计算机研究与发展 DOI: 10.7544/issn1000-1239.202220221
Journal of Computer Research and Development 60(12): 2806−2813, 2023

基于稳定性分析的非凸损失函数在线点对学习的遗憾界

郎璇聪[1] 李春生[1,2] 刘 勇[3,4] 王 梅[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]（黑龙江省石油大数据与智能分析重点实验室（东北石油大学） 黑龙江大庆 163318）

[3]（中国人民大学高瓴人工智能学院 北京 100872）

[4]（大数据管理与分析方法研究北京市重点实验室（中国人民大学） 北京 100071）

（xuancongl@163.com）

Regret Bounds for Online Pairwise Learning with Non-Convex Loss Functions Using Stability Analysis

Lang Xuancong[1], Li Chunsheng[1,2], Liu Yong[3,4], and Wang Mei[1,2]

[1]（*College of Computer and Information Technology*, *Northeastern Petroleum University*, *Daqing*, *Heilongjiang* 163318）

[2]（*Heilongjiang Provincial Key Laboratory of Petroleum Big Data and Intelligent Analysis* (*Northeastern Petroleum University*), *Daqing*, *Heilongjiang* 163318）

[3]（*Gaoling School of Artificial Intelligence*, *Renmin University of China*, *Beijing* 100872）

[4]（*Beijing Key Laboratory of Big Data Management and Analysis Methods* (*Renmin University of China*), *Beijing* 100071）

Abstract Pairwise learning refers to a learning task which involves a loss function depending on pairs of instances. Recently, there is a growing interest in studying pairwise learning since it includes many important machine learning tasks as specific examples, e.g., metric learning, AUC maximization and ranking. Regret bounds are particularly important for generalization analysis of online pairwise learning. The existing online pairwise learning analysis provides regret bounds only with convex loss functions. To fill the gap in the theoretical study of online pairwise learning with non-convex loss functions, we present a systematic study on the generalization analysis for online pairwise learning and propose regret bounds for non-convex online pairwise learning in this paper. We consider online learning in an adversarial, non-convex setting under the assumption that the learner has access to an offline optimization oracle and the learner’s prediction with expert advice. We first propose a general online pairwise learning framework and establish the stability of online pairwise learning with non-convex loss functions. Then, the regret bounds can be derived naturally from stability. Finally, we show that the general online pairwise learning framework with non-convex loss functions achieves optimal regret bounds of $O(T^{-1/2})$ when the learner has access to an offline optimization oracle.

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

摘 要 点对学习（pairwise learning）是指损失函数依赖于2个实例的学习任务.遗憾界对点对学习的泛

收稿日期：2022-03-14；**修回日期**：2023-01-30

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项目（LBH-Q20080）；黑龙江省自然科学基金项目（LH2020F003）；黑龙江省高校基本科研业务费项目（KYCXTD201903，YYYZX202105）

This work was supported by the National Natural Science Foundation of China (51774090, 62076234), the Heilongjiang Provincial Postdoctoral Research Startup Fund Project (LBH-Q20080), the Natural Science Foundation of Heilongjiang Province (LH2020F003), and the Basic Research Fund Project of Universities in Heilongjiang Province (KYCXTD201903, YYYZX202105).

通信作者：王梅（wangmei@nepu.edu.cn）

化分析尤为重要.现有的在线点对学习分析只提供了凸损失函数下的遗憾界.为了弥补非凸损失函数下在线点对学习理论研究的空白,提出了基于稳定性分析的非凸损失函数在线点对学习的遗憾界.首先提出了一个广义的在线点对学习框架,并给出了具有非凸损失函数的在线点对学习的稳定性分析;然后,根据稳定性和遗憾界之间的关系,对非凸损失函数下的遗憾界进行研究;最后证明了当学习者能够获得离线神谕（oracle）时,具有非凸损失函数的广义在线点对学习框架实现了最佳的遗憾界$O(T^{-1/2})$.

关键词 在线点对学习;非凸;稳定性;遗憾界;离线优化神谕

中图法分类号 TP301.5

点对学习(pairwise learning)在数据挖掘和机器学习中占有很重要的地位.在数据挖掘方面,主要的应用场景有运营式传统行业、互联网领域、物联网领域等[1];在机器学习方面,包括排序[2-5]、接收机操作特性曲线的面积计算[6]、度量学习[7]等.

关于在线点对学习泛化性的研究,Wang等人[8]建立的损失函数是一致有界条件下的泛化分析,给出遗憾界$O(T^{-1})$. Ying等人[9-10]基于非正则的再生核希尔伯特空间(reproducing kernel Hilbert space, RKHS)的在线点对学习算法,在损失函数没有强凸性和有界性的假设条件下,得到遗憾界$O\left(T^{-1/3}\log T\right)$. Chen等人[11]假设迭代序列满足更紧的一致约束,给出遗憾界$O\left(T^{-1/2}\right)$,同时提高了算法最后一次迭代的收敛率. Guo等人[12]基于正则的RKHS,对于特定的铰链损失函数有$O\left(T^{-1/4}\log T^{1/2}\right)$收敛率. Wang等人[13]分析了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最后一次迭代的收敛性,给出遗憾界$O\left(T^{-2/3}\right)$.文献[14]提出了基于分而治之策略的多正则化项的分布式在线点对学习的误差分析.

作为用来分析泛化性的重要工具之一,稳定性分析已经被广泛应用于单点学习算法的泛化边界研究[15-16].但是,除了Shen等人[17]和Lei等人[18]的工作以外,关于点对学习的稳定性研究较少. Shen等人[17]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,从而给出了泛化边界,并权衡了SGD下的泛化误差和优化误差, Lei等人[18]提供了一个更细化的稳定性分析,并将泛化边界改进为$O(\gamma\log n)$.

以上所述都是假设损失函数是强凸或一般凸,缺乏对非凸损失函数的在线点对学习的理论分析.针对这一问题,本文基于稳定性分析,提出了非凸损失函数的在线点对学习的遗憾界.本文包含2点贡献:1)提出了可以扩展到非凸损失函数的广义在线点对学习框架;2)建立了该框架下稳定性和遗憾界之间的关系,并给出了该框架的遗憾界及理论分析.

1 广义在线点对学习框架

1.1 广义在线点对学习

传统的机器学习与在线学习模式不同,传统的机器学习中,一次性地取全部训练样本进行学习,样本分为训练样本和测试样本.而在线学习没有训练样本和测试样本的区分,学习者(learner)实时获取样本,每获取到1个实例,就调整1次假设.

假设一个样本空间$Z=X\times Y$,其中$X\subset\mathbb{R}^d$是一个输入空间,$Y\subset\mathbb{R}$是一个输出空间,$\boldsymbol{x}\in X$和$y\in Y$分别是输入和输出.在线点对学习的学习过程迭代T轮,学习者每轮接收2个实例$(\boldsymbol{x}_t, y_t), (\boldsymbol{x}'_t, y'_t)$,并进行预测,然后接收标签,再依损失函数$\ell_t\in L$遭受的损失更新假设$\boldsymbol{w}_{t+1}\in W$.

广义在线点对学习表示为

$$\boldsymbol{w}_t\in\arg\min\left\{\sum_{i=1}^{t-1}\ell_i(\boldsymbol{w}, z, z')+\boldsymbol{\sigma}^{\mathrm{T}}\boldsymbol{w}:\boldsymbol{w}\in W\right\}.$$

显而易见,与在线点对学习模型相比,广义在线点对学习是一个更鲁棒、更广义的框架.该框架中包含一个$\boldsymbol{\sigma}$项,$\boldsymbol{\sigma}$项是一个从$\exp(\eta)$(参数为η的指数分布)采样的随机噪声.引入随机噪声项$\boldsymbol{\sigma}$避免过拟合,从而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,是一个更广义的在线框架. FTRL(follow the regularized leader)是一种广泛用于凸假设的非常有效的算法[19].事实上,当广义在线点对学习中的随机噪声为$\mu\boldsymbol{w}$时,FTRL即为广义在线点对学习的一个特例:

$$\boldsymbol{w}_t\in\arg\min\left\{\sum_{i=1}^{t-1}\ell_i(\boldsymbol{w}, z, z')+\mu\|\boldsymbol{w}\|^2\right\},$$
$$\mu\approx T^{\rho}, \rho\in(0, 1).$$

1.2 离线神谕模型

直觉上,度量在线学习质量的方式是学习者遭受的损失函数的累计加和,学习者的目标是使累计损失尽可能的小,这也是一般学习模式性能度量的

方式. 但是, 在线学习中的数据通常是对抗性生成的, 对手知道学习者的所有信息, 包括学习者给出的预测函数和损失函数等. 因此, 对手总是会给出与学习者预测值相反的标签, 使得学习者的预测总是错误的, 并遭受最大的损失. 在这种对抗环境中, 累计损失的度量方式难以奏效, 为了解决这个问题, 引入了专家(expert)这一概念. 专家就是一个映射集合$h: X \to Y$, 对输入$\boldsymbol{x} \in X$给出预测$h(\boldsymbol{x})$, 与学习者不同的是, 专家是固定的, 学习者会随着时间根据损失进行更新, 而专家给出的预测不会受到对手影响. 采用专家的损失作为参考, 矫正学习者的预测, 学习者在专家中选择的时候, 只需要选择对输入实例的累计损失函数值最小的专家, 即最优专家. 有了专家的参与, 在线学习性能度量方式就是学习者遭受的损失与最优专家的损失之差的累计加和, 以此避免对手的干扰.

有专家建议的在线点对学习是学习者和对手之间的重复游戏, 其特征是学习者可以从含有N个专家的有限集合H中进行选择, 对手从一组决策集合Y中选择, 还有一个损失函数$\ell_t \in L$. 首先, 在游戏开始前, 对手从Y中选择一个任意的决策序列$y_1, y_2, \cdots$, 在游戏的每一轮$t = 1, 2, \cdots$, 学习者必须选择(可能是随机的)一个专家$h_t \in H$, 然后对手揭示其决策$y_t \in Y$, 学习者遭受损失. 学习者每接收到一对实例$z, z' \in Z$后的目标是使学习者遭受的损失与最优假设$\boldsymbol{w}^* \in W$的损失之间的累积差尽可能的小, 因此, 遗憾界是衡量在线点对学习性能的标准, 对于算法$\mathcal{A}$定义为

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

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, 即学习者通过专家的建议和一个离线神谕进行预测. 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. 离线神谕模型是通过学习者提交一系列的损失函数, 返回使得累积损失最小的假设. 定义1给出离线神谕模型的一种近似神谕模型定义.

定义 1.[20] 一个离线神谕模型, 其输入是一个损失函数$\ell: W \to \mathbb{R}$和一个d维向量$\boldsymbol{\sigma}$, 输出是一个$\boldsymbol{w} \to \ell(\boldsymbol{w}, z, z') - \langle \boldsymbol{\sigma}, \boldsymbol{w} \rangle$的近似最小值. 若它返回的$\boldsymbol{w}^* \in W$满足

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

则称其为“(α, β)-近似神谕”.

为方便起见, 将一个“(α, β)-近似神谕”记为$ORAL_{\alpha,\beta}(\ell - \langle \boldsymbol{\sigma}, \cdot \rangle)$. 使用近似神谕是因为我们不可能知道一个假设是全局最小值还是仅为一个鞍点. $ORAL$可以输出一个$\boldsymbol{w}$而不需要保证其最优性. 在大多数情况下, $\boldsymbol{w}$实际上只是近似最优. 离线神谕模型(返回$\boldsymbol{w}^* \in \arg\min \ell(\boldsymbol{w})$)与$(\alpha, \beta)$-近似神谕模型的区别在于, 近似神谕模型有一个关于变量α, β, $\boldsymbol{\sigma}$的扰动项. 在指定变量α和β的大小时可以获得更好的遗憾界. 近似神谕模型关于在线单点学习的应用包括的实例有: 当 Query-GAN算法采用近似神谕时, 对于非凸损失函数以$T^{-1/2}$的速度收敛[21], $\alpha = \dfrac{\hat{R} + R_d(T)}{T}$($\hat{R}$为累积遗憾); FTRL 和 FTL(follow the leader)算法采用近似神谕时, 对于半凸损失函数可以达到T^{-1}的遗憾界[22], $\alpha = T^{-1/2}$. 在上述应用中, $\beta = 0$, 只有α被用作近似神谕的参数.

2 非凸广义在线点对学习的稳定性分析

2.1 非凸广义在线点对学习算法

非凸广义在线点对学习算法的原理是迭代T轮, 每轮产生一个随机向量$\boldsymbol{\sigma}$(该向量服从关于参数η的指数分布), 并获得一个假设$\boldsymbol{w}$, 该假设来自于(α, β)-近似离线神谕模型, 然后, 学习者会遭受相应损失并调整假设. 算法1给出当学习者可以访问(α, β)-近似离线神谕的非凸广义在线点对学习的算法.

算法 1.非凸的广义在线点对学习算法.

输入: 参数η, 近似神谕$ORAL_{\alpha,\beta}$;

输出: $\boldsymbol{w}^* \in W$.

① for $t = 1, 2, \cdots, T$ do

② $\{\sigma_{t,j}\}_{j=1}^{d} \overset{\text{i.i.d}}{\sim} \exp(\eta)$; /*生成随机向量$\boldsymbol{\sigma}_t$*/

③ 学习者在时刻t的假设:

④ $$\boldsymbol{w}_t = ORAL_{\alpha,\beta}\left(\sum_{i=1}^{t-1} \ell_i - \langle \boldsymbol{\sigma}_t, \cdot \rangle\right);$$

⑤ 学习者遭受损失ℓ_t;

⑥ end for

与在线点对学习相比, 广义在线点对学习中包含一个$\boldsymbol{\sigma}$项, 是一个具有更强鲁棒性、更广义的算法. 一些在线学习算法, 如在线镜像下降(online mirror descent, OMD)[23]、在线梯度下降(online gradient decent, OGD)[24]、FTL、FTRL 等, 通常需要在凸甚至强凸的条件下实现收敛. 文献 [20] 通过损失函数的随机扰动来保证遗憾消失, 这种随机扰动具有与 FTRL 和 OMD 中使用的显式正则项类似的作用, 将广义在线单点学习算法扩展到非凸设置中, 然而缺少关于非凸在线点对学习的内容. 针对这一问题, 本文将单点

学习扩展到点对设置中，通过稳定性分析将在线点对中的点对耦合进行解耦，从形式上把具有耦合的点对通过稳定性分析变换成 2 步假设之间的差，从而实现单点到点对学习的推广.

2.2 稳定性分析

算法稳定性是机器学习理论分析中一个重要的概念. 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. 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，对于在线学习的可学习性，稳定性条件也是同样有效的. 一种特别常用的稳定性度量方式是一致稳定性，已经被广泛应用到在线点对学习中，除此以外，还有由定义 2 给出的平均稳定性. 下文中用$\boldsymbol{a}^{\mathrm{T}}\boldsymbol{b}$或$\langle \boldsymbol{a}, \boldsymbol{b}\rangle$表示$\boldsymbol{a}$和$\boldsymbol{b}$之间的欧几里得点积. 用$\|\cdot\|$表示 2 范数，$\|\cdot\|_p$来表示一个特定的$\ell_p$范数. 若$|f(x)-f(y)|\leqslant G\|x-y\|, \forall x, y\in C$，则称$f: C\to\mathbb{R}$是关于$\|\cdot\|$范数$G$-Lipschitz 连续的.

定义 2.[18,25] 假设学习者遭受的损失序列是G-Lipschitz 连续的. 若$\exists\gamma>0$使得

$$\frac{1}{T}\sum_{t=1}^{T}E\|\ell_t(\boldsymbol{w}_t, z, z')-\ell_{t+1}(\boldsymbol{w}_{t+1}, z, z')\|\leqslant G\gamma, \qquad (1)$$

则称算法$\mathcal{A}$是γ–平均稳定的.

显然，平均稳定性比一致稳定性($\|\ell_t(\boldsymbol{w}_t, z, z')-\ell_{t+1}(\boldsymbol{w}_{t+1}, z, z')\|\leqslant G\gamma$)更弱，因为前者要求的条件仅仅是$\boldsymbol{w}_t$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，以此放松理论分析中的假设条件.

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. 假设D为决策集$W\subseteq\mathbb{R}^d$的界，学习者遭受的损失满足ℓ_1范数的G-Lipschitz 连续性. 则广义在线点对学习的遗憾上界可以表示为

$$\frac{1}{T}E\left[\sum_{t=1}^{T}\ell_t(\boldsymbol{w}_t, z, z')-\inf_{w\in W}\sum_{t=1}^{T}\ell_t(\boldsymbol{w}, z, z')\right]\leqslant \frac{G}{T}\sum_{t=1}^{T}\underbrace{E\left[\|\boldsymbol{w}_t-\boldsymbol{w}_{t+1}\|_1\right]}_{\text{Stability}}+\frac{d(\beta T+D)}{\eta T}+\alpha. \qquad (2)$$

证明. 在“遗忘的对手”模型中，假设对手的决策$\{\ell_t\}_{t=1}^{T}$与广义在线点对学习算法的假设$\{\boldsymbol{w}_t\}_{t=1}^{T}$无关，假设损失函数的序列$\{\ell_t\}_{t=1}^{T}$是预先固定的. 而在“非遗忘的对手”模型中，对手的决策取决于算法过去的假设，即每个ℓ_t由来自某函数$L_t: W^{t-1}\to F$的$\ell_t:=L_t[\boldsymbol{w}_{<t}]$给出，其中$F$是对手所有可能决策的集合，$\boldsymbol{w}_{<t}$是$\boldsymbol{w}_1, \boldsymbol{w}_2, \cdots, \boldsymbol{w}_{t-1}$的缩写，$L_t$是一个常数函数，其中函数$L_1, L_2, \cdots, L_T$决定了一个非遗忘的对手. 令$P_t$是广义在线点对学习基于之前的假设$\boldsymbol{w}_{<t}$给出的假设$\boldsymbol{w}_t$的条件分布，若在“遗忘的对手”中，$P_t$独立于$\boldsymbol{w}_{<t}$. 但无论是在“遗忘的对手”模型或是“非遗忘的对手”模型中，P_t都完全由对手之前的决策$\ell_{<t}$决定. 任何对“遗忘的对手”有界的算法对“非遗忘的对手”也有界[26]，令B是一个正常数，假设广义在线点对学习满足对“遗忘的对手”的遗憾约束$E\left[\sum_{t=1}^{T}\ell_t(\boldsymbol{w}_t)-\inf_{w\in W}\sum_{t=1}^{T}\ell_t(\boldsymbol{w})\right]\leqslant B, \forall\ell_1, \ell_2, \cdots, \ell_T\in F$，那么它也满足对“非遗忘的对手”的遗憾约束$\sum_{t=1}^{T}\ell_t(P_t)-\inf_{w\in W}\sum_{t=1}^{T}\ell_t(\boldsymbol{w})\leqslant B$.

本文研究“遗忘的对手”设定，因此只需使用单一的随机向量$\boldsymbol{\sigma}$，而不是在每次迭代中生成一个新的随机向量. $\boldsymbol{w}_t(\boldsymbol{\sigma})$为非凸广义在线点对学习基于随机扰动$\boldsymbol{\sigma}$，在第$t$次迭代时的假设.

对于任意$\boldsymbol{w}^*\in W$，都有

$$\begin{aligned}&\sum_{t=1}^{T}[\ell_t(\boldsymbol{w}_t, z, z')-\ell_t(\boldsymbol{w}^*, z, z')]=\\&\sum_{t=1}^{T}[\ell_t(\boldsymbol{w}_t, z, z')-\ell_t(\boldsymbol{w}_{t+1}, z, z')]+\\&\sum_{t=1}^{T}\ell_t(\boldsymbol{w}_{t+1}, z, z')-\ell_t(\boldsymbol{w}^*, z, z')\leqslant\\&\sum_{t=1}^{T}G\|\boldsymbol{w}_t-\boldsymbol{w}_{t+1}\|_1+\\&\sum_{t=1}^{T}[\ell_t(\boldsymbol{w}_{t+1}, z, z')-\ell_t(\boldsymbol{w}^*, z, z')]\ .\end{aligned}$$

令$\gamma(\boldsymbol{\sigma})=\alpha+\beta\|\boldsymbol{\sigma}\|_1$. 由归纳法证明

$$\begin{aligned}&\sum_{t=1}^{T}[\ell_t(\boldsymbol{w}_{t+1}, z, z')-\ell_t(\boldsymbol{w}^*, z, z')]\leqslant\\&\gamma(\boldsymbol{\sigma})T+\langle\boldsymbol{\sigma}, \boldsymbol{w}_2-\boldsymbol{w}^*\rangle.\end{aligned}$$

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

$$\begin{aligned}&\ell_1(\boldsymbol{w}_2, z, z')-\langle\boldsymbol{\sigma}, \boldsymbol{w}_2\rangle\leqslant\\&\min_{w\in W}\ell_1(\boldsymbol{w}, z, z')-\langle\boldsymbol{\sigma}, \boldsymbol{w}\rangle+\gamma(\boldsymbol{\sigma})\leqslant\\&\ell_1(\boldsymbol{w}^*, z, z')-\langle\boldsymbol{\sigma}, \boldsymbol{w}^*\rangle+\gamma(\boldsymbol{\sigma}).\end{aligned}\qquad (3)$$

式(3)中最后一个不等式对任何$\boldsymbol{w}^*\in W$均成立. 即

$$\ell_1(\boldsymbol{w}_2, z, z')-\ell_1(\boldsymbol{w}^*, z, z')\leqslant\gamma(\boldsymbol{\sigma})+\langle\boldsymbol{\sigma}, \boldsymbol{w}_2-\boldsymbol{w}^*\rangle.$$

归纳步骤：假设归纳对所有$T\leqslant T_0-1$成立，下面证明它对T_0也成立.

$$\sum_{t=1}^{T_0} \ell_t(\boldsymbol{w}_{t+1}, z, z') \overset{①}{\leqslant} \left[\begin{array}{c} \sum_{t=1}^{T_0-1} \ell_t\left(\boldsymbol{w}_{T_0+1}, z, z'\right) + \\ \langle \boldsymbol{\sigma}, \boldsymbol{w}_2 - \boldsymbol{w}_{T_0+1}\rangle + \gamma(\boldsymbol{\sigma})(T_0-1) \end{array} \right] + \ell_{T_0}\left(\boldsymbol{w}_{T_0+1}, z, z'\right) = \left[\begin{array}{c} \sum_{t=1}^{T_0} \ell_t\left(\boldsymbol{w}_{T_0+1}, z, z'\right) - \\ \langle \boldsymbol{\sigma}, \boldsymbol{w}_{T_0+1}\rangle \end{array} \right] + \langle \boldsymbol{\sigma}, \boldsymbol{w}_2\rangle + \gamma(\boldsymbol{\sigma})(T_0-1) \overset{②}{\leqslant} \sum_{t=1}^{T_0} \ell_t(\boldsymbol{w}^*, z, z') + \langle \boldsymbol{\sigma}, \boldsymbol{w}_2 - \boldsymbol{w}^*\rangle + \gamma(\boldsymbol{\sigma})T_0, \forall \boldsymbol{w}^* \in W,$$

其中①处是由于归纳对任何$T \leqslant T_0-1$都成立，而②处是由于$\boldsymbol{w}_{T_0+1}$的近似最优化.

由上述结果，得到非凸的广义在线点对学习的期望遗憾上界：

$$E\left[\sum_{t=1}^{T} \ell_t(\boldsymbol{w}_t, z, z') - \inf_{\boldsymbol{w}\in W} \sum_{t=1}^{T} \ell_t(\boldsymbol{w}, z, z')\right] \leqslant G\sum_{t=1}^{T} E[\|\boldsymbol{w}_t - \boldsymbol{w}_{t+1}\|_1] + E\left[\gamma(\boldsymbol{\sigma})T + \langle \boldsymbol{\sigma}, \boldsymbol{w}_2 - \boldsymbol{w}^*\rangle\right] \leqslant G\sum_{t=1}^{T} E[\|\boldsymbol{w}_t - \boldsymbol{w}_{t+1}\|_1] + (\beta T + D)\left(\sum_{i=1}^{d} E[\sigma_i]\right) + \alpha T,$$

由指数分布的属性$E[\sigma_i] = \dfrac{1}{\eta_i}$, 得证. 证毕.

定理1表明期望遗憾与平均稳定性相关. 式(2)中的稳定性即为定义2中的平均稳定性. 定理1的证明是受文献[26]的启发，证明了当平均稳定性的上界可得时，遗憾也可以实现收敛. 因此，定理2将着重于研究稳定项$E[\|\boldsymbol{w}_t - \boldsymbol{w}_{t+1}\|_1]$的上界.

定理2. $\boldsymbol{w}_t(\boldsymbol{\sigma})$为广义在线点对学习在第$t$次迭代时的假设，其中，$\boldsymbol{\sigma}$为随机扰动. $\boldsymbol{e}_i$表示第i个标准基向量，$\boldsymbol{w}_{t,i}$表示$\boldsymbol{w}_t$在i坐标上的假设. 对于任意$c>0$，都有单调性成立：

$$\boldsymbol{w}_{t,i}(\boldsymbol{\sigma} + c\boldsymbol{e}_i) \geqslant \boldsymbol{w}_{t,i}(\boldsymbol{\sigma}) - \frac{2(\alpha + \beta\|\boldsymbol{\sigma}\|_1)}{c} - \beta.$$

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

$$\begin{aligned} &\ell_{1:t-1}(\boldsymbol{w}_t(\boldsymbol{\sigma}), z, z') - \langle \boldsymbol{\sigma}, \boldsymbol{w}_t(\boldsymbol{\sigma})\rangle \leqslant \\ &\ell_{1:t-1}(\boldsymbol{w}_t(\boldsymbol{\sigma}'), z, z') - \langle \boldsymbol{\sigma}, \boldsymbol{w}_t(\boldsymbol{\sigma}')\rangle + \gamma(\boldsymbol{\sigma}) \overset{①}{=} \\ &\ell_{1:t-1}(\boldsymbol{w}_t(\boldsymbol{\sigma}'), z, z') - \langle \boldsymbol{\sigma}', \boldsymbol{w}_t(\boldsymbol{\sigma}')\rangle + \\ &c\boldsymbol{w}_{t,i}(\boldsymbol{\sigma}') + \gamma(\boldsymbol{\sigma})a \leqslant \\ &\ell_{1:t-1}(\boldsymbol{w}_t(\boldsymbol{\sigma}), z, z') - \langle \boldsymbol{\sigma}', \boldsymbol{w}_t(\boldsymbol{\sigma})\rangle + \\ &c\boldsymbol{w}_{t,i}(\boldsymbol{\sigma}') + \gamma(\boldsymbol{\sigma}) + \gamma(\boldsymbol{\sigma}') = \\ &\ell_{1:t-1}(\boldsymbol{w}_t(\boldsymbol{\sigma}), z, z') - \langle \boldsymbol{\sigma}, \boldsymbol{w}_t(\boldsymbol{\sigma})\rangle + \\ &c\left(\boldsymbol{w}_{t,i}(\boldsymbol{\sigma}') - \boldsymbol{w}_{t,i}(\boldsymbol{\sigma})\right) + \gamma(\boldsymbol{\sigma}) + \gamma(\boldsymbol{\sigma}'). \end{aligned} \tag{4}$$

其中①处是由于$\boldsymbol{w}_t(\boldsymbol{\sigma}')$的近似最优化. 结合式(4)中的第1项和最后1项，得

$$\boldsymbol{w}_{t,i}(\boldsymbol{\sigma}') \geqslant \boldsymbol{w}_{t,i}(\boldsymbol{\sigma}) - \frac{2\gamma(\boldsymbol{\sigma})}{c} - \beta.$$

证毕.

定理2证明了包含随机扰动项$\boldsymbol{\sigma}$的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2次迭代得到的假设之间的距离，因此，定理2得到的即为一维情况下的稳定性.

定理3. $\boldsymbol{w}_t(\boldsymbol{\sigma})$为广义在线点对学习在第$t$次迭代时的假设，其中$\boldsymbol{\sigma}$为随机扰动. $\boldsymbol{e}_i$表示第i个标准基向量，$\boldsymbol{w}_{t,i}$表示$\boldsymbol{w}_t$在i坐标上的假设. 假设$\|\boldsymbol{w}_t(\boldsymbol{\sigma}) - \boldsymbol{w}_{t+1}(\boldsymbol{\sigma})\|_1 \leqslant 10d\left|\boldsymbol{w}_{t,i}(\boldsymbol{\sigma}) - \boldsymbol{w}_{t+1,i}(\boldsymbol{\sigma})\right|$, 对于$\boldsymbol{\sigma}' = \boldsymbol{\sigma} + 100Gd\boldsymbol{e}_i$, 有单调性成立：

$$\min\left(\boldsymbol{w}_{t,i}(\boldsymbol{\sigma}'), \boldsymbol{w}_{t+1,i}(\boldsymbol{\sigma}')\right) \geqslant \max\left(\boldsymbol{w}_{t,i}(\boldsymbol{\sigma}), \boldsymbol{w}_{t+1,i}(\boldsymbol{\sigma})\right) - \frac{1}{10}\left|\boldsymbol{w}_{t,i}(\boldsymbol{\sigma}) - \boldsymbol{w}_{t+1,i}(\boldsymbol{\sigma})\right| - \frac{3(\alpha + \beta\|\boldsymbol{\sigma}\|_1)}{100Gd} - \beta.$$

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

$$\begin{aligned} &\ell_{1:t-1}(\boldsymbol{w}_t(\boldsymbol{\sigma}), z, z') - \langle \boldsymbol{\sigma}, \boldsymbol{w}_t(\boldsymbol{\sigma})\rangle + \ell_t(\boldsymbol{w}_t(\boldsymbol{\sigma}), z, z') \leqslant \\ &\ell_{1:t-1}(\boldsymbol{w}_{t+1}(\boldsymbol{\sigma}), z, z') - \langle \boldsymbol{\sigma}, \boldsymbol{w}_{t+1}(\boldsymbol{\sigma})\rangle + \\ &\ell_t(\boldsymbol{w}_t(\boldsymbol{\sigma}), z, z') + \gamma(\boldsymbol{\sigma}) \overset{①}{\leqslant} \\ &\ell_{1:t-1}(\boldsymbol{w}_{t+1}(\boldsymbol{\sigma}), z, z') - \langle \boldsymbol{\sigma}, \boldsymbol{w}_{t+1}(\boldsymbol{\sigma})\rangle + \\ &\ell_t(\boldsymbol{w}_{t+1}(\boldsymbol{\sigma}), z, z') + G\|\boldsymbol{w}_t(\boldsymbol{\sigma}) - \boldsymbol{w}_{t+1}(\boldsymbol{\sigma})\|_1 + \gamma(\boldsymbol{\sigma}) \overset{②}{\leqslant} \\ &\ell_{1:t-1}(\boldsymbol{w}_{t+1}(\boldsymbol{\sigma}), z, z') - \langle \boldsymbol{\sigma}, \boldsymbol{w}_{t+1}(\boldsymbol{\sigma})\rangle + \\ &\ell_t(\boldsymbol{w}_{t+1}(\boldsymbol{\sigma}), z, z') + 10Gd\left|\boldsymbol{w}_{t,i}(\boldsymbol{\sigma}) - \boldsymbol{w}_{t+1,i}(\boldsymbol{\sigma})\right| + \gamma(\boldsymbol{\sigma}), \end{aligned} \tag{5}$$

其中①处是由于$\ell_t(\cdot)$的 Lipschitz 连续性，②处是由于$\|\boldsymbol{w}_t(\boldsymbol{\sigma}) - \boldsymbol{w}_{t+1}(\boldsymbol{\sigma})\|_1$上的假设. 由$\boldsymbol{w}_{t+1}(\boldsymbol{\sigma}')$的近似最优化，得

$$\begin{aligned} &\ell_{1:t-1}(\boldsymbol{w}_t(\boldsymbol{\sigma}), z, z') - \langle \boldsymbol{\sigma}, \boldsymbol{w}_t(\boldsymbol{\sigma})\rangle + \ell_t(\boldsymbol{w}_t(\boldsymbol{\sigma}), z, z') = \\ &\ell_{1:t-1}(\boldsymbol{w}_t(\boldsymbol{\sigma}), z, z') - \langle \boldsymbol{\sigma}', \boldsymbol{w}_t(\boldsymbol{\sigma})\rangle + \ell_t(\boldsymbol{w}_t(\boldsymbol{\sigma}), z, z') + \\ &\langle 100Gd\boldsymbol{e}_i, \boldsymbol{w}_t(\boldsymbol{\sigma})\rangle \geqslant \\ &\ell_{1:t-1}(\boldsymbol{w}_{t+1}(\boldsymbol{\sigma}'), z, z') - \langle \boldsymbol{\sigma}', \boldsymbol{w}_{t+1}(\boldsymbol{\sigma}')\rangle + \\ &\ell_t(\boldsymbol{w}_{t+1}(\boldsymbol{\sigma}'), z, z') + 100Gd\boldsymbol{w}_{t,i}(\boldsymbol{\sigma}) - \gamma(\boldsymbol{\sigma}') = \\ &\ell_{1:t-1}(\boldsymbol{w}_{t+1}(\boldsymbol{\sigma}'), z, z') - \langle \boldsymbol{\sigma}, \boldsymbol{w}_{t+1}(\boldsymbol{\sigma}')\rangle - \\ &\gamma(\boldsymbol{\sigma}') + \ell_t(\boldsymbol{w}_{t+1}(\boldsymbol{\sigma}'), z, z') + \\ &100Gd\left(\boldsymbol{w}_{t,i}(\boldsymbol{\sigma}) - \boldsymbol{w}_{t+1,i}(\boldsymbol{\sigma}')\right) \overset{①}{\geqslant} \ell_{1:t-1}(\boldsymbol{w}_{t+1}(\boldsymbol{\sigma}), z, z') - \\ &\langle \boldsymbol{\sigma}, \boldsymbol{w}_{t+1}(\boldsymbol{\sigma})\rangle + \ell_t(\boldsymbol{w}_{t+1}(\boldsymbol{\sigma}), z, z') + \\ &100Gd\left(\boldsymbol{w}_{t,i}(\boldsymbol{\sigma}) - \boldsymbol{w}_{t+1,i}(\boldsymbol{\sigma}')\right) - \gamma(\boldsymbol{\sigma}') - \gamma(\boldsymbol{\sigma}), \end{aligned} \tag{6}$$

其中①处是由于$\boldsymbol{w}_{t+1}(\boldsymbol{\sigma})$的近似最优化. 结合式(5)和式(6)，得

$$\boldsymbol{w}_{t+1,i}(\boldsymbol{\sigma}')-\boldsymbol{w}_{t,i}(\boldsymbol{\sigma})\geqslant -\frac{1}{10}\left|\boldsymbol{w}_{t,i}(\boldsymbol{\sigma})-\boldsymbol{w}_{t+1,i}(\boldsymbol{\sigma})\right|-\frac{3\gamma(\boldsymbol{\sigma})}{100Gd}-\beta. \tag{7}$$

同理

$$\boldsymbol{w}_{t,i}(\boldsymbol{\sigma}')-\boldsymbol{w}_{t+1,i}(\boldsymbol{\sigma})\geqslant -\frac{1}{10}\left|\boldsymbol{w}_{t,i}(\boldsymbol{\sigma})-\boldsymbol{w}_{t+1,i}(\boldsymbol{\sigma})\right|-\frac{3\gamma(\boldsymbol{\sigma})}{100Gd}-\beta. \tag{8}$$

从定理 2 中的单调性，得

$$\boldsymbol{w}_{t+1,i}(\boldsymbol{\sigma}')-\boldsymbol{w}_{t+1,i}(\boldsymbol{\sigma})\geqslant -\frac{3\gamma(\boldsymbol{\sigma})}{100Gd}-\beta, \tag{9}$$

$$\boldsymbol{w}_{t,i}(\boldsymbol{\sigma}')-\boldsymbol{w}_{t,i}(\boldsymbol{\sigma})\geqslant -\frac{3\gamma(\boldsymbol{\sigma})}{100Gd}-\beta. \tag{10}$$

结合上述不等式(7)~(10)得证. 证毕.

定理 3 证明了d维情况下广义在线点对学习的稳定性. 虽然d维相较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来有效地将分析减少到一维. 同理，定理 2 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.

3 非凸广义在线点对学习的遗憾界

利用定理 1 所给出的非凸的广义在线点对学习稳定性分析，对其遗憾界进行研究. 由于定理 2 和定理 3 分别给出了一维和高维情况稳定性$E\left[\|\boldsymbol{w}_t-\boldsymbol{w}_{t+1}\|_1\right]$的界，结合定理 2 和定理 3 引导定理 4, 对定理 4 进行讨论.

定理 4. 假设D为决策集$W\subseteq\mathbb{R}^d$的界. 学习者遭受的损失满足ℓ_1范数的G-Lipschitz 连续性. 学习者可访问(α,β)-近似神谕. 对于任意η, 广义在线点对学习的假设都满足遗憾界：

$$\left[\frac{1}{T}\sum_{t=1}^{T}\ell_t(\boldsymbol{w}_t,z,z')-\frac{1}{T}\inf_{w\in W}\sum_{t=1}^{T}\ell_t(\boldsymbol{w},z,z')\right]\leqslant O\left(\eta d^2DG^2+\frac{d(\beta T+D)}{\eta T}+\alpha+\beta dG\right).$$

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

$$E\left[\|\boldsymbol{w}_t(\boldsymbol{\sigma})-\boldsymbol{w}_{t+1}(\boldsymbol{\sigma})\|_1\right]=\sum_{i=1}^{d}E\left[\left|\boldsymbol{w}_{t,i}(\boldsymbol{\sigma})-\boldsymbol{w}_{t+1,i}(\boldsymbol{\sigma})\right|\right]. \tag{11}$$

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

$$E_{-i}\left[\left|\boldsymbol{w}_{t,i}(\boldsymbol{\sigma})-\boldsymbol{w}_{t+1,i}(\boldsymbol{\sigma})\right|\right]:=E\left[\left|\boldsymbol{w}_{t,i}(\boldsymbol{\sigma})-\boldsymbol{w}_{t+1,i}(\boldsymbol{\sigma})\right|\middle|\left\{\sigma_j\right\}_{j\neq i}\right],$$

其中σ_j是$\boldsymbol{\sigma}$第j个坐标值.

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

$$E_{-i}\left[\left|\boldsymbol{w}_{t,i}(\boldsymbol{\sigma})-\boldsymbol{w}_{t+1,i}(\boldsymbol{\sigma})\right|\right]=E_{-i}\left[\boldsymbol{w}_{\max,i}(\boldsymbol{\sigma})\right]-E_{-i}\left[\boldsymbol{w}_{\min,i}(\boldsymbol{\sigma})\right].$$

定义

$$\varepsilon=\left\{\begin{matrix}\boldsymbol{\sigma}:\|\boldsymbol{w}_t(\boldsymbol{\sigma})-\boldsymbol{w}_{t+1}(\boldsymbol{\sigma})\|_1\leqslant\\ 10d\left|\boldsymbol{w}_{t,i}(\boldsymbol{\sigma})-\boldsymbol{w}_{t+1,i}(\boldsymbol{\sigma})\right|\end{matrix}\right\}.$$

对式(12)中的T_1, T_2求取下界：

$$E_{-i}\left[\boldsymbol{w}_{\min,i}(\boldsymbol{\sigma})\right]=P(\sigma_i<100Gd)\underbrace{E_{-i}\left[\boldsymbol{w}_{\min,i}(\boldsymbol{\sigma})|\sigma_i<100Gd\right]}_{T_1}+\underbrace{P(\sigma_i\geqslant 100Gd)E_{-i}\left[\boldsymbol{w}_{\min,i}(\boldsymbol{\sigma})|\sigma_i\geqslant 100Gd\right]}_{T_2}. \tag{12}$$

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$E_{-i}\left[\boldsymbol{w}_{\max,i}(\boldsymbol{\sigma})\right]$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$E_{-i}\left[\boldsymbol{w}_{\max,i}(\boldsymbol{\sigma})\right]-D$.将$T_2$重新定义为：

$$T_2=P(\sigma_i\geqslant 100Gd)E_{-i}\left[\boldsymbol{w}_{\min,i}(\boldsymbol{\sigma})|B_i\geqslant 100Gd\right]=\int_{\sigma_i=100Gd}^{\infty}\boldsymbol{w}_{\min,i}(\boldsymbol{\sigma})P(\sigma_i)\mathrm{d}\sigma_i=\int_{\sigma_i=100Gd}^{\infty}\boldsymbol{w}_{\min,i}(\boldsymbol{\sigma})\eta\exp(-\eta\sigma_i)\mathrm{d}\sigma_i.$$

改变积分中的 1 个变量，令$\sigma_i=\sigma_i'+100Gd$且$\boldsymbol{\sigma}'=\left(\sigma_1,\sigma_2,\cdots,\sigma_{i-1},\sigma_i',\sigma_{i+1},\cdots\right)$是将在第$i$个坐标的$\boldsymbol{\sigma}$替换为$\sigma_i'$得到的向量，得

$$\int_{\sigma_i=100Gd}^{\infty}\boldsymbol{w}_{\min,i}(\boldsymbol{\sigma})\eta\exp(-\eta\sigma_i)\mathrm{d}\sigma_i=\int_{\sigma_i'=0}^{\infty}\boldsymbol{w}_{\min,i}(\boldsymbol{\sigma}'+100Gd\boldsymbol{e}_i)\eta\exp\left(-\eta\left(\sigma_i'+100Gd\right)\right)\mathrm{d}\sigma_i'=\exp(-100\eta Gd)\times\int_{\sigma_i'=0}^{\infty}\boldsymbol{w}_{\min,i}(\boldsymbol{\sigma}'+100Gd\boldsymbol{e}_i)\eta\exp\left(-\eta\sigma_i'\right)\mathrm{d}\sigma_i'=\exp(-100\eta Gd)E_{-i}\left[\boldsymbol{w}_{\min,i}(\boldsymbol{\sigma}'+100Gd\boldsymbol{e}_i)\right].$$

则$T_2=\exp(-100\eta Gd)E_{-i}\left[\boldsymbol{w}_{\min,i}(\boldsymbol{\sigma}+100Gd\boldsymbol{e}_i)\right]$. 将$T_1$, T_2的下界代入式(12)中，可得

$$E_{-i}\left[\boldsymbol{w}_{\min,i}(\boldsymbol{\sigma})\right]\geqslant(1-\exp(-100\eta Gd))\left(E_{-i}\left[\boldsymbol{w}_{\max,i}(\boldsymbol{\sigma})\right]-D\right)+\exp(-100\eta Gd)E_{-i}\left[\boldsymbol{w}_{\min,i}(\boldsymbol{\sigma}+100Gd\boldsymbol{e}_i)\right].$$

则$E_{-i}\left[\boldsymbol{w}_{\min,i}(\boldsymbol{\sigma})\right]$下界：

$$E_{-i}\left[\boldsymbol{w}_{\min,i}(\boldsymbol{\sigma})\right]\geqslant(1-\exp(-100\eta Gd))\left(E_{-i}\left[\boldsymbol{w}_{\max,i}(\boldsymbol{\sigma})\right]-D\right)+\exp(-100\eta Gd)\times P_{-i}(\varepsilon)E_{-i}\left[\boldsymbol{w}_{\min,i}(\boldsymbol{\sigma}+100Gd\boldsymbol{e}_i)|\varepsilon\right]+\exp(-100\eta Gd)\times P_{-i}(\varepsilon^c)E_{-i}\left[\boldsymbol{w}_{\min,i}(\boldsymbol{\sigma}+100Gd\boldsymbol{e}_i)|\varepsilon^c\right],$$

其中$P_{-i}(\varepsilon):=P\left(\varepsilon|\left\{\sigma_j\right\}_{j\neq i}\right)$. 由定理 2 和定理 3 证明的单调性，得$E_{-i}\left[\boldsymbol{w}_{\min,i}(\boldsymbol{\sigma})\right]$下界. $\gamma(\boldsymbol{\sigma})=\alpha+\beta\|\boldsymbol{\sigma}\|_1$, 则

$$
\begin{aligned}
&E_{-i}\left[\boldsymbol{w}_{\min,i}(\boldsymbol{\sigma})\right] \geqslant \\
&(1-\exp(-100\eta Gd))\left(E_{-i}\left[\boldsymbol{w}_{\max,i}(\boldsymbol{\sigma})\right]-D\right)+ \\
&\exp(-100\eta Gd)P_{-i}(\varepsilon)\times \\
&E_{-i}\left[\begin{array}{l}\boldsymbol{w}_{\max,i}(\boldsymbol{\sigma})-\dfrac{1}{10}\left|\boldsymbol{w}_{t,i}(\boldsymbol{\sigma})-\boldsymbol{w}_{t+1,i}(\boldsymbol{\sigma})\right|- \\ \dfrac{3\gamma(\boldsymbol{\sigma})}{100Gd}-\beta|\varepsilon\end{array}\right]+ \\
&\exp(-100\eta Gd)P_{-i}(\varepsilon^c)E_{-i}\left[\begin{array}{l}\boldsymbol{w}_{\min,i}(\boldsymbol{\sigma})-\dfrac{2\gamma(\boldsymbol{\sigma})}{100Gd}- \\ \beta\mid\varepsilon^c\end{array}\right]\geqslant \\
&(1-\exp(-100\eta Gd))\left(E_{-i}\left[\boldsymbol{w}_{\max,i}(\boldsymbol{\sigma})\right]-D\right)+ \\
&\exp(-100\eta Gd)P_{-i}(\varepsilon)E_{-i}\left[\begin{array}{l}\boldsymbol{w}_{\max,i}(\boldsymbol{\sigma})- \\ \dfrac{1}{10}\left|\boldsymbol{w}_{t,i}(\boldsymbol{\sigma})-\boldsymbol{w}_{t+1,i}(\boldsymbol{\sigma})\right|- \\ \dfrac{3\gamma(\boldsymbol{\sigma})}{100Gd}-\beta|\varepsilon\end{array}\right]+ \\
&\exp(-100\eta Gd)\times \\
&P_{-i}(\varepsilon^c)E_{-i}\left[\begin{array}{l}\boldsymbol{w}_{\max,i}(\boldsymbol{\sigma})-\dfrac{1}{10d}\|\boldsymbol{w}_t(\boldsymbol{\sigma})-\boldsymbol{w}_{t+1}(\boldsymbol{\sigma})\|_1- \\ \dfrac{2\gamma(\boldsymbol{\sigma})}{100Gd}-\beta|\varepsilon^c\end{array}\right],
\end{aligned}
\tag{13}
$$

其中式(13)中第 1 个不等式来自于定理 2 和定理 3, 第 2 个不等式来自于ε^c的定义. 由 $P_{-i}(\varepsilon)\leqslant 1$, 可得

$$
\begin{aligned}
&E_{-i}\left[\boldsymbol{w}_{\min,i}(\boldsymbol{\sigma})\right] \geqslant \\
&(1-\exp(-100\eta Gd))\left(E_{-i}\left[\boldsymbol{w}_{\max,i}(\boldsymbol{\sigma})\right]-D\right)+ \\
&\exp(-100\eta Gd)E_{-i}\left[\boldsymbol{w}_{\max,i}(\boldsymbol{\sigma})-\frac{3\gamma(\boldsymbol{\sigma})}{100Gd}-\beta\right]- \\
&\exp(-100\eta Gd)E_{-i}\left[\begin{array}{l}\dfrac{1}{10}\left|\boldsymbol{w}_{t,i}(\boldsymbol{\sigma})-\boldsymbol{w}_{t+1,i}(\boldsymbol{\sigma})\right|+ \\ \dfrac{1}{10d}\|\boldsymbol{w}_t(\boldsymbol{\sigma})-\boldsymbol{w}_{t+1}(\boldsymbol{\sigma})\|_1\end{array}\right] \overset{①}{\geqslant} \\
&E_{-i}\left[\boldsymbol{w}_{\max,i}(\boldsymbol{\sigma})\right]-100\eta GdD-\frac{3\gamma(\boldsymbol{\sigma})}{100Gd}- \\
&\beta-E_{-i}\left[\begin{array}{l}\dfrac{1}{10}\left|\boldsymbol{w}_{t,i}(\boldsymbol{\sigma})-\boldsymbol{w}_{t+1,i}(\boldsymbol{\sigma})\right|+ \\ \dfrac{1}{10d}\|\boldsymbol{w}_t(\boldsymbol{\sigma})-\boldsymbol{w}_{t+1}(\boldsymbol{\sigma})\|_1\end{array}\right],
\end{aligned}
$$

其中①处是由于$\exp(\boldsymbol{w})\geqslant 1+\boldsymbol{w}$. 得

$$
\begin{aligned}
&E_{-i}\left[\left|\boldsymbol{w}_{t,i}(\boldsymbol{\sigma})-\boldsymbol{w}_{t+1,i}(\boldsymbol{\sigma})\right|\right]\leqslant \\
&\frac{1}{9d}E_{-i}\left[\|\boldsymbol{w}_t(\boldsymbol{\sigma})-\boldsymbol{w}_{t+1}(\boldsymbol{\sigma})\|_1\right]+ \\
&\frac{1000}{9}\eta GdD+\frac{E_{-i}[\gamma(\boldsymbol{\sigma})]}{30Gd}+\frac{10}{9}\beta.
\end{aligned}
\tag{14}
$$

由于式(14)对任意$\{\sigma_j\}_{j\neq i}$均成立, 因此可得无条件期望的界:

$$
\begin{aligned}
&E_{-i}\left[\left|\boldsymbol{w}_{t,i}(\boldsymbol{\sigma})-\boldsymbol{w}_{t+1,i}(\boldsymbol{\sigma})\right|\right]\leqslant \\
&\frac{1}{9d}E_{-i}\left[\|\boldsymbol{w}_t(\boldsymbol{\sigma})-\boldsymbol{w}_{t+1}(\boldsymbol{\sigma})\|_1\right]+ \\
&\frac{1000}{9}\eta GdD+\frac{E[\gamma(\boldsymbol{\sigma})]}{30Gd}+\frac{10}{9}\beta.
\end{aligned}
\tag{15}
$$

将式(15)代入到式(11)中, 可得稳定性界:

$$
\begin{aligned}
&E\left[\|\boldsymbol{w}_t(\boldsymbol{\sigma})-\boldsymbol{w}_{t+1}(\boldsymbol{\sigma})\|_1\right]\leqslant \\
&125\eta Gd^2D+\frac{\beta d}{20\eta G}+2\beta d+\frac{\alpha}{20G},
\end{aligned}
\tag{16}
$$

将式(16)代入式(2)中, 得证. 证毕.

由定理 4 知, 若取$\eta=\dfrac{d}{d^{3/2}T^{1/2}-d}$, 可得遗憾界$O\left(d^{3/2}T^{-1/2}+\alpha+\beta d^{3/2}T^{1/2}\right)$. 此外, 若取$\alpha=O\left(T^{-1/2}\right)$, $\beta=O\left(T^{-1}\right)$, 可得遗憾界$O\left(T^{-1/2}\right)$. 关于在线点对学习的理论分析的结论如表 1 所示.

Table 1 Comparison of Online Pairwise Learning Regret Bounds

表 1 在线点对学习遗憾界对比

基本假设	遗憾界
一致有界的强凸损失函数	$O(T^{-1})$[8]
具有最小平方损失函数的 RKHS	$O\left(T^{-1/3}\log T\right)$[10]
正则的 RKHS	$O\left(T^{-1/4}\log T^{1/2}\right)$[12]
具有最小平方损失函数的 RKHS	$O\left(T^{-1/2}\right)$[11]
具有最小平方损失函数的正则 RKHS	$O\left(T^{-2/3}\right)$[13]
具有非凸损失函数的广义在线学习（本文）	$O(T^{-1/2})$

由表 1 可知, 本文对具有非凸损失函数的广义在线点对学习得到了$O(T^{-1/2})$的遗憾界优于已有的遗憾界.

4 结束语

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, 得到广义在线点对学习的遗憾界$O(T^{-1/2})$.

基于在线梯度下降算法, 可进一步研究具有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.

作者贡献声明: 郎璇聪负责研究方案的设计、证明推导, 以及论文的撰写和修改; 李春生指导论文撰写与修改; 刘勇提出论文研究思路、设计研究方案; 王梅指导论文结构设计.

参 考 文 献

[1] Li Zhijie, Li Yuanxiang, Wang Feng, et al. Online learning algorithms for big data analytics: A survey[J]. Journal of Computer Research and Development, 2015, 52(8): 1707−1721 (in Chinese)

(李志杰, 李元香, 王峰, 等. 面向大数据分析的在线学习算法综述[J]. 计算机研究与发展, 2015, 52(8): 1707−1721)

[2] Clémençon S, Lugosi G, Vayatis N. Ranking and empirical minimization of U-statistics[J]. The Annals of Statistics, 2008, 36(2): 844−874

[3] Agarwal S, Niyogi P. Generalization bounds for ranking algorithms via algorithmic stability[J]. Journal of Machine Learning Research, 2009, 10(2): 441−474

[4] Rejchel W. On ranking and generalization bounds[J]. Journal of Machine Learning Research, 2012, 13(5): 1373−1392

[5] Rejchel W. Fast rates for ranking with large families[J]. Neurocomputing, 2015, 168: 1104−1110

[6] Zhao Peilin, Hoi S, Jin Rong, et al. Online AUC maximization[C] //Proc of the 28th Int Conf on Machine Learning. New York: ACM, 2011: 233−240

[7] Weinberger K Q, Saul L K. Distance metric learning for large margin nearest neighbor classification[J]. Journal of Machine Learning Research, 2009, 10(2): 207−244

[8] Wang Yuyang, Khardon R, Pechyony D, et al. Generalization bounds for online learning algorithms with pairwise loss functions[C] //Proc of the 25th Conf on Learning Theory. Berlin: Springer, 2012: 1−22

[9] Ying Yiming, Zhou Dingxuan. Unregularized online learning algorithms with general loss functions[J]. Applied and Computational Harmonic Analysis, 2017, 42(2): 224−244

[10] Ying Yiming, Zhou Dingxuan. Online pairwise learning algorithms[J]. Neural Computation, 2016, 28(4): 743−777

[11] Chen Xiaming, Lei Yunwen. Refined bounds for online pairwise learning algorithms[J]. Neurocomputing, 2018, 275: 2656−2665

[12] Guo Zhengchu, Ying Yiming, Zhou Dingxuan. Online regularized learning with pairwise loss functions[J]. Advances in Computational Mathematics, 2017, 43(1): 127−150

[13] Wang Cheng, Hu Ting. Online regularized pairwise learning with least squares loss[J]. Analysis and Applications, 2020, 18(1): 49−78

[14] Hu Ting, Fan Jun, Xiang Daohong. Convergence analysis of distributed multi-penalty regularized pairwise learning[J]. Analysis and Applications, 2020, 18(1): 109−127

[15] Bousquet O, Elisseeff A. Stability and generalization[J]. The Journal of Machine Learning Research, 2002, 2: 499−526

[16] Elisseeff A, Evgeniou T, Pontil M, et al. Stability of randomized learning algorithms[J]. Journal of Machine Learning Research, 2005, 6(1): 55−79

[17] Shen Wei, Yang Zhenhuan, Ying Yiming, et al. Stability and optimization error of stochastic gradient descent for pairwise learning[J]. Analysis and Applications, 2020, 18(5): 887−927

[18] Lei Yunwen, Ledent A, Kloft M. Sharper generalization bounds for pairwise learning[J]. Advances in Neural Information Processing Systems, 2020, 33: 21236−21246

[19] McMahan B. Follow-the-regularized-leader and mirror descent: Equivalence theorems and L1 regularization[C] //Proc of the 14th Int Conf on Artificial Intelligence and Statistics. Cambridge, MA: MIT , 2011: 525−533

[20] Agarwal N, Gonen A, Hazan E. Learning in non-convex games with an optimization oracle[C] //Proc of the 32nd Conf on Learning Theory. Berlin: Springer, 2019: 18−29

[21] Fine M. Certifiably accurate private data release with generative adversarial networks[J/OL]. 2020[2021-12-18].https://projects.iq.harvard.edu/files/privacytools/files/michael_fines_thesis_-_dp.pdf

[22] Grnarova P, Levy K Y, Lucchi A, et al. An online learning approach to generative adversarial networks[J]. arXiv preprint, arXiv: 1706.03269, 2017

[23] Srebro N, Sridharan K, Tewari A. On the universality of online mirror descent[J/OL]. Advances in Neural Information Processing Systems, 2011[2021-11-25].https://proceedings.neurips.cc/paper/2011/file/a8f8f60264024dca151f164729b76c0b-Paper.pdf

[24] Ying Yiming, Pontil M. Online gradient descent learning algorithms[J]. Foundations of Computational Mathematics, 2008, 8(5): 561−596

[25] Ross S, Bagnell J A. Stability conditions for online learnability[J]. arXiv preprint, arXiv: 1108.3154, 2011

[26] Suggala AS, Netrapalli P. Online non-convex learning: Following the perturbed leader is optimal[C] //Proc of the 31st Algorithmic Learning Theory. Berlin: Springer, 2020: 845−861

Lang Xuancong, born in 1998. PhD candidate. Student member of CCF. Her main research interest includes large-scale machine learning.

郎璇聪, 1998 年生. 博士研究生. CCF 学生会员. 主要研究方向为大规模机器学习.

Li Chunsheng, born in 1960. PhD, professor, PhD supervisor. Member of CCF. His main research interests include data mining and agent systems, multi-agent theory and application, and intelligent architecture.

李春生, 1960 年生. 博士, 教授, 博士生导师, CCF 会员. 主要研究方向为数据挖掘与智能系统、多 agent 理论与应用、智慧架构.

Liu Yong, born in 1986. PhD, accociate professor, PhD supervisor. His main research interests include machine learning, large-scale machine learning, and statistical learning theory.

刘　勇, 1986 年生. 博士, 副教授, 博士生导师. 主要研究方向为机器学习、大规模机器学习、统计学习理论.

Wang Mei, born in 1976. PhD, professor, master supervisor. Member of CCF. Her main research interests include machine learning, kernel methods, and model selection.

王　梅, 1976 年生. 博士, 教授, 硕士生导师. CCF 会员. 主要研究方向为机器学习、核方法、模型选择.