Learning with Interactive Models over Decision-Dependent Distributions

Man-Jie Yuan YUANMJ@LAMDA.NJU.EDU.CN
Wei Gao GAOW@NJU.EDU.CN
National Key Laboratory for Novel Software Technology, Nanjing University, Nanjing, China.

Editors: Emtiyaz Khan and Mehmet Gönen

Abstract

Classical supervised learning generally trains one model from an i.i.d. data according to an unknown yet fixed distribution. In some real applications such as finance, however, multiple models may be trained by different companies and interacted in a dynamic environment, where the data distribution may take shift according to different models' decisions. In this work, we study two models for simplicity, and formalize such scenario as a learning problem of two models over decision-dependent distributions. We develop the Repeated Risk Minimization (RRM) for two models, and present a sufficient condition to the existence of stable points for RRM, that is, an equilibrium notion. We further provide the theoretical analysis for the convergence of RRM to stable points based on data distribution and finite training sample, respectively. We also study more practical algorithms, such as gradient descent and stochastic gradient descent, to solve the RRM problem with convergence guarantees and we finally present some empirical studies to validate our theoretical analysis.

Keywords: Distributional Shift, Performative Prediction, Optimization.

1. Introduction

For classical supervised learning, we always learn one model from training data, drawn i.i.d. from some fixed and static distribution (Mitchell, 1997; Shalev-Shwartz and Ben-David, 2014; Goodfellow et al., 2016). In real applications, however, we may deal with multiple interactive models in a dynamic environment, where data distribution is decision-dependent (Maheshwari et al., 2022), i.e., it may take certain shift according to models' predictions. Here we take financial stocks as an example: many financial companies would train their quantitative trading model to make prediction for stock price and make trading decisions. In such environment, the stock data distribution makes fluctuations w.r.t. different models, while each model will also be influenced with stock data distribution and other models.

The interaction phenomenon between data distribution and model prediction has been known as performativity in real applications such as economics, social network and media platform (Healy, 2015; Ribeiro et al., 2020). Perdomo et al. (2020) proposed the novel performative prediction, which aims to learn one model from interactive data distribution. Mendler-Dünner et al. (2020) presented its convergence analysis based on stochastic gradient descent. Brown et al. (2020) studied the performativity from a decision-making perspective. Miller et al. (2021) and Izzo et al. (2021) considered the direct optimization of performative risk under strong conditions.

Drusvyatskiy and Xiao (2020) formalized performative prediction as an optimization problem with decision-dependent distributions, and gave online gradient descent algorithm. Wood et al. (2022) provided online projected gradient descent for stochastic optimization with decision-dependent distributions. Ray et al. (2022) studied the problem of decision-dependent risk minimization with distribution evolves according to a geometrically decaying process. Recent research in strategic classification (Hardt et al., 2016; Bechavod et al., 2020; Ghalme et al., 2021) also highlighted the prevalence of this phenomenon. Previous studies mostly focus on one single model over decision-dependent distribution, whereas we always confront multiple interactive models in some dynamic environments. This work tries to take one step on this direction.

A relevant work is the game theory with multiple interactive models (Fudenberg and Levine, 1998; Nisan et al., 2007), which tries to find some Nash equilibria, a stable status for players with respect to their payoffs. Recent years have witnessed increasing interests in combination between machine learning and game theory (Bravo et al., 2018; Chasnov et al., 2020). For example, the generative adversarial network has been formalized as a two-player zero-sum game (Heusel et al., 2017; Balduzzi et al., 2018), and Vlatakis-Gkaragkounis et al. (2019) studied the non-convex non-concave zero-sum games based on gradient descent ascent method. In most times, game theory tries to solve the minmax problem over some payoff function or fixed distribution, rather than the shifting decision-dependent data distribution.

Reinforcement learning is another related work (Srinivasan et al., 2018; Zhou and Xu, 2019), which involves interactions among multiple models and the shifted environment. Generally, reinforcement learning studies how to select actions and maximizes rewards, while we focus on risk minimization based on data sample or distribution and its convergence rate to stable points. In addition, our method can be viewed as a special formulation of off-policy learning, which may shed lights on when off-policy methods converge in a two-agent system.

This work tries to learn multiple interactive models under a decision-dependent dynamic environment. For simplicity, we formalize such scenario as a learning problem of two models over decision-dependent distributions, and the main contributions of this work can be summarized as follows:

- We develop the Repeated Risk Minimization (RRM), which is a simple iterative optimizing method for two models to learn with decision-dependent distribution. We present a sufficient condition to the existence of stable points of RRM, that is, an equilibrium among two interactive models and decision-dependent data distribution.
- We present the linear convergence rate to stable points for RRM under the full knowledge of data distribution, and an intuitive explanation is that data distribution should keep relatively stable w.r.t. models, while models should effectively adapt to decision-dependent distribution. We also present the linear convergence rate to a small neighborhood of stable points for RRM based on finite training sample.
- We propose more practical algorithms such as gradient descent and stochastic gradient descent for RRM. For gradient descent, we derive similar linear convergence rate but with different exponent based on data distribution and finite sample. For stochastic gradient descent, we present the $O(T^{-1/2})$ convergence rate, which is comparable to that of traditional stochastic gradient descent.

- We finally present empirical studies to verify our theoretical analysis both on synthetic and semi-synthetic datasets.

The rest of this paper is organized as follows: Section 2 formalizes the learning problem of two models over decision-dependent distributions. Section 3 presents theoretical analysis for RRM based on data distribution and Section 4 presents analysis base on finite sample. Section 5 provided convergence analysis based on gradient descent and stochastic gradient descent. Section 6 conducts empirical studies, and Section 7 concludes with future work.

2. Learning with Two models over Decision-Dependent Distributions

This section presents a framework on the learning problem of two models over decision-dependent distributions. For simplicity, denote by $\hat{1}$ and $\hat{2}$ two learning models, which can be further parameterized by $\theta^{\hat{1}}, \theta^{\hat{2}} \in \Theta$, respectively. Here, $\Theta \subseteq \mathbb{R}^m$ denotes some closed and convex parameter space, and we could make similar analysis when $\theta^{\hat{1}}$ and $\theta^{\hat{2}}$ belong to two different parameter spaces.

Traditional statistical learning focuses on a fixed joint distribution $\mathcal{D}$ over the product space $\mathcal{Z} = \mathcal{X} \times \mathcal{Y} \subseteq \mathbb{R}^d$ with input space $\mathcal{X}$ and output space $\mathcal{Y}$. However, the data distribution $\mathcal{D}$ may be affected by multiple learning models in many real applications, such as stock data, recommend systems, etc. We formalize such distribution as $\mathcal{D}(\theta^{\hat{1}}, \theta^{\hat{2}})$ in this work, i.e., data distribution may be changed w.r.t. different models.

For loss functions $\ell_{\hat{1}}$ and $\ell_{\hat{2}}$, we define the *decision-dependent risk* of learning models $\hat{1}$ and $\hat{2}$, respectively, as follows:

$$\begin{aligned} DDR_{\hat{1}}(\theta^{\hat{1}}, \theta^{\hat{2}}) &= \mathbb{E}_{\boldsymbol{z}\sim\mathcal{D}(\theta^{\hat{1}},\theta^{\hat{2}})}\left[\ell_{\hat{1}}(\boldsymbol{z};\theta^{\hat{1}})\right] , \\ DDR_{\hat{2}}(\theta^{\hat{1}}, \theta^{\hat{2}}) &= \mathbb{E}_{\boldsymbol{z}\sim\mathcal{D}(\theta^{\hat{1}},\theta^{\hat{2}})}\left[\ell_{\hat{2}}(\boldsymbol{z};\theta^{\hat{2}})\right] . \end{aligned}$$

It is not easy to directly optimize such risks because of its non-convexity, even for convex loss functions. Following Perdomo et al. (2020), we optimize the decision-dependent risks iteratively under the current fixed distribution and ignore the temporary changes of distribution. Formally, we define such process as *repeated risk minimization*:

Definition 1 *Given initial models $\theta_0^{\hat{1}}$ and $\theta_0^{\hat{2}}$, the Repeated Risk Minimization (RRM) performs the following updates, for $i \in [2]$ and $t \geq 0$:*

$$\theta_{t+1}^{\hat{i}} = \arg\min_{\theta^{\hat{i}}\in\Theta} \left\{\mathbb{E}_{\boldsymbol{z}\sim\mathcal{D}(\theta_t^{\hat{1}},\theta_t^{\hat{2}})}\left[\ell_{\hat{i}}(\boldsymbol{z};\theta^{\hat{i}})\right]\right\} .$$

In each iteration, two models update their parameters according to the current distribution, and then distribution $\mathcal{D}(\theta_t^{\hat{1}}, \theta_t^{\hat{2}})$ shifts according to the updated parameters. When two models reach a stable point without updates over parameters, we call such stable point as *decision-dependent stable point*, defined formally as:

Definition 2 *For RRM, we say $(\theta_{DS}^{\hat{1}}, \theta_{DS}^{\hat{2}})$ a decision-dependent stable point if it holds that*

$$\begin{aligned} \theta_{DS}^{\hat{1}} &= \arg\min_{\theta^{\hat{1}}\in\Theta}\mathbb{E}_{\boldsymbol{z}\sim\mathcal{D}(\theta_{DS}^{\hat{1}},\theta_{DS}^{\hat{2}})}\left[\ell_{\hat{1}}(\boldsymbol{z};\theta^{\hat{1}})\right] , \\ \theta_{DS}^{\hat{2}} &= \arg\min_{\theta^{\hat{2}}\in\Theta}\mathbb{E}_{\boldsymbol{z}\sim\mathcal{D}(\theta_{DS}^{\hat{1}},\theta_{DS}^{\hat{2}})}\left[\ell_{\hat{2}}(\boldsymbol{z};\theta^{\hat{2}})\right] . \end{aligned}$$

Intuitively, it is not easy to find some stable points for RRM when drastic changes occur to distribution $\mathcal{D}(\theta_t^{\hat{1}}, \theta_t^{\hat{2}})$, and we make the following assumption over distribution $\mathcal{D}(\theta^{\hat{1}}, \theta^{\hat{2}})$, which has been used by Perdomo et al. (2020) to analyze performative prediction.

Definition 3 *A distribution $\mathcal{D}(\cdot, \cdot)$ is said to be ϵ-Lipschitz continuous if it holds that, for every $\theta_1^{\hat{1}}, \theta_1^{\hat{2}}, \theta_2^{\hat{1}}, \theta_2^{\hat{2}} \in \Theta$,*

$$W_1\big(\mathcal{D}(\theta_1^{\hat{1}}, \theta_1^{\hat{2}}), \mathcal{D}(\theta_2^{\hat{1}}, \theta_2^{\hat{2}})\big) \leq \epsilon \big\|(\theta_1^{\hat{1}}, \theta_1^{\hat{2}}) - (\theta_2^{\hat{1}}, \theta_2^{\hat{2}})\big\|_2 \ ,$$

where W_1 denotes the Wasserstein-1 distance (Villani, 2009).

We now introduce some basic properties for loss functions in optimization problems (Nesterov, 2018) as follows:

Definition 4 *A loss function $\ell(\boldsymbol{z}; \theta)$ is said to be β-jointly smooth (with $\beta > 0$), if it holds that, for every $\theta_1, \theta_2 \in \Theta$ and $\boldsymbol{z}_1, \boldsymbol{z}_2 \in \mathcal{Z}$,*

$$\begin{aligned} \|\nabla_\theta \ell(\boldsymbol{z}_1; \theta_1) - \nabla_\theta \ell(\boldsymbol{z}_1; \theta_2)\|_2 &\leq \beta \|\theta_1 - \theta_2\|_2 \ , \\ \|\nabla_\theta \ell(\boldsymbol{z}_1; \theta_1) - \nabla_\theta \ell(\boldsymbol{z}_2; \theta_1)\|_2 &\leq \beta \|\boldsymbol{z}_1 - \boldsymbol{z}_2\|_2 \ . \end{aligned}$$

Definition 5 *A loss function $\ell(\boldsymbol{z}; \theta)$ is said to be γ-strongly convex (with $\gamma > 0$), if it holds that, for every $\theta_1, \theta_2 \in \Theta$ and $\boldsymbol{z} \in \mathcal{Z}$,*

$$\ell(\boldsymbol{z}; \theta_1) \geq \ell(\boldsymbol{z}; \theta_2) + \nabla_\theta \ell(\boldsymbol{z}; \theta_2)^{\mathrm{T}}(\theta_1 - \theta_2) + \frac{\gamma}{2}\|\theta_1 - \theta_2\|_2^2 \ .$$

3. Analysis on Repeated Risk Minimization over Data Distribution

We first present the existence analysis of decision-dependent stable points for Repeated Risk Minimization (RRM, given by Definition 1) as follows:

Theorem 6 *There is a decision-dependent stable point $(\theta_{DS}^{\hat{1}}, \theta_{DS}^{\hat{2}})$ for convex and compact space Θ and ϵ-Lipschitz continuous distribution $\mathcal{D}$, if loss functions $\ell_{\hat{1}}(\boldsymbol{z}; \theta^{\hat{1}})$ and $\ell_{\hat{2}}(\boldsymbol{z}; \theta^{\hat{2}})$ are convex and jointly continuous.*

Theorem 6 presents a sufficient condition for the existence of stable point for the learning problem of two models over decision-dependent distributions, and this theoretically guarantees the convergence of learning algorithms to some stable points, rather than training endlessly. The detailed proof is given in Appendix B.2, and the basic idea follows Kakutani's fixed point theorem (Kakutani, 1941).

We now present the convergence analysis on RRM (Definition 1) under the full knowledge of data distribution $\mathcal{D}$. Denote by $\theta_t = (\theta_t^{\hat{1}}, \theta_t^{\hat{2}})$ and $\theta_{\mathrm{DS}} = (\theta_{\mathrm{DS}}^{\hat{1}}, \theta_{\mathrm{DS}}^{\hat{2}})$ for simplicity, we have

Theorem 7 *For the RRM method after T iterations, we have*

$$\|\theta_T - \theta_{DS}\|_2 \leq \|\theta_0 - \theta_{DS}\|_2 \left(\epsilon(\beta_{\hat{1}}^2 + \beta_{\hat{2}}^2)^{\frac{1}{2}} / \min\{\gamma_{\hat{1}}, \gamma_{\hat{2}}\}\right)^T$$

for ϵ-Lipschitz continuous distribution $\mathcal{D}$, and for $\beta_{\hat{i}}$-jointly smooth, $\gamma_{\hat{i}}$-strongly convex function $\ell_{\hat{i}}(\boldsymbol{z}; \theta^{\hat{i}})$ with $i \in [2]$.

This theorem shows a linear convergence rate for the RRM method under the condition $\epsilon < \min\{\gamma_{\hat{1}}, \gamma_{\hat{2}}\}/(\beta_{\hat{1}}^2 + \beta_{\hat{2}}^2)^{1/2}$, and this condition could correlate data distribution with loss functions. An intuitive explanation for convergence is that, data distribution should keep relatively stable according to model parameters, while learning models should adapt to the changes of data distribution. Here, we present a proof sketch for Theorem 7, and the details can be found in Appendix B.3.

Proof sketch of Theorem 7. For simplicity, we denote by $G(\theta) = (G_{\hat{1}}(\theta), G_{\hat{2}}(\theta))$, where

$$G_{\hat{i}}(\theta) = \arg\min_{\theta^{\hat{i}} \in \Theta} \mathbb{E}_{\boldsymbol{z}\sim\mathcal{D}(\theta)}\left[\ell_{\hat{i}}(\boldsymbol{z}; \theta^{\hat{i}})\right] \quad \text{for} \quad i \in [2].$$

Here, $G_{\hat{i}}(\theta)$ denotes the output in one-step by RRM method with input $\theta = (\theta^{\hat{1}}, \theta^{\hat{2}})$. For $i \in [2]$ and $t \in [T]$, denote by

$$f_t^{\hat{i}}(\theta) = \mathbb{E}_{\boldsymbol{z}\sim\mathcal{D}(\theta_t)}[\ell_{\hat{i}}(\boldsymbol{z}, \theta)] \text{ and } f_{\mathrm{DS}}^{\hat{i}}(\theta) = \mathbb{E}_{\boldsymbol{z}\sim\mathcal{D}(\theta_{\mathrm{DS}})}[\ell_{\hat{i}}(\boldsymbol{z}, \theta)] \ .$$

From the strong convexity and the first-order optimality condition of $G_{\hat{i}}(\theta_t)$ and $G_{\hat{i}}(\theta_{\mathrm{DS}})$, we have, for $i \in [2]$,

$$\gamma_{\hat{i}} \|G_{\hat{i}}(\theta_t) - G_{\hat{i}}(\theta_{\mathrm{DS}})\|_2^2 \leq (G_{\hat{i}}(\theta_{\mathrm{DS}}) - G_{\hat{i}}(\theta_t))^{\mathrm{T}} \nabla f_t^{\hat{i}}(G_{\hat{i}}(\theta_{\mathrm{DS}})) \ .$$

This follows that, from the above inequality with $i \in [2]$,

$$\|G(\theta_t) - G(\theta_{\mathrm{DS}})\|_2^2 \min\{\gamma_{\hat{1}}, \gamma_{\hat{2}}\} \leq (G(\theta_{\mathrm{DS}}) - G(\theta_t))^{\mathrm{T}} \nabla f_t(G(\theta_{\mathrm{DS}})) \ ,$$

with $\nabla f_t(G(\theta_{\mathrm{DS}})) = (\nabla f_t^{\hat{1}}(G_{\hat{1}}(\theta_{\mathrm{DS}})); \nabla f_t^{\hat{2}}(G_{\hat{2}}(\theta_{\mathrm{DS}})))$. By Cauchy-Schwarz inequality, it is easy to check that

$$(G(\theta_t) - G(\theta_{\mathrm{DS}}))^{\mathrm{T}} \nabla_\theta \ell(\boldsymbol{z}; G(\theta_{\mathrm{DS}}))$$

is $(\beta_{\hat{1}}^2 + \beta_{\hat{2}}^2)^{1/2} \|G(\theta_t) - G(\theta_{\mathrm{DS}})\|_2$-Lipschitz continuous w.r.t. variable $\boldsymbol{z}$. From the ϵ-Lipschitz continuity of distribution $\mathcal{D}$ and its dual formulation of Wasserstein-1 distance, we have

$$\begin{aligned}(G(\theta_{\mathrm{DS}}) - G(\theta_t))^{\mathrm{T}} (\nabla f_t(\boldsymbol{z}; G(\theta_{\mathrm{DS}})) - \nabla f_{\mathrm{DS}}(\boldsymbol{z}; G(\theta_{\mathrm{DS}}))) \\ \leq \ \epsilon(\beta_{\hat{1}}^2 + \beta_{\hat{2}}^2)^{1/2} \|G(\theta_t) - G(\theta_{\mathrm{DS}})\|_2 \|\theta_t - \theta_{\mathrm{DS}}\|_2 \ .\end{aligned}$$

Under the first-order optimality condition, we further have

$$(G(\theta_{\mathrm{DS}}) - G(\theta_t))^{\mathrm{T}} \nabla f_t(\boldsymbol{z}; G(\theta_{\mathrm{DS}})) \leq \epsilon(\beta_{\hat{1}}^2 + \beta_{\hat{2}}^2)^{1/2} \|G(\theta_t) - G(\theta_{\mathrm{DS}})\|_2 \|\theta_t - \theta_{\mathrm{DS}}\|_2 \ .$$

This follows that, by combining with the above and Eqn. (3),

$$\|\theta_{t+1} - \theta_{\mathrm{DS}}\|_2 = \|G(\theta_t) - G(\theta_{\mathrm{DS}})\|_2 \leq \epsilon(\beta_{\hat{1}}^2 + \beta_{\hat{2}}^2)^{1/2} \|\theta_t - \theta_{\mathrm{DS}}\|_2 / \min\{\gamma_{\hat{1}}, \gamma_{\hat{2}}\} \ .$$

This completes the proof by multiplying both sides with $t = 0, 1, \cdots, T-1$. □

We now introduce an illustrative example for the learning problem of two models over decision-dependent distributions, and present the convergence curves of two models to support Theorems 6 and 7 empirically.

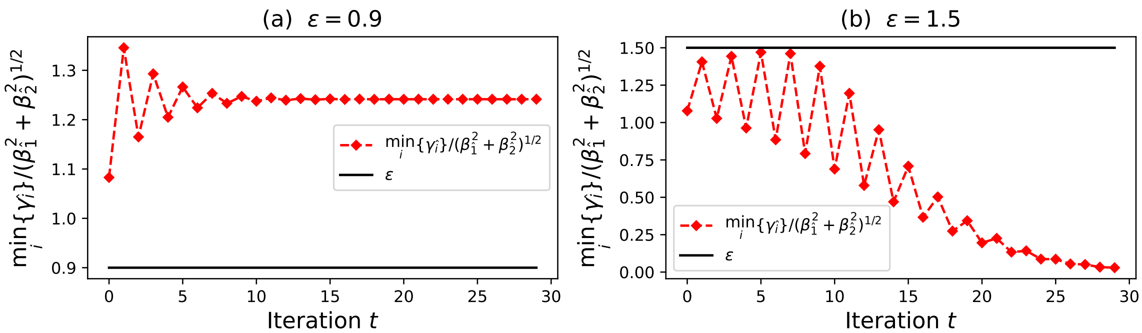

Figure 1: The convergence curves of $\min\{\gamma_{\hat{1}}, \gamma_{\hat{2}}\}/(\beta_{\hat{1}}^2 + \beta_{\hat{2}}^2)^{1/2}$ for RRM with $\epsilon = 0.9$ (left) and $\epsilon = 1.5$ (right) in Example 1.

Example 1 *We consider the one-dimensional linear regressions for two models as follows:*

$$h_{\hat{1}}(x; \theta^{\hat{1}}) = x\theta^{\hat{1}} - b \quad \text{ and } \quad h_{\hat{2}}(x; \theta^{\hat{2}}) = x\theta^{\hat{2}} + b \ ,$$

where $\theta^{\hat{1}}, \theta^{\hat{2}} \in \mathbb{R}$ *and* $b \in \mathbb{R}^+$*, and we also consider the ridge regression loss functions*

$$\ell_{\hat{i}}(\boldsymbol{z}; \theta^{\hat{i}}) = \frac{1}{2}\big(y - h_{\hat{i}}(x; \theta^{\hat{i}})\big)^2 + \frac{\gamma_{\hat{i}}}{2}(\theta^{\hat{i}})^2 \qquad (i \in [2]) \ ,$$

where $\gamma_{\hat{i}} > 0$*. Let input space* $\mathcal{X} = \mathbb{R}$ *with the standard Gaussian distribution* $\mathcal{N}(0, 1)$ *over* $\mathcal{X}$*, and suppose that the output value, with inputs* $x \neq 0$*,* $\theta^{\hat{1}}$ *and* $\theta^{\hat{2}}$*, is given by*

$$y = x\theta_* + \frac{\epsilon}{2|x|}\big(|x\theta_* - h_{\hat{1}}(x; \theta^{\hat{1}})| - |x\theta_* - h_{\hat{2}}(x; \theta^{\hat{2}})|\big) \ ,$$

for constants $\theta_* \in \mathbb{R}$ *and* $\epsilon > 0$*.*

This example can be viewed as a simplification of real problems such as recommend systems and companies bidding in a competitive environment. For example, suppose that there are two firms training different models to predict the customer's "willing to pay", and decide whether to send coupons. Hence, each customer wants to modify their features based on the predictions to get more coupons. In such application, two models interact in a decision-dependent environment. The term b is added to make two models have distinct "model space", and this gap will force them to compete with each other.

For $i \in [2]$, it is easy to check that $\ell_{\hat{i}}$ is $\gamma_{\hat{i}}$-strongly convex and $|\theta^{\hat{i}}|$-smooth w.r.t. the variables $\theta^{\hat{i}}$ and $\boldsymbol{z}$, respectively. We can also observe that the data distribution over input/output space satisfies the ϵ-Lipschitz continuity, as shown in Appendix B.1. Our theoretical results show that it suffices to reach a stable point after a few times retrain when ϵ is small. Note that we could also make similar analysis for d-dimensional regression with $d \geq 2$ in Example 1. For simplicity, the parameters are set as $\theta_* = 1$, $b = 100$ and $\gamma_{\hat{1}} = \gamma_{\hat{2}} = 30$. We simulate this example with $\epsilon = 0.9$ and 1.5 respectively.

Figure 1 shows the curves of $\min\{\gamma_{\hat{1}}, \gamma_{\hat{2}}\}/(\beta_{\hat{1}}^2 + \beta_{\hat{2}}^2)^{1/2}$ in Example 1 with respect to distribution Lipschitz parameter $\epsilon = 0.9$ and $\epsilon = 1.5$. According to the convergence

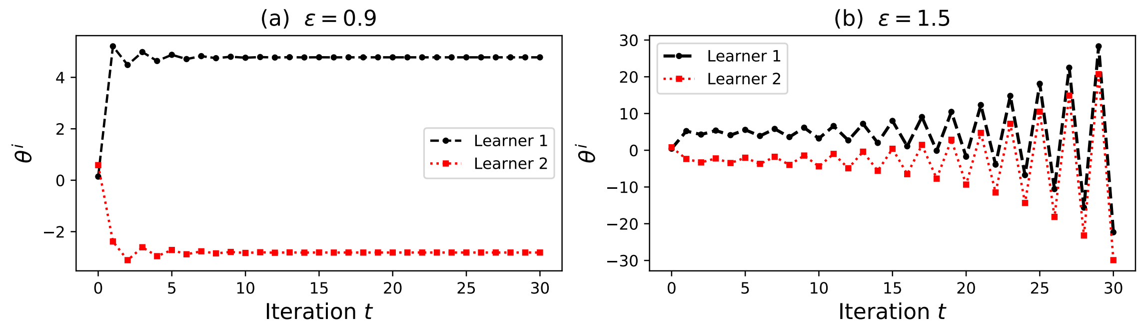

Figure 2: The convergence curves of models for RRM with $\epsilon = 0.9$ (left) and $\epsilon = 1.5$ (right) in Example 1.

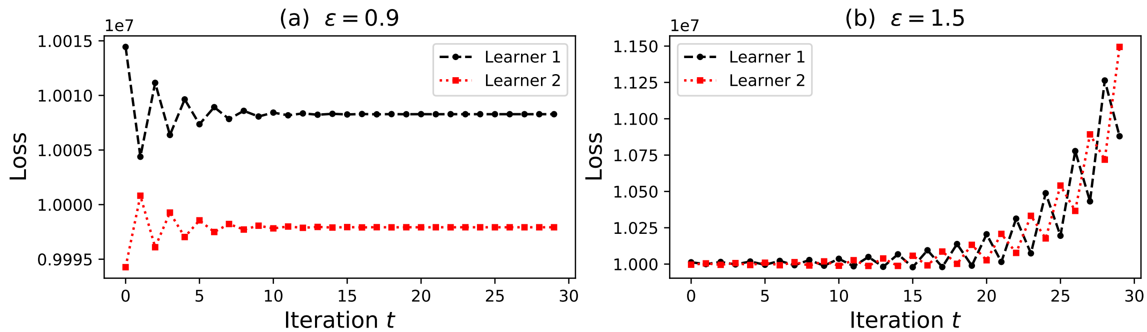

Figure 3: The convergence curves of losses for RRM with $\epsilon = 0.9$ (left) and $\epsilon = 1.5$ (right) in Example 1.

condition $\epsilon < \min\{\gamma_{\hat{1}}, \gamma_{\hat{2}}\}/(\beta_{\hat{1}}^2 + \beta_{\hat{2}}^2)^{1/2}$ in Theorem 7, we could conclude that the RRM method would converge to some stable point for $\epsilon = 0.9$, but diverge for $\epsilon = 1.5$.

Figure 2(a) shows the joint convergence curves of two models to some stable points for RRM after a few iterations when $\epsilon = 0.9$, whereas Figure 2(b) shows the divergence curves of two models for RRM when $\epsilon = 1.5$. This presents empirical supports to Theorem 7.

Figure 3 presents the curves of losses for two models. It is observable that two models compete against each other during iterations, i.e., one wins while the other loses. We can also see the convergence and divergence of the RRM method for $\epsilon = 0.9$ and $\epsilon = 1.5$, respectively. This is also in agreement with the theoretical argument in Theorem 7.

Moreover, Figure 3 also reveals that when two models diverge, their losses may keep increasing. This addresses the importance of finding a decision-dependent stable point, since two models could suffer great increase of loss when they fail to converge to the decision-dependent stable point.

4. Analysis on Repeated Risk Minimization over Finite Sample

We are required to know the data distribution $\mathcal{D}$ in the optimization of RRM method in previous theoretical analysis (Section 3). For most learning problems, however, we can only observe a finite training sample, rather than the whole data distribution. Hence, it is necessary to further exploit the learning algorithms based on finite training sample, and we formalize it as *repeated empirical risk minimization*:

Definition 8 *Given initial models $\theta_0^{\hat{1}}$ and $\theta_0^{\hat{2}}$, the Repeated Empirical Risk Minimization (RERM) performs the following updates, for every $t \geq 0$:*

$$\theta_{t+1}^{\hat{i}} = \arg\min_{\theta^{\hat{i}} \in \Theta} \left\{ \sum\nolimits_{z \in S_{n_t}^i} \ell_{\hat{i}}(\boldsymbol{z}; \theta^{\hat{i}})/n_t \right\} \qquad (i \in [2]) \ ,$$

where $S_{n_t}^1$ and $S_{n_t}^2$ are two training samples of size n_t, with each element drawn i.i.d. from distribution $\mathcal{D}(\theta_t^{\hat{1}}, \theta_t^{\hat{2}})$.

Due to the randomness of training sampling, it is difficult to guarantee an exact contraction to the decision-dependent stable point, whereas we could present the convergence to a neighborhood around the decision-dependent stable point for sufficient training samples with high probability.

Denote by $\theta_t = (\theta_t^{\hat{1}}, \theta_t^{\hat{2}})$ and $\theta_{\mathrm{DS}} = (\theta_{\mathrm{DS}}^{\hat{1}}, \theta_{\mathrm{DS}}^{\hat{2}})$, and we have

Theorem 9 *Let distribution $\mathcal{D}$ be ϵ-Lipschitz continuous such that $E_{\boldsymbol{z}\sim\mathcal{D}(\theta)}[\exp(\mu\|\boldsymbol{z}\|^{\alpha})] < +\infty$ for some positive α and μ and for every θ. Let loss $\ell_{\hat{i}}(\boldsymbol{z}; \theta^{\hat{i}})$ be $\beta_{\hat{i}}$-jointly smooth and $\gamma_{\hat{i}}$-strongly convex for $i \in [2]$. For real $r, \delta \in (0,1)$ and integer $t > 0$, the following holds for the RERM method in the t-th iteration with probability at least $1 - 6\delta/\pi^2 t^2$ over training samples $S_{n_t}^{\hat{1}}$ and $S_{n_t}^{\hat{2}}$*

$$\|\theta_t - \theta_{DS}\|_2 \leq \frac{2\epsilon(\beta_{\hat{1}}^2 + \beta_{\hat{2}}^2)^{1/2}}{\min\{\gamma_{\hat{1}}, \gamma_{\hat{2}}\}} \max\left\{r, \|\theta_{t-1} - \theta_{DS}\|_2\right\} \ ,$$

if the sample size $n_t = O(\log(t/\delta)/(\epsilon r)^d)$, where d is the dimensionality of training sample.

For $\epsilon < \min\{\gamma_{\hat{1}}, \gamma_{\hat{2}}\}/2(\beta_{\hat{1}}^2 + \beta_{\hat{2}}^2)^{1/2}$ and sufficient training sample, the models by the RERM method would converge to a small neighborhood of decision-dependent stable point at linear rate. Specifically speaking, we have

$$\|\theta_t - \theta_{\mathrm{DS}}\|_2 \leq \frac{2\epsilon(\beta_{\hat{1}}^2 + \beta_{\hat{2}}^2)^{1/2}}{\min\{\gamma_{\hat{1}}, \gamma_{\hat{2}}\}} \|\theta_{t-1} - \theta_{\mathrm{DS}}\|_2 \ ,$$

for $\|\theta_{t-1} - \theta_{\mathrm{DS}}\|_2 > r$, that is, our algorithm converges to a neighborhood of a decision-dependent stable point at linear rate when they are out of the neighborhood; we also have

$$\|\theta_t - \theta_{\mathrm{DS}}\|_2 \leq r \ ,$$

for $\|\theta_{t-1} - \theta_{\mathrm{DS}}\|_2 \leq r$, i.e., our algorithm would keep in the neighborhood after entrance.

The proof follows Theorem 7 but with Fournier and Guillin (2015)'s concentration inequality for Wasserstein distance, and the detailed proof is provided in Appendix B.4.

5. Repeated Risk Minimization by Gradient Descent and Stochastic Gradient Descent

It is not easy to directly solve the minimizers of RRM and RERM in Section 3 and 4. Hence, we resort to more practical algorithms such as gradient descent and stochastic gradient descent, which are popular optimization methods in machine learning.

We first consider the gradient descent for repeated risk minimization, and formalize it as *repeated gradient descent*:

Definition 10 *Given initial models $\theta_0^{\hat{1}}$ and $\theta_0^{\hat{2}}$, the Repeated Gradient Descent (RGD) performs the following updates, for every $t \geq 0$:*

$$\theta_{t+1}^{\hat{i}} = \Pi_\Theta\Big(\theta_t^{\hat{i}} - \eta \mathbb{E}_{\boldsymbol{z}\sim\mathcal{D}(\theta_t^{\hat{1}},\theta_t^{\hat{2}})}\big[\nabla_\theta \ell_{\hat{i}}(\boldsymbol{z};\theta_t^{\hat{i}})\big]\Big) \quad (i \in [2]) \ ,$$

where $\eta > 0$ and Π_Θ denote the step size and Euclidean projection operator, respectively.

The RGD method requires to know data distribution $\mathcal{D}$, and makes use of the expected gradient of loss functions to update models with step size η. We could also present the convergence to some decision-dependent stable point.

Denote by $\theta_t = (\theta_t^{\hat{1}}, \theta_t^{\hat{2}})$ and $\theta_{\rm DS} = (\theta_{\rm DS}^{\hat{1}}, \theta_{\rm DS}^{\hat{2}})$. We have

Theorem 11 *Let loss $\ell_{\hat{i}}(\boldsymbol{z};\theta^{\hat{i}})$ be $\beta_{\hat{i}}$-jointly smooth and $\gamma_{\hat{i}}$-strongly convex for $i \in [2]$, and distribution $\mathcal{D}$ is ϵ-Lipschitz continuous. For RGD with step size $\eta \leq 2\sqrt{2}/(\tilde{\beta} + \sqrt{2}\gamma_{\min})$, we have, after T iterations,*

$$\|\theta_T - \theta_{DS}\|_2 \leq \|\theta_0 - \theta_{DS}\|_2 \exp\left[-T\eta\Big(\frac{\tilde{\beta}\gamma_{\min}}{\tilde{\beta}+\sqrt{2}\gamma_{\min}} - \epsilon\tilde{\beta}(1+\eta\tilde{\beta}+0.5\eta\epsilon\tilde{\beta})\Big)\right] \quad ,$$

where $\gamma_{\min} = \min\{\gamma_{\hat{1}}, \gamma_{\hat{2}}\}$ and $\tilde{\beta} = (\beta_{\hat{1}}^2 + \beta_{\hat{2}}^2)^{1/2}$.

This theorem shows the convergence of the RGD method to some decision-dependent stable point $\theta_{\rm DS}$ at a linear rate when $\epsilon < 2\gamma_{\min}/[(\tilde{\beta}+\sqrt{2}\gamma_{\min})(3\eta\tilde{\beta}+2)]$. This is a stronger condition for convergence than that of Theorem 7, since gradient descent may fail to reach the minimizer of RRM. The detailed proof is given in Appendix B.5.

We could also consider the finite training sample for the gradient descent of repeated risk minimization, and formalize it as *repeated empirical gradient descent*:

Definition 12 *Given initial models $\theta_0^{\hat{1}}$ and $\theta_0^{\hat{2}}$, the Repeated Empirical Gradient Descent (REGD) performs the following updates with step size $\eta > 0$, for every $t \geq 0$:*

$$\theta_{t+1}^{\hat{i}} = \Pi_\Theta\Big(\theta_t^{\hat{i}} - \eta \sum\nolimits_{z\in S_{n_t}^i} \nabla_\theta \ell_{\hat{i}}(\boldsymbol{z};\theta_t^{\hat{i}})/n_t\Big) \quad (i \in [2]) \ ,$$

where Π_Θ denotes a Euclidean projection operator, and $S_{n_t}^1$ and $S_{n_t}^2$ are two training samples of size n_t with each element drawn i.i.d. from distribution $\mathcal{D}(\theta_t^{\hat{1}}, \theta_t^{\hat{2}})$.

We now analyze the convergence analysis for REGD method. Denote by $\theta_t = (\theta_t^{\hat{1}}, \theta_t^{\hat{2}})$ and $\theta_{\rm DS} = (\theta_{\rm DS}^{\hat{1}}, \theta_{\rm DS}^{\hat{2}})$, and we have

Theorem 13 *Let distribution $\mathcal{D}$ be ϵ-Lipschitz continuous s.t. $E_{\boldsymbol{z}\sim\mathcal{D}(\theta)}[\exp(\mu\|\boldsymbol{z}\|^{\alpha})] < +\infty$ for some positive α and μ and for every θ. Let loss $\ell_{\hat{i}}(\boldsymbol{z};\theta^{\hat{i}})$ be $\beta_{\hat{i}}$-jointly smooth and $\gamma_{\hat{i}}$-strongly convex for $i \in [2]$. For real $r, \delta \in (0,1)$ and integer $t > 0$, the following holds for REGD at the t-th iteration with probability at least $1 - 6\delta/\pi^2 t^2$ over samples $S^{\hat{1}}_{n_t}$ and $S^{\hat{2}}_{n_t}$*

$$\|\theta_t - \theta_{DS}\|_2 \ \leq \ \max\left\{r, \|\theta_{t-1} - \theta_{DS}\|_2\right\}\left(1 - \eta\left(\frac{\tilde{\beta}\gamma_{\min}}{\tilde{\beta} + \sqrt{2}\gamma_{\min}} - 2\epsilon(\tilde{\beta} + \eta\tilde{\beta}^2 + \eta\epsilon\tilde{\beta}^2)\right)\right) ,$$

if the step size $\eta \leq 2\sqrt{2}/(\tilde{\beta} + \sqrt{2}\gamma_{\min})$ and the sample size $n_t = O(\log(t/\delta)/(\epsilon r)^d)$. Here, $\gamma_{\min} = \min\{\gamma_{\hat{1}}, \gamma_{\hat{2}}\}$, $\tilde{\beta} = (\beta^2_{\hat{1}} + \beta^2_{\hat{2}})^{1/2}$ and d is the dimensionality of training sample.

If $\epsilon < \gamma_{\min}/[(\tilde{\beta} + \sqrt{2}\gamma_{\min})(4\eta\tilde{\beta} + 2)]$, then this theorem shows that, with probability at least $1-\delta$, the REGD method would converge to a neighborhood of the decision-dependent stable point at linear rate when they are out of the neighborhood, and would keep in the neighborhood after entrance. The proof is motivated from Theorems 9 and 11, i.e., when the sample size n_t is sufficiently large, they would behave similarly to that on the population level. We will present the detailed proof in Appendix B.6.

We finally consider the stochastic gradient descent for the repeated risk minimization, and formalize it as *repeated stochastic gradient descent*:

Definition 14 *Given initial models $\theta^{\hat{1}}_0$ and $\theta^{\hat{2}}_0$, the Repeated Stochastic Gradient Descent (RSGD) performs the following updates with step size $\eta_t > 0$, for every $t \geq 0$:*

$$\theta^{\hat{i}}_{t+1} = \Pi_\Theta\left(\theta^{\hat{i}}_t - \eta_t\nabla_\theta\ell_{\hat{i}}(\boldsymbol{z}^{(t)}_{\hat{i}};\theta^{\hat{i}}_t)\right) \quad (i \in [2]) ,$$

where $\boldsymbol{z}^{(t)}_{\hat{1}}$ and $\boldsymbol{z}^{(t)}_{\hat{2}}$ are drawn i.i.d. according to distribution $\mathcal{D}(\theta^{\hat{1}}_t, \theta^{\hat{2}}_t)$, and Π_Θ denotes a Euclidean projection operator.

For RSGD, two models only observe a stochastic gradient in each iteration respectively, and then update models. Denote by $\theta_t = (\theta^{\hat{1}}_t, \theta^{\hat{2}}_t)$ and $\theta_{\mathrm{DS}} = (\theta^{\hat{1}}_{\mathrm{DS}}, \theta^{\hat{2}}_{\mathrm{DS}})$. For the convergence of RSGD, it is necessary to make the second moment bounded assumption on the gradient of the loss functions as follows:

Assumption 1 *Suppose that loss function $\ell_{\hat{i}}(\boldsymbol{z};\theta^{\hat{i}})$ is $\beta_{\hat{i}}$-jointly smooth and $\gamma_{\hat{i}}$-strongly convex for $i \in [2]$, and there exist constants $\sigma^2_{\hat{i}}$ and $L^2_{\hat{i}}$ such that, for every $\theta^{\hat{i}}, \theta^{\hat{1}}_*, \theta^{\hat{2}}_* \in \Theta$,*

$$\mathbb{E}_{\boldsymbol{z}\sim\mathcal{D}(\theta^{\hat{1}}_*,\theta^{\hat{2}}_*)}\left[\|\nabla\ell^{\hat{i}}(\boldsymbol{z};\theta^{\hat{i}})\|^2_2\right] \leq \sigma^2_{\hat{i}} + L^2_{\hat{i}}\left\|\theta^{\hat{i}} - G_{\hat{i}}(\theta^{\hat{1}}_*, \theta^{\hat{2}}_*)\right\|^2_2 ,$$

with $G_{\hat{i}}(\theta^{\hat{1}}_*, \theta^{\hat{2}}_*) = \arg\min_{\theta^{\hat{i}}}\mathbb{E}_{\boldsymbol{z}\sim\mathcal{D}(\theta^{\hat{1}}_*,\theta^{\hat{2}}_*)}\left[\ell^{\hat{i}}(\boldsymbol{z};\theta^{\hat{i}})\right]$.

This assumption has been customarily made in stochastic optimization literature (Bottou et al., 2018; Mendler-Dünner et al., 2020). Under such assumption, we have the following theorem for the convergence of RSGD:

Theorem 15 *For ϵ-Lipschitz continuous distribution $\mathcal{D}$ and for the RSGD method after T iterations, we have*

$$\mathbb{E}\big[\|\theta_T - \theta_{DS}\|_2^2\big] \leq \frac{\max\{2\tilde{\sigma}^2, 8L_{\max}^2\|\theta_0 - \theta_{DS}\|_2^2\}}{(\gamma_{\min} - \epsilon\tilde{\beta})^2 T + 8L_{\max}^2} \ ,$$

if step size $\eta_t = \big((\gamma_{\min} - \tilde{\beta}\epsilon)t + 8L_{\max}^2/(\gamma_{\min} - \tilde{\beta}\epsilon)\big)^{-1}$ and $\epsilon < \gamma_{\min}/\tilde{\beta}$. Here, $\tilde{\beta} = (\beta_{\hat{1}}^2 + \beta_{\hat{2}}^2)^{1/2}$, $\tilde{\sigma} = (\sigma_{\hat{1}}^2 + \sigma_{\hat{2}}^2)^{1/2}$, $\gamma_{\min} = \min\{\gamma_{\hat{1}}, \gamma_{\hat{2}}\}$ and $L_{\max} = \max\{L_{\hat{1}}, L_{\hat{2}}\}$.

This theorem shows the convergence of RSGD to a decision-dependent stable point in expectation if $\epsilon < \gamma_{\min}/\tilde{\beta}$, and we get the $O(T^{-1/2})$ convergence rate, which is comparable to general stochastic gradient descent in machine learning. The detailed proof is presented in Appendix B.7.

6. Experiments

Following strategic classification (Dalvi et al., 2004) on interaction between classification rules and strategic agents, we present the empirical studies to verify our theoretical analysis in two-player strategic classification, where instances modify features to improve outcomes.

We use the hotel-booking dataset (Antonio et al., 2019) to simulate the game for customers and two hotels. It consists of two hotels' book information for customers, and the label indicates the cancel or not of a booking. We sample 4900 positive and 4900 negative instances for each hotel from original dataset. Here, two hotels $\hat{1}, \hat{2}$ try to predict the cancel or not for booking order of a customer, and they use logistic regression parameterized by $\theta^{\hat{1}}, \theta^{\hat{2}}$ with L_1-regularization to ensure the strong convexity for loss functions.

Given two models $\theta^{\hat{1}}$ and $\theta^{\hat{2}}$, each instance $(\boldsymbol{x}, y)$ could modify features via maximizing the utility $u(\boldsymbol{x}', \theta^{\hat{1}}, \theta^{\hat{2}})$ and cost $c(\boldsymbol{x}, \boldsymbol{x}')$ as follows:

$$\boldsymbol{x}_{\text{new}} = \arg\max_{\boldsymbol{x}'} \ u(\boldsymbol{x}', \theta^{\hat{1}}, \theta^{\hat{2}}) - c(\boldsymbol{x}, \boldsymbol{x}') \ .$$

Suppose that the customers hope to be classified as negative class for discounts by optimizing the following linear utility:

$$u(\boldsymbol{x}, \theta^{\hat{1}}, \theta^{\hat{2}}) = -\lambda \boldsymbol{x}^{\mathrm{T}}\theta^{\hat{1}} - (1-\lambda)\boldsymbol{x}^{\mathrm{T}}\theta^{\hat{2}} \ \text{ with } \ \lambda \in [0, 1] \ .$$

We consider the quadratic cost $c(\boldsymbol{x}, \boldsymbol{x}') = \frac{1}{2\epsilon}\|\boldsymbol{x} - \boldsymbol{x}'\|_2^2$, and the best response for customers can be written as

$$\begin{aligned} \boldsymbol{x}_{\text{new}} &= \arg\min_{\boldsymbol{x}'} \lambda \boldsymbol{x}'^{\mathrm{T}}\theta^{\hat{1}} + (1-\lambda)\boldsymbol{x}'^{\mathrm{T}}\theta^{\hat{2}} + \frac{1}{2\epsilon}\|\boldsymbol{x} - \boldsymbol{x}'\|_2^2 \\ &= \boldsymbol{x} - \epsilon(\lambda\theta^{\hat{1}} + (1-\lambda)\theta^{\hat{2}}) \ . \end{aligned}$$

It is easy to check the ϵ-Lipschitz continuity for data distribution, since for $\theta^{\hat{1}}, \theta^{\hat{2}}, \theta^{\hat{1}'}, \theta^{\hat{2}'} \in \Theta$, the following holds for every instance,

$$\begin{aligned} &\big\|\boldsymbol{x} - \epsilon(\lambda\theta^{\hat{1}} + (1-\lambda)\theta^{\hat{2}}) - \boldsymbol{x} + \epsilon(\lambda\theta^{\hat{1}'} + (1-\lambda)\theta^{\hat{2}'})\big\|_2 \\ &\leq \ \epsilon \max\{\lambda, 1-\lambda\}\big\|(\theta^{\hat{1}} - \theta^{\hat{1}'}) + (\theta^{\hat{2}} - \theta^{\hat{2}'})\big\|_2 \\ &\leq \ \epsilon\big\|(\theta^{\hat{1}}, \theta^{\hat{2}}) - (\theta^{\hat{1}'}, \theta^{\hat{2}'})\big\|_2 \ . \end{aligned}$$

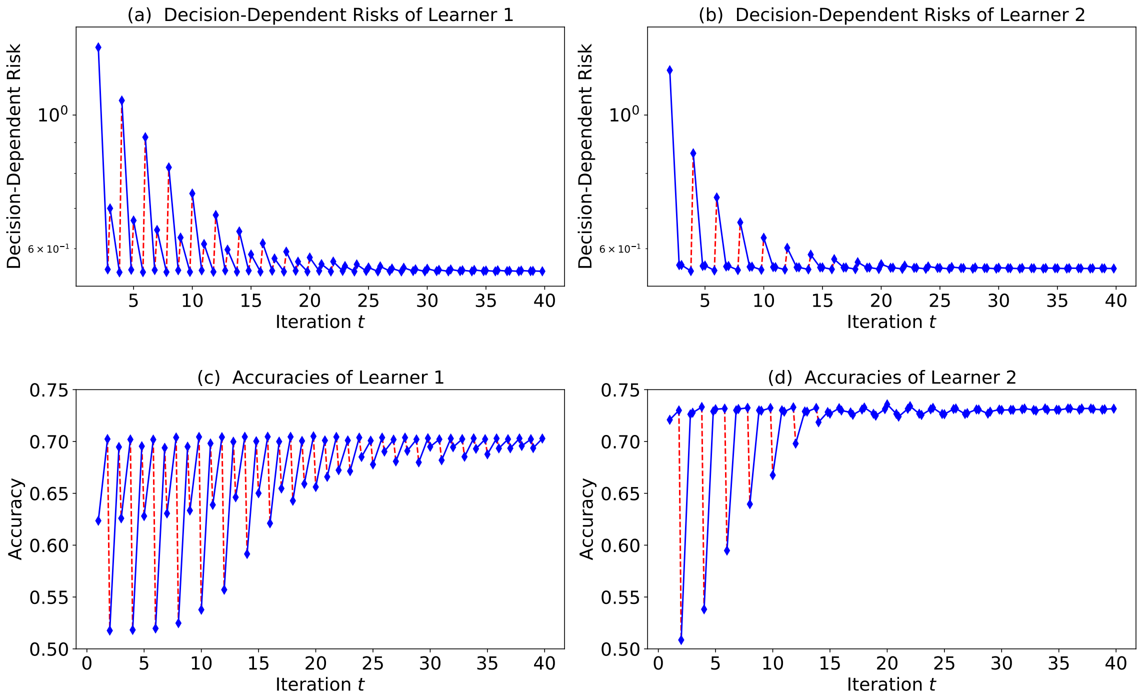

Figure 4: The convergence curves of decision-dependent risks (top) and accuracies (bottom) for RRM with $\epsilon = 13$. Blue line indicates the optimization for models, and red dotted line indicates the effect caused by decision-dependent distribution shift.

This bounds the transport distance for each instance. Therefore, we can also bound the Wasserstein-1 distance over the entire data distribution. In the following, we treat the points in the original dataset as the true distribution, and simulated the strategic classification with $\lambda = 0.5$ for RRM, RGD and RSGD, and adjust the parameter ϵ with different scales.

Figure 4 shows the evolution of the decision-dependent risks and accuracies for two models with RRM when the distribution Lipschitz parameter $\epsilon = 13$. The optimizing progress of RRM for two models is shown in blue lines, while red dotted lines denote the influences caused by updating features of strategic instances. We can clearly observe the interactions between models and instances, i.e., models try to minimize losses yet instances try to increase the losses in each iteration. We can also observe the stable convergence of RMM to a decision-dependent stable point, as expected in Theorem 7.

Figures 5 shows the convergence curves of the distances for RRM with two models after one iteration. As we can see, RRM could converge to a decision-dependent stable point with linear rate for small ϵ, but fail to converge for large ϵ. Moreover, the smaller the parameter ϵ, the faster the convergence. This is nicely in agreement with Theorem 7.

Figures 6 shows the convergence curves of the distances for RGD with two models after one iteration. We can also observe the convergence and divergence of RGD for small and large ϵ, respectively. In particularly, the RGD method would enter a chaos status for

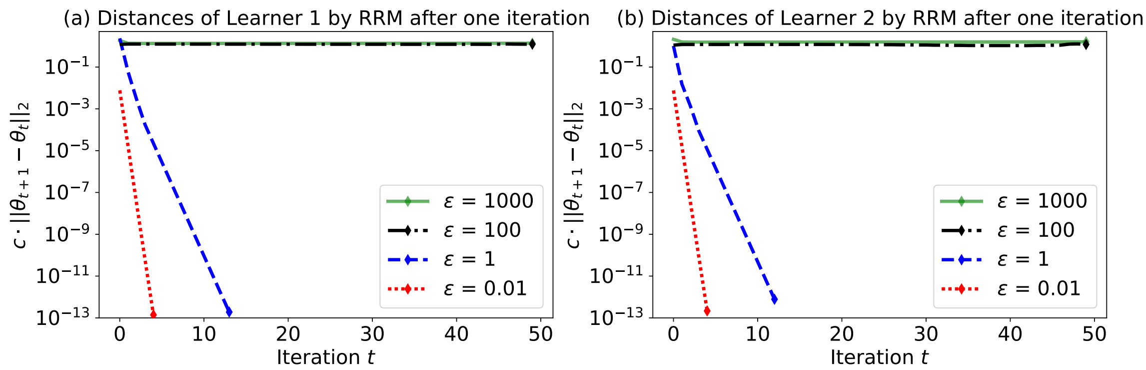

Figure 5: The convergence curves of the normalized distances for two models after one iteration by RRM with different Lipschitz continuity parameter ϵ.

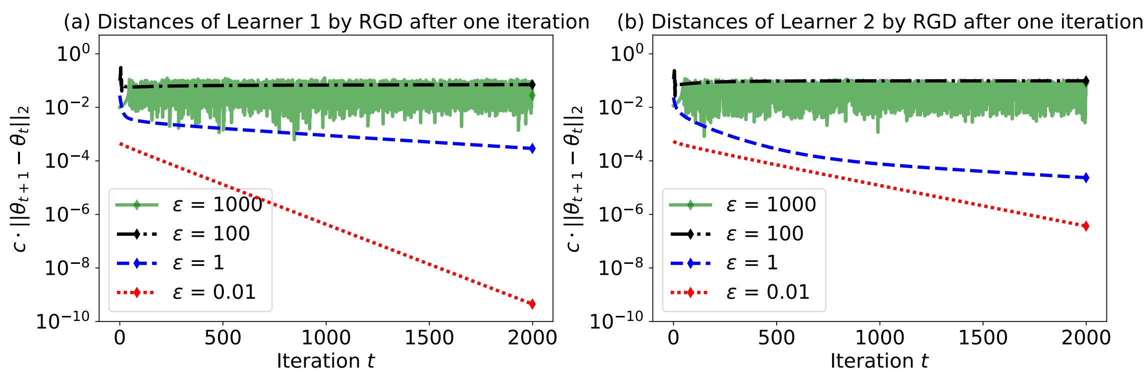

Figure 6: The convergence curves of the normalized distances for two models after one iteration by RGD with different Lipschitz continuity parameter ϵ.

extremely larger ϵ. This verifies Theorem 11 empirically. In comparisons with RRM, we can also find that RGD requires stronger condition for convergence with slower rate, which agrees with Theorems 7 and 11 empirically.

Figure 7 shows the convergence curves of accuracies for RSGD with different Lipschitz continuity parameter ϵ. As we can see, the RSGD method would also converge to some stable point for small ϵ, but diverge for large ϵ. This well supports our Theorem 15 empirically. Moreover, we can find that large ϵ result in hard learning problem for models, causing chaos learning dynamics, leading to low accuracies for both models. This also address the importance of fast convergence to the decision-dependent stable point.

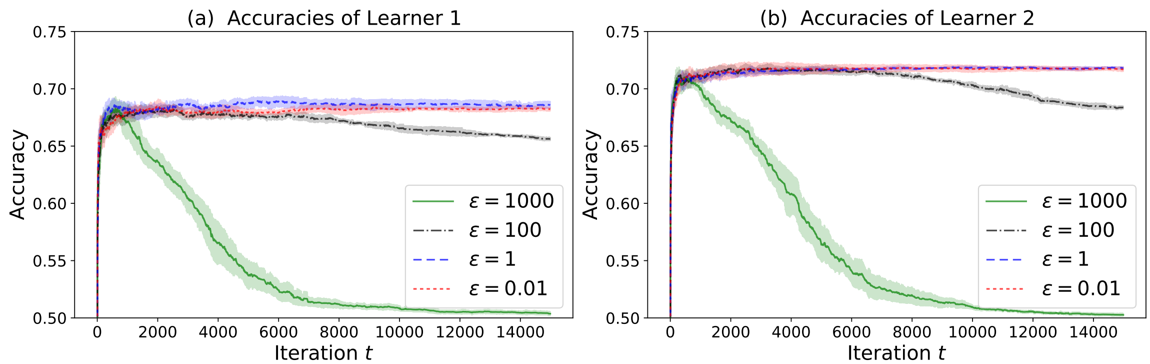

Figure 7: The convergence curves of accuracies for RSGD with different Lipschitz continuity parameter ϵ.

7. Conclusion

This work tries to learn multiple interactive models under a decision-dependent dynamic environment, where the data distribution may take shift according to different models. We formalize such scenario as a learning problem of two models over decision-dependent distributions, and develop the repeated risk minimization method for two models. We present the existence of stable points for RRM, and provide convergence analysis based on data distribution and finite training sample. We also study more practical algorithms with convergence analysis. An interesting future work is to study the learning problem over decision-dependent distributions with three or more models, and it is also interesting to exploit practical algorithms in the more general dynamic environment.

Acknowledgments

The authors want to thank the reviewers for helpful suggestions. This work is supported by National Key R&D Program of China (2021ZD0112802) and NSFC (61876078).

References

C. Aliprantis, K. Border, et al. *Infinite Dimensional Analysis.* Springer, 2006.

N. Antonio, A. D. Almeida, and L. Nunes. Hotel booking demand datasets. *Data in Brief*, 22:41–49, 2019.

D. Balduzzi, S. Racanière, J. Martens, J. N. Foerster, K. Tuyls, and T. Graepel. The mechanics of n-player differentiable games. In *Proceedings of the 35th International Conference on Machine Learning*, pages 363–372, 2018.

Y. Bechavod, K. Ligett, Z. S. Wu, and J. Ziani. Causal feature discovery through strategic modification. *CoRR/abstract*, 2002.07024, 2020.

L. Bottou, F. Curtis, and J. Nocedal. Optimization methods for large-scale machine learning. *Society for Industrial and Applied Mathematics Review*, 60(2):223–311, 2018.

M. Bravo, D. S. Leslie, and P. Mertikopoulos. Bandit learning in concave n-person games. In *Advances in Neural Information Processing Systems 31*, pages 5666–5676, 2018.

G. Brown, S. Hod, and I. Kalemaj. Performative prediction in a stateful world. *CoRR/abstract*, 2011.03885, 2020.

S. Bubeck. Convex optimization: Algorithms and complexity. *Foundations and Trends® in Machine Learning*, 8(3-4):231–357, 2015.

B. Chasnov, L. Ratliff, E. Mazumdar, and S. Burden. Convergence analysis of gradient-based learning in continuous games. In *Proceedings of the 36th Uncertainty in Artificial Intelligence*, pages 935–944, 2020.

N. Dalvi, P. Domingos, S. Sanghai, and D. Verma. Adversarial classification. In *Proceedings of the 10th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*, pages 99–108, 2004.

D. Drusvyatskiy and L. Xiao. Stochastic optimization with decision-dependent distributions. *CoRR/abstract*, 2011.11173, 2020.

N. Fournier and A. Guillin. On the rate of convergence in wasserstein distance of the empirical measure. *Probability Theory and Related Fields*, 162(3):707–738, 2015.

D. Fudenberg and D. K. Levine. *The Theory of Learning in Games*. MIT Press, 1998.

G. Ghalme, V. Nair, I. Eilat, I. Talgam-Cohen, and N. Rosenfeld. Strategic classification in the dark. In *Proceedings of the 38th International Conference on Machine Learning*, pages 3672–3681, 2021.

I. Goodfellow, Y. Bengio, and A. Courville. *Deep Learning*. MIT Press, 2016.

M. Hardt, N. Megiddo, C. H. Papadimitriou, and M. Wootters. Strategic classification. In *Proceedings of the 7th ACM Conference on Innovations in Theoretical Computer Science*, pages 111–122, 2016.

K. Healy. The performativity of networks. *European Journal of Sociology/Archives Européennes de Sociologie/Europäisches Archiv für Soziologie*, 56(2):175–205, 2015.

M. Heusel, H. Ramsauer, T. Unterthiner, B. Nessler, and S. Hochreiter. Gans trained by a two time-scale update rule converge to a local nash equilibrium. In *Advances in Neural Information Processing Systems 30*, pages 6626–6637, 2017.

Z. Izzo, L. Ying, and J. Zou. How to learn when data reacts to your model: Performative gradient descent. In *Proceedings of the 38th International Conference on Machine Learning*, pages 4641–4650, 2021.

S. Kakutani. A generalization of brouwer's fixed point theorem. *Duke Mathematical Journal*, 8(3):457–459, 1941.

C. Maheshwari, C.-Y. Chiu, E. Mazumdar, S. Sastry, and L. J. Ratliff. Zeroth-order methods for convex-concave min-max problems: Applications to decision-dependent risk minimization. In *The 25th International Conference on Artificial Intelligence and Statistics*, volume 151, pages 6702–6734, 2022.

C. Mendler-Dünner, J. Perdomo, T. Zrnic, and M. Hardt. Stochastic optimization for performative prediction. In *Advances in Neural Information Processing Systems 33*, 2020.

J. Miller, J. Perdomo, and T. Zrnic. Outside the echo chamber: Optimizing the performative risk. In *Proceedings of the 38th International Conference on Machine Learning*, pages 7710–7720, 2021.

T. Mitchell. *Machine Learning*. McGraw-Hill, 1997.

Y. Nesterov. *Lectures on Convex Optimization*. Springer, 2018.

N. Nisan, T. Roughgarden, É. Tardos, and V. V. Vazirani. *Algorithmic Game Theory*. Cambridge University Press, 2007.

J. Perdomo, T. Zrnic, C. Mendler-Dünner, and M. Hardt. Performative prediction. In *Proceedings of the 37th International Conference on Machine Learning*, pages 7599–7609, 2020.

M. Ray, D. Drusvyatskiy, M. Fazel, and L. J. Ratliff. Decision-dependent risk minimization in geometrically decaying dynamic environments. *CoRR/abstract*, 2204.08281, 2022.

M. H. Ribeiro, R. Ottoni, R. West, V. AF Almeida, and W. Meira Jr. Auditing radicalization pathways on youtube. In *Proceedings of the ACM Conference on Fairness, Accountability, and Transparency*, pages 131–141, 2020.

S. Shalev-Shwartz and S. Ben-David. *Understanding Machine Learning - From Theory to Algorithms*. Cambridge University Press, 2014.

S. Srinivasan, M. Lanctot, V. F. Zambaldi, J. Pérolat, K. Tuyls, R. Munos, and M. Bowling. Actor-Critic policy optimization in partially observable multiagent environments. In *Advances in Neural Information Processing Systems 31*, pages 3426–3439, 2018.

C. Villani. *Optimal Transport: Old and New*. Springer, 2009.

E.-V. Vlatakis-Gkaragkounis, L. Flokas, and G. Piliouras. Poincaré recurrence, cycles and spurious equilibria in gradient-descent-ascent for non-convex non-concave zero-sum games. In *Advances in Neural Information Processing Systems 32*, pages 10450–10461, 2019.

K. Wood, G. Bianchin, and E. Dall'Anese. Online projected gradient gescent for stochastic optimization with decision-dependent distributions. *IEEE Control Systems Letters*, 6: 1646–1651, 2022.

Z. Zhou and H. Xu. Deep reinforcement learning based intelligent decision making for two-player sequential game with uncertain irrational player. In *Proceedings of the IEEE Symposium Series on Computational Intelligence*, pages 9–15, 2019.