov Yu.E., Spokoiny V.G. On the Efficiency of a Randomized Mirror Descent Algorithm in Online Optimization Problems // Computational Mathematics and Mathematical Physics, vol. 55, no. 4, 2015, pp. 580–596.

6 Nesterov Y.Primal-dual subgradient methods for convex problems // Math. Program. Ser. B. 2009. V. 120(1). P. 261–283.