FORCES
FOUNDATIONS OF RESILIENT CYBER-PHYSICAL SYSTEMS

Online Learning on a Continuum

Walid Krichene
Maximilian Balandat **Claire Tomlin** **Alexandre Bayen**

Dept. of Electrical Engineering & Computer Sciences,
UC Berkeley, CA, USA

Berkeley University of California
Massachusetts Institute of Technology
NSF
Michigan
Vanderbilt University

Outline

Outline

1. The problem
2. Dual Averaging on $L^2(S)$
3. Dual averaging with ω potentials
4. An example

Online Learning over a finite set

A decision maker faces a sequential problem:

Online decision problem over a finite set $\{1, \ldots, N\}$.

---

for $t \in \mathbb{N}$ **do**
    Decision maker chooses distribution $x^{(t)}$ over $\{1, \ldots N\}$.
    A loss vector $\ell^{(t)} \in \mathbb{R}_+^N$ is revealed.
    The decision maker incurs expected loss $\sum_{n=1}^N \ell_n^{(t)} x_n^{(t)} = \left\langle x^{(t)}, \ell^{(t)} \right\rangle$
end for

FORCES
FOUNDATIONS OF RESILIENT CYBER-PHYSICAL SYSTEMS

Applications

Applications

- Convergence of player dynamics in games (Hannan, Blackwell)
  $\{1, \dots, N\}$ is the set of actions.
- Boosting in Machine Learning (Hazan, Shamir)
  $\{1, \dots, N\}$ is the training set.
- "Model-free" portfolio optimization (Cover, Blum)
  $\{1, \dots, N\}$ is the set of stocks.

Applications

Applications

- Convergence of player dynamics in games (Hannan, Blackwell)
  $\{1, \ldots, N\}$ is the set of actions.
- Boosting in Machine Learning (Hazan, Shamir)
  $\{1, \ldots, N\}$ is the training set.
- "Model-free" portfolio optimization (Cover, Blum)
  $\{1, \ldots, N\}$ is the set of stocks.

Applications

Applications

- Convergence of player dynamics in games (Hannan, Blackwell)
  $\{1, \ldots, N\}$ is the set of actions.
- Boosting in Machine Learning (Hazan, Shamir)
  $\{1, \ldots, N\}$ is the training set.
- "Model-free" portfolio optimization (Cover, Blum)
  $\{1, \ldots, N\}$ is the set of stocks.

Learning on a continuum

"What if the action set is infinite?"

Problem 1 Online decision problem on S.

for $t \in \mathbb{N}$ **do**
    Decision maker chooses distribution $x^{(t)}$ over S.
    A loss function $\ell^{(t)} : S \to \mathbb{R}_+$ is revealed.
    The decision maker incurs expected loss

$$\left\langle x^{(t)}, \ell^{(t)} \right\rangle = \int_S x^{(t)}(s)\ell^{(t)}(s)\lambda(ds) = \mathop{\mathbb{E}}_{s \sim x^{(t)}} [\ell^{(t)}(s)]$$

end for

Regret

$$R^{(T)}(x) = \sum_{t=1}^{T} \left\langle x^{(t)}, \ell^{(t)} \right\rangle - \left\langle x, \sum_{t=1}^{T} \ell^{(t)} \right\rangle$$

$$\sup_{(\ell^{(t)})} \sup_{x} R^{(T)}(x) = o(T)$$

Learning on a continuum

"What if the action set is infinite?"

Problem 2 Online decision problem on S.

for $t \in \mathbb{N}$ **do**
    Decision maker chooses distribution $x^{(t)}$ over S.
    A loss function $\ell^{(t)} : S \to \mathbb{R}_+$ is revealed.
    The decision maker incurs expected loss

$$\left\langle x^{(t)}, \ell^{(t)} \right\rangle = \int_S x^{(t)}(s)\ell^{(t)}(s)\lambda(ds) = \mathop{\mathbb{E}}_{s\sim x^{(t)}}[\ell^{(t)}(s)]$$

end for

Regret

$$R^{(T)}(x) = \sum_{t=1}^{T} \left\langle x^{(t)}, \ell^{(t)} \right\rangle - \left\langle x, \sum_{t=1}^{T} \ell^{(t)} \right\rangle$$

$$\sup_{(\ell^{(t)})} \sup_{x\in\Delta^N} R^{(T)}(x) = o(T)$$

Applications

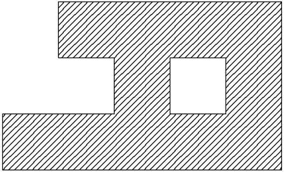
- Games with infinite action sets
  - Player dynamics
  - Computation of Nash equilibria
- Pricing problems:
  Action set is the price interval.
- Tracking on non-convex sets.

Applications

- Games with infinite action sets
  - Player dynamics
  - Computation of Nash equilibria
- Pricing problems:
  Action set is the price interval.
- Tracking on non-convex sets.

Applications

- Games with infinite action sets
  - Player dynamics
  - Computation of Nash equilibria
- Pricing problems:
  Action set is the price interval.
- Tracking on non-convex sets.



Results

Assumptions on $\ell^{(t)}$	convex	α-exp-concave	uniformly L-Lipschitz
Assumptions on S	convex	convex	v-uniformly fat
Method	Gradient (Zinkevich)	Hedge (Hazan et al.)	Dual Averaging (Krichene et al.)
Learning rates	$1/\sqrt{t}$	α	$1/\sqrt{t}$
$R^{(t)}$	$\mathcal{O}(\sqrt{t})$	$\mathcal{O}(\log t)$	$\mathcal{O}(\sqrt{t \log t})$

Table: Some regret upper bounds for different classes of losses.

Outline

A review of dual averaging (Nesterov)

Constrained convex optimization

$$\min_{x \in \mathcal{X}} f(x)$$

$\mathcal{X}$ closed, convex of a Hilbert $(H, \langle \cdot, \cdot \rangle)$. f convex.

Algorithm 3 Dual averaging method with dual sequence $(\ell^{(t)})$, learning rates (η_t), strongly convex regularizer ψ

for $t \in \mathbb{N}$ **do**
  Discover $\ell^{(t)} \in H^*$
  Define $L^{(t)} = \sum_{\tau=1}^{t} \ell^{(\tau)}$
  Update

$$x^{(t+1)} = \arg\min_{x \in \mathcal{X}} \left\langle L^{(t)}, x \right\rangle + \frac{1}{\eta_{t+1}} \psi(x) \quad (1)$$

end for

In convex optimization, $\ell^{(t)} = \nabla f(x^{(t)})$. But dual averaging has general guarantees, regardless of convexity.

A review of dual averaging (Nesterov)

Constrained convex optimization

$$\min_{x \in \mathcal{X}} f(x)$$

$\mathcal{X}$ closed, convex of a Hilbert $(H, \langle \cdot, \cdot \rangle)$. f convex.

Algorithm 4 Dual averaging method with dual sequence $(\ell^{(t)})$, learning rates (η_t), strongly convex regularizer ψ

---

for $t \in \mathbb{N}$ **do**

Discover $\ell^{(t)} \in H^*$

Define $L^{(t)} = \sum_{\tau=1}^{t} \ell^{(\tau)}$

Update

$$x^{(t+1)} = \arg\min_{x \in \mathcal{X}} \left\langle L^{(t)}, x \right\rangle + \frac{1}{\eta_{t+1}} \psi(x) \tag{1}$$

end for

In convex optimization, $\ell^{(t)} = \nabla f(x^{(t)})$. But dual averaging has general guarantees, regardless of convexity.

A review of dual averaging (Nesterov)

Dual averaging guarantee

$$\sum_{\tau=1}^{t} \left\langle \ell^{(\tau)}, x^{(\tau)} - x \right\rangle \leq \frac{1}{\eta_t}\psi(x) + \frac{M^2}{2\ell_\psi} \sum_{\tau=1}^{t} \eta_\tau$$

(here M is a bound on $\|\ell^{(t)}\|_*$)

Consequence

Convex optimization	Online learning
$f\left(\frac{1}{t}\sum_{\tau=1}^{t} x^{(\tau)}\right) - f^\star \to 0$	$\sup_{x\in\Delta^N} R^{(t)}(x) = o(t)$

Idea

1. Take $\mathcal{X} = \Delta(S)$.
2. Apply dual averaging.

A review of dual averaging (Nesterov)

Dual averaging guarantee

$$\sum_{\tau=1}^{t} \left\langle \ell^{(\tau)}, x^{(\tau)} - x \right\rangle \leq \frac{1}{\eta_t}\psi(x) + \frac{M^2}{2\ell_\psi} \sum_{\tau=1}^{t} \eta_\tau$$

(here M is a bound on $\|\ell^{(t)}\|_*$)

Consequence

Convex optimization	Online learning
$f\left(\frac{1}{t}\sum_{\tau=1}^{t} x^{(\tau)}\right) - f^\star \to 0$	$\sup_{x\in\Delta^N} R^{(t)}(x) = o(t)$

Idea

1. Take $\mathcal{X} = \Delta(S)$.
2. Apply dual averaging.

More precisely...

Assume S is compact subset of $\mathbb{R}^n$.

Set of Lebesgue continuous distributions over S

$$\mathcal{X} = \Delta(S) = \{x \in L^2(S) : x \geq 0 \text{ a.e. and } \int_S x(s)\lambda(ds) = 1\}$$

- $H = (L^2(S), \langle \cdot, \cdot \rangle)$ is Hilbert
- $H^* = H$, and since S is compact, $C(S) \subset L^2(S)$
- $\mathcal{X}$ is convex, closed

Even though S is not convex, $\Delta(S)$ is.

Problem solved?

More precisely...

Assume S is compact subset of $\mathbb{R}^n$.

Set of Lebesgue continuous distributions over S

$$\mathcal{X} = \Delta(S) = \{x \in L^2(S) : x \geq 0 \text{ a.e. and } \int_S x(s)\lambda(ds) = 1\}$$

- $H = (L^2(S), \langle \cdot, \cdot \rangle)$ is Hilbert
- $H^* = H$, and since S is compact, $C(S) \subset L^2(S)$
- $\mathcal{X}$ is convex, closed

Even though S is not convex, $\Delta(S)$ is.

Problem solved?

Well...

- $\Delta(S)$ is infinite dimensional. How do you solve

$$\min_{x\in\Delta(S)} \left\langle \sum_{\tau=1}^{t} \ell^{(\tau)}, x \right\rangle + \frac{1}{\eta_t}\psi(x)$$

- Can we obtain a meaningful regret bound?

$$R^{(t)}(x) \leq \frac{1}{\eta_t}\psi(x) + \frac{M}{2\ell_\psi}\sum_{\tau=1}^{t}\eta_\tau$$

E.g. the negative entropy $\psi(x) = \int_S x(s)\ln x(s)\lambda(ds)$ is unbounded (take $x = \frac{1}{\lambda(A)}1_A$, $A \subset S$, then $\psi(x) = -\ln\lambda(A)$)

Well...

- $\Delta(S)$ is infinite dimensional. How do you solve

$$\min_{x\in\Delta(S)} \left\langle \sum_{\tau=1}^{t} \ell^{(\tau)}, x \right\rangle + \frac{1}{\eta_t}\psi(x)$$

- Can we obtain a meaningful regret bound?

$$R^{(t)}(x) \leq \frac{1}{\eta_t}\psi(x) + \frac{M}{2\ell_\psi}\sum_{\tau=1}^{t}\eta_\tau$$

E.g. the negative entropy $\psi(x) = \int_S x(s)\ln x(s)\lambda(ds)$ is unbounded (take $x = \frac{1}{\lambda(A)}1_A$, $A \subset S$, then $\psi(x) = -\ln\lambda(A)$)

We are in luck!

For a class of regularizers ψ, induced by ω potentials,

- Can solve the dual averaging iteration.
- Have sufficient conditions for sublinear regret (when S has reasonable geometry)

Outline

ω potentials

Csiszár divergence induced by ω potential

Csiszár divergence, defined on $\mathcal{X}$

$$\psi_{f_\phi}(x) = \int_S f_\phi(x(s))\lambda(ds)$$

where $f_\phi(x) = \int_1^x \phi^{-1}(u)du$, and $\phi : (-\infty, a) \to (\omega, \infty)$ C^1 diffeomorphism such that $\lim_{u\to-\infty} \phi(u) = \omega$, $\lim_{u\to a} \phi(u) = +\infty$. ($f_\phi$ is convex and $f_\phi(1) = 0$.)

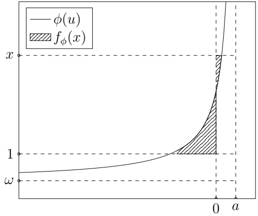


Figure: Illustration of an ω-potential.

Dual averaging iteration

$$x^{(t+1)} = \arg\min_{x \in \Delta(S)} \left\langle \sum_{\tau=1}^{t} \ell^{(\tau)}, x \right\rangle + \frac{1}{\eta_t}\psi(x)$$

Solution

$$x^{(t+1)}(s) = \phi(-\eta_{t+1}(L^{(t)}(s) + \nu^\star))_+$$

where $\nu^\star$ satisfies $\|x^{(t+1)}\|_1 = 1$.

Observing that $\|x^{(t+1)}\|_1$ is a monotone function of $\nu^\star$, this can be solved using a bisection method.

Dual averaging iteration

$$x^{(t+1)} = \underset{x \in \Delta(S)}{\arg\min} \left\langle \sum_{\tau=1}^{t} \ell^{(\tau)}, x \right\rangle + \frac{1}{\eta_t} \psi(x)$$

Solution

$$x^{(t+1)}(s) = \phi(-\eta_{t+1}(L^{(t)}(s) + \nu^\star))_+$$

where $\nu^\star$ satisfies $\|x^{(t+1)}\|_1 = 1$.

Observing that $\|x^{(t+1)}\|_1$ is a monotone function of $\nu^\star$, this can be solved using a bisection method.

Dual averaging iteration

$$x^{(t+1)} = \arg\min_{x \in \Delta(S)} \left\langle \sum_{\tau=1}^{t} \ell^{(\tau)}, x \right\rangle + \frac{1}{\eta_t}\psi(x)$$

Solution

$$x^{(t+1)}(s) = \phi(-\eta_{t+1}(L^{(t)}(s) + \nu^\star))_+$$

where $\nu^\star$ satisfies $\|x^{(t+1)}\|_1 = 1$.

Observing that $\|x^{(t+1)}\|_1$ is a monotone function of $\nu^\star$, this can be solved using a bisection method.

Example

L^2 projection:

- $\phi(u) = u$
- $\psi_{f_\phi}(x) = \frac{\|x\|_2^2 - 1}{2}$

Generalized entropy projection:

- $\phi(u) = e^{u+1} - \epsilon$
- $\psi_{f_\phi}(x) = -H(x+\epsilon) + H(1+\epsilon)$

Example

L^2 projection:

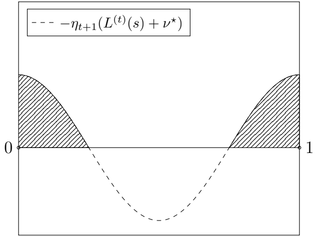
- $\phi(u) = u$
- $\psi_{f_\phi}(x) = \frac{\|x\|_2^2 - 1}{2}$

Generalized entropy projection:

- $\phi(u) = e^{u+1} - \epsilon$
- $\psi_{f_\phi}(x) = -H(x+\epsilon) + H(1+\epsilon)$

Dual averaging iteration

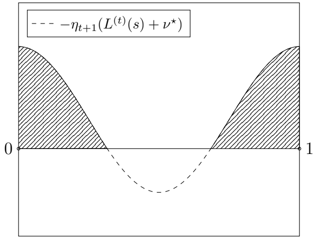
Example: L^2 projection



$$x^{(t+1)}(s) = \phi(-\eta_{t+1}(L^{(t)}(s) + \nu^\star))_+$$

Dual averaging iteration

Example: L^2 projection



$$x^{(t+1)}(s) = \phi(-\eta_{t+1}(L^{(t)}(s)+\nu^\star))_+$$

Regret bound

On which sets S can we learn?

Fat sets

S is v-uniformly fat if for all $s \in S$, $\exists K \subset S$ convex, with $s \in K$ and $\lambda(K) \geq v$

Intuitively, there is enough mass around each point of S.

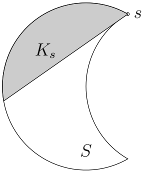


Figure: Illustration of uniform fatness

Regret bound

On which sets S can we learn?

Fat sets

S is v-uniformly fat if for all $s \in S$, $\exists K \subset S$ convex, with $s \in K$ and $\lambda(K) \geq v$

Intuitively, there is enough mass around each point of S.

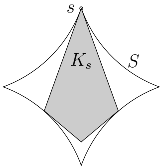


Figure: Illustration of uniform fatness

Regret bound

On which sets S can we learn?

Fat sets

S is v-uniformly fat if for all $s \in S$, $\exists K \subset S$ convex, with $s \in K$ and $\lambda(K) \geq v$

Intuitively, there is enough mass around each point of S.

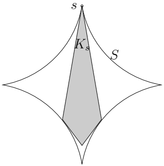


Figure: Illustration of uniform fatness

FORCES FOUNDATIONS OF RESILIENT CYBER-PHYSICAL SYSTEMS

Regret bound

On which sets S can we learn?

Fat sets

S is v-uniformly fat if for all $s \in S$, $\exists K \subset S$ convex, with $s \in K$ and $\lambda(K) \geq v$

Intuitively, there is enough mass around each point of S.

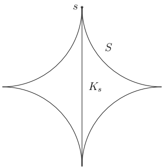


Figure: Illustration of uniform fatness

Regret bound

Regret rate

Suppose that S is v-uniformly fat, and that $\exists \epsilon > 0$ such that

$$f_\phi(x) = \mathcal{O}(x^{1+\epsilon}) \text{ as } x \to \infty.$$

Then DA with learning rates $\eta_t = \theta t^{-\alpha}$ satisfies

$$\frac{R^{(t)}}{t} = \mathcal{O}\left(t^{-\alpha} + t^{-\frac{1-\alpha}{1+n\epsilon}}\right)$$

FORCES
FOUNDATIONS OF RESILIENT CYBER-PHYSICAL SYSTEMS

Summary

For the family of Csiszár divergences

- Can compute the solution
- Regret bound
- (Also: sufficient conditions for strong convexity)

Outline

Numerical example

- $\ell^{(t)}$ are quadratics
- Hedge algorithm (ψ is the negative entropy)
- On the set S:

Results:

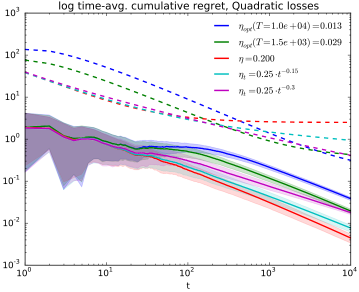


Figure: Mean time-average cumulative regret (solid), 10% and 90% quantiles (shaded regions) and worst-case bounds (dashed).

A second example

- $\ell^{(t)}$ are quadratics
- Dual averaging with a p-norm potential

(Loading Video...)

Figure: Evolution of the probability density $x^{(t)}$

Conclusion

We can learn on a continuum (when S has reasonable geometry).

Extensions and open questions

- Lower bounds on the regret.
- When can we sample efficiently? Depends on S and the family of loss functions.
- Extend to the bandit case: instead of observing the full loss function $\ell^{(t)}$, only observe $\ell^{(t)}(s^{(t)})$, where $s^{(t)}$ is sampled $\sim x^{(t)}$.

Conclusion

We can learn on a continuum (when S has reasonable geometry).

Extensions and open questions

- Lower bounds on the regret.
- When can we sample efficiently? Depends on S and the family of loss functions.
- Extend to the bandit case: instead of observing the full loss function $\ell^{(t)}$, only observe $\ell^{(t)}(s^{(t)})$, where $s^{(t)}$ is sampled $\sim x^{(t)}$.

Thank you.