Bandit Online Convex Optimization

Nadejda Drenska and Scott Yang

March 31, 2015

Outline

1. OCO vs Bandit OCO
2. Gradient Estimates
3. Oblivious Adversary
4. Reshaping for Improved Rates
5. Adaptive Adversary
6. Concluding Remarks

Review of (Online) Convex Optimization

Set-up

1. Sequence of convex functions $\{c_t\}_{t=1}^{\infty} : S \to \mathbb{R}$ over convex set S.
2. Learner chooses point $x_t \in S$ and receives $c_t(x_t)$
3. Goal: minimize regret $\sum_{t=1}^{n} c_t(x_t) - \min_{x \in S} \sum_{t=1}^{n} c_t(x)$
4. Update rule: $x_{t+1} = x_t - \eta \nabla c_t(x_t)$.

Scenarios:

1. Offline: $c_t \equiv c$ fixed function
2. Online Stochastic: $c_t(x) = c(x) + \epsilon_t(x)$ noisy estimate
3. Online Adversarial: c_t chosen adversarially

Review of (Online) Convex Optimization

Set-up

1. Sequence of convex functions $\{c_t\}_{t=1}^{\infty} : S \to \mathbb{R}$ over convex set S.
2. Learner chooses point $x_t \in S$ and receives $c_t(x_t)$
3. Goal: minimize regret $\sum_{t=1}^{n} c_t(x_t) - \min_{x \in S} \sum_{t=1}^{n} c_t(x)$
4. Update rule: $x_{t+1} = x_t - \eta \underline{\nabla c_t(x_t)}$.

Scenarios:

1. Offline: $c_t \equiv c$ fixed function
2. Online Stochastic: $c_t(x) = c(x) + \epsilon_t(x)$ noisy estimate
3. Online Adversarial: c_t chosen adversarially

Key Information: *knowledge of gradients!*

Bandit Setting

Bandit Scenario: Set-up

1. Sequence of convex functions $\{c_t\}_{t=1}^{\infty} : S \to \mathbb{R}$ over convex set S.
2. Learner chooses point $x_t \in S$ and receives *value* $c_t(x_t)$
3. Goal: minimize regret $\sum_{t=1}^{n} c_t(x_t) - \min_{x \in S} \sum_{t=1}^{n} c_t(x)$

Question: *can we perform gradient descent without gradients?*

Estimating Gradients - Multi-point

Multi-point $\longrightarrow$ use finite differences!

1. one dim: $f'(x) \approx \frac{1}{h}(f(x+h) - f(x))$
2. d dim:
   $\nabla f(x) \approx \frac{1}{h}\left((f(x+he_1) - f(x)), \ldots, (f(x+he_d) - f(x))\right)$

Theorem (Agarwal et al, 2010)

Querying $d+1$ points is enough to recover standard OCO bounds, and even querying 2 points will get you to within $\log(T)$ terms.

Estimating Gradients - One-point

One-point $\longrightarrow f(x) \approx \frac{1}{vol(\delta B_1)} \int_{\delta B_1} f(x+v)dv$

one dim:

1. $f(x) \approx \frac{1}{2\delta} \int_{-\delta}^{\delta} f(x+v)dv$
2. $f'(x) \approx \frac{1}{2\delta} \int_{-\delta}^{\delta} f'(x+v)dv = \frac{1}{2\delta}(f(x+\delta) - f(x-\delta))$

d dim:

1. $f(x) \approx \frac{\int_{\delta B_1} f(x+v)dv}{vol(\delta B_1)} = \mathbb{E}_{v \in B_1}[f(x+\delta v)] =: \hat{f}(x)$
2. $$\nabla f(x) \approx \nabla \hat{f}(x) = \frac{\int_{\delta B_1} \nabla f(x+v)dv}{vol(\delta B_1)} = \frac{\int_{\partial(\delta B_1)} f(x+u)\frac{u}{|u|}du}{vol(\delta B_1)}$$
$$= \frac{\mathbb{E}_{u \in \partial B_1}[f(x+\delta u)u]vol(\partial(\delta B_1))}{vol(\delta B_1)} = \mathop{\mathbb{E}}_{u \in \partial B_1}[f(x+\delta u)u]\frac{d}{\delta}$$

Algorithm

$$\nabla f(x) \approx \nabla \hat{f}(x) = \frac{d}{\delta} \mathop{\mathbb{E}}_{u \in \partial B_1} [f(x + \delta u)u] \text{ where } u \sim U(\partial B_1)$$

Algorithm [Bandit Gradient Descent - Flaxman et al, 2004]:

1. Draw $u_t \sim U(\partial B_1)$
2. Play $x_t = y_t + \delta u_t$ and receive value $c_t(x_t)$
3. Update: $y_{t+1} = P_S(y_t - \nu c_t(x_t) u_t)$

Algorithm

Algorithm [Bandit Gradient Descent - Flaxman et al, 2004]:

1. Draw $u_t \sim U(\partial B_1)$
2. Play $x_t = y_t + \delta u_t$ and receive value $c_t(x_t)$
3. Update: $y_{t+1} = P_S(y_t - \nu c_t(x_t) u_t)$

Subtle issue: $y_t + \delta u_t$ might fall out of S, so use $P_{(1-\alpha)S}$ for $\alpha \in (0, 1)$ instead

New Update: $y_{t+1} = P_{(1-\alpha)S}(y_t - \nu c_t(x_t) u_t)$

Review of OCO Results

Theorem (Zinkevich 2003)

Let $S \subset B_R$, $\{c_t\} : S \to \mathbb{R}$ *seq of convex functions,* $G = \sup_t \|\nabla c_t(x_t)\|$, *and* $x_{t+1} = P_S(x_t - \eta \nabla c_t(x_t))$. *Then*

$$\sum_{t=1}^{n} c_t(x_t) - \min_{x \in S} \sum_{t=1}^{n} c_t(x) \leq \frac{R^2}{\eta} + n\frac{\eta G^2}{2}$$

Lemma (Randomized Zinkevich)

Let $S \subset B_R$, $\{c_t\} : S \to \mathbb{R}$ *seq of convex functions,* $\{g_t\}$ *s.t.* $\mathbb{E}[g_t | x_t] = \nabla c_t(x_t)$, $G = \sup_t \|g_t\|$, *and* $x_{t+1} = P_S(x_t - \eta g_t)$. *Then*

$$\mathbb{E}[\sum_{t=1}^{n} c_t(x_t)] - \min_{x \in S} \sum_{t=1}^{n} c_t(x) \leq \frac{R^2}{\eta} + n\frac{\eta G^2}{2}$$

Bandit OCO

Algorithm [Bandit Gradient Descent - Flaxman et al, 2004]:

1. Draw $u_t \sim U(\partial B_1)$
2. Play $x_t = y_t + \delta u_t$ and receive value $c_t(x_t)$
3. Update: $y_{t+1} = P_{(1-\alpha)S}(y_t - \nu c_t(x_t) u_t)$

Theorem (Flaxman et al 2004)

Assume that $\underline{B_r \subset S} \subset B_R$ and that $\underline{|c_t| \leq C \text{ uniformly}}$. Then for sufficiently large n and suitable choices of ν, δ, and α, we have

$$\mathbb{E}[\sum_{t=1}^{n} c_t(x_t)] - \min_{x \in S} \sum_{t=1}^{n} c_t(x) \leq 3Cn^{5/6} \left(\frac{dR}{r}\right)^{1/3}$$

Extending Randomized Zinkevich

Need to show:

1. BGD's updates valid for randomized Zinkevich.
   1. $x_t \in S$ for chosen α and δ.
   2. $g_t = \frac{d}{\delta} c_t(x_t) u_t$ are valid
2. Upper bounds on G for randomized Zinkevich.
3. Bound for $\hat{c}_t$ can be extended to c_t.
4. Bound for $(1 - \alpha)S$ can be extended to S.

BGD's Updates Valid for Randomized Zinkevich

1. $x_t \in S$ for chosen α and δ.
   - $\delta \leq \alpha r \Rightarrow x_t = y_t + \delta u_t \in S$
2. $g_t = \frac{d}{\delta} c_t(x_t) u_t$ are valid
   - $\mathbb{E}[g_t | y_t] = \frac{d}{\delta} \mathbb{E}_{u \in \partial B_1} [c_t(y_t + \delta u) u] = \nabla \mathbb{E}_{v \in B_1} [c_t(y_t + \delta v)]$

Upper Bounds on G for Zinkevich's Theorem

Fact: $\|g_t\| = \|\frac{d}{\delta}c_t(x_t)u_t\| \leq \frac{dC}{\delta}$

Thus, for

$$\alpha \in (0,1), \quad \delta \leq \alpha r, \quad \eta = \nu\frac{\delta}{d}, \quad G = \frac{dC}{\delta},$$

randomized Zinkevich implies that

$$\begin{aligned} \mathbb{E}[\sum_{t=1}^{n}\hat{c}_t(y_t)] - \min_{x\in(1-\alpha)S}\sum_{t=1}^{n}\hat{c}_t(x) &\leq \frac{R^2}{\eta} + n\frac{\eta G^2}{2} \\ &= RG\sqrt{n} = R\frac{dC}{\delta}\sqrt{n} \\ &\text{for } \eta = \frac{R}{G}\sqrt{n} \end{aligned}$$

Extending from $\hat{c}_t$ to c_t

Lemma

$$|c_t(y) - c_t(x)| \leq \frac{2C}{\alpha r}|x - y|, \quad \forall y \in (1-\alpha)S, x \in S$$

$$|\hat{c}_t(y_t) - c_t(y_t)| \leq \delta \frac{2C}{\alpha r}, \quad |\hat{c}_t(y_t) - c_t(x_t)| \leq 2\delta \frac{2C}{\alpha r}$$

The previous regret bound implies

$$\mathbb{E}[\sum_{t=1}^{n} c_t(x_t) - 2\delta \frac{2C}{\alpha r}] - \min_{x \in (1-\alpha)S} \sum_{t=1}^{n} c_t(x) + \delta \frac{2C}{\alpha r} \leq \frac{RdC\sqrt{n}}{\delta}$$

or

$$\mathbb{E}[\sum_{t=1}^{n} c_t(x_t)] - \min_{x \in (1-\alpha)S} \sum_{t=1}^{n} c_t(x) \leq \frac{RdC\sqrt{n}}{\delta} + 3\delta \frac{2C}{\alpha r}$$

Extending from $(1-\alpha)S$ to S

Lemma

$$\min_{x \in (1-\alpha)S} \sum_{t=1}^{n} c_t(x) \leq \min_{x \in S} \sum_{t=1}^{n} c_t(x) + 2\alpha C n$$

Thus,

$$\mathbb{E}[\sum_{t=1}^{n} c_t(x_t)] - \min_{x \in S} \sum_{t=1}^{n} c_t(x) \leq \frac{RdC\sqrt{n}}{\delta} + 3\delta \frac{2C}{\alpha r} + 2\alpha C n$$

Recap

1) $$\mathbb{E}[\sum_{t=1}^{n} \hat{c}_t(y_t)] - \min_{x \in (1-\alpha)S} \sum_{t=1}^{n} \hat{c}_t(x) \leq R\frac{dC}{\delta}\sqrt{n}$$

(optimal η, and $\delta \leq \alpha r$, $\alpha < 1$, $B_r \subset S$, $|c_t| \leq C$)

2) $$\mathbb{E}[\sum_{t=1}^{n} c_t(x_t)] - \min_{x \in (1-\alpha)S} \sum_{t=1}^{n} c_t(x) \leq \frac{RdC\sqrt{n}}{\delta} + 3\delta\frac{2C}{\alpha r}$$

(Lip across $(1-\alpha)S$ and S, $B_r \subset S$, $|c_t| \leq C$)

3) $$\mathbb{E}[\sum_{t=1}^{n} c_t(x_t)] - \min_{x \in S} \sum_{t=1}^{n} c_t(x) \leq \frac{RdC\sqrt{n}}{\delta} + 3\delta\frac{2C}{\alpha r} + 2\alpha Cn$$

(min's are close, $B_r \subset S$, $|c_t| \leq C$)

Remarks

1. Optimize over δ and α to get the theorem:

$$\mathbb{E}[\sum_{t=1}^{n} c_t(x_t)] - \min_{x \in S} \sum_{t=1}^{n} c_t(x) \leq 3Cn^{5/6} \left(\frac{dR}{r}\right)^{1/3}$$

2. If $Lip(c_t) \leq L$ known, then get $\mathcal{O}(n^{3/4})$ bound.
3. Preconditioning $\Rightarrow$ can improve ratio $\frac{R}{r}$

Theorem for L-Lipschitz Cost Functions

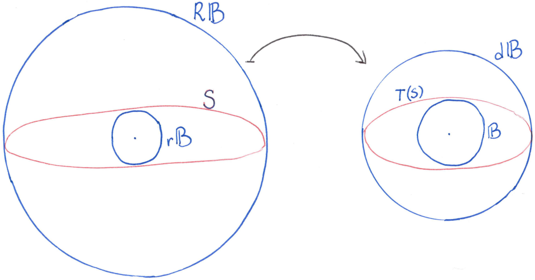
Theorem (Flaxman et al 2004 for L-Lipschitz cost functions)

If each c_t *is L-Lipschitz, then for* n *sufficiently large and* $\nu = \frac{R}{C\sqrt{n}}$, $\delta = n^{-.25}\sqrt{\frac{RdCr}{3(Lr+C)}}$, *and* $\alpha = \frac{\delta}{r}$, *we have*

$$\mathbb{E}[\sum_{t=1}^{n} c_t(x_t)] - \min_{x \in S} \sum_{t=1}^{n} c_t(x) \leq 2n^{3/4}\sqrt{3RdC(L + \frac{C}{r})}$$

.

Reshaping increases the accuracy of gradient descent!
Above regret bound depends on R/r - can be very large.
Idea: reshape the body to make it more 'round' - put it in isotropic position:



This amounts to finding an affine transformation T.

Isotropic Position and Algorithm

Isotropic position:

1. Estimate covariance of random samples from S (estimating r and R in $B_r \subseteq S \subseteq B_R$).
2. Find an affine transformation T so that the new covariance matrix is the *identity matrix*.
3. Apply T to $S \subseteq \mathbb{R}^d$ so $B_1 \subseteq T(S) \subseteq B_d$.
4. Then $R' = d$ and $r' = 1$.

Algorithm [Lovasz and Vempala, 2003]
- runs in time $O(d^4)$poly-log$(d, R/r)$ and
- puts the body in a *nearly isotropic* position: $R' = 1.01d$ and $r' = 1$.

Reshaping Properties

Lemma (New Lip constant $L' = LR$)

Let $c'_t(u) = c_t(T^{-1}(u))$. Then c'_t is LR-Lipschitz.

Proof outline:

Let $x_1, x_2 \in S$ and $u_1 = T(x_1), u_2 = T(x_2)$. Then,

$$|c'_t(u_1) - c'_t(u_2)| = |c_t(x_1) - c_t(x_2)| \leq L||x_1 - x_2||$$

Using that T is affine hence bounded, prove by contradiction that $||x_1 - x_2|| \leq R||u_1 - u_2||$, thus the LR-Lipschitz condition on c'_t. □

Reshaping and the BGD Algorithm

Corollary (Reshaping)

For a set S of diameter D, and c_t L-Lipschitz, after putting S into near-isotropic position, the BGD algorithm has expected regret

$$\mathbb{E}[\sum_{t=1}^{n} c_t(x_t)] - \min_{x \in S} \sum_{t=1}^{n} c_t(x) \leq 6n^{3/4} d(\sqrt{CLR} + C)$$

Without the L-Lipschitz condition

$$\mathbb{E}[\sum_{t=1}^{n} c_t(x_t)] - \min_{x \in S} \sum_{t=1}^{n} c_t(x) \leq 6n^{5/6} dC$$

Proof.

Use $r' = 1, R' = 1.01d, L' = LR$, and $C' = C$. □

Oblivious Adversary

So far, we have analyzed the algorithm in the case of an *oblivious adversary* who:

1. fixes the sequence of functions $c_1, c_2, ...$
2. knows the decision maker's algorithm
3. doesn't have knowledge of the random decisions of the algorithm

Adaptive Adversary

Consider an *adaptive adversary* who plays a game with the decision maker:

1. decision maker knows $x_1, c_1(x_1), x_2, c_2(x_2), ..., x_{t-1}, c_{t-1}(x_{t-1})$
2. decision maker chooses x_t
3. adaptive adversary knows $x_1, c_1, x_2, c_2, ..., x_{t-1}, c_{t-1}$
4. adaptive adversary chooses c_t

Main takeaway: *theorems against an oblivious adversary all hold against an adaptive adversary, up to changes of multiplicative constant by a factor of at most 3.*

Adaptive Adversary - Regret

Fact: *The bounds relating costs $c_t(x_t), c_t(y_t), \hat{c}_t(y_t)$ were all worst case bounds*, i.e. they hold for arbitrary c_t, regardless of whether the c_t are adaptively chosen or not. Thus it suffices to bound the regret:

$$\mathbb{E}[\sum_{t=1}^{n} \hat{c}_t(y_t) - \min_{y \in S} \sum_{t=1}^{n} \hat{c}_t(y)]$$

Idea: Need to show that the adversary's extra knowledge of $\{x_k\}_{k=1}^{t}$ cannot help to maximize the above regret.

Extending Randomized Zinkevich for an Adaptive Adversary

Need to show:

1. BGD's updates valid for randomized Zinkevich - next slides
    1. $x_t \in S$ for chosen α and δ.
    2. $g_t = \frac{d}{\delta} c_t(x_t) u_t$ are valid
2. Upper bounds on G for randomized Zinkevich. (from before)
3. Bound for $\hat{c}_t$ can be extended to c_t (from before)
4. Bound for $(1-\alpha)S$ can be extended to S - (from before)

Lemma - BGD Updates for an Adaptive Adversary

Lemma (BGD updates for adaptive costs)

Let $S \subset B_R$*,* $\{c_t\} : S \to \mathbb{R}$ *seq of convex differentiable functions,*$(c_{t+1}$ *possibly depending on* $z_1, z_2, ..., z_t)$*, where* $z_1, z_2, ..., z_t \in S$ *are defined by* $z_{t+1} = P_S(z_t - \eta g_t)$*. Here* $\{g_t\}$ *are vector-valued random variables s.t.*
$\mathbb{E}[g_t | z_1, c_1, z_2, c_2, ..., z_t, c_t] = \nabla c_t(z_t)$*,* $G = \sup_t \|g_t\|$*.*
Then, for $\eta = \frac{R}{G\sqrt{n}}$

$$\mathbb{E}[\sum_{t=1}^{n} c_t(z_t) - \min_{x \in S} \sum_{t=1}^{n} c_t(x)] \leq 3(\frac{R^2}{\eta} + n\frac{\eta G^2}{2}) = 3RG\sqrt{n}$$

Proof of Adaptive Adversary Lemma

Proof:

Let $h_t(x) := c_t(x) + x\xi_t$, where $\xi_t = g_t - \nabla c_t(z_t)$. Observe that

$$\nabla h_t(z_t) = \nabla c_t(z_t) + \xi_t = g_t$$

and
$||\xi_t|| \leq ||g_t|| + ||\nabla c_t(z_t)|| \leq 2G$. By Zinkevich's theorem, applied to h_t - which are deterministic at this point of the game

$$\sum_{t=1}^{n} h_t(z_t) \leq \min_{x \in S} \sum_{t=1}^{n} h_t(x) + RG\sqrt{n}$$

Since

$$\mathbb{E}[h_t(z_t)] = \mathbb{E}[c_t(z_t)] + \mathbb{E}[\xi_t . z_t] = \mathbb{E}[c_t(z_t)],$$

it suffices to show that

$$\mathbb{E}[\min_{x \in S} \sum_{t=1}^{n} h_t(x)] \leq \mathbb{E}[\min_{x \in S} \sum_{t=1}^{n} c_t(x)] + 2RG\sqrt{n}$$

Proof of Adaptive Adversary Lemma - part 2

Proof continued:

Left to show:

$$\mathbb{E}[\min_{x\in S}\sum_{t=1}^{n} h_t(x)] \le \mathbb{E}[\min_{x\in S}\sum_{t=1}^{n} c_t(x)] + 2RG\sqrt{n}$$

By Cauchy Schwartz
$|\sum_{t=1}^{n}(h_t(x) - c_t(x))| = |x\sum\xi_t| \le ||x||\cdot||\sum\xi_t|| \le R||\sum\xi_t||$.
This is in particular true for the minimal x. We take the expectation and bound the sum by using properties of i.i.d. vectors (recall: $||\xi_t|| \le 2G$):

$$(\mathbb{E}[||\sum\xi_t||])^2 \le \mathbb{E}[||\sum\xi_t||^2] = \sum\mathbb{E}[||\xi_t||^2] + 2\sum_{1\le s<t\le n}\mathbb{E}[\xi_s.\xi_t] \le 4nG^2$$

.

□

Summary

The paper extends Zinkevich's gradient descent idea to a problem in which one doesn't have access to the gradient.

Instead, the gradient of a function is approximated from a single sample.

Interpretation: approximation at each step is the gradient of a smoothed out version of the function at that step.

Analysis applies to both oblivious and adaptive adversaries (bounds change by a factor of 3).

Preconditioning ('reshaping') improves bounds significantly.

Possible Extensions

Extension of BGD to Zinkevich's model in the case of adaptive step size.

Extension of BGD to Zinkevich's model in the case of a non-stationary adversary, i.e. when regret is of the form:

$$\sum_{t=1}^{n} c_t(x_t) - \min_{w_1, w_2, \ldots, w_n \in S} \sum_{t=1}^{n} c_t(w_t)$$

Potential extension to minimizing any (convex?) function over a convex set by updates

$$y_{t+1} := y_t - \nu(c_t(x_t) - c_{t-1}(x_{t-1}))u_t,$$

even if we don't know the uniform bound on c_t.

References

Main paper:
A. D. Flaxman, A. T. Kalai, H. B. McMahan. Online convex optimization in the bandit setting: gradient descent without a gradient. 2004.

Zinkevich's paper:
M. Zinkevich. Online Convex Programming and Generalized Infinitesimal Gradient Ascent. In *Proceedings of the Twentieth International Conference on Machine Learning* pp. 928-936, 2003.

Multi-point paper:
A. Agarwal, O. Dekel, L. Xiao, (2010), Optimal Algorithms for Online Convex Optimization with Multi-Point Bandit Feedback., in *Adam Tauman Kalai and Mehryar Mohri, ed., 'COLT'* , Omnipress, , pp. 28-40 .

Thank you

Thank you!