§28.1 Online Linear Optimization

V. 0.1

2022.04.07

§28.1.1. Setting

Different from adversarial linear bandit:

y_t could be observed by the learner rather than $\langle y_t, a_t \rangle$.

§28.1.2. Mirror Descent

- $\eta > 0$: the learning rate.
- $F: \mathbb{R}^d \to \bar{\mathbb{R}}$ is a convex function: the potential function / regularizer (usually F is Legendre).

- $a_1 = \operatorname{argmin}_{a \in \mathcal{A}} F(a)$

- $a_{t+1} = \operatorname{argmin}_{a \in \mathcal{A}} \eta \langle y_t, a \rangle + D_F(a, a_t)$ (28.2)

Remark. A simple case when $(a_t)_{t=1}^n$ is well-defined is when $\mathcal{A}$ is compact and F is Legendre.

§28.1.3 Follow-the-Regularized-Leader

- $a_1 = \operatorname{argmin}_{a \in \mathcal{A}} F(a)$

- $a_{t+1} = \operatorname{argmin}_{a \in \mathcal{A}} \eta \cdot \sum_{s=1}^{t} \langle y_t, a \rangle + F(a).$ (28.3)

§ 28.1.4 Equivalence of Mirror Descent and Follow-the-Regularized Leader.

Assumption: F is Legendre on domain $D \subseteq \mathcal{A}$.

Remark. The above assumption implies that the minimize of

$$\Phi_t(a) = \eta \cdot \langle y_t, a \rangle + D_F(a, a_t).$$

occurs in the interior of $D \subseteq \mathcal{A}$ and $\nabla \Phi(a_{t+1}) = 0$.

- a_{t+1} in MD:

$$\nabla \Phi_t(a_{t+1}) = \eta \cdot y_t + \nabla F(a_{t+1}) - \nabla F(a_t) = 0,$$

which by induction implies that

$$\nabla F(a_{t+1}) = -\sum_{s=1}^{t} \eta \cdot y_s + \nabla F(a_1) = -\eta \cdot \sum_{s=1}^{t} y_s,$$

where the last inequality follows by that $a_1 = \operatorname{argmin}_{a \in \mathcal{A}} F(a)$,

and again the fact that F is Legendre on $\mathrm{dom}(F) = D \subseteq \mathcal{A}$.

- a_{t+1} in FTRL:

Let $\Phi_t'(a) = \eta \cdot \sum_{s=1}^{t} \langle y_s, a \rangle + F(a)$. Then it holds that

$$\nabla \Phi'_t(a_{t+1}) = \eta \cdot \sum_{s=1}^{t} y_s + \nabla F(a_{t+1}),$$

which implies that $\nabla F(a_{t+1}) = -\eta \cdot \sum_{s=1}^{t} y_s$.

Remark. The equivalence between MD and FTRL only holds when (a) F is Legendre on $D = \mathrm{dom}(F)$, (b) $D \subseteq \mathcal{A}$, and (c) learning rate η is fixed. If a time-varying learning rate $(\eta_t)_{t=1}^{n}$ is adopted, then in MD

$$\nabla F(a_{t+1}) = -\sum_{s=1}^{t} \eta_s \cdot y_s$$

but in FTRL

$$\nabla F(a_{t+1}) = -\eta_t \cdot \sum_{s=1}^{t} y_s.$$

Example 28.1 [MD as Online Gradient Descent]

Assume $\mathcal{A} = \mathbb{R}^d$ and $F(a) = \frac{1}{2}\|a\|_2^2$. Clearly F is Legendre, $\nabla F(a) = a$, and $D_F(a_1, a_2) = \frac{1}{2}\|a_1\|_2^2 - \frac{1}{2}\|a_2\|_2^2 - a_2^\top \cdot (a_1 - a_2)$. Then

$$a_{t+1} = \nabla F(a_{t+1}) = \nabla F(a_t) - \eta \cdot y_t = a_t - \eta \cdot y_t,$$

which is usually called online gradient descent.

Example 28.2 [MD as Online Projected Gradient Descent]

Assume $\mathcal{A}$ is a convex subset of $\mathbb{R}^d$, and $F(a) = \frac{1}{2}\|a\|_2^2$.

Then MD chooses

$$a_{t+1} = \underset{a\in\mathcal{A}}{\operatorname{argmin}}\ \eta\cdot\langle y_t, a\rangle + D_F(a, a_t).$$

$$= \underset{a\in\mathcal{A}}{\operatorname{argmin}}\ \eta\cdot\langle y_t, a\rangle + \frac{1}{2}\|a - a_t\|_2^2.$$

$$= \Pi_{\mathcal{A}}(a_t - \eta\cdot y_t),$$

where $\Pi_{\mathcal{A}}(a)$ is the Euclidean projection of a onto $\mathcal{A}$.

The above algorithm is online projected gradient descent.

Example 28.3 [FTRL as Exponential Weight Algorithm]

Assume $\mathcal{A}$ is a probability simplex, and

$$F(x) = \sum_i x_i\cdot\ln x_i - \sum_i x_i$$

is the normalized negative entropy. Then FTRL chooses

$$a_{t+1} = \underset{a\in\mathcal{A}}{\operatorname{argmin}}\ \eta\cdot\sum_{s=1}^{t}\langle y_s, a\rangle + \sum_{i=1}^{d} a_i\cdot\ln a_i - \sum_{i=1}^{d} a_i,$$

which can be shown to satisfy that

$$a_{t+1,i} = \frac{\exp(-\eta\cdot L_{t,i})}{\sum_{j=1}^{d}\exp(-\eta\cdot L_{t,j})},$$

where $L_t = \sum_{s=1}^{t} \cdot y_s$.

§28.1.5 Implementation

Assume

- $\mathcal{A}$ is compact, and non-empty
- F is Legendre
- $\nabla F(a) - \eta \cdot y \in \mathrm{int}(\mathrm{dom}(F^*))$, $\forall a \in \mathcal{A} \cap D$ and $y \in \mathcal{L}$.

$$(28.6)$$

A. Implementation of MD [A Two-Step Process]

The solution of Eq. (28.3) could be found as

$$\tilde{a}_{t+1} = \operatorname*{argmin}_{a \in D} \eta \cdot \langle y_t, a \rangle + D_F(a, a_t). \tag{28.7}$$

$$a_{t+1} = \operatorname*{argmin}_{a \in \mathcal{A}} D_F(a, \tilde{a}_{t+1}).$$

Remark. Since F is Legendre, ∇F is a bijection between $\mathrm{int}(\mathrm{dom}(F))$ and $\mathrm{int}(\mathrm{dom}(F^*))$, which together with Eq. (28.6) implies that

$$\tilde{a}_{t+1} = (\nabla F)^{-1}(\nabla F(a_t) - \eta \cdot y_t).$$

B. Implementation of FTRL.

Assume

- $\mathcal{A}$ is compact and non-empty
- F is Legendre
- $-\eta \sum_{s=1}^{t} \cdot y_s \in \mathrm{int}(\mathrm{dom}(F^*)) \quad \forall t = 1, \dots, n.$

The leader chosen by FTRL at time $t+1$ is

$$a_{t+1} = \Pi_{\mathcal{A}}^{F}\left(\nabla F^{-1}\left(-\eta \cdot \sum_{s=1}^{t} y_s\right)\right).$$

where $\Pi_{\mathcal{A}}^{F}(a)$ indicates the Bregman projection of a onto $\mathcal{A}$ w.r.t. $D_F(\cdot,\cdot)$.

§ 28.2 Regret Analysis.

Theorem. 28.4 [Mirror Descent Regret Bound]

Let $\eta > 0$, F be Legendre with $\mathcal{D} = \mathrm{dom}(F)$, $\mathcal{A}$ be a non-empty convex set with $\mathrm{int}(\mathcal{D}) \cap \mathcal{A} \neq \emptyset$. Let $(a_t)_{t=1}^{n+1}$ be the actions chosen by mirror descent, which are assumed to be well-defined. Then, for any $a \in \mathcal{A}$, the regret of MD is bounded by

$$R_n(a) \le \frac{F(a) - F(a_1)}{\eta} + \sum_{t=1}^{n} \cdot \langle a_t - a_{t+1}, y_t \rangle - \frac{1}{\eta} \sum_{t=1}^{n} \cdot D(a_{t+1}, a_t).$$

Furthermore, suppose Eq. (28.6) holds and $(\tilde{a}_t)_{t=2}^{n+1}$ are given by Eq. (28.7), then

$$R_n(a) \le \frac{1}{\eta}\left(F(a) - F(a_1) + \sum_{t=1}^{n} D(a_t, \tilde{a}_{t+1})\right).$$

Proof. (1) Let $\Phi_t(b) = \eta\langle y_t, b\rangle + D_F(b, a_t)$. We start by splitting the instantaneous regret as

$$\langle a_t - a, y_t\rangle = \langle a_t - a_{t+1}, y_t\rangle + \underline{\langle a_{t+1} - a, y_t\rangle}.$$

To bound the second term, note that the 1st-order optimality of a_{t+1} shows that

这样split的原因之一是 a_{t+1} 与 y_t 一起才能构成一阶最优条件

$$\langle a - a_{t+1}, \nabla\Phi(a_{t+1})\rangle \geq 0,$$

which implies that

$$\langle a - a_{t+1}, \nabla F(a_{t+1}) - \nabla F(a_t) + \eta \cdot y_t\rangle \geq 0.$$

Reording and using the definition of Bregman divergence leads to

$$\begin{aligned}\langle a_{t+1} - a, y_t\rangle &\leq \frac{1}{\eta}\cdot\langle a - a_{t+1}, \nabla F(a_{t+1}) - \nabla F(a_t)\rangle \\ &= \frac{1}{\eta}\cdot\big(D(a, a_t) - D(a, a_{t+1}) - D(a_{t+1}, a_t)\big),\end{aligned}$$

where the equality follows by Generalized Pythagoras Indentity.

Using the above display, along with the definition of regret,

$$\begin{aligned}R_n &= \sum_{t=1}^{n}\langle a_t - a, y_t\rangle \\ &= \sum_{t=1}^{n}\langle a_t - a_{t+1}, y_t\rangle + \sum_{t=1}^{n}\langle a_{t+1} - a, y_t\rangle \\ &\leq \sum_{t=1}^{n}\langle a_t - a_{t+1}, y_t\rangle \\ &\quad + \frac{1}{\eta}\sum_{t=1}^{n}\big(D(a, a_t) - D(a, a_{t+1}) - D(a_{t+1}, a_t)\big)\end{aligned}$$

$$= \sum_{t=1}^{n} \langle a_t - a_{t+1}, y_t \rangle + \frac{1}{\eta}\left[D(a, a_1) - D(a, a_{n+1}) - \sum_{t=1}^{n} D(a_{t+1}, a_t) \right]$$

$$\le \sum_{t=1}^{n} \langle a_t - a_{t+1}, y_t \rangle + \frac{F(a) - F(a_1)}{\eta} - \frac{1}{\eta} \sum_{t=1}^{n} D(a_{t+1}, a_t), \quad (28.10)$$

where the last inequality follows from the non-negativity of Bregman divergence, and $\nabla F(a_1) = 0$.

② To bound the first term,

$$\langle a_t - a_{t+1}, y_t \rangle = \frac{1}{\eta} \langle a_t - a_{t+1}, \nabla F(a_t) - \nabla F(\tilde{a}_{t+1}) \rangle$$

$$= \frac{1}{\eta} \left(D(a_{t+1}, a_t) + D(a_t, \tilde{a}_{t+1}) - D(a_{t+1}, \tilde{a}_{t+1}) \right)$$

$$\le \frac{1}{\eta} \cdot \left(D(a_{t+1}, a_t) + D(a_t, \tilde{a}_{t+1}) \right).$$

Substituting the above display into Eq. (28.10) concludes the proof of second part of the theorem.

Remark. The assumption that $a_1 = \operatorname{argmin}_{a \in \mathcal{A}} F(a)$ is only used to bound $D(a, a_1) \le F(a) - F(a_1)$. For a different initialisation it still holds that

$$R_n(a) \le \frac{1}{\eta} \cdot \left(D(a, a_1) + \sum_{t=1}^{n} D(a_t, \tilde{a}_{t+1}) \right).$$

Theorem 28.5 [Follow-the-Regularized-Leader Regret Bound]

Let $\eta>0$, F be convex with domain $D=\mathrm{dom}(F)$, $\mathcal{A}\subseteq\mathbb{R}^d$ is a non-empty convex set. Assume $(a_t)_{t=1}^{n+1}$ chosen by FTRL are well-defined. Then $\forall a\in\mathcal{A}$, the regret is bounded by

$$R_n(a)\leq \frac{F(a)-F(a_1)}{\eta}+\sum_{t=1}^{n}\langle a_t-a_{t+1}, y_t\rangle-\frac{1}{\eta}\sum_{t=1}^{n}D(a_{t+1},a_t).$$

Proof. We first split the regret as in MD as

$$\sum_{t=1}^{n}\langle a_t-a, y_t\rangle=\sum_{t=1}^{n}\langle a_t-a_{t+1}, y_t\rangle+\sum_{t=1}^{n}\langle a_{t+1}-a, y_t\rangle.$$

To bound the second term, let

$$\Phi_t(b)=\sum_{s=1}^{t}\langle b, y_s\rangle+\frac{F(b)}{\eta}.$$

Then

$$\begin{aligned}
&\sum_{t=1}^{n}\langle a_{t+1}-a, y_t\rangle\\
&=\sum_{t=1}^{n}\langle a_{t+1}, y_t\rangle-\Phi_n(a)+\frac{F(a)}{\eta}.\\
&=\sum_{t=1}^{n}\left(\Phi_t(a_{t+1})-\Phi_{t-1}(a_{t+1})\right)-\Phi_n(a)+\frac{F(a)}{\eta}\\
&=-\Phi_0(a_1)+\sum_{t=0}^{n-1}\left(\Phi_t(a_{t+1})-\Phi_t(a_{t+2})\right)+\Phi_n(a_{n+1})-\Phi_n(a)+\frac{F(a)}{\eta}\\
&\leq\frac{F(a)-F(a_1)}{\eta}+\sum_{t=0}^{n-1}\left(\Phi_t(a_{t+1})-\Phi_t(a_{t+2})\right). \qquad (28.\mathrm{Ex}.1)
\end{aligned}$$

where the inequality is due to that a_{n+1} is the minimizer of $\Phi_n(\cdot)$. Note that $D_{\Phi_t}(x,y) = D_{\frac{F}{\eta}}(x,y)$ since $D_g(x,y) = D_f(x,y)$ if $g(x) = f(x) + kx$. Then

$$\Phi_t(a_{t+1}) - \Phi_t(a_{t+2}) = -D_{\Phi_t}(a_{t+2}, a_{t+1}) - \nabla \Phi_t(a_{t+1})^T (a_{t+2} - a_{t+1})$$

$$\leq -D_{\frac{F}{\eta}}(a_{t+2}, a_{t+1}), \qquad (28.\text{Ex}.2)$$

where the inequality comes from the 1^{st}-order optimality condition of a_{t+1} w.r.t. Φ_t. Substituting Eq. (28. Ex. 2) into Eq. (28. Ex. 1) concludes the proof.

Proposition 28.6 [Regret on the Unit Ball]

TBC

Proposition 28.7 [Regret on the Simplex]

TBC

Corollary 28.8 [ρ-strongly Strongly Convex Potential Regret Bound]

Let F be a Legendre potential and be ρ-strongly convex w.r.t. norm $\|\cdot\|_t$. Then the regret of MD or FTRL satisfies

$$R_n \leq \frac{\operatorname{diam}_F(\mathcal{A})}{\eta} + \frac{\eta}{2\rho} \cdot \sum_{t=1}^{n} \|y_t\|_{t*}^2,$$

where $\|\cdot\|_{t*}$ is the dual norm of $\|\cdot\|_t$.