

“可预测的”非凸在线点对学习的遗憾界

郎璇聪¹, 王梅^{1,2}, 刘勇^{3,4}, 李春生^{1,2}

(1. 东北石油大学 计算机与信息技术学院 黑龙江 大庆 163318; 2. 东北石油大学 黑龙江省石油大数据与智能分析重点实验室 黑龙江 大庆 163318; 3. 中国人民大学 高瓴人工智能学院 北京 100049; 4. 中国人民大学 大数据管理与分析方法研究北京市重点实验室 北京 100049)

摘要: 在线点对学习是指损失函数依赖于一对实例的在线学习模式,泛化性是在线点对学习理论研究的重要内容。大多数关于在线点对学习使用对抗性损失函数且假设凸性,然而凸性并不符合通常的实际应用场景。针对非凸在线点对学习,给出基于稳定性分析的具有“可预测的”损失函数的在线点对学习遗憾界,作出相应的稳定性分析。通过稳定性和度量在线点对学习泛化能力的常用方式——遗憾之间的关系,建立“可预测的”非凸损失函数下的遗憾界。证明当学习者(learner)能够获得离线神谕(oracle)时,“可预测的”非凸广义在线点对学习达到遗憾界 $O(T^{-3/2})$ 。该结论丰富了非凸在线点对学习的理论研究,且优于现有理论保证。

关键词: 在线点对学习; 非凸; 稳定性; 遗憾界; 离线神谕

中图分类号: TP301.5

文献标志码: A

文章编号: 1671-6841(2025)04-0055-08

DOI: 10.13705/j.issn.1671-6841.2023199

Regret Bounds for Online Pairwise Learning with Predictable Non-convex Loss Functions

LANG Xuancong¹, WANG Mei^{1,2}, LIU Yong^{3,4}, LI Chunsheng^{1,2}

(1. College of Computer and Information Technology, Northeastern Petroleum University, Daqing 163318, China; 2. Heilongjiang Key Laboratory of Petroleum Big Data and Intelligent Analysis, Northeastern Petroleum University, Daqing 163318, China; 3. Gaoling School of Artificial Intelligence, Renmin University of China, Beijing 100049, China; 4. Beijing Key Laboratory of Big Data Management and Analysis Method at Renmin University of China, Beijing 100049, China)

Abstract: Online pairwise learning is a machine learning model in which the loss functions depend on a pair of instances. Generalization is an important aspect of online pairwise learning theory research. Most of the existing works on online pairwise learning used adversarial loss functions and provided regret bounds only with convex loss functions. However, convexity was not typically applicable in practical scenarios. For non-convex online pairwise learning, the regret bound of online pairwise learning with a "predictable" loss function based on stability analysis was provided and the corresponding stability analysis was conducted. Through the relationship between stability and regret, a common way to measure the generalization ability of online pairwise learning, the regret bound was established with a "predictable" non-convex loss function. It was proved that when the learner obtained an offline oracle, "predictable" non-convex generalized online pairwise learning reached the regret bound of $O(T^{-3/2})$. This study enriched the theoretical research on non-convex online pairwise learning and was superior to the existing theoretical guarantees.

Key words: online pairwise learning; non-convex; stability; regret bounds; offline optimization oracle

收稿日期: 2023-08-30

基金项目: 国家自然科学基金项目(51774090, 62076234); 黑龙江省博士后科研启动金资助项目(LBH-Q20080); 黑龙江省研究生精品课程建设项目(15141220103)

第一作者: 郎璇聪(1998—), 女, 博士研究生, 主要从事大规模机器学习研究, E-mail: xuancong@163.com。

通信作者: 刘勇(1986—), 男, 教授, 主要从事机器学习、大规模机器学习和统计学习理论研究, E-mail: liuyonggsai@ruc.edu.cn。

0 引言

在线点对学习主要应用于排序^[1-3]、度量学习^[4]等。相较于传统的批量学习,在线点对学习具有高效性、适应性和实时性,更适用于实际应用场景和大规模数据应用,近年来得到了广泛的关注。

在线点对学习的泛化能力是学习性能的重要表现之一,关于在线点对学习泛化性,有研究假设在线环境中有一系列独立同分布的数据且损失函数一致有界的条件,通过在线转批量的转换界,首次提出在线点对学习的超过泛化(excess generalization)界 $O(T^{-1})$ ^[5-6]。Kar等^[7]通过期望对称化改进了上述泛化界。Ying等^[8-9]基于非正则的再生核希尔伯特空间(RKHS)的在线点对学习算法,放松了损失函数强凸性和有界性的假设条件,得到遗憾界 $O(T^{-1/3} \log T)$ 。Chen等^[10]通过迭代序列更紧的一致约束,提高了算法最后迭代的收敛率,给出遗憾界 $O(T^{-1/2})$ 。Guo等^[11]基于正则的RKHS,对于特定的铰链损失函数有收敛结果 $O(T^{-1/4} \log T^{1/2})$ 。Wang等^[12]对具有多项式衰减步长和多个正则化参数的在线点对学习算法,给出遗憾界 $O(T^{-2/3})$ 。Hu等^[13]提出了基于分而治之策略的多正则化的分布式在线点对学习的误差分析。

作为分析泛化性的重要工具之一,稳定性已经广泛应用于单点学习算法的泛化边界研究^[14-15]。但是,除了Shen等^[16]和Lei等^[17]的工作,关于点对学习的稳定性研究较少。前者建立了随机梯度下降(stochastic gradient descent,SGD)在凸和强凸损失函数中的稳定性结果,从而给出泛化边界,并权衡了SGD的泛化误差和优化误差,后者提供了一个更细化的稳定性分析并改进泛化边界为 $O(\gamma \log n)$ 。

在线点对学习的理论研究中,通常假设损失函数是凸或强凸且基于对抗性选择的,关于非凸损失函数的研究较为匮乏。如果学习者(learner)接收一个已知的“可预测的”损失函数序列,“可预测的”是指具有关于损失函数序列先验知识的在线学习模式,比如方差^[18]和路径长度的界限^[19],那么,遗憾界优于对抗性选择下的界^[20]。

针对上述问题,本文提出了基于稳定性分析的“可预测的”非凸在线点对学习的遗憾界。本文主要工作有:1)把原始的在线点对学习改写为广义的在线点对学习框架,并基于广义在线点对学习提出“可预测的”非凸广义在线点对学习算法;2)建立了具有“可预测的”非凸损失函数的广义在线点对

学习的稳定性分析;3)根据稳定性和遗憾之间的关系,证明了具有“可预测的”非凸损失函数的广义在线点对学习可达到最优的遗憾界 $O(T^{-3/2})$ 。

1 在线点对学习

假设一个样本空间 $Z = X \times Y$,其中: $X \subset \mathbf{R}^d$ 是一个输入空间; $Y \subset \mathbf{R}$ 是一个输出空间; $x \in X$ 和 $y \in Y$ 分别是输入和输出。在线点对学习是学习一个预测函数 $p: X \times X \rightarrow \mathbf{R}, p(w \in W)$ 在一对实例 (z, z') 上的性能由非负损失函数 $\ell(w, z, z')$ 度量,即

$$w_t \in \operatorname{argmin} \left\{ \sum_{i=1}^{t-1} \ell_i(w, z, z'), w \in W \right\}.$$

排序和度量学习是点对学习经典的应用。对于排序,函数 $p: X \rightarrow Y$ 以与输出一致的方式对实例 $z = (x, y), z' = (x', y')$ 进行排序,即若 $y < y'$, 则 $p(x) < p(x')$ 。点对学习问题排序的形式化表述为 $\ell(w, z, z') = \psi(\operatorname{sgn}(y - y')(p(x) - p(x')))$, 其中: sgn 是符号函数; ψ 可以是铰链损失 $\psi(t) = \max\{1 - t, 0\}$, 也可以是逻辑损失 $\psi(t) = \log(1 + \exp(-t))$ 。对于 $Y = \{-1, 1\}$ 的监督的度量学习,距离函数对具有相同标签的实例是相似的,而具有不同标签的实例彼此分离,一个常用的距离函数是 $p(x, x') = \langle w, (x - x')(x - x')^T \rangle$ 。点对学习问题度量学习的形式化表述为 $\ell(w, z, z') = \psi(\tau(y, y')(1 - p(x, x')))$, 如果 $y = y'$, 则 $\tau(y, y') = 1$, 否则为 -1 。

2 “可预测的”非凸广义在线点对学习

2.1 广义在线点对学习

广义在线点对学习表示为

$$w_t \in \operatorname{argmin} \left\{ \sum_{i=1}^{t-1} \ell_i(w, z, z') + \sigma^T w; w \in W \right\}.$$

广义在线点对学习中包含一个 σ 项,它是一个从 $\exp(\eta)$ (参数为 η 的指数分布)采样的随机噪声。引入随机噪声项 σ 避免过拟合,提高在线点对学习的泛化能力。因此,广义在线点对学习是鲁棒的在线点对学习更广义的在线框架。FTRL(follow the regularized leader)是一种广泛用于凸假设的算法^[21]。当广义在线点对学习中的随机噪声为 μw 时,FTRL即为广义在线点对学习的一个特例,

$$w_t \in \operatorname{argmin} \left\{ \sum_{i=1}^{t-1} \ell_i(w, z, z') + \mu \|w\|^2 \right\},$$

$$\mu \approx T^p, \rho \in (0, 1).$$

基于广义在线点对学习框架,学习者接收的损

失序列是“可预测的”这种乐观变体,可以得到更优的遗憾界。在某些应用中,损失函数可能不是对抗性的,而是“可预测的”。文献[20]表明,当带有自洽正则器的 FTRL 算法和在线镜像下降(online mirroring descent, OMD)类型的算法包含一组可预测过程时,可以获得更好的遗憾界。在介绍广义在线点对学习算法之前,简要介绍算法涉及的基本概念。

2.2 离线神谕

有专家建议的在线点对学习的预测是学习者和对手之间的重复游戏,其特征是学习者从含有 N 个专家的有限集合 \mathbf{H} 中进行选择,对手从一组决策集合 \mathbf{Y} 中选择,损失函数 $\ell_t \in \mathbf{L}$ 。首先,在游戏开始前,对手从 \mathbf{Y} 中选择一个任意的决策序列 y_1, y_2, \dots , 在游戏的每一轮 $t = 1, 2, \dots$, 学习者必须选择(可能是随机的)一个专家 $h_t \in \mathbf{H}$, 然后对手揭示其决策 $y_t \in \mathbf{Y}$, 学习者遭受损失。学习者每接收一对实例 $z, z' \in \mathbf{Z}$ 后的目标是使学习者遭受的损失与最优假设 $w^* \in \mathbf{W}$ 之间的累积差异尽可能小,因此,遗憾是一种衡量在线点对学习性能的标准,定义为

$$\mathcal{R}_A(T) = \sum_{t=1}^T [\ell_t(w_t, z, z') - \ell_t(w^*, z, z')].$$

基于神谕的学习模型可称之为“可优化的专家”,该模型本质上是经典的在线学习模型。在“可优化的专家”模型中,假设最初学习者对损失函数 ℓ 是未知的,并允许学习者通过神谕来获取 ℓ 。离线神谕模型通过在线学习者提交一系列的损失函数,返回使累积损失最小的假设。下面给出离线神谕模型的近似神谕模型定义。

定义 1^[22] 一个离线神谕模型,其输入是一个损失函数 $\ell: \mathbf{W} \rightarrow \mathbf{R}$ 和一个 d 维向量 σ , 输出是一个 $w \mapsto \ell(w, z, z') - \langle \sigma, w \rangle$ 的近似最小值。若它返回的 $w^* \in \mathbf{W}$ 满足

$$\ell(w^*, z, z') - \langle \sigma, w^* \rangle \leq \inf_{w \in \mathbf{W}} [\ell(w, z, z') - \langle \sigma, w \rangle] + (\alpha + \beta \|\sigma\|_1),$$

则称其为“ (α, β) -近似神谕”。

方便起见,将一个 (α, β) -近似神谕记为 $ORAL_{\alpha, \beta}(\ell - \langle \sigma, \cdot \rangle)$ 。使用近似神谕是因为不可能知道一个假设是全局最小值或仅是一个鞍点。 $ORAL$ 可以输出一个 w 而不需要保证其最优性。在大多数情况下, w 只是近似最优。离线神谕模型(返回 $w^* \in \arg\min \ell(w)$)与 (α, β) -近似神谕模型的区别在于,后者有一个关于变量 α, β, σ 的扰动项。在指定 α, β 的大小时可获得更好的遗憾界。近似神谕在在线单点学习中的应用包括以下实例。

1) Query-GAN 算法。当采用近似神谕时, $\alpha = \frac{\hat{R} + R_d(T)}{T}$ (\hat{R} 为累积遗憾), 对于非凸损失函数以 $T^{-1/2}$ 的速度收敛^[23];

2) FTRL 和 FTL (follow the leader) 算法。当采用近似神谕时, $\alpha = T^{-1/2}$, 对于半凸损失函数可以达到 T^{-1} 的遗憾界^[24]。

在实例 1) 和 2) 中, $\beta = 0$, 只有 α 作为近似神谕的参数。

2.3 “可预测的”非凸广义在线点对学习算法

假设损失函数是“可预测的”且学习者可以访问一个 (α, β) -近似离线优化神谕。令 $g_t[\ell_1 \dots \ell_{t-1}]$ 为第 t 轮对 ℓ_t 的预测, $g_1 = 0$ 。给定 g_t , 学习者可以访问一个 (α, β) -近似离线优化神谕

$$w_t = ORAL_{\alpha, \beta}(\sum_{i=1}^{t-1} \ell_i + g_t - \langle \sigma_t, \cdot \rangle),$$

则“可预测的”非凸广义在线点对学习算法如算法 1。

算法 1 “可预测的”非凸广义在线点对学习算法

输入: $\eta, ORAL_{\alpha, \beta}$, 预测序列 $g_t[\ell_1 \dots \ell_{t-1}]$ 。

输出: $w^* \in \mathbf{W}$ 。

1) for $t = 1, \dots, T$ do

2) $\{\sigma_{t,j}\}_{j=1}^d \stackrel{i.i.d.}{\sim} \exp(\eta) / *$ 生成随机向量 σ_t $*$ /

3) 学习者在 t 时刻预测

4) $w_t = ORAL_{\alpha, \beta}(\sum_{i=1}^{t-1} \ell_i + g_t - \langle \sigma_t, \cdot \rangle)$

5) 学习者遭受损失 ℓ_t

6) end for

序列 g_t 相当于在线点对学习中的先验知识,显然当 g_t 近似 ℓ_t 时,这种非凸且“可预测的”乐观情况会有较小的遗憾。

3 稳定性分析

算法稳定性可以衡量一个学习算法的输出对训练数据集微小变化的敏感性,是机器学习理论分析中一个重要的概念。对于批量学习假设中独立同分布的样本,稳定性是可学习性的一个关键特性。类似地,稳定性对于在线学习的可学习性具有同样作用。除常用的一致稳定性以外,还有定义 2 所述的平均稳定性。接下来用 $a^T b$ 或 $\langle a, b \rangle$ 表示 a 和 b 之间的欧几里得点积; $\|\cdot\|$ 表示二范数; $\|\cdot\|_p$ 表示一个特定的 ℓ_p 范数。若 $|f(x) - f(y)| \leq L \|x - y\|$,

$\forall \mathbf{x}, \mathbf{y} \in C$, 则称 $f: C \rightarrow \mathbf{R}$ 是关于 $\|\cdot\|$ 范数 L -Lipschitz 连续的。

定义 2^[17,25] 假设学习者遭受的损失序列是 L -Lipschitz 连续的。若 $\exists \gamma > 0$ 使得

$$\frac{1}{T} \sum_{i=1}^T \mathbb{E} \|\ell_i(\mathbf{w}_i, \mathbf{z}, \mathbf{z}') - \ell_{i+1}(\mathbf{w}_{i+1}, \mathbf{z}, \mathbf{z}')\| \leq L\gamma, \quad (1)$$

则称算法 A 是 γ -平均稳定的。

显然, 平均稳定性比一致稳定性 ($\|\ell_i(\mathbf{w}_i, \mathbf{z}, \mathbf{z}') - \ell_{i+1}(\mathbf{w}_{i+1}, \mathbf{z}, \mathbf{z}')\| \leq L\gamma$) 更弱, 因为它成立的条件仅仅是 \mathbf{w}_i 的期望值和平均值满足不等式(1)。本文主要研究平均稳定性, 所有定理采用的也是平均稳定性, 以此作为理论分析中一种假设条件的放松。

稳定性分析首先将广义在线点对学习的期望遗憾与平均稳定性建立联系。如定理 1 所示, 构建了“可预测的”非凸广义在线点对学习的期望遗憾与平均稳定性的相关性。

定理 1 令 $\bar{\mathbf{w}}_i$ 是 $\sum_{i=1}^{i-1} \ell_i(\mathbf{w}) - \langle \boldsymbol{\sigma}, \mathbf{w} \rangle$ 最小化, 假设 D 为决策集 $\mathbf{W} \subseteq \mathbf{R}^d$ 的界, 学习者遭受的损失满足 ℓ_1 范数的 L_i -Lipschitz 连续性。则“可预测的”非凸广义在线点对学习的上界为

$$\begin{aligned} & \frac{1}{T} \mathbb{E} \left[\sum_{i=1}^T \ell_i(\mathbf{w}_i, \mathbf{z}, \mathbf{z}') - \inf_{\mathbf{w} \in \mathbf{W}} \sum_{i=1}^T \ell_i(\mathbf{w}, \mathbf{z}, \mathbf{z}') \right] \leq \\ & \sum_{i=1}^T \underbrace{\frac{L_i}{T} \mathbb{E} [\|\mathbf{w}_i - \bar{\mathbf{w}}_{i+1}\|_1]}_{\text{Stability}} + \frac{d(\beta T + D)}{\eta T} + \alpha. \end{aligned} \quad (2)$$

证明 令 $\Delta_i(\mathbf{w}) = \ell_i(\mathbf{w}, \mathbf{z}, \mathbf{z}') - g_i(\mathbf{w})$, $\gamma(\boldsymbol{\sigma}) = \alpha + \beta \|\boldsymbol{\sigma}\|_1$ 。对于任意的 $\mathbf{w}^* \in \mathbf{W}$,

$$\begin{aligned} & \sum_{i=1}^T [\ell_i(\mathbf{w}_i, \mathbf{z}, \mathbf{z}') - \ell_i(\mathbf{w}^*, \mathbf{z}, \mathbf{z}')] = \\ & \sum_{i=1}^T [\Delta_i(\mathbf{w}_i) - \Delta_i(\bar{\mathbf{w}}_{i+1})] + \sum_{i=1}^T [g_i(\mathbf{w}_i) - g_i(\bar{\mathbf{w}}_{i+1})] + \\ & \sum_{i=1}^T [\ell_i(\bar{\mathbf{w}}_{i+1}, \mathbf{z}, \mathbf{z}') - \ell_i(\mathbf{w}^*, \mathbf{z}, \mathbf{z}')] \leq \\ & \sum_{i=1}^T L_i \|\mathbf{w}_i - \bar{\mathbf{w}}_{i+1}\|_1 + \sum_{i=1}^T [g_i(\mathbf{w}_i) - g_i(\bar{\mathbf{w}}_{i+1})] + \\ & \sum_{i=1}^T [\ell_i(\bar{\mathbf{w}}_{i+1}, \mathbf{z}, \mathbf{z}') - \ell_i(\mathbf{w}^*, \mathbf{z}, \mathbf{z}')]. \end{aligned}$$

归纳法证明, 对于任意的 $T, \mathbf{w}^* \in \mathbf{W}$,

$$\sum_{i=1}^T [g_i(\mathbf{w}_i) - g_i(\bar{\mathbf{w}}_{i+1})] + \sum_{i=1}^T [\ell_i(\bar{\mathbf{w}}_{i+1}, \mathbf{z}, \mathbf{z}') - \ell_i(\mathbf{w}^*, \mathbf{z}, \mathbf{z}')] \leq \langle \boldsymbol{\sigma}, \bar{\mathbf{w}}_2 - \mathbf{w}^* \rangle + \gamma(\boldsymbol{\sigma})(T-1)$$

均成立。

$T=1$ 步骤: $g_1 = 0$, 由于 $\bar{\mathbf{w}}_2$ 是 $\ell_1(\mathbf{w}, \mathbf{z}, \mathbf{z}') - \langle \boldsymbol{\sigma}, \mathbf{w} \rangle$ 的近似最小化, 得

$$\begin{aligned} & \ell_1(\bar{\mathbf{w}}_2, \mathbf{z}, \mathbf{z}') - \langle \boldsymbol{\sigma}, \bar{\mathbf{w}}_2 \rangle \leq \ell_1(\mathbf{w}^*, \mathbf{z}, \mathbf{z}') - \langle \boldsymbol{\sigma}, \mathbf{w}^* \rangle, \\ & \forall \mathbf{w}^* \in \mathbf{W}. \end{aligned}$$

因此, $\ell_1(\bar{\mathbf{w}}_2, \mathbf{z}, \mathbf{z}') - \ell_1(\mathbf{w}^*, \mathbf{z}, \mathbf{z}') \leq \langle \boldsymbol{\sigma}, \bar{\mathbf{w}}_2 - \mathbf{w}^* \rangle$ 。

归纳步骤: 假设归纳对所有 $T \leq T_0 - 1$ 成立, 则它对 T_0 也成立

$$\begin{aligned} & \sum_{i=1}^{T_0} [g_i(\mathbf{w}_i) - g_i(\bar{\mathbf{w}}_{i+1})] + \sum_{i=1}^{T_0} \ell_i(\bar{\mathbf{w}}_{i+1}, \mathbf{z}, \mathbf{z}') \leq \\ & \left[\sum_{i=1}^{T_0-1} \ell_i(\mathbf{w}_{T_0}, \mathbf{z}, \mathbf{z}') + \langle \boldsymbol{\sigma}, \bar{\mathbf{w}}_2 - \mathbf{w}_{T_0} \rangle + \gamma(\boldsymbol{\sigma})(T_0 - 2) \right] + \\ & [g_{T_0}(\mathbf{w}_{T_0}) - g_{T_0}(\bar{\mathbf{w}}_{T_0+1}) + \ell_{T_0}(\bar{\mathbf{w}}_{T_0+1}, \mathbf{z}, \mathbf{z}')] = \\ & \left[\sum_{i=1}^{T_0-1} \ell_i(\mathbf{w}_{T_0}, \mathbf{z}, \mathbf{z}') + g_{T_0}(\mathbf{w}_{T_0}) - \langle \boldsymbol{\sigma}, \mathbf{w}_{T_0} \rangle \right] + \\ & [\langle \boldsymbol{\sigma}, \bar{\mathbf{w}}_2 \rangle - g_{T_0}(\bar{\mathbf{w}}_{T_0+1}) + \ell_{T_0}(\bar{\mathbf{w}}_{T_0+1}, \mathbf{z}, \mathbf{z}')] + \\ & \gamma(\boldsymbol{\sigma})(T_0 - 2) \leq \\ & \left[\sum_{i=1}^{T_0-1} \ell_i(\bar{\mathbf{w}}_{T_0+1}, \mathbf{z}, \mathbf{z}') + \right. \\ & \left. g_{T_0}(\bar{\mathbf{w}}_{T_0+1}) - \langle \boldsymbol{\sigma}, \bar{\mathbf{w}}_{T_0+1} \rangle \right] + [\langle \boldsymbol{\sigma}, \bar{\mathbf{w}}_2 \rangle - g_{T_0}(\bar{\mathbf{w}}_{T_0+1}) + \\ & \ell_{T_0}(\bar{\mathbf{w}}_{T_0+1}, \mathbf{z}, \mathbf{z}')] + \gamma(\boldsymbol{\sigma})(T_0 - 1) = \\ & \left[\sum_{i=1}^{T_0} \ell_i(\bar{\mathbf{w}}_{T_0+1}, \mathbf{z}, \mathbf{z}') - \langle \boldsymbol{\sigma}, \bar{\mathbf{w}}_{T_0+1} \rangle \right] + \\ & \langle \boldsymbol{\sigma}, \bar{\mathbf{w}}_2 \rangle + \gamma(\boldsymbol{\sigma})(T_0 - 1) \leq \\ & \sum_{i=1}^{T_0} \ell_i(\mathbf{w}^*, \mathbf{z}, \mathbf{z}') + \\ & \langle \boldsymbol{\sigma}, \bar{\mathbf{w}}_2 - \mathbf{w}^* \rangle + \gamma(\boldsymbol{\sigma})(T_0 - 1). \end{aligned}$$

其中: $\left[\sum_{i=1}^{T_0-1} \ell_i(\mathbf{w}_{T_0}, \mathbf{z}, \mathbf{z}') + \langle \boldsymbol{\sigma}, \bar{\mathbf{w}}_2 - \mathbf{w}_{T_0} \rangle + \gamma(\boldsymbol{\sigma})(T_0 - 2) \right] + [g_{T_0}(\mathbf{w}_{T_0}) - g_{T_0}(\bar{\mathbf{w}}_{T_0+1}) + \ell_{T_0}(\bar{\mathbf{w}}_{T_0+1}, \mathbf{z}, \mathbf{z}')]$ 是由于对任意的 $T \leq T_0 - 1$ 均成立; $\left[\sum_{i=1}^{T_0-1} \ell_i(\bar{\mathbf{w}}_{T_0+1}, \mathbf{z}, \mathbf{z}') + g_{T_0}(\bar{\mathbf{w}}_{T_0+1}) - \langle \boldsymbol{\sigma}, \bar{\mathbf{w}}_{T_0+1} \rangle \right] + [\langle \boldsymbol{\sigma}, \bar{\mathbf{w}}_2 \rangle - g_{T_0}(\bar{\mathbf{w}}_{T_0+1}) + \ell_{T_0}(\bar{\mathbf{w}}_{T_0+1}, \mathbf{z}, \mathbf{z}')] + \gamma(\boldsymbol{\sigma})(T_0 - 1)$ 是由于 \mathbf{w}_{T_0} 的近似最优性。

综上所述, 得到“可预测的”非凸广义在线点对学习有上界

$$\begin{aligned} & \sum_{i=1}^T \ell_i(\mathbf{w}_i, \mathbf{z}, \mathbf{z}') - \inf_{\mathbf{w} \in \mathbf{W}} \sum_{i=1}^T \ell_i(\mathbf{w}, \mathbf{z}, \mathbf{z}') \leq \\ & \sum_{i=1}^T L_i \|\mathbf{w}_i - \bar{\mathbf{w}}_{i+1}\|_1 + \frac{d(\beta T + D)}{\eta} + \alpha T, \end{aligned}$$

证毕。

由定理1建立了稳定性和“可预测的”非凸广义在线点对学习的遗憾界之间的关系,当平均稳定性的上界可得时,遗憾也可以实现收敛。下面求稳定性 $\|w_i - \bar{w}_{t+1}\|_1$ 的上界。

定理2 令 $w_i(\sigma)$ 是“可预测的”非凸广义在线点对学习在第 t 次的预测,对于任意的 $c > 0$ 有以下单调性成立

$$w_{t,i}(\sigma + ce_i) \geq w_{t,i}(\sigma) - \frac{2(\alpha + \beta \|\sigma\|_1)}{c} - \beta,$$

其中: σ 为随机扰动; e_i 表示第 i 个标准基向量; $w_{t,i}$ 表示 w_t 在 i 坐标上的假设。

证明 令 $\ell_{1:t}(\mathbf{w}, \mathbf{z}, \mathbf{z}') = \sum_{i=1}^t \ell_i(\mathbf{w}, \mathbf{z}, \mathbf{z}')$, $\sigma' = \sigma + ce_i$, $\gamma(\sigma) = \alpha + \beta \|\sigma\|_1$ 为离线神谕的近似误差,由 $w_i(\sigma)$ 的近似最优化,可得

$$\begin{aligned} & \ell_{1:t-1}(w_i(\sigma), \mathbf{z}, \mathbf{z}') + g_t(w_i(\sigma)) - \langle \sigma, w_i(\sigma) \rangle \leq \\ & \ell_{1:t-1}(w_i(\sigma'), \mathbf{z}, \mathbf{z}') + g_t(w_i(\sigma')) - \langle \sigma, w_i(\sigma') \rangle + \gamma(\sigma) = \\ & \ell_{1:t-1}(w_i(\sigma'), \mathbf{z}, \mathbf{z}') + g_t(w_i(\sigma')) - \langle \sigma', w_i(\sigma') \rangle + cw_{t,i}(\sigma') + \gamma(\sigma) \leq \\ & \ell_{1:t-1}(w_i(\sigma), \mathbf{z}, \mathbf{z}') + g_t(w_i(\sigma)) - \langle \sigma', w_i(\sigma) \rangle + cw_{t,i}(\sigma') + \gamma(\sigma) + \gamma(\sigma') = \\ & \ell_{1:t-1}(w_i(\sigma), \mathbf{z}, \mathbf{z}') + g_t(w_i(\sigma)) - \langle \sigma, w_i(\sigma) \rangle + c(w_{t,i}(\sigma') - w_{t,i}(\sigma)) + \gamma(\sigma) + \gamma(\sigma'), \end{aligned}$$

其中: $\ell_{1:t-1}(w_i(\sigma), \mathbf{z}, \mathbf{z}') + g_t(w_i(\sigma)) - \langle \sigma', w_i(\sigma) \rangle + cw_{t,i}(\sigma') + \gamma(\sigma) + \gamma(\sigma')$ 是由于 $w_i(\sigma')$ 的近似最优化,结合第一项和最后一项,得

$$w_{t,i}(\sigma') \geq w_{t,i}(\sigma) - \frac{2\gamma(\sigma)}{c} - \beta,$$

证毕。

定理2证明了“可预测的”非凸广义在线点对学习的稳定性。通过观察扰动向量的变化对广义在线点对学习的输出产生的影响,表明了其在一维情况下的单调性,由于在线学习中稳定性是指相邻两次迭代得到的假设之间的距离,所以定理2所得即为一维情况下的稳定性。

定理3 $w_i(\sigma)$ 为“可预测的”非凸广义在线点对学习在第 t 次迭代时的假设,假设 $\|w_i(\sigma) - \bar{w}_{t+1}(\sigma)\|_1 \leq 10d |w_{t,i}(\sigma) - \bar{w}_{t+1,i}(\sigma)|$, 对于任意的 $\sigma' = \sigma + 100L_d e_i$, 可得

$$\begin{aligned} & \min(w_{t,i}(\sigma'), \bar{w}_{t+1,i}(\sigma')) \geq \max(w_{t,i}(\sigma), \bar{w}_{t+1,i}(\sigma)) - \\ & \frac{1}{10} |w_{t,i}(\sigma) - \bar{w}_{t+1,i}(\sigma)| - \frac{3(\alpha + \beta \|\sigma\|_1)}{100L_d} - \beta, \end{aligned}$$

其中: σ 为随机扰动; e_i 表示第 i 个标准基向量; $w_{t,i}$ 表示 w_t 在 i 坐标上的假设。

证明 令 $\ell_{1:t}(\mathbf{w}, \mathbf{z}, \mathbf{z}') = \sum_{i=1}^t \ell_i(\mathbf{w}, \mathbf{z}, \mathbf{z}')$, $\gamma(\sigma) = \alpha + \beta \|\sigma\|_1$ 为离线神谕的近似误差,由 $w_i(\sigma)$ 的近似最优化,可得

$$\begin{aligned} & \ell_{1:t-1}(w_i(\sigma), \mathbf{z}, \mathbf{z}') - \langle \sigma, w_i(\sigma) \rangle + \ell_i(w_i(\sigma), \mathbf{z}, \mathbf{z}') \leq \\ & \ell_{1:t-1}(\bar{w}_{t+1}(\sigma), \mathbf{z}, \mathbf{z}') + g_t(\bar{w}_{t+1}(\sigma)) - \langle \sigma, \bar{w}_{t+1}(\sigma) \rangle + \\ & \ell_i(w_i(\sigma), \mathbf{z}, \mathbf{z}') - g_t(w_i(\sigma)) + \gamma(\sigma) \leq \\ & \ell_{1:t-1}(\bar{w}_{t+1}(\sigma), \mathbf{z}, \mathbf{z}') + g_t(\bar{w}_{t+1}(\sigma)) - \langle \sigma, \bar{w}_{t+1}(\sigma) \rangle + \\ & \ell_i(\bar{w}_{t+1}(\sigma), \mathbf{z}, \mathbf{z}') - g_t(\bar{w}_{t+1}(\sigma)) + \\ & L_t \|w_i(\sigma) - \bar{w}_{t+1}(\sigma)\|_1 + \gamma(\sigma) \leq \\ & \ell_{1:t-1}(\bar{w}_{t+1}(\sigma), \mathbf{z}, \mathbf{z}') + g_t(\bar{w}_{t+1}(\sigma)) - \langle \sigma, \bar{w}_{t+1}(\sigma) \rangle + \\ & \ell_i(\bar{w}_{t+1}(\sigma), \mathbf{z}, \mathbf{z}') - g_t(\bar{w}_{t+1}(\sigma)) + \\ & 10L_d d |w_{t,i}(\sigma) - \bar{w}_{t+1,i}(\sigma)| + \gamma(\sigma), \end{aligned} \quad (3)$$

其中: $\ell_{1:t-1}(\bar{w}_{t+1}(\sigma), \mathbf{z}, \mathbf{z}') + g_t(\bar{w}_{t+1}(\sigma)) - \langle \sigma, \bar{w}_{t+1}(\sigma) \rangle + \ell_i(\bar{w}_{t+1}(\sigma), \mathbf{z}, \mathbf{z}') - g_t(\bar{w}_{t+1}(\sigma)) + L_t \|w_i(\sigma) - \bar{w}_{t+1}(\sigma)\|_1 + \gamma(\sigma)$ 是由于 $\ell_i(\cdot)$ 的 Lipschitz 连续性; $\ell_{1:t-1}(\bar{w}_{t+1}(\sigma), \mathbf{z}, \mathbf{z}') + g_t(\bar{w}_{t+1}(\sigma)) - \langle \sigma, \bar{w}_{t+1}(\sigma) \rangle + \ell_i(\bar{w}_{t+1}(\sigma), \mathbf{z}, \mathbf{z}') - g_t(\bar{w}_{t+1}(\sigma)) + 10L_d d |w_{t,i}(\sigma) - \bar{w}_{t+1,i}(\sigma)| + \gamma(\sigma)$ 是由于 $\|w_i(\sigma) - \bar{w}_{t+1}(\sigma)\|_1$ 上的假设。由 $w_{t+1}(\sigma')$ 的近似最优化,得

$$\begin{aligned} & \ell_{1:t-1}(w_i(\sigma), \mathbf{z}, \mathbf{z}') - \langle \sigma, w_i(\sigma) \rangle + \ell_i(w_i(\sigma), \mathbf{z}, \mathbf{z}') = \\ & \ell_{1:t-1}(w_i(\sigma), \mathbf{z}, \mathbf{z}') - \langle \sigma', w_i(\sigma) \rangle + \\ & \ell_i(w_i(\sigma), \mathbf{z}, \mathbf{z}') + \langle 100L_d e_i, w_i(\sigma) \rangle \geq \\ & \ell_{1:t-1}(\bar{w}_{t+1}(\sigma'), \mathbf{z}, \mathbf{z}') - \langle \sigma', \bar{w}_{t+1}(\sigma') \rangle + \\ & \ell_i(\bar{w}_{t+1}(\sigma'), \mathbf{z}, \mathbf{z}') + 100L_d w_{t,i}(\sigma) = \\ & \ell_{1:t-1}(\bar{w}_{t+1}(\sigma'), \mathbf{z}, \mathbf{z}') - \langle \sigma, \bar{w}_{t+1}(\sigma') \rangle + \\ & \ell_i(\bar{w}_{t+1}(\sigma'), \mathbf{z}, \mathbf{z}') + 100L_d (w_{t,i}(\sigma) - \bar{w}_{t+1,i}(\sigma')) \geq \\ & \ell_{1:t-1}(\bar{w}_{t+1}(\sigma), \mathbf{z}, \mathbf{z}') - \langle \sigma, \bar{w}_{t+1}(\sigma) \rangle + \\ & \ell_i(\bar{w}_{t+1}(\sigma), \mathbf{z}, \mathbf{z}') + 100L_d (w_{t,i}(\sigma) - \bar{w}_{t+1,i}(\sigma')), \end{aligned} \quad (4)$$

其中: $\ell_{1:t-1}(\bar{w}_{t+1}(\sigma), \mathbf{z}, \mathbf{z}') - \langle \sigma, \bar{w}_{t+1}(\sigma) \rangle + \ell_i(\bar{w}_{t+1}(\sigma), \mathbf{z}, \mathbf{z}') + 100L_d (w_{t,i}(\sigma) - \bar{w}_{t+1,i}(\sigma'))$ 是由于 $w_{t+1}(\sigma)$ 的近似最优化。结合式(3)~(4),得

$$\begin{aligned} & \bar{w}_{t+1,i}(\sigma') - w_{t,i}(\sigma) \geq \\ & -\frac{1}{10} |w_{t,i}(\sigma) - \bar{w}_{t+1,i}(\sigma)| - \frac{\gamma(\sigma)}{100L_d d}. \end{aligned} \quad (5)$$

同理

$$\begin{aligned} & \mathbf{w}_{t,i}(\boldsymbol{\sigma}') - \bar{\mathbf{w}}_{t+1,i}(\boldsymbol{\sigma}) \geq \\ & -\frac{1}{10} \left| \mathbf{w}_{t,i}(\boldsymbol{\sigma}) - \bar{\mathbf{w}}_{t+1,i}(\boldsymbol{\sigma}) \right| - \frac{\gamma(\boldsymbol{\sigma})}{100L_t d}, \quad (6) \end{aligned}$$

由定理2的单调性,得

$$\begin{aligned} & \bar{\mathbf{w}}_{t+1,i}(\boldsymbol{\sigma}') - \bar{\mathbf{w}}_{t+1,i}(\boldsymbol{\sigma}) \geq 0, \\ & \mathbf{w}_{t,i}(\boldsymbol{\sigma}') - \mathbf{w}_{t,i}(\boldsymbol{\sigma}) \geq -\frac{3\gamma(\boldsymbol{\sigma})}{100L_t d} - \beta, \quad (7) \end{aligned}$$

结合式(4)~(7),得证。

定理3证明了 d 维情况下“可预测的”非凸广义在线点对学习的稳定性。虽然 d 维区别于一维的单调性证明会更具挑战性,但可以通过分别改变扰动项的每个坐标有效地将分析减少到一维的情况,同理于定理2中由单调性表明在线点对学习的稳定性,可得 d 维情况下的稳定性。

4 “可预测的”非凸广义在线点对学习的遗憾界

定理4 假设 D 为决策集 $\mathbf{W} \subseteq \mathbf{R}^d$ 的界。假设预测序列 g_t 对于任意的 $t \in [T]$, $(g_t - \ell_t)$ 都满足 ℓ_1 范数下的 L_t -Lipschitz 连续性。学习者可访问 (α, β) -近似优化神谕。对于任意的 η , “可预测的”非凸广义在线点对学习的预测都满足以下遗憾界

$$\begin{aligned} & \mathbb{E} \left[\frac{1}{T} \sum_{t=1}^T \ell_t(\mathbf{w}_t, \mathbf{z}, \mathbf{z}') - \frac{1}{T} \inf_{\mathbf{w} \in \mathbf{W}} \sum_{t=1}^T \ell_t(\mathbf{w}, \mathbf{z}, \mathbf{z}') \right] \leq \\ & O(\eta d^2 D \sum_{t=1}^T \frac{L_t^2}{T} + \frac{d(\beta T + D)}{\eta T} + \alpha + \beta d \sum_{t=1}^T \frac{L_t}{T}). \end{aligned}$$

证明 使用与定理2和定理3中相同的符号定义, $\mathbb{E}[\|\mathbf{w}_t(\boldsymbol{\sigma}) - \mathbf{w}_{t+1}(\boldsymbol{\sigma})\|_1]$ 也记作

$$\begin{aligned} & \mathbb{E}[\|\mathbf{w}_t(\boldsymbol{\sigma}) - \mathbf{w}_{t+1}(\boldsymbol{\sigma})\|_1] = \\ & \sum_{i=1}^d \mathbb{E}[\left| \mathbf{w}_{t,i}(\boldsymbol{\sigma}) - \mathbf{w}_{t+1,i}(\boldsymbol{\sigma}) \right|], \quad (8) \end{aligned}$$

因此,由 $\mathbb{E}[\left| \mathbf{w}_{t,i}(\boldsymbol{\sigma}) - \mathbf{w}_{t+1,i}(\boldsymbol{\sigma}) \right|], \forall i \in [d]$ 的界,可得 $\mathbb{E}[\|\mathbf{w}_t(\boldsymbol{\sigma}) - \mathbf{w}_{t+1}(\boldsymbol{\sigma})\|_1]$ 的界。对于任意的 $i \in [d]$, 定义 $\mathbb{E}_{-i}[\left| \mathbf{w}_{t,i}(\boldsymbol{\sigma}) - \mathbf{w}_{t+1,i}(\boldsymbol{\sigma}) \right|]$ 为

$$\begin{aligned} & \mathbb{E}_{-i}[\left| \mathbf{w}_{t,i}(\boldsymbol{\sigma}) - \mathbf{w}_{t+1,i}(\boldsymbol{\sigma}) \right|] := \\ & \mathbb{E}[\left| \mathbf{w}_{t,i}(\boldsymbol{\sigma}) - \mathbf{w}_{t+1,i}(\boldsymbol{\sigma}) \right| \mid \{\boldsymbol{\sigma}_j\}_{j \neq i}], \end{aligned}$$

其中: $\boldsymbol{\sigma}_j$ 是 $\boldsymbol{\sigma}$ 在坐标等于 j 的值。

令 $\mathbf{w}_{\max,i}(\boldsymbol{\sigma}) = \max(\mathbf{w}_{t,i}(\boldsymbol{\sigma}), \mathbf{w}_{t+1,i}(\boldsymbol{\sigma}))$ 。类似地,令 $\mathbf{w}_{\min,i}(\boldsymbol{\sigma}) = \min(\mathbf{w}_{t,i}(\boldsymbol{\sigma}), \mathbf{w}_{t+1,i}(\boldsymbol{\sigma}))$ 。则

$$\begin{aligned} & \mathbb{E}_{-i}[\left| \mathbf{w}_{t,i}(\boldsymbol{\sigma}) - \mathbf{w}_{t+1,i}(\boldsymbol{\sigma}) \right|] = \\ & \mathbb{E}_{-i}[\mathbf{w}_{\max,i}(\boldsymbol{\sigma})] - \mathbb{E}_{-i}[\mathbf{w}_{\min,i}(\boldsymbol{\sigma})], \end{aligned}$$

定义

$$\varepsilon = \{\boldsymbol{\sigma} : \|\mathbf{w}_t(\boldsymbol{\sigma}) - \mathbf{x}_{t+1}(\boldsymbol{\sigma})\|_1 \leq$$

$$10d \left| \mathbf{w}_{t,i}(\boldsymbol{\sigma}) - \mathbf{w}_{t+1,i}(\boldsymbol{\sigma}) \right| \},$$

对下式中的 T_1, T_2 求取下界

$$\begin{aligned} & \mathbb{E}_{-i}[\mathbf{w}_{\min,i}(\boldsymbol{\sigma})] \geq \\ & \mathbb{P}(\boldsymbol{\sigma}_i < 100L_t d) \underbrace{\mathbb{E}_{-i}[\mathbf{w}_{\min,i}(\boldsymbol{\sigma}) \mid \boldsymbol{\sigma}_i < 100L_t d]}_{T_1} + \\ & \underbrace{\mathbb{P}(\boldsymbol{\sigma}_i \geq 100L_t d) \mathbb{E}_{-i}[\mathbf{w}_{\min,i}(\boldsymbol{\sigma}) \mid \boldsymbol{\sigma}_i \geq 100L_t d]}_{T_2}, \quad (9) \end{aligned}$$

由于第 i 个坐标的域位于某个长度为 D 的区间内,并且由于 T_1 和 $\mathbb{E}_{-i}[\mathbf{w}_{\max,i}(\boldsymbol{\sigma})]$ 是这个区间的点,它们的差值被 D 所界限。所以 T_1 的下界为 $\mathbb{E}_{-i}[\mathbf{w}_{\max,i}(\boldsymbol{\sigma})] - D$ 。将 T_2 重新定义为

$$\begin{aligned} T_2 &= \mathbb{P}(\boldsymbol{\sigma}_i \geq 100L_t d) \mathbb{E}_{-i}[\mathbf{w}_{\min,i}(\boldsymbol{\sigma}) \mid \boldsymbol{\sigma}_i \geq 100L_t d] = \\ & \int_{\boldsymbol{\sigma}_i = 100L_t d}^{\infty} \mathbf{w}_{\min,i}(\boldsymbol{\sigma}) P(\boldsymbol{\sigma}_i) d\boldsymbol{\sigma}_i = \int_{\boldsymbol{\sigma}_i = 100L_t d}^{\infty} \mathbf{w}_{\min,i}(\boldsymbol{\sigma}) \eta e^{-\eta \boldsymbol{\sigma}_i} d\boldsymbol{\sigma}_i, \end{aligned}$$

改变上述积分中的一个变量,令 $\boldsymbol{\sigma}_i = \boldsymbol{\sigma}'_i + 100L_t d$ 且 $\boldsymbol{\sigma}' = [\boldsymbol{\sigma}_1, \dots, \boldsymbol{\sigma}_{i-1}, \boldsymbol{\sigma}'_i, \boldsymbol{\sigma}_{i+1}, \dots]$ 是将在第 i 个坐标的 $\boldsymbol{\sigma}$ 替换为 $\boldsymbol{\sigma}'_i$ 得到的向量,得

$$\begin{aligned} & \int_{\boldsymbol{\sigma}_i = 100L_t d}^{\infty} \mathbf{w}_{\min,i}(\boldsymbol{\sigma}) \eta e^{-\eta \boldsymbol{\sigma}_i} d\boldsymbol{\sigma}_i = \\ & \int_{\boldsymbol{\sigma}'_i = 0}^{\infty} \mathbf{w}_{\min,i}(\boldsymbol{\sigma}' + 100L_t d \mathbf{e}_i) \eta e^{-\eta(\boldsymbol{\sigma}'_i + 100L_t d)} d\boldsymbol{\sigma}'_i = \\ & e^{-100\eta L_t d} \int_{\boldsymbol{\sigma}'_i = 0}^{\infty} \mathbf{w}_{\min,i}(\boldsymbol{\sigma}' + 100L_t d \mathbf{e}_i) \eta e^{-\eta \boldsymbol{\sigma}'_i} d\boldsymbol{\sigma}'_i = \\ & e^{-100\eta L_t d} \mathbb{E}_{-i}[\mathbf{w}_{\min,i}(\boldsymbol{\sigma}' + 100L_t d \mathbf{e}_i)]. \end{aligned}$$

则 $T_2 = e^{-100\eta L_t d} \mathbb{E}_{-i}[\mathbf{w}_{\min,i}(\boldsymbol{\sigma} + 100L_t d \mathbf{e}_i)]$ 。将 T_1, T_2 的下界代入式(9)中,可得

$$\begin{aligned} & \mathbb{E}_{-i}[\mathbf{w}_{\min,i}(\boldsymbol{\sigma})] \geq \\ & (1 - \exp(-100\eta L_t d)) (\mathbb{E}_{-i}[\mathbf{w}_{\max,i}(\boldsymbol{\sigma})] - D) + \\ & \exp(-100\eta L_t d) \mathbb{E}_{-i}[\mathbf{w}_{\min,i}(\boldsymbol{\sigma} + 100L_t d \mathbf{e}_i)]. \end{aligned}$$

则 $\mathbb{E}_{-i}[\mathbf{w}_{\min,i}(\boldsymbol{\sigma})]$ 下界为

$$\begin{aligned} & \mathbb{E}_{-i}[\mathbf{w}_{\min,i}(\boldsymbol{\sigma})] \geq \\ & (1 - \exp(-100\eta L_t d)) (\mathbb{E}_{-i}[\mathbf{w}_{\max,i}(\boldsymbol{\sigma})] - D) + \\ & \exp(-100\eta L_t d) \mathbb{P}_{-i}(\varepsilon) \mathbb{E}_{-i}[\mathbf{w}_{\min,i}(\boldsymbol{\sigma} + 100L_t d \mathbf{e}_i) \mid \varepsilon] + \\ & \exp(-100\eta L_t d) \mathbb{P}_{-i}(\varepsilon^c) \mathbb{E}_{-i}[\mathbf{w}_{\min,i}(\boldsymbol{\sigma} + 100L_t d \mathbf{e}_i) \mid \varepsilon^c], \end{aligned}$$

其中: $\mathbb{P}_{-i}(\varepsilon) := \mathbb{P}(\varepsilon \mid \{\boldsymbol{\sigma}_j\}_{j \neq i})$ 。由定理2和定理3证明的单调性,得 $\mathbb{E}_{-i}[\mathbf{w}_{\min,i}(\boldsymbol{\sigma})]$ 下界。 $\gamma(\boldsymbol{\sigma}) = \alpha + \beta \|\boldsymbol{\sigma}\|_1$, 则

$$\begin{aligned} & \mathbb{E}_{-i}[\mathbf{w}_{\min,i}(\boldsymbol{\sigma})] \geq \\ & (1 - \exp(-100\eta L_t d)) (\mathbb{E}_{-i}[\mathbf{w}_{\max,i}(\boldsymbol{\sigma})] - D) + \end{aligned}$$

$$\begin{aligned} & \exp(-100\eta L_t d) \mathbb{P}_{-i}(\varepsilon) \mathbb{E}_{-i}[\mathbf{w}_{\max,i}(\boldsymbol{\sigma}) - \frac{1}{10} \|\mathbf{w}_{t,i}(\boldsymbol{\sigma}) - \mathbf{w}_{t+1,i}(\boldsymbol{\sigma})\| - \beta \mid \varepsilon] + \\ & \exp(-100\eta L_t d) \mathbb{P}_{-i}(\varepsilon^c) \mathbb{E}_{-i}[\mathbf{w}_{\min,i}(\boldsymbol{\sigma}) - \frac{2\gamma(\boldsymbol{\sigma})}{100L_t d} - \beta \mid \varepsilon^c] \geq \\ & (1 - \exp(-100\eta L_t d)) (\mathbb{E}_{-i}[\mathbf{w}_{\max,i}(\boldsymbol{\sigma})] - D) + \\ & \exp(-100\eta L_t d) \mathbb{P}_{-i}(\varepsilon) \mathbb{E}_{-i}[\mathbf{w}_{\max,i}(\boldsymbol{\sigma}) - \frac{1}{10} \|\mathbf{w}_{t,i}(\boldsymbol{\sigma}) - \mathbf{w}_{t+1,i}(\boldsymbol{\sigma})\| - \beta \mid \varepsilon] + \\ & \exp(-100\eta L_t d) \mathbb{P}_{-i}(\varepsilon^c) \mathbb{E}_{-i}[\mathbf{w}_{\max,i}(\boldsymbol{\sigma}) - \frac{1}{10d} \|\mathbf{w}_t(\boldsymbol{\sigma}) - \mathbf{w}_{t+1}(\boldsymbol{\sigma})\|_1 - \frac{2\gamma(\boldsymbol{\sigma})}{100L_t d} - \beta \mid \varepsilon^c], \end{aligned}$$

其中:不等式 $\mathbb{E}_{-i}[\mathbf{w}_{\min,i}(\boldsymbol{\sigma})] \geq (1 - \exp(-100\eta L_t d)) (\mathbb{E}_{-i}[\mathbf{w}_{\max,i}(\boldsymbol{\sigma})] - D) + \exp(-100\eta L_t d) \mathbb{P}_{-i}(\varepsilon) \mathbb{E}_{-i}[\mathbf{w}_{\max,i}(\boldsymbol{\sigma}) - \frac{1}{10} \|\mathbf{w}_{t,i}(\boldsymbol{\sigma}) - \mathbf{w}_{t+1,i}(\boldsymbol{\sigma})\| - \beta \mid \varepsilon] + \exp(-100\eta L_t d) \mathbb{P}_{-i}(\varepsilon^c) \mathbb{E}_{-i}[\mathbf{w}_{\min,i}(\boldsymbol{\sigma}) - \frac{2\gamma(\boldsymbol{\sigma})}{100L_t d} - \beta \mid \varepsilon^c]$ 来自于定理 2 和定理 3,不等式

$$\begin{aligned} & (1 - \exp(-100\eta L_t d)) (\mathbb{E}_{-i}[\mathbf{w}_{\max,i}(\boldsymbol{\sigma})] - D) + \\ & \exp(-100\eta L_t d) \mathbb{P}_{-i}(\varepsilon) \mathbb{E}_{-i}[\mathbf{w}_{\max,i}(\boldsymbol{\sigma}) - \frac{1}{10} \|\mathbf{w}_{t,i}(\boldsymbol{\sigma}) - \mathbf{w}_{t+1,i}(\boldsymbol{\sigma})\| - \beta \mid \varepsilon] + \\ & \exp(-100\eta L_t d) \mathbb{P}_{-i}(\varepsilon^c) \mathbb{E}_{-i}[\mathbf{w}_{\max,i}(\boldsymbol{\sigma}) - \frac{1}{10d} \|\mathbf{w}_t(\boldsymbol{\sigma}) - \mathbf{w}_{t+1}(\boldsymbol{\sigma})\|_1 - \frac{2\gamma(\boldsymbol{\sigma})}{100L_t d} - \beta \mid \varepsilon^c], \end{aligned}$$

来自 ε^c 的定义。由 $\mathbb{P}_{-i}(\varepsilon) \leq 1$,得 $\mathbb{E}_{-i}[\mathbf{w}_{\min,i}(\boldsymbol{\sigma})] \geq (1 - \exp(-100\eta L_t d)) (\mathbb{E}_{-i}[\mathbf{w}_{\max,i}(\boldsymbol{\sigma})] - D) + \exp(-100\eta L_t d) \mathbb{E}_{-i}[\mathbf{w}_{\max,i}(\boldsymbol{\sigma}) - \frac{3\gamma(\boldsymbol{\sigma})}{100L_t d} - \beta] - \exp(-100\eta L_t d) \mathbb{E}_{-i}[\frac{1}{10} \|\mathbf{w}_{t,i}(\boldsymbol{\sigma}) - \mathbf{w}_{t+1,i}(\boldsymbol{\sigma})\| + \frac{1}{10d} \|\mathbf{w}_t(\boldsymbol{\sigma}) - \mathbf{w}_{t+1}(\boldsymbol{\sigma})\|_1] \geq \mathbb{E}_{-i}[\mathbf{w}_{\max,i}(\boldsymbol{\sigma})] - 100\eta L_t d D - \frac{3\gamma(\boldsymbol{\sigma})}{100L_t d} - \beta - \mathbb{E}_{-i}[\frac{1}{10} \|\mathbf{w}_{t,i}(\boldsymbol{\sigma}) - \mathbf{w}_{t+1,i}(\boldsymbol{\sigma})\| + \frac{1}{10d} \|\mathbf{w}_t(\boldsymbol{\sigma}) - \mathbf{w}_{t+1}(\boldsymbol{\sigma})\|_1]$ 。

由于 $\exp(\mathbf{w}) \geq 1 + \mathbf{w}$, 所以上式的最后一步成立则

$$\begin{aligned} & \mathbb{E}_{-i}[\|\mathbf{w}_{t,i}(\boldsymbol{\sigma}) - \mathbf{w}_{t+1,i}(\boldsymbol{\sigma})\|] \leq \\ & \frac{1}{9d} \mathbb{E}_{-i}[\|\mathbf{w}_t(\boldsymbol{\sigma}) - \mathbf{w}_{t+1}(\boldsymbol{\sigma})\|_1] + \\ & \frac{1000}{9} \eta L_t d D + \frac{\mathbb{E}_{-i}[\gamma(\boldsymbol{\sigma})]}{30L_t d} + \frac{10}{9} \beta_0. \end{aligned}$$

由于上述界对任意 $\{\boldsymbol{\sigma}_j\}_{j \neq i}$ 均成立,因此可得无条件期望的界,

$$\begin{aligned} & \mathbb{E}_{-i}[\|\mathbf{w}_{t,i}(\boldsymbol{\sigma}) - \mathbf{w}_{t+1,i}(\boldsymbol{\sigma})\|] \leq \\ & \frac{1}{9d} \mathbb{E}_{-i}[\|\mathbf{w}_t(\boldsymbol{\sigma}) - \mathbf{w}_{t+1}(\boldsymbol{\sigma})\|_1] + \\ & \frac{1000}{9} \eta L_t d D + \frac{\mathbb{E}[\gamma(\boldsymbol{\sigma})]}{30L_t d} + \frac{10}{9} \beta_0. \end{aligned} \tag{10}$$

将式(10)代入到式(8)中,可得稳定性界,

$$\begin{aligned} & \mathbb{E}[\|\mathbf{w}_t(\boldsymbol{\sigma}) - \mathbf{w}_{t+1}(\boldsymbol{\sigma})\|_1] \leq \\ & 125\eta L_t d^2 D + \frac{\beta d}{20\eta L_t} + 2\beta d + \frac{\alpha}{20L_t}, \end{aligned} \tag{11}$$

将式(11)代入式(2)中,得证。

由于定理 4 的假设是学习者可接收 (α, β) -近似优化神谕、损失函数是可预测的且是 L_t -Lipschitz 连续的,因此当 $\eta = \frac{d}{d^{3/2} T^{1/2} - d}$, 定理 4 可得遗憾界 $O(d^{3/2} T^{-3/2} + \alpha + \beta d^{3/2} T^{-1/2})$ 。若取 $\alpha = O(T^{-3/2})$, $\beta = O(T^3)$, 可得最佳的遗憾界 $O(T^{-3/2})$ 。

5 结语

本文通过引入随机噪声提出了广义在线点对学习框架。基于文献[26]提出的近似神谕的近似最优假设,提出了损失是“可预测的”非凸广义在线点对学习算法。进一步,对广义在线点对学习进行泛化性研究,通过稳定性分析,得到“可预测的”广义在线点对学习的最佳遗憾界 $O(T^{-3/2})$ 。

基于在线梯度下降算法,可进一步研究具有“可预测的”非凸损失函数的在线点对学习的遗憾界。此外,本文中的遗憾界和平均稳定性结果是建立在期望意义下的,而如何得到高概率的界也是未来可进行的工作。

参考文献:

[1] 原晋江, 农庆琴. 平行批排序最小化最大完工时间在算法的一个注记 [J]. 郑州大学学报(理学版), 2006, 38(3): 1-3.
YUAN J J, NONG Q Q. A note on an on-line algorithm for the parallel-batching scheduling to minimize makespan [J]. Journal of Zhengzhou university (natural science

- edition), 2006, 38(3): 1-3.
- [2] 李文杰, 马冉. 最大化接收工件个数的在线分批排序问题研究[J]. 郑州大学学报(理学版), 2016, 48(2): 24-28.
LI W J, MA R. Research on the online batch scheduling to maximize total number of the accepted jobs[J]. Journal of Zhengzhou university (natural science edition), 2016, 48(2): 24-28.
- [3] REJCHEL W. On ranking and generalization bounds[J]. Journal of machine learning research, 2012, 13: 1373-1392.
- [4] WEINBERGER K Q, SAUL L K. Distance metric learning for large margin nearest neighbor classification[J]. Journal of machine learning research, 2009, 10: 207-244.
- [5] WANG B, KHARDON R, PECHYONY D, et al. Generalization bounds for online learning algorithms with pairwise loss functions[C]//Annual Conference Computational Learning Theory. Berlin: Springer Press, 2012:1-22.
- [6] WANG Y Y, KHARDON R, PECHYONY D, et al. Online learning with pairwise loss functions[J]. Journal of machine learning research, 2013, 14: 1-37.
- [7] KAR P, SRIPERUMBUDUR B K, JAIN P, et al. On the generalization ability of online learning algorithms for pairwise loss functions[EB/OL]. (2013-05-11) [2023-07-22]. <https://api.semanticscholar.org/CorpusID:8366093>.
- [8] YING Y M, ZHOU D X. Unregularized online learning algorithms with general loss functions[J]. Applied and computational harmonic analysis, 2017, 42(2): 224-244.
- [9] YING Y M, ZHOU D X. Online pairwise learning algorithms[J]. Neural computation, 2016, 28(4): 743-777.
- [10] CHEN X M, LEI Y W. Refined bounds for online pairwise learning algorithms[J]. Neurocomputing, 2018, 275: 2656-2665.
- [11] GUO Z C, YING Y M, ZHOU D X. Online regularized learning with pairwise loss functions[J]. Advances in computational mathematics, 2017, 43(1): 127-150.
- [12] WANG C, HU T. Online regularized pairwise learning with least squares loss[J]. Analysis and applications, 2020, 18(1): 49-78.
- [13] HU T, FAN J, XIANG D H. Convergence analysis of distributed multi-penalty regularized pairwise learning[J]. Analysis and applications, 2020, 18(1): 109-127.
- [14] BOUSQUET O, ELISSEEFF A. Stability and generalization[J]. Journal of machine learning research, 2002, 2: 499-526.
- [15] ELISSEEFF A, EVGENIOU T, PONTIL M. Stability of randomized learning algorithms[J]. Journal of machine learning research, 2005, 6: 55-79.
- [16] SHEN W, YANG Z H, YING Y M, et al. Stability and optimization error of stochastic gradient descent for pairwise learning[J]. Analysis and applications, 2020, 18(5): 887-927.
- [17] LEI Y W, LEDENT A, KLOFT M. Sharper generalization bounds for pairwise learning[J]. Advances in neural information processing systems, 2020, 33: 21236-21246.
- [18] HAZAN E, KALE S. Extracting certainty from uncertainty: regret bounded by variation incosts[J]. Machine learning, 2010, 80(2/3): 165-188.
- [19] CHIANG C K, YANG T B, LEE C J, et al. Online optimization with gradual variations [C] // Conference on Learning Theory. Berlin: Springer Press, 2012:1-20.
- [20] RAKHLIN A, SRIDHARAN K. Online learning with predictable sequences[EB/OL]. (2012-08-18) [2023-07-23]. <https://arxiv.org/abs/1208.3728>. pdf.
- [21] MCMAHAN B. Follow-the-regularized-leader and mirror descent: equivalence theorems and ℓ_1 regularization[C]// Proceedings of the Fourteenth International Conference on Artificial Intelligence and Statistics. Cambridge: MIT Press, 2011: 525-533.
- [22] AGARWAL N, GONEN A, HAZAN E. Learning in non-convex games with an optimization oracle [EB/OL]. (2018-10-17) [2023-07-23]. <https://arxiv.org/abs/1810.07362>. pdf.
- [23] FINE M. Certifiably accurate private data release with generative adversarial networks[D]. Boston: Harvard University, 2020.
- [24] GRNAROVA P, LEVY K Y, LUCCHI A, et al. An online learning approach to generative adversarial networks [EB/OL]. (2017-06-10) [2023-07-22]. <https://arxiv.org/pdf/1706.03269>.
- [25] ROSS S, BAGNELL J A. Stability conditions for online learnability [EB/OL]. (2011-08-16) [2023-07-22]. <https://arxiv.org/pdf/1108.3154>.
- [26] SUGGALA A S, NETRAPALLI P. Online non-convex learning: following the perturbed leader is optimal [EB/OL]. (2019-03-19) [2023-07-22]. <https://arxiv.org/abs/1903.08110>. pdf.