CSIE5410 Optimization algorithms

Lecture 9: Online learning & optimization

Yen-Huan Li (yenhuan.li@csie.ntu.edu.tw)

13.12.2018

Department of Computer Science and Information Engineering
National Taiwan University

Abstract ver. 1

In Lecture 1, we have seen that the standard theory of machine learning provides a framework to address i.i.d. data.

What if we would like to do *sequential decision making*?

Online learning provides an approach to sequential decision making.

Abstract ver. 2

In Lecture 1, we have seen that a standard approach to machine learning is empirical risk minimization (ERM).

However, the per-iteration computational complexity of the ERM scales at least linearly with the data size.

Online optimization provides an approach to circumvent this computational bottleneck.

Abstract ver. 3

Consider two players playing a game. We expect that some *equilibrium* would emerge.

Why and how does an equilibrium emerge?

The theory of *learning in games* provides an approach to addressing the question.

Recommended reading

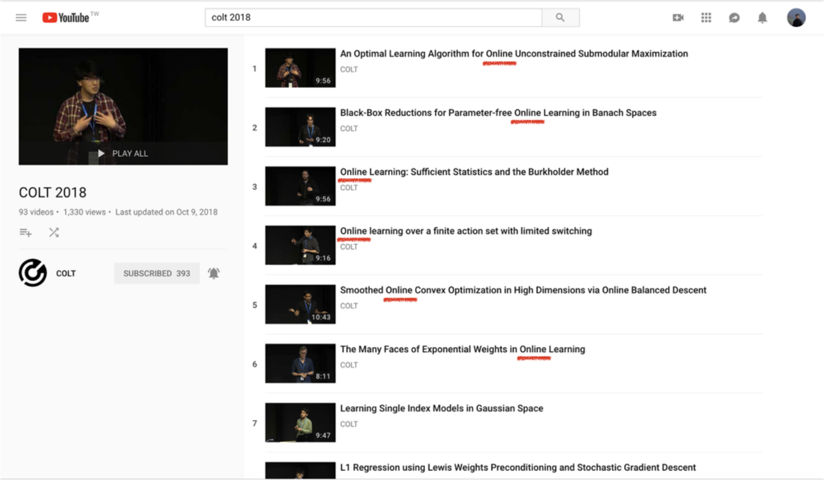
- Y. Freund and R. E. Schapire. 1997. A decision-theoretic generalization of on-line learning and an application to boosting.
- *V. Vovk. 1998. A game of prediction with expert advice.
- N. Cesa-Bianchi *et al.* 2004. On the generalization ability of on-line learning algorithms.
- S. Arora *et al.* 2012. The multiplicative weights update method: A meta-algorithm and applications.

Table of contents

Prelude

Conference on Learning Theory (COLT)

Popularity of online learning/optimization



Last COLT paper involving Taiwanese researchers

PMLR Proceedings of Machine Learning Research

Volume 23 All Volumes JMLR MLOSS FAQ Submission Format RSS

Online Optimization with Gradual Variations

[edit]

Chao-Kai Chiang, Tianbao Yang, Chia-Jung Lee, Mehrdad Mahdavi, Chi-Jen Lu, Rong Jin, Shenghuo Zhu ;
Proceedings of the 25th Annual Conference on Learning Theory, PMLR 23:6.1-6.20, 2012.

Abstract

We study the online convex optimization problem, in which an online algorithm has to make repeated decisions with convex loss functions and hopes to achieve a small regret. We consider a natural restriction of this problem in which the loss functions have a small deviation, measured by the sum of the distances between every two consecutive loss functions, according to some distance metrics. We show that for the linear and general smooth convex loss functions, an online algorithm modified from the gradient descend algorithm can achieve a regret which only scales as the square root of the deviation. For the closely related problem of prediction with expert advice, we show that an online algorithm modified from the multiplicative update algorithm can also achieve a similar regret bound for a different measure of deviation. Finally, for loss functions which are strictly convex, we show that an online algorithm modified from the online Newton step algorithm can achieve a regret which is only logarithmic in terms of the deviation, and as an application, we can also have such a logarithmic regret for the portfolio management problem.

Related Material

- Download PDF

Copy BibTeX | Download BibTeX

Copy Endnote | Download Endnote

Copy RIS | Download RIS

Copy APA | Download APA

This site last compiled Sun, 15 Jul 2018 13:29:10 +0000

Online mirror descent

Decision theoretic online learning (1/2)

A gambler, frustrated by persistent horse-racing losses and envious of his friends' winnings, decides to allow a group of his fellow gamblers to make bets on his behalf.

He decides he will wager a fixed sum of money in every race, but that he will apportion his money among his friends based on how well they are doing.

Y. Freund and R. E. Schapire. 1997. A decision-theoretic generalization of on-line learning and an application to boosting.

Decision theoretic online learning (2/2)

Certainly, if he knew psychically ahead of time which of his friends would win the most, he would naturally have that friend handle all his wagers.

Lacking such clairvoyance, however, he attempts to allocate each race's wager in such a way that his total winnings for the season will be reasonably close to what he would have won had he bet everything with the luckiest of his friends.

Y. Freund and R. E. Schapire. 1997. A decision-theoretic generalization of on-line learning and an application to boosting.

Formalization

Let $\mathcal{A}$ be a finite set of actions. Set the initial loss L_0 to be zero.
Set π_1 be a probability distribution on $\{ 1, \dots, m \}$.

For $t = 1, \dots, T$, the following happen in order.

1. EXPERT i announces their action $a_{i,t} \in \mathcal{A}$.
2. LEARNER *chooses an expert i_t randomly according to π_t.*
3. REALITY announces a loss $\ell_{i,t} \in [-1, 1]$ for EXPERT i.
4. $L_t = L_{t-1} + \mathsf{E}\left[\ell_{i_t,t}\right]$; LEARNER computes π_{t+1}.

Regret

Goal. To achieve a small *regret*:

$$R_T := L_T - \min_i \left\{ \sum_{t=1}^{T} \ell_{i,t} \;\middle|\; i = 1, \ldots, m \right\}.$$

Remark. Notice that the regret is a function of T, and the *best expert* can change with T.

Remakr. A regret is considered satisfactory, if it is sub-linear in T, i.e., $R_T = o(T)$.

Hedge algorithm

$\textsc{Learner}$'s strategy is defined by the probability distributions $\pi_1, \ldots, \pi_T$.

Algorithm Hedge algorithm

1: Let $w_1^{(i)} = 1, \quad \pi_1^{(i)} = \frac{w_1^{(i)}}{\sum_{i=1}^m w_1^{(i)}} \quad \forall i$. Set $\eta > 0$.
2: **for** $t = 1, \ldots, T$ **do**
3: $\quad w_{t+1}^{(i)} \leftarrow w_t^{(i)} \mathrm{e}^{-\eta \ell_{i,t}}$ for every $i = 1, \ldots, m$.
4: $\quad \pi_{t+1}^{(i)} \leftarrow \frac{w_{t+1}^{(i)}}{\sum_{i=1}^m w_{t+1}^{(i)}}$ for every $i = 1, \ldots, m$.
5: **end for**

Question. What algorithm do you think of?

Y. Freund and R. E. Schapire. 1997. A decision-theoretic generalization of on-line learning and an application to boosting.

Reformulation of the protocol

Denote by Δ the probability simplex in $\mathbb{R}^m$.

For $t = 1, \ldots, T$, the following happen in order.

1. $\textsc{Learner}$ announces $x_t \in \Delta$.
2. $\textsc{Reality}$ announces $\ell_t \in [-1, 1]^m$.

Definition. The regret is *equivalently (why?)* defined as

$$R_T \coloneqq \sum_{t=1}^{T} \langle \ell_t, x_t \rangle - \min_x \left\{ \sum_{t=1}^{T} \langle \ell_t, x \rangle \;\middle|\; x \in \Delta \right\}.$$

Generalization of the formulation

Let $\mathcal{X}$ be a bounded closed convex set in $\mathbb{R}^p$.

For $t = 1, \ldots, T$, the following happen in order.

1. $\textsc{Learner}$ announces $x_t \in \mathcal{X}$.
2. $\textsc{Reality}$ announces a *convex* function $f_t : \mathcal{X} \to \mathbb{R}$.

Definition. The regret is defined as

$$R_T := \sum_{t=1}^{T} f_t(x_t) - \min_{x} \left\{ \sum_{t=1}^{T} f_t(x) \;\middle|\; x \in \mathcal{X} \right\}.$$

M. Zinkevich. 2003. Online convex programming and generalized infinitesimal gradient ascent.

Back to DTOL

Let $\mathcal{X} = \Delta$.

For $t = 1, \dots, T$, the following happen in order.

1. LEARNER announces $x_t \in \mathcal{X}$.
2. REALITY announces $f_t := \langle \ell_t, \cdot \rangle$.

Remark. Then the condition $\ell_t \in [-1, +1]^{m \text{ (or } p)}$ translates to

$$\|\nabla f_t(x)\|_\infty \le 1, \quad \forall x \in \mathcal{X}.$$

This is the standard condition for entropic mirror descent.

Online mirror descent

Let h be a differentiable function, 1-strongly convex on $\mathcal{X}$ with respect to a norm $\|\cdot\|$. Denote the corresponding Bregman divergence by D_h.

Algorithm Online mirror descent

1: Let $x_1 \in \mathcal{X}$. Set $\eta > 0$.
2: **for** $t = 1, \ldots, T$ **do**
3: $\quad x_{t+1} \leftarrow \arg\min_x \{\, \eta \langle \nabla f_t(x_t), x - x_t \rangle + D_h(x, x_t) \mid x \in \mathcal{X} \,\}$
4: **end for**

Remark. Here, the notation $\nabla f_t(x_t)$ denotes a subgradient of f_t at x_t.

Proof of the regret bound (1/3)

Theorem. Define $R := \max_x \left\{ \sqrt{D_h(x, x_1)} \;\middle|\; x \in \mathcal{X} \right\}$. Suppose that $\|\nabla f(x)\|_* \leq L$ for all $x \in \mathcal{X}$. Then setting $\eta = \frac{\sqrt{2}R}{L\sqrt{T}}$, we have

$$R_T = O(LR\sqrt{T}).$$

Proof. For any $x \in \mathcal{X}$, we write

$$\begin{aligned}
\eta\left(f_t(x_t) - f_t(x)\right) &\leq \eta \left\langle \nabla f_t(x_t), x_t - x \right\rangle \\
&\leq \left[D_h(x, x_t) - D_h(x, x_{t+1})\right] + \\
&\quad \left[\eta \left\langle \nabla f_t(x_t), x_t - x_{t+1} \right\rangle - D_h(x_{t+1}, x_t)\right],
\end{aligned}$$

where the last inequality follows from the Bregman proximal inequality.

Proof of the regret bound (2/3)

Proof continued. By the strong convexity of h, we obtain

$$\begin{aligned}
&\eta \langle \nabla f_t(x_t), x_t - x_{t+1} \rangle - D_h(x_{t+1}, x_t) \\
&\quad \leq \eta \|\nabla f_t(x_t)\|_* \|x_t - x_{t+1}\| - \frac{1}{2}\|x_t - x_{t+1}\|^2 \\
&\quad \leq \frac{\eta^2}{2}\|\nabla f_t(x_t)\|_*^2.
\end{aligned}$$

Combined with the inequality in the previous slide, we have

$$(f_t(x_t) - f_t(x)) \leq \frac{1}{\eta}\left[D_h(x, x_t) - D_h(x, x_{t+1})\right] + \frac{\eta}{2}\|\nabla f_t(x_t)\|_*^2.$$

Proof of the regret bound (3/3)

Proof continued. Summing over all t, we get

$$\begin{aligned}\sum_{t=1}^{T}(f_t(x_t) - f_t(x)) &\le \frac{1}{\eta}D_h(x, x_1) + \frac{\eta}{2}\sum_{t=1}^{T}\|\nabla f_t(x_t)\|_*^2 \\ &\le \frac{1}{\eta}R^2 + \frac{\eta}{2}TL^2.\end{aligned}$$

The upper bound is minimized when

$$\eta = \frac{\sqrt{2}R}{L\sqrt{T}}.$$

Back to DTOL

Recall that DTOL corresponds to the case where $\mathcal{X}$ is the probability simplex in $\mathbb{R}^m$ and $\|\nabla f_t(x)\|_\infty \leq 1$.

Choosing h as the negative Shannon entropy and x_1 as the uniform probability distribution, we have

$$D_h(x, x_1) \leq \log m, \quad \forall x \in \mathcal{X}.$$

Corollary. Choosing $\eta = \sqrt{\frac{2 \log m}{T}}$, the hedge algorithm achieves

$$R_T = O(\sqrt{T \log m}).$$

Strongly convex case

Assume in addition that the norm is the 2-norm, and every f_t is μ-strongly convex on $\mathcal{X}$.

Algorithm Online projected gradient descent

1: Let $x_1 \in \mathcal{X}$. Set $\eta > 0$.
2: **for** $t = 1, \ldots, T$ **do**
3: $\quad x_{t+1} \leftarrow \mathrm{proj}_{\mathcal{X}}\left(x_t - \eta_t \nabla f_t(x_t)\right)$
4: **end for**

Theorem. The algorithm achieves $R_T = O\left(\frac{L^2 \log T}{\mu}\right)$.

Proof. You have done the proof in Homework 2 ;)

E. Hazan *et al.* 2007. Logarithmic regret algorithms for online convex optimization.

Online-to-batch conversion

Online-to-batch conversion (1/3)

Recall that the standard formulation of machine learning asks one to solve the *risk minimization problem*

$$w^\star \in \arg\min_{w} \{ \, \mathsf{E}\, f(w; z) \mid w \in \mathcal{W} \, \} \,,$$

for some loss function f and given set $\mathcal{W}$ parametrizing the hypothesis class, where z is a random variable representing the data.

Remark. Typically, the random variable z is a pair (x, y) for random variables x and y.

S. Shalev-Shwartz *et al.* 2010. Learnability, stability and uniform convergence.

Online-to-batch conversion (2/3)

Assumption. We do not know the exact probability distribution of z, but we have access to the data, modeled as i.i.d. random variables $z_1, \dots, z_n$ following the probability distribution of z.

Consider the online convex optimization problem with $f_t := f(\cdot; z_t)$. View an online convex optimization algorithm as a *learning algorithm* that outputs hypotheses sequentially.

Question. What is the resulting risk performance?

Online-to-batch conversion (3/3)

Theorem. Suppose that $f(\cdot; z)$ is convex and takes values in $[0, 1]$ for all z, and $\mathcal{W}$ is convex and closed. Suppose there exists an online convex optimization algorithm that for any sequence $z_1, \ldots, z_n$, achieves

$$\sum_{i=1}^{n} f(w_i; z_i) - \sum_{i=1}^{n} f(w; z_i) \leq R_n, \quad \forall w \in \mathcal{W},$$

for some number $R_n > 0$. Then with probability at least $1 - \delta$ (with respect to the data), it holds that

$$\mathsf{E}\, f(\bar{w}_n; z) - \mathsf{E}\, f(w^\star) \leq \frac{R_n}{n} + \sqrt{\frac{8\log(1/\delta)}{n}},$$

where $\bar{w}_n := \frac{1}{n}(w_1 + \cdots + w_n)$.

Preliminary knowledge (1/2)

Definition. A *martingale* is a sequence of random variables $(\xi_i)_{i\in\mathbb{N}}$ satisfying $\mathsf{E}|\xi_i| < +\infty$ and

$$\mathsf{E}\left[\xi_{i+1}|\xi_1,\dots,\xi_i\right] = \xi_i, \quad \forall i \in \mathbb{N}.$$

Definition. Let $(\eta_i)_{i\in\mathbb{N}}$ be a sequence of random variables. A *martingale difference sequence* with respect to $(\eta_i)_{i\in\mathbb{N}}$ is a sequence of random variables $(\zeta_i)_{i\in\mathbb{N}}$ satisfying $\mathsf{E}|\zeta_i| < +\infty$ and

$$\mathsf{E}\left[\zeta_{i+1}|\eta_1,\dots,\eta_i\right] = 0, \quad \forall i \in \mathbb{N}.$$

Proposition. If $(\xi_i)_{i\in\mathbb{N}}$ is a martingale, then $(\zeta_i := \xi_i - \xi_{i-1})_{i\in\mathbb{N}}$ is a martingale difference sequence.

Preliminary knowledge (2/2)

Theorem. (Hoeffding-Azuma inequality) Let $(\zeta_i)_{i\in\mathbb{N}}$ be a martingale difference sequence with respect to $(\xi_i)_{i\in\mathbb{N}}$. Suppose that $\zeta_i \in [-c, c]$ for all i, for some $c > 0$. Then, for any $\tau > 0$,

$$\mathsf{P}\left(\frac{1}{n}\sum_{i=1}^{n}\zeta_i \geq \tau\right) \leq \mathrm{e}^{-\frac{n\tau^2}{2c^2}},$$

$$\mathsf{P}\left(\frac{1}{n}\sum_{i=1}^{n}\zeta_i \leq -\tau\right) \leq \mathrm{e}^{-\frac{n\tau^2}{2c^2}}.$$

W. Hoeffding. 1963. Probability inequalities for sums of bounded random variables.
K. Azuma. 1967. Weighted sums of certain dependent random variables.

Proof of the risk bound (1/2)

Proof. Define

$$\zeta_i := [f(w_i; z_i) - \mathsf{E}_z\, f(w_i; z)] - [f(w^\star; z_i) - \mathsf{E}_z\, f(w^\star; z)]$$

Then, $(\zeta_i)_{i\in\mathbb{N}}$ is a martingale difference sequence with respect to $(z_i)_{i\in\mathbb{N}}$, taking values in $[-2, 2]$. Furthermore, we have

$$\begin{aligned}
&\sum_{i=1}^{n} [\mathsf{E}_z\, f(w_i; z) - \mathsf{E}_z\, f(w^\star; z)] \\
&\qquad = \sum_{i=1}^{n} [f(w_i; z_i) - f(w^\star; z_i)] - \sum_{i=1}^{n} \zeta_i \\
&\qquad \le R_n - \sum_{i=1}^{n} \zeta_i.
\end{aligned}$$

Proof of the risk bound (2/2)

Proof continued. Notice that $\mathsf{E}_z\, f(\cdot; z)$ is convex. We write

$$\begin{aligned}
&\mathsf{E}_z\, f(\bar{w}_n; z) - \mathsf{E}_z\, f(w^\star; z) \\
&\quad \le \frac{1}{n} \sum_{i=1}^{n} \left[\mathsf{E}_z\, f(w_i; z) - \mathsf{E}_z\, f(w^\star; z)\right] \\
&\quad \le \frac{R_n}{n} - \frac{1}{n} \sum_{i=1}^{n} \zeta_i.
\end{aligned}$$

It remains to apply the Hoeffding-Azuma inequality and obtain

$$\mathsf{P}\left(\frac{1}{n} \sum_{i=1}^{n} \zeta_i \le -\sqrt{\frac{8 \log(1/\delta)}{n}} \right) \le \delta.$$

Online-to-batch for the expected risk

Corollary. Suppose that $f(\cdot; z)$ is convex and takes values in $[0, 1]$ for all z, and $\mathcal{W}$ is convex and closed. Suppose there exists an online convex optimization algorithm that for any sequence $z_1, \ldots, z_n$, achieves

$$\sum_{i=1}^{n} f(w_i; z_i) - \sum_{i=1}^{n} f(w; z_i) \leq R_n, \quad \forall w \in \mathcal{W},$$

for some number $R_n > 0$. Then it holds that

$$\mathsf{E}_{z_1,\ldots,z_n} [\mathsf{E}_z\, f(\bar{w}_n; z)] - \mathsf{E}_z\, f(w^\star; z) \leq \frac{R_n}{n}.$$

Proof. Recall that in the proof of the theorem, we have shown that

$$\sum_{i=1}^{n} [\mathsf{E}_z\, f(w_i; z) - \mathsf{E}_z\, f(w^\star; z)] \leq R_n - \sum_{i=1}^{n} \zeta_i.$$

Risk of online mirror descent (1/2)

Corollary. Assume in addition that $\mathcal{W}$ is bounded and $\|\nabla f(w; z)\|_* \leq L$ for all $w \in \mathcal{W}$ and z. Then, with probability at least 0.9, the *excess risk* achieved by online mirror descent is $O(n^{-1/2})$.

Remark. This is already statistically optimal in most of the cases.

Remark. On the other hand, one may use existing statistical results to check if a regret bound is reasonable.

Remark. Notice that the per-iteration computational complexity of online mirror descent is *independent of the data size*.

Risk of online mirror descent (2/2)

Corollary. Assume that the assumptions above hold, and $f(\cdot; z)$ is μ-strongly convex for all z. Then, with probability at least 0.9, the excess risk achieved by the projected gradient descent is $O(n^{-1/2})$ (while $O(n^{-1} \log n)$ in expectation).

Remark. A variant of the online projected gradient descent achieves an $O(n^{-1})$ expected excess risk, and $O(n^{-1} \log \log n)$ excess risk with high probability.

E. Hazan and S. Kale. 2014. Beyond the regret minimization barrier: Optimal algorithms for stochastic strongly convex optimization.

Discussion

Question. Is the definition of the regret satisfactory?

Question. Is the online-to-batch conversion necessary for "machine learning"?

Conclusions

Summary

- We have introduced the problem of online convex optimization.
  - DTOL is a special case.
- OCO may be solved via online mirror descent.
  - Regret analysis similar to that for the standard convex optimization case.
- A sublinear regret implies “small” excess risk.
  - Online-to-batch conversion.

Next lecture

- Learning in games (Abstract ver. 3).
- Solving minimax problems.