Bandit Optimization with an Upper-Confident Frank-Wolfe

Vianney Perchet joint work with *Q. Berthet*

Last talk of the "In Memoriam of P. Rigollet" session
Optimization and Statistical Learning
Les Houches, April 2017

ENS Paris-Saclay
& Criteo Research, Paris

Original Motivations

The Original Problem

Heterogenous Source Estimation

- d different sources $X_k(t) \sim \mathcal{N}(\mu_k, \sigma_k^2)$ to estimate
- Total of N samples to allocate: $(N_1, N_2, \ldots, N_d)$
- Minimization of $\mathbb{E}\|\hat{\mu} - \mu\|^2 = \sum_k \frac{\sigma_k^2}{N_k} = \frac{1}{N}\sum_k \frac{\sigma_k^2}{p_k}$

Loss defined on Proportions

$$\mathcal{L}(p_1, \ldots, p_K) = \sum_k \frac{\sigma_k^2}{p_k}, \text{ with } p \in \Delta_d$$

The solution ?

Min. of $\mathcal{L}(p_1, \dots, p_d) = \sum_k \frac{\sigma_k^2}{p_k}$, constraint to $p \in \Delta_d$

- Easy to solve, $p_k^* = \frac{\sigma_k}{\sum \sigma_j}$ with error $\mathcal{L}(p^*) = (\sum \sigma_k)^2 \simeq \sigma^2 d^2$

The Question

What if the σ_k are also unknown?

- Sequentially estimate $\widehat{\sigma}_k^2 = \frac{1}{N_k} \sum_{t=1}^{N_k} \left(X_k(t) - \overline{X}_k(t)\right)^2$
  - Bigger N_k, better estimation of σ_k^2
  - Do not overshoot ! Smaller σ_k^2, smaller N_k

Sequential (simultaneous) Estimation **vs.** Optimization

Sounds Like Exploration vs Exploitation !

Stochastic Multi-Armed Bandit

- d different sources $X_k(t) \sim \mathcal{N}(\mu_k, \sigma_k^2)$
- Total of N samples to sequentially allocate: $(N_1, N_2, \ldots, N_d)$
- Minimization of $\frac{1}{N}\sum_k N_k \mu_k = \sum_k p_k \mu_k$

Loss defined on Proportions

$$\mathcal{L}(p_1, \ldots, p_d) = \sum_k p_k \mu_k = p^\top \mu, \text{ with } p \in \Delta_d$$

- Let's take $\sigma_k^2 = 1$ in bandits to simplify

The solution ?

Min. of $\mathcal{L}(p_1, \ldots, p_d) = p^\top \mu$, constrained to $p \in \Delta_k$

- Easiest ! $p^* = e_{k^*}$, where $\mu_{k^*} = \min \mu_k$ and error: $\mathcal{L}(p^*) = \mu_{k*}$

The Question in Bandit

What if the μ_k are unknown ?

- Use **U**pper-**C**onfidence **B**ound ! [Auer, Cesa-Bianchi, Freund, and Schapire]

  Or any variant [acronym]-UCB-[another acronym]:
  KL-UCB, UCB-V, UCB-2, Improved-UCB, UCB-+...
- Or variants of successive elimination (**E**xplore **T**hen **C**ommit)
  Worse performances than [X]-UCB-[Y], but simpler analysis (and interpretation)

Let's stick to the simplest

Upper-Confidence Bound - algorithm

1) Estimate μ_k by $\overline{\mu}_k(t) = \frac{1}{N_k(t)} \sum_{s=1}^{N_k(t)} X_k(s)$, but **biased**
2) "Un-bias it" with $\overline{\mu}_k(t) - \sqrt{2\frac{\log(t)}{N_k(t)}}$
3) Sample/pull the "arm" with smallest "unbiased" estimate

UCB-algo

$$\pi_{t+1} = \arg\max_k \left\{ \overline{\mu}_k(t) - \sqrt{2\frac{\log(t)}{N_k(t)}} \right\}$$

4) Enjoy Optimization error / "regret"

$$\mathcal{L}(p_N) - \mathcal{L}(p^*) \lesssim \frac{\log(N)}{N} \sum_k \frac{1}{\mu_k - \mu_{k^*}}$$

Ugly & useless but insightful 1 page proof

$$\mathcal{L}(p_{t+1}) - \mathcal{L}(p^*) = \mathcal{L}(p_t + \frac{1}{t+1}(e_{\pi_{t+1}} - p_t)) - \mathcal{L}(p^*)$$

Ugly & useless but insightful 1 page proof

$$\begin{aligned}\mathcal{L}(p_{t+1}) - \mathcal{L}(p^*) &= \mathcal{L}(p_t + \frac{1}{t+1}(e_{\pi_{t+1}} - p_t)) - \mathcal{L}(p^*) \\ &= \mathcal{L}(p_t) - \mathcal{L}(p^*) + \frac{1}{t+1}\nabla\mathcal{L}(p_t)^\top(e_{\pi_{t+1}} - p_t)\end{aligned}$$

Ugly & useless but insightful 1 page proof

$$
\begin{aligned}
\mathcal{L}(p_{t+1}) - \mathcal{L}(p^*) &= \mathcal{L}(p_t + \frac{1}{t+1}(e_{\pi_{t+1}} - p_t)) - \mathcal{L}(p^*) \\
&= \mathcal{L}(p_t) - \mathcal{L}(p^*) + \frac{1}{t+1}\nabla\mathcal{L}(p_t)^\top(e_{\pi_{t+1}} - p_t) \\
&= \mathcal{L}(p_t) - \mathcal{L}(p^*) + \frac{1}{t+1}\nabla\mathcal{L}(p_t)^\top(p^* - p_t) \\
&\qquad + \frac{1}{t+1}\nabla\mathcal{L}(p_t)^\top(e_{\pi_{t+1}} - p^*)
\end{aligned}
$$

Ugly & useless but insightful 1 page proof

$$
\begin{aligned}
\mathcal{L}(p_{t+1}) - \mathcal{L}(p^*) &= \mathcal{L}(p_t + \frac{1}{t+1}(e_{\pi_{t+1}} - p_t)) - \mathcal{L}(p^*) \\
&= \mathcal{L}(p_t) - \mathcal{L}(p^*) + \frac{1}{t+1}\nabla\mathcal{L}(p_t)^\top(e_{\pi_{t+1}} - p_t) \\
&= \mathcal{L}(p_t) - \mathcal{L}(p^*) + \frac{1}{t+1}\nabla\mathcal{L}(p_t)^\top(p^* - p_t) \\
&\qquad + \frac{1}{t+1}\nabla\mathcal{L}(p_t)^\top(e_{\pi_{t+1}} - p^*) \\
&\leq (1 - \frac{1}{t+1})\Big[\mathcal{L}(p_t) - \mathcal{L}(p^*)\Big] + \frac{1}{t+1}\nabla\mathcal{L}(p_t)^\top(e_{\pi_{t+1}} - p^*)
\end{aligned}
$$

Ugly & useless but insightful 1 page proof

$$
\begin{aligned}
\mathcal{L}(p_{t+1}) - \mathcal{L}(p^*) &= \mathcal{L}(p_t + \frac{1}{t+1}(e_{\pi_{t+1}} - p_t)) - \mathcal{L}(p^*) \\
&= \mathcal{L}(p_t) - \mathcal{L}(p^*) + \frac{1}{t+1}\nabla\mathcal{L}(p_t)^\top(e_{\pi_{t+1}} - p_t) \\
&= \mathcal{L}(p_t) - \mathcal{L}(p^*) + \frac{1}{t+1}\nabla\mathcal{L}(p_t)^\top(p^* - p_t) \\
&\qquad + \frac{1}{t+1}\nabla\mathcal{L}(p_t)^\top(e_{\pi_{t+1}} - p^*) \\
&\leq (1 - \frac{1}{t+1})\Big[\mathcal{L}(p_t) - \mathcal{L}(p^*)\Big] + \frac{1}{t+1}\nabla\mathcal{L}(p_t)^\top(e_{\pi_{t+1}} - p^*) \\
&\leq \frac{t}{t+1}\Big[\mathcal{L}(p_t) - \mathcal{L}(p^*)\Big] + \frac{1}{t+1}(\underbrace{\mu_{\pi_{t+1}} - \mu_{k^*}}_{:=\varepsilon_{t+1}})
\end{aligned}
$$

Ugly & useless but insightful 1 page proof

$$
\begin{aligned}
\mathcal{L}(p_{t+1}) - \mathcal{L}(p^*) &= \mathcal{L}(p_t + \frac{1}{t+1}(e_{\pi_{t+1}} - p_t)) - \mathcal{L}(p^*) \\
&= \mathcal{L}(p_t) - \mathcal{L}(p^*) + \frac{1}{t+1}\nabla\mathcal{L}(p_t)^\top(e_{\pi_{t+1}} - p_t) \\
&= \mathcal{L}(p_t) - \mathcal{L}(p^*) + \frac{1}{t+1}\nabla\mathcal{L}(p_t)^\top(p^* - p_t) \\
&\qquad + \frac{1}{t+1}\nabla\mathcal{L}(p_t)^\top(e_{\pi_{t+1}} - p^*) \\
&\leq (1 - \frac{1}{t+1})\Big[\mathcal{L}(p_t) - \mathcal{L}(p^*)\Big] + \frac{1}{t+1}\nabla\mathcal{L}(p_t)^\top(e_{\pi_{t+1}} - p^*) \\
&\leq \frac{t}{t+1}\Big[\mathcal{L}(p_t) - \mathcal{L}(p^*)\Big] + \frac{1}{t+1}(\underbrace{\mu_{\pi_{t+1}} - \mu_{k^*}}_{:=\varepsilon_{t+1}})
\end{aligned}
$$

$$
\mathcal{L}(p_N) - \mathcal{L}(p^*) \leq \frac{1}{N}\sum_{t=1}^{N}\mu_{\pi_t} - \mu_{k^*} := \frac{1}{N}\sum_{t=1}^{N}\varepsilon_t
$$

Still that proof !

$$\mathcal{L}(p_N) - \mathcal{L}(p^*) = \frac{1}{N}\sum_{t=1}^{N} \varepsilon_t = \frac{1}{N}\sum_k N_k(\mu_k - \mu_{k^*})$$

- $\pi_{t+1} = k$ if $\overline{X}_k(t) - \sqrt{\frac{\log(t)}{N_k(t)}} \simeq \mu_k - \sqrt{\frac{\log(N)}{N_k(t)}} \leq \mu_{k^*} \Rightarrow \varepsilon_t \lesssim \sqrt{\frac{\log(t)}{N_k(t)}}$

Slow rate of convergence

$$\mathcal{L}(p_N) - \mathcal{L}(p^*) \lesssim \frac{1}{N}\sum_k \sum_{s=1}^{N_k} \sqrt{\frac{\log(N)}{s}} \lesssim \frac{1}{N}\sum_k \sqrt{\log(N) N_k} \leq \sqrt{\frac{d\log(N)}{N}}$$

- Second equality + Cauchy-Schwartz

Fast rate of convergence

$$\left(\frac{1}{N}\sum_k N_k(\mu_k - \mu_{k^*})\right)^2 \leq \left(\sum_k \frac{\log(N)/N}{\mu_k - \mu_{k^*}}\right)\left(\frac{1}{N}\sum_k N_k(\mu_k - \mu_{k^*})\right)$$

$$\left(\mathcal{L}(p_N) - \mathcal{L}(p^*)\right)^2 \leq \left(\sum_k \frac{\log(N)/N}{\mu_k - \mu_{k^*}}\right)\left(\mathcal{L}(p_N) - \mathcal{L}(p^*)\right)$$

What did we learn with UCB ?

Optimistic Estimation of $\nabla\mathcal{L}(p)$ or "unbiased"

$$\overline{\mu}_k(t) - \sqrt{2\frac{\log(t)}{N_k(t)}} = \widehat{\nabla}_k^- \mathcal{L}(p_t) \text{ and } e_{t+1} = \arg\min_{p\in\Delta_d} \widehat{\nabla}^- \mathcal{L}(p_t)^\top p$$

Variant of Frank-Wolfe, because of proportions

$$\begin{aligned} p_{t+1} &= (1-\tfrac{1}{t+1})p_t + \tfrac{1}{t+1}e_{t+1} \\ &= (1-\tfrac{1}{t+1})p_t + \tfrac{1}{t+1}\arg\min_{p\in\Delta_d} \widehat{\nabla}^- \mathcal{L}(p_t)^\top p \end{aligned}$$

FW-algo: $p_{t+1} = (1-\gamma_t)p_t + \gamma_t \arg\min_{p\in\Delta_d} \nabla\mathcal{L}(p_t)^\top p$

From Slow to Fast Rates with (heavy) computations

$$\mathcal{L}(p_N) - \mathcal{L}(p^*) \lesssim \sqrt{\frac{\log(N)}{N}} \text{ vs. } \frac{\log(N)}{N}$$

Our model & results

More General Model

Optimization of convex loss $\mathcal{L}(p_N)$ on Δ_d, think of $\mathcal{L}(p) = \sum_k \frac{\sigma_k^2}{p_k}$

- Typical parametric form: $\mathcal{L}_\theta(p) = \sum_k f_k(\theta_k, p_k)$ with θ_k unknown

Main assumption (typical case)

f_k is smooth w.r.t. p and θ

- $\|\nabla f_k(\theta_k, p_k) - \nabla f_k(\theta'_k, p'_k)\| \leq C|p_k - p'_k| + C'\|\theta_k - \theta'_k\|$
- At stage t, choose e_{π_t} and observe $X_{\pi_t}(t) \sim \mathcal{N}(\theta_{\pi_t}, 1)$

After $N_k(t)$ observations, $\overline{X}_k(t) \simeq \theta_k \pm \sqrt{\frac{\log(t/\delta)}{N_k(t)}}$

- Noisy information on $\nabla_k \mathcal{L}(\cdot)$ only when sampling process k

Other examples

- Utility maximization Optim. basket of substitutes goods

Other examples

- Utility maximization Optim. basket of substitutes goods
  - V. thinks of “cardio”, “bench-press” and “squats” for fitness training



Other examples

- Utility maximization Optim. basket of substitutes goods
  - V. thinks of “cardio”, “bench-press” and “squats” for fitness training



  - Q. thinks of “wine”, “bread” and “cheese” for his breakfast



Other examples

- **Utility maximization** Optim. **basket of substitutes goods**
  - V. thinks of “cardio”, “bench-press” and “squats” for fitness training



  - Q. thinks of “wine”, “bread” and “cheese” for his breakfast



  - Kobb-Douglas utility $\mathcal{U}(x_1, \ldots, x_d) = x_1^{\beta_1} x_2^{\beta_2} \ldots x_d^{\beta_d}$
  - Use/buy one good (same price 1), estimate log-utility increase
- **online Markovitz portfolio optimization**
  - Optimize $\mathcal{L}(p) = p^\top \Sigma p - \lambda \mu^\top p$ with Σ known, μ unknown
- **General Case**
  - $\mathcal{L}$ is C-smooth w.r.t. p and $\left|\widehat{\nabla}_k \mathcal{L}(p) - \nabla_k \mathcal{L}(p)\right| \leq C' \sqrt{\frac{\log(t/\delta)}{N_k(t)}}$

The algorithm - Stochastic/Online optimization

FwUC: Frank-Wolfe Upper Confident

- Optimistic/Unbiased grad. $\widehat{\nabla}_k^- \mathcal{L}(p) = \widehat{\nabla}_k \mathcal{L}(p) - C' \sqrt{\frac{\log(t/\delta)}{N_k(t)}}$
- Frank-Wolfe: $e_{\pi_{t+1}} = \arg\min_{p \in \Delta_k} p^\top \widehat{\nabla}_k^- \mathcal{L}(p_t)$, with $\delta = 1/t$

First result (rather easy) Slow Rate of FwUC

$$\mathbb{E}\mathcal{L}(p_N) - \mathcal{L}(p^*) \lesssim C' \sqrt{\frac{d \log(N)}{N}} + O\left(\frac{\log(T)}{T}\right)$$

- **Proof ?** (almost) identical to UCB !

The proof. Identical !!

$$\mathcal{L}(p_{t+1}) - \mathcal{L}^* = \mathcal{L}(p_t + \frac{1}{t+1}(e_{\pi_{t+1}} - p_t)) - \mathcal{L}^*$$

$$\leq \mathcal{L}(p_t) - \mathcal{L}^* + \frac{1}{t+1}\nabla\mathcal{L}(p_t)^\top(e_{\pi_{t+1}} - p_t) + \frac{C}{(t+1)^2}$$

$$= \mathcal{L}(p_t) - \mathcal{L}^* + \frac{1}{t+1}\nabla\mathcal{L}(p_t)^\top(p^* - p_t)$$

$$+ \frac{1}{t+1}\nabla\mathcal{L}(p_t)^\top(e_{\pi_{t+1}} - p^*) + \frac{C}{(t+1)^2}$$

$$\leq \frac{t}{t+1}\Big[\mathcal{L}(p_t) - \mathcal{L}^*\Big] + \frac{1}{t+1}\nabla\mathcal{L}(p_t)^\top(e_{\pi_{t+1}} - p^*) + \frac{C}{(t+1)^2}$$

$$\leq \frac{t}{t+1}\Big[\mathcal{L}(p_t) - \mathcal{L}^*\Big] + \frac{1}{t+1}\varepsilon_t + \frac{C}{(t+1)^2}$$

$$\mathcal{L}(p_N) - \mathcal{L}^* \leq \frac{1}{N}\sum_{t=1}^{N}\varepsilon_t + C\frac{\log(N)}{N} \text{ and } \sum\varepsilon_t \simeq C\sum_k\sum_t\sqrt{\frac{\log(t)}{N_k(t)}}$$

Similar results/techniques

- Stochastic Frank Wolfe (errors independent of algorithms)
  - [Jaggi], [Lacoste-Julien et al.], [Lafond et al.]
- Global Cost. Specific $\mathcal{L}(p) = f(\theta^\top p)$ with θ unknown, f known
  - **Adversarial:** [Even-Dar et al.], [Blackwell], [Mannor et al.], etc.
  - **Stochastic:** [Agrawal and Devanur], [Agrawal et al] Also use stochastic Frank Wolfe
- Specific Cases. with pb tailored algorithm
  - [Carpentier et al.], [the bandit community]

Fast Rates

What about Fast Rates ?

- **Multi armed bandit** $\sqrt{\frac{d\log(N)}{N}}$ transformed into $\frac{d\log(N)}{N}$
- **General Case.** Can we do the same ?
  - Without more assumption, **no**.
  - Maybe with strong convexity

Strong convexity

$$f(y) \geq f(x) + \nabla f(x)^\top (y - x) + \mu \|y - x\|^2$$

- **Positive Results.** Fast rates sometimes possible
  - without noise [Garben and Hazan] [Jaggi][...]
  - with decaying noise [Lafond et al.]
  - in (almost) online convex optim. [Polyak-Tsybakov], [...]
- **Negative Results**
  - Cannot leverage strong convexity in online convex optim. [Shamir], [Jamieson et al.]
  - No choice of parameter in FW, has to be $\frac{1}{t+1}$

The model for fast rates

- On top of the previous assumptions

Assumptions

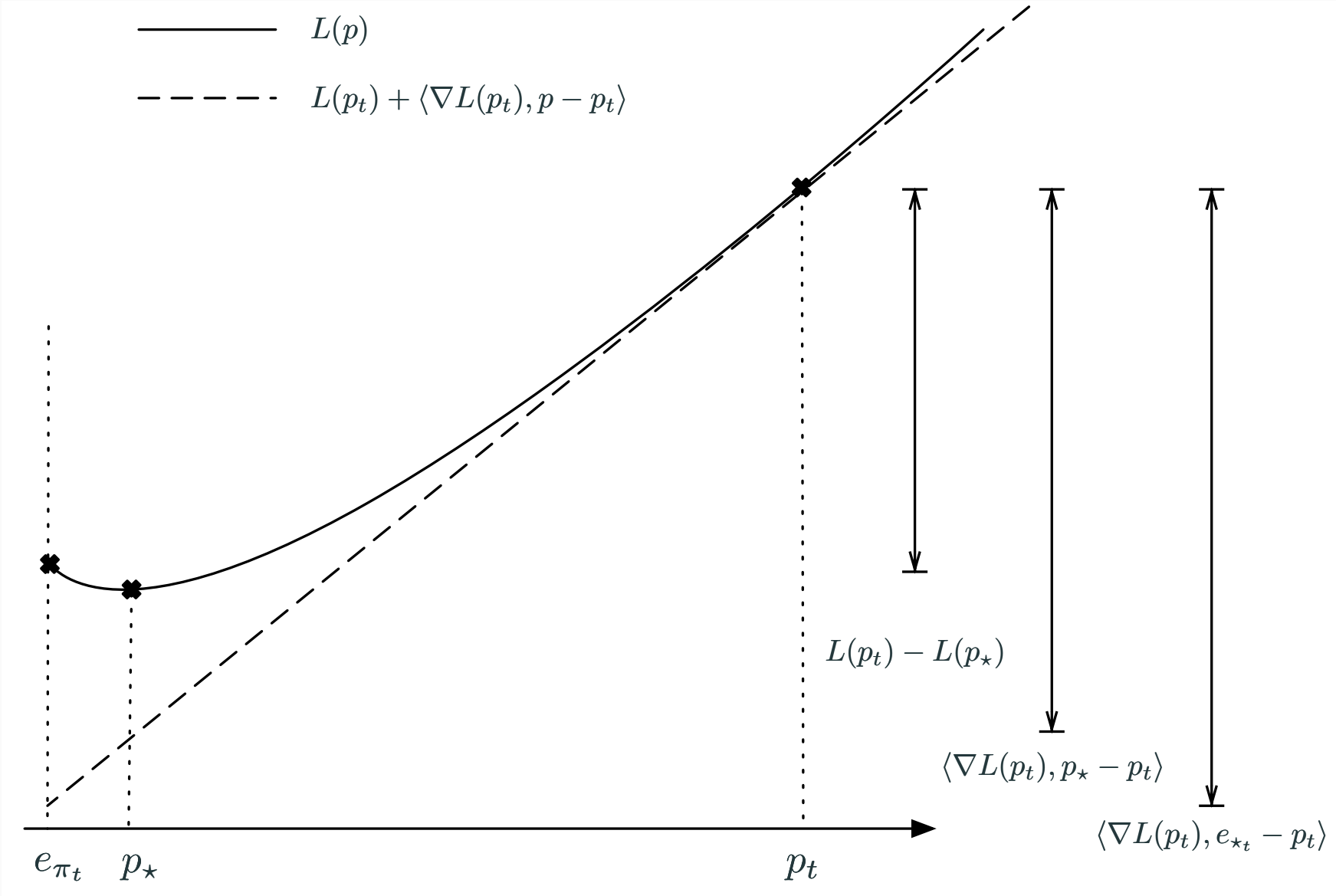
$\mathcal{L}$ is μ-strongly convex and minimized in the interior of Δ_d

$\eta := d(\partial\Delta_d, p^*)$ will play a role [Lacoste-Julien & Jaggi]

$$\mathcal{L}(p) - \mathcal{L}(p^*) \leq \frac{1}{2\mu\eta^2} |\nabla\mathcal{L}(p)^\top (e_{\star,p} - p)|^2$$

where $e_{\star,p} = \arg\min_{q \in \Delta_d} \mathcal{L}(p)^\top q$

The model for fast rates



The model for fast rates

- On top of the previous assumptions

Assumptions

$\mathcal{L}$ is μ-strongly convex and minimized in the interior of Δ_d

$\eta := d(\partial\Delta_d, p^*)$ will play a role [Lacoste-Julien & Jaggi]

$$\mathcal{L}(p) - \mathcal{L}(p^*) \leq \frac{1}{2\mu\eta^2} |\nabla\mathcal{L}(p)^\top (e_{\star,p} - p)|^2$$

where $e_{\star,p} = \arg\min_{q\in\Delta_d} \mathcal{L}(p)^\top q$

- **Main idea - change in proofs**
  - Before $\frac{1}{t+1}\nabla\mathcal{L}(p_t)^\top (e_{\star,p_t} - p_t) \leq -\frac{1}{t+1}(\mathcal{L}(p_t) - \mathcal{L}(p^*))$
  - Now $\frac{1}{t+1}\nabla\mathcal{L}(p_t)^\top (e_{\star,p_t} - p_t) \leq -\frac{\sqrt{2\mu\eta^2}}{t+1}\sqrt{(\mathcal{L}(p_t) - \mathcal{L}(p^*))}$

Fast rates, our result

FwUC: $e_{\pi_{t+1}} = \arg\min_{p \in \Delta_k} p^\top \widehat{\nabla}_k^- \mathcal{L}(p_t)$, with $\delta = 1/t$

Assumptions:

- C-smoothness/gradient estimation,
- μ-strong convexity,
- η-interior minimum

Main result, Fast rates of FwUC

$$\mathbb{E}\mathcal{L}(p_N) - \mathcal{L}(p^*) \leq c_1 \frac{\log^2(N)}{N} + c_2 \frac{\log(N)}{N} + c_3 \frac{1}{N}$$

with $c_1 = 3\frac{d(C')^2}{\mu\eta^2}$, $c_2 = 3\frac{dC'\|\mathcal{L}\|_\infty}{(\mu\eta^2)^3}$, $c_3 = dC'\|L\|_\infty + C$

Some remarks

- FwUC Fully adaptive to
  - The strong/non-strong convexity **and** the parameter μ
  - The horizon N
  - And any other constants/parameters except C'
- Parameters dependencies (**Leading Term**)
  - **Linear** in the ambiant dimension d
  - **inverse-Linear** in the strongly-convexity parameter μ
  - **inverse-square** in the distance to the boundary η (but $\frac{1}{d}$ on Δ_d)
- Generalizations
  - Gradients errors $\left(\frac{\log(t/\delta)}{N_k(t)}\right)^\beta$ with $\beta \leq 1/2$

    Slow rate $\left(\frac{\log(N)}{N}\right)^\beta$, and fast rates $\frac{\log(N)}{N^{2\beta}}$
  - (Non-strongly convex) without interior minimum but $\nabla\mathcal{L}(p^*) \ll 0$
- **Lower bounds** matching in N (classic in bandits/stoc. optim)

Ideas of proof

- Recall the first upper-bound

$$\begin{aligned}\mathcal{L}(p_{t+1}) - \mathcal{L}^* &\leq \mathcal{L}(p_t) - \mathcal{L}^* + \frac{1}{t+1}\nabla\mathcal{L}(p_t)^\top(e_{\pi_{t+1}} - p_t) + \frac{C}{(t+1)^2} \\ &\leq \mathcal{L}(p_t) - \mathcal{L}^* - \frac{\sqrt{2\mu\eta^2}}{t+1}\sqrt{\mathcal{L}(p_t) - \mathcal{L}^*} + \frac{\varepsilon_t}{t+1} + \frac{C}{(t+1)^2}\end{aligned}$$

- Introducing $\rho_t = \mathcal{L}(p_t) - \mathcal{L}^*$ and $\psi(x) = x^2 - \sqrt{2\mu\eta^2}x$, we get

$$(t+1)\rho_{t+1} \leq t\rho_t + \left[\psi(\rho_t) - \psi(\tfrac{\varepsilon_t}{\alpha})\right] + \tfrac{\varepsilon_t^2}{\alpha^2} + \tfrac{C}{t+1}$$

- if $\psi(\rho_t) - \psi(\frac{\varepsilon_t}{\alpha}) \leq 0$ then ok. **but not always...**

  more or less only asymptotically, if everything goes right.

Back to the original question

$$\mathcal{L}(p) = \sum_k \frac{\sigma_k^2}{p_k}, \quad p_k^* = \frac{\sigma_k}{\sum_j \sigma_j}, \quad , \mathcal{L}(p^*) = \left(\sum_j \sigma_j\right)^2$$

- Main issue: $\mathcal{L}$ not smooth in p nor σ^2...
  - But smooth "**around**" p^*, with $C' \simeq \frac{\sum \sigma_j}{\sigma_{\min}}$ and $C \simeq \frac{\left(\sum \sigma_j\right)^3}{\sigma_{\min}}$
- First phase of rough estimation of σ_k^2
  - Difficult to estimate $\sigma_k^2 \pm \varepsilon$, easy for $[\frac{\sigma_k^2}{2}, \frac{3\sigma_k^2}{2}]$
  - $X_t \sim \mathcal{N}(\theta_k, 1)$, sample as long as $\overline{X}_\tau \leq \sqrt{\frac{\log(T/\delta)}{\tau}}$
  - Need roughly $\frac{\log(T/\delta)}{\theta^2} = o(T)$ samples
- Second phase of sampling linear time
  - k sampled $N \frac{\widehat{\sigma}_k/2}{\sum_j 3\widehat{\sigma}_j/2} \leq N_k$ times
- Third phase of optimization, using FwUC
  - p_t far from boundary, close to p^*. Valid upper-bounds on C, C'