Optimization, learning and games
“Replicator dynamics: old and new”

Sylvain Sorin
sylvain.sorin@imj-prg.fr

Institut de Mathématiques Jussieu-PRG
Sorbonne Université, UPMC - Paris 6
CNRS UMR 7586

GDO 2018
University of Vienna, Faculty of Mathematics
COST Action CA16228 Workshop
March 13-15, 2018

Abstract
The purpose of this talk is to underline links between algorithms used in convex optimization, games and on-line learning, using both discrete and continuous time approaches.
We will introduce the unilateral version associated to the replicator dynamics and describe its connection to on-line learning procedures and classical gradient.
We will survey recent results on extension of this dynamics and some applications : regularization functions and variable weights, equilbria and variational inequalities, potential and dissipative games, Hessian Riemannian metrics.

Part of this research is supported by a PGMO grant: "COGLED - Convergence of gradient-like and evolutionary dynamics", with Jerôme Bolte.

Basic properties

Origin: one population (Taylor and Jonker, 1978)

Consider the evolution of the composition of a large population with K types.
The interaction is represented by a $K \times K$ fitness matrix A.
A_{ij} is the outcome of "i" facing "j" (amount of offsprings).
Discrete dynamics:

$$N_{m+1}^k = N_m^k (1 + h\, e^k A x_m)$$

N_m^k : number of elements of type k at stage m
x_m : corresponding proportion of types
e^k : k-th unit vector
h : time step size.

Continuous version: x_t^k is the frequency of type k at time t.

Replicator dynamics on the simplex $\Delta(K)$ of $\mathbb{R}^K$:

$$\dot{x}_t^k = x_t^k \left(e^k A x_t - x_t A x_t\right), \quad k \in K \qquad (RD)$$

Strong links with Evolutionary Stable Strategies
(Maynard Smith, 1982; Hofbauer and Sigmund, 1998)

Continuous version: x_t^k is the frequency of type k at time t.

Replicator dynamics on the simplex $\Delta(K)$ of $I\!R^K$:

$$\dot{x}_t^k = x_t^k \left(e^k A x_t - x_t A x_t \right), \quad k \in K \qquad (RD)$$

Strong links with Evolutionary Stable Strategies
(Maynard Smith, 1982; Hofbauer and Sigmund, 1998)

Two populations (cross matching)

Two $S \times T$ fitness matrices A and B, $x_t \in \Delta(S)$, $y_t \in \Delta(T)$:

$$\dot{x}_t^p = x_t^p[e^p A y_t - x_t A y_t], \qquad p \in S$$

and

$$\dot{y}_t^q = y_t^q[x_t B e^q - x_t B y_t], \qquad q \in T.$$

I populations

$F^i : S = \prod_{j \in I} S^j \to I\!R$ with multilinear extension to $\prod_j \Delta(S^j)$, $x_t^i \in \Delta(S^i)$:

$$\dot{x}_t^{ip} = x_t^{ip}[F^i(e^{ip}, x_t^{-i}) - F^i(x_t^i, x_t^{-i})], \qquad p \in S^i, i \in I$$

natural interpretation in terms of game: $x_t^i \in \Delta(S^i)$ is a mixed strategy of player i.
Set of rest points= $\{NE(F'); F' \subset F\}$, F' restricted game.

Two populations (cross matching)

Two $S \times T$ fitness matrices A and B, $x_t \in \Delta(S)$, $y_t \in \Delta(T)$:

$$\dot{x}_t^p = x_t^p [e^p A y_t - x_t A y_t], \qquad p \in S$$

and

$$\dot{y}_t^q = y_t^q [x_t B e^q - x_t B y_t], \qquad q \in T.$$

I populations

$F^i : S = \prod_{j \in I} S^j \to I\!R$ with multilinear extension to $\prod_j \Delta(S^j)$, $x_t^i \in \Delta(S^i)$:

$$\boxed{\dot{x}_t^{ip} = x_t^{ip} [F^i(e^{ip}, x_t^{-i}) - F^i(x_t^i, x_t^{-i})], \qquad p \in S^i, i \in I}$$

natural interpretation in terms of game: $x_t^i \in \Delta(S^i)$ is a mixed strategy of player i.
Set of rest points= $\{NE(F'); F' \subset F\}$, F' restricted game.

Unilateral replicator dynamics for one participant

Given a bounded measurable process $U = \{U_t \in I\!R^K\}$, the U-replicator process $(U - RP)$ is specified by the following equation on $\Delta(K)$:

$$\dot{x}_t^k = x_t^k[U_t^k - \langle x_t, U_t \rangle], \qquad k \in K \qquad (U - RP)$$

Logit representation

Define the logit map L from $I\!R^K$ to $\Delta(K)$ by:

$$L^k(V) = \frac{\exp V^k}{\sum_j \exp V^j}.$$

Rustichini (1999), Hofbauer, Sorin, Viossat (2009):

Proposition

$$x_t = L(\int_0^t U_s\, ds) \quad \textit{satisfies} \quad (U-RP):$$

$$\dot{x}_t^k = x_t^k[U_t^k - \langle x_t, U_t\rangle], \qquad k \in K.$$

Explicit representation of the dynamics

Logit representation

Define the logit map L from $I\!R^K$ to $\Delta(K)$ by:

$$L^k(V) = \frac{\exp V^k}{\sum_j \exp V^j}.$$

Rustichini (1999), Hofbauer, Sorin, Viossat (2009):

Proposition

$$x_t = L(\int_0^t U_s\, ds) \quad \textit{satisfies} \quad (U-RP):$$

$$\dot{x}_t^k = x_t^k [U_t^k - \langle x_t, U_t \rangle], \qquad k \in K.$$

Explicit representation of the dynamics

Variational approach

Let $H(x) = \sum_k x^k \log x^k$ be the entropy function on $\Delta(K)$.

Proposition

$L(V)$ is the argmax of $x \mapsto \langle V, x \rangle - H(x)$ on $\Delta(K)$.

Model and no-regret

Consider a discrete time process $\{U_n\}$ of vectors in $\mathscr{U} = [-1, 1]^K$.

At each stage n, an agent having observed the past realizations of the vectors $U_1, ..., U_{n-1}$, chooses a component k_n in K. Let

$$\omega_n = U_n^{k_n}.$$

This generates a (random) trajectory in K. One compares the average of $\{U_n\}$ on it and on any constant trajectory $\hat{k}_n = k \in K$.

Formally a strategy σ is a map from histories $h_{n-1} = (k_1, U_1, ..., k_{n-1}, U_{n-1})$ to probability distributions on K. $x_n = \sigma(h_{n-1})$ is the law of k_n.

Model and no-regret

Consider a discrete time process $\{U_n\}$ of vectors in $\mathscr{U} = [-1,1]^K$.
At each stage n, an agent having observed the past realizations of the vectors $U_1, ..., U_{n-1}$, chooses a component k_n in K. Let

$$\omega_n = U_n^{k_n}.$$

This generates a (random) trajectory in K. One compares the average of $\{U_n\}$ on it and on any constant trajectory $\hat{k}_n = k \in K$.

Formally a strategy σ is a map from histories $h_{n-1} = (k_1, U_1, ..., k_{n-1}, U_{n-1})$ to probability distributions on K.
$x_n = \sigma(h_{n-1})$ is the law of k_n.

Model and no-regret

Consider a discrete time process $\{U_n\}$ of vectors in $\mathscr{U} = [-1, 1]^K$.
At each stage n, an agent having observed the past realizations of the vectors $U_1, ..., U_{n-1}$, chooses a component k_n in K. Let

$$\omega_n = U_n^{k_n}.$$

This generates a (random) trajectory in K. One compares the average of $\{U_n\}$ on it and on any constant trajectory $\hat{k}_n = k \in K$.

Formally a strategy σ is a map from histories $h_{n-1} = (k_1, U_1, ..., k_{n-1}, U_{n-1})$ to probability distributions on K.
$x_n = \sigma(h_{n-1})$ is the law of k_n.

Model and no-regret

Consider a discrete time process $\{U_n\}$ of vectors in $\mathscr{U} = [-1, 1]^K$.
At each stage n, an agent having observed the past realizations of the vectors $U_1, ..., U_{n-1}$, chooses a component k_n in K. Let

$$\omega_n = U_n^{k_n}.$$

This generates a (random) trajectory in K. One compares the average of $\{U_n\}$ on it and on any constant trajectory $\hat{k}_n = k \in K$.

Formally a strategy σ is a map from histories $h_{n-1} = (k_1, U_1, ..., k_{n-1}, U_{n-1})$ to probability distributions on K.
$x_n = \sigma(h_{n-1})$ is the law of k_n.

Definition

A strategy σ satisfies no external regret (ER) if, for every process $\{U_m\}$:

$$\bar{R}_n^k = \frac{1}{n} R_n^k = \frac{1}{n} \sum_{m=1}^{n} (U_m^k - \omega_m) \leq A_n^k \longrightarrow 0, \text{ as } n \to \infty, \quad \forall k \in K, \text{ a.s.}$$

Foster and Vohra (1999) , Fudenberg and Levine (1995)

Similarly the average internal regret

$$\overline{S}_n^{k\ell} = \frac{1}{n} \sum_{m=1, k_m=k}^{n} (U_m^{\ell} - U_m^{k})$$

is the comparison for each component $k \in K$, of the average payoff obtained on the dates where k was selected, to the payoff for an alternative choice $\ell \in K$.

Definition

A strategy σ satisfies no internal regret (IR) if, for every process $\{U_m\}$ and every couple $k, \ell \in K$:

$$[\overline{S}_n^{k\ell}] \leq B_n^{k\ell} \longrightarrow 0 \text{ as } n \to +\infty, \quad \text{a.s.}$$

see Foster and Vohra (1999) , Fudenberg and Levine (1999)

There exist algorithms that generates IR procedures from ER ones, by using paralell inputs:
Blum and Mansour (2007), Stoltz and Lugosi (2005).

Application to games

Finite games:
set of players I,
action spaces S^i,
payoff function $G^i : S = S^i \times S^{-i} \to I\!R, i \in I$.
repeated interaction and standard signalling
each player i knows G^i
observation after each stage n: vector of actions of his opponents, $s_n^{-i} \in S^{-i}$.

Fix $i \in$ and let $K = S^i$.
Player i knows after each stage n:
the stage payoff $\omega_n = G^i(k_n, s_n^{-i})$ as well as
the vector payoff $U_n = G^i(., s_n^{-i}) \in I\!R^K$.

Application to games

Finite games:
set of players I,
action spaces S^i,
payoff function $G^i : S = S^i \times S^{-i} \to I\!R, i \in I$.
repeated interaction and standard signalling
each player i knows G^i
observation after each stage n: vector of actions of his opponents, $s_n^{-i} \in S^{-i}$.

Fix $i \in$ and let $K = S^i$.
Player i knows after each stage n:
the stage payoff $\omega_n = G^i(k_n, s_n^{-i})$ as well as
the vector payoff $U_n = G^i(., s_n^{-i}) \in I\!R^K$.

Linearity allows to translate the properties from the set of payoffs to the set of moves.
Introduce $z_n = \frac{1}{n}\sum_{m=1}^{n} s_m \in \Delta(S)$ with $s_m = \{s_m^j\}, j \in I$: empirical distribution on moves up to stage n.

$$\begin{aligned}
\overline{R}_n &= [\frac{1}{n}\sum_{m=1}^{n}\{G^i(k, s_m^{-i}) - G^i(s_m)\}; k \in K] \\
&= [G^i(k, \frac{1}{n}\sum_{m=1}^{n} s_m^{-i}) - G^i(\frac{1}{n}\sum_{m=1}^{n} s_m); k \in K] \\
&= [G^i(k, z_n^{-i}) - G^i(z_n); k \in K]
\end{aligned}$$

Then σ satisfies no ER is equivalent to :
$d(z_n, H^i) \to 0$ a.s., with:

$$H^i = \{z \in \Delta(S); G^i(k, z^{-i}) - G^i(z) \leq 0, \forall k \in K\},$$

Hannan's set for player i, Hannan (1955).

Similarly $\overline{S}_n = \mathbf{S}(z_n)$ is the $K \times K$ matrix with:

$$\mathbf{S}^{k,j}(z) = \sum_{\ell \in S^{-i}} [G^i(j,\ell) - G^i(k,\ell)] z(k,\ell)$$

and σ satisfies no IR is equivalent to $d(z_n, C^i) \to 0$ a.s. with:

$$C^i = \{z \in \Delta(S); \mathbf{S}^{k,j}(z) \leq 0, \forall k,j \in K\}.$$

Note that $\cap_i C^i$ is the set of correlated equilibrium distributions (Aumann, 1974).

Alternative proof of existence through the existence of no IR procedures.

No similar result for Nash equilibria.

Similarly $\overline{S}_n = \mathbf{S}(z_n)$ is the $K \times K$ matrix with:

$$\mathbf{S}^{k,j}(z) = \sum_{\ell \in S^{-i}} [G^i(j,\ell) - G^i(k,\ell)] z(k,\ell)$$

and σ satisfies no IR is equivalent to $d(z_n, C^i) \to 0$ a.s. with:

$$C^i = \{z \in \Delta(S); \mathbf{S}^{k,j}(z) \leq 0, \forall k,j \in K\}.$$

Note that $\cap_i C^i$ is the set of correlated equilibrium distributions (Aumann, 1974).

Alternative proof of existence through the existence of no IR procedures.

No similar result for Nash equilibria.

Exponential weight algorithm

Conditional expectation

Total regret at stage n, that one wants to control, is:

$$R_n = \{ \sum_{m=1}^{n} U_m^k - \omega_m, \quad k \in K \}$$

where $\omega_m = U_m^{k_m}$ is the random payoff at stage m.
Let $x_m \in \Delta(K)$ be the law of k_m, then:

$$\mathsf{E}(\omega_m | h_{m-1}) = \langle U_m, x_m \rangle$$

so that $\omega_m - \langle U_m, x_m \rangle$ is a bounded martingale difference.

Thus we will study quantities of the form:

$$\sum_{m=1}^{n} U_m^k - \langle U_m, x_m \rangle, \quad k \in K$$

or equivalently:

$$\sum_{m=1}^{n} \langle U_m, x \rangle - \langle U_m, x_m \rangle, \quad x \in \Delta(K)$$

because of the linearity.
The no ER condition becomes:

$$ER_n(x) = \sum_{m=1}^{n} \langle U_m, x - x_m \rangle \leq C_n = o(n), \quad \forall x \in \Delta(K).$$

Dynamical variational inequality:

$$\limsup \frac{1}{n} \sum_{m=1}^{n} \langle U_m, x - x_m \rangle \leq 0, \quad \forall x \in \Delta(K).$$

Thus we will study quantities of the form:

$$\sum_{m=1}^{n} U_m^k - \langle U_m, x_m \rangle, \quad k \in K$$

or equivalently:

$$\sum_{m=1}^{n} \langle U_m, x \rangle - \langle U_m, x_m \rangle, \quad x \in \Delta(K)$$

because of the linearity.
The no ER condition becomes:

$$\boxed{ER_n(x) = \sum_{m=1}^{n} \langle U_m, x - x_m \rangle \leq C_n = o(n), \quad \forall x \in \Delta(K).}$$

Dynamical variational inequality:

$$\limsup \frac{1}{n} \sum_{m=1}^{n} \langle U_m, x - x_m \rangle \leq 0, \quad \forall x \in \Delta(K).$$

Exponential weight algorithm (EW) in discrete time:

$$\sigma^k(h_n) = x_{n+1}^k = \frac{\exp(A\sum_{m=1}^n U_m^k)}{\sum_{j=1}^K \exp(A\sum_{m=1}^n U_m^j)}$$

where A is a positive parameter.
Vovk (1990), Littlestone and Warmuth (1994), Freund and Schapire (1999).

Main result (Auer and alii, 1995):

Proposition

For $A = 1/\sqrt{n}$, the exponential weight algorithm satisfies:

$$ER_n(x) \leq M\sqrt{n}.$$

Exponential weight algorithm (EW) in discrete time:

$$\sigma^k(h_n) = x_{n+1}^k = \frac{\exp(A\sum_{m=1}^n U_m^k)}{\sum_{j=1}^K \exp(A\sum_{m=1}^n U_m^j)}$$

where A is a positive parameter.
Vovk (1990), Littlestone and Warmuth (1994), Freund and Schapire (1999).

Main result (Auer and alii, 1995):

Proposition

For $A = 1/\sqrt{n}$, the exponential weight algorithm satisfies:

$$ER_n(x) \leq M\sqrt{n}.$$

Continuous time approach

Given a measurable process $U_t, t \geq 0$, with values in $[0,1]^K$, define: continuous time exponential weight algorithm $(CTEW)$ = replicator dynamics = measurable process $x_t \in \Delta(K)$, by:

$$x_t^k = L^k(\int_0^t U_s \, ds) = \frac{\exp \int_0^t U_s^k ds}{\sum_{j \in K} \exp \int_0^t U_s^j ds}$$

Sorin (2009)
(Similar continuous time approaches: Cesa-Bianchi and Lugosi (2003), Hart and Mas-Colell (2003)).

Theorem

Conditional expected ER holds for $x_t = L(\int_0^t U_s ds)$:

$$\int_0^T \langle U_s, x - x_s \rangle ds \leq C_T = o(T).$$

Continuous time approach

Given a measurable process $U_t, t \geq 0$, with values in $[0,1]^K$, define: continuous time exponential weight algorithm $(CTEW)$ = replicator dynamics = measurable process $x_t \in \Delta(K)$, by:

$$x_t^k = L^k\left(\int_0^t U_s\, ds\right) = \frac{\exp \int_0^t U_s^k ds}{\sum_{j \in K} \exp \int_0^t U_s^j ds}$$

Sorin (2009)

(Similar continuous time approaches: Cesa-Bianchi and Lugosi (2003), Hart and Mas-Colell (2003)).

Theorem

Conditional expected ER holds for $x_t = L(\int_0^t U_s ds)$:

$$\int_0^T \langle U_s, x - x_s \rangle ds \leq C_T = o(T).$$

Application for discrete process:

Given a discrete process $\{U_m\}$ and a corresponding EW algorithm $\{x_m\}$ the aim is to get a bound on:

$$\frac{1}{n}\sum_{m=1}^{n}\langle U_m, x - x_m\rangle$$

from an evaluation of:

$$\frac{1}{T}\int_0^T \langle V_s, y - y_s\rangle ds$$

where V_t is a continuous process constructed from U_m and y_t is the $CTEW$ algorithm associated to V_t.

Alternative proof of:

$$\frac{1}{n}\sum_{m=1}^{n}\langle x - x_m, U_m\rangle \leq Mn^{-1/2}$$

Given n, choose $T = \sqrt{n}$ so that:

- the bound in the continuous version is of the order $1/T = 1/\sqrt{n}$

$$\frac{1}{T}\int_0^T \langle y - y_t, V_t\rangle dt \leq \frac{\log K}{\sqrt{n}}$$

- the error term with the discrete approximation with step size $T/n = 1/\sqrt{n}$ is:

$$|\frac{1}{n}\sum_{m=1}^{n}\langle x_m, U_m\rangle - \frac{1}{T}(\int_0^T \langle y_t, V_t\rangle dt)| \leq L/\sqrt{n}$$

Alternative proof of:

$$\frac{1}{n}\sum_{m=1}^{n}\langle x - x_m, U_m\rangle \leq Mn^{-1/2}$$

Given n, choose $T = \sqrt{n}$ so that:
- the bound in the continuous version is of the order $1/T = 1/\sqrt{n}$

$$\frac{1}{T}\int_0^T \langle y - y_t, V_t\rangle dt \leq \frac{\log K}{\sqrt{n}}$$

- the error term with the discrete approximation with step size $T/n = 1/\sqrt{n}$ is:

$$|\frac{1}{n}\sum_{m=1}^{n}\langle x_m, U_m\rangle - \frac{1}{T}(\int_0^T \langle y_t, V_t\rangle dt)| \leq L/\sqrt{n}$$

Extensions: regularization and variable parameter

We follow the analysis in Kwon and Mertikopoulos (2017).
- extend the analysis from $\Delta(K)$ to any compact convex set $X \subset I\!R^n$,
- replace the entropy function by any strictly convex l.s.c. regularization function F with $domF = X$.
- use time variable positive parameters η_t.
Let:

$$S_F(V) = argmax_X[\langle V|x\rangle - F(x)], \tag{1}$$

then COLD (Cumulative On Line Dynamics) is defined by:

$$x_t = S_F(\eta_t \int_0^t U_s ds). \tag{2}$$

(compare $x_t = L(\int_0^t U_s ds)$ for the replicator dynamics.)

Extensions: regularization and variable parameter

We follow the analysis in Kwon and Mertikopoulos (2017).
- extend the analysis from $\Delta(K)$ to any compact convex set $X \subset I\!R^n$,
- replace the entropy function by any strictly convex l.s.c. regularization function F with $domF = X$.
- use time variable positive parameters η_t.

Let:

$$S_F(V) = argmax_X[\langle V|x\rangle - F(x)], \tag{1}$$

then COLD (Cumulative On Line Dynamics) is defined by:

$$x_t = S_F(\eta_t \int_0^t U_s ds). \tag{2}$$

(compare $x_t = L(\int_0^t U_s ds)$ for the replicator dynamics.)

Extensions: regularization and variable parameter

We follow the analysis in Kwon and Mertikopoulos (2017).
- extend the analysis from $\Delta(K)$ to any compact convex set $X \subset I\!R^n$,
- replace the entropy function by any strictly convex l.s.c. regularization function F with $domF = X$.
- use time variable positive parameters η_t.

Let:

$$S_F(V) = argmax_X[\langle V|x\rangle - F(x)], \tag{1}$$

then COLD (Cumulative On Line Dynamics) is defined by:

$$x_t = S_F(\eta_t \int_0^t U_s ds). \tag{2}$$

(compare $x_t = L(\int_0^t U_s ds)$ for the replicator dynamics.)

Note that:

$$S_F(V) = \nabla F^*(V)$$

where $F^*(V) = \max_X[\langle V|x\rangle - F(x)]$ is the Fenchel conjugate of F.

Continuous time bound

Assume η_t positive, decreasing in t and let $r_X(F) = \sup_X F(x) - \inf_X F(x)$.

Proposition (Kwon and Mertikopoulos (2017))

$$R_t(x) = \int_0^t \langle U_s|x - x_s\rangle ds \leq \frac{1}{\eta_t} r_X(F).$$

Application to the discrete process

Assume now that F is L-strongly convex for some norm $\|.\|$ on $V = I\!R^n$.
Consider a positive decreasing sequence η_k and a discrete time process $\{U_k\} \in V^*$. Let:

$$x_{k+1} = S_F(\eta_k \sum_{j=1}^{k} U_j) = \nabla F^*(\eta_k \sum_{j=1}^{k} U_j).$$

Proposition (Kwon and Mertikopoulos, 2017)

$$R_k(x) = \sum_{j=1}^{k} \langle U_j | x - x_j \rangle \leq \frac{r_X(F)}{\eta_k} + \frac{\sum_{j=1}^{k} \eta_{j-1} \|U_j\|_*^2}{2L}$$

where the first term corresponds to the continuous trajectory bound and the second to the approximation error.

i) $\eta_k = k^{-1/2}$ gives convergence of the mean regret R_k/k to 0 with speed $0(k^{-1/2})$.
ii) $\eta_k = 1$ corresponds to the replicator dynamics with best continuous time convergence speed but bad discrete approximation: $\sum_{j=1}^{k} \eta_{j-1} \sim k$.
iii) $\eta_k = \frac{1}{k\varepsilon}$ gives R_k/k of the order of ε.
It corresponds to the continuous process:

$$x_t = S_F(\frac{1}{\varepsilon} \times \frac{1}{t} \int_0^t U_s ds)$$

and the discrete associated dynamics is the smooth fictitious play since with $\overline{U}_k = \frac{1}{k} \sum_{j=1}^{k} U_j$, $x_{k+1} \in \mathbf{br}^{\varepsilon}(\overline{U}_k)$ i.e. maximizes:

$$\langle x, \overline{U}_k \rangle - \varepsilon F(x).$$

One recovers the bound of Fudenberg and Levine (1995).

Many alternative approaches to no ER or no IR procedures: Hannan, Blackwell, Cover, Foster and Vohra, Hart and Mas-Colell, Kalai and Vempala, Fudenberg and Levine, Freund and Shapire, Cesa-Bianchi and Lugosi, Benaim, Hofbauer and Sorin ....

Gradient descent and extensions

Let $f : \mathbf{R}^n \to \mathbf{R}$ be a C^1 convex function together with its gradient dynamics in $I\!R^n$:

$$\dot{x}_t + \nabla f(x_t) = 0. \tag{3}$$

The Euler scheme (with step size λ_k) is:

$$x_{k+1} - x_k = -\lambda_k \nabla f(x_k) \tag{4}$$

which corresponds to:

$$x_{k+1} = argmin\{\langle \nabla f(x_k), x\rangle + (1/2\lambda_k)\|x - x_k\|^2\}. \tag{5}$$

For the problem under constraints, with X compact and convex:

$$\min f(x), \quad x \in X \tag{6}$$

one aims at a point $x^* \in S = argmin_X f(x)$, i.e. satisfying:

$$\langle \nabla f(x^*), x - x^* \rangle \geq 0, \quad \forall x \in X. \tag{7}$$

A standard discrete dynamics is (Levitin and Polyak (1966), Polyak (1987)):

$$x_{k+1} = argmin_X \{ \langle \nabla f(x_k), x \rangle + (1/2\lambda_k) \|x - x_k\|^2 \}, \tag{8}$$

the gradient projection algorithm which corresponds to:

$$x_{k+1} = \Pi_X [x_k - \lambda_k \nabla f(x_k)], \tag{9}$$

or with variational characterization:

$$\langle x_k - \lambda_k \nabla f(x_k) - x_{k+1}, x - x_{k+1} \rangle \leq 0, \forall x \in X. \tag{10}$$

Letting $x_k = x_{t_k}$, with $t_{k+1} = t_k + \lambda_k$, leads to the continuous formulation:

$$\langle -\nabla f(x_t) - \dot{x}_t, x - x_t - \dot{x}_t \rangle \leq 0, \qquad \forall x \in X, \tag{11}$$

or

$$\langle -\nabla f(x_t) - \dot{x}_t, z - \dot{x}_t \rangle \leq 0, \qquad \forall z \in TC_X(x_t), \tag{12}$$

where $TC_X(x)$ is the tangent cône to X at $x \in X$,
which gives:

$$\boxed{\dot{x}_t = \Pi_{TC_X(x_t)}[-\nabla f(x_t)]. \qquad (13)}$$

Extensions

Still in the framework of convex optimization this algorithm has been extended in two directions: mirror approach where the increment with the gradient occurs in the dual space (Nemirovski and Yudin, 1983), on-line extension where one faces a sequence of convex functions.

Mirror approach

Given a duality (V, V^*) on $I\!R^n$ represented by $\langle .|.\rangle$ and a strictly convex differentiable function F with $X \subset domF$, the associated Bregman distance is :

$$D_F(x,y) = F(x) - F(y) - \langle \nabla F(y) | x - y \rangle.$$

The algorithm is now (Beck and Teboulle, 2003):

$$x_{k+1} = argmin_X \{ \langle \nabla f(x_k) | x \rangle + (1/\lambda_k) D_F(x, x_k) \}. \quad (14)$$

The variational expression takes the form:

$$\langle \nabla F(x_k) - \lambda_k \nabla f(x_k) - \nabla F(x_{k+1}) | x - x_{k+1} \rangle \leq 0, \forall x \in X, \quad (15)$$

which corresponds to the "greedy mirror descent algorithm".

The "lazy" variant (dual averaging, Nesterov 2009) starts with $y_0 = 0 \in V^*$, then $y_k = y_{k-1} - \lambda_k \nabla f(x_k)$ and:

$$\langle y_k - \nabla F(x_{k+1}) | x - x_{k+1} \rangle \leq 0, \forall x \in X, \qquad (16)$$

or :

$$\langle -\sum_{j=1}^{k} \lambda_j \nabla f(x_j) - \nabla F(x_{k+1}) | x - x_{k+1} \rangle \leq 0, \forall x \in X. \qquad (17)$$

A third variant (that coincides with the previous one in the case of constant step size) is :

$$\langle -\eta_k \sum_{j=1}^{k} \nabla f(x_j) - \nabla F(x_{k+1}) | x - x_{k+1} \rangle \leq 0, \forall x \in X, \qquad (18)$$

and corresponds to the discrete COLD for F differentiable and $U_k = -\nabla f(x_k)$. (Recall $x_{k+1} = argmax_X [\langle \eta_k \sum_{j=1}^{k} U_j, x \rangle - F(x)]$.)

The aim is to control, for $x \in X$ and f convex Lipschitz:

$$
\begin{aligned}
f(\bar{x}) - f(x) &\leq \sum_1^K \lambda_k/\Lambda_K [f(x_k) - f(x)] && (19) \\
&\leq \sum_1^K \lambda_k/\Lambda_K \langle -\nabla f(x_k), x - x_k \rangle && (20) \\
&= 1/\Lambda_K \sum_1^K \langle -\lambda_k \nabla f(x_k), x - x_k \rangle && (21)
\end{aligned}
$$

with $\Lambda_K = \sum_1^K \lambda_k$ and $\Lambda_K \bar{x} = \sum_1^K \lambda_k x_k$.

The previous analysis and majoration apply.

On-line approach

Starting with Zinkevich (2003) on-line procedures similar to the gradient projection agorithm (8) have being developped to control the quantity $\sum_k f_k(x_k) - f_k(x)$ where f_k is a sequence of (unknown) equi-Lipschitz convex functions.
One has:

$$\sum_1^K [f_k(x_k) - f_k(x)] \leq \sum_1^K \langle -\nabla f_k(x_k), x - x_k \rangle.$$

and $-\nabla f_k(x_k)$ plays the rôle of U_k.

Summary

There is a huge recent litterature on the links between on-line learning and convex optimization.

Cesa-Bianchi and Lugosi (2006) Prediction, Learning and Games, Cambridge UP.

Rakhlin A. (2009) Lecture notes on On Line Learning.

Hazan E. (2011)The convex optimization approach to regret minimization, *Optimization for machine learning*, S. Sra, S. Nowozin, S. Wright eds, 287-303, MIT Press.

Bubeck S. (2011) Introduction to online optimization, Lecture Notes.

Shalev-Shwartz S. (2012) Online Learning and Online Convex Optimization, *Foundations and Trends in Machine Learning,* **4**, *107-194, 2012.*

...

COLD approach:

Given a strictly convex l.s.c. penalization/regularization function F define a dynamics on X:

$$x_t = \nabla F^* [\int_0^t U_s \, ds] \qquad (OLD)$$

that satisfies, for any bounded adapted process U_t:

$$R_t(x) = \int_0^t \langle U_s | x - x_s \rangle ds \leq 0(t), \quad \forall x \in X.$$

1) Does not assume F differentiable.

2) Replace an incremental approach:

$$x_{k+1} - x_k = M(x_k, U_k; \lambda_k)$$

leading to an ODE:

$$\dot{x}_t = \overline{M}(x_t, U_t),$$

by a cumulative one:

$$x_{k+1} = Q(U_k, \sum_{j=1}^{k-1} U_j; \lambda_k)$$

leading to an explicit representation:

$$x_t = \overline{Q}(\int_0^t U_s ds).$$

3) Give the same (or better) bounds than greedy /lazy algorithms

4) Apply to any bounded (in V^*) sequence $\{U_k\}$ that plays the rôle of $-\nabla f_k(x_k)$.

Typical examples are:
- replicator dynamics on the simplex $\Delta(K)$:

$$\dot{x}_t^k = x_t^k[U_t^k - \langle U_t | x_t \rangle], k \in K,$$

- local projection dynamics:

$$\dot{x}_t = \Pi_{TC_X(x_t)}[U_t].$$

Global dynamics

Properties of COLD in finite multilinear games have been studied in Coucheney P., B. Gaujal and Mertikopoulos (2015), Mertikopoulos and Sandholm (2016) : extensions of the “Folk theorem of evolutionary dynamics”...

Variational inequalities

The replicator dynamics as well as COLD can be extended to other configurations were equilibrium conditions take the form of variational inequalities, Sorin and Wan (2016).

The framework is defined by a finite set I and
evaluation functions Φ^i
with for each $i \in I, X^i \subset I\!R^{k^i}$ compact convex, $X = \prod X^i$,
$\Phi^i : X \longrightarrow I\!R^{k^i}$ continuous.
$NE(\Phi)$ = set of equilibria of $\Gamma(\Phi)$ = set of solutions of the variational inequality:

$$\langle \Phi(x), x - y \rangle = \sum_i \langle \Phi^i(x), x^i - y^i \rangle \geq 0, \qquad \forall y \in X. \tag{22}$$

Examples:
- finite games : $\Phi^i = VG^i$
- concave continuous games : $\Phi^i = \nabla^i g^i$
- population games : $\Phi^i = F^i$

The framework is defined by a finite set I and
evaluation functions Φ^i
with for each $i \in I, X^i \subset I\!R^{k^i}$ compact convex, $X = \prod X^i$,
$\Phi^i : X \longrightarrow I\!R^{k^i}$ continuous.
$NE(\Phi)$ = set of equilibria of $\Gamma(\Phi)$ = set of solutions of the variational inequality:

$$\langle \Phi(x), x - y \rangle = \sum_i \langle \Phi^i(x), x^i - y^i \rangle \geq 0, \qquad \forall y \in X. \tag{22}$$

Examples:
- finite games : $\Phi^i = VG^i$
- concave continuous games : $\Phi^i = \nabla^i g^i$
- population games : $\Phi^i = F^i$

The replicator dynamics is now

$$\dot{x}_t^{ip} = x_t^{ip}[\Phi_t^{ip}(x_t) - \overline{\Phi}^i(x_t)], \quad p \in S^i, i \in I,$$

where

$$\overline{\Phi}^i(x) = \langle x^i, \Phi^i(x) \rangle = \sum_{p \in S^i} x^{ip} \Phi^{ip}(x)$$

and for COLD, the U_t process for player i is $\Phi^i(x_t)$ so that:

$$x_t^i = \nabla F^* (\int_0^t \Phi^i(x_s)).$$

Note that if x_t converges to x^*,

$$\int_0^t \langle \Phi(x_s), x - x_s \rangle ds \leq o(t), \forall x \in X$$

implies

$$\langle \Phi(x^*), x - x^* \rangle \leq 0, \forall x \in X$$

so $x^* \in NE(\Phi)$.

A dynamics

$$\dot{x}_t = \mathscr{B}_\Phi(x_t)$$

satisfies positive correlation (PC) (Sandholm, 2010) if:

$$\langle \mathscr{B}^i_\Phi(x), \Phi^i(x) \rangle > 0, \quad \forall i \in I, \forall x \in X \text{ s.t. } \mathscr{B}^i_\Phi(x) \neq 0.$$

This corresponds to MAD (myopic adjustment dynamics, Swinkels (1993)).

Proposition

Replicator Dynamics satisfies (PC).
More generally COLD satisfies (PC).

Specific cases: Friesz, Bernstein, Mehta, Tobin and Ganjalizadeh (1994), Lahkar and Sandholm (2008), Sandholm , Dokumaci and Lahkar (2008).

Potential games

A real function W, of class $\mathscr{C}^1$ on a neighborhood Ω of X, is a potential for Φ if for each $i \in I$, there exists a strictly positive function $\mu^i(x)$ defined on X such that

$$\left\langle \nabla^i W(x) - \mu^i(x)\Phi^i(x), z^i \right\rangle = 0, \quad \forall x \in X, \forall z^i \in TX^i(x_i), \forall i \in I, \tag{23}$$

where $TX^i(x^i)$ is the tangent space to X^i at x^i.

The game $\Gamma(\Phi)$ is then called a potential game, Monderer and Shapley (1996), Sandholm (2001, 2009).

Theorem

Let $\Gamma(\Phi)$ be a game with potential W.

1. Every local maximum of W is an equilibrium of $\Gamma(\Phi)$.

2. If W is concave on X, then any equilibrium of $\Gamma(\Phi)$ is a global maximum of W on X.

Potential games

A real function W, of class $\mathscr{C}^1$ on a neighborhood Ω of X, is a potential for Φ if for each $i \in I$, there exists a strictly positive function $\mu^i(x)$ defined on X such that

$$\left\langle \nabla^i W(x) - \mu^i(x)\Phi^i(x), z^i \right\rangle = 0, \quad \forall x \in X, \forall z^i \in TX^i(x_i), \forall i \in I, \tag{23}$$

where $TX^i(x^i)$ is the tangent space to X^i at x^i.

The game $\Gamma(\Phi)$ is then called a potential game, Monderer and Shapley (1996), Sandholm (2001, 2009).

Theorem

Let $\Gamma(\Phi)$ be a game with potential W.

1. Every local maximum of W is an equilibrium of $\Gamma(\Phi)$.

2. If W is concave on X, then any equilibrium of $\Gamma(\Phi)$ is a global maximum of W on X.

Potential games

A real function W, of class $\mathscr{C}^1$ on a neighborhood Ω of X, is a potential for Φ if for each $i \in I$, there exists a strictly positive function $\mu^i(x)$ defined on X such that

$$\langle \nabla^i W(x) - \mu^i(x)\Phi^i(x), z^i \rangle = 0, \quad \forall x \in X, \forall z^i \in TX^i(x_i), \forall i \in I, \tag{23}$$

where $TX^i(x^i)$ is the tangent space to X^i at x^i.

The game $\Gamma(\Phi)$ is then called a potential game, Monderer and Shapley (1996), Sandholm (2001, 2009).

Theorem

Let $\Gamma(\Phi)$ be a game with potential W.

1. Every local maximum of W is an equilibrium of $\Gamma(\Phi)$.

2. If W is concave on X, then any equilibrium of $\Gamma(\Phi)$ is a global maximum of W on X.

Proposition

Consider a potential game $\Gamma(\Phi)$ *with potential function* W*. If the dynamics* $\dot{x} = \mathscr{B}_\Phi(x)$ *satisfies (PC), then* W *is a strict Lyapunov function for* $\mathscr{B}_\Phi$*. All* ω*-limit points are rest points of* $\mathscr{B}_\Phi$*.*

$$\frac{d}{dt}W(x_t) = \sum_i \langle \nabla^i W(x_t), \dot{x}_t^i \rangle = \sum_i h^i(x_t) \langle \Phi^i(x_t), \dot{x}_t^i \rangle > 0$$

In particular this applies to RD and COLD.

Note that this is a product of unilateral properties: for each i the dynamics for x^i belongs to COLD.

Proposition

Consider a potential game $\Gamma(\Phi)$ with potential function W. If the dynamics $\dot{x} = \mathscr{B}_\Phi(x)$ satisfies (PC), then W is a strict Lyapunov function for $\mathscr{B}_\Phi$. All ω-limit points are rest points of $\mathscr{B}_\Phi$.

$$\frac{d}{dt}W(x_t) = \sum_i \langle \nabla^i W(x_t), \dot{x}_t^i \rangle = \sum_i h^i(x_t) \langle \Phi^i(x_t), \dot{x}_t^i \rangle > 0$$

In particular this applies to RD and COLD.

Note that this is a product of unilateral properties: for each i the dynamics for x^i belongs to COLD.

Dissipative games

The game $\Gamma(\Phi)$ is dissipative if Φ satisfies:

$$\langle \Phi(x) - \Phi(y), x - y \rangle \leq 0, \qquad \forall\, (x,y) \in X \times X.$$

In the framework of population games, Hofbauer and Sandholm (2009) studied this class under the name "stable games".
Let $SNE(\Phi)$ be the set of $x \in X$ satisfying:

$$\langle \Phi(y), x - y \rangle \geq 0, \qquad \forall y \in X.$$

Proposition

If $\Gamma(\Phi)$ is dissipative

$$SNE(\Phi) = NE(\Phi).$$

in particular $NE(\Phi)$ is convex.

Results similar to potential games hold for dissipative games and several evolutionary dynamics with ad hoc Lyapunov functions.

Proposition

Consider a dissipative game $\Gamma(\Phi)$. Let $x^* \in NE(\Phi)$. Define:

$$H(x) = \sum_{i \in I} \sum_{p \in supp(x^{i*})} x_p^{i*} \ln \frac{x_p^{i*}}{x_p^i}.$$

For the replicator dynamics, H is a local Lyapunov function. If $\Gamma(\Phi)$ is strictly dissipative, then H is a local strict Lyapunov function.

A general formation is $H(x,p) = \sum_i \langle \nabla F(x^i), x^i - p^i \rangle - F(x^i)$ which is a Lyapounov function if $p = x^*$.

$H(x_t,p)$ can also be written as:
$G(y_t,p) = \sum_i F^*(y_t^i) - \langle y_t^i, p^i \rangle$
with $y_t^i = \int_0^t U_s^i ds$ and $U_s^i = \Phi^i(x_s)$.
Then:

$$\frac{d}{dt} G(y_t,p) = \sum_i \nabla F^*(y_t^i)\dot{y}_t^i - \langle \dot{y}_t^i, p^i \rangle = \langle \Phi(x_t), x_t - p$$

which is negative when $p = x^*$. Note finally the interest to use as variable the cumulative evaluation in the dual:

$$\dot{y}_t^i = \Phi^i(x_t) = \Phi^i(\{\nabla F^*(y_t^j)\}).$$

Akin(1979), Hofbauer and Sigmund (1998)
Consider the initial one population RD with symmetric interaction: $A =^t A$.
Then Ax derives from the potential $W(x) = \frac{1}{2}xAx$ and the replicator dynamics is a gradient for the Shashahani metric (1979) $(.|.)_x$, for x in the interior of $\Delta(K)$, defined on the tangent space by:

$$(u|v)_x = \sum_k \frac{1}{x^k} u^k v^k.$$

Then RD writes:

$$\dot{x}_t^k = grad_{x_t} W(x_t)$$

since:

$$(grad_{x_t} W(t)|v)_{x_t} = \sum_k \frac{1}{x_t^k} x_t^k [e^k A x_t - x_t A x_t] v^k = \langle Ax, v \rangle = DW(x).v$$

shows that:

$$grad_{x_t} W(x_t) = \{x_t^k [e^k A x_t - x_t A x_t]\}.$$

More generally assume F $\mathscr{C}^2$ strictly convex with $\|\nabla F(x)\| \to \infty$ as $x \to \partial X$.
Then COLD gives:

$$x_t = \nabla F^*(\int_0^t U_s ds)$$

thus:

$$\nabla F(x_t) = \int_0^t U_s ds$$

and:

$$\nabla^2 F(x_t)\dot{x}_t = U_t$$

finally :

$$\dot{x}_t = [\nabla^2 F(x_t)]^{-1} U_t$$

which corresponds when: $U_t = -\nabla f(x_t)$ to Hessian Riemannian Gradients Flows
Alvarez, Bolte and Brahic (2004), Attouch, Teboulle ...

Concluding comments

A class of dynamics adapted to a process U:
- cv for U gradient in convex analysis
- cv for U evaluation class of games (potential, dissipative)
- cv for U paiement vectoriel to Hannan set for finite games
-no regret properties in on-line learning.

Bibliography I

- AKIN E. (1979) *The Geometry of Population Genetics*, Springer.
- AKIN E. AND J. HOFBAUER (1982) Recurrence of the unfit, *Mathematical Biosciences*, **61**, 51-62.
- ALVAREZ F., J. BOLTE AND O. BRAHIC (2004) Hessian Riemannian gradient flows in convex programming, *SIAM Journal on Control and Optimization*, **43**, 477-501.
- AUER P., CESA-BIANCHI N., FREUND Y. AND R.E. SHAPIRE (2002) The nonstochastic multiarmed bandit problem, *SIAM J. Comput.*, **32**, 48-77.
- AUMANN R.J (1974) Subjectivity and correlation in randomized strategies, *Journal of Mathematical Economics*, **1**, 67-96.

Bibliography II

BENAIM M. (2015) On gradient like properties of population games, learning models and self reinforced processes, J.-P. Bourguignon et al (eds.), *Dynamic, Games and Science*, Springer, 117-152.

BENAIM M. AND M. FAURE (2013) Consistency of vanishingly smooth fictitious play, *Mathematics of Operations Research*, **38**, 437-450.

BENAIM M., J. HOFBAUER AND S. SORIN (2005) Stochastic approximations and differential inclusions, *SIAM Journal on Control and Optimization*, **44**, 328-348.

BENAIM M., J. HOFBAUER AND S. SORIN (2006) Stochastic approximations and differential inclusions. Part II: applications, *Mathematics of Operations Research*, **31**, 673-695.

Bibliography III

- BLACKWELL D. (1956) An analog of the minimax theorem for vector payoffs, *Pacific Journal of Mathematics*, **6**, 1-8.
- BLUM A. AND Y. MANSOUR (2007) From external to internal regret, *Journal of Machine Learning Reserach*, **8**, 1307-1324.
- BOLTE, J. (2003) Continuous gradient projection method in Hilbert spaces, *JOTA*, **119**, 235-259.
- 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.
- BROWN G. W. (1951) Iterative solution of games by fictitious play, in Koopmans T.C. (ed.) *Activity Analysis of Production and Allocation*, Wiley, 374-376.

Bibliography IV

- 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.
- 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.
- COUCHENEY P., B. GAUJAL AND P. MERTIKOPOULOS (2015) Penalty-regulated dynamics and robust learning procedures in games, *Mathematics of Operations Research*, **40**, 611-633.

Bibliography V

- Cover T. (1991) Universal Portfolios, *Math. Finance*, **1**, 1-29.
- Crandall, M. G. and T. M. Liggett (1971) Generation of semi-groups of nonlinear transformations on general Banach spaces, *Amer. J. Math.*, **93**, 265-298.
- Dafermos S.C. (1980) Traffic equilibrium and variational inequalities, *Transportation Sci.*, **14** , 42–54.
- Dupuis P. and A. Nagurney (1993) Dynamical systems and variational inequalities, *Ann. Oper. Res.*, **44**, 9–42.
- Foster D. and R. Vohra, (1993) A randomization rule for selecting forecasts, *Operations Research*, **41**, 704-707.
- Foster D. and R. Vohra, (1997) Calibrated leaning and correlated equilibria, *Games and Economic Behavior*, **21**, 40-55.

Bibliography VI

- FOSTER D. AND R. VOHRA, (1998) Asymptotic calibration, *Biometrika*, **85**, 379-390.
- FOSTER D. AND R. VOHRA, (1999) Regret in the on-line decision problem, *Games and Economic Behavior*, **29**, 7-35.
- FOSTER D. AND P. YOUNG, (2003) Learning, hypothesis testing, and Nash equilibria, *Games and Economic Behavior*, **45**, 73-96.
- FOSTER D. AND P. YOUNG, (2005) Regret testing, *Games and Economic Behavior*, **25**, 79-96.
- FREUND Y. AND R.E. SCHAPIRE (1999) Adaptive game playing using multiplicative weights, *Games and Economic Behavior*, **29**, 79-103.

Bibliography VII

FRIESZ T.L., D. BERNSTEIN, N. J. MEHTA, R. L. TOBIN AND S. GANJALIZADEH (1994) Day-to-day dynamic network disequilibria and idealized traveler information systems, *Oper. Res.*, **42**, 1120–1136.

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.

GAUNERSDORFER A. AND J. HOFBAUER (1995) Fictitious play, Shapley polygons and the replicator equation, *Games and Economic Behavior*, **11**, 279-303.

Bibliography VIII

- GILBOA I. AND A. MATSUI (1991) Social stability and equilibrium, *Econometrica*, **59**, 859-867.
- 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. (2005) Adaptive heuristics, *Econometrica*, **73**, 1401-1430.
- 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 (2001b) A reinforcement procedure leading to correlated equilibria, in *Economic Essays: A Festschrift for W. Hildenbrandt* ed. by G. Debreu, W. Neuefeind and W. Trocknel, Springer, 181-200.

Bibliography IX

- 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.
- HOFBAUER J. (2000) From Nash and Brown to Maynard Smith: Equilibria, dynamics and ESS, *Selection*, **1**, 81–88.
- HOFBAUER J. (2011) Deterministic evolutionary game dynamics, in *Evolutionary Game Dynamics*, K. Sigmund ed., Proceedings of Symposia in Applied Mathematics, **69**, A.M.S., 61-79.

Bibliography X

- Hofbauer J. and W. H. Sandholm (2002) On the global convergence of stochastic fictitious play, *Econometrica*, **70**, 2265-2294.
- 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.
- Hofbauer J. and S. Sorin (2006) Best response dynamics for continuous zero-sum games, *Discrete and Continuous Dynamical Systems-series B*, **6**, 215-224.

Bibliography XI

- HOFBAUER J., S. SORIN AND Y. VIOSSAT (2009) Time average replicator and best reply dynamics, *Mathematics of Operations Research*, **34**, 263-269.
- KALAI A. AND S. VEMPALA (2005) Efficient algorithms for online decision problems, *Journal of Computer and System Sciences*, **71**, 291-307.
- 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.
- LEVITIN E.S. AND B.T. POLYAK (1966) Constrained minimization methods, *USSR Computational Mathematics and Mathematical Physics*, **6**, 1-50.

Bibliography XII

- LITTLESTONE N. AND M.K. WARMUTH (1994) The weighted majority algorithm, *Information and Computation*, **108**, 212-261.
- MARDEN J. R., J. S. SHAMMA (2015) Game theory and distributed control. H. P. Young, S. Zamir, eds., *Handbook of Game Theory IV*, Elsevier, 861-899.
- MAYNARD SMITH J. (1982) *Evolution and the Theory of Games*, Cambridge U.P.
- 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 (2016) Riemannian game dynamics, preprint.

Bibliography XIII

- MONDERER D. AND L. S. SHAPLEY (1996) Potential games, *Games Econom. Behav.*, **14**, 124–143.
- MONDERER D. AND L.S. SHAPLEY (1996) Fictitious Play property for games with identical interests, *Journal of Economic Theory*, **68**, 258-265.
- NAGURNEY A.AND D. ZHANG (1997) Projected dynamical systems in the formulation, stability analysis, and computation of fixed demand traffic network equilibria, *Transportation Sci.*, **31**, 147–158.
- NEMIROVSKI A. AND D. YUDIN (1983) *Problem Complexity and Method Efficiency in Optimization*, Wiley.
- 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.

Bibliography XIV

- POLYAK B. (1987) *Introduction to optimization*, Optimization software.
- RAKHLIN A. (2009) *Lecture notes on On Line Learning*.
- ROBINSON J. (1951) An iterative method of solving a game, *Annals of Mathematics*, **54**, 296-301.
- RUSTICHINI A. (1999) Optimal properties of stimulus-response learning models, *Games and Economic Behavior*, **29**, 230-244.
- SANDHOLM W. H. (2001) Potential games with continuous player sets, *Journal of Economic Theory*, **97**, 81-108.
- SANDHOLM W. H. (2009) Large population potential games, *Journal of Economic Theory*, **144**, 1710-1725.
- SANDHOLM W. H. (2010) *Population Games and Evolutionary Dynamics*, MIT Press.

Bibliography XV

- 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.
- SHAHSHAHANI S. (1979) A new mathematical framework for the study of linkage and selection, *Memoirs of the American Mathematical Society*, 211.
- 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 XVI

- 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. AND C. WANG (2016) Finite composite games: equilibria and dynamics, *Journal of Dynamics and Games*, **3**, 101-120.
- STOLTZ G. AND G. LUGOSI (2005) Internal Regret in On-Line Portfolio Selection, *Machine Learning*, **59**, 125-159.
- SWINKELS J.M. (1983) Adjustment dynamics and rational play in games, *Games and Economic Behavior*, **5**, 455–484.

Bibliography XVII

- TAYLOR P. AND L. JONKER (1978) Evolutionary stable strategies and game dynamics, *Mathematical Biosciences*, **40**, 145-156.

- VIOSSAT Y. (2007) The replicator dynamics does not lead to correlated equilibria, *Games and Economic Behavior*, **59**, 397-407

- 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.

- WEIBULL J. W. (1995) *Evolutionary Game Theory*, MIT Press.

Bibliography XVIII

YOUNG P., *Strategic Learning and Its Limits*, Oxford U. P. , 2004.

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.