ON NO-REGRET ALGORITHMS

Sylvain Sorin
sylvain.sorin@imj-prg.fr

Institut de Mathématiques Jussieu-PRG
Sorbonne Université, Campus P&M Curie - Paris 6
CNRS UMR 7586

International Conference on
Challenges and Advances in Modern Variational Analysis,
dedicated to Professor Jean-Paul Penot's 80th birthday
March 4-15, XLIM Laboratory, Université de Limoges

Abstract

We study the implications of the « no-regret property » on algorithms, in continuous and discrete time, for on-line learning, game theory dynamics and convex optimization.
The analysis cover first order dynamics like projected gradient, mirror descent and dual averaging.

1.1. Model

V normed vector space, finite dimensional,
dual V^* and duality map $\langle .|.\rangle$,
$X \subset V$, compact and convex.

We study algorithms that associate to:
- a trajectory of parameters $\{u_t \in V^*, t \geq 0\}$ in the dual space,
- a process of actions $\{x_t \in X, t \geq 0\}$ in the primal space,
where the actions are “adapted” to the parameters in the sense that x_t depends on $\{(x_s, u_s), 0 \leq s < t\}$.

The evaluation of the adequation of the action process to the parameter trajectory relies on a the family of functions “cumulative regret up to time t facing y”:

$$R_t(y) = \int_0^t \langle u_s | y - x_s \rangle ds, \qquad t \geq 0, \quad y \in X \tag{1}$$

or in discrete time, $\{x_m\}$ depending on $\{x_1, u_1, \ldots, x_{m-1}, u_{m-1}\}$:

$$R_n(y) = \sum_{m=1}^{n} \langle u_m | y - x_m \rangle, \qquad y \in X. \tag{2}$$

The procedure satisfies the no-regret property if:

$$R_t(y) \le o(t), \qquad \forall y \in X, \tag{3}$$

or

$$R_n(y) \le o(n), \qquad \forall y \in X. \tag{4}$$

This means that the time average regret is asymptotically less than 0.

A) We compare the performance of the algorithms at three levels:
(I) general case: $\{u_t\}$ is a bounded measurable process in V^*,
(II) closed form: $u_t = \phi(x_t)$ for a continuous vector field $\phi : X \to V^*$,
(III) convex gradient: $u_t = -\nabla f(x_t)$, for a $\mathscr{C}^1$ convex function $f : X \to I\!R$.
(with similar properties in discrete time).

B) We consider three different procedures:
a) Projected dynamics (PD),
b) Mirror descent (MD),
c) Dual averaging (DA).

C) We analyze the relations between the continuous and discrete time framework.

D) We also study the convergence of the trajectories of $\{x_t\}$ or $\{x_n\}$ (in classes (II) and (III)).

1.2. The three levels

Level (I) corresponds to the usual model of on-line learning where the agent observes $\{u_s, s < t\}$ and chooses x_t.

Note that since no hypothesis is made on the process u_t, no prediction makes sense, but the no-regret condition expresses a desirable a-posteriori property.

The notion of regret appears in Hannan, 1957 [32], Blackwell,1954 [12] 1956 [13], in a game theoretical set-up.

Algorithms and properties are studied in this spirit in Foster and Vohra, 1993 [26], Fudenberg and Levine, 1995 [29], Foster and Vohra, 1999 [27], Hart and Mas-Colell, 2000 [33], Lehrer, 2003 [49] , Benaim, Hofbauer and Sorin, 2005 [11], Cesa-Bianchi and Lugosi, 2006 [20], Viossat and Zapechelnyuk, 2013 [91], ... among others.

This topic is analyzed a.o. in the following books:

Cesa-Bianchi and Lugosi (2006) *Prediction, Learning and Games*, Cambridge University Press.
Hart and Mas-Colell (2013) *Simple Adaptive Strategies: From Regret-Matching to Uncoupled Dynamics*, World Scientific Publishing.

and the connection with related notions of approachability and calibrating is well presented in the survey:

Perchet (2014) *Approachability, regret and calibration: implications and equivalences*.

Similar tools and properties are present in statistics and in the learning community:

Vvok, 1990 [92], Cover, 1991 [23], Littlestone and Warmuth, 1994 [51], Freund and Shapire, 1999 [28], Auer, Cesa-Bianchi, Freund and Shapire, 2002 [6], Cesa-Bianchi and Lugosi, 2003 [19], Stoltz and Lugosi, 2005 [85], Kalai and Vempala, 2005 [43], Blum and Mansour, 2007 [14], ...

Note that this literature deals also with the random approach while we consider here the deterministic (conditional expectation) case.

The next two levels (II) and (III), describe more specific cases where the parameter u_t is a function of the action x_t.

Level (II), *closed form*, is relevant for game dynamics and variational inequalities.

Consider a strategic game $\Gamma(\phi)$ with a finite set of players I, where the equilibrium set E is given by the solutions $x \in X$ of the following variational inequalities:

$$\langle \phi^i(x) | x^i - y^i \rangle \geq 0, \qquad \forall y^i \in X^i, \forall i \in I.$$

Here $X^i \subset V^i$ is the strategy set of player $i \in I$, $X = \prod_i X^i$, and $\phi^i : X \to V^{i*}$ is her evaluation function.

Examples include:
- finite games (with mixed extension): ϕ^i is the vector payoff function VG^i.
- continuous games with payoff function G^i, $\mathscr{C}^1$ and concave w.r.t. x^i, $\forall i \in I$
then ϕ^i is the gradient of G^i w.r.t. x^i.
- population games (Wardrop equilibria), X^i is the simplex $\Delta(S^i)$ and ϕ^i corresponds to the outcome function $F^i : S^i \times X \longrightarrow I\!R$.

For each player i, the reference process is $u_t^i = \phi^i(x_t)$ which, as a function of x_t, is determined by the behavior of all players.
Hence the overall global dynamics of $\{x_t\}$ is generated by a family of unilateral procedures since for each i, x_t^i depends on (u^i, x^i) only.
In particular for each player i, the knowledge of $\phi^j, j \neq i$, is not assumed.

Thus for each player individually the situation is like *general case* (I), while the private parameters of the players are linked via x_t.

We will analyze the consequences on the process $\{x_t\}$ of the assumption that each player uses a procedure satisfying the no-regret condition.

Obviously the (global) algorithm associated to the (global) parameter function $\phi = \{\phi^i\}$ will also share the no-regret property since:

$$\int_0^t \langle \phi^i(x_s) | y^i - x_s^i \rangle ds \leq o(t), \qquad \forall y^i \in X^i, \qquad \forall i \in I,$$

implies:

$$\int_0^t \langle \phi(x_s) | y - x_s \rangle ds \leq o(t), \qquad \forall y \in X.$$

But in addition it is decentralized in the sense that $\{x^i\}$ depends upon ϕ^i only.

Level (III) covers the case of convex optimization where the parameter, after the choice x_t, is the gradient of the (unknown) convex function and $u_t = -\nabla f(x_t)$.

The research in this area is extremely active;
it links basic optimization algorithms, Polyak, 1987 [68], Nemirovski and Yudin, 1983 [60], Nesterov, 2004 [62], to on-line procedures, see e.g. Zinkevich, 2003 [97].

Recent books and lecture notes include a.o.:

Bubeck S. (2015) Convex optimization: Algorithms and complexity, *Fondations and Trends in Machine Learning*, **8**, 231-357.
Hazan E. (2015) Introduction to Online Convex Optimization, *Fondations and Trends in Optimization*, **2**, 157-325.
Hazan E. (2019) Optimization for Machine Learning , https://arxiv.org/pdf/1909.03550.pdf.
Shalev-Shwartz S. (2012) Online Learning and Online Convex Optimization, *Foundations and Trends in Machine Learning*, **4**, 107-194.

Related algorithms have also been developped in Operations Research (transportation, networks), see e.g. Dupuis and Nagurney, 1993 [24], Nagurney and Zhang, 1996 [59], Smith, 1984 [77].

Note that each community (learning, game theory, optimization) has its own terminology and point of view.

One of the aims of the current work is to clarify the relations between several approaches and results and to unify the analysis.

2.1. Definitions and notations

We describe here some relations with variational inequalities when the parameter process has a *closed form*: $u = \phi(x)$.

Notation 2.1

$NE(\phi)$ is the set of (internal) solutions, in X, of the variational inequality:

$$\langle \phi(x) | y - x \rangle \leq 0, \qquad \forall y \in X. \tag{5}$$

a) If ϕ is the evaluation function in a game $\Gamma(\phi)$, $NE(\phi)$ corresponds to the set of equilibria.

b) The minimization of a $\mathscr{C}^1$ convex function f on X corresponds to the variational inequality (5) with $\phi = -\nabla f$.
This case presents two properties:
ϕ is dissipative ($-\phi$ is monotone),
ϕ is a gradient.
The general definitions are as follows.

Definition 2.1

$\phi : X \to V^*$ *is* *dissipative* *if it satisfies:*

$$\langle \phi(x) - \phi(y) | x - y \rangle \leq 0, \qquad \forall x, y \in X. \tag{6}$$

Notation 2.2

$SE(\phi)$ *is the set of (external) solutions, in* X*, of the variational inequality:*

$$\langle \phi(y) | y - x \rangle \leq 0, \qquad \forall y \in X. \tag{7}$$

Notice that $SE(\phi)$ is convex.

Recall, see Minty, 1967 [56], that if ϕ is dissipative, then :

$$NE(\phi) \subset SE(\phi) \neq \emptyset$$

and if ϕ is continuous the reverse inclusion is satisfied:

$$SE(\phi) \subset NE(\phi) \neq \emptyset.$$

If $NE(\phi) = SE(\phi)$ we will also use the notation $E(\phi) = E$ for this set.

Fundamental example: 0-sum game

If $F : X = X^1 \times X^2 \to I\!R$ is $\mathscr{C}^1$ and concave/convex, the vector field $\phi = (\nabla^1 F, -\nabla^2 F)$ is dissipative, Rockafellar (1970) [70].
The elements of $NE(\phi) = SE(\phi) = E$ are optimal strategies of the associated 0-sum game.

We define a potential for a vector field, see e.g. Sorin and Wan (2016) [84].

Definition 2.2

A real function W of class $\mathscr{C}^1$ on X, is a potential for ϕ if there exist strictly positive functions μ^i on X, $i \in I$, such that:

$$\left\langle \nabla^i W(x) - \mu^i(x)\phi^i(x), y^i - x^i \right\rangle = 0, \quad \forall x \in X, \forall y^i \in X^i, \forall i \in I. \quad (8)$$

where $\nabla^i W$ denotes the gradient w.r.t. x^i.

The following result is classical, see e.g. Sandholm (2001) [72].

Proposition 2.1

Let ϕ be a vector field with potential W.
1. Every local maximum of W belongs to $NE(\phi)$.
2. If W is concave on X, then any element in $NE(\phi)$ is a global maximum of W on X.

2.2. Results

Assume that the procedure satisfies the no-regret property:

$$R_t(y) \leq o(t), \qquad \forall y \in X,$$

where:

$$R_t(y) = \int_0^t \langle \phi(x_s) | y - x_s \rangle ds, \qquad t \geq 0, y \in X$$

A first property deals with convergent trajectories $\{x_t\}$.

Lemma 2.1

If ϕ is continuous and $x_s \to x$, then $x \in NE(\phi)$.

Proof:

Since $R_t(y) = \int_0^t \langle \phi(x_s) | y - x_s \rangle ds$:

$$\frac{R_t(y)}{t} \to \langle \phi(x) | y - x \rangle, \qquad \forall y \in X. \tag{9}$$

and $R_t(y) \leq o(t)$ implies $x \in NE(\phi)$. ■

In particular, if x is a stationary point for the discrete or continuous time procedure, then $x \in NE(\phi)$.

Define the time average trajectories :

$$\bar{x}_t = \frac{1}{t}\int_0^t x_s ds \qquad \text{and} \qquad \bar{x}_n = \frac{1}{n}\sum_{m=1}^{n} x_m.$$

Lemma 2.2

If ϕ is dissipative, the accumulation points of $\{\bar{x}_t\}$ or $\{\bar{x}_n\}$ are in $SE(\phi)$.

Proof:

$$\frac{R_t(y)}{t} = \frac{1}{t}\int_0^t \langle \phi(x_s)|y - x_s\rangle \geq \frac{1}{t}\int_0^t \langle \phi(y)|y - x_s\rangle = \langle \phi(y)|y - \bar{x}_t\rangle.$$

Hence under the no-regret property any accumulation point $\hat{x}$ of $\{\bar{x}_t\}$ will satisfy $\langle \phi(y)|y - \hat{x}\rangle \leq 0$. ∎

This result implies the non-emptiness of $SE(\phi)$ for dissipative ϕ. In particular the minmax theorem (in the $\mathscr{C}^1$ case) follows from the existence of no-regret procedures.

Class (III): convex gradient.

$u_t = -\nabla f(x_t)$ with f $\mathscr{C}^1$ convex, this corresponds to a specific case of dissipative and continuous vector field ϕ, hence: $SE(\phi) = NE(\phi) = E = argmin_X f$.

Use that:

$$\langle \nabla f(x_t) | y - x_t \rangle \leq f(y) - f(x_t)$$

to obtain

$$\int_0^t [f(x_s) - f(y)]\, ds \leq \int_0^t \langle -\nabla f(x_s) | y - x_s \rangle ds = R_t(y)$$

which implies by Jensen's inequality:

$$f(\bar{x}_t) - f(y) \leq \frac{1}{t} \int_0^t [f(x_s) - f(y)]\, ds \leq \frac{R_t(y)}{t}. \qquad (10)$$

In particular one obtains:

Lemma 2.3

i) The accumulation points of $\{\bar{x}_t\}$ *or* $\{\bar{x}_n\}$ *belong to* E*.*
ii) If $t \mapsto f(x_t)$ *(resp.* $n \mapsto f(x_n)$*) is decreasing, the accumulation points of* $\{x_t\}$ *or* $\{x_n\}$ *belong to* E*.*

3.1. Cover (level) functions

Definition 3.1

$P : I\!R^+ \times X \to I\!R^+$ *is a* *cover function* *(for* $\{u_t, x_t\}$*) if:*

$$\langle u_t, x_t - y \rangle \geq \frac{d}{dt} P(t;y), \quad \forall t \in I\!R^+, \forall y \in X. \tag{11}$$

Proposition 3.1

The existence of a cover function implies:
$R_t(y) = \int_0^t \langle u_s | y - x_s \rangle ds \leq P(0;y) - P(t;y)$ *is bounded.*

(1) *no-regret property**: Rate of convergence* $1/t$*.*

(2) Class (II): Assume $y^* \in SE(\phi)$*, then* $P(t;y^*)$ *is decreasing:*

$$\frac{d}{dt} P(t;y^*) \leq \langle \phi(x_t), x_t - y^* \rangle \leq 0.$$

3.2. Positive correlation

We follow Sandholm (2010) [73].
Positive correlation, (between the dynamics and the vector field) holds if:

$$\dot{x}_t^i \neq 0 \Longrightarrow \langle \phi^i(x_t), \dot{x}_t^i \rangle > 0.$$

This notion is important for potential vector fields:

Proposition 3.2

Consider a vector field ϕ with potential W.
If the dynamics satisfies positive correlation, then W is a strict Lyapunov function.
All ω-limit points are rest points.

We will show that this property holds for the three dynamics defined below.

3.3 Hilbertian framework: Projected Dynamics

V Hilbert space, $X \subset V$, convex closed.

Dynamics

Analogous to projected gradient descent (Levitin and Polyak, 1966) and defined, as projected dynamics (PD), by $x_t \in X$ with:

$$\langle u_t - \dot{x}_t, y - x_t \rangle \leq 0, \; \forall y \in X. \tag{12}$$

which is:

$$\dot{x}_t = \Pi_{T_X(x_t)}(u_t). \tag{13}$$

since $T_C(x)$ is a cône.

Values

Proposition 3.3

$$V(t;y) = \frac{1}{2}\|x_t - y\|^2, \quad y \in X, \tag{14}$$

is a cover function.

Proof:
(12) gives:

$$\langle u_t, y - x_t \rangle \leq \langle \dot{x}_t, y - x_t \rangle = -\frac{d}{dt} V(t;y).$$

∎

Trajectories

Proposition 3.4

Assume ϕ dissipative and $E \neq \emptyset$.
$\{\bar{x}_t\}$ converges weakly to a point in E.

Proof:

- $\{\bar{x}_t\}$ is bounded hence has weak accumulation points.
- The weak limit points of $\{\bar{x}_t\}$ are in E
- $\|x_t - y^*\|$ converges when $y^* \in E$

Hence by Opial's lemma, $\bar{x}_t$ converges weakly to a point in E. ■

Lemma 3.1

Positive correlation holds.

Proof:

$$\langle \phi(x_t), \dot{x}_t \rangle = \|\dot{x}_t\|^2$$

since $\langle u_t - \dot{x}_t, \dot{x}_t \rangle = 0$ by (12) and Moreau's decomposition, Moreau, 1965 [58]. ■

Consider class (III): $u_t = -\nabla f(x_t)$.

Proposition 3.5

$f(x_t)$ *is decreasing and converges to* $f^* = \min_X f$ *at speed* $1/t$
Assume $E \neq \emptyset$. $\{x_t\}$ *weakly converges to a point in* E.

Proof:
Weak accumulation points of $\{x_t\}$ are in E.
Then Opial's lemma applies. ■

3.4. Mirror descent

Continuous version of "Mirror descent algorithm",
Nemirovski and Yudin (1983), Beck and Teboulle (2003).

Dynamics

$H : V \to \overline{I\!R}$, strictly convex, $\mathscr{C}^2$,
X (compact, convex) $\subset dom\ H$.
The continuous time process mirror descent (MD) satisfies, $x_t \in X$ and:

$$\langle u_t - \frac{d}{dt}\nabla H(x_t) | y - x_t \rangle \leq 0, \forall y \in X. \tag{15}$$

The previous analysis corresponds to the case: $H(x) = \frac{1}{2}\|x\|^2$.

Values

Bregman distance associated to H:

$$D_H(y,x) = H(y) - H(x) - \langle \nabla H(x) | y - x \rangle \ (\geq 0).$$

$$\frac{d}{dt} D_H(y, x_t) = \langle -\frac{d}{dt} \nabla H(x_t) | y - x_t \rangle, \tag{16}$$

so that (15) implies:

$$\langle u_t | y - x_t \rangle \leq -\frac{d}{dt} D_H(y, x_t).$$

Proposition 3.6

$P(t;y) = D_H(y, x_t)$ is a cover function.

Trajectories

The use of special functions H adapted to X allows to produce a trajectory that remains in $int X$ hence to get rid of the normal cône .
This leads to:

$$\frac{d}{dt}\nabla H(x_t) = u_t \tag{17}$$

$$\dot{x}_t = \nabla^2 H(x_t)^{-1} u_t. \tag{18}$$

which corresponds to a Riemannian metric, see Bolte and Teboulle, 2003 [15], Alvarez, Bolte and Brahic, 2004 [1], Mertikopoulos and Sandholm, 2018 [54].

Lemma 3.2

Positive correlation holds.

Proof :

$$\langle \phi(x_t) | \dot{x}_t \rangle = \langle \phi(x_t) | \nabla^2 H(x_t)^{-1} \phi(x_t) \rangle \geq 0.$$

Consider now class (III).

By Lemma 2.3 the accumulation points of $\{x_t\}$ are in E.

To prove convergence one introduces the following :

Hypothesis [H1]: if $z^k \to y^* \in S$ then $D_H(y^*, z^k) \to 0$.
For example H is L-smooth (see e.g. Nesterov, 2004 [62] Section 1.2.2.) and then:

$$0 \leq D_H(x,y) \leq \frac{L}{2}\|x-y\|^2.$$

Hypothesis [H2]: if $D_H(y^*, z^k) \to 0, y^* \in S$ then $z^k \to y^*$.
For example H is β-strongly convex (see e.g. Nesterov, 2004 [62] Section 2.1.3.) and then:

$$D_H(x,y) \geq \frac{\beta}{2}\|x-y\|^2.$$

Proposition 3.7

Consider class (III). If H is smooth and strongly convex, $\{x_t\}$ converges weakly to some $x^* \in E$.

Proof:

Let x^* be an accumulation point of $\{x_t\}$. Then $x^* \in E$ by Lemma 2.3 and thus $D_H(x^*, x_t)$ is decreasing. Since this sequence is decreasing to 0 on a subsequence $x_{t_k} \to x^*$ by [H1], it is decreasing to 0, hence by [H2] $x_t \to x^*$. ∎

3.5. Dual averaging

Continuous version of dual averaging Nesterov, 2009 [63].
We follow Kwon and Mertikopoulos, 2017 [47].

Dynamics

Assume $h : V \to \overline{IR}$ bounded strictly convex s.c.i. on $dom\, h = X$ (convex compact) $\subset V$.
Let $h^*(w) = \sup_{x \in V} \langle w|x\rangle - h(x)$ be the Fenchel conjugate. h^* is differentiable.
Introduce :

$$U_t = \int_0^t u_s ds$$

and let the dual averaging (DA) dynamics be defined by:

$$x_t = argmax\{\langle U_t|x\rangle - h(x); x \in V\} = argmax\{\langle U_t|x\rangle - h(x); x \in X\}.$$

The dynamics can be written as:

$$x_t = \nabla h^*(U_t) \in X \qquad (19)$$

Values

Consider the Fenchel coupling between U_t and y:

$$W(t;y) = h^*(U_t) + h(y) - \langle U_t | y \rangle \qquad (\geq 0). \tag{20}$$

$$\frac{d}{dt} h^*(U_t) = \langle u_t | \nabla h^*(U_t) \rangle = \langle u_t | x_t \rangle \tag{21}$$

thus:

$$\frac{d}{dt} W(t;y) = \langle u_t | x_t - y \rangle$$

Proposition 3.8

W is a level function.

Trajectories

Lemma 3.3

Positive correlation holds.

Proof:

$$\langle \phi(x_t) | \dot{x}_t \rangle = \langle \phi(x_t) | \nabla^2 h^*(U_t)(u_t) \rangle$$

with $u_t = \phi(x_t)$. ∎

Hence in class (III), using Lemma 2.3 the accumulation points of $\{x_t\}$ are in E.

3.6. Comments on the continuous time case

1) Existence of a level function and same speed of convergence of the no-regret quantities in classes (I), (II) or (III): $O(\frac{1}{t})$.
2) In the framework of games the entropy function:

$$h(x) = \sum_{p \in S} x^p Log x^p$$

defined on the simplex $X = \Delta(S)$ leads (via (MD) or (DA)) to the replicator dynamics on $int\ X$, Taylor and Jonker, 1978 [89], Hofbauer and Sigmund, 1998 [41], Sorin, 2009 [79], 2020 [81]. The corresponding Riemannian metric is introduced in Shahshahani, 1979 [76].

On the other hand, $h(x) = \frac{1}{2}\|x^2\|$ leads to the local/direct projection dynamics, for a comparison, see Lahkar and Sandholm, 2008 [48], Sandholm, Dokumaci and Lahkar, 2008 [75].

Recal that the replicator dynamics is the continuous version of the multiplicative weight algorithm, Littlestone and Warmuth, 1994 [51], Vovk, 1990 [92], Sorin, 2009 [79], 2020 [81].

3) There is an important literature on continuous time dynamics with similar features, see e.g. :
- in convex optimization: Attouch and Teboulle, 2004 [2], Attouch, Bolte, Redont and Teboulle, 2004 [3], Auslender and Teboulle, 2006 [7], 2009 [8]... Teboulle, 2018 [90],
- in game theory: Hofbauer and Sandholm, 2009 [40], Coucheney, Gaujal and Mertikopoulos, 2015 [22], Mertikopoulos and Sandholm, 2016 [53], Mertikopoulos and Sandholm (2018) [54], Mertikopoulos and Zhou (2019) [55] ...

We consider now discrete time algorithms.

Remark that the dynamics depends on an additional parameter, the step size.

4.1. Hilbertian framework: Projected Dynamics

Dynamics

Levitin and Polyak (1966) Polyak (1987) gradient projection method:

$$
\begin{aligned}
x_{m+1} &= argmin_X\{-\langle u_m, x\rangle + \frac{1}{2\eta_m}\|x - x_m\|^2\}, \\
&= argmax_X\{\langle u_m, x\rangle - \frac{1}{2\eta_m}\|x - x_m\|^2\},
\end{aligned} \qquad (22)
$$

with η_m decreasing, which corresponds to:

$$
x_{m+1} = \Pi_X[x_m + \eta_m u_m], \qquad (23)
$$

or with variational characterization:

$$
\langle x_m + \eta_m u_m - x_{m+1}, y - x_{m+1}\rangle \leq 0, \forall y \in X. \qquad (24)
$$

Values

Let $m(X)$ be the diameter of X. Assume $\|u_m\| = \|u_m\|_* \leq M$.

Proposition 4.1

$$R_n(x) \leq \frac{1}{2\eta_n} m(X)^2 + \frac{M^2}{2} \sum_{m=1}^{n} \eta_m$$

hence with $\eta_n = 1/\sqrt{n}$:

$$R_n(x) \leq O(\sqrt{n}).$$

Trajectories

Lemma 4.1

For $x^* \in SE(\phi)$, $\|x_m - x^*\|$ *converges if* $\eta_n \in \ell^2$.

Proposition 4.2

If $\eta_n \in \ell^2$ *and* g *is dissipative,* $\{\bar{x}_n\}$ *converges to a point in* $SE(\phi)$.

4.2. Mirror descent

Assumption:
H, L*-strongly convex for some norm* $\|.\|$ *on* $V = I\!R^n$.
$\|u_n\|_* \leq M$.

Dynamics

Nemirovski and Yudin (1983), Beck and Teboulle (2003)
The mirror descent algorithm is given by :

$$x_{m+1} = argmin_X\{-\langle u_m | \, x\rangle + \frac{1}{\eta_m} D_H(x, x_m)\}, \qquad (25)$$

Variational formulation:

$$\langle \nabla H(x_m) + \eta_m u_m - \nabla H(x_{m+1}) | x - x_{m+1}\rangle \leq 0, \forall x \in X. \qquad (26)$$

Values

Proposition 4.3

$$R_n(x) \leq \frac{D_H(x, x_1)}{\eta} + n\eta \frac{M^2}{2L}.$$

Then $\eta = 1/\sqrt{n}$ and $R_n(x) \leq O(\sqrt{n})$.

Same property with $\eta_n = 1/\sqrt{n}$ via double trick.

4.3. Dual averaging

Assumptions:
a) h *is a l.s.c. function from* V *to* $I\!R \cup \{+\infty\}$, L*-strongly convex for some norm* $\|.\|$ *on* $V = I\!R^n$, *with* $dom\, h = X$.
b) $\|u_m\|_* \leq M, \forall n \in$ ***N****.*

Dynamics

Dual averaging, Nesterov (2009).
Let $U_m = \sum_{k=1}^{m} u_k$
The algorithm is again given by a maximization property:

$$\begin{aligned} x_{m+1} &= argmin_X\{-\langle U_m|x\rangle + (1/\eta_m)h(x)\}, \\ &= argmax_X\{\langle U_m|x\rangle - (1/\eta_m)h(x)\} \end{aligned} \quad (27)$$

which is:

$$x_{m+1} = \nabla h^*(\eta_m U_m).$$

where $\{\eta_m\}$ is decreasing.

Values

Xiao (2010) or discrete approximation of (19) Kwon and Mertikopoulos (2017).

Proposition 4.4

$$R_n(x) = \sum_{m=1}^{n} \langle u_m | x - x_m \rangle \leq \frac{r_X(h)}{\eta_n} + \frac{\sum_{m=1}^{n} \eta_{m-1} \|u_m\|_*^2}{2L}. \tag{28}$$

Assume: $\|u_m\|_* \leq M$.
Cconvergence rate $O(\sqrt{n})$ with time varying parameters $\eta_m = 1/\sqrt{m}$.

4.4. Comments on the discrete time case

1) The three algorithms achieve the same bound $O(1/\sqrt{n})$ for the speed of convergence of the average regret, which is optimal already in class (III), Nesterov, 2004 [62] , using time varying step sizes $\eta_n = 1/\sqrt{n}$.

2) More precise properties concerning the trajectories are available only in the (PD) set-up. The results are similar to the ones in the continuous case, Section 3.2, if $\eta_n \in \ell^2$. (Compare to the analysis in Peypouquet and Sorin, 2010 [67] for dynamics induced by maximal monotone operators in discrete and continuous time.)

3) For vector fields ϕ with potential W one does not have the property $W(x_n)$ decreasing.

This section deals with class (III) *convex gradient*, where in addition f satisfies some regularity properties.

Recall that f is β smooth if:

$$|f(y)-f(x)-\langle\nabla f(x)|y-x\rangle| \leq \frac{\beta}{2}\|x-y\|^2. \tag{29}$$

Equivalently, ∇f is β-Lipschitz.

5.1. Hilbertian framework: Projected Dynamics

Assumption: f is β smooth.

Same procedure with constant steps:

$$x_{m+1} = \Pi_X(x_m - \eta \nabla f(x_m)).$$

The analysis in this section is standard, see e.g. Nesterov, 2004 [62].
Take $\eta = 1/\beta$ and define $v_n = \beta(x_{n+1} - x_n)$.

The main tool is the following:

Lemma 5.1 (Descent lemma)

$$f(x_{n+1}) - f(y) \leq \langle v_n, y - x_n \rangle - \frac{1}{2\beta} \|v_n\|^2.$$

In particular $f(x_n)$ decreasing and $\{\|v_n\|\} \in \ell^2$.

Values

$$n[f(x_{n+1})-f(y)] \leq R_n(y) - \frac{1}{2\beta}\|\sum_{m=1}^{n}\|v_m\|^2 = \frac{\beta}{2}\|y-x_1\|^2.$$

Hence convergence rate of the order $\frac{1}{n}$ with constant step size.

Trajectories

Lemma 5.2

Let $y^* \in E$. Then $\|x_n - y^*\|$ decreases.

Proposition 5.1

$\{x_n\}$ converges to a point in E.

5.2. Mirror descent

The dynamics is still:

$$\langle \nabla H(x_n) - \lambda \nabla f(x_n) - \nabla H(x_{n+1}) | x - x_{n+1} \rangle \leq 0, \forall x \in X.$$

We follow Bauschke, Bolte and Teboulle, 2017 [9]

H and f are $\mathscr{C}^1$
Hypothesis [A]:
there exists $L > 0$ such that:

$$L\, D_H - D_f \geq 0$$

(preorder: $L\, H - f$ convex, Nguyen, 2017 [64])

If H is strongly convex and f is smooth, [A] holds.

Values

One has, by [A]:

$$f(x) \leq f(y) + \langle \nabla f(z) | x - y \rangle + L D_h(x,z) - D_f(y,z)$$

(the last term is ≤ 0 when f is convex).
Take $2\lambda L = 1$

Proposition 5.2

Assume H convex.
1) $f(x_n)$ is decreasing.
2) $\sum D_H(x_{n+1}, x_n) < +\infty$.
3) Assume f convex, lower bounded.

$$f(x_n) - f(y) \leq \frac{2L}{n} D_H(y, x_1)$$

Trajectories

Proposition 5.3

Assume f convex.
1) $y^* \in E$ implies $D_H(y^*, x_n)$ decreases.
2) Assume:
$[H1] : x^k \to x^* \in E \Rightarrow D_H(x^*, x^k) \to 0$
$[H2] : x^* \in E, D_H(x^*, x^k) \to 0 \Rightarrow x^k \to x^*$
Then $\{x_n\}$ converges to a point in E.

5.3. Dual averaging

We follow Lu, Freund and Nesterov (2018)

Dual averaging with constant step size under Hypothesis $[A]$:
$L\,h - f$ convex
f convex and $\mathscr{C}^1$
$h : V \to I\!R \cup \{+\infty\}$ l.s.c. with $dom\, h = X$.

$$x_{m+1} = argmax_X \{\langle U_m | x \rangle - L\,h(x)\} \tag{30}$$

with $u_k = -\nabla f(x_k)$.

Values

Proposition 5.4

f convex, lower bounded.

$$f(\bar{x}_n) - f(y) \leq \frac{L}{n} h(y), \quad \forall y \in X.$$

5.4. Comments on the regular case

1) In the three cases (PD), (MD) and (DA) the speed of convergence of the values is $O(1/n)$ and the algorithms use a constant step parameter.

2) Using (PD) with f smooth implies $f(x_n)$ decreasing and the convergence of $\{x_n\}$.

3) The approach in Section 5.2 shows that similar results can be obtained using (MD) without assuming f with Lipschitz gradient if the regularization function H is adapted to f : condition $[A]$.

4) Analogous results for the values are much simpler to obtain in the (DA) framework. However the properties concern the value at the average $f(\bar{x}_n)$ and no result is available on the trajectories.

For the three dynamics (PG), (MD) and (DA) the following 1), 2) and 3) holds:

1) In continuous time the speed of convergence of the average regret to 0, of the order $O(1/t)$ is not better in the general gradient convex case than in on-line learning.

2) In discrete time the speed of convergence of the average regret to 0, of the order $O(1/\sqrt{n})$ is not better in the general gradient convex case than in on-line learning.

3) Adding a regularity hypothesis on the convex function does not change the convergence rate in continuous time but allow a better convergence in discrete time from $O(1/\sqrt{n})$ to $O(1/n)$.

4) A similar phenomena appears with the "acceleration procedures" following Nesterov, 1983 [61].

In the continuous time case a second order ODE leads to a speed of convergence $f(x_t) - f(x^*) \leq O(\frac{1}{t^2})$ with no further hypothesis on f, see Su, Boyd and Candes, 2014 [86], 2016 [87], Krichene, Bayen and Bartlett, 2015 [45], 2016 [46], Wibisono, Wilson and Jordan, 2016 [93], Attouch and Peypouquet, 2016 [4], Attouch, Chbani, Peypouquet and Redont, 2018 [5]...

To obtain a similar property in discrete time, namely $f(x_n) - f(x^*) \leq O(\frac{1}{n^2})$ one has to assume f smooth.

The same remark apply to the (weak) convergence of the trajectory, where the smooth hypothesis on f is needed in discrete time and not in continuous time, Chambolle and Dossal, 2015[21], Attouch, Chbani, Peypouquet, Redont 2018 [5]...

5) Concerning the link between discrete and continuous time dynamics, there are no direct results of the form: no-regret property in continuous time imply no-regret property in discrete time but analogy of the tools used and ad-hoc choice of the stage step size, see Sorin, 2009 [79], Kwon and Mertikopoulos, 2017 [47] and the Lyapounov functions in Krichene, Bayen and Bartlett, 2015 [45], 2016 [46], Wibisono, Wilson and Jordan, 2016 [93].

6) The Hilbert framework for (PD) allows to obtain convergence results on the trajectories. The two other algorithms are more flexible and can achieve better explicit speed of convergence of the values by choosing an adequate norm, adapted to the problem, see the discussion in Bauschke, Bolte and Teboulle, 2017 [9]. For (MD), specific regularization functions H can also lead to convergence of the trajectories. (DA) is much simpler to implement due to its integral formulation. However no convergence properties of the trajectories are in general available.

Bibliography I

- Alvarez F., J. Bolte and O. Brahic (2004) Hessian Riemannian gradient flows in convex programming, *SIAM Journal on Control and Optimization*, **43**, 477-501.
- Attouch H. and M. Teboulle (2004) Regularized Lotka-Volterra dynamical system as continuous proximal-like method in optimization, *JOTA*, **121**, 541-570.
- Attouch H., Bolte J., Redont P. and M. Teboulle (2004) Singular Riemannian barrier methods and gradient-projection dynamical systems for constrained optimization, *Optimization*, **53**, 435-454.
- Attouch H. and J. Peypouquet (2016) The rate of convergence of Nesterov's accelerated forward backward method is actually faster than $1/k^2$, *SIAM Journal on Optimization*, **26**, 1824-1834.

Bibliography II

- Attouch H., Z. Chbani, J. Peypouquet and P. Redont (2018) Fast convergence of inertial dynamics and algorithms with asymptotic vanishing damping, *Mathematical Programming*, **168**, 123-175.
- Auer P., Cesa-Bianchi N., Freund Y. and R.E. Shapire (2002) The nonstochastic multiarmed bandit problem, *SIAM J. Comput.*, **32**, 48-77.
- Auslender A. and M. Teboulle (2006) Interior gradient and proximal methods for convex and conic optimization, *SIAM J Optimization*, **16**, 697-725.
- Auslender A. and M. Teboulle (2009) Projected subgradient methods with non Euclidean distances for non-differentiable convex minimization and variational inequalities, *Mathematical Programming*, **120**, 27-48.

Bibliography III

- Bauschke H., Bolte J. and M. Teboulle (2017) A descent Lemma beyond Lipschitz gradient continuty: first-order methods revisited and applications, *Mathematics of Operations Research*, **42**, 330-348.
- Beck A. and M. Teboulle (2003) Mirror descent and nonlinear projected subgradient methods for convex optimization, *Oper. Res. Lett.*, **31**, 167-175.
- Benaim M., J. Hofbauer and S. Sorin (2005) Stochastic approximations and differential inclusions, *SIAM Journal on Control and Optimization*, **44**, 328-348.
- Blackwell D. (1954) Controlled random walks, Erven P. Noordhoff, N.V. Groningen (Eds.), Proc. Int. Congress of Mathematicians, vol. 3, North-Holland, 335-338.
- Blackwell D. (1956) An analog of the minimax theorem for vector payoffs, *Pacific Journal of Mathematics*, **6**, 1-8.

Bibliography IV

- Blum A. and Y. Mansour (2007) From external to internal regret, *Journal of Machine Learning Reserach*, **8**, 1307-1324.
- Bolte J. and M. Teboulle (2003) Barrier operators and associated gradient-like dynamical systems for constrained minimization problems, *SIAM Journal on Control and Optimization*, **42**, 1266-1292.
- Brézis H. (1973) *Opérateurs maximaux monotones et semi-groupes de contractions dans les espaces de Hilbert*, North Holland.
- Bubeck S. (2011) *Introduction to online optimization*, Lecture Notes.
- Bubeck S. (2015) Convex optimization: Algorithms and complexity, *Fondations and Trends in Machine Learning*, **8**, 231-357.

Bibliography V

- Cesa-Bianchi N. and G. Lugosi (2003) Potential-based algorithms in on-line prediction and game theory, *Machine Learning*, **51**, 239-261.
- Cesa-Bianchi N. and G. Lugosi (2006) *Prediction, Learning and Games*, Cambridge University Press.
- Chambolle A. and C. Dossal (2015) On the convergence of the iterates of FISTA, *Journal of Optimization Theory and Applications*, **166**, 968-982.
- Coucheney P., B. Gaujal and P. Mertikopoulos (2015) Penalty-regulated dynamics and robust learning procedures in games, *Mathematics of Operations Research*, **40**, 611-633.
- Cover T. (1991) Universal Portfolios, *Math. Finance*, **1**, 1-29.

Bibliography VI

- Dupuis P. and A. Nagurney (1993) Dynamical systems and variational inequalities, *Ann. Oper. Res.*, **44**, 9–42.
- Facchinei F. and J. Pang (2007) *Finite-Dimensional Variational Inequalities and Complementarity Problems*, Springer.
- Foster D. and R. Vohra (1993) A randomization rule for selecting forecasts, *Operations Research*, **41**, 704-707.
- Foster D. and R. Vohra (1999) Regret in the on-line decision problem, *Games and Economic Behavior*, **29**, 7-35.
- Freund Y. and R.E. Schapire (1999) Adaptive game playing using multiplicative weights, *Games and Economic Behavior*, **29**, 79-103.

Bibliography VII

- Fudenberg D. and D. K. Levine (1995) Consistency and cautious fictitious play, *Journal of Economic Dynamics and Control*, **19**, 1065-1089.
- Fudenberg D. and D. K. Levine (1998) *The Theory of Learning in Games*, MIT Press.
- Fudenberg D. and D. K. Levine (1999) Conditional universal consistency, *Games and Economic Behavior*, **29**, 104-130.
- Hannan J. (1957) Approximation to Bayes risk in repeated plays, *Contributions to the Theory of Games, III*, Drescher M., A.W. Tucker and P. Wolfe eds., Princeton University Press, 97-139.
- Hart S. and A. Mas-Colell (2000) A simple adaptive procedure leading to correlated equilibria, *Econometrica*, **68**, 1127-1150.

Bibliography VIII

- Hart S. and A. Mas-Colell (2001a) A general class of adaptive strategies, *Journal of Economic Theory*, **98**, 26-54.
- Hart S. and A. Mas-Colell (2003) Regret-based continuous time dynamics, *Games and Economic Behavior*, **45**, 375-394.
- Hart S. and A. Mas Colell (2013) *Simple Adaptive Strategies: From Regret-Matching to Uncoupled Dynamics*, World Scientific Publishing.
- Hazan E. (2011)The convex optimization approach to regret minimization, *Optimization for machine learning*, S. Sra, S. Nowozin, S. Wright eds, MIT Press, 287-303.
- Hazan E. (2015) Introduction to Online Convex Optimization, *Fondations and Trends in Optimization*, **2**, 157-325.

Bibliography IX

Hazan E. (2019) Optimization for Machine Learning , https://arxiv.org/pdf/1909.03550.pdf.

Hofbauer J. and W. H. Sandholm (2009) Stable games and their dynamics, *Journal of Economic Theory*, **144**, 1710-1725.

Hofbauer J. and K. Sigmund (1998) *Evolutionary Games and Population Dynamics*, Cambridge U.P.

Hofbauer J. and K. Sigmund (2003) Evolutionary Game Dynamics, *Bulletin of the A.M.S.*, **40**, 479-519.

Kalai A. and S. Vempala (2005) Efficient algorithms for online decision problems, *Journal of Computer and System Sciences*, **71**, 291-307.

Bibliography X

- Kinderlehrer D. and G. Stampacchia (1980) *An Introduction to Variational Inequalities and their Applications*, Academic Press.
- Krichene W., A. Bayen, and P. Bartlett (2015) Accelerated mirror descent in continuous and discrete time, *NIPS 2015*.
- Krichene W., A. Bayen, and P. Bartlett (2016) Adaptive Averaging in Accelerated Descent Dynamics, *NIPS 2016*.
- Kwon J. and P. Mertikopoulos (2017) A continuous time approach to on-line optimization, *Journal of Dynamics and Games*, **4**, 125-148.
- Lahkar R. and W. H. Sandholm (2008) The projection dynamic and the geometry of population games, *Games Econom. Behav.*, **64**, 565–590.
- Lehrer E. (2003) A wide range no-regret theorem, *Games and Economic Behavior*, **42**, 101-115.

Bibliography XI

- Levitin E.S. and B.T. Polyak (1966) Constrained minimization methods, *USSR Computational Mathematics and Mathematical Physics*, **6**, 1-50.
- Littlestone N. and M.K. Warmuth (1994) The weighted majority algorithm, *Information and Computation*, **108**, 212-261.
- Lu H., Freund and Y. Nesterov (2018) Relatively smooth convex optimization by first-order methods, and applications, Siam J Opt, 28, 333-354.
- Mertikopoulos P. and W. H. Sandholm (2016) Learning in games via reinforcement and regularization, *Mathematics of Operations Research*, **41**, 1297-1324.
- Mertikopoulos P. and W. H. Sandholm (2018) Riemannian game dynamics, *Journal of Economic Theory*, **177**, 315-364.

Bibliography XII

- Mertikopoulos P. and Z. Zhou (2019) Learning in games with continuous action sets and unknown payoff functions, *Mathematical Programming*, **173**, 465-507.
- Minty G. J.(1967) On the generalization of a direct method of the calculus of variations, *Bulletin A.M.S.*, **73**, 315–321.
- Monderer D. and L. S. Shapley (1996) Potential games, *Games Econom. Behav.*, **14**, 124–143.
- Moreau J.J. (1965) Proximité et dualité dans un espace Hilbertien, *Bull. Soc. Math. France*, **93**, 273–299.
- Nagurney A. and D. Zhang (1996) *Projected Dynamical Systems and Variational Inequalities with Applications* , Kluwer.
- Nemirovski A. and D. Yudin (1983) *Problem Complexity and Method Efficiency in Optimization*, Wiley.

Bibliography XIII

- Nesterov Y. (1983) A method of solving a convex programming problem with convergence rate o(1/k2), *Soviet Mathematics Doklady*, **27**, 372-376.
- Nesterov Y. (2004) *Introductory Lectures on Convex Optimization*, Kluwer.
- Nesterov Y. (2009) Primal-dual subgradient methods for convex problems, *Mathematical Programming*, **120**, 221-259.
- Nguyen Q.V. (2017) Forward-Backward splitting with Bregman distances, *Vietnam J. Math.*, **45**, 519-539.
- Opial Z. (1967) Weak Convergence of the sequence of successive approximations for nonexpansive mappings, *Bull. Amer. Math. Soc.*, **73**, 591-597.

Bibliography XIV

- Perchet V. (2014) Approachability, regret and calibration: implications and equivalences, *Journal of Dynamics and Games*, **1**, 181-254.
- Peypouquet J. and S. Sorin (2010) Evolution equations for maximal monotone operators: asymptotic analysis in continuous and discrete time, *Journal of Convex Analysis*, **17**, 1113-1163.
- Polyak B. (1987) *Introduction to optimization*, Optimization software.
- Rakhlin A. (2009) *Lecture notes on On Line Learning*.
- Rockafellar R.T. (1970) Monotone operators associated with saddle-functions and minmax problems, *Nonlinear Functional Analysis, F. Browder, ed.*, Proceedings of Symposia in Pure Math, **18**, AMS , 241-250.

Bibliography XV

- Rosen J.B. (1965) Existence and uniqueness of equilibrium points for concave N-person games, *Econometrica*, **33**, 520-534.
- Sandholm W. H. (2001) Potential games with continuous player sets, *Journal of Economic Theory*, **97**, 81-108.
- Sandholm W. H. (2010) *Population Games and Evolutionary Dynamics*, MIT Press.
- Sandholm W. H. (2015) Population games and deterministic evolutionary dynamics, H. P. Young, S. Zamir, eds *Handbook of Game Theory IV*, Elsevier, 703-778.
- Sandholm W. H., E. Dokumaci and R. Lahkar (2008) The projection dynamic and the replicator dynamic, *Games and Economic Behavior*, **64**, 666-683.

Bibliography XVI

- Shahshahani S. (1979) A new mathematical framework for the study of linkage and selection, *Memoirs of the American Mathematical Society*, 211.
- Smith M.J. (1984) A descent algorithm for solving monotone variational inequalities and monotone complementarity problems, *J. Optim. Theory Appl.*, **44**, 485–496.
- Shalev-Shwartz S. (2012) Online Learning and Online Convex Optimization, *Foundations and Trends in Machine Learning*, **4** , 107-194.
- Sorin S. (2009) Exponential weight algorithm in continuous time, *Mathematical Programming*, Ser. B, **116**, 513-528.

Bibliography XVII

- Sorin S. (2011) On some global and unilateral adaptive dynamics, in *Evolutionary Game Dynamics*, K. Sigmund ed., Proceedings of Symposia in Applied Mathematics, **69**, A.M.S., 81-109.
- Sorin S. (2020) Replicator dynamics: Old and new, *Journal of Dynamics and Games*, **7**, 365-385.
- Sorin S. (2022) On some continuous time algorithms in optimization and game theory, *Dynamic Games and Applications*, https://doi.org/10.1007/s13235-021-00423-x.
- Sorin S. (2023) No-regret algorithms in on-line learning, games and convex optimization, *Mathematical Programming*, to appear.
- Sorin S. and C. Wan (2016) Finite composite games: equilibria and dynamics, *Journal of Dynamics and Games*, **3**, 101-120.

Bibliography XVIII

- Stoltz G. and G. Lugosi (2005) Internal Regret in On-Line Portfolio Selection, *Machine Learning*, **59**, 125-159.
- Su W., S. Boyd, and E. Candes (2014) A differential equation for modeling Nesterov's accelerated gradient method: Theory and insights, *NIPS 2014*.
- Su W., S. Boyd, and E. Candes (2016) A differential equation for modeling Nesterov's accelerated gradient method: Theory and insights, *Journal of Machine Learning Research*, **17**, 1-43.
- Swinkels J.M. (1983) Adjustment dynamics and rational play in games, *Games and Economic Behavior*, **5**, 455–484.
- Taylor P. and L. Jonker (1978) Evolutionary stable strategies and game dynamics, *Mathematical Biosciences*, **40**, 145-156.

Bibliography XIX

- Teboulle M. (2018) A simplified view of first order methods for optimization, *Math Programming*, **170**, 67-96.
- Viossat Y. and A. Zapechelnyuk (2013) No-regret dynamics and fictitious play, *Journal of Economic Theory*, **148**, 825-842.
- Vovk V. (1990) Aggregating strategies, *Proceedings of the 3rd Annual Conference on Computational Learning Theory*, 371-383.
- Wibisono A., A. C. Wilson, and M. I. Jordan (2016) A variational perspective on accelerated methods in optimization, *PNAS* , **113**, 7351-7358.
- Xiao L. (2010) Dual averaging methods for regularized stochastic learning and online optimization, *Journal of Machine Learning Research*, **11**, 2543-2596.

Bibliography XX

- Young P. (2004) *Strategic Learning and Its Limits*, Oxford U. P.
- Zhang D. and A. Nagurney (1997) Formulation, stability, and computation of traffic network equilibria as projected dynamical systems, *J. Optim. Theory Appl.*, **93**, 417–444.
- Zinkevich M. (2003) Online convex programming and generalized infinitesimal gradient ascent, *Proceedings of the 20th International Conference on Machine Learning*, 928-936.