Online Learning with a Hint

Jordana Mecler

Based on the article by Ofer Dekel, Arthur Flajolet, Nika Haghtalab and Patrick Jaillet, 2017

November 13, 2018

Outline

Preliminaries

- Notion of hint
- Online linear optimization
- Uniform convexity
- Exp-concavity
- Regret using exp-concavity

Improved Regret Bounds for Strongly Convex $\mathcal{K}$

- Naive strategy
- Virtual loss function
- Relationship between $\mathcal{K}$ and $\hat{c}_t$
- $\mathcal{A}_{hint}$ algorithm
- Analysis of $\mathcal{A}_{hint}$ algorithm
- Conclusion on $\mathcal{A}_{hint}$ algorithm

Lower Bounds

- Curvature of $\mathcal{K}$
- Optimal regret for strongly convex $\mathcal{K}$

Notion of hint

Other approaches:

- First coordinate of loss function is known (Hazan and Megiddo, 2007)
- Predictability around the gradient of the convex loss functions (Rahklin and Sridharan, 2013)

This work:

- Correlation between loss vector and another vector when the set is uniformly convex

Online linear optimization

Without hints: Let $\mathcal{A}$ be the player's algorithm for choosing actions. On round t player uses $\mathcal{A}$ and information observed so far to choose an action x_t in a convex compact set $\mathcal{K} \subseteq \mathbb{R}^d$. The adversary chooses a loss vector c_t in a compact set $\mathcal{C} \subseteq \mathbb{R}^d$. The player incurs the loss $c_t^T x_t$. The player's regret is:

$$R(\mathcal{A}, c_{1:T}) = \sum_{t=1}^{T} c_t^T x_t - \min_{x \in \mathcal{K}} \sum_{t=1}^{T} c_t^T x$$

With hints: Player observes $v_t \in \mathbb{R}^d with \|v_t\|_2 = 1$ before choosing x_t, with the guarantee that $v_t^T c_t \geq \alpha \|c_t\|_2$, for some $\alpha > 0$.

Uniform convexity

Degree of convexity of the playing set $\mathcal{K}$. The convex combination of any two points x and y on the boundary of $\mathcal{K}$ has to be sufficiently far from the boundary.

$\mathcal{K}$ is (C, q)-uniformly convex when $\|\frac{x+y}{2}\|_{\mathcal{K}} \leq 1 - C\epsilon^q$.

Examples

L_p balls for any $1 < p < \infty$ with modulus of convexity $\delta_{L_p}(\epsilon) = C_p \epsilon^p$ for $p \geq 2$ and a constant C_p and $\delta_{L_p}(\epsilon) = (p-1)\epsilon^2$ for $1 < p \leq 2$.

When $q = 2$, we say that $\mathcal{K}$ is *strongly convex*.

Exp-concavity

A function $f : \mathcal{K} \to \mathbb{R}$ is exp-concave with parameter $\beta > 0$ if $\exp(-\beta f(x))$ is a concave function of $x \in \mathcal{K}$.

Formally:
$\forall x, y \in \mathcal{K}, f(y) \geq f(x) + (y-x)^T \nabla f(x) + \frac{\beta^2}{2}((y-x)^T \nabla f(x))^2$

Regret using exp-concavity

Proposition 1 (Hazan et al., 2007): *Consider online convex optimization on a sequence of loss functions* $f_1, ..., f_T$ *over a feasible set* $\mathcal{K} \in \mathbb{R}^d$*, such that all* t*,* $f_t : \mathcal{K} \to \mathbb{R}$ *is exp-concave with parameter* $\beta > 0$*. There is an algorithm, with runtime polynomial in* d*, which we call* $\mathcal{A}_{EXP}$*, with a regret that is at most* $\frac{d}{\beta}(1 + \log(T + 1))$*.*

Naive strategy

First thought: since $c_t^T v_t \geq \alpha \|c_t\|_2$, move action x in the direction of $-v_t$, i.e., choose x_t as the extremal point in $\mathcal{K}$ in the direction of $-v_t$. (Linear regret)

Example

$c_t = (1, \frac{1}{2})$ and $v_t = (0, 1)$ for all t. Let $\mathcal{K}$ be the Euclidian ball.
Choosing $x_t = -v_t$ incurs a loss of $-\frac{T}{2}$.
Choosing the best fixed action in hindsight $x_t = (\frac{-2}{\sqrt{5}}, \frac{-1}{\sqrt{5}})$ incurs a loss of $\frac{-\sqrt{5}}{2}T$, leading to a regret of $\frac{\sqrt{5}-1}{2}T$

Intuitively it doesn't work because the hint doesn't give information about the $(d-1)$-dimensional subspace orthogonal to v_t.

Virtual loss function

The virtual loss simulates the process of moving x as far as possible in the direction of $-v_t$ without changing its value in the other directions, and applying the original cost function on that point, instead of x.

$$\hat{c}_t(x) = \min_{w \in \mathcal{K}} c_t^T w \;\; \text{s.t.} \;\; w^{\perp v_t} = x^{\perp v_t}$$

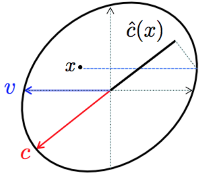

Figure 1: Virtual function $\hat{c}(.)$.

Relationship between $\mathcal{K}$ and $\hat{c}_t$

As the set of actions $\mathcal{C}$ is compact, there exists $G \geq 0$ such that $\|c\|_2 \leq G$. r and R are respectively the radius of the largest ball that fits into $\mathcal{K}$ and the radius of the smallest ball that contains $\mathcal{K}$.

Lemma 2: *If $\mathcal{K}$ is $(C, 2)$-uniformly convex, then $\hat{c}_t(.)$ is $8\frac{\alpha \cdot C \cdot r}{G \cdot R^2}$-exp-concave.*

Proof: Let $\gamma = 8\frac{\alpha \cdot C \cdot r}{G \cdot R^2}$. To prove that $\hat{c}_t(.)$ is exp-concave, it is sufficient to show that:

$$exp(-\gamma \cdot \hat{c}_t(\frac{x+y}{2})) \geq \frac{1}{2} exp(-\gamma \cdot \hat{c}_t(x)) + \frac{1}{2} exp(-\gamma \cdot \hat{c}_t(y))$$
$$\forall (x, y) \in \mathcal{K}.$$

Relationship between $\mathcal{K}$ and $\hat{c}_t$

Choose $(\hat{x}, \hat{y}) \in \mathcal{K}$ such that $\hat{c}_t(x) = c_t^T(\hat{x})$ and $\hat{c}_t(y) = c_t^T(\hat{y})$. We have $\|\hat{x}\|_{\mathcal{K}} = \|\hat{y}\|_{\mathcal{K}} = 1$. Since $exp(-\gamma \hat{c}_t(.))$ is decreasing in $\hat{c}_t(.)$:

$$exp(-\gamma \cdot \hat{c}_t(\frac{x+y}{2})) = \max_{\|w\|_{\mathcal{K}} \leq 1, w^{\perp v_t} = (\frac{x+y}{2})^{\perp v_t}} exp(-\gamma \cdot c_t^T w)$$

Using $w = \frac{\hat{x}+\hat{y}}{2} - \delta_{\mathcal{K}}(\|\hat{x} - \hat{y}\|_{\mathcal{K}}) \frac{v_t}{\|v_t\|_{\mathcal{K}}}$:

$$exp(-\gamma \cdot \hat{c}_t(\frac{x+y}{2})) \geq exp(-\frac{\gamma}{2} \cdot (c_t^T \hat{x} + c_t^T \hat{y}) + \gamma \cdot \frac{c_t^T v_t}{\|v_t\|_{\mathcal{K}}} \cdot \delta_{\mathcal{K}}(\|\hat{x} - \hat{y}\|_{\mathcal{K}}))$$

Relationship between $\mathcal{K}$ and $\hat{c}_t$

Since $\|v_t\|_{\mathcal{K}} \leq \frac{1}{r}\|v_t\|_2 = \frac{1}{r}$ and $\|\hat{x} - \hat{y}\|_{\mathcal{K}} \geq \frac{1}{R}\|\hat{x} - \hat{y}\|_2$ and $\delta_{\mathcal{K}}(\epsilon) = C\epsilon^q$:

$$exp(-\gamma \cdot \frac{c_t^T v_t}{\|v_t\|_{\mathcal{K}}} \cdot \delta_{\mathcal{K}}(\|\hat{x} - \hat{y}\|_{\mathcal{K}})) \geq$$

$$exp(\gamma \cdot r \cdot \alpha \cdot \|c_t\|_2 \cdot C \cdot \frac{1}{R^2}\|\hat{x} - \hat{y}\|_2^2) \geq$$

$$exp(\gamma \cdot \frac{\alpha \cdot C \cdot r}{R^2} \cdot \|c_t\|_2 \cdot (\frac{c_t^T \hat{x}}{\|c_t\|_2} - \frac{c_t^T \hat{y}}{\|c_t\|_2})^2) \geq$$

$$exp(\frac{(\frac{\gamma}{2})^2 \cdot (c_t^T \hat{x} - c_t^T \hat{y})^2}{2}) \geq$$

$$\frac{1}{2} \cdot exp(\frac{\gamma}{2} \cdot (c_t^T \hat{x} - c_t^T \hat{y})) + \frac{1}{2} exp(\frac{\gamma}{2} \cdot (c_t^T \hat{y} - c_t^T \hat{x}))$$

Relationship between $\mathcal{K}$ and $\hat{c}_t$

$$exp(-\gamma \cdot \hat{c}_t(\frac{x+y}{2})) \geq \frac{1}{2} exp(-\frac{\gamma}{2}({c_t}^T \hat{x} + {c_t}^T \hat{y})) \cdot$$
$$\{exp(\frac{\gamma}{2} \cdot ({c_t}^T \hat{x} - {c_t}^T \hat{y})) + exp(\frac{\gamma}{2} \cdot ({c_t}^T \hat{y} - {c_t}^T \hat{x}))\} =$$
$$\frac{1}{2} exp(-\gamma \cdot {c_t}^T \hat{y}) + \frac{1}{2} exp(-\gamma \cdot {c_t}^T \hat{x}) =$$
$$\frac{1}{2} exp(-\gamma \hat{c}_t(y)) + \frac{1}{2} exp(-\gamma \hat{c}_t(x))$$

$\mathcal{A}_{hint}$ algorithm

Algorithm 1 $\mathcal{A}_{hint}$ for strongly convex $\mathcal{K}$

for $t = 1, ..., T$ **do**

  Use algorithm $\mathcal{A}_{EXP}$ with the history $\hat{c}_\tau(.)$ for $\tau < t$. We move from x_{t-1} to $\hat{x}_t$ (the current chosen action).

  Let $x_t = \arg\min_{w \in \mathcal{K}}(v_t^T w)$ s.t. $w^{\perp v_t} = \hat{x}_t^{\perp v_t}$. Play x_t and receive c_t as feedback.

end for

Analysis of $\mathcal{A}_{hint}$ algorithm

Lemma 3: *For any sequence of loss functions $c_1, ...c_T$, let $R(\mathcal{A}_{hint}, c_{1:T})$ be the regret of algorithm $\mathcal{A}_{hint}$ on the sequence $c_1, ..., c_T$, and $R(\mathcal{A}_{EXP}, \hat{c}_{1:T})$ be the regret of algorithm $\mathcal{A}_{EXP}$ on the sequence of virtual loss functions $\hat{c}_1, ..., \hat{c_T}$. Then, $R(\mathcal{A}_{hint}, c_{1:T}) \leq R(\mathcal{A}_{EXP}, \hat{c}_{1:T})$.*

Analysis of $\mathcal{A}_{hint}$ algorithm

Proof: Start by showing that the loss of algorithm $\mathcal{A}_{hint}$ on the sequence $c_{1:T}$ is the same as the loss of algorithm $\mathcal{A}_{EXP}$ on the sequence $\hat{c}_{1:T}$.

$$\sum_{t=1}^{T} \hat{c}_t^T \hat{x}_t = \sum_{t=1}^{T} \min_{w \in \mathcal{K}: w^\perp = \hat{x}_t^\perp} c_t^T w =$$

$$\sum_{t=1}^{T} c_t^T (\arg\min_{w \in \mathcal{K}: w^\perp = \hat{x}_t^\perp} c_t^T w) = \sum_{t=1}^{T} c_t^T x_t$$

Analysis of $\mathcal{A}_{hint}$ algorithm

Now we show that the offline optimal on the sequence $\hat{c}_{1:T}$ is more competitive than the offline optimal on the sequence $c_{1:T}$. Note that $\hat{c}_t(x) = \min_{w \in \mathcal{K}: w^{\perp} = x^{\perp}} c_t^T w \leq c_t^T x$, therefore $\min_{x \in \mathcal{K}} \sum_{t=1}^{T} \hat{c}_t(x) \leq \min_{x \in \mathcal{K}} \sum_{t=1}^{T} c_t^T x$.

$$R(\mathcal{A}_{hint}, c_{1:T}) = \sum_{t=1}^{T} c_t^T x_t - \min_{x \in \mathcal{K}} \sum_{t=1}^{T} c_t^T x$$

$$\leq \sum_{t=1}^{T} \hat{c}_t(\hat{x}_t) - \min_{x \in \mathcal{K}} \sum_{t=1}^{T} \hat{c}_t(x) = R(\mathcal{A}_{EXP}, \hat{c}_{1:T})$$

Conclusion on $\mathcal{A}_{hint}$ algorithm

By applying Lemmas 3 and 4:

Theorem 4: *Suppose that $\mathcal{K} \in \mathbb{R}^d$ is a $(C, 2)$-uniformly convex set that is symmetric around the origin, and $B_r \subseteq \mathcal{K} \subseteq B_R$ for some r and R. Consider online linear optimization with hints were the cost function at round t is $\|c_t\|_2 \leq G$ and the hint v_t is such that $c_t^T v_t \geq \alpha \|c_t\|_2$, while $\|v_t\|_2 = 1$. Algorithm 1 in combination with $\mathcal{A}_{EXP}$ has a worst case regret of:*

$$R(\mathcal{A}_{hint}, c_{1:T}) \leq \frac{d \cdot G \cdot R^2}{8\alpha \cdot C \cdot r} \cdot (1 + \log(T + 1))$$

Curvature of $\mathcal{K}$

The curvature of the player's decision set $\mathcal{K}$ is necessary to get rates better than $O(\sqrt{T})$ even in the presence of a hint.

Theorem 5: *If the set of feasible actions is a polyhedron then, depending on the set $\mathcal{C}$, either there exists a trivial algorithm that achieves zero regret or every online algorithm has worst-case regret $\Omega(\sqrt{T})$. This is true even if the adversary is restricted to pick a fixed hint $v_t = v$ for all $t = 1,...,T$.*

Optimal regret for strongly convex $\mathcal{K}$

$O(\log(T))$ is the optimal achievable regret when $\mathcal{K}$ is strongly convex in OLO with a hint.

Theorem 6: *If $\mathcal{K}$ is a L_2 ball then, depending on the set $\mathcal{C}$, either there exists a trivial algorithm that achieves zero regret or every online algorithm has worst-case regret $\Omega(\log(T))$. This is true even if the adversary is restricted to pick a fixed hint $v_t = v$ for all $t =$ 1,...,T.*

References

O. Dekel, A. Flajolet, N. Haghtalab, P. Jaillet. Online Learning with a Hint. 2017.

E. Hazan and N. Megiddo. Online learning with prior knowledge. 2007.

A. Rakhlin and K. Sridharan. Online learning with predictable sequences. 2013.

G. Pisier. Martingales in banach spaces (in connection with type and cotype). 2011.

E. Hazan, A. Agarwal, S. Kale. Logarithmic regret algorithms for online convex optimization. 2007.