Online Learning
and Online Convex Optimization

Nicolò Cesa-Bianchi

Università degli Studi di Milano
1. My beautiful regret

2. A supposedly fun game I’ll play again

3. The joy of convex
1. My beautiful regret

2. A supposedly fun game I’ll play again

3. The joy of convex
Machine learning

Classification/regression tasks

- Predictive models \( h \) mapping data instances \( X \) to labels \( Y \) (e.g., binary classifier)
- Training data \( S_T = \left( (X_1, Y_1), \ldots, (X_T, Y_T) \right) \) (e.g., email messages with spam vs. nonspam annotations)
- Learning algorithm \( A \) (e.g., Support Vector Machine) maps training data \( S_T \) to model \( h = A(S_T) \)

Evaluate the risk of the trained model \( h \) with respect to a given loss function
Two notions of risk

View data as a statistical sample: **statistical risk**

\[ \mathbb{E} \left[ \ell \left( \mathcal{A}(S_T), (X, Y) \right) \right] \]

Training set \( S_T = ((X_1, Y_1), \ldots, (X_T, Y_T)) \) and test example \((X, Y)\) drawn i.i.d. from the same unknown and fixed distribution

Training set

\( S_T = ((X_1, Y_1), \ldots, (X_T, Y_T)) \)
Two notions of risk

View data as a statistical sample: **statistical risk**

\[ \mathbb{E}
\left[
\ell\left(A\left(S_T\right), (X, Y)\right)
\right]
\]

Training set \( S_T \) = \((X_1, Y_1), \ldots, (X_T, Y_T)\) and test example \((X, Y)\) drawn i.i.d. from the same unknown and fixed distribution

View data as an arbitrary sequence: **sequential risk**

\[
\sum_{t=1}^{T} \ell\left(A\left(S_{t-1}\right), (X_t, Y_t)\right)
\]

Sequence of models trained on growing prefixes \( S_t \) = \((X_1, Y_1), \ldots, (X_t, Y_t)\) of the data sequence
Learning algorithm $A$ maps datasets to models in a given class $\mathcal{H}$.

**Variance error in statistical learning**

$$\mathbb{E}\left[\ell(A(S_T), (X, Y))\right] - \inf_{h \in \mathcal{H}} \mathbb{E}\left[\ell(h, (X, Y))\right]$$

compare to expected loss of best model in the class.
Learning algorithm $A$ maps datasets to models in a given class $\mathcal{H}$.

**Variance error in statistical learning**

$$
\mathbb{E}\left[\ell(A(S_T), (X, Y))\right] - \inf_{h \in \mathcal{H}} \mathbb{E}\left[\ell(h, (X, Y))\right]
$$

compare to expected loss of best model in the class

**Regret in online learning**

$$
\sum_{t=1}^{T} \ell(A(S_{t-1}), (X_t, Y_t)) - \inf_{h \in \mathcal{H}} \sum_{t=1}^{T} \ell(h, (X_t, Y_t))
$$

compare to cumulative loss of best model in the class
Incremental model update

A natural blueprint for online learning algorithms

For $t = 1, 2, \ldots$

1. Apply current model $h_{t-1}$ to next data element $(X_t, Y_t)$
2. Update current model: $h_{t-1} \rightarrow h_t \in \mathcal{H}$ (local optimization)
Incremental model update

A natural blueprint for online learning algorithms

For $t = 1, 2, \ldots$

1. Apply current model $h_{t-1}$ to next data element $(X_t, Y_t)$
2. Update current model: $h_{t-1} \rightarrow h_t \in \mathcal{H}$ (local optimization)

Goal: control regret

$$\sum_{t=1}^{T} \ell(h_{t-1}, (X_t, Y_t)) - \inf_{h \in \mathcal{H}} \sum_{t=1}^{T} \ell(h, (X_t, Y_t))$$
Incremental model update

A natural blueprint for online learning algorithms

For $t = 1, 2, \ldots$

1. Apply current model $h_{t-1}$ to next data element $(X_t, Y_t)$
2. Update current model: $h_{t-1} \rightarrow h_t \in \mathcal{H}$ (local optimization)

Goal: control regret

$$
\sum_{t=1}^{T} \ell(h_{t-1}, (X_t, Y_t)) - \inf_{h \in \mathcal{H}} \sum_{t=1}^{T} \ell(h, (X_t, Y_t))
$$

View this as a repeated game between a player generating predictors $h_t \in \mathcal{H}$ and an opponent generating data $(X_t, Y_t)$
1. My beautiful regret

2. A supposedly fun game I’ll play again

3. The joy of convex
Theory of repeated games

James Hannan (1922–2010)

David Blackwell (1919–2010)

Learning to play a game (1956)

Play a game repeatedly against a possibly suboptimal opponent
Zero-sum 2-person games played more than once

<table>
<thead>
<tr>
<th></th>
<th>1</th>
<th>2</th>
<th>...</th>
<th>M</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>$\ell(1,1)$</td>
<td>$\ell(1,2)$</td>
<td>...</td>
<td></td>
</tr>
<tr>
<td>2</td>
<td>$\ell(2,1)$</td>
<td>$\ell(2,2)$</td>
<td>...</td>
<td></td>
</tr>
<tr>
<td>...</td>
<td>...</td>
<td>...</td>
<td>...</td>
<td>...</td>
</tr>
<tr>
<td>N</td>
<td>...</td>
<td>...</td>
<td>...</td>
<td>...</td>
</tr>
</tbody>
</table>

*N \times M* known loss matrix

- Row player *(player)* has *N* actions
- Column player *(opponent)* has *M* actions

For each game round $t = 1, 2, \ldots$

- Player chooses action $i_t$ and opponent chooses action $y_t$
- The player suffers loss $\ell(i_t, y_t)$ (= gain of opponent)

Player can learn from opponent’s history of past choices $y_1, \ldots, y_{t-1}$
Opponent’s moves $y_1, y_2, \ldots$ define a sequential prediction problem with a time-varying loss function $\ell(i_t, y_t) = \ell_t(i_t)$.
Playing the experts game

A sequential decision problem

- $N$ actions
- Unknown deterministic assignment of losses to actions
  \[ \ell_t = (\ell_t(1), \ldots, \ell_t(N)) \in [0, 1]^N \text{ for } t = 1, 2, \ldots \]

For $t = 1, 2, \ldots$
Playing the experts game

A sequential decision problem

- N actions
- Unknown deterministic assignment of losses to actions
  \[ \ell_t = (\ell_t(1), \ldots, \ell_t(N)) \in [0, 1]^N \text{ for } t = 1, 2, \ldots \]

For \( t = 1, 2, \ldots \)

- Player picks an action \( I_t \) (possibly using randomization) and incurs loss \( \ell_t(I_t) \)
Playing the experts game

A sequential decision problem

- $N$ actions
- Unknown deterministic assignment of losses to actions
  \[ \ell_t = (\ell_t(1), \ldots, \ell_t(N)) \in [0, 1]^N \] for $t = 1, 2, \ldots$

For $t = 1, 2, \ldots$

1. Player picks an action $I_t$ (possibly using randomization) and incurs loss $\ell_t(I_t)$
2. Player gets feedback information: $\ell_t(1), \ldots, \ell_t(N)$
Regret analysis

Regret

\[ R_T \overset{\text{def}}{=} \mathbb{E} \left[ \sum_{t=1}^{T} \ell_t(I_t) \right] - \min_{i=1,\ldots,N} \sum_{t=1}^{T} \ell_t(i) \overset{\text{want}}{=} o(T) \]
Regret analysis

Regret

\[ R_T \overset{\text{def}}{=} \mathbb{E} \left[ \sum_{t=1}^{T} \ell_t(I_t) \right] - \min_{i=1,\ldots,N} \sum_{t=1}^{T} \ell_t(i) = o(T) \]

Lower bound using random losses

- \( \ell_t(i) \rightarrow L_t(i) \in \{0, 1\} \) independent random coin flip
- For any player strategy \( \mathbb{E} \left[ \sum_{t=1}^{T} L_t(I_t) \right] = \frac{T}{2} \)
- Then the expected regret is
  \[ \mathbb{E} \left[ \max_{i=1,\ldots,N} \sum_{t=1}^{T} \left( \frac{1}{2} - L_t(i) \right) \right] = (1 - o(1)) \sqrt{\frac{T \ln N}{2}} \]
  for \( N, T \rightarrow \infty \)
Exponentially weighted forecaster (Hedge)

At time $t$ pick action $I_t = i$ with probability proportional to

$$\exp \left( -\eta \sum_{s=1}^{t-1} \ell_s(i) \right)$$

the sum at the exponent is the total loss of action $i$ up to now

Regret bound [Experts’ paper, 1997]

- If $\eta = \sqrt{\ln N / (8T)}$ then
  $$R_T \leq \sqrt{\frac{T \ln N}{2}}$$
- Matching lower bound including constants
- Dynamic choice $\eta_t = \sqrt{\ln N / (8t)}$ only loses small constants
The nonstochastic bandit problem

For $t = 1, 2, \ldots$

Player picks an action $I_t$ (possibly using randomization) and incurs loss $\ell_t(I_t)$.

Player gets partial information: Only $\ell_t(I_t)$ is revealed.

Player still competing against the best offline action $R_T = \mathbb{E} \left[ \sum_{t=1}^{T} \ell_t(I_t) \right] - \min_{i = 1, \ldots, N} \mathbb{E} \left[ \sum_{t=1}^{T} \ell_t(i) \right]$. 
The nonstochastic bandit problem

For $t = 1, 2, \ldots$

1. Player picks an action $I_t$ (possibly using randomization) and incurs loss $\ell_t(I_t)$.
The nonstochastic bandit problem

For $t = 1, 2, \ldots$

1. Player picks an action $I_t$ (possibly using randomization) and incurs loss $\ell_t(I_t)$

2. Player gets partial information: Only $\ell_t(I_t)$ is revealed
The nonstochastic bandit problem

For $t = 1, 2, \ldots$

1. Player picks an action $I_t$ (possibly using randomization) and incurs loss $\ell_t(I_t)$
2. Player gets partial information: Only $\ell_t(I_t)$ is revealed

Player still competing against best offline action

$$R_T = \mathbb{E} \left[ \sum_{t=1}^{T} \ell_t(I_t) \right] - \min_{i=1,\ldots,N} \sum_{t=1}^{T} \ell_t(i)$$
The Exp3 algorithm

Hedge with estimated losses

\[ P_t(I_t = i) \propto \exp \left( -\eta \sum_{s=1}^{t-1} \hat{\ell}_s(i) \right) \quad i = 1, \ldots, N \]

\[ \hat{\ell}_t(i) = \begin{cases} \ell_t(i) & \text{if } I_t = i \\ \frac{P_t(\ell_t(i) \text{ observed})}{P_t(\ell_t(i) \text{ observed})} & \text{otherwise} \end{cases} \]

Only one non-zero component in \( \hat{\ell}_t \)
The Exp3 algorithm

Hedge with estimated losses

- \( P_t(I_t = i) \propto \exp \left( -\eta \sum_{s=1}^{t-1} \hat{\ell}_s(i) \right) \) \( i = 1, \ldots, N \)

- \( \hat{\ell}_t(i) = \begin{cases} \begin{array}{ll}
\ell_t(i) & \text{if } I_t = i \\
\frac{P_t(\ell_t(i) \text{ observed})}{P_t(\ell_t(i) \text{ observed})} & \text{otherwise}
\end{array} \end{cases} \)

Only one non-zero component in \( \hat{\ell}_t \)

Properties of importance weighting estimator

- \( \mathbb{E}_t \left[ \hat{\ell}_t(i) \right] = \ell_t(i) \) unbiasedness

- \( \mathbb{E}_t \left[ \hat{\ell}_t(i)^2 \right] \leq \frac{1}{P_t(\ell_t(i) \text{ observed})} \) variance control
\[ R_T \leq \frac{\ln N}{\eta} + \frac{\eta}{2} \mathbb{E} \left[ \sum_{t=1}^{T} \sum_{i=1}^{N} P_t(I_t = i) \mathbb{E}_t \left[ \hat{\ell}_t(i)^2 \right] \right] \]

\[ \leq \frac{\ln N}{\eta} + \frac{\eta}{2} \mathbb{E} \left[ \sum_{t=1}^{T} \sum_{i=1}^{N} \frac{P_t(I_t = i)}{P_t(\ell_t(i) \text{ is observed})} \right] \]

\[ = \frac{\ln N}{\eta} + \frac{\eta}{2} \sqrt{NT} = \sqrt{NT \ln N} \quad \text{lower bound } \Omega(\sqrt{NT}) \]
Exp3 regret bound

\[ R_T \leq \frac{\ln N}{\eta} + \frac{\eta}{2} \mathbb{E} \left[ \sum_{t=1}^{T} \sum_{i=1}^{N} P_t(I_t = i) \mathbb{E}_t \left[ \hat{\ell}_t(i)^2 \right] \right] \]

\leq \frac{\ln N}{\eta} + \frac{\eta}{2} \mathbb{E} \left[ \sum_{t=1}^{T} \sum_{i=1}^{N} \frac{P_t(I_t = i)}{P_t(\ell_t(i) \text{ is observed})} \right]

= \frac{\ln N}{\eta} + \frac{\eta}{2} NT = \sqrt{NT \ln N} \quad \text{lower bound} \quad \Omega(\sqrt{NT})

Improved matching upper bound by [Audibert and Bubeck, 2009]
Exp3 regret bound

\[ R_T \leq \frac{\ln N}{\eta} + \frac{\eta}{2} \mathbb{E} \left[ \sum_{t=1}^{T} \sum_{i=1}^{N} P_t(I_t = i) \mathbb{E}_t[\hat{\ell}_t(i)^2] \right] \]

\leq \frac{\ln N}{\eta} + \frac{\eta}{2} \mathbb{E} \left[ \sum_{t=1}^{T} \sum_{i=1}^{N} \frac{P_t(I_t = i)}{P_t(\ell_t(i) \text{ is observed})} \right] \]

\[ = \frac{\ln N}{\eta} + \frac{\eta}{2} NT = \sqrt{NT \ln N} \quad \text{lower bound } \Omega(\sqrt{NT}) \]

Improved matching upper bound by [Audibert and Bubeck, 2009]

The full information (experts) setting

- Player observes vector of losses \( \ell_t \) after each play
- \( P_t(\ell_t(i) \text{ is observed}) = 1 \)
- \( R_T \leq \sqrt{T \ln N} \)
The adaptive adversary

- The loss of action $i$ at time $t$ depends on the player’s past $m$ actions $\ell_t(i) \rightarrow \ell_t(I_{t-m}, \ldots, I_{t-1}, i)$
Nonoblivious opponents

The adaptive adversary

- The loss of action $i$ at time $t$ depends on the player’s past $m$ actions $\ell_t(i) \rightarrow \ell_t(I_{t-m}, \ldots, I_{t-1}, i)$
- Examples: bandits with switching cost
Nonoblivious opponents

The adaptive adversary

- The loss of action $i$ at time $t$ depends on the player’s past $m$ actions $\ell_t(i) \to \ell_t(I_{t-m}, \ldots, I_{t-1}, i)$
- Examples: bandits with switching cost

Nonoblivious regret

$$R_{T}^{\text{non}} = \mathbb{E} \left[ \sum_{t=1}^{T} \ell_t(I_{t-m}, \ldots, I_{t-1}, I_t) - \min_{i=1, \ldots, N} \sum_{t=1}^{T} \ell_t(I_{t-m}, \ldots, I_{t-1}, i) \right]$$
Nonoblivious opponents

The adaptive adversary

- The loss of action $i$ at time $t$ depends on the player’s past $m$ actions $\ell_t(i) \rightarrow \ell_t(I_{t-m}, \ldots, I_{t-1}, i)$
- Examples: bandits with switching cost

Nonoblivious regret

$$R^{\text{non}}_T = \mathbb{E} \left[ \sum_{t=1}^{T} \ell_t(I_{t-m}, \ldots, I_{t-1}, I_t) - \min_{i=1,\ldots,N} \sum_{t=1}^{T} \ell_t(I_{t-m}, \ldots, I_{t-1}, i) \right]$$

Policy regret

$$R^{\text{pol}}_T = \mathbb{E} \left[ \sum_{t=1}^{T} \ell_t(I_{t-m}, \ldots, I_{t-1}, I_t) - \min_{i=1,\ldots,N} \sum_{t=1}^{T} \ell_t(i, \ldots, i, i) \right]$$
Bandits and reactive opponents

Bounds on the nonoblivious regret (even when $m$ depends on $T$)

$$R_T^{\text{non}} = \Theta\left(\sqrt{TN \ln N}\right)$$

- Exp3 with biased loss estimates
- Is the $\sqrt{\ln N}$ factor necessary?
Bandits and reactive opponents

**Bounds on the nonoblivious regret (even when \( m \) depends on \( T \))**

\[
R_T^{\text{non}} = \Theta(\sqrt{T N \ln N})
\]

- Exp3 with biased loss estimates
- Is the \( \sqrt{\ln N} \) factor necessary?

**Bounds on the policy regret for any constant \( m \geq 1 \)**

\[
R_T^{\text{pol}} = \Theta((N \ln N)^{1/3} T^{2/3})
\]

- Achieved by a very simple player strategy
- Optimal up to log factors!

[Dekel, Koren, and Peres, 2014]
Partial monitoring: not observing any loss

Dynamic pricing: Perform as the best fixed price

1. Post a T-shirt price
2. Observe if next customer buys or not
3. Adjust price

Feedback does not reveal the player’s loss

<table>
<thead>
<tr>
<th></th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
<th>5</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>0</td>
<td>1</td>
<td>2</td>
<td>3</td>
<td>4</td>
</tr>
<tr>
<td>2</td>
<td>c</td>
<td>0</td>
<td>1</td>
<td>2</td>
<td>3</td>
</tr>
<tr>
<td>3</td>
<td>c</td>
<td>c</td>
<td>0</td>
<td>1</td>
<td>2</td>
</tr>
<tr>
<td>4</td>
<td>c</td>
<td>c</td>
<td>c</td>
<td>0</td>
<td>1</td>
</tr>
<tr>
<td>5</td>
<td>c</td>
<td>c</td>
<td>c</td>
<td>c</td>
<td>0</td>
</tr>
</tbody>
</table>

Loss matrix

<table>
<thead>
<tr>
<th></th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
<th>5</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
</tr>
<tr>
<td>2</td>
<td>0</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
</tr>
<tr>
<td>3</td>
<td>0</td>
<td>0</td>
<td>1</td>
<td>1</td>
<td>1</td>
</tr>
<tr>
<td>4</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>1</td>
<td>1</td>
</tr>
<tr>
<td>5</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>1</td>
</tr>
</tbody>
</table>

Feedback matrix
A characterization of minimax regret

Special case

Multiarmed bandits: loss and feedback matrix are the same

A general gap theorem

[Bartok, Foster, Pál, Rakhlin and Szepesvári, 2013]

A constructive characterization of the minimax regret for any pair of loss/feedback matrix

Only three possible rates for nontrivial games:

1. Easy games (e.g., bandits): $\Theta(\sqrt{T})$
2. Hard games (e.g., revealing action): $\Theta(T^{2/3})$
3. Impossible games: $\Theta(T)$
A characterization of minimax regret

Special case
Multiarmed bandits: loss and feedback matrix are the same

A general gap theorem [Bartok, Foster, Pál, Rakhlin and Szepesvári, 2013]

- A constructive characterization of the minimax regret for any pair of loss/feedback matrix
- Only three possible rates for nontrivial games:
  1. Easy games (e.g., bandits): $\Theta(\sqrt{T})$
  2. Hard games (e.g., revealing action): $\Theta(T^{2/3})$
  3. Impossible games: $\Theta(T)$
A game equivalent to prediction with expert advice

Online linear optimization in the simplex

1. Play $p_t$ from the $N$-dimensional simplex $\Delta_N$
2. Incur linear loss $\mathbb{E}[\ell_t(I_t)] = p_t^T \ell_t$
3. Observe loss gradient $\ell_t$

Regret: compete against the best point in the simplex

$$\sum_{t=1}^{T} p_t^T \ell_t - \min_{q \in \Delta_N} \sum_{t=1}^{T} q^T \ell_t$$

$$= \min_{i=1,\ldots,N} \frac{1}{T} \sum_{t=1}^{T} \ell_t(i)$$
Opponent’s moves $y_t$ are viewed as values or labels assigned to observations $x_t \in \mathbb{R}^d$ (e.g., categories of documents)

A repeated game between the player choosing an element $w_t$ of a linear space and the opponent choosing a label $y_t$ for $x_t$

Regret with respect to best element in the linear space
Summary

1. My beautiful regret

2. A supposedly fun game I’ll play again

3. The joy of convex
Online convex optimization

1. Play $\mathbf{w}_t$ from a convex and compact subset $S$ of a linear space
2. Observe convex loss $\ell_t : S \to \mathbb{R}$ and pay $\ell_t(\mathbf{w}_t)$
3. Update: $\mathbf{w}_t \rightarrow \mathbf{w}_{t+1} \in S$
Online convex optimization

1. Play $w_t$ from a **convex and compact subset $S$** of a linear space
2. Observe convex loss $\ell_t : S \to \mathbb{R}$ and pay $\ell_t(w_t)$
3. Update: $w_t \rightarrow w_{t+1} \in S$

Example

- Regression with square loss: $\ell_t(w) = (w^\top x_t - y_t)^2 \quad y_t \in \mathbb{R}$
- Classification with hinge loss: $\ell_t(w) = [1 - y_t w^\top x_t]_+ \quad y_t \in \{-1, +1\}$
Online convex optimization

1. Play $\mathbf{w}_t$ from a **convex and compact subset** $S$ of a linear space.
2. Observe convex loss $\ell_t : S \to \mathbb{R}$ and pay $\ell_t(\mathbf{w}_t)$.
3. Update: $\mathbf{w}_t \to \mathbf{w}_{t+1} \in S$.

**Example**

- **Regression with square loss**: $\ell_t(\mathbf{w}) = (\mathbf{w}^\top \mathbf{x}_t - y_t)^2 \quad y_t \in \mathbb{R}$
- **Classification with hinge loss**: $\ell_t(\mathbf{w}) = \left[ 1 - y_t \mathbf{w}^\top \mathbf{x}_t \right]_+ \quad y_t \in \{-1, +1\}$

**Regret**

$$R_T(\mathbf{u}) = \sum_{t=1}^{T} \ell_t(\mathbf{w}_t) - \sum_{t=1}^{T} \ell_t(\mathbf{u}) \quad \mathbf{u} \in S$$
Follow the leader

\[ \mathbf{w}_{t+1} = \arg \inf_{\mathbf{w} \in S} \sum_{s=1}^{t} \ell_s(\mathbf{w}) \]

Regret can be linear due to lack of stability

\[ S = [-1, +1] \quad \ell_1(\mathbf{w}) = \frac{\mathbf{w}}{2} \quad \ell_t(\mathbf{w}) = \begin{cases} -\mathbf{w} & \text{if } t \text{ is even} \\ +\mathbf{w} & \text{if } t \text{ is odd} \end{cases} \]
Finding a good online algorithm

Follow the leader

\[ \mathbf{w}_{t+1} = \operatorname{argmin}_{\mathbf{w} \in S} \sum_{s=1}^{t} \ell_s(\mathbf{w}) \]

Regret can be linear due to lack of stability

\[ S = [-1, +1] \quad \ell_1(\mathbf{w}) = \frac{\mathbf{w}}{2} \quad \ell_t(\mathbf{w}) = \begin{cases} -\mathbf{w} & \text{if } t \text{ is even} \\ +\mathbf{w} & \text{if } t \text{ is odd} \end{cases} \]

Note:

\[ \sum_{s=1}^{t} \ell_s(\mathbf{w}) = \begin{cases} -\frac{\mathbf{w}}{2} & \text{if } t \text{ is even} \\ +\frac{\mathbf{w}}{2} & \text{if } t \text{ is odd} \end{cases} \]

Hence \( \ell_{t+1}(\mathbf{w}_{t+1}) = 1 \) for all \( t = 1, 2, \ldots \)
Follow the regularized leader

[Shalev-Shwartz, 2007; Abernethy, Hazan and Rakhlin, 2008]

\[ w_{t+1} = \arg\min_{w \in S} \left[ \eta \sum_{s=1}^{t} \ell_s(w) + \Phi(w) \right] \]

\( \Phi \) is a strongly convex regularizer and \( \eta > 0 \) is a scale parameter
Convexity, smoothness, and duality

**Strong convexity**

Φ : S → R is β-strongly convex w.r.t. a norm \( \| \cdot \| \) if for all \( u, v \in S \)

\[
\Phi(v) \geq \Phi(u) + \nabla \Phi(u)^\top (v - u) + \frac{\beta}{2} \| u - v \|^2
\]
Convexity, smoothness, and duality

**Strong convexity**

\( \Phi : S \rightarrow \mathbb{R} \) is \( \beta \)-strongly convex w.r.t. a norm \( \| \cdot \| \) if for all \( u, v \in S \)

\[
\Phi(v) \geq \Phi(u) + \nabla \Phi(u)^\top (v - u) + \frac{\beta}{2} \| u - v \|^2
\]

**Smoothness**

\( \Phi : S \rightarrow \mathbb{R} \) is \( \alpha \)-smooth w.r.t. a norm \( \| \cdot \| \) if for all \( u, v \in S \)

\[
\Phi(v) \leq \Phi(u) + \nabla \Phi(u)^\top (v - u) + \frac{\alpha}{2} \| u - v \|^2
\]
Convexity, smoothness, and duality

### Strong convexity

\( \Phi : S \to \mathbb{R} \) is \( \beta \)-strongly convex w.r.t. a norm \( \| \cdot \| \) if for all \( u, v \in S \)

\[
\Phi(v) \geq \Phi(u) + \nabla \Phi(u)^\top (v - u) + \frac{\beta}{2} \| u - v \|^2
\]

### Smoothness

\( \Phi : S \to \mathbb{R} \) is \( \alpha \)-smooth w.r.t. a norm \( \| \cdot \| \) if for all \( u, v \in S \)

\[
\Phi(v) \leq \Phi(u) + \nabla \Phi(u)^\top (v - u) + \frac{\alpha}{2} \| u - v \|^2
\]

- If \( \Phi \) is \( \beta \)-strongly convex w.r.t. \( \| \cdot \|_2 \), then \( \nabla^2 \Phi \succeq \beta I \)
- If \( \Phi \) is \( \alpha \)-smooth w.r.t. \( \| \cdot \|_2 \), then \( \nabla^2 \Phi \preceq \alpha I \)
Examples

- Euclidean norm: $\Phi = \frac{1}{2} \| \cdot \|_2^2$ is $1$-strongly convex w.r.t. $\| \cdot \|_2$

- $p$-norm: $\Phi = \frac{1}{2} \| \cdot \|_2^p$ is $(p - 1)$-strongly convex w.r.t. $\| \cdot \|_p$ (for $1 < p \leq 2$)

- Entropy: $\Phi(p) = d \sum_{i=1}^{d} p_i \ln p_i$ is $1$-strongly convex w.r.t. $\| \cdot \|_1$ (for $p$ in the probability simplex)

- Power norm: $\Phi(w) = \frac{1}{2} w^\top A w$ is $1$-strongly convex w.r.t. $\| w \| = \sqrt{w^\top A w}$ (for $A$ symmetric and positive definite)
Examples

- Euclidean norm: $\Phi = \frac{1}{2} \| \cdot \|_2^2$ is $1$-strongly convex w.r.t. $\| \cdot \|_2$

- $p$-norm: $\Phi = \frac{1}{2} \| \cdot \|_p^2$ is $(p-1)$-strongly convex w.r.t. $\| \cdot \|_p$
  (for $1 < p \leq 2$)
Examples

- **Euclidean norm:** $\Phi = \frac{1}{2} \| \cdot \|^2_2$ is 1-strongly convex w.r.t. $\| \cdot \|_2$

- **$p$-norm:** $\Phi = \frac{1}{2} \| \cdot \|^2_p$ is $(p - 1)$-strongly convex w.r.t. $\| \cdot \|_p$  
  (for $1 < p \leq 2$)

- **Entropy:** $\Phi(p) = \sum_{i=1}^{d} p_i \ln p_i$ is 1-strongly convex w.r.t. $\| \cdot \|_1$  
  (for $p$ in the probability simplex)
Examples

- **Euclidean norm:** $\Phi = \frac{1}{2} \| \cdot \|_2^2$ is 1-strongly convex w.r.t. $\| \cdot \|_2$
- **p-norm:** $\Phi = \frac{1}{2} \| \cdot \|_p^2$ is $(p - 1)$-strongly convex w.r.t. $\| \cdot \|_p$ (for $1 < p \leq 2$)
- **Entropy:** $\Phi(p) = \sum_{i=1}^{d} p_i \ln p_i$ is 1-strongly convex w.r.t. $\| \cdot \|_1$ (for $p$ in the probability simplex)
- **Power norm:** $\Phi(w) = \frac{1}{2} w^\top A w$ is 1-strongly convex w.r.t. $\| w \| = \sqrt{w^\top A w}$ (for $A$ symmetric and positive definite)
Convex duality

**Definition**

The **convex dual** of $\Phi$ is

$$\Phi^*(\theta) = \max_{w \in S} \left( \theta^T w - \Phi(w) \right)$$

**1-dimensional example**

- Convex $f : \mathbb{R} \to \mathbb{R}$ such that $f(0) = 0$
- $f^*(\theta) = \max_{w \in \mathbb{R}} (w \times \theta - f(w))$
- The maximizer is $w_0$ such that $f'(w_0) = \theta$
- This gives $f^*(\theta) = w_0 \times f'(w_0) - f(w_0)$
- As $f(0) = 0$, $f^*(\theta)$ is the error in approximating $f(0)$ with a first-order expansion around $f(w_0)$
Convex duality

(thanks to Shai Shalev-Shwartz for the image)
Examples

- **Euclidean norm:** \( \Phi = \frac{1}{2} \| \cdot \|_2^2 \) and \( \Phi^* = \Phi \)
Convexity, smoothness, and duality

Examples

- Euclidean norm: $\Phi = \frac{1}{2} \| \cdot \|_2^2$ and $\Phi^* = \Phi$

- p-norm: $\Phi = \frac{1}{2} \| \cdot \|_p^2$ and $\Phi^* = \frac{1}{2} \| \cdot \|_q^2$ where $\frac{1}{p} + \frac{1}{q} = 1$
Convexity, smoothness, and duality

Examples

- **Euclidean norm**: $\Phi = \frac{1}{2} \| \cdot \|_2^2$ and $\Phi^* = \Phi$

- **p-norm**: $\Phi = \frac{1}{2} \| \cdot \|_p^2$ and $\Phi^* = \frac{1}{2} \| \cdot \|_q^2$ where $\frac{1}{p} + \frac{1}{q} = 1$

- **Entropy**: $\Phi(p) = \sum_{i=1}^{d} p_i \ln p_i$ and $\Phi^*(\theta) = \ln \left( e^{\theta_1} + \cdots + e^{\theta_d} \right)$
Convexity, smoothness, and duality

Examples

- **Euclidean norm:** $\Phi = \frac{1}{2} \| \cdot \|_2^2$ and $\Phi^* = \Phi$

- **p-norm:** $\Phi = \frac{1}{2} \| \cdot \|_p^2$ and $\Phi^* = \frac{1}{2} \| \cdot \|_q^2$ where $\frac{1}{p} + \frac{1}{q} = 1$

- **Entropy:** $\Phi(p) = \sum_{i=1}^{d} p_i \ln p_i$ and $\Phi^*(\theta) = \ln \left( e^{\theta_1} + \cdots + e^{\theta_d} \right)$

- **Power norm:** $\Phi(w) = \frac{1}{2} w^\top A w$ and $\Phi^*(\theta) = \frac{1}{2} \theta^\top A^{-1} \theta$
Some useful properties

If $\Phi : S \to \mathbb{R}$ is $\beta$-strongly convex w.r.t. $\| \cdot \|$, then

- Its convex dual $\Phi^*$ is everywhere differentiable and $\frac{1}{\beta}$-smooth w.r.t. $\| \cdot \|_*$ (the dual norm of $\| \cdot \|$)

- $\nabla \Phi^*(\theta) = \arg\max_{w \in S} \left( \theta^T w - \Phi(w) \right)$
Some useful properties

If $\Phi : S \to \mathbb{R}$ is $\beta$-strongly convex w.r.t. $\| \cdot \|$, then

- Its convex dual $\Phi^*$ is everywhere differentiable and $\frac{1}{\beta}$-smooth w.r.t. $\| \cdot \|_*$ (the dual norm of $\| \cdot \|$)

- $\nabla \Phi^*(\theta) = \arg\max_{w \in S} \left( \theta^T w - \Phi(w) \right)$

Recall: Follow the regularized leader (FTRL)

$$w_{t+1} = \arg\min_{w \in S} \left[ \eta \sum_{s=1}^{t} \ell_s(w) + \Phi(w) \right]$$

$\Phi$ is a strongly convex regularizer and $\eta > 0$ is a scale parameter
Using the loss gradient

**Linearization of convex losses**

\[
\ell_t(w_t) - \ell_t(u) \leq \nabla \ell_t(w_t)^\top w_t - \nabla \ell_t(w_t)^\top u
\]

**FTRL with linearized losses**

\[
w_{t+1} = \arg\min_{w \in S} \left( \eta \sum_{s=1}^{t} \tilde{\ell}_s^\top w + \Phi(w) \right) = \arg\max_{w \in S} \left( \theta_{t+1}^\top w - \Phi(w) \right)
\]

\[
= \nabla \Phi^*(\theta_{t+1})
\]

**Note:** \(w_{t+1} \in S\) always holds
The Mirror Descent algorithm [Nemirovsky and Yudin, 1983]

Recall: \( \mathbf{w}_{t+1} = \nabla \Phi^*(\theta_t) = \nabla \Phi^*(-\eta \sum_{s=1}^{t} \nabla \ell_s(\mathbf{w}_s)) \)

Online Mirror Descent (FTRL with linearized losses)

Parameters: Strongly convex regularizer \( \Phi \) with domain \( S \), \( \eta > 0 \)
Initialize: \( \theta_1 = 0 \) // primal parameter

For \( t = 1, 2, \ldots \)

1. Use \( \mathbf{w}_t = \nabla \Phi^*(\theta_t) \) // dual parameter (via mirror step)
2. Suffer loss \( \ell_t(\mathbf{w}_t) \)
3. Observe loss gradient \( \nabla \ell_t(\mathbf{w}_t) \)
4. Update \( \theta_{t+1} = \theta_t - \eta \nabla \ell_t(\mathbf{w}_t) \) // gradient step
An equivalent formulation

Under some assumptions on the regularizer $\Phi$, OMD can be equivalently written in terms of **projected gradient descent**

**Online Mirror Descent (alternative version)**

**Parameters:** Strongly convex regularizer $\Phi$ and learning rate $\eta > 0$

**Initialize:** $z_1 = \nabla \Phi^* (0)$ and $w_1 = \arg\min_{w \in S} D_{\Phi} (w \parallel z_1)$

For $t = 1, 2, \ldots$

1. Use $w_t$ and suffer loss $\ell_t (w_t)$
2. Observe loss gradient $\nabla \ell_t (w_t)$
3. Update $z_{t+1} = \nabla \Phi^* \left( \nabla \Phi (z_t) - \eta \nabla \ell_t (w_t) \right)$ // gradient step
4. $w_{t+1} = \arg\min_{w \in S} D_{\Phi} (w \parallel z_{t+1})$ // projection step
An equivalent formulation

Under some assumptions on the regularizer $\Phi$, OMD can be equivalently written in terms of **projected gradient descent**

**Online Mirror Descent (alternative version)**

**Parameters:** Strongly convex regularizer $\Phi$ and learning rate $\eta > 0$

**Initialize:** $z_1 = \nabla \Phi^*(0)$ and $w_1 = \arg\min_{w \in S} D_\Phi(w \| z_1)$

For $t = 1, 2, \ldots$

1. Use $w_t$ and suffer loss $\ell_t(w_t)$
2. Observe loss gradient $\nabla \ell_t(w_t)$
3. Update $z_{t+1} = \nabla \Phi^* \left( \nabla \Phi(z_t) - \eta \nabla \ell_t(w_t) \right)$ // gradient step
4. $w_{t+1} = \arg\min_{w \in S} D_\Phi(w \| z_{t+1})$ // projection step

$D_\Phi$ is the **Bregman divergence** induced by $\Phi$
### Some examples

**Online Gradient Descent (OGD)**

- \( \Phi(w) = \frac{1}{2} \|w\|^2 \)
- **Update:** \( w' = w_t - \eta \nabla \ell_t(w_t) \)

**p-norm version:** \( \Phi(w) = \frac{1}{2} \|w\|^2_p \)

- \( w_{t+1} = \text{argmin}_{w \in S} \|w - w'\|_2 \)

[Zinkevich, 2003; Gentile, 2003]
Some examples

**Online Gradient Descent (OGD)**  
[Zinkevich, 2003; Gentile, 2003]

- $\Phi(w) = \frac{1}{2} \|w\|^2$
- Update: $w' = w_t - \eta \nabla \ell_t(w_t)$
- $p$-norm version: $\Phi(w) = \frac{1}{2} \|w\|^2_p$
- $w_{t+1} = \text{arg inf}_{w \in S} \|w - w'\|_2$

**Exponentiated gradient (EG)**  
[Kivinen and Warmuth, 1997]

- $\Phi(p) = \sum_{i=1}^{d} p_i \ln p_i$
- $p \in S \equiv$ simplex
- $p_{t+1,i} = \frac{p_{t,i} e^{-\eta \nabla \ell_t(p_t)_i}}{\sum_{j=1}^{d} p_{t,j} e^{-\eta \nabla \ell_t(p_t)_j}}$

**Note:** when losses are linear this is Hedge
Regret analysis

Regret bound

\[ R_T(u) \leq \frac{\Phi(u) - \min_{w \in S} \Phi(w)}{\eta} + \frac{\eta}{2} \sum_{t=1}^{T} \frac{\|\nabla \ell_t(w_t)\|_*^2}{\beta} \]

for all \( u \in S \), where \( \ell_1, \ell_2, \ldots \) are arbitrary convex losses

- \( R_T(u) \leq GD \sqrt{T} \) for all \( u \in S \) when \( \eta \) is tuned w.r.t.

\[ \sup_{w \in S} \|\nabla \ell_t(w)\|_* \leq G \quad \sqrt{\sup_{u,w \in S} \left( \Phi(u) - \Phi(w) \right)} \leq D \]

- Boundedness of gradients of \( \ell_t \) w.r.t. \( \|\cdot\|_* \) equivalent to Lipschitzess of \( \ell_t \) w.r.t. \( \|\cdot\| \)

- Regret bound optimal for general convex losses \( \ell_t \)
Analysis relies on smoothness of $\Phi^*$

$$
\Phi^*(\theta_{t+1}) - \Phi^*(\theta_t) \leq \underbrace{\nabla \Phi^*(\theta_t)^\top}_{\mathbf{w}_t} \underbrace{(\theta_{t+1} - \theta_t)}_{\eta \nabla \ell_t(w_t)} + \frac{1}{2\beta} \|\theta_{t+1} - \theta_t\|^2_*
$$
Analysis relies on smoothness of $\Phi^*$

\[
\Phi^*(\theta_{t+1}) - \Phi^*(\theta_t) \leq \nabla \Phi^*(\theta_t) \quad (\theta_{t+1} - \theta_t) + \frac{1}{2\beta} \| \theta_{t+1} - \theta_t \|^2
\]

\[
\sum_{t=1}^{T} \left( -\eta \nabla \ell_t(w_t) - \Phi(u) = u^\top \theta_{T+1} - \Phi(u) \right)
\]

\[
\leq \Phi^*(\theta_{T+1}) \quad \text{Fenchel-Young inequality}
\]

\[
= \sum_{t=1}^{T} (\Phi^*(\theta_{t+1}) - \Phi^*(\theta_t)) + \Phi^*(\theta_1)
\]

\[
\leq \sum_{t=1}^{T} \left( -\eta w_t^\top \nabla \ell_t(w_t) + \frac{\eta^2}{2\beta} \| \nabla \ell_t(w_t) \|^2 \right) + \Phi^*(0)
\]
Analysis relies on smoothness of $\Phi^*$

$$
\Phi^*(\theta_{t+1}) - \Phi^*(\theta_t) \leq \nabla \Phi^*(\theta_t)^\top \left( \frac{\theta_{t+1} - \theta_t}{w_t} \right) + \frac{1}{2\beta} \|\theta_{t+1} - \theta_t\|_*^2
$$

$$
\sum_{t=1}^{T} -\eta u^\top \nabla \ell_t(w_t) - \Phi(u) = u^\top \theta_{T+1} - \Phi(u)
$$

$$
\leq \Phi^*(\theta_{T+1}) \quad \text{Fenchel-Young inequality}
$$

$$
= \sum_{t=1}^{T} (\Phi^*(\theta_{t+1}) - \Phi^*(\theta_t)) + \Phi^*(\theta_1)
$$

$$
\leq \sum_{t=1}^{T} \left( -\eta w_t^\top \nabla \ell_t(w_t) + \frac{\eta^2}{2\beta} \|\nabla \ell_t(w_t)\|_*^2 \right) + \Phi^*(0)
$$

$$
\Phi^*(0) = \max_{w \in S} \left( w^\top 0 - \Phi(w) \right) = -\min_{w \in S} \Phi(w)
$$
Some examples

\[ l_t(w) \rightarrow l_t(w^T x_t) \quad \max_t |l'_t| \leq L \quad \max_t \|x_t\|_p \leq X_p \]
Some examples

\[
\ell_t(w) \rightarrow \ell_t(w^\top x_t) \quad \max_t |\ell'_t| \leq L \quad \max_t \|x_t\|_p \leq X_p
\]

Bounds for OGD with convex losses

\[
R_T(u) \leq B L X_2 \sqrt{T} = O(dL \sqrt{T})
\]

for all \( u \) such that \( \|u\|_2 \leq B \)
Some examples

$$\ell_t(w) \to \ell_t(w^\top x_t) \quad \max_t |\ell'_t| \leq L \quad \max_t \|x_t\|_p \leq X_p$$

Bounds for OGD with convex losses

$$R_T(u) \leq BLX_2 \sqrt{T} = \Theta(dL \sqrt{T})$$

for all $u$ such that $\|u\|_2 \leq B$

Bounds logarithmic in the dimension

- Regret bound for EG run in the simplex, $S = \Delta_d$
  $$R_T(q) \leq LX_\infty \sqrt{(\ln d)T} = \Theta(L \sqrt{(\ln d)T}) \quad p \in \Delta_d$$
- Same bound for $p$-norm regularizer with $p = \frac{\ln d}{\ln d - 1}$
- If losses are linear with $[0, 1]$ coefficients then we recover the bound for Hedge
Training set \((x_1, y_1), \ldots, (x_m, y_m) \in \mathbb{R}^d \times \{-1, +1\}\)

SVM objective \(F(w) = \frac{1}{m} \sum_{t=1}^{m} \left[ 1 - y_t w^\top x_t \right]_+ + \frac{\lambda}{2} \|w\|^2\) over \(\mathbb{R}^d\)

Rewrite \(F(w) = \frac{1}{m} \sum_{t=1}^{m} \ell_t(w)\) where \(\ell_t(w) = h_t(w) + \frac{\lambda}{2} \|w\|^2\)

Each loss \(\ell_t\) is \(\lambda\)-strongly convex
Exploiting curvature: minimization of SVM objective

- Training set \((x_1, y_1), \ldots, (x_m, y_m) \in \mathbb{R}^d \times \{-1, +1\}\)

- SVM objective \(F(w) = \frac{1}{m} \sum_{t=1}^{m} \left[ 1 - y_t w^\top x_t \right]_+ + \frac{\lambda}{2} \|w\|^2\) over \(\mathbb{R}^d\)

- Rewrite \(F(w) = \frac{1}{m} \sum_{t=1}^{m} \ell_t(w)\) where \(\ell_t(w) = h_t(w) + \frac{\lambda}{2} \|w\|^2\)

- Each loss \(\ell_t\) is \(\lambda\)-strongly convex

The Pegasos algorithm

- Run OGD on random sequence of \(T\) training examples

\[ \mathbb{E} \left[ F \left( \frac{1}{T} \sum_{t=1}^{T} w_t \right) \right] \leq \min_{w \in \mathbb{R}^d} F(w) + \frac{G^2 \ln T + 1}{2\lambda T} \]

\(O(\ln T)\) rates hold for any sequence of strongly convex losses
Exp-concave losses

Exp-concavity (strong convexity along the gradient direction)

- A convex $\ell : S \rightarrow \mathbb{R}$ is $\alpha$-exp-concave when $g(w) = e^{-\alpha\ell(w)}$ is concave.
- For twice-differentiable losses:
  \[ \nabla^2 \ell(w) \succeq \alpha \nabla \ell(w) \nabla \ell(w)^\top \] for all $w \in S$.
- $\ell_t(w) = -\ln (w^\top x_t)$ is exp-concave.
Online Newton Step

[Hazan, Agarwal and Kale, 2007]

- Update: \( \mathbf{w}' = A_t^{-1} \nabla \ell_t(\mathbf{w}_t) \) \( \mathbf{w}_{t+1} = \arg\min_{\mathbf{w} \in S} \|\mathbf{w} - \mathbf{w}'\|_{A_t} \)

- Where \( A_t = \epsilon I + \sum_{s=1}^{t} \nabla \ell_s(\mathbf{w}_s) \nabla \ell_s(\mathbf{w}_s)^\top \)

Note: Not an instance of OMD
Online Newton Step

[Hazan, Agarwal and Kale, 2007]

- **Update:**
  \[
  \mathbf{w}' = A_t^{-1} \nabla \ell_t(\mathbf{w}_t)
  \]
  \[
  \mathbf{w}_{t+1} = \arg\min_{\mathbf{w} \in S} \|\mathbf{w} - \mathbf{w}'\|_{A_t}
  \]

- **Where**
  \[
  A_t = \varepsilon I + \sum_{s=1}^{t} \nabla \ell_s(\mathbf{w}_s) \nabla \ell_s(\mathbf{w}_s)^	op
  \]

**Note:** Not an instance of OMD

Logarithmic regret bound for exp-concave losses

\[
R_T(\mathbf{u}) \leq 5d \left( \frac{1}{\alpha} + GD \right) \ln(T + 1) \quad \mathbf{u} \in S
\]
Online Newton Step

- Update: \( \mathbf{w}' = A_t^{-1} \nabla \ell_t(\mathbf{w}_t) \) \( \mathbf{w}_{t+1} = \text{argmin}_{\mathbf{w} \in S} \| \mathbf{w} - \mathbf{w}' \|_{A_t} \)

- Where \( A_t = \varepsilon I + \sum_{s=1}^{t} \nabla \ell_s(\mathbf{w}_s) \nabla \ell_s(\mathbf{w}_s)^\top \)

Note: Not an instance of OMD

Logarithmic regret bound for exp-concave losses

\[ R_T(\mathbf{u}) \leq 5d \left( \frac{1}{\alpha} + \text{GD} \right) \ln(T+1) \quad \mathbf{u} \in S \]

Extension of ONS to convex losses

\[ \ell_t(\mathbf{w}) \rightarrow \ell_t(\mathbf{w}^\top \mathbf{x}_t) \quad \max_t |\ell'_t| \leq L \]

\[ R_T(\mathbf{u}) \leq \tilde{O}(CL \sqrt{dT}) \quad \text{for all } \mathbf{u} \text{ s.t. } |\mathbf{u}^\top \mathbf{x}_t| \leq C \]

Invariance to linear transformations of the data
Online Ridge Regression \cite{Vovk:2001, Azoury:2001}

Logarithmic regret for square loss

\[ \ell_t(u) = (u^\top x_t - y_t)^2 \quad Y = \max_{t=1,...,T} |y_t| \quad X = \max_{t=1,...,T} \|x_t\| \]

- OMD with adaptive regularizer \( \Phi_t(w) = \frac{1}{2} \|w\|^2_{A_t} \)

- Where \( A_t = I + \sum_{s=1}^{t} x_s x_s^\top \) and \( \theta_t = \sum_{s=1}^{t} -y_s x_s \)

- Regret bound (oracle inequality)

\[
\sum_{t=1}^{T} \ell_t(w_t) \leq \inf_{u \in \mathbb{R}^d} \left( \sum_{t=1}^{T} \ell_t(u) + \|u\|^2 \right) + dY^2 \ln \left( 1 + \frac{TX^2}{d} \right)
\]

- Parameterless
- Scale-free: unbounded comparison set
Scale free algorithm for convex losses

- OMD with adaptive regularizer

\[ \Phi_t(w) = \Phi_0(w) \sqrt{\sum_{s=1}^{t-1} \| \nabla \ell_s(w_s) \|^2_*} \]

- \( \Phi_0 \) is a \( \beta \)-strongly convex base regularizer

- Regret bound (oracle inequality) for convex loss functions \( \ell_t \)

\[ \sum_{t=1}^{T} \ell_t(w_t) \leq \inf_{u \in \mathbb{R}^d} \sum_{t=1}^{T} \ell_t(u) + \left( \Phi_0(u) + \frac{1}{\beta} + \max_{t} \| \nabla \ell_t(w_t) \|_* \right) \sqrt{T} \]
Regularization via stochastic smoothing

\[ \mathbf{w}_{t+1} = \mathbb{E}_Z \left[ \arg\min_{\mathbf{w} \in \mathcal{S}} \sum_{s=1}^{t} \left( \eta \nabla \ell_s(\mathbf{w}_s) + \mathbf{Z} \right)^\top \mathbf{w} \right] \]

- The distribution of \( \mathbf{Z} \) must be “stable” (small variance and small average sensitivity)
- Regret bound similar to FTRL/OMD
- For some choices of \( \mathbf{Z} \), FPL becomes equivalent to OMD \cite{Abernethy2014}
- Linear losses: Follow the Perturbed Leader algorithm \cite{Kalai2005}
Nonstationarity

- If data source is not fitted well by any model in the class, then comparing to the best model $u \in S$ is trivial.
- Compare instead to the best sequence $u_1, u_2, \cdots \in S$ of models.

Shifting Regret for OMD

$$\sum_{t=1}^{T} \ell_t(w_t) \leq \inf_{u_1, \ldots, u_T \in S} \sum_{t=1}^{T} \ell_t(u_t) + \sum_{t=1}^{T} \|u_t - u_{t-1}\| + \text{diam}(S) + \square$$

- $\sum_{t=1}^{T} \ell_t(w_t)$: cumulative loss
- $\sum_{t=1}^{T} \ell_t(u_t)$: model fit
- $\sum_{t=1}^{T} \|u_t - u_{t-1}\|$: shifting model cost
- $\text{diam}(S)$: diameter of the set $S$
Definition

For all intervals $I = \{r, \ldots, s\}$ with $1 \leq r < s \leq T$

$$R_{T,I}(u) = \sum_{t \in I} l_t(w_t) - \sum_{t \in I} l_t(u)$$
Strongly adaptive regret

[Daniely, Gonen, Shalev-Shwartz, 2015]

**Definition**

For all intervals $I = \{r, \ldots, s\}$ with $1 \leq r < s \leq T$

$$R_{T,I}(u) = \sum_{t \in I} l_t(w_t) - \sum_{t \in I} l_t(u)$$

**Regret bound for strongly adaptive OGD**

$$R_{T,I}(u) \leq \left( BLX_2 + \ln(T + 1) \right) \sqrt{|I|} \quad \text{for all } u \text{ such that } \|u\|_2 \leq B$$
Strongly adaptive regret

[Daniely, Gonen, Shalev-Shwartz, 2015]

Definition

For all intervals $I = \{r, \ldots, s\}$ with $1 \leq r < s \leq T$

$$R_{T,I}(u) = \sum_{t \in I} \ell_t(w_t) - \sum_{t \in I} \ell_t(u)$$

Regret bound for strongly adaptive OGD

$$R_{T,I}(u) \leq \left( BLX_2 + \ln(T+1) \right) \sqrt{|I|} \quad \text{for all } u \text{ such that } \|u\|_2 \leq B$$

Remarks

- Generic black-box reduction applicable to any online learning algorithm
- It runs a logarithmic number of instances of the base learner
Online bandit convex optimization

1. Play $w_t$ from a convex and compact subset $S$ of a linear space
2. Observe $\ell_t(w_t)$, where $\ell: S \to \mathbb{R}$ is unobserved convex loss
3. Update: $w_t \rightarrow w_{t+1} \in S$

Regret: $R_T(u) = \sum_{t=1}^{T} \ell_t(w_t) - \sum_{t=1}^{T} \ell_t(u) \quad u \in S$
Online bandit convex optimization

1. Play \( w_t \) from a convex and compact subset \( S \) of a linear space
2. Observe \( \ell_t(w_t) \), where \( \ell : S \rightarrow \mathbb{R} \) is unobserved convex loss
3. Update: \( w_t \rightarrow w_{t+1} \in S \)

Regret: \( R_T(u) = \sum_{t=1}^{T} \ell_t(w_t) - \sum_{t=1}^{T} \ell_t(u) \quad u \in S \)

Results

- Linear losses: \( \Omega(d \sqrt{T}) \)  
  [Dani, Hayes, and Kakade, 2008]
Online bandit convex optimization

1. Play $w_t$ from a convex and compact subset $S$ of a linear space
2. Observe $\ell_t(w_t)$, where $\ell : S \to \mathbb{R}$ is unobserved convex loss
3. Update: $w_t \rightarrow w_{t+1} \in S$

Regret: $R_T(u) = \sum_{t=1}^{T} \ell_t(w_t) - \sum_{t=1}^{T} \ell_t(u)$, $u \in S$

Results

- Linear losses: $\Omega(d \sqrt{T})$ [Dani, Hayes, and Kakade, 2008]
- Linear losses: $\tilde{O}(d \sqrt{T})$ [Bubeck, C-B, and Kakade, 2012]

N. Cesa-Bianchi (UNIMI)
Online bandit convex optimization

1. Play $\mathbf{w}_t$ from a convex and compact subset $S$ of a linear space
2. Observe $\ell_t(\mathbf{w}_t)$, where $\ell : S \rightarrow \mathbb{R}$ is unobserved convex loss
3. Update: $\mathbf{w}_t \rightarrow \mathbf{w}_{t+1} \in S$

Regret: $R_T(\mathbf{u}) = \sum_{t=1}^{T} \ell_t(\mathbf{w}_t) - \sum_{t=1}^{T} \ell_t(\mathbf{u}) \quad \mathbf{u} \in S$

Results

- Linear losses: $\Omega(d \sqrt{T})$ [Dani, Hayes, and Kakade, 2008]
- Linear losses: $\tilde{O}(d \sqrt{T})$ [Bubeck, C-B, and Kakade, 2012]
- Strongly convex and smooth losses: $\tilde{O}(d^{3/2} \sqrt{T})$ [Hazan and Levy, 2014]
Online bandit convex optimization

1. Play $w_t$ from a convex and compact subset $S$ of a linear space
2. Observe $\ell_t(w_t)$, where $\ell : S \to \mathbb{R}$ is unobserved convex loss
3. Update: $w_t \rightarrow w_{t+1} \in S$

Regret: $R_T(u) = \sum_{t=1}^{T} \ell_t(w_t) - \sum_{t=1}^{T} \ell_t(u) \quad u \in S$

Results

- Linear losses: $\Omega(d \sqrt{T})$ [Dani, Hayes, and Kakade, 2008]
- Linear losses: $\tilde{O}(d \sqrt{T})$ [Bubeck, C-B, and Kakade, 2012]
- Strongly convex and smooth losses: $\tilde{O}(d^{3/2} \sqrt{T})$ [Hazan and Levy, 2014]
- Convex losses: $\tilde{O}(d^{9.5} \sqrt{T})$ [Bubeck, Eldan, and Lee, 2016]