Online Learning and Online Convex Optimization

Dimitris Fotakis

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

Agnostic PAC Learning

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

Mostly **binary classification** $\mathcal{Y} = \{-1, 1\}$

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

Agnostic PAC Learning

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

Mostly **binary classification** $\mathcal{Y} = \{-1, 1\}$

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

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

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

In general, (possibly surrogate) **loss function** $\ell : \mathcal{H} \times (\mathcal{X} \times \mathcal{Y}) \to \mathbb{R}_{\geq 0}$:

1. 0-1 loss: $\ell(h, (x, y)) = \begin{cases} 1 & \text{if } h(x) \neq y \\ 0 & \text{if } h(x) = y \end{cases}$
2. Absolute-value loss: $\ell(h, (x, y)) = |h(x) - y|$
3. Cost-sensitive loss: $\ell(h, (x, y)) = \mathrm{Cost}(h(x), y)$.

Agnostic PAC Learning

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

Mostly **binary classification** $\mathcal{Y} = \{-1, 1\}$

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

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

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

In general, (possibly surrogate) **loss function** $\ell : \mathcal{H} \times (\mathcal{X} \times \mathcal{Y}) \to \mathbb{R}_{\geq 0}$:

1. 0-1 loss: $\ell(h, (x, y)) = \begin{cases} 1 & \text{if } h(x) \neq y \\ 0 & \text{if } h(x) = y \end{cases}$
2. Absolute-value loss: $\ell(h, (x, y)) = |h(x) - y|$
3. Cost-sensitive loss: $\ell(h, (x, y)) = \mathrm{Cost}(h(x), y)$.
4. Squared loss (linear regression): $\ell(h, (x, y)) = (h(x) - y)^2$
5. Hinge loss (SVM): $\ell(h, (x, y)) = \max\{1 - y \cdot h(x), 0\}$
6. Exponential loss (logistic regression): $\ell(h, (x, y)) = \ln(1 + e^{-y \cdot h(x)})$

Agnostic PAC Learning

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

Mostly **binary classification** $\mathcal{Y} = \{-1, 1\}$

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

In general, (possibly surrogate) **loss function** $\ell : \mathcal{H} \times (\mathcal{X} \times \mathcal{Y}) \to \mathbb{R}_{\geq 0}$:

1. 0-1 loss: $\ell(h, (x, y)) = \begin{cases} 1 & \text{if } h(x) \neq y \\ 0 & \text{if } h(x) = y \end{cases}$
2. Absolute-value loss: $\ell(h, (x, y)) = |h(x) - y|$
3. Cost-sensitive loss: $\ell(h, (x, y)) = \mathrm{Cost}(h(x), y)$.
4. Squared loss (linear regression): $\ell(h, (x, y)) = (h(x) - y)^2$
5. Hinge loss (SVM): $\ell(h, (x, y)) = \max\{1 - y \cdot h(x), 0\}$
6. Exponential loss (logistic regression): $\ell(h, (x, y)) = \ln(1 + e^{-y \cdot h(x)})$

Loss of hypothesis $h \in \mathcal{H}$ (wrt. ℓ): $L_{\mathcal{D}}(h) = \mathbb{E}\mathrm{xp}_{(x,y)\sim\mathcal{D}}[\ell(h, (x, y))]$

Agnostic PAC Learning

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 any $\mathcal{D}$,

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

Agnostic PAC Learning

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 any $\mathcal{D}$,

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

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

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

Training set S ε-representative if $\forall h \in \mathcal{H}$, $|L_S(h) - L_{\mathcal{D}}(h)| \leq \varepsilon$.

$L_{\mathcal{D}}(\mathrm{ERM}_{\mathcal{H}}(S)) = \varepsilon_{\mathrm{app}} + \varepsilon_{\mathrm{est}}$

- $\varepsilon_{\mathrm{app}}$ due to restriction to (possibly too simple) hypothesis class $\mathcal{H}$
- $\varepsilon_{\mathrm{est}}$ due to misrepresentation of S wrt class $\mathcal{H}$

Agnostic PAC Learning

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 any $\mathcal{D}$,

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

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

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

Training set S ε-representative if $\forall h \in \mathcal{H}$, $|L_S(h) - L_{\mathcal{D}}(h)| \leq \varepsilon$.

$L_{\mathcal{D}}(\mathrm{ERM}_{\mathcal{H}}(S)) = \varepsilon_{\mathrm{app}} + \varepsilon_{\mathrm{est}}$

- $\varepsilon_{\mathrm{app}}$ due to restriction to (possibly too simple) hypothesis class $\mathcal{H}$
- $\varepsilon_{\mathrm{est}}$ due to misrepresentation of S wrt class $\mathcal{H}$

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.

Agnostic PAC Learning and Convex Optimization

ERM on representative training set S wrt (surrogate) **convex loss** ℓ:
convex optimization !

$L_{\mathcal{D}}(\mathrm{ERM}_{\mathcal{H}}(S)) = \varepsilon_{\text{app}} + \varepsilon_{\text{opt}} + \varepsilon_{\text{est}}$

- $\varepsilon_{\text{opt}} = |\min_h L_{\mathcal{D}}^{\text{sur}}(h) - \min_h L_{\mathcal{D}}^{0\text{-}1}(h)|$ (estimation of 0-1 loss by ℓ).

Agnostic PAC Learning and Convex Optimization

ERM on representative training set S wrt (surrogate) **convex loss** ℓ: **convex optimization** !

$L_{\mathcal{D}}(\mathrm{ERM}_{\mathcal{H}}(S)) = \varepsilon_{\mathrm{app}} + \varepsilon_{\mathrm{opt}} + \varepsilon_{\mathrm{est}}$

- $\varepsilon_{\mathrm{opt}} = |\min_h L_{\mathcal{D}}^{\mathrm{sur}}(h) - \min_h L_{\mathcal{D}}^{0\text{-}1}(h)|$ (estimation of 0-1 loss by ℓ).

(Projected) Gradient Descent of convex $f : S \to \mathbb{R}$ on convex $\subseteq \mathbb{R}^d$:

$$\boldsymbol{x}_{t+1} = \arg\min_{\boldsymbol{x} \in S} \left\| \left[\boldsymbol{x}_t - \eta \nabla f(\boldsymbol{x}_t)\right] - \boldsymbol{x} \right\|$$

Agnostic PAC Learning and Convex Optimization

ERM on representative training set S wrt (surrogate) **convex loss** ℓ: **convex optimization** !

$L_{\mathcal{D}}(\mathrm{ERM}_{\mathcal{H}}(S)) = \varepsilon_{\mathrm{app}} + \varepsilon_{\mathrm{opt}} + \varepsilon_{\mathrm{est}}$

- $\varepsilon_{\mathrm{opt}} = |\min_h L_{\mathcal{D}}^{\mathrm{sur}}(h) - \min_h L_{\mathcal{D}}^{0\text{-}1}(h)|$ (estimation of 0-1 loss by ℓ).

(Projected) Gradient Descent of convex $f : S \to \mathbb{R}$ on convex $\subseteq \mathbb{R}^d$:

$$\boldsymbol{x}_{t+1} = \arg\min_{\boldsymbol{x} \in S} \left\| \left[\boldsymbol{x}_t - \eta \nabla f(\boldsymbol{x}_t)\right] - \boldsymbol{x} \right\|$$

Theorem : For step size $\eta = \varepsilon / G^2$ and #steps $T \geq D^2 G^2 / \varepsilon^2$,

$$f\left(\sum_t \boldsymbol{x}_t / T\right) \leq f(\boldsymbol{x}^*) + \varepsilon$$

Agnostic PAC Learning and Convex Optimization

ERM on representative training set S wrt (surrogate) **convex loss** ℓ: **convex optimization** !

$L_{\mathcal{D}}(\mathrm{ERM}_{\mathcal{H}}(S)) = \varepsilon_{\mathrm{app}} + \varepsilon_{\mathrm{opt}} + \varepsilon_{\mathrm{est}}$

- $\varepsilon_{\mathrm{opt}} = \left| \min_h L_{\mathcal{D}}^{\mathrm{sur}}(h) - \min_h L_{\mathcal{D}}^{0\text{-}1}(h) \right|$ (estimation of 0-1 loss by ℓ).

(Projected) Gradient Descent of convex $f : S \to \mathbb{R}$ on convex $\subseteq \mathbb{R}^d$:

$$\boldsymbol{x}_{t+1} = \arg\min_{x \in S} \left\| \left[\boldsymbol{x}_t - \eta \nabla f(\boldsymbol{x}_t) \right] - \boldsymbol{x} \right\|$$

Theorem : For step size $\eta = \varepsilon / G^2$ and #steps $T \geq D^2 G^2 / \varepsilon^2$,

$$f\left(\sum_t \boldsymbol{x}_t / T \right) \leq f(\boldsymbol{x}^*) + \varepsilon$$

Gradient Descent step on all training data is **too expensive** : **online learning** through online convex optimization!

Online Learning as a Game

We fix $\mathcal{H}$ and loss ℓ (known to algo) [and $\mathcal{D}$ (unknown to algo)].

On each step $t = 1, \ldots, T$:

1. Learner picks hypothesis $h_t \in \mathcal{H}$
2. Training example $(\boldsymbol{x}_t, y_t)$ is chosen (may be from $\mathcal{D}$, but even by adversary)
3. Learner incurs loss $\ell_t(h_t, (\boldsymbol{x}_t, y_t))$

Online Learning as a Game

We fix $\mathcal{H}$ and loss ℓ (known to algo) [and $\mathcal{D}$ (unknown to algo)].

On each step $t = 1, \ldots, T$:

1. Learner picks hypothesis $h_t \in \mathcal{H}$
2. Training example $(\boldsymbol{x}_t, y_t)$ is chosen (may be from $\mathcal{D}$, but even by adversary)
3. Learner incurs loss $\ell_t(h_t, (\boldsymbol{x}_t, y_t))$

Goal is to minimize **regret**:

$$\mathrm{Regret}(T) = \sup_{(\boldsymbol{x}_1,y_1),\ldots(\boldsymbol{x}_T,y_T)} \left(\sum_{t=1}^{T} \ell_t(h_t, (\boldsymbol{x}_t, y_t)) - \min_{h^* \in \mathcal{H}} \sum_{t=1}^{T} \ell_t(h^*, (\boldsymbol{x}_t, y_t)) \right)$$

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

Online Learning as a Game

We fix $\mathcal{H}$ and loss ℓ (known to algo) [and $\mathcal{D}$ (unknown to algo)].

On each step $t = 1, \ldots, T$:

1. Learner picks hypothesis $h_t \in \mathcal{H}$
2. Training example $(\boldsymbol{x}_t, y_t)$ is chosen (may be from $\mathcal{D}$, but even by adversary)
3. Learner incurs loss $\ell_t(h_t, (\boldsymbol{x}_t, y_t))$

Goal is to minimize **regret**:

$$\text{Regret}(T) = \sup_{(\boldsymbol{x}_1,y_1),\ldots(\boldsymbol{x}_T,y_T)} \left(\sum_{t=1}^{T} \ell_t(h_t, (\boldsymbol{x}_t, y_t)) - \min_{h^* \in \mathcal{H}} \sum_{t=1}^{T} \ell_t(h^*, (\boldsymbol{x}_t, y_t)) \right)$$

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

Any no-regret online algorithm can be used for learning!

We focus on **regret minimization** for this and next lecture.

Online Learning: Basic Setting

Two actions: H and L (binary classification).

On each day $t = 1, \ldots, T$:

1. Learner **picks action** $i_t \in \{H, L\}$
2. Adversary **picks loss** vector $\boldsymbol{\ell}_t = (\ell_t^H, \ell_t^L) \in [0, 1]^2$
3. Learner learns $\boldsymbol{\ell}_t$ and **incurs loss** $\ell_t^{i_t}$

Online Learning: Basic Setting

Two actions: H and L (binary classification).

On each day $t = 1, \dots, T$:

1. Learner **picks action** $i_t \in \{H, L\}$
2. Adversary **picks loss** vector $\boldsymbol{\ell}_t = (\ell_t^H, \ell_t^L) \in [0,1]^2$
3. Learner learns $\boldsymbol{\ell}_t$ and **incurs loss** $\ell_t^{i_t}$

Goal is to minimize **regret** (loss wrt. **best fixed** action in hindsight):

$$\text{Regret}(T) = \sup_{\boldsymbol{\ell}_1, \dots, \boldsymbol{\ell}_T} \left(\sum_{t=1}^{T} \ell_t^{i_t} - \min_{i \in \{H,T\}} \sum_{t=1}^{T} \ell_t^{i} \right)$$

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

Online Learning: Follow the Leader

Follow the Leader (FTL):

$$i_t = \arg\min_{i \in \{H,L\}} \sum_{\tau=1}^{t-1} \ell_\tau^i$$

Online Learning: Follow the Leader

Follow the Leader (FTL):

$$i_t = \arg \min_{i \in \{H,L\}} \sum_{\tau=1}^{t-1} \ell_\tau^i$$

Two obvious concerns with FTL:

1. **Deterministic** action choice, given the past (randomness always helps against the unknown).
2. Action choices can be very **unstable** (different choice each day).

Online Learning: Follow the Leader

Follow the Leader (FTL):

$$i_t = \arg\min_{i \in \{H,L\}} \sum_{\tau=1}^{t-1} \ell_\tau^i$$

Two obvious concerns with FTL:

1. **Deterministic** action choice, given the past (randomness always helps against the unknown).
2. Action choices can be very **unstable** (different choice each day).

Lower bound : Any deterministic algorithm has **linear**, i.e., $\Omega(T)$, regret.

Online Learning: Follow the Leader

Follow the Leader (FTL):

$$i_t = \arg \min_{i \in \{H,L\}} \sum_{\tau=1}^{t-1} \ell_\tau^i$$

Two obvious concerns with FTL:

1. **Deterministic** action choice, given the past (randomness always helps against the unknown).
2. Action choices can be very **unstable** (different choice each day).

Lower bound : Any deterministic algorithm has **linear**, i.e., $\Omega(T)$, regret.

Proof : loss for action i_t (chosen by the algorithm) $= 1$, and loss for other action $= 0$.

Any deterministic algorithm incurs loss $= T$, while best action incures loss $\leq T/2$.

Online Learning: Randomization

Two actions: H and L (binary classification).

On each day $t = 1, \ldots, T$:

1. Learner picks action H with probability p_t (and L with probability $1 - p_t$).
2. Adversary picks loss vector $\boldsymbol{\ell}_t = (\ell_t^H, \ell_t^L) \in [0, 1]^2$
3. Learner learns $\boldsymbol{\ell}_t$ and incurs **expected loss**

$$f(p_t; \boldsymbol{\ell}_t) = p_t \ell_t^H + (1 - p_t) \ell_t^L$$

Online Learning: Randomization

Two actions: H and L (binary classification).

On each day $t = 1, \ldots, T$:

1. Learner picks action H with probability p_t (and L with probability $1 - p_t$).
2. Adversary picks loss vector $\boldsymbol{\ell}_t = (\ell_t^H, \ell_t^L) \in [0,1]^2$
3. Learner learns $\boldsymbol{\ell}_t$ and incurs **expected loss**

$$f(p_t; \boldsymbol{\ell}_t) = p_t \ell_t^H + (1 - p_t)\ell_t^L$$

Goal is to minimize **expected regret** :

$$\text{Exp-Regret}(T) = \sup_{\boldsymbol{\ell}_1, \ldots, \boldsymbol{\ell}_T} \left(\sum_{t=1}^{T} f(p_t; \boldsymbol{\ell}_t) - \min_{p \in [0,1]} \sum_{t=1}^{T} f(p; \boldsymbol{\ell}_t) \right)$$

Randomization potentially allows for improved **stability** .

Online Learning: (Randomized) Follow the Leader

Follow the Leader (FTL):

$$p_t = \arg\min_{p \in [0,1]} \sum_{\tau=1}^{t-1} f(p; \ell_\tau) = \arg\min_{p \in [0,1]} F_{t-1}(p)$$

Online Learning: (Randomized) Follow the Leader

Follow the Leader (FTL):

$$p_t = \arg\min_{p \in [0,1]} \sum_{\tau=1}^{t-1} f(p; \ell_\tau) = \arg\min_{p \in [0,1]} F_{t-1}(p)$$

Is randomized FTL **really different** from deterministic FTL?

Online Learning: (Randomized) Follow the Leader

Follow the Leader (FTL):

$$p_t = \arg\min_{p\in[0,1]} \sum_{\tau=1}^{t-1} f(p;\ell_\tau) = \arg\min_{p\in[0,1]} F_{t-1}(p)$$

Is randomized FTL **really different** from deterministic FTL?

Theorem : For any loss sequence $\ell_1, \ldots, \ell_T$, FTL has:

$$\text{Exp-Regret}_{FTL}(T) = \underbrace{\sum_{t=1}^{T} f(p_t;\ell_t) - \min_{p\in[0,1]} \sum_{t=1}^{T} f(p;\ell_t)}_{\text{expected regret}} \le \underbrace{\sum_{t=1}^{T} |p_t - p_{t+1}|}_{\text{instability}}$$

Online Learning: (Randomized) Follow the Leader

Follow the Leader (FTL):

$$p_t = \arg\min_{p\in[0,1]} \sum_{\tau=1}^{t-1} f(p;\boldsymbol{\ell}_\tau) = \arg\min_{p\in[0,1]} F_{t-1}(p)$$

Is randomized FTL **really different** from deterministic FTL?

Theorem : For any loss sequence $\boldsymbol{\ell}_1, \ldots, \boldsymbol{\ell}_T$, FTL has:

$$\text{Exp-Regret}_{FTL}(T) = \underbrace{\sum_{t=1}^{T} f(p_t;\boldsymbol{\ell}_t) - \min_{p\in[0,1]} \sum_{t=1}^{T} f(p;\boldsymbol{\ell}_t)}_{\text{expected regret}} \leq \underbrace{\sum_{t=1}^{T} |p_t - p_{t+1}|}_{\text{instability}}$$

For the analysis, we define **Be the Leader** (BTL):

$$p_t^* = \arg\min_{p\in[0,1]} \sum_{\tau=1}^{t} f(p;\boldsymbol{\ell}_\tau) = \arg\min_{p\in[0,1]} F_t(p)$$

Regret of Be the Leader

Lemma : For any loss sequence $\boldsymbol{\ell}_1, \ldots, \boldsymbol{\ell}_T$, $\text{Regret}_{BTL}(T) \leq 0$

Regret of Be the Leader

Lemma: For any loss sequence $\boldsymbol{\ell}_1, \ldots, \boldsymbol{\ell}_T$, $\text{Regret}_{BTL}(T) \leq 0$

By **induction** on t, we show that for any $t \geq 1$:

$$\underbrace{\sum_{\tau=1}^{t} f(p_\tau^*; \boldsymbol{\ell}_t)}_{\text{loss of BTL up to } t} \leq \underbrace{\min_{p \in [0,1]} F_t(p)}_{\text{loss of best fixed action up to } t} = \underbrace{F_t(p_t^*)}_{\text{by definition of } p_t^*}$$

Regret of Be the Leader

Lemma: For any loss sequence $\boldsymbol{\ell}_1, \dots, \boldsymbol{\ell}_T$, $\text{Regret}_{BTL}(T) \le 0$

By **induction** on t, we show that for any $t \ge 1$:

$$\underbrace{\sum_{\tau=1}^{t} f(p_\tau^*; \boldsymbol{\ell}_t)}_{\text{loss of BTL up to } t} \le \underbrace{\min_{p \in [0,1]} F_t(p)}_{\text{loss of best fixed action up to } t} = \underbrace{F_t(p_t^*)}_{\text{by definition of } p_t^*}$$

$$\sum_{\tau=1}^{t+1} f(p_\tau^*; \boldsymbol{\ell}_\tau) = f(p_{t+1}^*; \boldsymbol{\ell}_{t+1}) + \sum_{\tau=1}^{t} f(p_\tau^*; \boldsymbol{\ell}_\tau)$$

Regret of Be the Leader

Lemma : For any loss sequence $\boldsymbol{\ell}_1, \ldots, \boldsymbol{\ell}_T$, $\text{Regret}_{BTL}(T) \leq 0$

By **induction** on t, we show that for any $t \geq 1$:

$$\underbrace{\sum_{\tau=1}^{t} f(p_\tau^*; \boldsymbol{\ell}_t)}_{\text{loss of BTL up to } t} \leq \underbrace{\min_{p \in [0,1]} F_t(p)}_{\text{loss of best fixed action up to } t} = \underbrace{F_t(p_t^*)}_{\text{by definition of } p_t^*}$$

$$\begin{aligned}\sum_{\tau=1}^{t+1} f(p_\tau^*; \boldsymbol{\ell}_\tau) &= f(p_{t+1}^*; \boldsymbol{\ell}_{t+1}) + \sum_{\tau=1}^{t} f(p_\tau^*; \boldsymbol{\ell}_\tau) \\ &\leq f(p_{t+1}^*; \boldsymbol{\ell}_{t+1}) + \min_{p \in [0,1]} F_t(p) \qquad \text{induction hypth.}\end{aligned}$$

Regret of Be the Leader

Lemma : For any loss sequence $\boldsymbol{\ell}_1, \ldots, \boldsymbol{\ell}_T$, $\text{Regret}_{BTL}(T) \leq 0$

By **induction** on t, we show that for any $t \geq 1$:

$$\underbrace{\sum_{\tau=1}^{t} f(p_\tau^*; \boldsymbol{\ell}_t)}_{\text{loss of BTL up to } t} \leq \underbrace{\min_{p \in [0,1]} F_t(p)}_{\text{loss of best fixed action up to } t} = \underbrace{F_t(p_t^*)}_{\text{by definition of } p_t^*}$$

$$\begin{aligned}
\sum_{\tau=1}^{t+1} f(p_\tau^*; \boldsymbol{\ell}_\tau) &= f(p_{t+1}^*; \boldsymbol{\ell}_{t+1}) + \sum_{\tau=1}^{t} f(p_\tau^*; \boldsymbol{\ell}_\tau) \\
&\leq f(p_{t+1}^*; \boldsymbol{\ell}_{t+1}) + \min_{p \in [0,1]} F_t(p) && \text{induction hypth.} \\
&\leq f(p_{t+1}^*; \boldsymbol{\ell}_{t+1}) + F_t(p_{t+1}^*) && F_t(p_t^*) \leq F_t(p_{t+1}^*)
\end{aligned}$$

Regret of Be the Leader

Lemma : For any loss sequence $\boldsymbol{\ell}_1, \ldots, \boldsymbol{\ell}_T$, $\text{Regret}_{BTL}(T) \leq 0$

By **induction** on t, we show that for any $t \geq 1$:

$$\underbrace{\sum_{\tau=1}^{t} f(p^*_\tau; \boldsymbol{\ell}_t)}_{\text{loss of BTL up to } t} \leq \underbrace{\min_{p \in [0,1]} F_t(p)}_{\text{loss of best fixed action up to } t} = \underbrace{F_t(p^*_t)}_{\text{by definition of } p^*_t}$$

$$\begin{aligned}
\sum_{\tau=1}^{t+1} f(p^*_\tau; \boldsymbol{\ell}_\tau) &= f(p^*_{t+1}; \boldsymbol{\ell}_{t+1}) + \sum_{\tau=1}^{t} f(p^*_\tau; \boldsymbol{\ell}_\tau) & \\
&\leq f(p^*_{t+1}; \boldsymbol{\ell}_{t+1}) + \min_{p \in [0,1]} F_t(p) & \text{induction hypth.} \\
&\leq f(p^*_{t+1}; \boldsymbol{\ell}_{t+1}) + F_t(p^*_{t+1}) & F_t(p^*_t) \leq F_t(p^*_{t+1}) \\
&= F_{t+1}(p^*_{t+1}) & \text{by dfn of } F_{t+1}(p)
\end{aligned}$$

Regret of FTL Against BTL

Lemma : For any loss sequence $\boldsymbol{\ell}_1, \ldots, \boldsymbol{\ell}_T$,

$$\text{Regret}_{FTL}(T) \le \text{Regret}_{BTL}(T) + \underbrace{\sum_{t=1}^{T} |p_t - p_{t+1}|}_{\text{instability}}$$

Regret of FTL Against BTL

Lemma : For any loss sequence $\boldsymbol{\ell}_1, \ldots, \boldsymbol{\ell}_T$,

$$\mathrm{Regret}_{FTL}(T) \leq \mathrm{Regret}_{BTL}(T) + \underbrace{\sum_{t=1}^{T} |p_t - p_{t+1}|}_{\text{instability}}$$

$$\sum_{t=1}^{T} f(p_t; \boldsymbol{\ell}_t) = \sum_{t=1}^{T} f(p_t^*; \boldsymbol{\ell}_t) + \sum_{t=1}^{T} \Big(f(p_t; \boldsymbol{\ell}_t) - f(p_t^*; \boldsymbol{\ell}_t) \Big)$$

Regret of FTL Against BTL

Lemma : For any loss sequence $\boldsymbol{\ell}_1, \ldots, \boldsymbol{\ell}_T$,

$$\text{Regret}_{FTL}(T) \leq \text{Regret}_{BTL}(T) + \underbrace{\sum_{t=1}^{T} |p_t - p_{t+1}|}_{\text{instability}}$$

$$\begin{aligned}
\sum_{t=1}^{T} f(p_t; \boldsymbol{\ell}_t) &= \sum_{t=1}^{T} f(p_t^*; \boldsymbol{\ell}_t) + \sum_{t=1}^{T} \Big(f(p_t; \boldsymbol{\ell}_t) - f(p_t^*; \boldsymbol{\ell}_t) \Big) \\
&= \sum_{t=1}^{T} f(p_t^*; \boldsymbol{\ell}_t) + \sum_{t=1}^{T} (p_t - p_t^*)(\ell_t^H - \ell_t^L) \qquad \text{by dfn of } f(p_t; \boldsymbol{\ell}_t)
\end{aligned}$$

Regret of FTL Against BTL

Lemma : For any loss sequence $\boldsymbol{\ell}_1, \ldots, \boldsymbol{\ell}_T$,

$$\text{Regret}_{FTL}(T) \leq \text{Regret}_{BTL}(T) + \underbrace{\sum_{t=1}^{T} |p_t - p_{t+1}|}_{\text{instability}}$$

$$\begin{aligned}
\sum_{t=1}^{T} f(p_t; \boldsymbol{\ell}_t) &= \sum_{t=1}^{T} f(p_t^*; \boldsymbol{\ell}_t) + \sum_{t=1}^{T} \Big(f(p_t; \boldsymbol{\ell}_t) - f(p_t^*; \boldsymbol{\ell}_t) \Big) \\
&= \sum_{t=1}^{T} f(p_t^*; \boldsymbol{\ell}_t) + \sum_{t=1}^{T} (p_t - p_t^*)(\ell_t^H - \ell_t^L) && \text{by dfn of } f(p_t; \boldsymbol{\ell}_t) \\
&\leq \sum_{t=1}^{T} f(p_t^*; \boldsymbol{\ell}_t) + \sum_{t=1}^{T} |p_t - p_t^*| && \text{losses } \boldsymbol{\ell}_t \in [0,1]^2
\end{aligned}$$

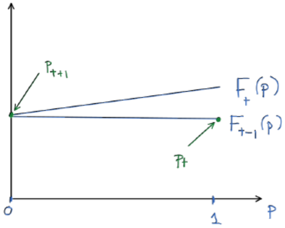
Regret of FTL Against BTL

Lemma : For any loss sequence $\boldsymbol{\ell}_1, \ldots, \boldsymbol{\ell}_T$,

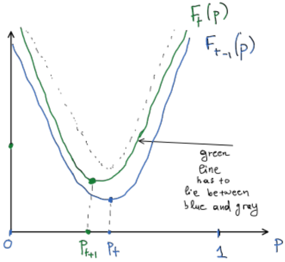
$$\text{Regret}_{FTL}(T) \leq \text{Regret}_{BTL}(T) + \underbrace{\sum_{t=1}^{T} |p_t - p_{t+1}|}_{\text{instability}}$$

$$\begin{aligned}
\sum_{t=1}^{T} f(p_t; \boldsymbol{\ell}_t) &= \sum_{t=1}^{T} f(p_t^*; \boldsymbol{\ell}_t) + \sum_{t=1}^{T} \Big(f(p_t; \boldsymbol{\ell}_t) - f(p_t^*; \boldsymbol{\ell}_t) \Big) \\
&= \sum_{t=1}^{T} f(p_t^*; \boldsymbol{\ell}_t) + \sum_{t=1}^{T} (p_t - p_t^*)(\ell_t^H - \ell_t^L) && \text{by dfn of } f(p_t; \boldsymbol{\ell}_t) \\
&\leq \sum_{t=1}^{T} f(p_t^*; \boldsymbol{\ell}_t) + \sum_{t=1}^{T} |p_t - p_t^*| && \text{losses } \boldsymbol{\ell}_t \in [0,1]^2 \\
&= \sum_{t=1}^{T} f(p_t^*; \boldsymbol{\ell}_t) + \sum_{t=1}^{T} |p_t - p_{t+1}| && \text{by dfn, } p_t^* = p_{t+1}
\end{aligned}$$

Convexity and Stability

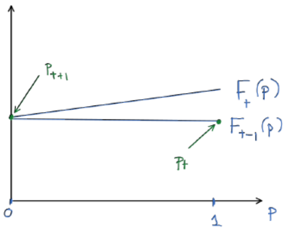


(a) Two linear functions that are close to each other can have very far minima.

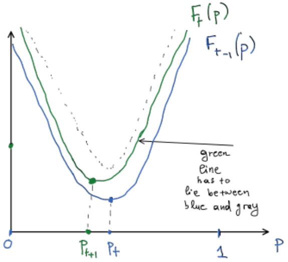


(b) For convex functions, closeness of the functions implies closeness of their minima.

Convexity and Stability



(a) Two linear functions that are close to each other can have very far minima.

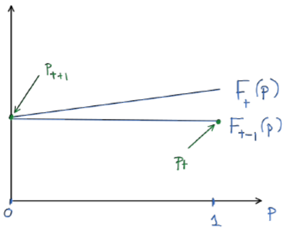


(b) For convex functions, closeness of the functions implies closeness of their minima.

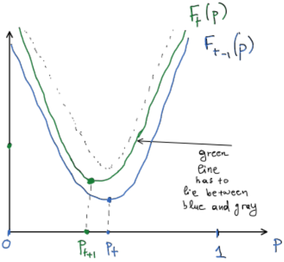
$1/\eta$-**strongly convex** function $f : S \to \mathbb{R}$ wrt norm $\|\cdot\|$, if $\forall\, x, y \in S$:

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

Convexity and Stability



(a) Two linear functions that are close to each other can have very far minima.

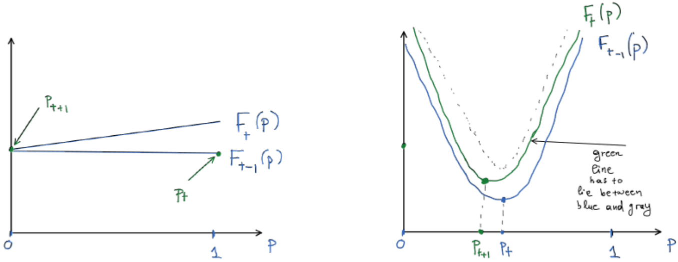


(b) For convex functions, closeness of the functions implies closeness of their minima.

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

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

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 $\|\cdot\|$.



(a) Two linear functions that are close to each other can have very far minima.

(b) For convex functions, closeness of the functions implies closeness of their minima.

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

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

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.

Convexity and Stability

Functions $f, g : [0,1] \to \mathbb{R}$ be $1/\eta$-strongly convex and $h(x) = g(x) - f(x)$ be L-Lipschitz.

Then, $|p_f - p_g| \leq \eta \cdot L$, with p_f, p_g minimizers of f, g.

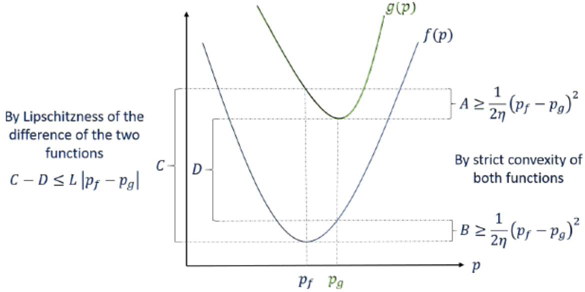


Figure 3: The proof of Lemma 3 follows immediately by noting that $C - D = A + B$ in the above figure, together with the fact that $C - D \leq L|p_f - p_g|$ by Lipschitzness of the difference of the two functions and $A + B \geq \frac{1}{\eta}(p_f - p_g)^2$ by the strict convexity of the two functions.

Convexity Through Regularization

If **cumulative loss** $F_t(\cdot)$ was $1/\eta$-strongly convex (for all t), stability could be bounded as:

$$\sum_{t=1}^{T} |p_t - p_{t+1}| \leq \eta \cdot T \,,$$

because $F_t(p) - F_{t-1}(p) = f(p; \boldsymbol{\ell}_t)$ is 1-Lipschitz (due to $\boldsymbol{\ell}_t \in [0, 1]^2$).

Convexity Through Regularization

If **cumulative loss** $F_t(\cdot)$ was $1/\eta$-strongly convex (for all t), stability could be bounded as:

$$\sum_{t=1}^{T} |p_t - p_{t+1}| \leq \eta \cdot T \,,$$

because $F_t(p) - F_{t-1}(p) = f(p; \boldsymbol{\ell}_t)$ is 1-Lipschitz (due to $\boldsymbol{\ell}_t \in [0,1]^2$).

But our cumulative loss $F_t(\cdot)$ **is not** strongly convex!

Convexity Through Regularization

If **cumulative loss** $F_t(\cdot)$ was $1/\eta$-strongly convex (for all t), stability could be bounded as:

$$\sum_{t=1}^{T} |p_t - p_{t+1}| \leq \eta \cdot T \,,$$

because $F_t(p) - F_{t-1}(p) = f(p; \boldsymbol{\ell}_t)$ is 1-Lipschitz (due to $\boldsymbol{\ell}_t \in [0,1]^2$).

But our cumulative loss $F_t(\cdot)$ **is not** strongly convex!

Make it strongly convex through **regularization** !

$\tilde{F}_t(p) = F_t(p) + R(p)/\eta$, where $R(\cdot)$ any 1-strongly convex function:

- $R(p) = p^2/2$
- $R(p) = p\ln(p) + (1-p)\ln(1-p)$
- $R(p) = \ln(\frac{p}{1-p})$

Follow / Be the Regularized Leader

$F_t(p) = \sum_{\tau=1}^{t} f(p; \boldsymbol{\ell}_\tau)$ and $\tilde{F}_t(p) = \sum_{\tau=1}^{t} f(p; \boldsymbol{\ell}_\tau) + R(p)/\eta$

FTRL: $\tilde{p}_t = \arg\min_{p \in [0,1]} \tilde{F}_{t-1}(p)$

BTRL: $\tilde{p}_t^* = \arg\min_{p \in [0,1]} \tilde{F}_t(p)$

Follow / Be the Regularized Leader

$F_t(p) = \sum_{\tau=1}^{t} f(p;\boldsymbol{\ell}_\tau)$ and $\tilde{F}_t(p) = \sum_{\tau=1}^{t} f(p;\boldsymbol{\ell}_\tau) + R(p)/\eta$

FTRL: $\tilde{p}_t = \arg\min_{p\in[0,1]} \tilde{F}_{t-1}(p)$

BTRL: $\tilde{p}_t^* = \arg\min_{p\in[0,1]} \tilde{F}_t(p)$

Theorem :

$$\text{Regret}_{FTRL}(T) \leq \eta \cdot T + \frac{2\max_{p\in[0,1]} |R(p)|}{\eta}$$

Follow / Be the Regularized Leader

$F_t(p) = \sum_{\tau=1}^{t} f(p; \boldsymbol{\ell}_\tau)$ and $\tilde{F}_t(p) = \sum_{\tau=1}^{t} f(p; \boldsymbol{\ell}_\tau) + R(p)/\eta$

FTRL: $\tilde{p}_t = \arg\min_{p \in [0,1]} \tilde{F}_{t-1}(p)$

BTRL: $\tilde{p}^*_t = \arg\min_{p \in [0,1]} \tilde{F}_t(p)$

Theorem :

$$\text{Regret}_{FTRL}(T) \leq \eta \cdot T + \frac{2 \max_{p \in [0,1]} |R(p)|}{\eta}$$

Let $R^* = \max_{p \in [0,1]} |R(p)|$.

Setting $\eta = \sqrt{2R^*/T}$, we get $\text{Regret}_{FTRL}(T) \leq 2\sqrt{2R^* T}$

Follow / Be the Regularized Leader

$F_t(p) = \sum_{\tau=1}^{t} f(p; \boldsymbol{\ell}_\tau)$ and $\tilde{F}_t(p) = \sum_{\tau=1}^{t} f(p; \boldsymbol{\ell}_\tau) + R(p)/\eta$

FTRL: $\tilde{p}_t = \arg\min_{p \in [0,1]} \tilde{F}_{t-1}(p)$

BTRL: $\tilde{p}_t^* = \arg\min_{p \in [0,1]} \tilde{F}_t(p)$

Theorem :

$$\mathrm{Regret}_{FTRL}(T) \le \eta \cdot T + \frac{2 \max_{p \in [0,1]} |R(p)|}{\eta}$$

Let $R^* = \max_{p \in [0,1]} |R(p)|$.

Setting $\eta = \sqrt{2R^*/T}$, we get $\mathrm{Regret}_{FTRL}(T) \le 2\sqrt{2R^* T}$

Lower bound on $\mathrm{Regret}_A(T)$ for any online (even randomized) optimization algorithm A?

Regret of FTRL Against BTRL

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

Regret of FTRL Against BTRL

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

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

Regret of FTRL Against BTRL

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

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

$$\begin{aligned}\text{Regret}_{FTRL}(T) - \text{Regret}_{BTRL}(T) &= \sum_{t=1}^{T} (f(\tilde{p}_t; \boldsymbol{\ell}_t) - f(\tilde{p}_t^*; \boldsymbol{\ell}_t)) \\ &\leq \sum_{t=1}^{T} |\tilde{p}_t - \tilde{p}_t^*| \\ &= L \sum_{t=1}^{T} |\tilde{p}_t - \tilde{p}_{t+1}|\end{aligned}$$

Regret of Be the Regularized Leader

$$\text{Regret}_{BTRL}(T) \leq \frac{2\max_{p\in[0,1]}|R(p)|}{\eta}$$

Regret of Be the Regularized Leader

$$\text{Regret}_{BTRL}(T) \leq \frac{2\max_{p\in[0,1]}|R(p)|}{\eta}$$

Proof : Let $f_t(p) = f(p; \boldsymbol{\ell}_t)$ for brevity.

- Let $f_0(p) = R(p)/\eta$ and $\tilde{p}_0^* = \arg\min_{p\in[0,1]} R(p)/\eta$.
- 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

$$\text{Regret}_{BTRL}(T) \leq \frac{2\max_{p\in[0,1]}|R(p)|}{\eta}$$

Proof : Let $f_t(p) = f(p; \boldsymbol{\ell}_t)$ for brevity.

- Let $f_0(p) = R(p)/\eta$ and $\tilde{p}_0^* = \arg\min_{p\in[0,1]} R(p)/\eta$.
- 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,

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

Regret of Be the Regularized Leader

$$\text{Regret}_{BTRL}(T) \leq \frac{2\max_{p\in[0,1]}|R(p)|}{\eta}$$

Proof: Let $f_t(p) = f(p; \boldsymbol{\ell}_t)$ for brevity.

- Let $f_0(p) = R(p)/\eta$ and $\tilde{p}_0^* = \arg\min_{p\in[0,1]} R(p)/\eta$.
- Using induction on t, we show that for all $t \geq 1$,

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

- Then, using the claim above,

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

- Hence, by rearranging:

$$\sum_{t=1}^{T} f_t(\tilde{p}_t^*) - \min_{p\in[0,1]} \sum_{t=1}^{T} f_t(p) \leq \max_{p\in[0,1]} R(p)/\eta - \min_{p\in[0,1]} R(p)/\eta \leq \max_{p\in[0,1]} |R(p)|/\eta$$

Regret of Follow the Regularized Leader

Theorem :

$$\text{Regret}_{FTRL}(T) \le \eta \cdot T + \frac{2 \max_{p \in [0,1]} |R(p)|}{\eta}$$

Regret of Follow the Regularized Leader

Theorem :

$$\text{Regret}_{FTRL}(T) \leq \eta \cdot T + \frac{2 \max_{p \in [0,1]} |R(p)|}{\eta}$$

Let $R^* = \max_{p \in [0,1]} |R(p)|$.

Setting $\eta = \sqrt{2R^*/T}$, we get $\text{Regret}_{FTRL}(T) \leq 2\sqrt{2R^* T}$

Regret of Follow the Regularized Leader

Theorem :

$$\text{Regret}_{FTRL}(T) \leq \eta \cdot T + \frac{2 \max_{p \in [0,1]} |R(p)|}{\eta}$$

Let $R^* = \max_{p \in [0,1]} |R(p)|$.

Setting $\eta = \sqrt{2R^*/T}$, we get $\text{Regret}_{FTRL}(T) \leq 2\sqrt{2R^* T}$

Multiplicative weight updates :

- Negative entropy $E^-(p) = p \ln(p) + (1-p)\ln(1-p)$ is 1-strongly convex wrt L_1 norm.
- Using $E^-(p)$ as regularizer, results in the following update rule for expected loss $f(p_t; \boldsymbol{\ell}_t) = p_t \ell_t^H + (1-p_t)\ell_t^L$:

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

Regret of Follow the Regularized Leader

Theorem :

$$\text{Regret}_{FTRL}(T) \le \eta \cdot T + \frac{2 \max_{p \in [0,1]} |R(p)|}{\eta}$$

Let $R^* = \max_{p \in [0,1]} |R(p)|$.

Setting $\eta = \sqrt{2R^*/T}$, we get $\text{Regret}_{FTRL}(T) \le 2\sqrt{2R^* T}$

Multiplicative weight updates :

- Negative entropy $E^-(p) = p \ln(p) + (1-p)\ln(1-p)$ is 1-strongly convex wrt L_1 norm.
- Using $E^-(p)$ as regularizer, results in the following update rule for expected loss $f(p_t; \boldsymbol{\ell}_t) = p_t \ell_t^H + (1-p_t)\ell_t^L$:

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

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