Online Learning and Online Convex Optimization

Nicolò Cesa-Bianchi

Università degli Studi di Milano

Summary

Summary

Machine learning

Classification/regression tasks

- Predictive models h mapping data instances X to labels Y (e.g., binary classifier)
- Training data $S_T = \big((X_1, Y_1), \dots, (X_T, Y_T)\big)$ (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

$$\mathbb{E}\Big[\ell\big(\underbrace{A(S_T)}_{\text{trained model}}, \underbrace{(X,Y)}_{\text{test example}}\big)\Big]$$

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

Two notions of risk

View data as a statistical sample: statistical risk

$$\mathbb{E}\Big[\ell\big(\underbrace{A(S_T)}_{\text{trained model}},\ \underbrace{(X,Y)}_{\text{test example}}\big)\Big]$$

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

View data as an arbitrary sequence: sequential risk

$$\sum_{t=1}^{T} \ell\big(\underbrace{A(S_{t-1})}_{\text{trained model}},\ \underbrace{(X_t,Y_t)}_{\text{test example}}\big)$$

Sequence of models trained on growing prefixes
$S_t = \big((X_1,Y_1),\dots,(X_t,Y_t)\big)$ of the data sequence

Regrets, I had a few

Learning algorithm A maps datasets to models in a given class $\mathcal{H}$

Variance error in statistical learning

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

compare to expected loss of best model in the class

Regrets, I had a few

Learning algorithm A maps datasets to models in a given class $\mathcal{H}$

Variance error in statistical learning

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

compare to expected loss of best model in the class

Regret in online learning

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

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: $h_{t-1} \to h_t \in \mathcal{H}$ (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: $h_{t-1} \to h_t \in \mathcal{H}$ (local optimization)

Goal: control regret

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

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: $h_{t-1} \to h_t \in \mathcal{H}$ (local optimization)

Goal: control regret

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

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)

Summary

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

	1	2	$\dots$	M
1	$\ell(1,1)$	$\ell(1,2)$	$\dots$	
2	$\ell(2,1)$	$\ell(2,2)$	$\dots$	
$\vdots$	$\vdots$	$\vdots$	$\ddots$	
N				

$N \times M$ known loss matrix

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

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

- 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 $y_1, \dots, y_{t-1}$

Volodya Vovk

Manfred Warmuth

	$t = 1$	$t = 2$	$\dots$
1	$\ell_1(1)$	$\ell_2(1)$	$\dots$
2	$\ell_1(2)$	$\ell_2(2)$	$\dots$
$\vdots$	$\vdots$	$\vdots$	$\ddots$
N	$\ell_1(N)$	$\ell_2(N)$	

Opponent's moves $y_1, y_2, \dots$ 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
  $\ell_t = \big(\ell_t(1), \dots, \ell_t(N)\big) \in [0,1]^N$ for $t = 1, 2, \dots$

For $t = 1, 2, \dots$

Playing the experts game

A sequential decision problem

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

For $t = 1, 2, \dots$

1. 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
  $\ell_t = \big(\ell_t(1), \dots, \ell_t(N)\big) \in [0,1]^N$ for $t = 1, 2, \dots$

.7 .3 .2 .4 .1 .6 .7 .4 .9

For $t = 1, 2, \dots$

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

Regret analysis

Regret

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

Regret analysis

Regret

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

Lower bound using random losses [Experts' paper, 1997]

- $\ell_t(i) \to L_t(i) \in \{0,1\}$ 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

$$\mathbb{E}\left[\max_{i=1,\dots,N} \sum_{t=1}^{T} \left(\frac{1}{2} - L_t(i)\right)\right] = (1 - o(1))\sqrt{\frac{T \ln N}{2}}$$

for $N, T \to \infty$

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 $R_T \leqslant \sqrt{\frac{T \ln N}{2}}$
- Matching lower bound including constants
- Dynamic choice $\eta_t = \sqrt{(\ln N)/(8t)}$ only loses small constants

The nonstochastic bandit problem

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 $t = 1, 2, \dots$

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 $t = 1, 2, \dots$

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 agaist best offline action

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

The Exp3 algorithm [Auer et al., 2002]

Hedge with estimated losses

- $\mathbb{P}_t(I_t = i) \quad \propto \quad \exp\left(-\eta \sum_{s=1}^{t-1} \widehat{\ell}_s(i)\right) \qquad i = 1, \dots, N$
- $\widehat{\ell}_t(i) = \begin{cases} \dfrac{\ell_t(i)}{\mathbb{P}_t\big(\ell_t(i) \text{ observed}\big)} & \text{if } I_t = i \\ 0 & \text{otherwise} \end{cases}$

Only one non-zero component in $\widehat{\ell}_t$

The Exp3 algorithm [Auer et al., 2002]

Hedge with estimated losses

- $\mathbb{P}_t(I_t = i) \quad \propto \quad \exp\left(-\eta \sum_{s=1}^{t-1} \widehat{\ell}_s(i)\right) \qquad i = 1, \dots, N$
- $\widehat{\ell}_t(i) = \begin{cases} \dfrac{\ell_t(i)}{\mathbb{P}_t\big(\ell_t(i) \text{ observed}\big)} & \text{if } I_t = i \\ 0 & \text{otherwise} \end{cases}$

Only one non-zero component in $\widehat{\ell}_t$

Properties of importance weighting estimator

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

$\mathbb{E}_t\left[\widehat{\ell}_t(i)^2\right] \leqslant \dfrac{1}{\mathbb{P}_t\big(\ell_t(i) \text{ observed}\big)}$ variance control

Exp3 regret bound

$$
\begin{aligned}
R_T &\leqslant \frac{\ln N}{\eta} + \frac{\eta}{2}\,\mathbb{E}\left[\sum_{t=1}^{T}\sum_{i=1}^{N}\mathbb{P}_t(I_t = i)\mathbb{E}_t\left[\widehat{\ell}_t(i)^2\right]\right] \\
&\leqslant \frac{\ln N}{\eta} + \frac{\eta}{2}\,\mathbb{E}\left[\sum_{t=1}^{T}\sum_{i=1}^{N}\frac{\mathbb{P}_t(I_t = i)}{\mathbb{P}_t\big(\ell_t(i) \text{ is observed}\big)}\right] \\
&= \frac{\ln N}{\eta} + \frac{\eta}{2}NT = \sqrt{NT\ln N}
\end{aligned}
$$

lower bound $\Omega\big(\sqrt{NT}\big)$

Exp3 regret bound

$$R_T \leqslant \frac{\ln N}{\eta} + \frac{\eta}{2}\,\mathbb{E}\left[\sum_{t=1}^{T}\sum_{i=1}^{N}\mathbb{P}_t(I_t = i)\mathbb{E}_t\left[\widehat{\ell}_t(i)^2\right]\right]$$

$$\leqslant \frac{\ln N}{\eta} + \frac{\eta}{2}\,\mathbb{E}\left[\sum_{t=1}^{T}\sum_{i=1}^{N}\frac{\mathbb{P}_t(I_t = i)}{\mathbb{P}_t\big(\ell_t(i) \text{ is observed}\big)}\right]$$

$$= \frac{\ln N}{\eta} + \frac{\eta}{2}NT = \sqrt{NT\ln N}$$

lower bound $\Omega\big(\sqrt{NT}\big)$

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

Exp3 regret bound

$$R_T \leqslant \frac{\ln N}{\eta} + \frac{\eta}{2}\,\mathbb{E}\left[\sum_{t=1}^{T}\sum_{i=1}^{N}\mathbb{P}_t(I_t = i)\mathbb{E}_t\left[\widehat{\ell}_t(i)^2\right]\right]$$

$$\leqslant \frac{\ln N}{\eta} + \frac{\eta}{2}\,\mathbb{E}\left[\sum_{t=1}^{T}\sum_{i=1}^{N}\frac{\mathbb{P}_t(I_t = i)}{\mathbb{P}_t\big(\ell_t(i)\text{ is observed}\big)}\right]$$

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

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

The full information (experts) setting

- Player observes vector of losses ℓ_t after each play
- $\mathbb{P}_t\big(\ell_t(i)\text{ is observed}\big) = 1$
- $R_T \leqslant \sqrt{T\ln N}$

Nonoblivious opponents

The adaptive adversary

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

Nonoblivious opponents

The adaptive adversary

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

Nonoblivious regret

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

Nonoblivious opponents

The adaptive adversary

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

Nonoblivious regret

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

Policy regret

$$R_T^{\mathrm{pol}} = \mathbb{E}\left[\sum_{t=1}^{T} \ell_t(I_{t-m}, \dots, I_{t-1}, I_t) - \min_{i=1,\dots,N} \sum_{t=1}^{T} \ell_t(\underbrace{i, \dots, i}_{m \text{ times}}, i)\right]$$

Bandits and reactive opponents

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

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

- 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)

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

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

Bounds on the policy regret for any constant $m \geqslant 1$

$$R_T^{\text{pol}} = \mathcal{O}\left((N \ln N)^{1/3} T^{2/3}\right)$$

- 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

	1	2	3	4	5
1	0	1	2	3	4
2	c	0	1	2	3
3	c	c	0	1	2
4	c	c	c	0	1
5	c	c	c	c	0

Loss matrix

	1	2	3	4	5
1	1	1	1	1	1
2	0	1	1	1	1
3	0	0	1	1	1
4	0	0	0	1	1
5	0	0	0	0	1

Feedback matrix

A characterization of minimax regret

Special case

Multiarmed bandits: loss and feedback matrix are the same

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

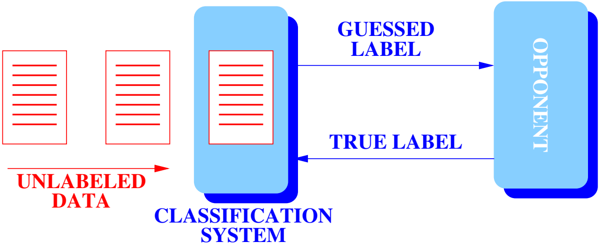
1. Play $\mathbf{p}_t$ from the N-dimensional simplex Δ_N
2. Incur linear loss $\mathbb{E}\big[\ell_t(I_t)\big] = \mathbf{p}_t^\top \boldsymbol{\ell}_t$
3. Observe loss gradient $\boldsymbol{\ell}_t$

Regret: compete against the best point in the simplex

$$\sum_{t=1}^{T} \mathbf{p}_t^\top \boldsymbol{\ell}_t - \underbrace{\min_{\mathbf{q}\in\Delta_N} \sum_{t=1}^{T} \mathbf{q}^\top \boldsymbol{\ell}_t}$$

$$= \min_{i=1,\dots,N} \frac{1}{T} \sum_{t=1}^{T} \ell_t(i)$$

From game theory to machine learning



- 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 $\boldsymbol{w}_t$ of a linear space and the opponent choosing a label y_t for $\boldsymbol{x}_t$
- Regret with respect to best element in the linear space

Summary

Online convex optimization [Zinkevich, 2003]

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

Online convex optimization [Zinkevich, 2003]

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

Example

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

Online convex optimization [Zinkevich, 2003]

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

Example

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

Regret

$$R_T(\boldsymbol{u}) = \sum_{t=1}^{T} \ell_t(\boldsymbol{w}_t) - \sum_{t=1}^{T} \ell_t(\boldsymbol{u}) \qquad \boldsymbol{u} \in S$$

Finding a good online algorithm

Follow the leader

$$w_{t+1} = \operatorname*{arginf}_{w \in S} \sum_{s=1}^{t} \ell_s(w)$$

Regret can be linear due to lack of stability

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

Finding a good online algorithm

Follow the leader

$$w_{t+1} = \operatorname*{arginf}_{w \in S} \sum_{s=1}^{t} \ell_s(w)$$

Regret can be linear due to lack of stability

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

- Note: $\sum_{s=1}^{t} \ell_s(w) = \begin{cases} -\frac{w}{2} & \text{if t is even} \\ +\frac{w}{2} & \text{if t is odd} \end{cases}$
- Hence $\ell_{t+1}(w_{t+1}) = 1$ for all $t = 1, 2 \ldots$

Follow the regularized leader

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

$$\boldsymbol{w}_{t+1} = \operatorname*{argmin}_{\boldsymbol{w} \in S} \left[\eta \sum_{s=1}^{t} \ell_s(\boldsymbol{w}) + \Phi(\boldsymbol{w}) \right]$$

Φ is a strongly convex regularizer and $\eta > 0$ is a scale parameter

Convexity, smoothness, and duality

Strong convexity

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

$$\Phi(\boldsymbol{v}) \geqslant \Phi(\mathbf{u}) + \nabla\Phi(\mathbf{u})^\top(\boldsymbol{v} - \mathbf{u}) + \frac{\beta}{2}\|\mathbf{u} - \boldsymbol{v}\|^2$$

Convexity, smoothness, and duality

Strong convexity

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

$$\Phi(\boldsymbol{v}) \geqslant \Phi(\mathbf{u}) + \nabla\Phi(\mathbf{u})^\top(\boldsymbol{v} - \mathbf{u}) + \frac{\beta}{2}\|\mathbf{u} - \boldsymbol{v}\|^2$$

Smoothness

$\Phi : S \to \mathbb{R}$ is α-smooth w.r.t. a norm $\|\cdot\|$ if for all $\mathbf{u}, \boldsymbol{v} \in S$

$$\Phi(\boldsymbol{v}) \leqslant \Phi(\mathbf{u}) + \nabla\Phi(\mathbf{u})^\top(\boldsymbol{v} - \mathbf{u}) + \frac{\alpha}{2}\|\mathbf{u} - \boldsymbol{v}\|^2$$

Convexity, smoothness, and duality

Strong convexity

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

$$\Phi(\mathbf{v}) \geqslant \Phi(\mathbf{u}) + \nabla\Phi(\mathbf{u})^\top(\mathbf{v} - \mathbf{u}) + \frac{\beta}{2}\|\mathbf{u} - \mathbf{v}\|^2$$

Smoothness

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

$$\Phi(\mathbf{v}) \leqslant \Phi(\mathbf{u}) + \nabla\Phi(\mathbf{u})^\top(\mathbf{v} - \mathbf{u}) + \frac{\alpha}{2}\|\mathbf{u} - \mathbf{v}\|^2$$

- If Φ is β-strongly convex w.r.t. $\|\cdot\|_2$, then $\nabla^2\Phi \succeq \beta I$
- If Φ is α-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. $\|\cdot\|_2$

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 $1 < p \leqslant 2$)

Examples

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

Convex duality

Definition

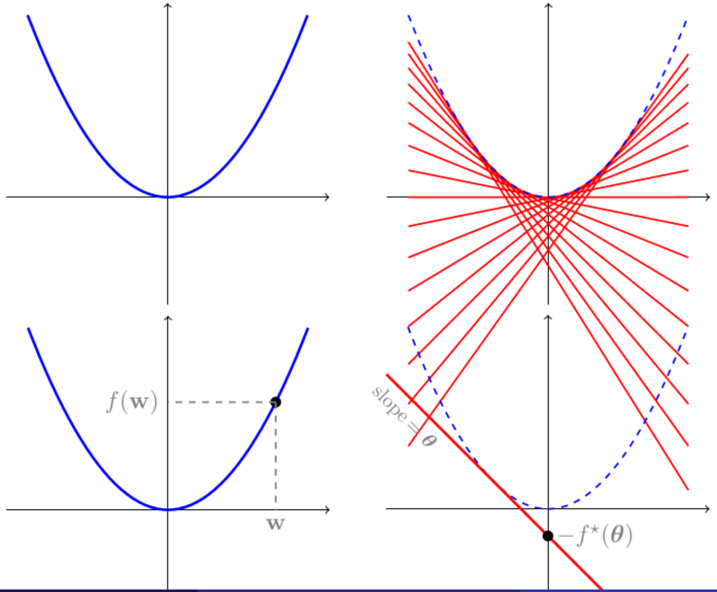
The convex dual of Φ is $\Phi^*(\theta) = \max_{w \in S} \left(\theta^\top w - \Phi(w) \right)$

1-dimensional example

- Convex $f : \mathbb{R} \to \mathbb{R}$ such that $f(0) = 0$
- $f^*(\theta) = \max_{w \in \mathbb{R}} \left(w \times \theta - f(w) \right)$
- 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)



Convexity, smoothness, and duality

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

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: $\Phi(\mathbf{p}) = \sum_{i=1}^{d} p_i \ln p_i$ and $\Phi^*(\theta) = \ln\left(e^{\theta_1} + \cdots + e^{\theta_d}\right)$
- Power norm: $\Phi(\boldsymbol{w}) = \frac{1}{2}\boldsymbol{w}^\top A \boldsymbol{w}$ and $\Phi^*(\theta) = \frac{1}{2}\theta^\top A^{-1}\theta$

Some useful properties

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

- Its convex dual Φ^* is everywhere differentiable and $\frac{1}{\beta}$-smooth w.r.t. $\|\cdot\|_*$ (the dual norm of $\|\cdot\|$)
- $\nabla\Phi^*(\theta) = \operatorname*{argmax}_{\boldsymbol{w}\in S}\left(\theta^\top \boldsymbol{w} - \Phi(\boldsymbol{w})\right)$

Some useful properties

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

- Its convex dual Φ^* is everywhere differentiable and $\frac{1}{\beta}$-smooth w.r.t. $\|\cdot\|_*$ (the dual norm of $\|\cdot\|$)
- $\nabla\Phi^*(\theta) = \underset{w \in S}{\operatorname{argmax}} \left(\theta^\top w - \Phi(w)\right)$

Recall: Follow the regularized leader (FTRL)

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

Φ is a strongly convex regularizer and $\eta > 0$ is a scale parameter

Using the loss gradient

Linearization of convex losses

$$\ell_t(\boldsymbol{w}_t) - \ell_t(\mathbf{u}) \leqslant \underbrace{\nabla \ell_t(\boldsymbol{w}_t)}_{\widetilde{\ell}_t}{}^\top \boldsymbol{w}_t - \underbrace{\nabla \ell_t(\boldsymbol{w}_t)}_{\widetilde{\ell}_t}{}^\top \mathbf{u}$$

FTRL with linearized losses

$$\boldsymbol{w}_{t+1} = \underset{\boldsymbol{w} \in S}{\operatorname{argmin}} \left(\underbrace{\eta \sum_{s=1}^{t} \widetilde{\ell}_s}_{-\theta_{t+1}}{}^\top \boldsymbol{w} + \Phi(\boldsymbol{w}) \right) = \underset{\boldsymbol{w} \in S}{\operatorname{argmax}} \left(\theta_{t+1}^\top \boldsymbol{w} - \Phi(\boldsymbol{w}) \right)$$

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

Note: $\boldsymbol{w}_{t+1} \in S$ always holds

The Mirror Descent algorithm [Nemirovsky and Yudin, 1983]

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

Online Mirror Descent (FTRL with linearized losses)

Parameters: Strongly convex regularizer Φ with domain S, $\eta > 0$

Initialize: $\theta_1 = \mathbf{0}$ `// primal parameter`

For $t = 1, 2, \ldots$

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

An equivalent formulation

Under some assumptions on the regularizer Φ, OMD can be equivalently written in terms of projected gradient descent

Online Mirror Descent (alternative version)

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

Initialize: $z_1 = \nabla\Phi^*(\mathbf{0})$ and $w_1 = \operatorname*{argmin}_{w \in S} D_\Phi(w \| z_1)$

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 $z_{t+1} = \nabla\Phi^*\Big(\nabla\Phi(z_t) - \eta\nabla\ell_t(w_t)\Big)$ `// gradient step`
4. $w_{t+1} = \operatorname*{argmin}_{w \in S} D_\Phi(w \| z_{t+1})$ `// projection step`

An equivalent formulation

Under some assumptions on the regularizer Φ, OMD can be equivalently written in terms of projected gradient descent

Online Mirror Descent (alternative version)

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

Initialize: $z_1 = \nabla\Phi^*(\mathbf{0})$ and $w_1 = \operatorname*{argmin}_{w \in S} D_\Phi(w \| z_1)$

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 $z_{t+1} = \nabla\Phi^*\Big(\nabla\Phi(z_t) - \eta\nabla\ell_t(w_t)\Big)$ `// gradient step`
4. $w_{t+1} = \operatorname*{argmin}_{w \in S} D_\Phi(w \| z_{t+1})$ `// projection step`

D_Φ is the Bregman divergence induced by Φ

Some examples

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

- $\Phi(\boldsymbol{w}) = \frac{1}{2}\|\boldsymbol{w}\|^2$ p-norm version: $\Phi(\boldsymbol{w}) = \frac{1}{2}\|\boldsymbol{w}\|_p^2$
- Update: $\boldsymbol{w}' = \boldsymbol{w}_t - \eta \nabla \ell_t(\boldsymbol{w}_t)$ $\boldsymbol{w}_{t+1} = \operatorname*{arginf}_{\boldsymbol{w} \in S} \|\boldsymbol{w} - \boldsymbol{w}'\|_2$

Some examples

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

- $\Phi(\boldsymbol{w}) = \frac{1}{2}\|\boldsymbol{w}\|^2$ — p-norm version: $\Phi(\boldsymbol{w}) = \frac{1}{2}\|\boldsymbol{w}\|_p^2$
- Update: $\boldsymbol{w}' = \boldsymbol{w}_t - \eta\nabla\ell_t(\boldsymbol{w}_t)$ $\quad \boldsymbol{w}_{t+1} = \operatorname*{arginf}_{\boldsymbol{w}\in S}\left\|\boldsymbol{w} - \boldsymbol{w}'\right\|_2$

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

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

Note: when losses are linear this is Hedge

Regret analysis

Regret bound [Kakade, Shalev-Shwartz and Tewari, 2012]

$$R_T(\mathbf{u}) \leqslant \frac{\Phi(\mathbf{u}) - \min_{\boldsymbol{w}\in S}\Phi(\boldsymbol{w})}{\eta} + \frac{\eta}{2}\sum_{t=1}^{T}\frac{\|\nabla\ell_t(\boldsymbol{w}_t)\|_*^2}{\beta}$$

for all $\mathbf{u} \in S$, where $\ell_1, \ell_2, \ldots$ are arbitrary convex losses

- $R_T(\mathbf{u}) \leqslant GD\sqrt{T}$ for all $\mathbf{u} \in S$ when η is tuned w.r.t.

$$\sup_{\boldsymbol{w}\in S}\|\nabla\ell_t(\boldsymbol{w})\|_* \leqslant G \qquad \sqrt{\sup_{\mathbf{u},\boldsymbol{w}\in S}\Big(\Phi(\mathbf{u}) - \Phi(\boldsymbol{w})\Big)} \leqslant D$$

- Boundedness of gradients of ℓ_t w.r.t. $\|\cdot\|_*$ equivalent to Lipschitzess of ℓ_t w.r.t. $\|\cdot\|$
- Regret bound optimal for general convex losses ℓ_t

Analysis relies on smoothness of Φ^*

$$\Phi^*(\theta_{t+1}) - \Phi^*(\theta_t) \leqslant \underbrace{\nabla\Phi^*(\theta_t)}_{w_t}^\top \Big(\underbrace{\theta_{t+1} - \theta_t}_{-\eta\nabla\ell_t(w_t)} \Big) + \frac{1}{2\beta} \|\theta_{t+1} - \theta_t\|_*^2$$

Analysis relies on smoothness of Φ^*

$$\Phi^*(\theta_{t+1}) - \Phi^*(\theta_t) \leqslant \underbrace{\nabla\Phi^*(\theta_t)}_{w_t}{}^\top \Big(\underbrace{\theta_{t+1} - \theta_t}_{-\eta\nabla\ell_t(w_t)} \Big) + \frac{1}{2\beta}\|\theta_{t+1} - \theta_t\|_*^2$$

$$\begin{aligned}
\sum_{t=1}^{T} -\eta \mathbf{u}^\top \nabla\ell_t(w_t) - \Phi(\mathbf{u}) &= \mathbf{u}^\top \theta_{T+1} - \Phi(\mathbf{u}) \\
&\leqslant \Phi^*(\theta_{T+1}) \qquad \text{Fenchel-Young inequality} \\
&= \sum_{t=1}^{T} \big(\Phi^*(\theta_{t+1}) - \Phi^*(\theta_t)\big) + \Phi^*(\theta_1) \\
&\leqslant \sum_{t=1}^{T} \left(-\eta w_t^\top \nabla\ell_t(w_t) + \frac{\eta^2}{2\beta}\|\nabla\ell_t(w_t)\|_*^2 \right) + \Phi^*(\mathbf{0})
\end{aligned}$$

Analysis relies on smoothness of Φ^*

$$\Phi^*(\theta_{t+1}) - \Phi^*(\theta_t) \leqslant \underbrace{\nabla\Phi^*(\theta_t)}_{w_t}^\top \Big(\underbrace{\theta_{t+1} - \theta_t}_{-\eta\nabla\ell_t(w_t)}\Big) + \frac{1}{2\beta}\|\theta_{t+1} - \theta_t\|_*^2$$

$$\begin{aligned}
\sum_{t=1}^{T} -\eta \mathbf{u}^\top \nabla\ell_t(w_t) - \Phi(\mathbf{u}) &= \mathbf{u}^\top \theta_{T+1} - \Phi(\mathbf{u}) \\
&\leqslant \Phi^*(\theta_{T+1}) \qquad \text{Fenchel-Young inequality} \\
&= \sum_{t=1}^{T} \big(\Phi^*(\theta_{t+1}) - \Phi^*(\theta_t)\big) + \Phi^*(\theta_1) \\
&\leqslant \sum_{t=1}^{T} \left(-\eta w_t^\top \nabla\ell_t(w_t) + \frac{\eta^2}{2\beta}\|\nabla\ell_t(w_t)\|_*^2\right) + \Phi^*(\mathbf{0})
\end{aligned}$$

$$\Phi^*(\mathbf{0}) = \max_{w\in S}\Big(w^\top \mathbf{0} - \Phi(w)\Big) = -\min_{w\in S}\Phi(w)$$

Some examples

$\ell_t(\boldsymbol{w}) \to \ell_t\big(\boldsymbol{w}^\top \boldsymbol{x}_t\big) \qquad \max_t |\ell_t'| \leqslant L \qquad \max_t \|\boldsymbol{x}_t\|_p \leqslant X_p$

Some examples

$\ell_t(\boldsymbol{w}) \to \ell_t(\boldsymbol{w}^\top \boldsymbol{x}_t)$ $\quad \max_t |\ell_t'| \leqslant L \quad \max_t \|\boldsymbol{x}_t\|_p \leqslant X_p$

Bounds for OGD with convex losses

$$R_T(\mathbf{u}) \leqslant BLX_2\sqrt{T} = \mathcal{O}\big(dL\sqrt{T}\big)$$

for all $\mathbf{u}$ such that $\|\mathbf{u}\|_2 \leqslant B$

Some examples

$\ell_t(\boldsymbol{w}) \to \ell_t(\boldsymbol{w}^\top \boldsymbol{x}_t) \qquad \max_t |\ell_t'| \leqslant L \qquad \max_t \|\boldsymbol{x}_t\|_p \leqslant X_p$

Bounds for OGD with convex losses

$$R_T(\boldsymbol{u}) \leqslant BLX_2\sqrt{T} = \mathcal{O}\left(dL\sqrt{T}\right)$$

for all $\boldsymbol{u}$ such that $\|\boldsymbol{u}\|_2 \leqslant B$

Bounds logarithmic in the dimension

- Regret bound for EG run in the simplex, $S = \Delta_d$

$$R_T(\boldsymbol{q}) \leqslant LX_\infty\sqrt{(\ln d)T} = \mathcal{O}\left(L\sqrt{(\ln d)T}\right) \qquad \boldsymbol{p} \in \Delta_d$$

- Same bound for p-norm regularizer with $p = \dfrac{\ln d}{\ln d - 1}$
- If losses are linear with $[0, 1]$ coefficients then we recover the bound for Hedge

Exploiting curvature: minimization of SVM objective

- Training set $(x_1, y_1), \dots, (x_m, y_m) \in \mathbb{R}^d \times \{-1, +1\}$
- SVM objective $F(w) = \dfrac{1}{m} \sum_{t=1}^{m} \underbrace{\left[1 - y_t\, w^\top x_t\right]_+}_{\text{hinge loss } h_t(w)} + \dfrac{\lambda}{2} \|w\|^2$ over $\mathbb{R}^d$
- Rewrite $F(w) = \dfrac{1}{m} \sum_{t=1}^{m} \ell_t(w)$ where $\ell_t(w) = h_t(w) + \frac{\lambda}{2} \|w\|^2$
- Each loss ℓ_t is λ-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 $F(\boldsymbol{w}) = \frac{1}{m} \sum_{t=1}^{m} \underbrace{\left[1 - y_t \boldsymbol{w}^\top x_t\right]_+}_{\text{hinge loss } h_t(\boldsymbol{w})} + \frac{\lambda}{2} \|\boldsymbol{w}\|^2$ over $\mathbb{R}^d$
- Rewrite $F(\boldsymbol{w}) = \frac{1}{m} \sum_{t=1}^{m} \ell_t(\boldsymbol{w})$ where $\ell_t(\boldsymbol{w}) = h_t(\boldsymbol{w}) + \frac{\lambda}{2} \|\boldsymbol{w}\|^2$
- Each loss ℓ_t is λ-strongly convex

The Pegasos algorithm

- Run OGD on random sequence of T training examples
- $\mathbb{E}\left[F\left(\frac{1}{T}\sum_{t=1}^{T} \boldsymbol{w}_t\right)\right] \leqslant \min_{\boldsymbol{w} \in \mathbb{R}^d} F(\boldsymbol{w}) + \frac{G^2}{2\lambda} \frac{\ln T + 1}{T}$
- $\mathcal{O}(\ln T)$ rates hold for any sequence of strongly convex losses

Exp-concave losses

Exp-concavity (strong convexity along the gradient direction)

- A convex $\ell : S \to \mathbb{R}$ is α-exp-concave when $g(\boldsymbol{w}) = e^{-\alpha \ell(\boldsymbol{w})}$ is concave
- For twice-differentiable losses: $\nabla^2 \ell(\boldsymbol{w}) \succeq \alpha \nabla \ell(\boldsymbol{w}) \nabla \ell(\boldsymbol{w})^\top$ for all $\boldsymbol{w} \in S$
- $\ell_t(\boldsymbol{w}) = -\ln\left(\boldsymbol{w}^\top \boldsymbol{x}_t\right)$ is exp-concave

Online Newton Step [Hazan, Agarwal and Kale, 2007]

- Update: $\boldsymbol{w}' = A_t^{-1} \nabla \ell_t(\boldsymbol{w}_t) \qquad \boldsymbol{w}_{t+1} = \operatorname*{argmin}_{\boldsymbol{w} \in S} \left\| \boldsymbol{w} - \boldsymbol{w}' \right\|_{A_t}$
- Where $A_t = \varepsilon I + \sum_{s=1}^{t} \nabla \ell_s(\boldsymbol{w}_s) \nabla \ell_s(\boldsymbol{w}_s)^\top$

Note: Not an instance of OMD

Online Newton Step [Hazan, Agarwal and Kale, 2007]

- Update: $\boldsymbol{w}' = A_t^{-1} \nabla \ell_t(\boldsymbol{w}_t) \qquad \boldsymbol{w}_{t+1} = \operatorname*{argmin}_{\boldsymbol{w} \in S} \|\boldsymbol{w} - \boldsymbol{w}'\|_{A_t}$
- Where $A_t = \varepsilon I + \sum_{s=1}^{t} \nabla \ell_s(\boldsymbol{w}_s) \nabla \ell_s(\boldsymbol{w}_s)^\top$

Note: Not an instance of OMD

Logarithmic regret bound for exp-concave losses

$$R_T(\boldsymbol{u}) \leqslant 5d \left(\frac{1}{\alpha} + GD \right) \ln(T+1) \qquad \boldsymbol{u} \in S$$

Online Newton Step [Hazan, Agarwal and Kale, 2007]

- Update: $\boldsymbol{w}' = A_t^{-1}\nabla\ell_t(\boldsymbol{w}_t) \qquad \boldsymbol{w}_{t+1} = \operatorname*{argmin}_{\boldsymbol{w}\in S} \left\|\boldsymbol{w}-\boldsymbol{w}'\right\|_{A_t}$
- Where $A_t = \varepsilon I + \sum_{s=1}^{t} \nabla\ell_s(\boldsymbol{w}_s)\,\nabla\ell_s(\boldsymbol{w}_s)^\top$

Note: Not an instance of OMD

Logarithmic regret bound for exp-concave losses

$$R_T(\mathbf{u}) \leqslant 5d\left(\frac{1}{\alpha}+GD\right)\ln(T+1) \qquad \mathbf{u}\in S$$

Extension of ONS to convex losses [Luo, Agarwal, C-B, Langford, 2016]

$\ell_t(\boldsymbol{w}) \to \ell_t\left(\boldsymbol{w}^\top \boldsymbol{x}_t\right) \qquad \max_t |\ell_t'| \leqslant L$

$$R_T(\mathbf{u}) \leqslant \widetilde{\mathcal{O}}\left(CL\sqrt{dT}\right) \quad \text{for all } \mathbf{u} \text{ s.t. } \left|\mathbf{u}^\top \boldsymbol{x}_t\right| \leqslant C$$

Invariance to linear transformations of the data

Online Ridge Regression [Vovk, 2001; Azoury and Warmuth, 2001]

Logarithmic regret for square loss

$\ell_t(\mathbf{u}) = \left(\mathbf{u}^\top \mathbf{x}_t - y_t\right)^2 \qquad Y = \max_{t=1,\dots,T} |y_t| \qquad X = \max_{t=1,\dots,T} \|\mathbf{x}_t\|$

- OMD with adaptive regularizer $\Phi_t(\boldsymbol{w}) = \frac{1}{2}\|\boldsymbol{w}\|^2_{A_t}$
- Where $A_t = I + \sum_{s=1}^{t} \mathbf{x}_s \mathbf{x}_s^\top$ and $\theta_t = \sum_{s=1}^{t} -y_s \mathbf{x}_s$
- Regret bound (oracle inequality)

$$\sum_{t=1}^{T} \ell_t(\boldsymbol{w}_t) \leqslant \inf_{\mathbf{u}\in\mathbb{R}^d} \left(\sum_{t=1}^{T} \ell_t(\mathbf{u}) + \|\mathbf{u}\|^2 \right) + dY^2 \ln\left(1 + \frac{TX^2}{d}\right)$$

- Parameterless
- Scale-free: unbounded comparison set

Scale free algorithm for convex losses [Orabona and Pál, 2015]

Scale free algorithm for convex losses

- OMD with adaptive regularizer

$$\Phi_t(\boldsymbol{w}) = \Phi_0(\boldsymbol{w}) \sqrt{\sum_{s=1}^{t-1} \|\nabla \ell_s(\boldsymbol{w}_s)\|_*^2}$$

- Φ_0 is a β-strongly convex base regularizer
- Regret bound (oracle inequality) for convex loss functions ℓ_t

$$\sum_{t=1}^{T} \ell_t(\boldsymbol{w}_t) \leqslant \inf_{\mathbf{u} \in \mathbb{R}^d} \sum_{t=1}^{T} \ell_t(\mathbf{u}) + \left(\Phi_0(\mathbf{u}) + \frac{1}{\beta} + \max_t \|\nabla \ell_t(\boldsymbol{w}_t)\|_* \right) \sqrt{T}$$

Regularization via stochastic smoothing

$$w_{t+1} = \mathbb{E}_Z \left[\underset{w \in S}{\operatorname{argmin}} \sum_{s=1}^{t} \left(\eta \nabla \ell_s(w_s) + Z \right)^\top w \right]$$

- The distribution of Z must be "stable" (small variance and small average sensitivity)
- Regret bound similar to FTRL/OMD
- For some choices of Z, FPL becomes equivalent to OMD [Abernethy, Lee, Sinha and Tewari, 2014]
- Linear losses: Follow the Perturbed Leader algorithm [Kalai and Vempala, 2005]

Shifting regret [Herbster and Warmuth, 2001]

Nonstationarity

- If data source is not fitted well by any model in the class, then comparing to the best model $\mathbf{u} \in S$ is trivial
- Compare instead to the best sequence $\mathbf{u}_1, \mathbf{u}_2, \dots \in S$ of models

Shifting Regret for OMD [Zinkevich, 2003]

$$\underbrace{\sum_{t=1}^{T} \ell_t(\boldsymbol{w}_t)}_{\text{cumulative loss}} \leqslant \inf_{\mathbf{u}_1,\dots,\mathbf{u}_T \in S} \underbrace{\sum_{t=1}^{T} \ell_t(\mathbf{u}_t)}_{\text{model fit}} + \underbrace{\sum_{t=1}^{T} \|\mathbf{u}_t - \mathbf{u}_{t-1}\|}_{\text{shifting model cost}} + \operatorname{diam}(S) + \square$$

Strongly adaptive regret [Daniely, Gonen, Shalev-Shwartz, 2015]

Definition

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

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

Strongly adaptive regret [Daniely, Gonen, Shalev-Shwartz, 2015]

Definition

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

$$R_{T,I}(\mathbf{u}) = \sum_{t \in I} \ell_t(\boldsymbol{w}_t) - \sum_{t \in I} \ell_t(\mathbf{u})$$

Regret bound for strongly adaptive OGD

$$R_{T,I}(\mathbf{u}) \leqslant \Big(BLX_2 + \ln(T+1)\Big)\sqrt{|I|} \qquad \text{for all } \mathbf{u} \text{ such that } \|\mathbf{u}\|_2 \leqslant B$$

Strongly adaptive regret [Daniely, Gonen, Shalev-Shwartz, 2015]

Definition

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

$$R_{T,I}(\mathbf{u}) = \sum_{t \in I} \ell_t(\boldsymbol{w}_t) - \sum_{t \in I} \ell_t(\mathbf{u})$$

Regret bound for strongly adaptive OGD

$R_{T,I}(\mathbf{u}) \leqslant \big(BLX_2 + \ln(T+1)\big)\sqrt{|I|}$ for all $\mathbf{u}$ such that $\|\mathbf{u}\|_2 \leqslant B$

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

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

Online bandit convex optimization

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

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

Results

- Linear losses: $\Omega\big(d\sqrt{T}\big)$ [Dani, Hayes, and Kakade, 2008]

Online bandit convex optimization

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

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

Results

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

Online bandit convex optimization

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

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

Results

- Linear losses: $\Omega\big(d\sqrt{T}\big)$ [Dani, Hayes, and Kakade, 2008]
- Linear losses: $\widetilde{O}\big(d\sqrt{T}\big)$ [Bubeck, C-B, and Kakade, 2012]
- Strongly convex and smooth losses: $\widetilde{O}\big(d^{3/2}\sqrt{T}\big)$ [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: $w_t \to w_{t+1} \in S$

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

Results

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