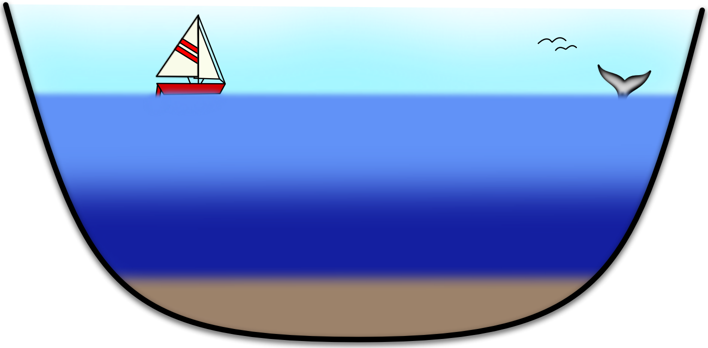
Improved Regret for Zeroth Order Bandit Convex Optimisation

Tor Lattimore

DeepMind

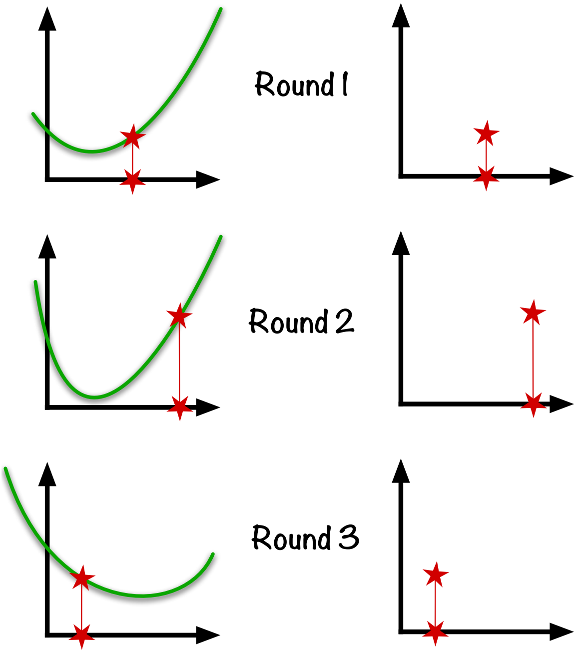


ADVERSARIAL BANDIT CONVEX OPTIMISATION

- $\mathcal{K}$ convex subset of $\mathbb{R}^d$
- Adversary chooses a sequence of convex functions $(f_t)_{t=1}^n$
- $f_t : \mathcal{K} \to [0,1]$
- Learner sequentially chooses $(X_t)_{t=1}^n$
- Observes $f_t(X_t)$
- Minimax regret

$$\mathfrak{R}_n^* = \inf_{\text{policies}} \sup_{(f_t)_{t=1}^n} \max_{x \in \mathcal{K}} \mathbb{E}\left[\sum_{t=1}^n f_t(X_t) - f_t(x)\right]$$

- Policy : history $\to$ distribution over actions
- Expectation with respect to randomness in policy


Round 1
Round 2
Round 3

Previous work

Authors	Result	Notes
Flaxman et al. (2005)	$dn^{4/5}$	
Agarwal et al. (2013)	$d^{18}\sqrt{n}$	stochastic
Bubeck et al. (2015)	$\sqrt{n}$	$d = 1$, inefficient
Hazan and Li (2016)	$O(\exp(d)\sqrt{n})$	
Bubeck and Eldan (2018)	$d^{11}\sqrt{n}$	inefficient
Bubeck et al. (2017)	$d^{9.5}\sqrt{n}$	inefficient
Bubeck et al. (2017)	$d^{10.5}\sqrt{n}$	efficient
Hazan and Levy (2014)	$d(d+\beta/m)^{1/2}\sqrt{n}$	β-smooth and m-strongly convex
This work	$\boldsymbol{d^{2.5}\sqrt{n}}$	**inefficient**
Dani et al. (2008)	$d\sqrt{n}$	**Lower bound**
Hu et al. (2016)	naive stuff fails	**Lower bound**

Theorem (L, 2020) Suppose that

(a) $\mathcal{K}$ contains a unit-width Euclidean ball

(b) $(f_t)_{t=1}^n$ are convex functions $f_t : \mathcal{K} \to [0, 1]$

Then the minimax regret is bounded by

$$\mathfrak{R}_n^* \leq \mathrm{const} \cdot d^{2.5} \sqrt{n} \log(n \, \mathrm{diam}(\mathcal{K}))$$

Approach

- **Reduction** Assume (f_t) are n-Lipschitz and $1/n$-strongly convex
- Bayesian regret

$$\mathfrak{BR}_n^* = \sup_{\text{prior on } (F_t)_{t=1}^n} \inf_{\text{policies}} \mathbb{E}\left[\max_{x \in \mathcal{K}} \sum_{t=1}^n F_t(X_t) - F_t(x)\right]$$

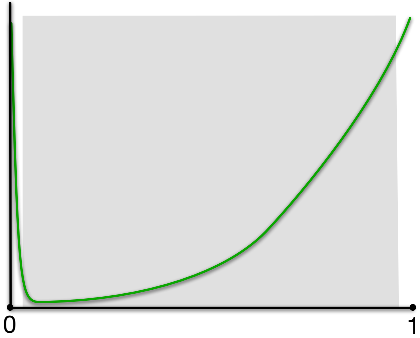
- Minimax theorem (Bubeck et al. (2015) and L & Szepesvári (2019))

$$\mathfrak{R}_n^* = \mathfrak{BR}_n^*$$

- Bound the Bayesian regret using information-theoretic machinery developed by Russo and Van Roy (2014)
- **Downside** No algorithm

REDUCTION

- Bounded convex functions must be Lipschitz deep inside the domain
- Adding small quadratic does not effect the regret much
- Play on $\{x \in \mathcal{K} : \mathrm{dist}(x, \partial\mathcal{K}) \geq 1/n\}$
- Add $\frac{1}{2n} \|X_t\|^2$ to all observed losses



Let $\mathcal{F}$ be the space of convex functions from $\mathcal{K}$ to $[0, 1]$

Theorem (Russo and Van Roy, 2014; Bubeck and Eldan, 2018) Suppose that for all finitely supported probability distributions μ on $\mathcal{F}$, there exists a probability measure ρ on $\mathcal{K}$ such that

$$\underbrace{\int_{\mathcal{K}} \bar{f}(x)\, \mathrm{d}\rho(x) - \int_{\mathcal{F}} f_\star\, \mathrm{d}\mu(f)}_{\text{regret}} \leq \alpha + \left(\beta \underbrace{\int_{\mathcal{F}} \int_{\mathcal{K}} (\bar{f}(x) - f(x))^2\, \mathrm{d}\rho(x)\, \mathrm{d}\mu(f)}_{\text{information gain}}\right)^{\frac{1}{2}}$$

where $\bar{f} = \int_{\mathcal{F}} f\, \mathrm{d}\mu(f)$ and $f_\star = \min_{x \in \mathcal{K}} f(x)$

Then there exists a policy such that

$$\mathfrak{BR}_n \leq n\alpha + O\left(\sqrt{dn\beta \log(n)}\right)$$

Combining exploratory distributions

Let μ be a finitely supported distribution on $\mathcal{F}$ and $\bar{f} = \int f \, \mathrm{d}\mu(f)$

Lemma Let $(\rho_i)_{i=1}^k$ be a collection of measures on $\mathcal{K}$. Suppose that for all $f \in \mathcal{F}$, there exists an i such that

$$\int_{\mathcal{K}} \bar{f} \, \mathrm{d}\rho_i - f_\star \leq \alpha + \left(\beta \int_{\mathcal{K}} (\bar{f} - f)^2 \, \mathrm{d}\rho_i \right)^{1/2}$$

Then, there exists a ρ such that

$$\int_{\mathcal{K}} \bar{f} \, \mathrm{d}\rho - \int_{\mathcal{F}} f_\star \, \mathrm{d}\mu(f) \leq \alpha + \left(k\beta \int_{\mathcal{F}} \int_{\mathcal{K}} (\bar{f} - f)^2 \, \mathrm{d}\rho \, \mathrm{d}\mu(f) \right)^{1/2}$$

Combining exploratory distributions

Let μ be a finitely supported distribution on $\mathcal{F}$ and $\bar{f} = \int f \, \mathrm{d}\mu(f)$

Lemma Let $(\rho_i)_{i=1}^k$ be a collection of measures on $\mathcal{K}$. Suppose that for all $f \in \mathcal{F}$, there exists an i such that

$$\int_{\mathcal{K}} \bar{f} \, \mathrm{d}\rho_i - f_\star \leq \alpha + \left(\beta \int_{\mathcal{K}} (\bar{f} - f)^2 \, \mathrm{d}\rho_i\right)^{1/2}$$

Then, there exists a ρ such that

$$\int_{\mathcal{K}} \bar{f} \, \mathrm{d}\rho - \int_{\mathcal{F}} f_\star \, \mathrm{d}\mu(f) \leq \alpha + \left(k\beta \int_{\mathcal{F}} \int_{\mathcal{K}} (\bar{f} - f)^2 \, \mathrm{d}\rho \, \mathrm{d}\mu(f)\right)^{1/2}$$

Implication: If for all $\bar{f} \in \mathcal{F}$ we can find $(\rho_i)_{i=1}^k$ satisfying the above, then

$$\mathfrak{R}_n^* = O\left(n\alpha + \sqrt{dnk\beta \log(n)}\right)$$

Proof Let $(\mathcal{F}_i)_{i=1}^k$ be a partition of $\mathcal{F}$ such that for all $f \in \mathcal{F}_i$,

$$\int_{\mathcal{K}} \bar{f} \, \mathrm{d}\rho_i - f_\star \leq \alpha + \left(\beta \int_{\mathcal{K}} (g - f)^2 \, \mathrm{d}\rho_i \right)^{1/2}$$

Let $\mu_i = \mu(\cdot \mid \mathcal{F}_i)$ and $\rho = \sum_{i=1}^k \mu(\mathcal{F}_i)\rho_i$

Proof Let $(\mathcal{F}_i)_{i=1}^k$ be a partition of $\mathcal{F}$ such that for all $f \in \mathcal{F}_i$,

$$\int_{\mathcal{K}} \bar{f} \, \mathrm{d}\rho_i - f_\star \leq \alpha + \left(\beta \int_{\mathcal{K}} (g - f)^2 \, \mathrm{d}\rho_i \right)^{1/2}$$

Let $\mu_i = \mu(\cdot \mid \mathcal{F}_i)$ and $\rho = \sum_{i=1}^k \mu(\mathcal{F}_i)\rho_i$

$$\int_{\mathcal{K}} \bar{f} \, \mathrm{d}\rho - \int_{\mathcal{F}} f_\star \, \mathrm{d}\mu(f) = \sum_{i=1}^k \mu(\mathcal{F}_i) \int_{\mathcal{F}} \left(\int_{\mathcal{K}} \bar{f} \, \mathrm{d}\rho_i - f_\star \right) \mathrm{d}\mu_i(f)$$

Proof Let $(\mathcal{F}_i)_{i=1}^k$ be a partition of $\mathcal{F}$ such that for all $f \in \mathcal{F}_i$,

$$\int_{\mathcal{K}} \bar{f} \, \mathrm{d}\rho_i - f_\star \leq \alpha + \left(\beta \int_{\mathcal{K}} (g - f)^2 \, \mathrm{d}\rho_i\right)^{1/2}$$

Let $\mu_i = \mu(\cdot \mid \mathcal{F}_i)$ and $\rho = \sum_{i=1}^k \mu(\mathcal{F}_i)\rho_i$

$$\int_{\mathcal{K}} \bar{f} \, \mathrm{d}\rho - \int_{\mathcal{F}} f_\star \, \mathrm{d}\mu(f) = \sum_{i=1}^k \mu(\mathcal{F}_i) \int_{\mathcal{F}} \left(\int_{\mathcal{K}} \bar{f} \, \mathrm{d}\rho_i - f_\star\right) \mathrm{d}\mu_i(f)$$

(by assumption) $\leq \alpha + \sum_{i=1}^k \mu(\mathcal{F}_i) \int_{\mathcal{F}} \left(\beta \int_{\mathcal{K}} (\bar{f} - f)^2 \, \mathrm{d}\rho_i\right)^{1/2} \mathrm{d}\mu_i(f)$

Proof Let $(\mathcal{F}_i)_{i=1}^k$ be a partition of $\mathcal{F}$ such that for all $f \in \mathcal{F}_i$,

$$\int_{\mathcal{K}} \bar{f} \,\mathrm{d}\rho_i - f_\star \leq \alpha + \left(\beta \int_{\mathcal{K}} (g - f)^2 \,\mathrm{d}\rho_i\right)^{1/2}$$

Let $\mu_i = \mu(\cdot \mid \mathcal{F}_i)$ and $\rho = \sum_{i=1}^k \mu(\mathcal{F}_i)\rho_i$

$$\int_{\mathcal{K}} \bar{f} \,\mathrm{d}\rho - \int_{\mathcal{F}} f_\star \,\mathrm{d}\mu(f) = \sum_{i=1}^k \mu(\mathcal{F}_i) \int_{\mathcal{F}} \left(\int_{\mathcal{K}} \bar{f} \,\mathrm{d}\rho_i - f_\star\right) \mathrm{d}\mu_i(f)$$

$$\text{(by assumption)} \leq \alpha + \sum_{i=1}^k \mu(\mathcal{F}_i) \int_{\mathcal{F}} \left(\beta \int_{\mathcal{K}} (\bar{f} - f)^2 \,\mathrm{d}\rho_i\right)^{1/2} \mathrm{d}\mu_i(f)$$

$$\text{(by Jensen's)} \leq \alpha + \sum_{i=1}^k \mu(\mathcal{F}_i) \left(\beta \int_{\mathcal{F}} \int_{\mathcal{K}} (\bar{f} - f)^2 \,\mathrm{d}\rho_i \,\mathrm{d}\mu_i(f)\right)^{1/2}$$

Proof Let $(\mathcal{F}_i)_{i=1}^k$ be a partition of $\mathcal{F}$ such that for all $f \in \mathcal{F}_i$,

$$\int_{\mathcal{K}} \bar{f} \, \mathrm{d}\rho_i - f_\star \leq \alpha + \left(\beta \int_{\mathcal{K}} (g - f)^2 \, \mathrm{d}\rho_i \right)^{1/2}$$

Let $\mu_i = \mu(\cdot \mid \mathcal{F}_i)$ and $\rho = \sum_{i=1}^k \mu(\mathcal{F}_i)\rho_i$

$$\int_{\mathcal{K}} \bar{f} \, \mathrm{d}\rho - \int_{\mathcal{F}} f_\star \, \mathrm{d}\mu(f) = \sum_{i=1}^k \mu(\mathcal{F}_i) \int_{\mathcal{F}} \left(\int_{\mathcal{K}} \bar{f} \, \mathrm{d}\rho_i - f_\star \right) \mathrm{d}\mu_i(f)$$

$$\text{(by assumption)} \leq \alpha + \sum_{i=1}^k \mu(\mathcal{F}_i) \int_{\mathcal{F}} \left(\beta \int_{\mathcal{K}} (\bar{f} - f)^2 \, \mathrm{d}\rho_i \right)^{1/2} \mathrm{d}\mu_i(f)$$

$$\text{(by Jensen's)} \leq \alpha + \sum_{i=1}^k \mu(\mathcal{F}_i) \left(\beta \int_{\mathcal{F}} \int_{\mathcal{K}} (\bar{f} - f)^2 \, \mathrm{d}\rho_i \, \mathrm{d}\mu_i(f) \right)^{1/2}$$

$$\text{(by Cauchy-Schwarz)} \leq \alpha + \sqrt{k\beta \sum_{i=1}^k \mu(\mathcal{F}_i)^2 \int_{\mathcal{F}} \int_{\mathcal{K}} (\bar{f} - f)^2 \, \mathrm{d}\rho_i \, \mathrm{d}\mu_i(f)}$$

Proof Let $(\mathcal{F}_i)_{i=1}^k$ be a partition of $\mathcal{F}$ such that for all $f \in \mathcal{F}_i$,

$$\int_{\mathcal{K}} \bar{f}\,\mathrm{d}\rho_i - f_\star \leq \alpha + \left(\beta \int_{\mathcal{K}} (g - f)^2\,\mathrm{d}\rho_i\right)^{1/2}$$

Let $\mu_i = \mu(\cdot \mid \mathcal{F}_i)$ and $\rho = \sum_{i=1}^k \mu(\mathcal{F}_i)\rho_i$

$$\int_{\mathcal{K}} \bar{f}\,\mathrm{d}\rho - \int_{\mathcal{F}} f_\star\,\mathrm{d}\mu(f) = \sum_{i=1}^k \mu(\mathcal{F}_i) \int_{\mathcal{F}} \left(\int_{\mathcal{K}} \bar{f}\,\mathrm{d}\rho_i - f_\star\right) \mathrm{d}\mu_i(f)$$

$$\text{(by assumption)} \leq \alpha + \sum_{i=1}^k \mu(\mathcal{F}_i) \int_{\mathcal{F}} \left(\beta \int_{\mathcal{K}} (\bar{f} - f)^2\,\mathrm{d}\rho_i\right)^{1/2} \mathrm{d}\mu_i(f)$$

$$\text{(by Jensen's)} \leq \alpha + \sum_{i=1}^k \mu(\mathcal{F}_i) \left(\beta \int_{\mathcal{F}} \int_{\mathcal{K}} (\bar{f} - f)^2\,\mathrm{d}\rho_i\,\mathrm{d}\mu_i(f)\right)^{1/2}$$

$$\text{(by Cauchy-Schwarz)} \leq \alpha + \sqrt{k\beta \sum_{i=1}^k \mu(\mathcal{F}_i)^2 \int_{\mathcal{F}} \int_{\mathcal{K}} (\bar{f} - f)^2\,\mathrm{d}\rho_i\,\mathrm{d}\mu_i(f)}$$

$$\left(\textstyle\sum_i a_i b_i \leq \sum_{i,j} a_i b_j\right) \leq \alpha + \sqrt{k\beta \int_{\mathcal{F}} \int_{\mathcal{K}} (\bar{f} - f)^2\,\mathrm{d}\rho\,\mathrm{d}\mu}$$

Goal

For any $\bar{f} \in \mathcal{F}$ find $(\rho_i)_{i=1}^k$ such that for all $f \in \mathcal{F}$,

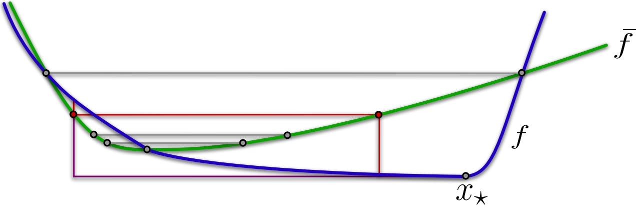
$$\int_{\mathcal{K}} \bar{f} \, \mathrm{d}\rho_\delta - f_\star \le \alpha + \left(\beta \int_{\mathcal{K}} (\bar{f} - f)^2 \, \mathrm{d}\rho_\delta\right)^{1/2}$$

When $d = 1$

- $\alpha = O(1/n)$
- $\beta = O(1)$
- $k = O(\log(n))$

When $d > 1$

- $\alpha = O(1/n)$
- $\beta = O(d^3)$
- $k = O(d \log(n))$



Notation

$K_\delta = \{x : \bar{f}(x) \leq \bar{f}_\star + \delta\}$

$\varepsilon :\; K_\varepsilon \subset \bar{x}_\star \pm \frac{1}{2n^2}$

$\mathcal{E} = \{0, \varepsilon, 2\varepsilon, \ldots, 2^{k-1}\varepsilon\}$

ρ_δ uniform on $\mathrm{bd}(K_\delta)$

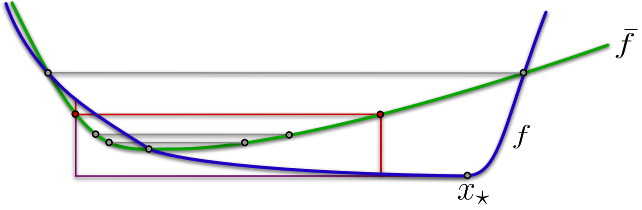
$k \approx \log(n)$, $\varepsilon = 1/\operatorname{poly}(n)$

Goal: $\forall f \in \mathcal{F}\ \exists \delta \in \mathcal{E} : \int_{\mathcal{K}} \bar{f}\, \mathrm{d}\rho_\delta - f_\star \leq \alpha + \left(\beta \int_{\mathcal{K}} (\bar{f} - f)^2\, \mathrm{d}\rho_\delta\right)^{1/2}$

Case 1: $f_\star \geq \bar{f}_\star - 1/n$

Case 2: $f_\star < \bar{f}_\star - 1/n$ and $\|x_\star - \bar{x}_\star\| \leq 1/(2n^2)$

Case 3: $f_\star < \bar{f}_\star - 1/n$ and $\|x_\star - \bar{x}_\star\| > 1/(2n^2)$



Notation

$K_\delta = \{x : \bar{f}(x) \leq \bar{f}_\star + \delta\}$

$\varepsilon :\ K_\varepsilon \subset \bar{x}_\star \pm \frac{1}{2n^2}$

$\mathcal{E} = \{0, \varepsilon, 2\varepsilon, \ldots, 2^{k-1}\varepsilon\}$

ρ_δ uniform on $\operatorname{bd}(K_\delta)$

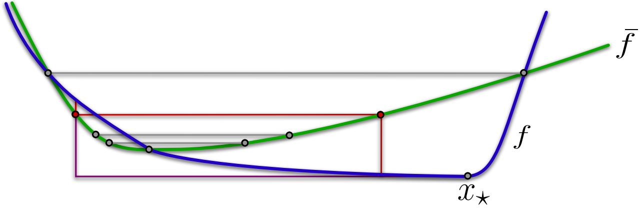
$k \approx \log(n),\ \varepsilon = 1/\operatorname{poly}(n)$

Goal: $\forall f \in \mathcal{F}\ \exists \delta \in \mathcal{E} : \int_{\mathcal{K}} \bar{f}\, d\rho_\delta - f_\star \leq \alpha + \left(\beta \int_{\mathcal{K}} (\bar{f} - f)^2\, d\rho_\delta\right)^{1/2}$

Case 1: $f_\star \geq \bar{f}_\star - 1/n$ Trivial with $\delta = 0$, $\alpha = 1/n$ and $\beta = 0$

Case 2: $f_\star < \bar{f}_\star - 1/n$ and $\|x_\star - \bar{x}_\star\| \leq 1/(2n^2)$

Case 3: $f_\star < \bar{f}_\star - 1/n$ and $\|x_\star - \bar{x}_\star\| > 1/(2n^2)$



Notation

$K_\delta = \{x : \bar f(x) \leq \bar f_\star + \delta\}$

$\varepsilon :\ K_\varepsilon \subset \bar x_\star \pm \frac{1}{2n^2}$

$\mathcal{E} = \{0, \varepsilon, 2\varepsilon, \ldots, 2^{k-1}\varepsilon\}$

ρ_δ uniform on $\operatorname{bd}(K_\delta)$

$k \approx \log(n),\ \varepsilon = 1/\operatorname{poly}(n)$

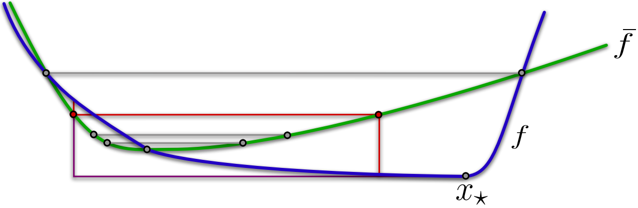
Goal: $\forall f \in \mathcal{F}\ \exists \delta \in \mathcal{E} : \int_{\mathcal{K}} \bar f \,\mathrm{d}\rho_\delta - f_\star \leq \alpha + \left(\beta \int_{\mathcal{K}} (\bar f - f)^2 \,\mathrm{d}\rho_\delta\right)^{1/2}$

Case 1: $f_\star \geq \bar f_\star - 1/n$ Trivial with $\delta = 0$, $\alpha = 1/n$ and $\beta = 0$

Case 2: $f_\star < \bar f_\star - 1/n$ and $\|x_\star - \bar x_\star\| \leq 1/(2n^2)$

$$\bar f(\bar x_\star) - f(\bar x_\star) \overset{\text{Lipschitzness}}{\geq} \bar f(\bar x_\star) - f_\star - n\|x_\star - \bar x_\star\| \geq \bar f(\bar x_\star) - f_\star - \frac{1}{2n} \geq \frac{1}{2}(\bar f(\bar x_\star) - f_\star)$$

Case 3: $f_\star < \bar f_\star - 1/n$ and $\|x_\star - \bar x_\star\| > 1/(2n^2)$



Notation

$K_\delta = \{x : \bar{f}(x) \leq \bar{f}_\star + \delta\}$

$\varepsilon : \; K_\varepsilon \subset \bar{x}_\star \pm \frac{1}{2n^2}$

$\mathcal{E} = \{0, \varepsilon, 2\varepsilon, \ldots, 2^{k-1}\varepsilon\}$

ρ_δ uniform on $\mathrm{bd}(K_\delta)$

$k \approx \log(n)$, $\varepsilon = 1/\operatorname{poly}(n)$

Goal: $\forall f \in \mathcal{F} \; \exists \delta \in \mathcal{E} : \int_{\mathcal{K}} \bar{f} \, \mathrm{d}\rho_\delta - f_\star \leq \alpha + \left(\beta \int_{\mathcal{K}} (\bar{f} - f)^2 \, \mathrm{d}\rho_\delta\right)^{1/2}$

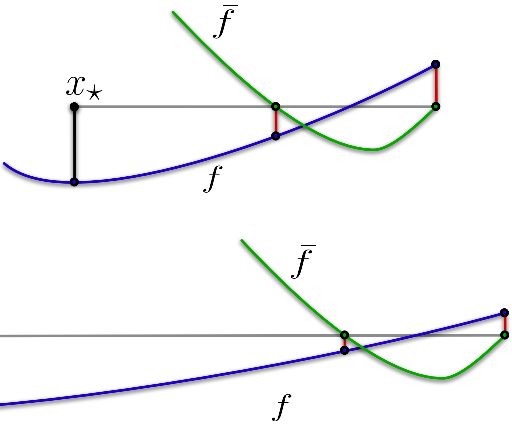
Case 1: $f_\star \geq \bar{f}_\star - 1/n$ Trivial with $\delta = 0$, $\alpha = 1/n$ and $\beta = 0$

Case 2: $f_\star < \bar{f}_\star - 1/n$ and $\|x_\star - \bar{x}_\star\| \leq 1/(2n^2)$

$$\bar{f}(\bar{x}_\star) - f(\bar{x}_\star) \overset{\text{Lipschitzness}}{\geq} \bar{f}(\bar{x}_\star) - f_\star - n\|x_\star - \bar{x}_\star\| \geq \bar{f}(\bar{x}_\star) - f_\star - \frac{1}{2n} \geq \frac{1}{2}(\bar{f}(\bar{x}_\star) - f_\star)$$

Case 3: $f_\star < \bar{f}_\star - 1/n$ and $\|x_\star - \bar{x}_\star\| > 1/(2n^2)$

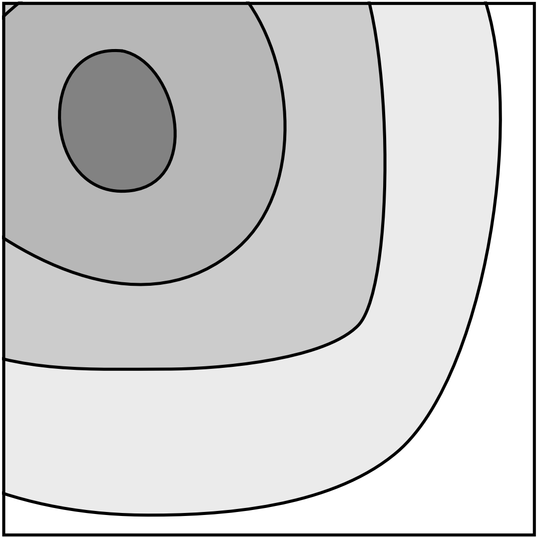
See figure



λ uniform on level set

$$\int_{\mathcal{K}} \bar f \, \mathrm{d}\lambda - f_\star = \bar f(x) - f_\star \lesssim \left(\frac{\text{distance}}{\text{width}}\right) \sqrt{\int_{\mathcal{K}} (\bar f - f)^2 \, \mathrm{d}\lambda}$$

Multi-dimensional ideas



Notation

$K_\delta = \{x : \bar{f}(x) \le \bar{f}_\star + \delta\}$

$\varepsilon :\ K_\varepsilon \subset \bar{x}_\star \pm \frac{1}{2n^2}$

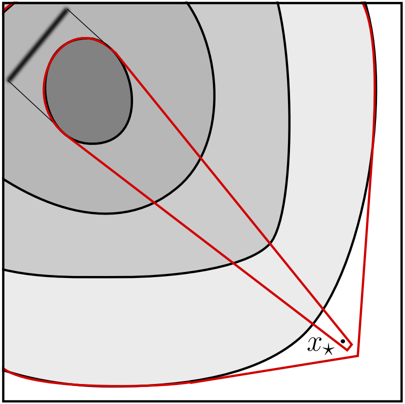
$\mathcal{E} = \{0, \varepsilon, \gamma\varepsilon, \ldots, \gamma^{k-1}\varepsilon\}$

ρ_δ non-uniform

$\gamma = (1+1/d),\ k \approx d\log(n),$

$\varepsilon = 1/\operatorname{poly}(n)$

Multi-dimensional ideas



Notation

$K_\delta = \{x : \bar{f}(x) \leq \bar{f}_\star + \delta\}$

$\varepsilon : \; K_\varepsilon \subset \bar{x}_\star \pm \frac{1}{2n^2}$

$\mathcal{E} = \{0, \varepsilon, \gamma\varepsilon, \ldots, \gamma^{k-1}\varepsilon\}$

ρ_δ non-uniform

$\gamma = (1+1/d),\; k \approx d\log(n),$

$\varepsilon = 1/\operatorname{poly}(n)$

Minimal surface area position

If K is in minimal surface area position, then $\lambda(\text{shadow}) \geq 1/d$

Conclusion

$\alpha = 1/n,\; \beta \approx d^3,\; k \approx d\log(n)$

$$\implies \mathfrak{R}_n^* = O(\sqrt{\beta d n k \log(n)}) = O(d^{2.5}\sqrt{n}\log(n))$$

$$\int_{\mathcal{K}} \bar{f}\,\mathrm{d}\lambda - f_\star \lesssim \underbrace{\left(\frac{\text{distance}}{\text{width}}\right)}_{\approx d} \left(\frac{1}{\lambda(\text{shadow})}\int_{\mathcal{K}} (\bar{f} - f)^2\,\mathrm{d}\lambda\right)^{1/2}$$

From information theory to mirror descent

$$Q_t(x) = \frac{\exp\left(-\eta \sum_{s=1}^{t-1} \hat{f}_s(x)\right)}{\int_{\mathcal{K}} \exp\left(-\eta \sum_{s=1}^{t-1} \hat{f}_s(x)\right) \mathrm{d}x} \qquad \Psi(u) = \exp(-u) - 1 + u$$

Find P_t and $G_t : \mathcal{K} \times \mathcal{K} \times [0,1] \to \mathbb{R}$ minimising

$$\max_{z \in \mathcal{K}} \max_{f \in \mathcal{F}} \Bigg[\underbrace{\int_{\mathcal{K}} f(x)\, \mathrm{d}(P_t - Q_t)(x)}_{\text{relative loss}}$$

$$+ \underbrace{\int_{\mathcal{K}} \left(f(y) - \int_{\mathcal{K}} G_t(y \mid x, f(x))\, \mathrm{d}x \right) \mathrm{d}(Q_t - \delta_z)(y)}_{\text{bias}}$$

$$+ \underbrace{\frac{1}{\eta} \int_{\mathcal{K}} \int_{\mathcal{K}} \Psi\left(\frac{\eta G_t(y \mid x, f(x))}{P_t(x)} \right) \mathrm{d}Q_t(y)\, \mathrm{d}P_t(x)}_{\text{estimation variance}} \Bigg]$$

Sample $X_t \sim P_t$ and $\hat{f}_t(\cdot) = G_t(\cdot | X_t, f_t(X_t))/P_t(X_t)$

Theorem (L & György): $\mathfrak{R}_n \leq \mathrm{const}\, d^{2.5} \sqrt{n} \log(n\, \mathrm{diam}(\mathcal{K}))$

THE END

- L & György, Mirror Descent and the Information Ratio, 2020
- L, Improved Regret for Adversarial Zeroth Order Bandit Convex Optimisation, 2020
- Conjecture is $\mathfrak{R}_n^* = \Theta(d^{1.5}\sqrt{n})$
- Best lower bound is $\mathfrak{R}_n^* = \Omega(d\sqrt{n})$

Agarwal, A., Foster, D. P., Hsu, D., Kakade, S. M., and Rakhlin, A. (2013). Stochastic convex optimization with bandit feedback. *SIAM Journal on Optimization*, 23(1):213–240.

Bubeck, S., Dekel, O., Koren, T., and Peres, Y. (2015). Bandit convex optimization: $\sqrt{T}$ regret in one dimension. In *Proceedings of the 28th Conference on Learning Theory*, pages 266–278, Paris, France. JMLR.org.

Bubeck, S. and Eldan, R. (2018). Exploratory distributions for convex functions. *Mathematical Statistics and Learning*, 1(1):73–100.

Bubeck, S., Lee, Y.-T., and Eldan, R. (2017). Kernel-based methods for bandit convex optimization. In *Proceedings of the 49th Annual ACM SIGACT Symposium on Theory of Computing*, pages 72–85.

Dani, V., Hayes, T. P., and Kakade, S. M. (2008). Stochastic linear optimization under bandit feedback. In *Proceedings of the 21st Conference on Learning Theory*, pages 355–366.

Flaxman, A., Kalai, A., and McMahan, H. (2005). Online convex optimization in the bandit setting: Gradient descent without a gradient. In *SODA'05: Proceedings of the sixteenth annual ACM-SIAM symposium on Discrete algorithms*, pages 385–394.

Hazan, E. and Levy, K. (2014). Bandit convex optimization: Towards tight bounds. In *Advances in Neural Information Processing Systems*, pages 784–792.

Hazan, E. and Li, Y. (2016). An optimal algorithm for bandit convex optimization. *arXiv preprint arXiv:1603.04350*.

Hu, X., Prashanth, L., György, A., and Szepesvári, C. (2016). (bandit) convex optimization with biased noisy gradient oracles. In *Artificial Intelligence and Statistics*, pages 819–828.

Russo, D. and Van Roy, B. (2014). Learning to optimize via information-directed sampling. In *Advances in Neural Information Processing Systems*, pages 1583–1591. Curran Associates, Inc.