Online Convex Optimization

Dimitris Fotakis

School of Electrical and Computer Engineering
National Technical University of Athens, Greece

PAC Learning

Domain $\mathcal{X}$, labels $\mathcal{Y}$, hypothesis class $\mathcal{H} = \{h : (h : \mathcal{X} \to \mathcal{Y})\}$

(Fixed unknown) distribution $\mathcal{D}$ over $\mathcal{X}$

Training set $S \sim \mathcal{D}^m$

Realizability assumption: $\exists f \in \mathcal{H}$ that labels all $x \in \mathcal{X}$

Loss of hypothesis $h \in \mathcal{H}$: $L_{\mathcal{D},f}(h) = \mathbb{P}\mathrm{r}_{x\sim\mathcal{D}}[h(x) \neq f(x)]$

Class $\mathcal{H}$ is **PAC learnable** if for all ε, δ, there is # samples = $m_{\mathcal{H}}(\varepsilon, \delta)$ and algorithm A so that for any $m \geq m_{\mathcal{H}}(\varepsilon, \delta)$, $\mathcal{D}$ and f,

$$\mathbb{P}\mathrm{r}_{S\sim\mathcal{D}^m}\left[L_{\mathcal{D},f}(A(S)) \leq \varepsilon\right] \geq 1 - \delta$$

Empirical Risk Minimization (ERM): output hypothesis h not suffering any loss on S

PAC Learning

Domain $\mathcal{X}$, labels $\mathcal{Y}$, hypothesis class $\mathcal{H} = \{h : (h : \mathcal{X} \to \mathcal{Y})\}$

(Fixed unknown) distribution $\mathcal{D}$ over $\mathcal{X}$

Training set $S \sim \mathcal{D}^m$

Realizability assumption: $\exists f \in \mathcal{H}$ that labels all $x \in \mathcal{X}$

Loss of hypothesis $h \in \mathcal{H}$: $L_{\mathcal{D},f}(h) = \mathbb{P}\mathrm{r}_{x\sim\mathcal{D}}[h(x) \neq f(x)]$

Class $\mathcal{H}$ is **PAC learnable** if for all ε, δ, there is # samples = $m_{\mathcal{H}}(\varepsilon, \delta)$ and algorithm A so that for any $m \geq m_{\mathcal{H}}(\varepsilon, \delta)$, $\mathcal{D}$ and f,

$$\mathbb{P}\mathrm{r}_{S\sim\mathcal{D}^m}\left[L_{\mathcal{D},f}(A(S)) \leq \varepsilon\right] \geq 1 - \delta$$

Empirical Risk Minimization (ERM): output hypothesis h not suffering any loss on S

VC dimension :

- $\mathcal{H}$ **shatters** $C \subseteq \mathcal{X}$ if each of the $2^{|C|}$ possible labelings of C can be produced by some $h \in \mathcal{H}$.
- VC dimension of $\mathcal{H} = \sup\{|C| : \mathcal{H} \text{ shatters } C\}$

Agnostic PAC Learning

Domain $\mathcal{X}$, labels $\mathcal{Y}$, hypothesis class $\mathcal{H} = \{h : (h : \mathcal{X} \to \mathcal{Y})\}$

(Fixed unknown) distribution $\mathcal{D}$ over $\mathcal{X} \times \mathcal{Y}$

Training set $S = \{(x_1, y_1), \ldots, (x_m, y_m)\} \sim \mathcal{D}^m$

Loss of hypothesis $h \in \mathcal{H}$: $L_{\mathcal{D}}(h) = \mathbb{P}\mathrm{r}_{(x,y)\sim\mathcal{D}}[h(x) \neq y]$

Class $\mathcal{H}$ is **agnostically PAC learnable** if for all ε, δ, there is #samples = $m_{\mathcal{H}}(\varepsilon, \delta)$ and algorithm A so that for any $m \geq m_{\mathcal{H}}(\varepsilon, \delta)$ and $\mathcal{D}$,

$$\mathbb{P}\mathrm{r}_{S\sim\mathcal{D}^m}\left[L_{\mathcal{D}}(A(S)) \leq \varepsilon + \min_{f\in\mathcal{H}} L_{\mathcal{D}}(f)\right] \geq 1 - \delta$$

Empirical Risk Minimization (ERM): $\arg\min_{h\in\mathcal{H}} L_S(h)$

Uniform convergence: ERM on $\frac{\varepsilon}{2}$-representative training sets

For finite hypothesis class $\mathcal{H}$, $\lceil \frac{2\ln(2|\mathcal{H}|/\delta)}{\varepsilon^2} \rceil$ samples suffice for $\frac{\varepsilon}{2}$-representative training set.

Online Convex Optimization

General framework: convex set $S \subseteq \mathbb{R}^d$ On each day $t = 1, \ldots, T$:

1. Learner picks vector $p_t \in S$
2. Adversary picks convex **loss** function $f_t : S \to \mathbb{R}$, with f_t differentiable and L-Lipschitz wrt some norm $\| \cdot \|$, i.e., $|f_t(p) - f_t(p')| \leq L \cdot \|p - p'\|$
3. Learner learns f_t and incurs loss $f_t(p_t)$

Online Convex Optimization

General framework: convex set $S \subseteq \mathbb{R}^d$ On each day $t = 1, \ldots, T$:

1. Learner picks vector $p_t \in S$
2. Adversary picks convex **loss** function $f_t : S \to \mathbb{R}$, with f_t differentiable and L-Lipschitz wrt some norm $\| \cdot \|$, i.e., $|f_t(p) - f_t(p')| \leq L \cdot \|p - p'\|$
3. Learner learns f_t and incurs loss $f_t(p_t)$

Goal is to minimize **regret**:

$$\text{Regret}(T) = \sup_{f_1,\ldots,f_T} \left(\sum_{t=1}^{T} f_t(p_t) - \min_{p \in S} \sum_{t=1}^{T} f_t(p) \right)$$

(Online) algorithm is **no-regret** if $\text{Regret}(T)/T \to 0$ at $T \to \infty$

Online Convex Optimization

General framework: convex set $S \subseteq \mathbb{R}^d$ On each day $t = 1, \ldots, T$:

1. Learner picks vector $p_t \in S$
2. Adversary picks convex **loss** function $f_t : S \to \mathbb{R}$, with f_t differentiable and L-Lipschitz wrt some norm $\| \cdot \|$, i.e., $|f_t(p) - f_t(p')| \leq L \cdot \|p - p'\|$
3. Learner learns f_t and incurs loss $f_t(p_t)$

Goal is to minimize **regret**:

$$\text{Regret}(T) = \sup_{f_1, \ldots, f_T} \left(\sum_{t=1}^{T} f_t(p_t) - \min_{p \in S} \sum_{t=1}^{T} f_t(p) \right)$$

(Online) algorithm is **no-regret** if $\text{Regret}(T)/T \to 0$ at $T \to \infty$

- Experts: $S = \Delta_d$, $f_t(p) = \langle p, \ell_t \rangle$ (linear expected loss)
- Online Quadratic Optimization: learner $p_t \in S$, adversary $z_t \in S$, $f_t(p) = \|p - z_t\|_2^2$.
- Online Least Squares Linear Regression: learner $p_t \in \mathbb{R}^d$, $\|p_t\| \leq B$, adversary $(x_t, y_t) \in \mathbb{R}^d \times \mathbb{R}$, $f_t(p) = (\langle p, x_t \rangle - y_t)^2$

Follow / Be the Regularized Leader

$F_t(p) = \sum_{\tau=1}^{t} f_\tau(p)$ and $\tilde{F}_t(p) = \sum_{\tau=1}^{t} f_\tau(p) + R(p)/\eta$

FTRL: $\tilde{p}_t = \arg\min_{p \in S} \tilde{F}_{t-1}(p)$

BTRL: $\tilde{p}_t^* = \arg\min_{p \in S} \tilde{F}_t(p)$

Follow / Be the Regularized Leader

$F_t(p) = \sum_{\tau=1}^{t} f_\tau(p)$ and $\tilde{F}_t(p) = \sum_{\tau=1}^{t} f_\tau(p) + R(p)/\eta$

FTRL: $\tilde{p}_t = \arg\min_{p \in S} \tilde{F}_{t-1}(p)$

BTRL: $\tilde{p}_t^* = \arg\min_{p \in S} \tilde{F}_t(p)$

$1/\eta$-**strongly convex** function $R : S \to \mathbb{R}$ wrt norm $\| \cdot \|$, if $\forall\, x, y \in S$:

$$R(x) \geq R(y) + \langle \nabla R(y), x - y \rangle + \frac{1}{2\eta} \|x - y\|^2$$

$F_t(p) = \sum_{\tau=1}^{t} f_\tau(p)$ and $\tilde{F}_t(p) = \sum_{\tau=1}^{t} f_\tau(p) + R(p)/\eta$

FTRL: $\tilde{p}_t = \arg\min_{p \in S} \tilde{F}_{t-1}(p)$

BTRL: $\tilde{p}_t^* = \arg\min_{p \in S} \tilde{F}_t(p)$

$1/\eta$-**strongly convex** function $R : S \to \mathbb{R}$ wrt norm $\| \cdot \|$, if $\forall\, x, y \in S$:

$$R(x) \geq R(y) + \langle \nabla R(y), x - y \rangle + \frac{1}{2\eta} \|x - y\|^2$$

Functions $f, g : S \to \mathbb{R}$ be $1/\eta$-strongly convex wrt some norm $\| \cdot \|$ and $h(x) = g(x) - f(x)$ be L-Lipschitz wrt $\| \cdot \|$. Then, $\|x_f^* - x_g^*\| \leq \eta \cdot L$, with x_f^*, x_g^* minimizers of f, g.

Regret of FTRL Against BTRL

$$\text{Regret}_{FTRL}(T) \leq \text{Regret}_{BTRL}(T) + L \cdot \sum_{t=1}^{T} \|\tilde{p}_t - \tilde{p}_{t+1}\|$$
$$\leq \text{Regret}_{BTRL}(T) + \eta \cdot L^2 \cdot T$$

Regret of FTRL Against BTRL

$$\begin{aligned}\text{Regret}_{FTRL}(T) &\leq \text{Regret}_{BTRL}(T) + L \cdot \sum_{t=1}^{T} \|\tilde{p}_t - \tilde{p}_{t+1}\| \\ &\leq \text{Regret}_{BTRL}(T) + \eta \cdot L^2 \cdot T\end{aligned}$$

Proof :

- Second inequality from strong convexity, because $\tilde{p}_t, \tilde{p}_{t+1}$ are minimizers of $1/\eta$-strong convex functions $\tilde{F}_{t-1}(p)$ and $\tilde{F}_t(p)$ with difference $f_t(p)$ which is L-Lipschitz.

Regret of FTRL Against BTRL

$$\text{Regret}_{FTRL}(T) \le \text{Regret}_{BTRL}(T) + L \cdot \sum_{t=1}^{T} \|\tilde{p}_t - \tilde{p}_{t+1}\|$$
$$\le \text{Regret}_{BTRL}(T) + \eta \cdot L^2 \cdot T$$

Proof :

- Second inequality from strong convexity, because $\tilde{p}_t, \tilde{p}_{t+1}$ are minimizers of $1/\eta$-strong convex functions $\tilde{F}_{t-1}(p)$ and $\tilde{F}_t(p)$ with difference $f_t(p)$ which is L-Lipschitz.
- We observe that:

$$\text{Regret}_{FTRL}(T) - \text{Regret}_{BTRL}(T) = \sum_{t=1}^{T} (f_t(\tilde{p}_t) - f_t(\tilde{p}_t^*))$$
$$\le L \sum_{t=1}^{T} \|\tilde{p}_t - \tilde{p}_t^*\|$$
$$= L \sum_{t=1}^{T} \|\tilde{p}_t - \tilde{p}_{t+1}\|$$

Regret of Be the Regularized Leader

$$\text{Regret}_{BTRL}(T) \leq \frac{1}{\eta}\left(\max_{p\in S} R(p) - \min_{p\in S} R(p)\right)$$

Regret of Be the Regularized Leader

$$\text{Regret}_{BTRL}(T) \leq \frac{1}{\eta}\left(\max_{p \in S} R(p) - \min_{p \in S} R(p)\right)$$

Proof :

- Let $f_0(p) = R(p)/\eta$ and $\tilde{p}_0^* = \arg\min_{p \in S} R(p)/\eta$.
- Using induction on t, we show that for all $t \geq 1$,

$$\sum_{\tau=0}^{t} f_\tau(\tilde{p}_\tau^*) \leq \tilde{F}_t(\tilde{p}_t^*) \quad \text{(including fake action } \tilde{p}_0^* \text{ at } \tau = 0\text{)}$$

Regret of Be the Regularized Leader

$$\text{Regret}_{BTRL}(T) \leq \frac{1}{\eta}\left(\max_{p\in S} R(p) - \min_{p\in S} R(p)\right)$$

Proof :

- Let $f_0(p) = R(p)/\eta$ and $\tilde{p}_0^* = \arg\min_{p\in S} R(p)/\eta$.
- Using induction on t, we show that for all $t \geq 1$,

$$\sum_{\tau=0}^{t} f_\tau(\tilde{p}_\tau^*) \leq \tilde{F}_t(\tilde{p}_t^*) \quad \text{(including fake action } \tilde{p}_0^* \text{ at } \tau = 0\text{)}$$

- Then, using the claim above,

$$\sum_{t=0}^{T} f_t(\tilde{p}_t^*) \leq \min_{p\in S} \sum_{t=0}^{T} f_t(p) \leq \max_{p\in S} f_0(p) + \min_{p\in S} \sum_{t=1}^{T} f_t(p)$$

Regret of Be the Regularized Leader

$$\text{Regret}_{BTRL}(T) \leq \frac{1}{\eta}\left(\max_{p\in S} R(p) - \min_{p\in S} R(p)\right)$$

Proof :

- Let $f_0(p) = R(p)/\eta$ and $\tilde{p}_0^* = \arg\min_{p\in S} R(p)/\eta$.
- Using induction on t, we show that for all $t \geq 1$,

$$\sum_{\tau=0}^{t} f_\tau(\tilde{p}_\tau^*) \leq \tilde{F}_t(\tilde{p}_t^*) \quad \text{(including fake action } \tilde{p}_0^* \text{ at } \tau = 0\text{)}$$

- Then, using the claim above,

$$\sum_{t=0}^{T} f_t(\tilde{p}_t^*) \leq \min_{p\in S} \sum_{t=0}^{T} f_t(p) \leq \max_{p\in S} f_0(p) + \min_{p\in S} \sum_{t=1}^{T} f_t(p)$$

- Hence, by rearranging:

$$\sum_{t=1}^{T} f_t(\tilde{p}_t^*) - \min_{p\in S} \sum_{t=1}^{T} f_t(p) \leq \max_{p\in S} R(p)/\eta - \min_{p\in S} R(p)/\eta$$

Regret of Follow the Regularized Leader

$$\text{Regret}_{FTRL}(T) \leq \eta \cdot L^2 \cdot T + \frac{\left(\max_{p \in S} R(p) - \min_{p \in S} R(p)\right)}{\eta}$$

Regret of Follow the Regularized Leader

$$\text{Regret}_{FTRL}(T) \le \eta \cdot L^2 \cdot T + \frac{\left(\max_{p \in S} R(p) - \min_{p \in S} R(p)\right)}{\eta}$$

Let $R^* = \max_{p \in S} R(p) - \min_{p \in S} R(p)$. Setting $\eta = \sqrt{R^*/T}$ yields $\text{Regret}_{FTRL}(T) \le (L^2 + 1)\sqrt{R^* T}$

Regret of Follow the Regularized Leader

$$\text{Regret}_{FTRL}(T) \leq \eta \cdot L^2 \cdot T + \frac{\left(\max_{p \in S} R(p) - \min_{p \in S} R(p)\right)}{\eta}$$

Regret of Follow the Regularized Leader

$$\text{Regret}_{FTRL}(T) \leq \eta \cdot L^2 \cdot T + \frac{\left(\max_{p \in S} R(p) - \min_{p \in S} R(p)\right)}{\eta}$$

Let $R^* = \max_{p \in S} R(p) - \min_{p \in S} R(p)$.
Setting $\eta = \sqrt{R^*/T}$ yields $\text{Regret}_{FTRL}(T) \leq (L^2 + 1)\sqrt{R^* T}$

Regret of Follow the Regularized Leader

$$\text{Regret}_{FTRL}(T) \leq \eta \cdot L^2 \cdot T + \frac{\left(\max_{p\in S} R(p) - \min_{p\in S} R(p)\right)}{\eta}$$

Let $R^* = \max_{p\in S} R(p) - \min_{p\in S} R(p)$.
Setting $\eta = \sqrt{R^*/T}$ yields $\text{Regret}_{FTRL}(T) \leq (L^2+1)\sqrt{R^* T}$

Multiplicative weight updates:

- Negative entropy $E^-(p) = \sum_{i=1}^{d} p_i \ln(p_i)$ is 1-strongly convex wrt L_1 norm.
- Using $E^-(p)$ as regularizer, results in the following update rule for linear losses $f_t(p) = \langle p, \ell_t \rangle$:

$$p_{t+1}(i) = p_t(i) \cdot e^{-\eta \ell_t(i)} \approx p_t(i)(1 - \eta \ell_t(i))$$

Regret of Follow the Regularized Leader

$$\text{Regret}_{FTRL}(T) \leq \eta \cdot L^2 \cdot T + \frac{\left(\max_{p\in S} R(p) - \min_{p\in S} R(p)\right)}{\eta}$$

Let $R^* = \max_{p\in S} R(p) - \min_{p\in S} R(p)$.
Setting $\eta = \sqrt{R^*/T}$ yields $\text{Regret}_{FTRL}(T) \leq (L^2+1)\sqrt{R^* T}$

Multiplicative weight updates:

- Negative entropy $E^-(p) = \sum_{i=1}^{d} p_i \ln(p_i)$ is 1-strongly convex wrt L_1 norm.
- Using $E^-(p)$ as regularizer, results in the following update rule for linear losses $f_t(p) = \langle p, \ell_t \rangle$:

$$p_{t+1}(i) = p_t(i) \cdot e^{-\eta \ell_t(i)} \approx p_t(i)(1 - \eta \ell_t(i))$$

- If $\ell_t \in [0,1]^d$, setting $\eta = \sqrt{ln(d)/T}$, yields regret $2\sqrt{T \ln(d)}$