Gradient-Variation Online Learning:

Theory and Applications

Peng Zhao

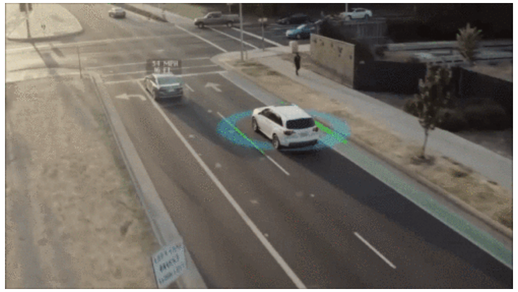
LAMDA Group

School of AI, Nanjing University

June 07, 2024

Online Learning

- In more and more applications, data are coming in an ***online*** fashion



autonomous driving

intelligence manufacturing

base station communicate

- Online learning/optimization

  - update the model in an iterated optimization fashion
  - try to equip with theoretical guarantees for the online update

Online Convex Optimization

- View online learning as a game between ***learner*** and ***environment***.

At each round $t = 1, 2, \ldots, T$:

- the learner submits a decision (model) $\mathbf{x}_t \in \mathcal{X} \subseteq \mathbb{R}^d$
- at the same time, environments decide an online function $f_t : \mathcal{X} \mapsto \mathbb{R}$
- the learner suffers $f_t(\mathbf{x}_t)$ and receives gradient information

Example: Considering ***online classification***, we have

(i) predictive loss $\ell : \widehat{\mathcal{Y}} \times \mathcal{Y} \mapsto \mathbb{R}$, and

(ii) hypothesis function $h : \mathcal{X} \times \Psi \mapsto \widehat{\mathcal{Y}}$.

$$\Longrightarrow \quad \begin{aligned} f_t(\mathbf{x}) &= \ell(h(\mathbf{x}; \boldsymbol{\psi}_t), y_t) \\ &= \ell(\mathbf{x}^\top \boldsymbol{\psi}_t, y_t) \quad \textit{for simplicity} \end{aligned}$$

Online Convex Optimization

- **Regret**: online prediction as good as the best offline model

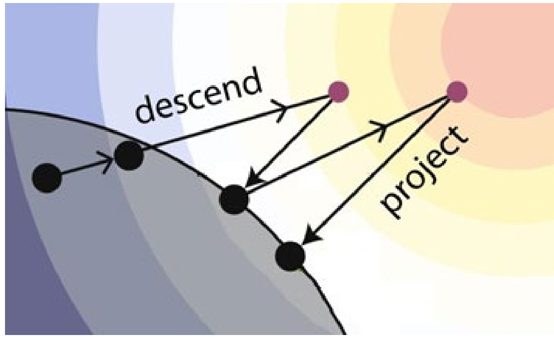
$$\mathrm{Reg}_T = \sum_{t=1}^{T} f_t(\mathbf{x}_t) - \min_{\mathbf{x}\in\mathcal{X}} \sum_{t=1}^{T} f_t(\mathbf{x})$$

learner's excess loss compared to the best offline model

- Many efforts have been devoted to regret minimization.
  - ❑ Online Gradient Descent (OGD)

$$\mathbf{x}_{t+1} = \Pi_{\mathcal{X}}\left[\mathbf{x}_t - \eta_t \nabla f_t(\mathbf{x}_t)\right]$$

$\Pi_{\mathcal{X}}[\cdot]$ denotes Euclidean projection onto feasible domain $\mathcal{X}$.



https://www.nature.com/articles/s41534-017-0043-1

  - ❑ Other frameworks include online mirror descent (OMD) and Follow-the-Regularized-Leader (FTRL).

Online Convex Optimization

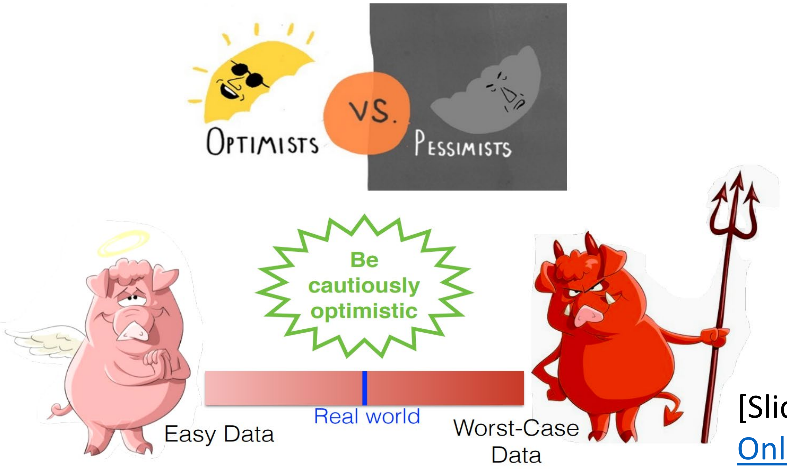
- **Existing efforts**: obtain the ***minimax optimal*** regret guarantees

e.g., OGD with step size $\eta_t = 1/\sqrt{t}$

$$\mathbf{x}_{t+1} = \Pi_{\mathcal{X}}\left[\mathbf{x}_t - \eta_t \nabla f_t(\mathbf{x}_t)\right] \Longrightarrow \sum_{t=1}^{T} f_t(\mathbf{x}_t) - \min_{\mathbf{x}\in\mathcal{X}} \sum_{t=1}^{T} f_t(\mathbf{x}) \leq \mathcal{O}(\sqrt{T})$$

This regret bound is optimal in the worst case --- there exists a hard instance that no online algorithm can behave better than this regret rate.

- **However**, in many cases, the environments may not so hard!



[Slides from Dylan Foster, Adaptive Online Learning @NIPS'15 workshop]

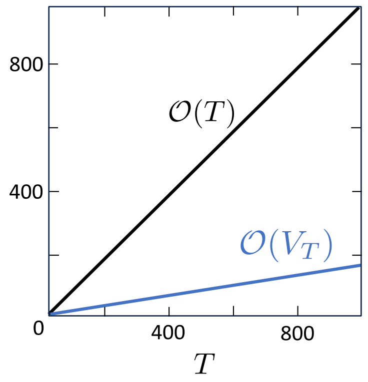
Beyond Worst-Case Analysis

- ***Problem-dependent*** methods aim to achieve more adaptive results:
  (i) regret guarantee should be much improved for ***easy*** environments ;
  (ii) meanwhile can safeguard the minimax optimal rate in the ***worst case***.

Gradient variation:

$$V_T \triangleq \sum_{t=2}^{T} \sup_{\mathbf{x}\in\mathcal{X}} \|\nabla f_t(\mathbf{x}) - \nabla f_{t-1}(\mathbf{x})\|_2^2$$

which measures cumulative variations in gradients.



can be $O(1)$ when functions are stable, and is at most $O(T)$ in the worst case (under standard assumption)

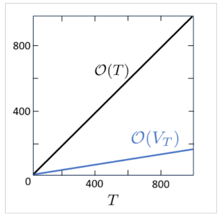
Why Gradient Variation?

- Consider convex functions, if we achieve an $\mathcal{O}(\sqrt{V_T})$ regret,
  - ❑ become even $\mathcal{O}(1)$ when functions are stable
  - ❑ safeguard the $\mathcal{O}(\sqrt{T})$ rate in the worst case

Gradient variation:

$$V_T \triangleq \sum_{t=2}^{T} \sup_{\mathbf{x}\in\mathcal{X}} \|\nabla f_t(\mathbf{x}) - \nabla f_{t-1}(\mathbf{x})\|_2^2$$

which measures cumulative variations in gradients.



- More importantly, it has profound connections to other areas.
  - ❑ **Game theory:** GV regret is essential for obtaining fast convergence in games.
    [Syrgkanis et al., NIPS'15 best paper; Wei & Luo, COLT'18]
  - ❑ **Adaptive Optimization:** GV regret can bridge stochastic and adversarial optimization.
    [Sachs et al., NeurIPS'22; Chen et al., ICML'23]

Implications: Game Theory

- **Gradient Variation in *Games*:** [Syrgkanis et al., NIPS'15 best paper]

Two-Player Zero-Sum Games

y-player decision $\mathbf{y}_t = (\frac{1}{2}, \frac{1}{2}, 0)^\top$

x-player decision $\mathbf{x}_t = \begin{pmatrix} 0 \\ 1/2 \\ 1/2 \end{pmatrix}$

Game payoff matrix *A*

	Rock	Scissors	Paper
Rock	(0,0)	(1,−1)	(−1,1)
Scissors	(−1,1)	(0,0)	(−1,1)
Paper	(1,−1)	(−1,1)	(0,0)

*A good situation for two players is to achieve **Nash equilibrium**; but its computation is often hard.*

Implications: Game Theory

- **Gradient Variation in *Games***: [Syrgkanis et al., NIPS'15 best paper]

Online Game Protocol

The environments decide a payoff matrix A

At each round $t = 1, 2, \ldots, T$:

- x-player submits $\mathbf{x}_t \in \Delta_d$ and y-player submits $\mathbf{y}_t \in \Delta_d$
- x-player suffers loss $\mathbf{x}_t^\top A \mathbf{y}_t$ and receives gradient $A\mathbf{y}_t$, and y-player receives reward $\mathbf{x}_t^\top A \mathbf{y}_t$ and receives gradient $\mathbf{x}_t^\top A$

The online function that x-player receives is $f_t^x(\cdot) \triangleq \cdot^\top A \mathbf{y}_t$.

consider a relaxed gradient feedback (we don't even require full knowledge of A)

Implications: Game Theory

- Gradient-variation online learning for ***fast convergence in games***.

Deploying ***gradient-variation algorithm*** (e.g., optimistic mirror descent with last-round gradient) attains:

$$f_t^x(\mathbf{x}) \triangleq \mathbf{x}^\top A \mathbf{y}_t \qquad f_{t-1}^x(\mathbf{x}) \triangleq \mathbf{x}^\top A \mathbf{y}_{t-1}$$

$$\mathrm{Reg}_T^x \lesssim 1 + \underbrace{\sum_{t=2}^T \|A\mathbf{y}_t - A\mathbf{y}_{t-1}\|_\infty^2}_{\text{gradient variation}} \underbrace{- \sum_{t=2}^T \|\mathbf{x}_t - \mathbf{x}_{t-1}\|_1^2}_{\text{negative stability}}$$

Deploying ***gradient-variation algorithm*** (e.g., optimistic mirror descent with last-round gradient) attains:

$$f_t^y(\mathbf{y}) \triangleq \mathbf{x}_t^\top A \mathbf{y} \qquad f_{t-1}^y(\mathbf{y}) \triangleq \mathbf{x}_{t-1}^\top A \mathbf{y}$$

$$\mathrm{Reg}_T^y \lesssim 1 + \underbrace{\sum_{t=2}^T \|\mathbf{x}_t^\top A - \mathbf{x}_{t-1}^\top A\|_\infty^2}_{\text{gradient variation}} \underbrace{- \sum_{t=2}^T \|\mathbf{y}_t - \mathbf{y}_{t-1}\|_1^2}_{\text{negative stability}}$$

Regret summation *is usually related to global metrics in games,*
such as Nash equilibrium regret (measuring the quality of approximated solution).

Implications: Game Theory

- Gradient-variation online learning for ***fast convergence in games***.

Deploying ***gradient-variation algorithm*** (e.g., optimistic mirror descent with last-round gradient) attains:

$$f_t^x(\mathbf{x}) \triangleq \mathbf{x}^\top A \mathbf{y}_t \qquad f_{t-1}^x(\mathbf{x}) \triangleq \mathbf{x}^\top A \mathbf{y}_{t-1}$$

$$\mathrm{Reg}_T^x \lesssim 1 + \underbrace{\sum_{t=2}^T \|A\mathbf{y}_t - A\mathbf{y}_{t-1}\|_\infty^2}_{\text{gradient variation}} \underbrace{- \sum_{t=2}^T \|\mathbf{x}_t - \mathbf{x}_{t-1}\|_1^2}_{\text{negative stability}}$$

Deploying ***gradient-variation algorithm*** (e.g., optimistic mirror descent with last-round gradient) attains:

$$f_t^y(\mathbf{y}) \triangleq \mathbf{x}_t^\top A \mathbf{y} \qquad f_{t-1}^y(\mathbf{y}) \triangleq \mathbf{x}_{t-1}^\top A \mathbf{y}$$

$$\mathrm{Reg}_T^y \lesssim 1 + \underbrace{\sum_{t=2}^T \|\mathbf{x}_t^\top A - \mathbf{x}_{t-1}^\top A\|_\infty^2}_{\text{gradient variation}} \underbrace{- \sum_{t=2}^T \|\mathbf{y}_t - \mathbf{y}_{t-1}\|_1^2}_{\text{negative stability}}$$

$\Longrightarrow$ $\mathrm{Reg}_T^x + \mathrm{Reg}_T^y \leq \mathcal{O}(1)$ *which leads to* $\mathcal{O}(\frac{1}{T})$ *fast convergence, as opposed to* $\mathcal{O}(\frac{1}{\sqrt{T}})$.

Implications: min-max application

Learning And Mining from DatA

Similar idea is further used for training GAN (Generative Adversarial Networks) and its theoretical explanations.

ICLR 2018

TRAINING GANS WITH OPTIMISM

Constantinos Daskalakis*
MIT, EECS
costis@mit.edu

Andrew Ilyas*
MIT, EECS
ailyas@mit.edu

Vasilis Syrgkanis*
Microsoft Research
vasy@microsoft.com

Haoyang Zeng*
MIT, EECS
haoyangz@mit.edu

ABSTRACT

We address the issue of limit cycling behavior in training Generative Adversarial Networks and propose the use of Optimistic Mirror Decent (OMD) for training Wasserstein GANs. Recent theoretical results have shown that optimistic mirror decent (OMD) can enjoy faster regret rates in the context of zero-sum games. WGANs is exactly a context of solving a zero-sum game with simultaneous no-regret dynamics. Moreover, we show that optimistic mirror decent addresses the limit cycling problem in training WGANs. We formally show that in the case of bi-linear zero-sum games the last iterate of OMD dynamics converges to an equilibrium, in contrast to GD dynamics which are bound to cycle. We also portray the huge qualitative difference between GD and OMD dynamics with toy examples, even when GD is modified with many adaptations proposed in the recent literature, such as gradient penalty or momentum. We apply OMD WGAN training to a bioinformatics problem of generating DNA sequences. We observe that models trained with OMD achieve consistently smaller KL divergence with respect to the true underlying distribution, than models trained with GD variants. Finally, we introduce a new algorithm, Optimistic Adam, which is an optimistic variant of Adam. We apply it to WGAN training on CIFAR10 and observe improved performance in terms of inception score as compared to Adam.

ICLR 2019

OPTIMISTIC MIRROR DESCENT IN SADDLE-POINT PROBLEMS: GOING THE EXTRA (GRADIENT) MILE

Panayotis Mertikopoulos
Univ. Grenoble Alpes, CNRS, Inria, Grenoble INP, LIG
38000 Grenoble, France
panayotis.mertikopoulos@imag.fr

Bruno Lecouat, Houssam Zenati, Chuan-Sheng Foo, Vijay Chandrasekhar
Institute for Infocomm Research, A*STAR
1 Fusionopolis Way, #21-01 Connexis (South Tower), Singapore
{bruno_lecouat,houssam_zenati,foocs,vijay}@i2r.a-star.edu.sg

Georgios Piliouras
Singapore University of Technology and Design
8 Somapah Road, Singapore
georgios@sutd.edu.sg

ABSTRACT

Owing to their connection with generative adversarial networks (GANs), saddle-point problems have recently attracted considerable interest in machine learning and beyond. By necessity, most theoretical guarantees revolve around convex-concave (or even linear) problems; however, making theoretical inroads towards efficient GAN training depends crucially on moving beyond this classic framework. To make piecemeal progress along these lines, we analyze the behavior of mirror descent (MD) in a class of non-monotone problems whose solutions coincide with those of a naturally associated variational inequality – a property which we call *coherence*. We first show that ordinary, "vanilla" MD converges under a strict version of this condition, but not otherwise; in particular, it may fail to converge even in bilinear models with a unique solution. We then show that this deficiency is mitigated by optimism: by taking an "extra-gradient" step, optimistic mirror descent (OMD) converges in *all* coherent problems. Our analysis generalizes and extends the results of Daskalakis et al. [2018] for optimistic gradient descent (OGD) in bilinear problems, and makes concrete headway for provable convergence beyond convex-concave games. We also provide stochastic analogues of these results, and we validate our analysis by numerical experiments in a wide array of GAN models (including Gaussian mixture models, and the CelebA and CIFAR-10 datasets).

Implications: Adaptive Optimization

Learning And Mining from DatA

- **Gradient Variation in *Stochastic/Adversarial Optimization* :**

[Sachs et al., NeurIPS'22]

$$\min_{\mathbf{x}} f_t(\mathbf{x})$$

How is loss function f_t generated?

$\xrightarrow{\text{stochastic}}$ stochastic optimization $f \sim \mathcal{D}$

$\xrightarrow{\text{adversarial}}$ adversarial optimization $\{f_t\}_{t=1}^{T}$

➢ These two fields are previously studied ***separately***.

➢ Recent works reveal the essential role of ***gradient variation***, which provides an ***adaptive interpolation*** between stochastic and adversarial optimization.

Implications: Adaptive Optimization

LAMDA
Learning And Mining from DatA

- **SEA (Stochastically Extended Adversarial) model** [Sachs et al., NeurIPS'22]

Setup: at round $t \in [T]$, SEA optimizes $\min_{\mathbf{x} \in \mathcal{X}} f_t(\mathbf{x})$

f_t is the *randomized function* sampled from underlying distribution $\mathcal{D}_t$: $f_t \sim \mathcal{D}_t$

F_t is the *expected function* of f_t: $F_t(\cdot) \triangleq \mathbb{E}_{f_t \sim \mathcal{D}_t}[f_t(\cdot)]$

Two complexity measures:

$$\sigma_{1:T}^2 \triangleq \underbrace{\sum_{t=1}^{T} \max_{\mathbf{x} \in \mathcal{X}} \mathbb{E}_{f_t \sim \mathcal{D}_t}[\|\nabla f_t(\mathbf{x}) - \nabla F_t(\mathbf{x})\|^2]}_{\textit{stochastic variance}}, \quad \Sigma_{1:T}^2 \triangleq \underbrace{\mathbb{E}\left[\sum_{t=2}^{T} \sup_{\mathbf{x} \in \mathcal{X}} \|\nabla F_t(\mathbf{x}) - \nabla F_{t-1}(\mathbf{x})\|^2\right]}_{\textit{adversarial change}}$$

Implications: Adaptive Optimization

Learning And Mining from DatA

- **SEA (Stochastically Extended Adversarial) model** [Sachs et al., NeurIPS'22]

Setup: at round $t \in [T]$, SEA optimizes $\min_{\mathbf{x}\in\mathcal{X}} f_t(\mathbf{x})$

f_t is the *randomized function* sampled from underlying distribution $\mathcal{D}_t$: $f_t \sim \mathcal{D}_t$

F_t is the *expected function* of f_t: $F_t(\cdot) \triangleq \mathbb{E}_{f_t\sim\mathcal{D}_t}[f_t(\cdot)]$

$\Rightarrow$ SEA model can be solved by deploying gradient-variation algorithm over the randomized function $\{f_t\}_{t=1}^T$.

$$\underbrace{\nabla f_t(\mathbf{x})-\nabla f_{t-1}(\mathbf{x})}_{\text{gradient variation}} = \underbrace{[\nabla f_t(\mathbf{x})-\nabla F_t(\mathbf{x})]}_{\text{stochastic variance}}+\underbrace{[\nabla F_t(\mathbf{x})-\nabla F_{t-1}(\mathbf{x})]}_{\text{adversarial change}}+\underbrace{[\nabla F_{t-1}(\mathbf{x})-\nabla f_{t-1}(\mathbf{x})]}_{\text{stochastic variance}}$$

Approximately $V_T \approx \sigma_{1:T}^2 + \Sigma_{1:T}^2$.

For stochastic optimization, $\sigma_{1:T}^2 = \sigma^2 T$ and $\Sigma_{1:T}^2 = 0$.

For adversarial optimization, $\sigma_{1:T}^2 = 0$ and $\Sigma_{1:T}^2 = V_T$.

Implications: Online Label Shift

Similar idea was also used for dealing with online label shift adaptation.

NeurIPS 2022

Adapting to Online Label Shift with Provable Guarantees

Yong Bai[1*], **Yu-Jie Zhang**[2,1*], **Peng Zhao**[1], **Masashi Sugiyama**[3,2], **Zhi-Hua Zhou**[1†]
[1] National Key Laboratory for Novel Software Technology, Nanjing University, Nanjing, China
[2] The University of Tokyo, Chiba, Japan
[3] RIKEN AIP, Tokyo, Japan

Abstract

The standard supervised learning paradigm works effectively when training data shares the same distribution as the upcoming testing samples. However, this stationary assumption is often violated in real-world applications, especially when testing data appear in an online fashion. In this paper, we formulate and investigate the problem of *online label shift* (OLaS): the learner trains an initial model from the labeled offline data and then deploys it to an unlabeled online environment where the underlying label distribution changes over time but the label-conditional density does not. The non-stationarity nature and the lack of supervision make the problem challenging to be tackled. To address the difficulty, we construct a new unbiased risk estimator that utilizes the unlabeled data, which exhibits many benign properties albeit with potential non-convexity. Building upon that, we propose novel online ensemble algorithms to deal with the non-stationarity of the environments. Our approach enjoys optimal *dynamic regret*, indicating that the performance is competitive with a clairvoyant who knows the online environments in hindsight and then chooses the best decision for each round. The obtained dynamic regret bound scales with the intensity and pattern of label distribution shift, hence exhibiting the adaptivity in the OLaS problem. Extensive experiments are conducted to validate the effectiveness and support our theoretical findings.

NeurIPS 2023

Online Label Shift: Optimal Dynamic Regret meets Practical Algorithms

Dheeraj Baby*
UC Santa Barbara
dheeraj@ucsb.edu

Saurabh Garg*
Carnegie Mellon University
sgarg2@andrew.cmu.edu

Tzu-Ching Yen*
Carnegie Mellon University
tzuchiny@andrew.cmu.edu

Sivaraman Balakrishnan
Carnegie Mellon University
sbalakri@andrew.cmu.edu

Zachary C. Lipton
Carnegie Mellon University
zlipton@andrew.cmu.edu

Yu-Xiang Wang
UC Santa Barbara
yuxiangw@cs.ucsb.edu

Abstract

This paper focuses on supervised and unsupervised online label shift, where the class marginals $Q(y)$ varies but the class-conditionals $Q(x|y)$ remain invariant. In the unsupervised setting, our goal is to adapt a learner, trained on some offline labeled data, to changing label distributions given unlabeled online data. In the supervised setting, we must both learn a classifier and adapt to the dynamically evolving class marginals given only labeled online data. We develop novel algorithms that reduce the adaptation problem to online regression and guarantee optimal dynamic regret without any prior knowledge of the extent of drift in the label distribution. Our solution is based on bootstrapping the estimates of *online regression oracles* that track the drifting proportions. Experiments across numerous simulated and real-world online label shift scenarios demonstrate the superior performance of our proposed approaches, often achieving 1-3% improvement in accuracy while being sample and computationally efficient. Code is publicly available at this url.

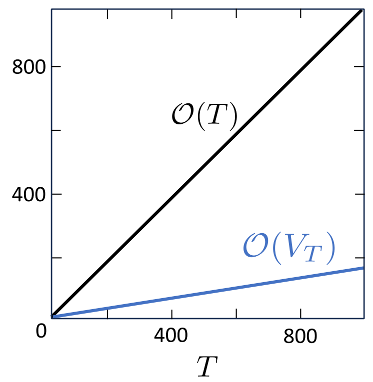
Gradient-Variation Online Learning

- The importance of problem-dependent adaptivity.
  - (i) regret guarantee should be much improved for ***easy*** environments ;
  - (ii) meanwhile can safeguard the minimax optimal rate in the ***worst case***.

Gradient variation:

$$V_T \triangleq \sum_{t=2}^{T} \sup_{\mathbf{x}\in\mathcal{X}} \|\nabla f_t(\mathbf{x}) - \nabla f_{t-1}(\mathbf{x})\|^2$$

which measures cumulative variations in gradients.



- Profound connections to other important problems.

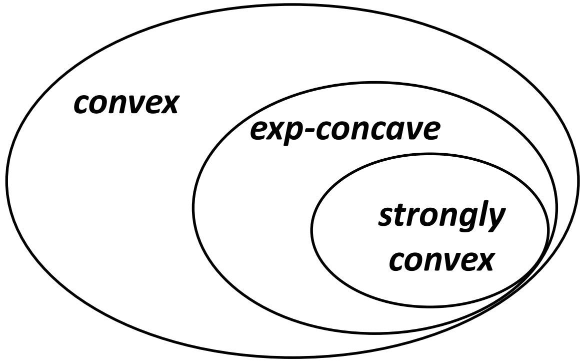
  (i) Game theory, (ii) Adaptive optimization, etc.

Modern Online Learning

Key requirement: ***robustness*** to uncertain environments

- **Universal online learning**
  - For different function cases, targeting algorithm should be used (which may confuse users);

Function type	Algorithm	Regret
convex	OGD with $\eta_t \approx \frac{1}{\sqrt{t}}$	$\mathcal{O}(\sqrt{T})$
λ-strongly convex	OGD with $\eta_t = \frac{1}{\lambda t}$	$\mathcal{O}(\frac{1}{\lambda}\log T)$
α-exp-concave	Online Newton Step with α	$\mathcal{O}(\frac{1}{\alpha} d \log T)$



  - Design a ***single*** algorithm capable of handling different types of functions, while achieving the same regret as if they were known.

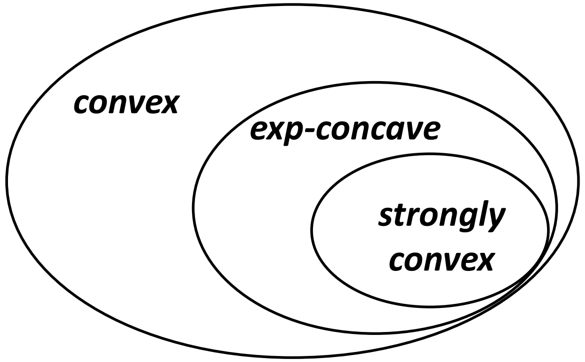
⇨ A single algorithm achieves $\mathcal{O}(\sqrt{T})$, $\mathcal{O}(d\log T)$, and $\mathcal{O}(\log T)$ regret for convex/ exp-concave/strongly convex functions, respectively.

Modern Online Learning

Key requirement: ***robustness*** to uncertain environments

- **Universal online learning**
  - ❑ For different function cases, targeting algorithm should be used (which may confuse users);

Function type	Algorithm	Regret
convex	OGD with $\eta_t \approx \frac{1}{\sqrt{t}}$	$\mathcal{O}(\sqrt{T})$
λ-strongly convex	OGD with $\eta_t = \frac{1}{\lambda t}$	$\mathcal{O}(\frac{1}{\lambda}\log T)$
α-exp-concave	Online Newton Step with α	$\mathcal{O}(\frac{1}{\alpha} d \log T)$



  - ❑ Design a ***single*** algorithm capable of handling different types of functions, while achieving the same regret as if they were known.

GV regret ⇒ A single algorithm achieves $\mathcal{O}(\sqrt{V_T})$, $\mathcal{O}(d \log V_T)$, and $\mathcal{O}(\log V_T)$ regret for convex/ exp-concave/strongly convex functions, respectively.

Modern Online Learning

LAMDA Learning And Mining from DatA

Key requirement: ***robustness*** to uncertain environments

- **Non-stationary online learning**
  - For many scenarios, the environments exhibit non-stationarity, such as distribution shift.

$$\mathrm{Reg}_T \triangleq \sum_{t=1}^{T} f_t(\mathbf{x}_t) - \min_{\mathbf{x}\in\mathcal{X}} \sum_{t=1}^{T} f_t(\mathbf{x})$$

cumulative loss of best offline model

  - Design a robust algorithm capable of minimizing ***dynamic regret***

$$\text{D-Reg}_T(\mathbf{u}_1,\ldots,\mathbf{u}_T) \triangleq \sum_{t=1}^{T} f_t(\mathbf{x}_t) - \sum_{t=1}^{T} f_t(\mathbf{u}_t)$$

optimal model changes in non-stationary environments

$P_T = \sum_{t=2}^{T} \|\mathbf{u}_t - \mathbf{u}_{t-1}\|_2$ *measures the non-stationary level*

⇨ An algorithm equips with dynamic regret scaling with P_T, such as $\mathcal{O}(\sqrt{T(1+P_T)})$.

Modern Online Learning

Key requirement: ***robustness*** to uncertain environments

- **Non-stationary online learning**
  - For many scenarios, the environments exhibit non-stationarity, such as distribution shift.

$$\mathrm{Reg}_T \triangleq \sum_{t=1}^{T} f_t(\mathbf{x}_t) - \min_{\mathbf{x}\in\mathcal{X}} \sum_{t=1}^{T} f_t(\mathbf{x})$$

cumulative loss of best offline model

  - Design a robust algorithm capable of minimizing ***dynamic regret***

$$\text{D-Reg}_T(\mathbf{u}_1,\ldots,\mathbf{u}_T) \triangleq \sum_{t=1}^{T} f_t(\mathbf{x}_t) - \sum_{t=1}^{T} f_t(\mathbf{u}_t)$$

optimal model changes in non-stationary environments

$P_T = \sum_{t=2}^{T} \|\mathbf{u}_t - \mathbf{u}_{t-1}\|_2$ *measures the non-stationary level*

GV regret

⇨ An algorithm equips with GV-dynamic regret, such as $\mathcal{O}(\sqrt{V_T(1+P_T)})$.

Guiding Questions

- Developing ***gradient-variation regret*** for modern online learning.

❑ Universal online learning

⇨ A single algorithm achieves $\mathcal{O}(\sqrt{V_T})$, $\mathcal{O}(d \log V_T)$, and $\mathcal{O}(\log V_T)$ regret for convex/ exp-concave/strongly convex functions, respectively.

❑ Non-stationary online learning

⇨ An algorithm equips with GV-dynamic regret, such as $\mathcal{O}(\sqrt{V_T(1+P_T)})$.

We successfully achieved above gradient-variation results and obtained many implications.

Review: GV regret in simple case

- **Gentle start:** convex functions *(not universal)*, standard regret *(not non-stationary)*

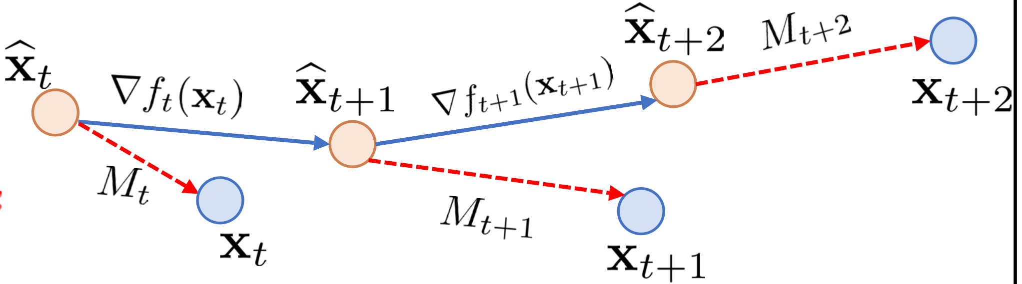
Optimistic Update: to leverage historical (gradient) information

Optimistic OGD [Chiang et al., COLT'12; Rakhlin & Sridharan, COLT'13]

$$\mathbf{x}_t = \Pi_{\mathcal{X}}\left[\widehat{\mathbf{x}}_t - \eta M_t\right]$$

$$\widehat{\mathbf{x}}_{t+1} = \Pi_{\mathcal{X}}\left[\widehat{\mathbf{x}}_t - \eta \nabla f_t(\mathbf{x}_t)\right]$$

where $\{M_t \in \mathbb{R}^d\}_{t=1}^T$ is a sequence of *optimistic vectors* serving as available prediction of *future gradients*



$$\Longrightarrow \sum_{t=1}^T f_t(\mathbf{x}_t) - \sum_{t=1}^T f_t(\mathbf{u}) \leq \mathcal{O}\left(\sqrt{1 + \sum_{t=1}^T \|\nabla f_t(\mathbf{x}_t) - M_t\|_2^2}\right)$$

regret performance depends on the quality of the optimism M_t

Review: GV regret in simple case

Learning And Mining from DatA

- A general optimistic bound

$$\sum_{t=1}^{T} f_t(\mathbf{x}_t) - \sum_{t=1}^{T} f_t(\mathbf{u}) \leq \mathcal{O}\left(\sqrt{1 + \sum_{t=1}^{T} \|\nabla f_t(\mathbf{x}_t) - M_t\|_2^2}\right)$$

choose last-round gradient as optimistic vector
$M_t = \nabla f_{t-1}(\mathbf{x}_{t-1})$

$$\bar{V}_T \triangleq \sum_{t=2}^{T} \|\nabla f_t(\mathbf{x}_t) - \nabla f_{t-1}(\mathbf{x}_{t-1})\|_2^2$$

(empirical) gradient variation

- By smoothness (of online functions), empirical gradient variation can be upper bounded by

$$\sum_{t=2}^{T} \|\mathbf{x}_t - \mathbf{x}_{t-1}\|^2$$

algorithmic stability

and

$$V_T \triangleq \sum_{t=2}^{T} \sup_{\mathbf{x}\in\mathcal{X}} \|\nabla f_t(\mathbf{x}) - \nabla f_{t-1}(\mathbf{x})\|^2$$

gradient variation

key quantity to handle!

Review: GV regret in simple case

Learning And Mining from DatA

- Surprisingly, many online algorithms enjoy ***negative terms*** in regret analysis.

Optimistic OGD

$$\mathrm{Reg}_T \leq \eta \sum_{t=1}^{T} \|\nabla f_t(\mathbf{x}_t) - \nabla f_{t-1}(\mathbf{x}_{t-1})\|_2^2 + \frac{D^2}{2\eta} - \frac{1}{4\eta} \sum_{t=1}^{T} \|\mathbf{x}_{t+1} - \mathbf{x}_t\|_2^2$$

$$\leq 2\eta \sum_{t=2}^{T} \|\nabla f_t(\mathbf{x}_t) - \nabla f_{t-1}(\mathbf{x}_t)\|_2^2 + 2\eta \sum_{t=2}^{T} \|\nabla f_{t-1}(\mathbf{x}_t) - \nabla f_{t-1}(\mathbf{x}_{t-1})\|_2^2 + \frac{D^2}{2\eta} - \frac{1}{4\eta} \sum_{t=1}^{T} \|\mathbf{x}_{t+1} - \mathbf{x}_t\|_2^2$$

$$\leq 2\eta V_T + 2\eta L^2 \sum_{t=2}^{T} \|\mathbf{x}_t - \mathbf{x}_{t-1}\|_2^2 + \frac{D^2}{2\eta} - \frac{1}{4\eta} \sum_{t=1}^{T} \|\mathbf{x}_{t+1} - \mathbf{x}_t\|_2^2 \quad \text{(by } L\text{-smoothness of online functions)}$$

positive terms are cancelled out when step size is small enough ($\eta \leq 1/L$)

COLT 2012
best student paper

Chiang, et al. Online optimization with gradual variations. COLT 2012.

Challenge In Modern Online Learning

- Developing *gradient-variation regret* for modern online learning.

Universal online learning

⇨ A single algorithm achieves $\mathcal{O}(\sqrt{V_T})$, $\mathcal{O}(d \log V_T)$, and $\mathcal{O}(\log V_T)$ regret for convex/ exp-concave/strongly convex functions, respectively.

Non-stationary online learning

⇨ An algorithm equips with GV-dynamic regret, such as $\mathcal{O}(\sqrt{V_T(1+P_T)})$.

Key challenge: ***robustness*** to uncertain environments and ***adaptivity*** to gradient variation

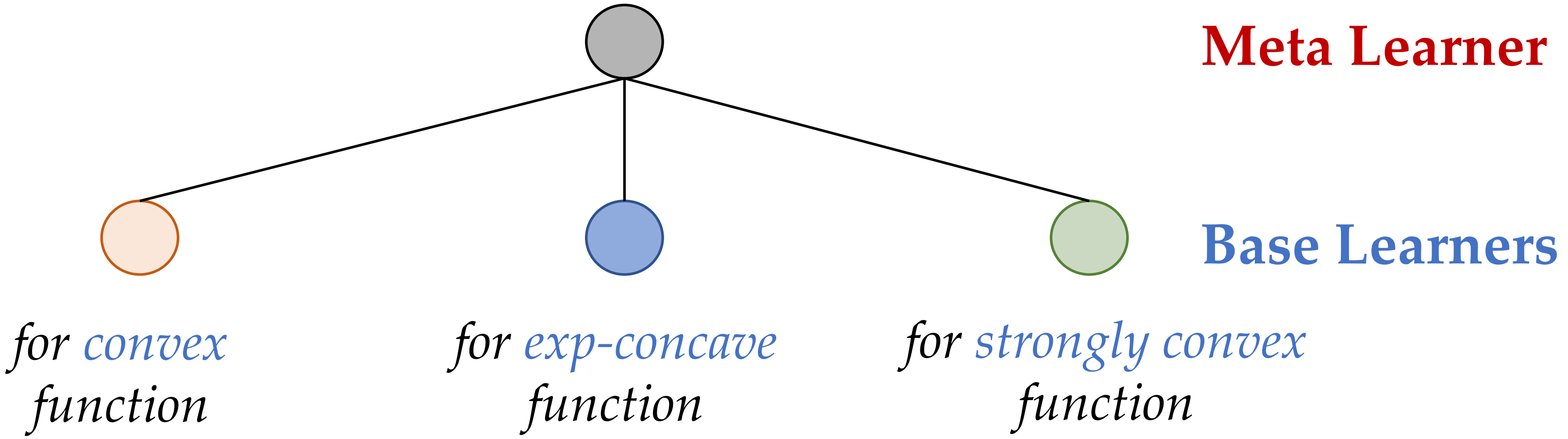
Our Framework: Online Ensemble

- **Robustness:** combine multiple base models to combat with uncertainty

$$\mathbf{x}_t = \sum_{i=1}^{N} p_{t,i} \mathbf{x}_{t,i}$$

- $\boldsymbol{p}_t = [p_{t,1}, \ldots, p_{t,N}]^\top$ is the meta weight;
- $\{\mathbf{x}_{t,i}\}_{t=1}^{T}$ is the base decisions of the i-th base learners, $i \in [N]$.

Universal online learning: dealing with unknown function curvatures



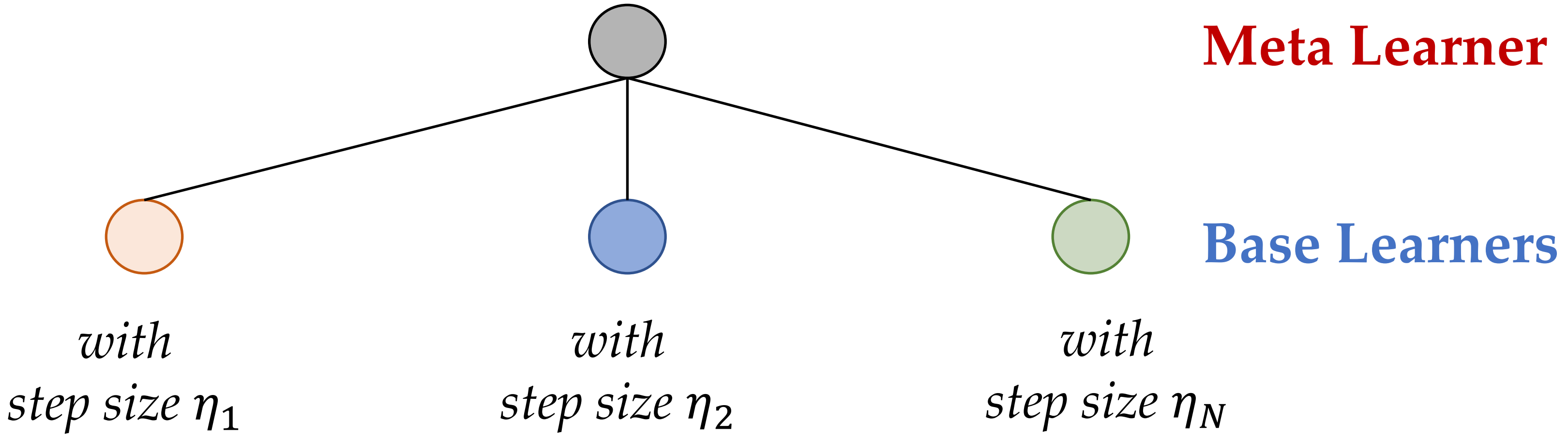
Our Framework: Online Ensemble

- **Robustness:** combine multiple base models to combat with uncertainty

$$\mathbf{x}_t = \sum_{i=1}^{N} p_{t,i}\mathbf{x}_{t,i}$$

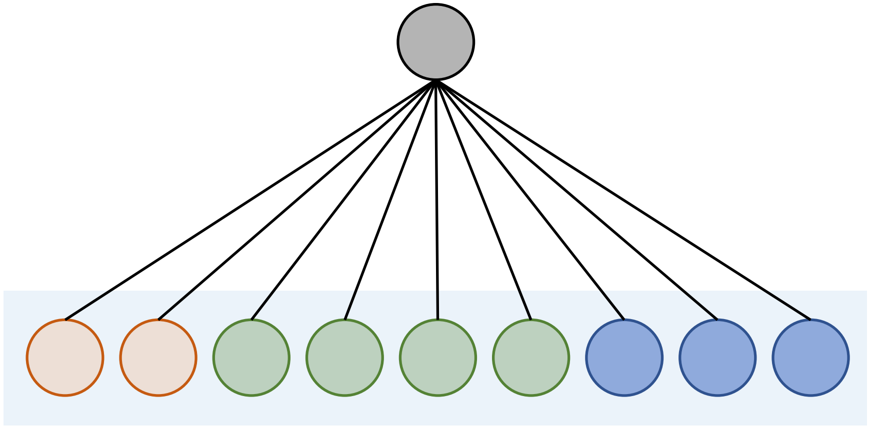
- $\boldsymbol{p}_t = [p_{t,1}, \ldots, p_{t,N}]^\top$ is the meta weight;
- $\{\mathbf{x}_{t,i}\}_{t=1}^{T}$ is the base decisions of the i-th base learners, $i \in [N]$.

Non-stationary online learning: dealing with unknown non-stationarity level



GV in Universal/Non-stationary Regret

- How to handle algorithmic stability in online ensemble structure?



$$\mathbf{x}_t = \sum_{i=1}^{N} p_{t,i}\mathbf{x}_{t,i}$$

How to ensure ***stability*** of an online ensemble?

$$\sum_{t=2}^{T} \|\mathbf{x}_t - \mathbf{x}_{t-1}\|^2$$

⇒ We need to decompose two layers for a refined analysis.

$$\textsc{Reg}_T = \left[\sum_{t=1}^{T} f_t(\mathbf{x}_t) - \sum_{t=1}^{T} f_t(\mathbf{x}_{t,i^\star})\right] \quad \textit{meta regret}$$

$$+ \left[\sum_{t=1}^{T} f_t(\mathbf{x}_{t,i^\star}) - \min_{\mathbf{x}\in\mathcal{X}} \sum_{t=1}^{T} f_t(\mathbf{x})\right] \quad \textit{base regret}$$

GV in Universal/Non-stationary Regret

Learning And Mining from DatA

- How to handle algorithmic stability in online ensemble structure?

$$\sum_{t=2}^{T} \|\mathbf{x}_t - \mathbf{x}_{t-1}\|^2 \implies \|\mathbf{x}_t - \mathbf{x}_{t-1}\|_2^2 \leq \underbrace{2D^2 \|\boldsymbol{p}_t - \boldsymbol{p}_{t-1}\|_1^2}_{\text{meta stability}} + \underbrace{2\sum_{i=1}^{N} p_{t,i} \|\mathbf{x}_{t,i} - \mathbf{x}_{t-1,i}\|_2^2}_{\text{weighted combine of base stability}}$$

❑ base algorithm

(consider Optimistic OGD for simplicity)

$$\mathbf{x}_{t,i} = \Pi_{\mathcal{X}}\left[\widehat{\mathbf{x}}_{t,i} - \eta_i \nabla f_{t-1}(\mathbf{x}_{t-1,i})\right]$$

$$\mathbf{x}_{t+1,i} = \Pi_{\mathcal{X}}\left[\mathbf{x}_{t,i} - \eta_i \nabla f_t(\mathbf{x}_{t,i})\right]$$

$\Rightarrow$

$$\sum_{t=1}^{T} f_t(\mathbf{x}_{t,i}) - \min_{\mathbf{x}\in\mathcal{X}} \sum_{t=1}^{T} f_t(\mathbf{x})$$

$$\lesssim \frac{1}{\eta_i} + \eta_i \bar{V}_T - \frac{1}{\eta_i} \sum_{t=1}^{T} \|\mathbf{x}_{t,i} - \mathbf{x}_{t-1,i}\|_2^2$$

base regret still keeps the negative term

GV in Universal/Non-stationary Regret

Learning And Mining from DatA

- How to handle algorithmic stability in online ensemble structure?

$$\sum_{t=2}^{T} \|\mathbf{x}_t - \mathbf{x}_{t-1}\|^2 \Longrightarrow \|\mathbf{x}_t - \mathbf{x}_{t-1}\|_2^2 \leq 2D^2 \underbrace{\|\boldsymbol{p}_t - \boldsymbol{p}_{t-1}\|_1^2}_{\text{meta stability}} + \underbrace{2\sum_{i=1}^{N} p_{t,i} \|\mathbf{x}_{t,i} - \mathbf{x}_{t-1,i}\|_2^2}_{\text{weighted combine of base stability}}$$

❑ **meta algorithm**

(consider Optimistic Hedge for simplicity)

$$p_{t+1,i} \propto \exp\left(-\varepsilon\left(\sum_{s=1}^{t} \ell_{s,i} + m_{t+1,s}\right)\right) \Longrightarrow$$

with meta loss as $\ell_{s,i} = f_s(\mathbf{x}_{s,i}), \forall s \in [t]$

and meta optimism as $m_{t+1,i} = f_t(\mathbf{x}_{t,i})$

$$\sum_{t=1}^{T} f_t(\mathbf{x}_t) - \sum_{t=1}^{T} f_t(\mathbf{x}_{t,i}) \lesssim \frac{1}{\varepsilon} + \varepsilon \bar{V}_T - \frac{1}{\varepsilon} \sum_{t=1}^{T} \|\boldsymbol{p}_t - \boldsymbol{p}_{t-1}\|_1^2$$

meta regret also keeps the negative term

GV in Universal/Non-stationary Regret

LAMDA Learning And Mining from DatA

- How to handle algorithmic stability in online ensemble structure?

$$\sum_{t=2}^{T} \|\mathbf{x}_t - \mathbf{x}_{t-1}\|^2 \quad \Rightarrow \quad \|\mathbf{x}_t - \mathbf{x}_{t-1}\|_2^2 \leq \underbrace{2D^2 \|\boldsymbol{p}_t - \boldsymbol{p}_{t-1}\|_1^2}_{\text{meta stability}} + \underbrace{2\sum_{i=1}^{N} p_{t,i} \|\mathbf{x}_{t,i} - \mathbf{x}_{t-1,i}\|_2^2}_{\text{weighted combine of base stability}}$$

$$\sum_{t=1}^{T} f_t(\mathbf{x}_t) - \sum_{t=1}^{T} f_t(\mathbf{x}_{t,i}) \lesssim \frac{1}{\varepsilon} + \varepsilon \bar{V}_T - \frac{1}{\varepsilon}\sum_{t=1}^{T} \|\boldsymbol{p}_t - \boldsymbol{p}_{t-1}\|_1^2$$

meta regret also keeps the negative term

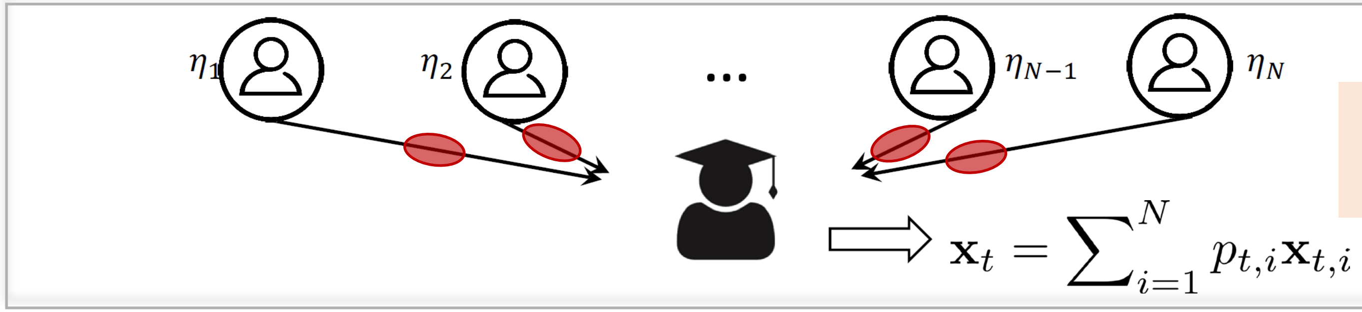
$$\sum_{t=1}^{T} f_t(\mathbf{x}_{t,i}) - \min_{\mathbf{x}\in\mathcal{X}} \sum_{t=1}^{T} f_t(\mathbf{x}) \lesssim \frac{1}{\eta_i} + \eta_i \bar{V}_T - \frac{1}{\eta_i}\sum_{t=1}^{T} \|\mathbf{x}_{t,i} - \mathbf{x}_{t-1,i}\|_2^2$$

base regret still keeps the negative term

negative term only for a particular base learner, insufficient for cancellation

Collaborative Online Ensemble

- **Stablization**: meta algorithm $p_{t+1,i} \propto \exp\left(-\varepsilon(L_{t,i} + m_{t+1,i})\right)$ with
  - (corrected) meta loss $\boldsymbol{\ell}_t \in \mathbb{R}^N$ with $\ell_{t,i} = \langle \nabla f_t(\mathbf{x}_t), \mathbf{x}_{t,i} \rangle + \lambda \|\mathbf{x}_{t,i} - \mathbf{x}_{t-1,i}\|_2^2$;
  - (corrected) meta optimism $\boldsymbol{m}_{t+1} \in \mathbb{R}^N$ with $m_{t+1,i} = \langle M_{t+1}, \mathbf{x}_{t+1,i} \rangle + \lambda \|\mathbf{x}_{t+1,i} - \mathbf{x}_{t,i}\|_2^2$.



***correction**: penalizing unstable base learners*

- Technically, stability term will cancel out by

$$\sum_{t=2}^{T} \|\mathbf{x}_t - \mathbf{x}_{t-1}\|^2 \begin{cases} -\sum_{t=2}^T \|\boldsymbol{p}_t - \boldsymbol{p}_{t-1}\|_1^2 \\ -\sum_{t=2}^T \|\mathbf{x}_{t,i} - \mathbf{x}_{t-1,i}\|_2^2 \\ -\sum_{t=2}^T \sum_{i=1}^N p_{t,i} \|\mathbf{x}_{t,i} - \mathbf{x}_{t-1,i}\|_2^2 \end{cases}$$

Collaborations between meta and base learners: *simultaneously exploiting*

- ✷ *negative terms in regret analysis*
- ✷ *correction terms in algorithm design*

Results: Non-stationary Online Learning

- The first ***non-stationary*** online algorithm with ***gradient-variation regret***.

Theorem 1 (Z-Zhang-Zhang-Zhou; NeurIPS 2020; JMLR 2024). *Under standard assumptions, our online-ensemble algorithm ensures that*

$$\sum_{t=1}^{T} f_t(\mathbf{x}_t) - \sum_{t=1}^{T} f_t(\mathbf{u}_t) \leq \mathcal{O}\left(\sqrt{V_T(1+P_T)} + P_T\right),$$

where $V_T = \sum_{t=2}^{T} \sup_{\mathbf{x}\in\mathcal{X}} \|\nabla f_t(\mathbf{x}) - \nabla f_{t-1}(\mathbf{x})\|^2$ *is gradient variation and* $P_T = \sum_{t=2}^{T} \|\mathbf{u}_t - \mathbf{u}_{t-1}\|_2$ *is the path length measuring the non-stationarity.*

- ❑ Recovering the existing $\mathcal{O}(\sqrt{T(1+P_T)})$ optimal dynamic regret in the worst case such that $V_T = \mathcal{O}(T)$.
- ❑ Recovering the existing $\mathcal{O}(\sqrt{V_T})$ static regret when $\mathbf{u}_t$ is fixed such that $P_T = 0$

Results: Universal Online Learning

- The first ***universal*** online algorithm with ***gradient-variation regret***.

Theorem 2 (Yan-Z-Zhou; NeurIPS 2023). *Under standard assumptions, our online-ensemble algorithm ensures that*

- *it achieves* $\mathcal{O}(\frac{1}{\lambda}\log V_T)$ *regret for* λ*-strongly convex functions;*
- *it achieves* $\mathcal{O}(\frac{d}{\alpha}\log V_T)$ *regret for* α*-exp-concave functions;*
- *it achieves* $\widehat{\mathcal{O}}(\sqrt{V_T})$ *regret for convex functions.*

Here, $V_T = \sum_{t=2}^{T} \sup_{\mathbf{x}\in\mathcal{X}} \|\nabla f_t(\mathbf{x}) - \nabla f_{t-1}(\mathbf{x})\|^2$ *is gradient variation and* $\widehat{\mathcal{O}}(\cdot)$ *omits* $\log V_T$ *factors.*

- ❑ (Almostly) recovering the existing $\mathcal{O}(\sqrt{T})$, $\mathcal{O}(d\log T)$, $\mathcal{O}(\log T)$ optimal universal regret (for convex/ exp-concave/ strongly convex functions) in the worst case such that $V_T = \mathcal{O}(T)$.
- ❑ Still exhibits a *logarithmic* gap for convex functions (we are working on that).

Implications: Games and Adaptive Opt.

Learning And Mining from DatA

❑ Non-stationary Online Learning

- **Game theory**: time-varying games [Zhang-**Z**-Luo-Zhou, ICML'22]

  At each round $t = 1, 2, \ldots, T$:

  - x-player submits $\mathbf{x}_t \in \Delta_d$ and y-player submits $\mathbf{y}_t \in \Delta_d$
  - the x-player suffers loss $\mathbf{x}_t^\top A_t \mathbf{y}_t$ and receives gradient $A_t \mathbf{y}_t$, and the y-player receives reward $\mathbf{x}_t^\top A_t \mathbf{y}_t$ and receives gradient $A_t \mathbf{x}_t$

$$\Rightarrow \quad \text{DynNE-Reg}_T = \left| \sum_{t=1}^T x_t^\top A_t y_t - \sum_{t=1}^T \min_{x \in \Delta_m} \max_{y \in \Delta_n} x^\top A_t y \right| \leq \tilde{\mathcal{O}}\big(\min\{\sqrt{V_T(1+P_T)}, W_T\}\big)$$

- **Adaptive Optimization**: SEA model [Chen-Zhang-Tu-**Z**-Zhang, ICML'23 & JMLR'24]

$$\mathbb{E}\left[\sum^T F_t(\mathbf{x}_t) - \sum_{t=1}^T F_t(\mathbf{u}_t)\right] \leq \mathcal{O}\left(P_T + \sqrt{1+P_T}\left(\sqrt{\sigma_{1:T}^2} + \sqrt{\Sigma_{1:T}^2}\right)\right)$$

Implications: Games and Adaptive Opt.

❑ Universal Online Learning

- **Game theory**: min-max optimization [Yan-**Z**-Zhou, NeurIPS'23]

  Under standard assumptions, for bilinear and strongly convex-concave games, our algorithm enjoys $\mathcal{O}(1)$ regret summation in the honest case, $\widehat{\mathcal{O}}(\sqrt{T})$ and $\mathcal{O}(\log T)$ bounds respectively in the dishonest case.

- **Adaptive Optimization**: SEA model [Yan-**Z**-Zhou, NeurIPS'23]

  Under standard assumptions, our algorithm obtains $\mathcal{O}((\sigma_{\max}^2 + \Sigma_{\max}^2) \log(\sigma_{1:T}^2 + \Sigma_{1:T}^2))$ regret for strongly convex functions, $\mathcal{O}(d \log(\sigma_{1:T}^2 + \Sigma_{1:T}^2))$ regret for exp-concave functions and $\widehat{\mathcal{O}}(\sqrt{\sigma_{1:T}^2 + \Sigma_{1:T}^2})$ regret for convex functions.

Conclusion

Learning And Mining from DatA

❑ Gradient-variation Online Learning

- *Universality*: a single algorithm simultaneously optimal for different function families
- *Non-stationarity*: an algorithm optimizing dynamic regret with changing comparators
- *Collaborative online ensemble*: online ensemble with optimistic update, exploiting negative terms in regret analysis and injecting corrections in algorithmic design
- *Applications*: useful for game theory, adaptive optimization, etc

❑ Open Problems

- Consider exp-concave and strongly convex functions for non-stationary online learning
- How to enhance universality to more challenging with heterogenous curvature info.?
- Connection to continual learning, beyond the convexity assumption.

Thanks!

References

- P. Zhao, Y.-J. Zhang, L. Zhang, and Z.-H. Zhou. Adaptivity and non-stationarity: Problem-dependent dynamic regret for online convex optimization. JMLR 2024.
  (major ref for non-stationary online learning)
- Y.-H. Yan, P. Zhao, and Z.-H. Zhou. Universal online learning with gradient variations: A multi-layer online ensemble approach. NeurIPS 2023.
  (major ref for universal online learning)
- M. Zhang, P. Zhao, H. Luo, and Z.-H. Zhou. No-regret learning in time-varying zero-sum games. ICML 2022.
- Y. Bai, Y.-J. Zhang, P. Zhao, M. Sugiyama, and Z.-H. Zhou. Adapting to online label shift with provable guarantees. NeurIPS 2022.
- S. Chen, W.-W. Tu, P. Zhao, and L. Zhang. Optimistic online mirror descent for bridging stochastic and adversarial online convex optimization. ICML 2023.

Thanks!

References

- C. Chiang, T. Yang, C. Lee, M. Mahdavi, C. Lu, R. Jin, and S. Zhu. Online optimization with gradual variations. COLT 2012.
- V. Syrgkanis, A. Agarwal, H. Luo, and R. E. Schapire. Fast convergence of regularized learning in games. NIPS 2015.
- T. van Erven and W. M. Koolen. MetaGrad: Multiple learning rates in online learning. NIPS 2016.
- L. Zhang, G. Wang, J. Yi, and T. Yang. A simple yet universal strategy for online convex optimization. ICML 2022.
- S. Sachs, H. Hadiji, T. van Erven, and C. Guzmán. Between stochastic and adversarial online convex optimization: Improved regret bounds via smoothness. NeurIPS 2022.

Thanks!