The Hedge Algorithm on a Continuum

ICML 2015 - Lille, France

Walid Krichene Maximilian Balandat Claire Tomlin Alexandre Bayen

Outline

Outline

Online Learning over a finite set

A decision maker faces a sequential problem:

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

1: **for** $t \in \mathbb{N}$ **do**
2: Decision maker chooses distribution $x^{(t)}$ over $\{1, \dots N\}$.
3: A loss vector $\ell^{(t)} \in [0, M]^N$ is revealed.
4: The decision maker incurs expected loss $\sum_{n=1}^{N} \ell_n^{(t)} x_n^{(t)} = \left\langle x^{(t)}, \ell^{(t)} \right\rangle$
5: **end for**

Applications

Applications

- Convergence of player dynamics in games (Blackwell [3], Hannan[8]) $\{1, \ldots, N\}$ is the set of actions.

---

[3]David Blackwell. An analog of the minimax theorem for vector payoffs. *Pacific Journal of Mathematics*, 6(1):1–8, 1956

[8]James Hannan. Approximation to Bayes risk in repeated plays. *Contributions to the Theory of Games*, 3:97–139, 1957

[7]Thomas M. Cover. Universal portfolios. *Mathematical Finance*, 1(1):1–29, 1991

[4]Avrim Blum and Adam Kalai. Universal portfolios with and without transaction costs. *Machine Learning*, 35(3):193–205, 1999

Applications

Applications

- Convergence of player dynamics in games (Blackwell [3], Hannan[8])
  $\{1, \ldots, N\}$ is the set of actions.
- Machine Learning
  $\{1, \ldots, N\}$ is the training set.

[3]David Blackwell. An analog of the minimax theorem for vector payoffs. *Pacific Journal of Mathematics*, 6(1):1–8, 1956

[8]James Hannan. Approximation to Bayes risk in repeated plays. *Contributions to the Theory of Games*, 3:97–139, 1957

[7]Thomas M. Cover. Universal portfolios. *Mathematical Finance*, 1(1):1–29, 1991

[4]Avrim Blum and Adam Kalai. Universal portfolios with and without transaction costs. *Machine Learning*, 35(3):193–205, 1999

Applications

Applications

- Convergence of player dynamics in games (Blackwell [3], Hannan[8])
  $\{1, \dots, N\}$ is the set of actions.
- Machine Learning
  $\{1, \dots, N\}$ is the training set.
- "Model-free" portfolio optimization (Cover [7], Blum [4])
  $\{1, \dots, N\}$ is the set of stocks.
- Many others

---

[3]David Blackwell. An analog of the minimax theorem for vector payoffs. *Pacific Journal of Mathematics*, 6(1):1–8, 1956

[8]James Hannan. Approximation to Bayes risk in repeated plays. *Contributions to the Theory of Games*, 3:97–139, 1957

[7]Thomas M. Cover. Universal portfolios. *Mathematical Finance*, 1(1):1–29, 1991

[4]Avrim Blum and Adam Kalai. Universal portfolios with and without transaction costs. *Machine Learning*, 35(3):193–205, 1999

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 $\ell^{(t)} : S \to [0, M]$ is revealed.

The decision maker incurs expected loss

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

end for

Learning on a continuum

"What if the action set is infinite?"

Problem 2 Online decision problem on S.

1: **for** $t \in \mathbb{N}$ **do**
2: Decision maker chooses distribution $x^{(t)}$ over S.
3: A loss function $\ell^{(t)} : S \to [0, M]$ is revealed.
4: The decision maker incurs expected loss

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

5: **end for**

Regret

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

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

Results

Variant of this problem: Online optimization on convex sets.

Assumptions on $\ell^{(t)}$	convex	α-exp-concave	uniformly L-Lipschitz
Assumptions on S	convex	convex	v-uniformly fat
Method	Gradient (Zinkevich [14])	Hedge (Hazan et al. [9])	Hedge (This talk)
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.

[14]Martin Zinkevich. Online convex programming and generalized infinitesimal gradient ascent.
In *ICML*, pages 928–936, 2003

[9]Elad Hazan, Amit Agarwal, and Satyen Kale. Logarithmic regret algorithms for online convex optimization.
Machine Learning, 69(2-3):169–192, 2007

Outline

1. The Problem
2. Hedge on a Continuum
3. Numerical Examples

Hedge on a finite set

A.k.a.

- Weighted majority algorithm [11]
- Exponentiated gradient [10]
- Entropic descent [2]
- Multiplicative weights [1]
- Exponentially weighted forecaster [6]

---

[11]Nick Littlestone and Manfred K Warmuth. The weighted majority algorithm. In *Foundations of Computer Science, 1989., 30th Annual Symposium on*, pages 256–261. IEEE, 1989

[10]Jyrki Kivinen and Manfred K. Warmuth. Exponentiated gradient versus gradient descent for linear predictors.
Information and Computation, 132(1):1 – 63, 1997

[2]Amir Beck and Marc Teboulle. Mirror descent and nonlinear projected subgradient methods for convex optimization.
Oper. Res. Lett., 31(3):167–175, May 2003

[1]Sanjeev Arora, Elad Hazan, and Satyen Kale. The multiplicative weights update method: a meta-algorithm and applications.
Theory of Computing, 8(1):121–164, 2012

[6]Nicolò Cesa-Bianchi and Gábor Lugosi. *Prediction, learning, and games.*
Cambridge University Press, 2006

Hedge on a finite set

Hedge algorithm with learning rates (η_t).

1: **for** $t \in \mathbb{N}$ **do**
2: Play $x^{(t)}$
3: Reveal $\ell^{(t)} \in [0, M]^N$, call $L^{(t)} = \sum_{\tau=1}^{t} \ell^{(\tau)}$
4: Update

$$x_n^{(t+1)} \propto x_n^{(0)} e^{-\eta_{t+1} L_n^{(t)}}$$

5: **end for**

One interpretation: instance of the dual averaging method [13]

$$x^{(t+1)} \in \arg\min_{x \in \Delta^N} \left\langle L^{(t)}, x \right\rangle + \frac{1}{\eta_{t+1}} \psi(x)$$

with $\psi(x) = \sum_{n=1}^{N} x_n \ln x_n$.

[13]Yurii Nesterov. Primal-dual subgradient methods for convex problems. *Mathematical Programming*, 120(1):221–259, 2009

Hedge on a finite set

Basic Regret Bound

For all $x \in \Delta^N$,

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

Take $\eta_t = t^{-\frac{1}{2}}$, then $\sum_1^t \eta_\tau = O(\sqrt{t})$ and $\frac{1}{t} = O(\sqrt{t})$
It suffices to bound ψ on Δ^N.
When $\psi(x) = \sum_i x_i \ln x_i$, $\psi(x) \leq \ln N$ on Δ^N. So

$$\sup_{x \in \Delta^N} R^{(T)}(x) \leq \frac{M^2}{2} \sum_{\tau=1}^{t} \eta_{\tau+1} + \frac{\ln N}{\eta_{t+1}}$$

Hedge on a continuum

Hedge on S with learning rates (η_t).

1: **for** $t \in \mathbb{N}$ **do**
2: Play $\sim x^{(t)}$
3: Reveal $\ell^{(t)} : S \to [0, M]$
4: Update

$$x^{(t+1)}(s) \propto x^{(0)}(s) e^{-\eta_{t+1} L^{(t)}(s)}$$

5: **end for**

Hedge on a continuum

Hedge on S with learning rates (η_t).

1: **for** $t \in \mathbb{N}$ **do**
2: Play $\sim x^{(t)}$
3: Reveal $\ell^{(t)} : S \to [0, M]$
4: Update

$$x^{(t+1)}(s) \propto x^{(0)}(s) e^{-\eta_{t+1} L^{(t)}(s)}$$

5: **end for**

One interpretation: instance of the dual averaging method

$$x^{(t+1)} \in \arg\min_{x \in \Delta(S)} \left\langle L^{(t)}, x \right\rangle + \frac{1}{\eta_{t+1}} \psi(x)$$

with

- Hilbert space $\mathcal{H} = L^2(S)$, $\langle \ell, x \rangle = \int_S \ell(s) x(s) \lambda(ds)$
- $\Delta(S) = \{x \in L^2(S) : x \geq 0, \ \|x\|_1 = 1\}$
- $\psi(x) = \int_S x(s) \ln x(s) \lambda(ds)$

Hedge on a continuum

Basic Regret Bound

For all $x \in \Delta(S)$,

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

But ψ is unbounded on $\Delta(S)$.

Hedge on a continuum

Basic Regret Bound

For all $x \in \Delta(S)$,

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

But ψ is unbounded on $\Delta(S)$.
Take $x = \frac{1}{\lambda(A)} 1_A$ for some $A \subset S$. Then

$$\psi(x) = \int_S x(s) \ln x(s) \lambda(ds) = \ln \frac{1}{\lambda(A)}$$

can be arbitrarily large for arbitrarily small A.

Working around unbounded regularizers

Idea:

- Call $s_t^\star \in \arg\min_{s \in S} L^{(t)}(s)$ (L supposed continuous).
- Let $\mathcal{B}_t$ be a set of distributions supported near $s_t^\star$
- For any $y \in \mathcal{B}_t$,

$$
\begin{aligned}
R^{(t)}(x) &= \sum_{\tau=1}^{t} \left\langle \ell^{(\tau)}, x^{(\tau)} - x \right\rangle \\
&\leq \sum_{\tau=1}^{t} \left\langle \ell^{(\tau)}, x^{(\tau)} - \delta_{s_t^\star} \right\rangle \\
&= \sum_{\tau=1}^{t} \left\langle \ell^{(\tau)}, x^{(\tau)} - y \right\rangle + \sum_{\tau=1}^{t} \left\langle \ell^{(\tau)}, y - \delta_{s_t^\star} \right\rangle \\
&= R^{(t)}(y) + \left\langle L^{(t)}, y - \delta_{s_t^\star} \right\rangle
\end{aligned}
$$

Revised regret bound

$$
\sup_{x \in \Delta(S)} R^{(t)}(x) \leq R^{(t)}(y_0) + \sup_{y \in \mathcal{B}_t} \left\langle L^{(t)}, y - \delta_{s_t^\star} \right\rangle
$$

Working around unbounded regularizers

Take $y_0 = \frac{1}{\lambda(A_t)} 1_{A_t}$

$$\sup_{x \in \Delta(S)} R^{(t)}(x) \leq R^{(t)}(y_0) + \sup_{y \in \mathcal{B}_t} \left\langle L^{(t)}, y - \delta_{s_t^\star} \right\rangle$$

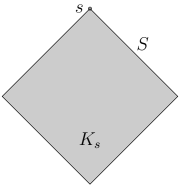
$$R^{(t)}(y_0) \leq \frac{M^2}{2} \sum_{\tau=1}^{t} \eta_{\tau+1} + \frac{\psi(y_0)}{\eta_{t+1}} \leq \frac{M^2}{2} \sum_{\tau=1}^{t} \eta_{\tau+1} + \frac{1}{\eta_{t+1}} \ln \frac{1}{\lambda(A_t)}$$

$$\left\langle L^{(t)}, y - \delta_{s_t^\star} \right\rangle = \int_{A_t} y(s)(L^{(t)}(s) - L^{(t)}(s_t^\star))\lambda(ds) \leq Ltd(A_t)$$

Uniformly fat sets

Uniform fatness

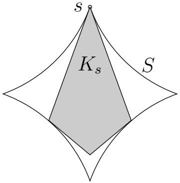
S is v-uniformly fat (w.r.t. the measure λ) if
$\forall s \in S$, $\exists$ convex $K_s \subset S$, with $s \in K_s$ and $\lambda(K_s) \geq v$.



Uniformly fat sets

Uniform fatness

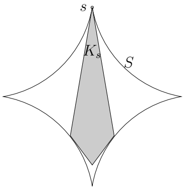
S is v-uniformly fat (w.r.t. the measure λ) if
$\forall s \in S$, $\exists$ convex $K_s \subset S$, with $s \in K_s$ and $\lambda(K_s) \geq v$.



Uniformly fat sets

Uniform fatness

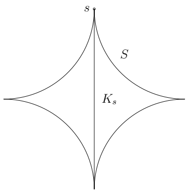
S is v-uniformly fat (w.r.t. the measure λ) if
$\forall s \in S$, $\exists$ convex $K_s \subset S$, with $s \in K_s$ and $\lambda(K_s) \geq v$.



Uniformly fat sets

Uniform fatness

S is v-uniformly fat (w.r.t. the measure λ) if
$\forall s \in S$, $\exists$ convex $K_s \subset S$, with $s \in K_s$ and $\lambda(K_s) \geq v$.



Regret bound on uniformly fat sets

$$\sup_{x\in\Delta(S)} R^{(t)}(x) \leq \frac{M^2}{2}\sum_{\tau=1}^{t}\eta_{\tau+1} + \frac{1}{\eta_{t+1}}\ln\frac{1}{\lambda(A_t)} + Ltd(A_t)$$

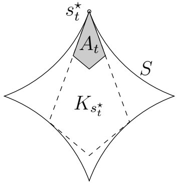


Figure: $A_t = s_t^\star + d_t(K_{s_t^\star} - s_t^\star)$. Then $\lambda(A_t) \geq d_t^n v$ and $d(A_t) \leq d_t d(S)$.

$$\sup_{x\in\Delta(S)} R^{(t)}(x) \leq \frac{M^2}{2}\sum_{\tau=1}^{t}\eta_{\tau+1} + \frac{1}{\eta_{t+1}}\ln\frac{1}{vd_t^n} + Ltd_t d(S)$$

Regret bound on uniformly fat sets

Final bound

$$\sup_{x\in\Delta(S)} R^{(t)}(x) \leq \frac{M^2}{2}\sum_{\tau=1}^{t}\eta_{\tau+1} + \frac{\ln\frac{1}{v}}{\eta_{t+1}} + \frac{n\ln t}{\eta_{t+1}} + Ld(S)$$

Can optimize over η_t to get

$$\sup_{x\in\Delta(S)} R^{(t)}(x) \leq Ld(S) + M\sqrt{\frac{nt\ln t}{2}}$$

Hedge Vs. learning on a cover

- Given a horizon T and a cover $\mathcal{A}_T$ with $d(A) \leq d_T d(S)$ for all $A \in \mathcal{A}_T$.
- Run discrete Hedge on elements of the cover.

Hedge Vs. learning on a cover

- Given a horizon T and a cover $\mathcal{A}_T$ with $d(A) \leq d_T d(S)$ for all $A \in \mathcal{A}_T$.
- Run discrete Hedge on elements of the cover.
- Then

$$R^{(T)}(x) \leq \underbrace{\frac{M^2 T\eta}{8} + \frac{\ln|\mathcal{A}_T|}{\eta}}_{\text{Discrete Hedge}} + \underbrace{Ld(S)d_T}_{\text{Additional regret}}$$

- With $|\mathcal{A}_T| \approx \frac{1}{d_T^n}$,

$$R^{(T)}(x) \leq \frac{M^2 T\eta}{8} + \frac{\ln \frac{1}{d_T^n}}{\eta} + LD(S)d_T$$

Hedge Vs. learning on a cover

- Given a horizon T and a cover $\mathcal{A}_T$ with $d(A) \leq d_T d(S)$ for all $A \in \mathcal{A}_T$.
- Run discrete Hedge on elements of the cover.
- Then

$$R^{(T)}(x) \leq \underbrace{\frac{M^2 T\eta}{8} + \frac{\ln|\mathcal{A}_T|}{\eta}}_{\text{Discrete Hedge}} + \underbrace{Ld(S)d_T}_{\text{Additional regret}}$$

- With $|\mathcal{A}_T| \approx \frac{1}{d_T^n}$,

$$R^{(T)}(x) \leq \frac{M^2 T\eta}{8} + \frac{\ln \frac{1}{d_T^n}}{\eta} + LD(S)d_T$$

- Have to explicitly compute a (hierarchical) cover.

Beyond Hedge

Dual averaging with learning rates (η_t), strongly convex regularizer ψ

1: **for** $t \in \mathbb{N}$ **do**
2: Play $x^{(t)}$
3: Discover $\ell^{(t)} \in \mathcal{H}^*$
4: Update

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

5: **end for**

Beyond Hedge

Dual averaging with learning rates (η_t), strongly convex regularizer ψ

for $t \in \mathbb{N}$ **do**
  Play $x^{(t)}$
  Discover $\ell^{(t)} \in \mathcal{H}^*$
  Update

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

end for

- $\mathcal{H}$ is infinite dimensional. Can we solve

$$\min_{x \in \Delta(S)} \left\langle L^{(t)}, x \right\rangle + \frac{1}{\eta_{t+1}} \psi(x)$$

Beyond Hedge

Dual averaging with learning rates (η_t), strongly convex regularizer ψ

1: **for** $t \in \mathbb{N}$ **do**
2: Play $x^{(t)}$
3: Discover $\ell^{(t)} \in \mathcal{H}^*$
4: Update

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

5: **end for**

- $\mathcal{H}$ is infinite dimensional. Can we solve

$$\min_{x \in \Delta(S)} \left\langle L^{(t)}, x \right\rangle + \frac{1}{\eta_{t+1}} \psi(x)$$

- Can we obtain a sublinear regret bound?

$$\sup_{x \in \Delta(S)} R^{(t)}(x) \leq \frac{M^2}{2} \sum_{\tau=1}^{t} \eta_{\tau+1} + \frac{1}{\eta_{t+1}} \psi(y) + Ltd(A_t)$$

Outline

Numerical Example

Conclusion

Summary

- Can learn on a continuum, when losses are Lipschitz and S has reasonable geometry.
- Similar guarantee to learning on a cover, but do not need to maintain a cover.
- Can generalize to the dual averaging method.

Extensions and open questions

- Bandit formulation, e.g. [5].
- Regret lower bound.
- When is it easy to sample from the Hedge distribution?
  Partial answer provided by [12].

---

[5]Sébastien Bubeck, Rémi Munos, Gilles Stoltz, and Csaba Szepesvari. X-armed bandits. *Journal of Machine Learning Research (JMLR)*, 12(12):1587–1627, 2011

[12]Chris J Maddison, Daniel Tarlow, and Tom Minka. A$*$ sampling. In Z. Ghahramani, M. Welling, C. Cortes, N.D. Lawrence, and K.Q. Weinberger, editors, *Advances in Neural Information Processing Systems 27*, pages 3086–3094. Curran Associates, Inc., 2014

Thank you

Thank you.

walid@eecs.berkeley.edu
balandat@eecs.berkeley.edu

References I

[1] Sanjeev Arora, Elad Hazan, and Satyen Kale. The multiplicative weights update method: a meta-algorithm and applications. *Theory of Computing*, 8(1):121–164, 2012.

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

[3] David Blackwell. An analog of the minimax theorem for vector payoffs. *Pacific Journal of Mathematics*, 6(1):1–8, 1956.

[4] Avrim Blum and Adam Kalai. Universal portfolios with and without transaction costs. *Machine Learning*, 35(3):193–205, 1999.

[5] Sébastien Bubeck, Rémi Munos, Gilles Stoltz, and Csaba Szepesvari. X-armed bandits. *Journal of Machine Learning Research (JMLR)*, 12(12): 1587–1627, 2011.

[6] Nicolò Cesa-Bianchi and Gábor Lugosi. *Prediction, learning, and games*. Cambridge University Press, 2006.

[7] Thomas M. Cover. Universal portfolios. *Mathematical Finance*, 1(1):1–29, 1991.

References II

[8] James Hannan. Approximation to Bayes risk in repeated plays. *Contributions to the Theory of Games*, 3:97–139, 1957.

[9] Elad Hazan, Amit Agarwal, and Satyen Kale. Logarithmic regret algorithms for online convex optimization. *Machine Learning*, 69(2-3): 169–192, 2007.

[10] Jyrki Kivinen and Manfred K. Warmuth. Exponentiated gradient versus gradient descent for linear predictors. *Information and Computation*, 132 (1):1 – 63, 1997.

[11] Nick Littlestone and Manfred K Warmuth. The weighted majority algorithm. In *Foundations of Computer Science, 1989., 30th Annual Symposium on*, pages 256–261. IEEE, 1989.

[12] Chris J Maddison, Daniel Tarlow, and Tom Minka. A∗ sampling. In Z. Ghahramani, M. Welling, C. Cortes, N.D. Lawrence, and K.Q. Weinberger, editors, *Advances in Neural Information Processing Systems 27*, pages 3086–3094. Curran Associates, Inc., 2014.

[13] Yurii Nesterov. Primal-dual subgradient methods for convex problems. *Mathematical Programming*, 120(1):221–259, 2009.

[14] Martin Zinkevich. Online convex programming and generalized infinitesimal gradient ascent. In *ICML*, pages 928–936, 2003.