中国科学: 信息科学

SCIENTIA SINICA Informationis

论文

组分布鲁棒优化的随机近似方法

庄镇华[1]

1. 南京大学, 南京 210023
E-mail: zhuangzh@lamda.nju.edu.cn

基金资助

摘要 本文研究了组分布鲁棒优化（GDRO），其目的是学习一个在 m 个不同分布表现良好的模型。预备知识部分介绍了相关假设与基本设定，GDRO 可以形式化为随机凸凹鞍点问题，进而利用随机镜像下降（SMD）算法解决。文献综述部分，第一篇文章提出 SMD(1) 算法，从实验方面表明了组分布鲁棒优化可以提高过参数化模型的鲁棒性，但由于构建的随机梯度方差上界较大，仅能获得次优的样本复杂度 $O(m^2(\log m)/\epsilon^2)$；第二篇文章将组分布鲁棒优化看作是一个双人零和博弈问题，即两个玩家使用在线学习算法迭代地更新他们的参数，并提出了 GDRO-EXP3 算法，使得样本复杂度降低到 $O(m(\log m)/\epsilon^2)$，但由于没有考虑在线学习的非遗忘性质，其理论证明存在依赖性问题；第三篇文章从减小随机梯度的对偶范数角度出发，通过每轮使用 m 个样本，提出了 SMD(m) 算法，将样本复杂度降低到 $O(m(\log m)/\epsilon^2)$，这几乎与该问题的下界 $\Omega(m/\epsilon^2)$ 相匹配。然后，又考虑到一个更实际的问题，即从每个分布中采样的数量是不同的，并提出了一种新的加权组分布鲁棒优化（weighted DRO）情景，可以获得分布相关的收敛速度。用 n_i 表示第 i 个分布的样本预算，并假设 $n_1 \geqslant n_2 \geqslant \cdots \geqslant n_m$。为了解决这个问题，其在 SMD 中引入非均匀抽样，使得样本预算满足期望，并证明了第 i 个分布的超额风险以 $O(\sqrt{n_1 \log m}/n_i)$ 的速率收敛。最后，给出未来展望，即可以将 SA 算法扩展到具有类似形式的问题，例如联邦学习和极小极大遗憾优化等问题。

关键词 在线优化, 组分布鲁棒优化, 随机近似, 镜像下降

1 引言

在经典统计机器学习中，我们的目标是最小化固定分布 $\mathcal{P}_0$ 的风险, 即

$$\min_{\mathbf{w}\in\mathcal{W}} \left\{R_0(\mathbf{w}) = \mathrm{E}_{\mathbf{z}\sim\mathcal{P}_0}\left[\ell(\mathbf{w};\mathbf{z})\right]\right\}, \tag{1}$$

其中 $\mathbf{z}\in\mathcal{Z}$ 是从 $\mathcal{P}_0$ 中抽取的样本, $\mathcal{W}$ 代表假设类, $\ell(\mathbf{w};\mathbf{z})$ 是度量模型 $\mathbf{w}$ 在 $\mathbf{z}$ 上预测误差的损失函数。近年来, 用于优化 (1) 的各种算法相继被提出, 并且可以分为两类：样本平均近似 (SAA) 和随机近似 (SA) 方法. 在 SAA 方法中，我们最小化经验风险，即从 $\mathcal{P}_0$ 中采样的一组样本的平均风险, 而在 SA 方法中, 我们利用目标函数 $R_0(\cdot)$ 的随机观测值直接解决原始问题。

然而，在单一分布上训练的模型可能缺乏鲁棒性，这是因为 (i) 尽管平均损失很低，但其仍可能在少数类中表现糟糕; (ii) 当在另一个分布上测试时，其性能可能会急剧下降。分布式鲁棒优化 (DRO) 提供了一种原理性的方法通过最小化分布 $\mathcal{P}_0$ 邻域最坏情况下的损失来解决这些问题 (Ben-Tal et al., 2013)。最近, 其已经被广泛应用于优化、统计学、运筹学和机器学习等领域。在本文中, 我们考虑最新出现的问题，即组分布鲁棒优化 (GDRO) 问题，其优化有限分布上的最大风险 (Sagawa et al., 2020)。GDRO 可以表述为一个随机极小极大问题:

$$\min_{\mathbf{w}\in\mathcal{W}}\max_{i\in[m]}\ \left\{R_i(\mathbf{w})=\mathrm{E}_{\mathbf{z}\sim\mathcal{P}_i}\big[\ell(\mathbf{w};\mathbf{z})\big]\right\} \tag{2}$$

这里 $\mathcal{P}_1,\ldots,\mathcal{P}_m$ 表示 m 个分布。一个启发性的例子是联邦学习，即在多个客户端部署一个集中式模型，并且每个客户端都面临着不同的数据分布。

由 (Nemirovski et al., 2009, §3.2) 所启发, 我们可以将 (2) 转化为随机凸凹鞍点问题:

$$\min_{\mathbf{w}\in\mathcal{W}}\max_{\mathbf{q}\in\Delta_m}\ \left\{\phi(\mathbf{w},\mathbf{q})=\sum_{i=1}^{m}q_iR_i(\mathbf{w})\right\} \tag{3}$$

其中 $\Delta_m=\{\mathbf{q}\in\mathbb{R}^m:\mathbf{q}\geqslant 0,\sum_{i=1}^m q_i=1\}$ 是 $(m{-}1)$ 维单纯形，然后可以利用随机镜像下降方法（SMD）来求解 (3) 。

2 相关工作

分布鲁棒优化（DRO）源于 (Scarf, 1958) 的开创性工作，并且随着鲁棒优化的发展而得到了广泛的关注。它已经广泛地应用于各种机器学习任务，包括对抗训练 (Sinha et al., 2018)、算法公平性 (Hashimoto et al., 2018)、类别不均衡 (Xu et al., 2020)、长尾分布 (Samuel and Chechik, 2021)、标签移位 (Zhang et al., 2021) 等问题。

一般来说，DRO 反应了我们对目标分布的不确定性。为确保在分布扰动下的良好性能, 它最小化不确定集中最坏分布的风险, 即

$$\min_{\mathbf{w}\in\mathcal{W}}\ \sup_{\mathcal{P}\in\mathcal{S}(\mathcal{P}_0)}\left\{\mathrm{E}_{\mathbf{z}\sim\mathcal{P}}\big[\ell(\mathbf{w};\mathbf{z})\big]\right\} \tag{4}$$

其中 $\mathcal{S}(\mathcal{P}_0)$ 表示 $\mathcal{P}_0$ 周围概率分布的集合。

本文的研究重点是 (2)/(3) 中的 GDRO 问题, 而不是 (4) 中的传统 DRO 问题。(Sagawa et al., 2020) 已经应用了 SMD(Nemirovski et al., 2009) 到 (3)，但由于梯度的方差上界较大，仅获得次优的样本复杂度。在后续工作中, (Haghtalab et al., 2022) 和 (Soma et al., 2022) 已经通过重用样本和应用来自 MAB 的技术来降低样本复杂度，但他们的分析存在依赖性问题。为了处理不同分布的非均匀噪声, (Agarwal and Zhang, 2022) 提出了 GDRO 算法的变种——即极小极大遗憾优化算法 (MRO), 它将风险 $R_i(\mathbf{w})$ 替换为“超额风险“ $R_i(\mathbf{w})-\min_{\mathbf{w}\in\mathcal{W}}R_i(\mathbf{w})$。更一般地，可以在 DRO 中引入校准项，以防止任何单一分布支配最大值 (Słowik and Bottou, 2022)。

3 预备知识

首先，我们重温随机镜像下降（SMD）方法的设定 (Nemirovski et al., 2009)。我们为定义域 $\mathcal{W}$ 配置距离生成函数 $\nu_w(\cdot)$, 其关于范数 $\|\cdot\|_w$ 是 1-强凸的。我们定义与 $\nu_w(\cdot)$ 相关的 Bregman 距离为

$$B_w(\mathbf{u},\mathbf{v}) = \nu_w(\mathbf{u}) - \big[\nu_w(\mathbf{v}) + \langle \nabla \nu_w(\mathbf{v}), \mathbf{u}-\mathbf{v}\rangle\big].$$

对于单纯形 Δ_m, 我们选择熵函数 $\nu_q(\mathbf{q}) = \sum_{i=1}^m q_i \ln q_i$ 作为距离生成函数, 其关于向量 ℓ_1-norm $\|\cdot\|_1$ 是 1-强凸的。类似的, $B_q(\cdot,\cdot)$ 是与 $\nu_q(\cdot)$ 关联的 Bregman 距离。

然后，我们介绍关于定义域和损失函数的一般假设。

假设 1 定义域 $\mathcal{W}$ 是凸的并且根据 $\nu_w(\cdot)$ 计算的直径上界为 D, 即

$$\max_{\mathbf{w}\in\mathcal{W}} \nu_w(\mathbf{w}) - \min_{\mathbf{w}\in\mathcal{W}} \nu_w(\mathbf{w}) \leqslant D^2. \tag{5}$$

对于 Δ_m, 容易验证其由熵函数计算的直径以 $\sqrt{\ln m}$ 为界。为了简化表示，我们假设损失属于 $[0,1]$, 并且其梯度也是有界的。

假设 2 对所有 $i\in[m]$, 我们有

$$0 \leqslant \ell(\mathbf{w};\mathbf{z}) \leqslant 1,\ \forall \mathbf{w}\in\mathcal{W},\ \mathbf{z}\sim\mathcal{P}_i. \tag{6}$$

假设 3 对所有 $i\in[m]$, 我们有

$$\|\nabla_{\mathbf{w}}\ell(\mathbf{w};\mathbf{z})\|_{w,*} \leqslant G,\ \forall \mathbf{w}\in\mathcal{W},\ \mathbf{z}\sim\mathcal{P}_i \tag{7}$$

其中 $\|\cdot\|_{w,*}$ 是 $\|\cdot\|_w$ 的对偶范数。

最后，我们讨论性能度量标准。为了分析收敛性，我们通过如下误差来衡量 (3) 的近似解 $(\bar{\mathbf{w}},\bar{\mathbf{q}})$ 的质量

$$\epsilon_\phi(\bar{\mathbf{w}},\bar{\mathbf{q}}) = \max_{\mathbf{q}\in\Delta_m} \phi(\bar{\mathbf{w}},\mathbf{q}) - \min_{\mathbf{w}\in\mathcal{W}} \phi(\mathbf{w},\bar{\mathbf{q}}) \tag{8}$$

它可以支配 $\bar{\mathbf{w}}$ 对于原始问题 (2) 的最优性, 因为

$$\begin{aligned}&\max_{i\in[m]} R_i(\bar{\mathbf{w}}) - \min_{\mathbf{w}\in\mathcal{W}}\max_{i\in[m]} R_i(\mathbf{w}) = \max_{\mathbf{q}\in\Delta_m}\sum_{i=1}^m q_i R_i(\bar{\mathbf{w}}) - \min_{\mathbf{w}\in\mathcal{W}}\max_{\mathbf{q}\in\Delta_m}\sum_{i=1}^m q_i R_i(\mathbf{w})\\ \leqslant&\max_{\mathbf{q}\in\Delta_m}\sum_{i=1}^m q_i R_i(\bar{\mathbf{w}}) - \min_{\mathbf{w}\in\mathcal{W}}\sum_{i=1}^m \bar{q}_i R_i(\mathbf{w}) = \epsilon_\phi(\bar{\mathbf{w}},\bar{\mathbf{q}}).\end{aligned} \tag{9}$$

4 文献综述

4.1 针对组偏移的分布鲁棒神经网络 (Sagawa et al., 2020)

4.1.1 主要内容：GDRO 的在线优化方法

本文是第一篇应用 SMD(Nemirovski et al., 2009) 到 (3) 的工作，为了应用 SMD，关键是构造 (3) 中函数 $\phi(\mathbf{w},\mathbf{q})$ 的随机梯度。我们首先给出它关于 $\mathbf{w}$ 和 $\mathbf{q}$ 的真实梯度：

$$\nabla_{\mathbf{w}}\phi(\mathbf{w},\mathbf{q}) = \sum_{i=1}^m q_i \nabla R_i(\mathbf{w}),\ \text{and}\ \nabla_{\mathbf{q}}\phi(\mathbf{w},\mathbf{q}) = [R_1(\mathbf{w}),\ldots,R_m(\mathbf{w})]^\top. \tag{10}$$

在第 t 轮中, 用 $\mathbf{w}_t$ 和 $\mathbf{q}_t$ 表示当前解。我们首先从 $[m]$ 中均匀选择一个索引 i_t，然后从分布 $\mathcal{P}_i$ 中抽取一个样本 $\mathbf{z}_t^{(i_t)}$, 并将随机梯度定义为

$$\mathbf{g}_w(\mathbf{w}_t,\mathbf{q}_t)=q_{t,i_t}m\nabla\ell(\mathbf{w}_t;\mathbf{z}_t^{(i_t)}),\ \text{and}\ \mathbf{g}_q(\mathbf{w}_t,\mathbf{q}_t)=[0,\dots,m\ell(\mathbf{w}_t;\mathbf{z}_t^{(i_t)}),\dots,0]^\top \tag{11}$$

显然，它们是真实梯度的无偏估计:

$$\mathrm{E}_{t-1}[\mathbf{g}_w(\mathbf{w}_t,\mathbf{q}_t)]=\nabla_{\mathbf{w}}\phi(\mathbf{w}_t,\mathbf{q}_t),\ \text{and}\ \mathrm{E}_{t-1}[\mathbf{g}_q(\mathbf{w}_t,\mathbf{q}_t)]=\nabla_{\mathbf{q}}\phi(\mathbf{w}_t,\mathbf{q}_t)$$

其中 $\mathrm{E}_{t-1}[\cdot]$ 表示直到第 $t-1$ 轮以随机性为条件的期望值。

然后，我们应用 SMD 来更新 $\mathbf{w}_t$ 和 $\mathbf{q}_t$:

$$\mathbf{w}_{t+1}=\operatorname*{argmin}_{\mathbf{w}\in\mathcal{W}}\left\{\eta_w\langle\mathbf{g}_w(\mathbf{w}_t,\mathbf{q}_t),\mathbf{w}-\mathbf{w}_t\rangle+B_w(\mathbf{w},\mathbf{w}_t)\right\}, \tag{12}$$

$$\mathbf{q}_{t+1}=\operatorname*{argmin}_{\mathbf{q}\in\Delta_m}\left\{\eta_q\langle-\mathbf{g}_q(\mathbf{w}_t,\mathbf{q}_t),\mathbf{q}-\mathbf{q}_t\rangle+B_q(\mathbf{q},\mathbf{q}_t)\right\} \tag{13}$$

其中 $\eta_w>0$ 和 $\eta_q>0$ 是两个步长。$\mathbf{w}_t$ 的更新规则取决于距离生成函数 $\nu_w(\cdot)$ 的选择。例如, 如果 $\nu_w(\mathbf{w})=\frac{1}{2}\|\mathbf{w}\|_2^2$, 则 (12) 变为随机梯度下降（SGD)，即

$$\mathbf{w}_{t+1}=\Pi_{\mathcal{W}}\left[\mathbf{w}_t-\eta_w\mathbf{g}_w(\mathbf{w}_t,\mathbf{q}_t)\right]$$

其中 $\Pi_{\mathcal{W}}[\cdot]$ 表示到 $\mathcal{W}$ 中最近点的欧式投影。由于 $B_q(\mathbf{q},\mathbf{q}_t)$ 是根据负熵定义的，因此 (13) 等价于

$$q_{t+1,i}=\frac{q_{t,i}\exp\left(\eta_q\cdot\mathbf{g}_q(\mathbf{w}_t,\mathbf{q}_t)_i\right)}{\sum_{j=1}^m q_{t,j}\exp\left(\eta_q\cdot\mathbf{g}_q(\mathbf{w}_t,\mathbf{q}_t)_j\right)},\ \forall i\in[m] \tag{14}$$

即为应用于最大化问题的 Hedge 算法 (Freund and Schapire, 1997)。一开始, 我们设定 $\mathbf{w}_1=\operatorname{argmin}_{\mathbf{w}\in\mathcal{W}}\nu_w(\mathbf{w})$, $\mathbf{q}_1=\frac{1}{m}\mathbf{1}_m$, 其中 $\mathbf{1}_m$ 是由 1 组成的 m 维向量。在最后一步，我们返回平均迭代 $\bar{\mathbf{w}}=\frac{1}{T}\sum_{t=1}^T\mathbf{w}_t$ 和 $\bar{\mathbf{q}}=\frac{1}{T}\sum_{t=1}^T\mathbf{q}_t$ 作为最终解。整体的程序流程在算法 1中。

算法 1 GDRO 的在线优化算法

输入: 两个步长: η_w 和 η_q

初始化 $\mathbf{w}_1=\operatorname{argmin}_{\mathbf{w}\in\mathcal{W}}\nu_w(\mathbf{w})$, $\mathbf{q}_1=[1/m,\dots,1/m]^\top\in\mathbb{R}^m$
for $t=1$ to T **do**
  $i_t\sim\text{Uniform}(1,\dots,m)$
  从分布 $\mathcal{P}_{i_t}$ 中抽取一个样本 $\mathbf{z}_t^{(i_t)}$
  构造公式 (11) 中定义的随机梯度
  分别根据 (12) 和 (13) 更新 $\mathbf{w}_t$ 和 $\mathbf{q}_t$
end for
return $\bar{\mathbf{w}}=\frac{1}{T}\sum_{t=1}^T\mathbf{w}_t$ 和 $\bar{\mathbf{q}}=\frac{1}{T}\sum_{t=1}^T\mathbf{q}_t$

基于随机凸凹优化的 SMD 理论保证 (Nemirovski et al., 2009, §3.1), 对于算法 1，我们有如下定理:

定理 1 基于假设 1, 2 和 3, 并且设定算法 1中的 $\eta_w=\frac{D^2}{m}\sqrt{\frac{8}{5T(D^2G^2+\ln m)}}$, $\eta_q=\frac{\ln m}{m}\sqrt{\frac{8}{5T(D^2G^2+\ln m)}}$, 我们有

$$\mathrm{E}\left[\epsilon_\phi(\bar{\mathbf{w}},\bar{\mathbf{q}})\right]\leqslant 2m\sqrt{\frac{10(D^2G^2+\ln m)}{T}}$$

并且以至少 $1-\delta$ 的概率有,

$$\epsilon_{\phi}(\bar{\mathbf{w}},\bar{\mathbf{q}}) \leqslant 2m\left(\sqrt{10}+2\sqrt{2\ln 1/\delta}\right)\sqrt{\frac{D^2G^2+\ln m}{T}}.$$

注释 1 定理 1 表明算法 1达到 $O(m\sqrt{(\log m)/T})$ 的收敛速度。由于每次迭代消耗 1 个样本, 因此样本复杂度为 $O(m^2(\log m)/\epsilon^2)$。

4.1.2 文章点评

本文是第一篇应用 SMD 到组分布鲁棒优化的工作, 从实验方面表明了组分布鲁棒优化可以防止模型学习预设定的虚假相关性, 提高过参数化模型的鲁棒性; 但其缺点也比较明显, 即理论分析中由于构建的随机梯度方差上界较大, 仅能获得次优的样本复杂度 $O(m^2(\log m)/\epsilon^2)$。

4.2 广义组分布鲁棒优化的最优算法 (Soma et al., 2022)

4.2.1 主要内容: GDRO 的双人零和博弈在线优化方法

本文将 (3) 看作是一个双人零和博弈问题, 即一个玩家选择 $\mathbf{w}\in\mathcal{W}$, 另一个玩家选择 $\mathbf{q}\in\Delta_m$。两个玩家使用在线学习算法迭代地更新他们的参数; 特别地, 本文分别对 $\mathbf{w}$ 玩家和 $\mathbf{q}$ 玩家使用在线梯度下降 (OGD) 和在线镜像下降 (OMD) 算法。为了达到最佳的收敛速度, 我们必须仔细选择正则化方法以及随机梯度的构建方式, 尤其需要平衡随机梯度的方差以及 OGD、OMD 中定义域的直径。本文基于对抗性多臂赌博机算法 EXP3(Auer et al., 2002) 提出其 GDRO 算法 (GDRO-EXP3)。

在本节中, 我们将描述 GDRO-EXP3 算法。设 $\phi(\mathbf{w},\mathbf{q})=\sum_{i=1}^m q_i\mathrm{E}_{\mathbf{z}\sim\mathcal{P}_i}\left[\ell(\mathbf{w};\mathbf{z})\right]$, 假设 $\mathbf{w}$ 玩家和 $\mathbf{q}$ 玩家分别运行 OGD 和 OMD 在线算法, 以解决极小极大问题 $\min_{\mathbf{w}\in\mathcal{W}}\max_{\mathbf{q}\in\Delta_m}\phi(\mathbf{w},\mathbf{q})$。在第 t 轮中, 用 $\mathbf{w}_t$ 和 $\mathbf{q}_t$ 表示当前解。我们首先从 $[m]$ 中选择一个索引 $i_t\sim\mathbf{q}_t$, 然后从分布 $\mathcal{P}_{i_t}$ 中抽取一个样本 $\mathbf{z}_t^{(i_t)}$, 并将随机梯度定义为

$$\mathbf{g}_w(\mathbf{w}_t,\mathbf{q}_t)=\nabla\ell(\mathbf{w}_t;\mathbf{z}_t^{(i_t)}),\ \text{and}\ \mathbf{g}_q(\mathbf{w}_t,\mathbf{q}_t)=[0,\dots,\frac{\ell(\mathbf{w}_t;\mathbf{z}_t^{(i_t)})}{q_{t,i_t}},\dots,0]^\top \tag{15}$$

显然, 它们是真实梯度的无偏估计。然后, 我们分别应用 OGD 和 SMD 来更新 $\mathbf{w}_t$ 和 $\mathbf{q}_t$:

$$\mathbf{w}_{t+1}=\Pi_{\mathcal{W}}\left[\mathbf{w}_t-\eta_w\mathbf{g}_w(\mathbf{w}_t,\mathbf{q}_t)\right] \tag{16}$$

$$\mathbf{q}_{t+1}=\underset{\mathbf{q}\in\Delta_m}{\operatorname{argmin}}\left\{\eta_q\langle-\mathbf{g}_q(\mathbf{w}_t,\mathbf{q}_t),\mathbf{q}-\mathbf{q}_t\rangle+B_q(\mathbf{q},\mathbf{q}_t)\right\} \tag{17}$$

其中 $\eta_w>0$ 和 $\eta_q>0$ 是两个步长。由于 $B_q(\mathbf{q},\mathbf{q}_t)$ 是根据负熵定义的, 相当于 $\mathbf{q}$ 玩家运行 EXP3 算法, 因此 (17) 等价于 (14)。在最后一步, 我们返回平均迭代 $\bar{\mathbf{w}}=\frac{1}{T}\sum_{t=1}^T\mathbf{w}_t$ 作为最终解。整体的程序流程在算法 2中。

令

$$\epsilon_{\phi}(\bar{\mathbf{w}})=\max_{\mathbf{q}\in\Delta_m}\phi(\bar{\mathbf{w}},\mathbf{q})-\min_{\mathbf{w}\in\mathcal{W}}\max_{\mathbf{q}\in\Delta_m}\phi(\mathbf{w},\mathbf{q})$$

$$R_w(T)=\sum_{t=1}^T\phi(\mathbf{w}_t,\mathbf{q}_t)-\min_{\mathbf{w}\in\mathcal{W}}\sum_{t=1}^T\phi(\mathbf{w},\mathbf{q}_t)$$

算法 2 GDRO-EXP3 算法

输入: 两个步长: η_w 和 η_q

1: 初始化 $\mathbf{w}_1 = \operatorname{argmin}_{\mathbf{w}\in\mathcal{W}} \nu_w(\mathbf{w})$, $\mathbf{q}_1 = [1/m, \ldots, 1/m]^\top \in \mathbb{R}^m$
2: **for** $t = 1$ to T **do**
3: $\quad i_t \sim \mathbf{q}_t$
4: $\quad$ 从分布 $\mathcal{P}_{i_t}$ 中抽取一个样本 $\mathbf{z}_t^{(i_t)}$
5: $\quad$ 构造公式 (15) 中定义的随机梯度
6: $\quad$ 分别根据 (16) 和 (17) 更新 $\mathbf{w}_t$ 和 $\mathbf{q}_t$
7: **end for**
8: **return** $\bar{\mathbf{w}} = \frac{1}{T}\sum_{t=1}^{T} \mathbf{w}_t$

$$R_q(T) = \max_{\mathbf{q}\in\Delta_m} \sum_{t=1}^{T} \phi(\mathbf{w}_t, \mathbf{q}) - \sum_{t=1}^{T} \phi(\mathbf{w}_t, q_t)$$

然后，根据遗憾的定义和 Jensen 不等式，我们有

$$\begin{aligned}
\epsilon_\phi(\bar{\mathbf{w}}) &\leqslant \max_{\mathbf{q}\in\Delta_m} \frac{1}{T}\sum_{t=1}^{T} \phi(\mathbf{w}_t, \mathbf{q}) - \min_{\mathbf{w}\in\mathcal{W}} \max_{\mathbf{q}\in\Delta_m} \phi(\mathbf{w}, \mathbf{q}) \\
&= \frac{R_q(T)}{T} + \frac{1}{T}\sum_{t=1}^{T} \phi(\mathbf{w}_t, \mathbf{q}_t) - \min_{\mathbf{w}\in\mathcal{W}} \max_{\mathbf{q}\in\Delta_m} \phi(\mathbf{w}, \mathbf{q}) \\
&\leqslant \frac{R_q(T)}{T} + \frac{1}{T}\sum_{t=1}^{T} \phi(\mathbf{w}_t, \mathbf{q}_t) - \min_{\mathbf{w}\in\mathcal{W}} \frac{1}{T}\sum_{t=1}^{T} \phi(\mathbf{w}, \mathbf{q}_t) \\
&= \frac{R_q(T) + R_w(T)}{T}.
\end{aligned}$$

即我们可以通过在线算法的误差 R_w 和 R_q 来约束期望 $E[\epsilon_\phi(\bar{\mathbf{w}})]$ 的收敛速度

$$\mathrm{E}[\epsilon_\phi(\bar{\mathbf{w}})] \leqslant \frac{\mathrm{E}[R_{\mathbf{w}}(T)] + \mathrm{E}[R_{\mathbf{q}}(T)]}{T}. \tag{18}$$

对于算法 2，我们有如下定理：

定理 2 基于假设 1, 2 和 3, 并且设定算法 2中的 $\eta_w = \frac{D}{G\sqrt{T}}$, $\eta_q = \sqrt{\frac{2\log m}{mM^2T}}$ ，我们有

$$\mathrm{E}[\epsilon_\phi(\bar{\mathbf{w}})] \leqslant \sqrt{2}\frac{\sqrt{G^2D^2 + 2M^2 m\log m}}{\sqrt{T}}.$$

注释 2 定理 2 表明算法 2达到 $O(\sqrt{(m\log m)/T})$ 的收敛速度。由于每次迭代消耗 1 个样本，因此样本复杂度为 $O(m(\log m)/\epsilon^2)$。

4.2.2 文章点评

本文将 (3) 看作是一个双人零和博弈问题，即一个玩家选择 $\mathbf{w}\in\mathcal{W}$，另一个玩家选择 $\mathbf{q}\in\Delta_m$，两个玩家使用在线学习算法迭代地更新他们的参数，并基于 EXP3 提出了 GDRO-EXP3 算法，使得样本复杂度降低到 $O(m(\log m)/\epsilon^2)$；但其缺点也比较明显，即没有考虑非遗忘性质，因此基于遗忘在线学习建立的理论保证无法证明其对于 (3) 的最优性。具体来说，该结果意味着，对于任何独

立于 $\bar{\mathbf{w}}$ 和 $\bar{\mathbf{q}}$ 的固定 $\mathbf{w}$ 和 $\mathbf{q}$，有

$$\mathrm{E}\left[\phi(\bar{\mathbf{w}},\mathbf{q})-\phi(\mathbf{w},\bar{\mathbf{q}})\right]=O\left(\sqrt{\frac{m\log m}{T}}\right). \tag{19}$$

然而，由于依赖性问题，(19) 不能用于约束 (8) 中的 $\epsilon_\phi(\bar{\mathbf{w}},\bar{\mathbf{q}})$。更明确地说，我们有

$$\epsilon_\phi(\bar{\mathbf{w}},\bar{\mathbf{q}})=\max_{\mathbf{q}\in\Delta_m}\phi(\bar{\mathbf{w}},\mathbf{q})-\min_{\mathbf{w}\in\mathcal{W}}\phi(\mathbf{w},\bar{\mathbf{q}})=\phi(\bar{\mathbf{w}},\widehat{\mathbf{q}})-\phi(\widehat{\mathbf{w}},\bar{\mathbf{q}}),$$

其中 $\widehat{\mathbf{w}}=\operatorname{argmin}_{\mathbf{w}\in\mathcal{W}}\phi(\mathbf{w},\bar{\mathbf{q}})$ 并且 $\widehat{\mathbf{q}}=\operatorname{argmax}_{\mathbf{q}\in\Delta_m}\phi(\bar{\mathbf{w}},\mathbf{q})$ 依赖于 $\bar{\mathbf{w}}$ 和 $\bar{\mathbf{q}}$.

4.3 组分布鲁棒优化的随机近似（SA）方法 (Zhang et al., 2023)

4.3.1 主要内容：组分布鲁棒优化的随机近似方法

GDRO 的随机镜像下降方法

接下来，我们采用另一种方式构造 (3) 中函数 $\phi(\mathbf{w},\mathbf{q})$ 的随机梯度，以降低样本复杂度。在第 t 轮中, 用 $\mathbf{w}_t$ 和 $\mathbf{q}_t$ 表示当前解。我们从每个分布 $\mathcal{P}_i$ 中抽取一个样本 $\mathbf{z}_t^{(i)}$, 并将随机梯度定义为

$$\mathbf{g}_w(\mathbf{w}_t,\mathbf{q}_t)=\sum_{i=1}^m q_{t,i}\nabla\ell(\mathbf{w}_t;\mathbf{z}_t^{(i)}),\ \text{and}\ \mathbf{g}_q(\mathbf{w}_t,\mathbf{q}_t)=[\ell(\mathbf{w}_t;\mathbf{z}_t^{(1)}),\ldots,\ell(\mathbf{w}_t;\mathbf{z}_t^{(m)})]^\top. \tag{20}$$

显然，它们也是真实梯度的无偏估计：

$$\mathrm{E}_{t-1}[\mathbf{g}_w(\mathbf{w}_t,\mathbf{q}_t)]=\nabla_{\mathbf{w}}\phi(\mathbf{w}_t,\mathbf{q}_t),\ \text{and}\ \mathrm{E}_{t-1}[\mathbf{g}_q(\mathbf{w}_t,\mathbf{q}_t)]=\nabla_{\mathbf{q}}\phi(\mathbf{w}_t,\mathbf{q}_t)$$

其中 $\mathrm{E}_{t-1}[\cdot]$ 表示直到第 $t-1$ 轮以随机性为条件的期望值。

然后,我们根据 (12) 和 (13) 应用 SMD 来更新 $\mathbf{w}_t$ 和 $\mathbf{q}_t$。初始时, 我们仍设定 $\mathbf{w}_1=\operatorname{argmin}_{\mathbf{w}\in\mathcal{W}}\nu_w(\mathbf{w})$, $\mathbf{q}_1=\frac{1}{m}\mathbf{1}_m$。在最后一步，我们返回平均迭代 $\bar{\mathbf{w}}=\frac{1}{T}\sum_{t=1}^T\mathbf{w}_t$ 和 $\bar{\mathbf{q}}=\frac{1}{T}\sum_{t=1}^T\mathbf{q}_t$ 作为最终解。整体的程序流程在算法 3中。

算法 3 GDRO 的随机镜像下降算法

输入: 两个步长: η_w 和 η_q

1: 初始化 $\mathbf{w}_1=\operatorname{argmin}_{\mathbf{w}\in\mathcal{W}}\nu_w(\mathbf{w})$, $\mathbf{q}_1=[1/m,\ldots,1/m]^\top\in\mathbb{R}^m$
2: **for** $t=1$ to T **do**
3: 　对于每个 $i\in[m]$, 从分布 $\mathcal{P}_i$ 中抽取一个样本 $\mathbf{z}_t^{(i)}$
4: 　构造公式 (20) 中定义的随机梯度
5: 　分别根据 (12) 和 (13) 更新 $\mathbf{w}_t$ 和 $\mathbf{q}_t$
6: **end for**
7: **return** $\bar{\mathbf{w}}=\frac{1}{T}\sum_{t=1}^T\mathbf{w}_t$ 和 $\bar{\mathbf{q}}=\frac{1}{T}\sum_{t=1}^T\mathbf{q}_t$

类似的, 对于算法 3，我们有如下定理：

定理 3 基于假设 1, 2 和 3, 并且设定算法 3中的 $\eta_w=D^2\sqrt{\frac{8}{5T(D^2G^2+\ln m)}}$, $\eta_q=(\ln m)\sqrt{\frac{8}{5T(D^2G^2+\ln m)}}$, 我们有

$$\mathrm{E}\big[\epsilon_\phi(\bar{\mathbf{w}},\bar{\mathbf{q}})\big]\leqslant 2\sqrt{\frac{10(D^2G^2+\ln m)}{T}}$$

并且以至少 $1-\delta$ 的概率有，

$$\epsilon_\phi(\bar{\mathbf{w}}, \bar{\mathbf{q}}) \leqslant 2\left(\sqrt{10} + 2\sqrt{2\ln 1/\delta}\right)\sqrt{\frac{D^2G^2 + \ln m}{T}}.$$

注释 3　定理 3 表明算法 3达到 $O(\sqrt{(\log m)/T})$ 的收敛速度。由于每次迭代消耗 m 个样本，因此样本复杂度为 $O(m(\log m)/\epsilon^2)$，几乎与下界 $\Omega(m/\epsilon^2)$ 匹配 (Soma et al., 2022, Theorem 5)。

注释 4　算法 1的收敛速度比算法 3慢，并且其随机梯度 (11) 的对偶范数是算法 3的 m 倍，导致其样本复杂度是算法 3的 m 倍，为 $O(m^2(\log m)/\epsilon^2)$ 。

4.3.2 主要内容：加权组分布鲁棒优化（Weighted GDRO）和随机近似方法

正如我们在第 4.3.1节中所做的那样，在为 GDRO 问题设计 SA 方法时，通常假设算法可以自由地从每个分布中抽取样本。然而，由于从不同分布收集数据的成本不同，这种假设在实践中可能不成立 (Radivojac et al., 2004)。在本节中，我们将研究从每个分布中采样的样本数量可能不同的情况，用 n_i 表示可以从分布 $\mathcal{P}_i$ 中采样的数量，不失一般性，我们假设 $n_1 \geqslant n_2 \geqslant \cdots \geqslant n_m$。注意，我们有一个简单的**基线**，它只运行 n_m 次算法 3，并且优化误差为 $\epsilon_\phi(\bar{\mathbf{w}}, \bar{\mathbf{q}}) = O(\sqrt{(\log m)/n_m})$。

非均匀采样的随机镜像下降算法

为了满足预算，我们建议将非均匀采样纳入 SMD 算法。在进入技术细节之前，我们首先解释使用非均匀采样的主要思想。一种方法是在每次迭代中以 $p_i = n_i/n_1$ 的概率从每个分布 $\mathcal{P}_i$ 中抽取 1 个样本。那么，在 n_1 次迭代之后，从 $\mathcal{P}_i$ 中采样的期望数目将是 $n_1 p_i = n_i$，因此预算在期望意义上满足。

具体地，在第 t 轮中，我们首先生成一组伯努利随机变量 $\{b_t^{(1)}, \ldots, b_t^{(m)}\}$，其中 $\Pr[b_t^{(i)} = 1] = p_i$，用于确定是否从第 i 分布采样。如果 $b_t^{(i)} = 1$，我们从 $\mathcal{P}_i$ 中提取样本 $\mathbf{z}_t^{(i)}$。现在，问题变成如何利用这些样本构建随机梯度，令 $\mathcal{C}_t = \{i | b_t^{(i)} = 1\}$ 表示所选分布的索引集合。如果我们求解 (3) 中的原始问题，那么随机梯度应按以下方式构造以确保无偏性。

$$\mathbf{g}_w(\mathbf{w}_t, \mathbf{q}_t) = \sum_{i \in C_t} \frac{q_{t,i}}{p_i} \nabla \ell(\mathbf{w}_t; \mathbf{z}_t^{(i)}), \text{ and } [\mathbf{g}_q(\mathbf{w}_t, \mathbf{q}_t)]_i = \begin{cases} \ell(\mathbf{w}_t; \mathbf{z}_t^{(i)})/p_i, \ i \in \mathcal{C}_t \\ 0, \qquad\qquad \text{otherwise} \end{cases} \tag{21}$$

然后，可以利用 SMD 来更新 $\mathbf{w}_t$ 和 $\mathbf{q}_t$。为了分析优化误差，我们需要限制 (21) 中随机梯度的范数。为此，我们有 $\|\mathbf{g}_w(\mathbf{w}_t, \mathbf{q}_t)\|_{w,*} \leqslant G n_1/n_m$ 和 $\|\mathbf{g}_q(\mathbf{w}_t, \mathbf{q}_t)_i\|_\infty \leqslant n_1/n_m$。根据定理 3的论证，我们可以证明误差为 $\epsilon_\phi(\bar{\mathbf{w}}, \bar{\mathbf{q}}) = O(\sqrt{(\log m)/n_1} \cdot n_1/n_m) = O(\sqrt{n_1 \log m}/n_m)$，甚至大于基线的 $O(\sqrt{(\log m)/n_m})$ 误差。

在下文中，我们证明了上述过程的简单转化仍可以产生与基线互补的有意义的结果。我们观察到 (21) 中随机梯度的范数较大是由概率的倒数 $1/p_i$ 引起的。一个自然的想法是忽略 $1/p_i$，并定义以下随机梯度：

$$\mathbf{g}_w(\mathbf{w}_t, \mathbf{q}_t) = \sum_{i \in C_t} q_{t,i} \nabla \ell(\mathbf{w}_t; \mathbf{z}_t^{(i)}), \text{ and } [\mathbf{g}_q(\mathbf{w}_t, \mathbf{q}_t)]_i = \begin{cases} \ell(\mathbf{w}_t; \mathbf{z}_t^{(i)}), \ i \in \mathcal{C}_t \\ 0, \qquad\qquad \text{otherwise.} \end{cases} \tag{22}$$

通过这种方式，它们不再是 (3) 的随机梯度，而是可以被视为加权 GDRO 问题的随机梯度:

$$\min_{\mathbf{w}\in\mathcal{W}}\max_{\mathbf{q}\in\Delta_m}\left\{\varphi(\mathbf{w},\mathbf{q})=\sum_{i=1}^{m}q_ip_iR_i(\mathbf{w})\right\} \tag{23}$$

其中每个风险 $R_i(\cdot)$ 通过因子 p_i 缩放。基于 (22) 中的梯度，我们仍然使用 (12) 和 (13) 来更新 $\mathbf{w}_t$ 和 $\mathbf{q}_t$，并在算法 4中总结了完整的过程。

算法 4 加权 GDRO 的随机镜像下降算法

输入: 两个步长: η_w 和 η_q

1: 初始化 $\mathbf{w}_1=\mathrm{argmin}_{\mathbf{w}\in\mathcal{W}}\,\nu_w(\mathbf{w})$, $\mathbf{q}_1=[1/m,\ldots,1/m]^\top\in\mathbb{R}^m$
2: **for** $t=1$ to n_1 **do**
3: 对于每个 $i\in[m]$, 生成一个伯努利随机变量 $b_t^{(i)}$ ，其中 $\Pr[b_t^{(i)}=1]=p_i$, 如果 $b_t^{(i)}=1$, 则从分布 $\mathcal{P}_i$ 中抽取一个样本 $\mathbf{z}_t^{(i)}$
4: 构造 (22) 中定义的随机梯度
5: 分别根据 (12) 和 (13) 更新 $\mathbf{w}_t$ 和 $\mathbf{q}_t$
6: **end for**
7: **return** $\bar{\mathbf{w}}=\frac{1}{n_1}\sum_{t=1}^{n_1}\mathbf{w}_t$ 和 $\bar{\mathbf{q}}=\frac{1}{n_1}\sum_{t=1}^{n_1}\mathbf{q}_t$

我们省略算法 4对于 (23) 的优化误差，因为它具有与定理 3完全相同的形式。我们真正感兴趣的是它在多重分布上解的理论保证。为此，我们有以下定理:

定理 4 基于假设 1, 2 和 3, 并且设定算法 4中的 $\eta_w=D^2\sqrt{\frac{8}{5n_1(D^2G^2+\ln m)}}$, $\eta_q=(\ln m)\sqrt{\frac{8}{5n_1(D^2G^2+\ln m)}}$, 则以至少 $1-\delta$ 的概率, 有

$$R_i(\bar{\mathbf{w}})-\frac{n_1}{n_i}p_\varphi^*\leqslant\frac{1}{p_i}\mu(\delta)\sqrt{\frac{D^2G^2+\ln m}{n_1}}=\mu(\delta)\frac{\sqrt{(D^2G^2+\ln m)n_1}}{n_i},\ \forall i\in[m]$$

其中 p_φ^* 是式 (23) 的最优值，并且 $\mu(\delta)=2\left(\sqrt{10}+2\sqrt{2\ln 1/\delta}\right)$.

注释 5 我们观察到算法 4 表现出与分布相关的收敛性: 样本数 n_i 越大，目标风险 $n_1p_\varphi^*/n_i$ 越小，收敛速度 $O(\sqrt{n_1\log m}/n_i)$ 越快。注意，它的收敛率总是优于使用 (21) 作为梯度的 SMD 方法 $O(\sqrt{n_1\log m}/n_m)$ 的收敛率。此外，当 $n_i\geqslant\sqrt{n_1n_m}$ 时，它比基线收敛得更快。特别的，对于分布 $\mathcal{P}_1$，算法 4获得了 $O(\sqrt{(\log m)/n_1})$ 的收敛速率，几乎和从单个分布学习的最优 $O(\sqrt{1/n_1})$ 收敛速率相匹配。我们注意到，很难描述 p_φ^* 的阶，以前的研究通常将这些量视为极小值甚至为 0(Agarwal and Zhang, 2022, Corollary 9)。

4.3.3 文章点评

对于 GDRO 设定，本文从减小随机梯度的对偶范数出发，通过每轮使用 m 个样本，提出了算法3，将样本复杂度降低到 $O(m(\log m)/\epsilon^2)$，并且不存在算法2中的依赖性问题；并且又基于在线不平衡数据提出了 Weighted GDRO 设定，进而提出算法4，当 $n_i\geqslant\sqrt{n_1n_m}$ 时，它比基线算法1收敛得更快。特别的，对于分布 $\mathcal{P}_1$ 算法 4获得了 $O(\sqrt{(\log m)/n_1})$ 的收敛速率，几乎和从单个分布学习的最优 $O(\sqrt{1/n_1})$ 收敛速率相匹配。本文思路简洁，并且也得到了目前最优的理论结果，基本没有缺点。

5 定理证明

在这一节中，我们给出主要定理的证明。

5.1 随机镜像下降算法定理的证明

定理 1和 3的证明基于 (Nemirovski et al., 2009) 的 §2.3 节——随机镜像下降算法，因此为了完整性，我们首先证明 §2.3 节中的主要结果。

算法 5 随机镜像下降算法

输入: 步长: η

初始化 $\mathbf{w}_1 = \operatorname{argmin}_{\mathbf{w}\in\mathcal{W}} \nu_w(\mathbf{w})$
for $t = 1$ to T **do**
  从分布 $\mathcal{P}$ 中抽取一个样本 $\mathbf{z}_t$
  构造随机梯度 $\nabla\ell(\mathbf{w}_t;\mathbf{z}_t)$
  根据下式更新 $\mathbf{w}_t$

$$\mathbf{w}_{t+1} = \operatorname*{argmin}_{\mathbf{w}\in\mathcal{W}} \{\eta\langle\nabla\ell(\mathbf{w}_t;\mathbf{z}_t), \mathbf{w}-\mathbf{w}_t\rangle + B_w(\mathbf{w},\mathbf{w}_t)\}$$

end for
return $\bar{\mathbf{w}} = \frac{1}{T}\sum_{t=1}^{T}\mathbf{w}_t$

定理 5 基于假设 1, 2和 3, 在算法 5中设定 $\eta = \frac{\sqrt{2}D}{G\sqrt{T}}$, 我们有

$$\mathrm{E}\left[R(\bar{\mathbf{w}}) - R^*\right] \leqslant DG\sqrt{\frac{2}{T}}. \tag{24}$$

此外，以至少 $1-\delta$ 的概率有,

$$R(\bar{\mathbf{w}}) - R^* \leqslant DG\sqrt{\frac{2}{T}}\left(1 + 4\sqrt{\ln\frac{1}{\delta}}\right). \tag{25}$$

证明: 设 $\mathbf{w}_* \in \operatorname{argmin}_{\mathbf{w}\in\mathcal{W}} R(\mathbf{w})$ 是最小化 $R(\cdot)$ 的最优解。根据镜像下降的性质，即Nemirovski et al. (2009) 的引理 2.1，我们有

$$\begin{aligned}
&\eta\langle\nabla\ell(\mathbf{w}_t;\mathbf{z}_t), \mathbf{w}_t - \mathbf{w}_*\rangle\\
\leqslant& B_w(\mathbf{w}_*,\mathbf{w}_t) - B_w(\mathbf{w}_*,\mathbf{w}_{j+1}) + \frac{\eta^2}{2}\|\nabla\ell(\mathbf{w}_t;\mathbf{z}_t)\|_{w,*}^2\\
\overset{(7)}{\leqslant}& B_w(\mathbf{w}_*,\mathbf{w}_t) - B_w(\mathbf{w}_*,\mathbf{w}_{j+1}) + \frac{\eta^2G^2}{2}.
\end{aligned} \tag{26}$$

因此，我们有

$$\begin{aligned}
&\eta\langle\nabla R(\mathbf{w}_t), \mathbf{w}_t - \mathbf{w}_*\rangle\\
=&\eta\langle\nabla\ell(\mathbf{w}_t;\mathbf{z}_t), \mathbf{w}_t - \mathbf{w}_*\rangle + \eta\langle\nabla R(\mathbf{w}_t) - \nabla\ell(\mathbf{w}_t;\mathbf{z}_t), \mathbf{w}_t - \mathbf{w}_*\rangle\\
\overset{(26)}{\leqslant}& B_w(\mathbf{w}_*,\mathbf{w}_t) - B_w(\mathbf{w}_*,\mathbf{w}_{j+1}) + \frac{\eta^2G^2}{2} + \eta\langle\nabla R(\mathbf{w}_t) - \nabla\ell(\mathbf{w}_t;\mathbf{z}_t), \mathbf{w}_t - \mathbf{w}_*\rangle.
\end{aligned}$$

对于 $t = 1, \dots, T$, 将上述不等式求和，我们有

$$\begin{aligned}
&\sum_{t=1}^{T} \eta \langle \nabla R(\mathbf{w}_t), \mathbf{w}_t - \mathbf{w}_* \rangle \\
\leqslant & B_w(\mathbf{w}_*, \mathbf{w}_1) + \frac{G^2}{2} \sum_{t=1}^{T} \eta^2 + \sum_{t=1}^{T} \eta \langle \nabla R(\mathbf{w}_t) - \nabla \ell(\mathbf{w}_t; \mathbf{z}_t), \mathbf{w}_t - \mathbf{w}_* \rangle \\
\overset{(5)}{\leqslant} & D^2 + \frac{G^2}{2} \sum_{t=1}^{T} \eta^2 + \sum_{t=1}^{T} \eta \langle \nabla R(\mathbf{w}_t) - \nabla \ell(\mathbf{w}_t; \mathbf{z}_t), \mathbf{w}_t - \mathbf{w}_* \rangle.
\end{aligned} \tag{27}$$

由于风险函数的凸性，我们有

$$\begin{aligned}
& R(\bar{\mathbf{w}}) - R(\mathbf{w}_*) = R\left(\sum_{t=1}^{T} \frac{\mathbf{w}_t}{T}\right) - R(\mathbf{w}_*) \\
\leqslant & \left(\sum_{t=1}^{T} \frac{1}{T} R(\mathbf{w}_t)\right) - R(\mathbf{w}_*) = \sum_{t=1}^{T} \frac{1}{T} \left(R(\mathbf{w}_t) - R(\mathbf{w}_*)\right) \\
\leqslant & \sum_{t=1}^{T} \frac{1}{T} \langle \nabla R(\mathbf{w}_t), \mathbf{w}_t - \mathbf{w}_* \rangle \\
\overset{(27)}{\leqslant} & \frac{D^2 + \frac{G^2}{2} \sum_{t=1}^{T} \eta^2 + \sum_{t=1}^{T} \eta \langle \nabla R(\mathbf{w}_t) - \nabla \ell(\mathbf{w}_t; \mathbf{z}_t), \mathbf{w}_t - \mathbf{w}_* \rangle}{\eta T}.
\end{aligned} \tag{28}$$

定义

$$\delta_t = \eta \langle \nabla R(\mathbf{w}_t) - \nabla \ell(\mathbf{w}_t; \mathbf{z}_t), \mathbf{w}_t - \mathbf{w}_* \rangle. \tag{29}$$

由于 $\mathbf{w}_t$ 和 $\mathbf{w}_*$ 不依赖于 $\mathbf{z}_t$，因此

$$\mathrm{E}_{t-1}[\delta_t] = \eta \left\langle \mathrm{E}_{t-1}\left[\nabla R(\mathbf{w}_t) - \nabla \ell(\mathbf{w}_t; \mathbf{z}_t)\right], \mathbf{w}_t - \mathbf{w}_* \right\rangle = 0.$$

因此，$\delta_1, \dots, \delta_t$ 是鞅差序列。

5.1.1 期望上界

对 (28) 式左右两边取期望, 我们有

$$\mathrm{E}\left[R(\bar{\mathbf{w}}) - R(\mathbf{w}_*)\right] \leqslant \frac{2D^2 + TG^2\eta^2}{2\eta T}. \tag{30}$$

设 $\eta = \frac{\sqrt{2}D}{G\sqrt{T}}$, 根据基本不等式，我们有

$$\mathrm{E}\left[R(\bar{\mathbf{w}}) - R(\mathbf{w}_*)\right] \leqslant DG\sqrt{\frac{2}{T}}. \tag{31}$$

因而 (24) 得证。

5.1.2 高概率上界

我们不加证明的给出如下引理，即鞅的 Hoeffding-Azuma 不等式。

引理 1 设 $V_1, V_2, \ldots$ 是关于某个序列 $X_1, X_2, \ldots$ 的鞅差序列，其中 $V_i \in [A_i, A_i + c_i]$，A_i 是随机变量，c_i 是常数。如果 $S_n = \sum_{i=1}^n V_i$, 那么对于任意 $t > 0$,

$$\Pr[S_n > t] \leqslant \exp\left(-\frac{2t^2}{\sum_{i=1}^n c_i^2}\right).$$

为了应用上面的引理，我们需要证明 $|\delta_t|$ 是有界的：

$$\begin{aligned}
&\|\nabla R(\mathbf{w}_t) - \nabla \ell(\mathbf{w}_t; \mathbf{z}_t)\|_{w,*} \leqslant \|\nabla R(\mathbf{w}_t)\|_{w,*} + \|\nabla \ell(\mathbf{w}_t; \mathbf{z}_t)\|_{w,*} \\
\leqslant & \mathrm{E}_{t-1}\left[\|\nabla \ell(\mathbf{w}_t; \mathbf{z}_t)\|_{w,*}\right] + \|\nabla \ell(\mathbf{w}_t; \mathbf{z}_t)\|_{w,*} \overset{(7)}{\leqslant} 2G,
\end{aligned} \tag{32}$$

$$\begin{aligned}
&\|\mathbf{w}_t - \mathbf{w}_*\|_w \leqslant \|\mathbf{w}_t - \mathbf{w}_1\|_w + \|\mathbf{w}_1 - \mathbf{w}_*\|_w \\
\leqslant & \left(\sqrt{B_w(\mathbf{w}_t, \mathbf{w}_1)} + \sqrt{B_w(\mathbf{w}_*, \mathbf{w}_1)}\right) \overset{(5)}{\leqslant} 2D.
\end{aligned} \tag{33}$$

因此,

$$\begin{aligned}
|\delta_t| =& \eta \left|\langle \nabla R(\mathbf{w}_t) - \nabla \ell(\mathbf{w}_t; \mathbf{z}_t), \mathbf{w}_t - \mathbf{w}_* \rangle\right| \\
\leqslant & \eta \|\nabla R(\mathbf{w}_t) - \nabla \ell(\mathbf{w}_t; \mathbf{z}_t)\|_{w,*} \|\mathbf{w}_t - \mathbf{w}_*\|_w \overset{(32),(33)}{\leqslant} 4\eta DG.
\end{aligned} \tag{34}$$

根据引理 1，以至少 $1-\delta$ 的概率，我们有

$$\sum_{t=1}^T \delta_t \leqslant 4DG \sqrt{2 \sum_{t=1}^T \eta^2 \ln \frac{1}{\delta}}. \tag{35}$$

将 (35) 代入 (28)，以至少 $1-\delta$ 的概率，我们有

$$\begin{aligned}
R(\bar{\mathbf{w}}) - R(\mathbf{w}_*) \leqslant & \frac{D^2 + \frac{G^2}{2} \sum_{t=1}^T \eta^2 + 4DG \sqrt{2 \sum_{t=1}^T \eta^2 \left(\ln \frac{1}{\delta}\right)}}{\sum_{t=1}^T \eta} \\
\leqslant & DG \sqrt{\frac{2}{T}} + 4DG \sqrt{\frac{2\left(\ln \frac{1}{\delta}\right)}{T}} \\
= & DG \sqrt{\frac{2}{T}} \left(1 + 4\sqrt{\ln \frac{1}{\delta}}\right)
\end{aligned}$$

5.2 定理 1和 3的证明

首先证明定理 3。证明基于 (Nemirovski et al., 2009) 的引理 3.1 和命题 3.2，即 §3.1 节——求解随机凸凹鞍点问题的镜像下降算法。

5.2.1 合并 (12) 和 (13) 中的更新规则

虽然 $\mathbf{w}$ 和 $\mathbf{q}$ 是通过 SMD 的两个实例来更新的，但可以通过将 $\mathbf{w}$ 和 $\mathbf{q}$ 连接为单个变量 $[\mathbf{w};\mathbf{q}] \in \mathcal{W}\times\Delta_m$，并重新定义范数和距离生成函数，以达到只使用 1 个 SMD 实例的目的 (Nemirovski et al., 2009, §3.1)。设 $\mathcal{E}$ 是 $\mathcal{W}$ 所在的空间。我们给笛卡尔积 $\mathcal{E}\times\mathbb{R}^m$ 配备以下范数和对偶范数:

$$\big\|[\mathbf{w};\mathbf{q}]\big\| = \sqrt{\frac{1}{2D^2}\|\mathbf{w}\|_w^2 + \frac{1}{2\ln m}\|\mathbf{q}\|_1^2}, \text{ and } \big\|[\mathbf{u},\mathbf{v}]\big\|_* = \sqrt{2D^2\|\mathbf{u}\|_{w,*}^2 + 2\|\mathbf{v}\|_\infty^2 \ln m}. \tag{36}$$

我们使用符号 $\mathbf{x} = [\mathbf{w};\mathbf{q}]$，并为集合 $\mathcal{W}\times\Delta_m$ 选择距离生成函数

$$\nu(\mathbf{x}) = \nu([\mathbf{w};\mathbf{q}]) = \frac{1}{2D^2}\nu_w(\mathbf{w}) + \frac{1}{2\ln m}\nu_q(\mathbf{q}). \tag{37}$$

易知 $\nu(\mathbf{x})$ 关于范数 $\|\cdot\|$ 是 1-强凸的。设 $B(\cdot,\cdot)$ 是与 $\nu(\cdot)$ 相关联的 Bregman 距离:

$$\begin{aligned} B(\mathbf{x}_1,\mathbf{x}_2) =& \nu(\mathbf{x}_1) - \big[\nu(\mathbf{x}_2) + \langle\nabla\nu(\mathbf{x}_2), \mathbf{x}_1 - \mathbf{x}_2\rangle\big] \\ =& \frac{1}{2D^2}\left(\nu_w(\mathbf{w}_1) - \big[\nu_w(\mathbf{w}_2) + \langle\nabla\nu_w(\mathbf{w}_2), \mathbf{w}_1 - \mathbf{w}_2\rangle\big]\right) \\ &+ \frac{1}{2\ln m}\left(\nu_q(\mathbf{q}_1) - \big[\nu_q(\mathbf{q}_2) + \langle\nabla\nu_q(\mathbf{q}_2), \mathbf{q}_1 - \mathbf{q}_2\rangle\big]\right) \\ =& \frac{1}{2D^2}B_w(\mathbf{w}_1,\mathbf{w}_2) + \frac{1}{2\ln m}B_q(\mathbf{q}_1,\mathbf{q}_2) \end{aligned} \tag{38}$$

其中 $\mathbf{x}_1 = [\mathbf{w}_1, \mathbf{q}_1]$ 并且 $\mathbf{x}_2 = [\mathbf{w}_2, \mathbf{q}_2]$.

然后，我们可以证明域 $\mathcal{W}\times\Delta_m$ 是有界的，因为

$$\max_{(\mathbf{w},\mathbf{q})\in\mathcal{W}\times\Delta_m} B([\mathbf{w},\mathbf{q}],[\mathbf{w}_1,\mathbf{q}_1]) = \frac{1}{2D^2}\max_{\mathbf{w}\in\mathcal{W}} B_w(\mathbf{w},\mathbf{w}_1) + \frac{1}{2\ln m}\max_{\mathbf{q}\in\Delta_m} B_q(\mathbf{q},\mathbf{q}_1) \overset{(5)}{\leqslant} 1. \tag{39}$$

接着，我们考虑用以下 SMD 来更新 $\mathbf{x}_t$:

$$\mathbf{x}_{t+1} = \operatorname*{argmin}_{\mathbf{x}\in\mathcal{W}\times\Delta_m} \Big\{\eta\big\langle[\mathbf{g}_w(\mathbf{w}_t,\mathbf{q}_t); -\mathbf{g}_q(\mathbf{w}_t,\mathbf{q}_t)], \mathbf{x}-\mathbf{x}_t\big\rangle + B(\mathbf{x},\mathbf{x}_t)\Big\} \tag{40}$$

这里 $\eta > 0$ 是步长。初始时，我们设置 $\mathbf{x}_1 = \operatorname{argmin}_{\mathbf{x}\in\mathcal{W}\times\Delta_m}\nu(\mathbf{x}) = [\mathbf{w}_1,\mathbf{q}_1]$。从 (38) 中的 Bregman 距离的分解过程，我们得出通过设置如下步长，(40) 等价于 (12) 和 (13)。

$$\eta_w = 2\eta D^2, \text{ and } \eta_q = 2\eta\ln m. \tag{41}$$

5.2.2 无偏随机梯度的 SMD 算法收敛性分析

为了简化符号，我们定义

$$\begin{aligned} F(\mathbf{w}_t,\mathbf{q}_t) &= [\nabla_{\mathbf{w}}\phi(\mathbf{w}_t,\mathbf{q}_t), -\nabla_{\mathbf{q}}\phi(\mathbf{w}_t,\mathbf{q}_t)] \\ &= \left[\sum_{i=1}^m q_{t,i}\nabla R_i(\mathbf{w}_t), -\big[R_1(\mathbf{w}_t),\dots,R_m(\mathbf{w}_t)\big]^\top\right] \end{aligned} \tag{42}$$

其包含在 $(\mathbf{w}_t, \mathbf{q}_t)$ 处的 $\phi(\cdot,\cdot)$ 的真实梯度，并且

$$\begin{aligned}\mathbf{g}(\mathbf{w}_t, \mathbf{q}_t) &= [\mathbf{g}_w(\mathbf{w}_t, \mathbf{q}_t); -\mathbf{g}_q(\mathbf{w}_t, \mathbf{q}_t)] \\ &\overset{(20)}{=} \left[\sum_{i=1}^m q_{t,i} \nabla \ell(\mathbf{w}_t; \mathbf{z}_t^{(i)}), -\left[\ell(\mathbf{w}_t; \mathbf{z}_t^{(1)}), \dots, \ell(\mathbf{w}_t; \mathbf{z}_t^{(m)})\right]^\top\right]\end{aligned} \tag{43}$$

即式 (40) 中的随机梯度，且其范数的上界为:

$$\begin{aligned}\|\mathbf{g}_w(\mathbf{w}_t, \mathbf{q}_t)\|_{w,*} &= \left\|\sum_{i=1}^m q_{t,i} \nabla \ell(\mathbf{w}_t; \mathbf{z}_t^{(i)})\right\|_{w,*} \leqslant \sum_{i=1}^m q_{t,i} \left\|\nabla \ell(\mathbf{w}_t; \mathbf{z}_t^{(i)})\right\|_{w,*} \overset{(7)}{\leqslant} \sum_{i=1}^m q_{t,i} G = G, \\ \|\mathbf{g}_q(\mathbf{w}_t, \mathbf{q}_t)\|_\infty &= \left\|\left[\ell(\mathbf{w}_t; \mathbf{z}_t^{(1)}), \dots, \ell(\mathbf{w}_t; \mathbf{z}_t^{(m)})\right]^\top\right\|_\infty \overset{(6)}{\leqslant} 1\end{aligned}$$

因而

$$\left\|\mathbf{g}(\mathbf{w}_t, \mathbf{q}_t)\right\|_* = \sqrt{2D^2 \|\mathbf{g}_w(\mathbf{w}_t, \mathbf{q}_t)\|_{w,*}^2 + 2\|\mathbf{g}_q(\mathbf{w}_t, \mathbf{q}_t)\|_\infty^2 \ln m} \leqslant \underbrace{\sqrt{2D^2G^2 + 2\ln m}}_{:=M}. \tag{44}$$

由 $\phi(\cdot,\cdot)$ 的凸凹性，我们有 (Nemirovski et al., 2009, (3.9))

$$\begin{aligned}&\max_{\mathbf{q}\in\Delta_m} \phi(\bar{\mathbf{w}}, \mathbf{q}) - \min_{\mathbf{w}\in\mathcal{W}} \phi(\mathbf{w}, \bar{\mathbf{q}}) \\ =& \max_{\mathbf{q}\in\Delta_m} \phi\left(\sum_{t=1}^T \frac{\mathbf{w}_t}{T}, \mathbf{q}\right) - \min_{\mathbf{w}\in\mathcal{W}} \phi\left(\mathbf{w}, \sum_{t=1}^T \frac{q_t}{T}\right) \\ \leqslant& \frac{1}{T}\left[\max_{\mathbf{q}\in\Delta_m} \sum_{t=1}^T \phi(\mathbf{w}_t, \mathbf{q}) - \min_{\mathbf{w}\in\mathcal{W}} \sum_{t=1}^T \phi(\mathbf{w}, \mathbf{q}_t)\right] \\ \leqslant& \frac{1}{T} \max_{(\mathbf{w},\mathbf{q})\in\mathcal{W}\times\Delta_m} \sum_{t=1}^T \left[\left\langle \nabla_{\mathbf{w}} \phi(\mathbf{w}_t, \mathbf{q}_t), \mathbf{w}_t - \mathbf{w}\right\rangle - \left\langle \nabla_{\mathbf{q}} \phi(\mathbf{w}_t, \mathbf{q}_t), \mathbf{q}_t - \mathbf{q}\right\rangle\right] \\ \overset{(42)}{=}& \frac{1}{T} \max_{\mathbf{x}\in\mathcal{W}\times\Delta_m} \sum_{t=1}^T \left\langle F(\mathbf{w}_t, \mathbf{q}_t), \mathbf{x}_t - \mathbf{x}\right\rangle.\end{aligned} \tag{45}$$

因此，我们可以将优化误差分解如下：

$$\begin{aligned}&\max_{\mathbf{q}\in\Delta_m} \phi(\bar{\mathbf{w}}, \mathbf{q}) - \min_{\mathbf{w}\in\mathcal{W}} \phi(\mathbf{w}, \bar{\mathbf{q}}) \\ \leqslant& \underbrace{\frac{1}{T} \max_{\mathbf{x}\in\mathcal{W}\times\Delta_m} \sum_{t=1}^T \left\langle \mathbf{g}(\mathbf{w}_t, \mathbf{q}_t), \mathbf{x}_t - \mathbf{x}\right\rangle}_{:=E_1} \\ &+ \frac{1}{T} \underbrace{\max_{\mathbf{x}\in\mathcal{W}\times\Delta_m} \sum_{t=1}^T \left\langle F(\mathbf{w}_t, \mathbf{q}_t) - \mathbf{g}(\mathbf{w}_t, \mathbf{q}_t), \mathbf{x}_t - \mathbf{x}\right\rangle}_{:=E_2}.\end{aligned} \tag{46}$$

我们继续约束 (46) 中的两项。为了约束第一项 E_1，我们仿照 (26) 的推导，有

$$\begin{aligned}\eta\left\langle \mathbf{g}(\mathbf{w}_t, \mathbf{q}_t), \mathbf{x}_t - \mathbf{x}\right\rangle \leqslant& B(\mathbf{x}, \mathbf{x}_t) - B(\mathbf{x}, \mathbf{x}_{t+1}) + \frac{\eta^2}{2}\|\mathbf{g}(\mathbf{w}_t, \mathbf{q}_t)\|_*^2 \\ \overset{(44)}{\leqslant}& B(\mathbf{x}, \mathbf{x}_t) - B(\mathbf{x}, \mathbf{x}_{t+1}) + \frac{M^2}{2}\eta^2.\end{aligned} \tag{47}$$

对于 $t=1,\ldots,T$, 对上述不等式求和，我们有

$$\sum_{t=1}^{T}\eta\big\langle\mathbf{g}(\mathbf{w}_t,\mathbf{q}_t),\mathbf{x}_t-\mathbf{x}\big\rangle\leqslant B(\mathbf{x},\mathbf{x}_1)+\frac{M^2}{2}\sum_{t=1}^{T}\eta^2\overset{(39)}{\leqslant}1+\frac{M^2}{2}\sum_{t=1}^{T}\eta^2. \tag{48}$$

接下来，我们考虑第二项 E_2。由于最大化运算，变量 $\mathbf{x}$ 受到算法随机性的影响，因此我们不能将 $\eta\langle\mathrm{E}_{t-1}[\mathbf{g}(\mathbf{w}_t,\mathbf{q}_t)]-\mathbf{g}(\mathbf{w}_t,\mathbf{q}_t),\mathbf{x}_t-\mathbf{x}\rangle$, $t=1,\ldots,T$ 视为鞅差序列。为了解决这一挑战，我们使用Nemirovski et al. (2009) 中的“幽灵迭代”技术，得出以下引理:

引理 2　在定理 3的条件下，我们有

$$\mathrm{E}\left[\max_{\mathbf{x}\in\mathcal{W}\times\Delta_m}\sum_{t=1}^{T}\eta\big\langle F(\mathbf{w}_t,\mathbf{q}_t)-\mathbf{g}(\mathbf{w}_t,\mathbf{q}_t),\mathbf{x}_t-\mathbf{x}\big\rangle\right]\leqslant 1+2M^2\sum_{t=1}^{T}\eta^2. \tag{49}$$

此外，以至少 $1-\delta$ 的概率，我们有

$$\begin{aligned}&\max_{\mathbf{x}\in\mathcal{W}\times\Delta_m}\sum_{t=1}^{T}\eta\big\langle F(\mathbf{w}_t,\mathbf{q}_t)-\mathbf{g}(\mathbf{w}_t,\mathbf{q}_t),\mathbf{x}_t-\mathbf{x}\big\rangle\\ \leqslant&1+2M^2\sum_{t=1}^{T}\eta^2+4M\sqrt{2\sum_{t=1}^{T}\eta^2\left(\ln\frac{1}{\delta}\right)}.\end{aligned} \tag{50}$$

证明: 我们通过以 $F(\mathbf{w}_t,\mathbf{q}_t)-\mathbf{g}(\mathbf{w}_t,\mathbf{q}_t)$ 作为梯度执行 SMD 来创建虚拟序列:

$$\mathbf{y}_{t+1}=\underset{\mathbf{x}\in\mathcal{W}\times\Delta_m}{\operatorname{argmin}}\left\{\eta\big\langle F(\mathbf{w}_t,\mathbf{q}_t)-\mathbf{g}(\mathbf{w}_t,\mathbf{q}_t),\mathbf{x}-\mathbf{y}_t\big\rangle+B(\mathbf{x},\mathbf{y}_t)\right\} \tag{51}$$

其中 $\mathbf{y}_1=\mathbf{x}_1$。然后，我们进一步将误差项分解为

$$\begin{aligned}&\max_{\mathbf{x}\in\mathcal{W}\times\Delta_m}\sum_{t=1}^{T}\eta\big\langle F(\mathbf{w}_t,\mathbf{q}_t)-\mathbf{g}(\mathbf{w}_t,\mathbf{q}_t),\mathbf{x}_t-\mathbf{x}\big\rangle\\ \leqslant&\max_{\mathbf{x}\in\mathcal{W}\times\Delta_m}\underbrace{\sum_{t=1}^{T}\eta\big\langle F(\mathbf{w}_t,\mathbf{q}_t)-\mathbf{g}(\mathbf{w}_t,\mathbf{q}_t),\mathbf{y}_t-\mathbf{x}\big\rangle}_{:=A}\\ &+\underbrace{\sum_{t=1}^{T}\eta\big\langle F(\mathbf{w}_t,\mathbf{q}_t)-\mathbf{g}(\mathbf{w}_t,\mathbf{q}_t),\mathbf{x}_t-\mathbf{y}_t\big\rangle}_{:=B}.\end{aligned} \tag{52}$$

为了约束项 A，我们重复 (47) 的分析，有

$$\begin{aligned}&\eta\big\langle F(\mathbf{w}_t,\mathbf{q}_t)-\mathbf{g}(\mathbf{w}_t,\mathbf{q}_t),\mathbf{y}_t-\mathbf{x}\big\rangle\\ \leqslant&B(\mathbf{x},\mathbf{y}_t)-B(\mathbf{x},\mathbf{y}_{t+1})+\frac{\eta^2}{2}\big\|F(\mathbf{w}_t,\mathbf{q}_t)-\mathbf{g}(\mathbf{w}_t,\mathbf{q}_t)\big\|_*^2\\ \leqslant&B(\mathbf{x},\mathbf{y}_t)-B(\mathbf{x},\mathbf{y}_{t+1})+2M^2\eta^2\end{aligned} \tag{53}$$

其中最后一步利用以下不等式

$$\big\|F(\mathbf{w}_t,\mathbf{q}_t)-\mathbf{g}(\mathbf{w}_t,\mathbf{q}_t)\big\|_*\leqslant\big\|F(\mathbf{w}_t,\mathbf{q}_t)\big\|_*+\|\mathbf{g}(\mathbf{w}_t,\mathbf{q}_t)\|_*\overset{(44)}{\leqslant}2M. \tag{54}$$

对 $t=1,\ldots,T$ 累加 (53), 我们有

$$\begin{aligned} A=&\sum_{t=1}^{T}\eta\big\langle F(\mathbf{w}_t,\mathbf{q}_t)-\mathbf{g}(\mathbf{w}_t,\mathbf{q}_t),\mathbf{y}_t-\mathbf{x}\big\rangle \\ \leqslant&B(\mathbf{x},\mathbf{y}_1)+2M^2\sum_{t=1}^{T}\eta^2\overset{(39)}{\leqslant}1+2M^2\sum_{t=1}^{T}\eta^2. \end{aligned} \tag{55}$$

为了约束 (52) 中的项 B，我们定义

$$\delta_t=\eta\big\langle F(\mathbf{w}_t,\mathbf{q}_t)-\mathbf{g}(\mathbf{w}_t,\mathbf{q}_t),\mathbf{x}_t-\mathbf{y}_t\big\rangle.$$

由于 $\mathbf{x}_t$ 和 $\mathbf{y}_t$ 独立于用于在 (43) 中构造梯度 $\mathbf{g}(\mathbf{w}_t,\mathbf{q}_t)$ 的随机样本 $\mathbf{z}_t^{(1)},\ldots,\mathbf{z}_t^{(m)}$，$\delta_1,\ldots,\delta_T$ 构成鞅差序列。因此，我们有

$$\mathrm{E}[B]=\mathrm{E}\left[\sum_{t=1}^{T}\delta_t\right]=0. \tag{56}$$

对 (52) 两边求期望, 我们有

$$\begin{aligned} &\mathrm{E}\left[\max_{\mathbf{x}\in\mathcal{W}\times\Delta_m}\sum_{t=1}^{T}\eta\big\langle F(\mathbf{w}_t,\mathbf{q}_t)-\mathbf{g}(\mathbf{w}_t,\mathbf{q}_t),\mathbf{x}_t-\mathbf{x}\big\rangle\right] \\ \leqslant&\mathrm{E}\left[\max_{\mathbf{x}\in\mathcal{W}\times\Delta_m}A\right]+\mathrm{E}[B]\overset{(55),(56)}{\leqslant}1+2M^2\sum_{t=1}^{T}\eta^2 \end{aligned}$$

为了建立高概率上界，我们遵循 5.1.2节中的分析，并利用引理 1来确定 B 的上界。为此，我们首先证明 $|\delta_t|$ 有界：

$$\begin{aligned} |\delta_t|=&\Big|\eta\big\langle F(\mathbf{w}_t,\mathbf{q}_t)-\mathbf{g}(\mathbf{w}_t,\mathbf{q}_t),\mathbf{x}_t-\mathbf{y}_t\big\rangle\Big| \\ \leqslant&\eta\big\|F(\mathbf{w}_t,\mathbf{q}_t)-\mathbf{g}(\mathbf{w}_t,\mathbf{q}_t)\big\|_*\|\mathbf{x}_t-\mathbf{y}_t\| \\ \overset{(54)}{\leqslant}&2M\eta\big(\|\mathbf{x}_t-\mathbf{x}_1\|+\|\mathbf{x}_1-\mathbf{y}_t\|\big) \\ \leqslant&2M\eta\left(\sqrt{B(\mathbf{x}_t,\mathbf{x}_1)}+\sqrt{B(\mathbf{y}_t,\mathbf{x}_1)}\right)\overset{(39)}{\leqslant}4M\eta. \end{aligned}$$

根据引理 1和联合界不等式，以至少 $1-\delta$ 的概率有

$$B=\sum_{t=1}^{T}\delta_t\leqslant 4M\sqrt{2\sum_{t=1}^{T}\eta^2\left(\ln\frac{1}{\delta}\right)}. \tag{57}$$

我们把 (55) 和 (57) 代入 (52) 即可得到 (50)。

至此，引理2证明完毕。

结合 (46) , (48) 和 (49), 我们有

$$\mathrm{E}\left[\max_{\mathbf{q}\in\Delta_m}\phi(\bar{\mathbf{w}},\mathbf{q})-\min_{\mathbf{w}\in\mathcal{W}}\phi(\mathbf{w},\bar{\mathbf{q}}_t)\right]\leqslant\frac{2}{\eta T}+\frac{5M^2}{2}\eta.$$

通过设定

$$\eta = \frac{2}{M\sqrt{5T}} = \sqrt{\frac{2}{5T(D^2G^2 + \ln m)}}, \tag{58}$$

我们有

$$\mathrm{E}\left[\max_{\mathbf{q}\in\Delta_m} \phi(\bar{\mathbf{w}}, \mathbf{q}) - \min_{\mathbf{w}\in\mathcal{W}} \phi(\mathbf{w}, \bar{\mathbf{q}}_t)\right] \leqslant 2M\sqrt{\frac{5}{T}} = 2\sqrt{\frac{10(D^2G^2 + \ln m)}{T}}.$$

结合 (46) , (48) 和 (50), 以至少 $1-\delta$ 的概率, 我们有

$$\max_{\mathbf{q}\in\Delta_m} \phi(\bar{\mathbf{w}}, \mathbf{q}) - \min_{\mathbf{w}\in\mathcal{W}} \phi(\mathbf{w}, \bar{\mathbf{q}}_t) \leqslant 2M\sqrt{\frac{5}{T}} + 4M\sqrt{\frac{\ln 1/\delta}{T}} = 2\left(\sqrt{10} + 2\sqrt{2\ln 1/\delta}\right)\sqrt{\frac{D^2G^2 + \ln m}{T}}.$$

至此，定理3证明完毕。

接下来证明定理 1，易知算法 1和算法 3的不同之处仅在于随机梯度的构建方式，因此根据算法 1的假设，我们有

$$\|\mathbf{g}_w(\mathbf{w}_t, \mathbf{q}_t)\|_{w,*} = \left\|q_{t,i_t} m \nabla \ell(\mathbf{w}_t; \mathbf{z}_t^{(i)})\right\|_{w,*} \leqslant q_{t,i_t} m \left\|\nabla \ell(\mathbf{w}_t; \mathbf{z}_t^{(i)})\right\|_{w,*} \overset{(7)}{\leqslant} q_{t,i_t} mG \leqslant mG,$$

$$\|\mathbf{g}_q(\mathbf{w}_t, \mathbf{q}_t)\|_\infty = \left\|[0, \ldots, m\ell(\mathbf{w}_t; \mathbf{z}_t^{(i_t)}), \ldots, 0]^\top\right\|_\infty \overset{(6)}{\leqslant} m.$$

$$\begin{aligned}\left\|[\mathbf{g}_w(\mathbf{w}_t, \mathbf{q}_t); -\mathbf{g}_q(\mathbf{w}_t, \mathbf{q}_t)]\right\|_* &= \sqrt{2D^2\|\mathbf{g}_w(\mathbf{w}_t, \mathbf{q}_t)\|_{w,*}^2 + 2\|\mathbf{g}_q(\mathbf{w}_t, \mathbf{q}_t)\|_\infty^2 \ln m} \\ &\leqslant \underbrace{\sqrt{2m^2D^2G^2 + 2m^2\ln m}}_{:=M}.\end{aligned}$$

通过设定

$$\eta = \frac{2}{M\sqrt{5T}} = \sqrt{\frac{2}{5Tm^2(D^2G^2 + \ln m)}},$$

并仿照定理 3的证明过程即可完成证明。

至此，定理1证明完毕。

5.3 定理 2的证明

我们首先给出一些引理:

在线镜像下降 设 $\mathcal{W} \subseteq \mathbb{R}^d$ 是一个紧凸集并且 $\nu_w: \mathcal{W} \to \mathbb{R}$ 是严格凸函数。在线镜像下降（OMD）算法的过程如下，对于 $t = 1, \ldots, T$:

1. 令 $\tilde{\mathbf{w}}_{t+1} \in \mathbb{R}^n$ 为 $\nabla\nu_w(\tilde{\mathbf{w}}_{t+1}) = \nabla\nu_w(\mathbf{w}_t) - \eta_t\nabla\nu_w(\nabla_t)$ 的解, 其中 $\eta_t > 0$ 是步长并且 $\nabla_t = \nabla f_t(\mathbf{w}_t)$ 是第 t 轮的梯度反馈。

2. 令 $\mathbf{w}_{t+1} \in \operatorname{argmin}_{\mathbf{w}\in\mathcal{W}} B_w(\mathbf{w}, \tilde{\mathbf{w}}_{t+1})$, 其中 $B_w(\mathbf{u}, \mathbf{v}) = \nu_w(\mathbf{u}) - \left[\nu_w(\mathbf{v}) + \langle\nabla\nu_w(\mathbf{v}), \mathbf{u} - \mathbf{v}\rangle\right]$ 是与 $\nu_w(\cdot)$ 相关的 Bregman 距离。

引理 3 (OMD 的遗憾界; Orabona (2019, Theorem 6.8)) OMD 对于任何 $\mathbf{w}^* \in \mathcal{W}$，满足

$$\sum_{t=1}^{T} f_t(\mathbf{w}_t) - \sum_{t=1}^{T} f_t(\mathbf{w}^*) \leqslant \frac{1}{2}\sum_{t=1}^{T} \eta_t \|\nabla_t\|_{t,*}^2 + \frac{B_w(\mathbf{w}^*, \mathbf{w}_1)}{\eta_1} + \sum_{t=2}^{T}\left(\frac{1}{\eta_t} - \frac{1}{\eta_{t-1}}\right) B_w(\mathbf{w}^*, \mathbf{w}_t), \tag{59}$$

其中 $\|\mathbf{w}\|_t$ 表示局部范数, 即 $\|\mathbf{w}\|_t = \sqrt{\mathbf{w}^\top \nabla^2 \nu_w(\mathbf{x}_t)\mathbf{w}}$ 对于 $\mathbf{x}_t \in [\mathbf{w}_t, \tilde{\mathbf{w}}_{t+1}]$ 并且 $\|\mathbf{w}\|_{t,*} = \sqrt{\mathbf{w}^\top \nabla^2 \nu_w(\mathbf{x}_t)^{-1}\mathbf{w}}$ 为其对偶范数。

在线梯度下降 当 $\nu_w(\mathbf{w}) = \frac{1}{2}\mathbf{w}_2^2$ 时，OMD 可以简化为在线梯度下降 (OGD)：

$$\mathbf{w}_{t+1} = \Pi_{\mathcal{W}}\Big[\mathbf{w}_t - \eta_t \nabla_t\Big].$$

注意到 $B_w(\mathbf{u}, \mathbf{v}) = \frac{1}{2}\|\mathbf{u} - \mathbf{v}\|_2^2$，最小化 Bregman 散度等价于正交投影。

引理 4 (OGD 的遗憾界) 对于任意 $\mathbf{w}^* \in \mathcal{W}$，OMD 满足

$$\sum_{t=1}^{T} f_t(\mathbf{w}_t) - \sum_{t=1}^{T} f_t(\mathbf{w}^*) \leqslant \frac{1}{2}\sum_{t=1}^{T} \eta_t \|\nabla_t\|^2 + \frac{\|\mathbf{w}^* - \mathbf{w}_1\|_2^2}{2\eta_1} + \frac{1}{2}\sum_{t=2}^{T}\left(\frac{1}{\eta_t} - \frac{1}{\eta_{t-1}}\right)\|\mathbf{w}^* - \mathbf{w}_t\|_2^2. \tag{60}$$

如果使用递减步长并且 $\max_{t=1}^T \|\mathbf{w}^* - \mathbf{w}_t\| \leqslant D$, 我们有

$$\sum_{t=1}^{T} f_t(\mathbf{w}_t) - \sum_{t=1}^{T} f_t(\mathbf{w}^*) \leqslant \frac{1}{2}\sum_{t=1}^{T} \eta_t \|\nabla_t\|^2 + \frac{D^2}{2\eta_T}. \tag{61}$$

Hedge 当 $\nu_w(x) = \sum_i (x_i \log x_i - x_i)$ 时，OMD 等价于 Hedge 算法

$$\tilde{\mathbf{w}}_{t+1} = \mathbf{w}_t \exp(-\eta_t \nabla_t), \quad \mathbf{w}_{t+1} = \frac{\tilde{\mathbf{w}}_{t+1}}{\|\tilde{\mathbf{w}}_{t+1}\|_1}.$$

注意到 $\nabla^2 \nu_w(\mathbf{w}) = \mathrm{diag}(1/\mathbf{w}_i)$。如果 $\nabla_t \geqslant 0$, 那么 $\tilde{\mathbf{w}}_{t+1} \leqslant \mathbf{w}_t$ 并且 $\|\nabla_t\|_{t,*} \leqslant \|\nabla_t\|_{\nabla^2 \nu_w(\mathbf{w}_t)^{-1}}$。对于任意 $\mathbf{w}^*$，$\mathbf{w}_1 = \mathbf{1}/d$, 有 $\nu_w(\mathbf{w}^*, \mathbf{w}_1) \leqslant \log d$。

引理 5 (Hedge 的遗憾界) 对于 $\nabla_t \geqslant 0$ $(t = 1, \ldots, 1)$, 固定步长 $\eta > 0$ 的 Hedge 算法满足

$$\sum_{t=1}^{T} f_t(\mathbf{w}_t) - \sum_{t=1}^{T} f_t(\mathbf{w}^*) \leqslant \frac{\eta}{2}\sum_{t=1}^{T} \|\nabla_t\|_{\nabla^2 \nu_w(\mathbf{w}_t)^{-1}}^2 + \frac{\log d}{\eta}. \tag{62}$$

定理 2的证明 设 i_t 和 $\mathbf{z}_t^{(i_t)}$ 分别为第 t 轮迭代时所选的组和样本。观察到算法2分别使用了在线学习 OGD 和 Hedge，并且其随机梯度构建为

$$\mathbf{g}_w(\mathbf{w}_t, \mathbf{q}_t) = \nabla \ell(\mathbf{w}_t; \mathbf{z}_t^{(i_t)}), \text{ and } \mathbf{g}_q(\mathbf{w}_t, \mathbf{q}_t) = [0, \ldots, \frac{\ell(\mathbf{w}_t; \mathbf{z}_t^{(i_t)})}{q_{t,i_t}}, \ldots, 0]^\top.$$

对于 OGD 算法, 假设 $\|\nabla \ell(\mathbf{w}; \mathbf{z})\|_2^2 \leqslant G$，我们使用引理 4得出

$$\mathrm{E}[R_{\mathbf{w}}(T)] \leqslant \frac{G^2}{2}\sum_{t=1}^{T} \eta_{\mathbf{w},t} + \frac{D^2}{2\eta_{\mathbf{w},T}}$$

对于 OMD 算法, 我们使用引理 3，得到

$$\|\mathbf{g}_q(\mathbf{w}_t,\mathbf{q}_t)\|^2_{\nabla^2\nu_w(\mathbf{q}_t)^{-1}} = \frac{\ell(\mathbf{w}_t,\mathbf{z}_t)^2(\nabla^2\nu_w(\mathbf{q}_t)^{-1})_{i_t,i_t}}{\mathbf{q}^2_{t,i_t}} \leqslant \frac{(\nabla^2\nu_w(\mathbf{q}_t)^{-1})_{i_t,i_t}}{\mathbf{q}^2_{t,i_t}}.$$

因此得到,

$$\mathrm{E}[R_q(T)] \leqslant \frac{1}{2}\sum_{t=1}^{T}\eta_q \mathrm{E}_{i_t}\left[\frac{(\nabla^2\nu_w(\mathbf{q}_t))^{-1}_{i_t,i_t}}{\mathbf{q}^2_{t,i_t}}\right] + \frac{B_w(\mathbf{q}^*,\mathbf{q}_1)}{\eta_q}$$

因为 $\nabla^2\nu_w(\mathbf{q}_t) = \mathrm{diag}(q_t^{-1})$, 我们得到

$$\mathrm{E}_{i_t}\left[\frac{(\nabla^2\nu_w(\mathbf{q}_t))^{-1}_{i_t,i_t}}{\mathbf{q}^2_{t,i_t}}\right] = \sum_{i=1}^{m}\Pr(i_t=i)\cdot\frac{1}{\mathbf{q}_{t,i}} = m,$$

在 $i_1,\ldots,i_{t-1}$ 的条件下。因此

$$\begin{aligned}\mathrm{E}[\epsilon_\phi(\bar{\mathbf{w}})] &\leqslant \frac{\mathrm{E}[R_{\mathbf{w}}(T)]+\mathrm{E}[R_{\mathbf{q}}(T)]}{T}\\ &\leqslant \frac{1}{T}\left(\frac{G^2}{2}\sum_{t=1}^{T}\eta_{\mathbf{w},t} + \frac{D^2}{2\eta_{\mathbf{w},T}} + \frac{1}{2}\sum_{t=1}^{T}\eta_q m + \frac{B_w(\mathbf{q}^*,\mathbf{q}_1)}{\eta_q}\right)\\ &\leqslant \frac{1}{T}\left(\frac{G^2}{2}\sum_{t=1}^{T}\eta_{\mathbf{w},t} + \frac{D^2}{2\eta_{\mathbf{w},T}} + \frac{1}{2}\sum_{t=1}^{T}\eta_q m + \frac{\log m}{\eta_q}\right)\end{aligned}$$

设 $\eta_{w,t} = \frac{D}{G\sqrt{T}}$, $\eta_q = \sqrt{\frac{2\log m}{mM^2T}}$, 得到

$$\mathrm{E}[\epsilon_\phi(\bar{\mathbf{w}})] \leqslant \sqrt{2}\frac{\sqrt{G^2D^2+2M^2m\log m}}{\sqrt{T}}.$$

至此，定理2证明完毕。

5.4 定理 4的证明

对于 (22) 中的随机梯度，其范数上界可以通过和 (20) 类似方式得出，即

$$\|\mathbf{g}_w(\mathbf{w}_t,\mathbf{q}_t)\|_{w,*} = \left\|\sum_{i\in C_t} q_{t,i}\nabla\ell(\mathbf{w}_t;\mathbf{z}_t^{(i)})\right\|_{w,*} \leqslant \sum_{i\in C_t} q_{t,i}\left\|\nabla\ell(\mathbf{w}_t;\mathbf{z}_t^{(i)})\right\|_{w,*} \overset{(7)}{\leqslant} \sum_{i\in C_t} q_{t,i}G = G,$$

$$\|\mathbf{g}_q(\mathbf{w}_t,\mathbf{q}_t)\|_\infty = \max_{i\in C_t}|\ell(\mathbf{w}_t;\mathbf{z}_t^{(i)})| \overset{(6)}{\leqslant} 1.$$

因此，通过与定理 3完全相同的分析，我们有

$$\mathrm{E}\big[\epsilon_\varphi(\bar{\mathbf{w}},\bar{\mathbf{q}})\big] \leqslant 2\sqrt{\frac{10(D^2G^2+\ln m)}{n_1}}$$

以至少为 $1-\delta$ 的概率，有

$$\epsilon_\varphi(\bar{\mathbf{w}},\bar{\mathbf{q}}) \leqslant 2\left(\sqrt{10}+2\sqrt{2\ln 1/\delta}\right)\sqrt{\frac{D^2G^2+\ln m}{n_1}}. \tag{63}$$

接下来，我们将讨论如何在每个分布 $\mathcal{P}_i$ 上限定 $\bar{\mathbf{w}}$ 的风险，即 $R_i(\bar{\mathbf{w}})$。在 (9) 中的推导后，我们知道

$$\max_{i\in[m]} p_i R_i(\bar{\mathbf{w}}) - \min_{\mathbf{w}\in\mathcal{W}}\max_{i\in[m]} p_i R_i(\mathbf{w}) \leqslant \epsilon_\varphi(\bar{\mathbf{w}},\bar{\mathbf{q}}).$$

因此，对于所有分布 $\mathcal{P}_i$，$R_i(\bar{\mathbf{w}})$ 的上界为：

$$R_i(\bar{\mathbf{w}}) \leqslant \frac{1}{p_i}\min_{\mathbf{w}\in\mathcal{W}}\max_{i\in[m]} p_i R_i(\mathbf{w}) + \frac{1}{p_i}\epsilon_\varphi(\bar{\mathbf{w}},\bar{\mathbf{q}}).$$

以 (63) 中的高概率界为例，我们以至少 $1-\delta$ 的概率有

$$\begin{aligned} R_i(\bar{\mathbf{w}}) \leqslant& \frac{1}{p_i}\min_{\mathbf{w}\in\mathcal{W}}\max_{i\in[m]} p_i R_i(\mathbf{w}) + \frac{1}{p_i}2\left(\sqrt{10}+2\sqrt{2\ln 1/\delta}\right)\sqrt{\frac{D^2G^2+\ln m}{n_1}} \\ =& \frac{n_1}{n_i}\min_{\mathbf{w}\in\mathcal{W}}\max_{i\in[m]} p_i R_i(\mathbf{w}) + 2\left(\sqrt{10}+2\sqrt{2\ln 1/\delta}\right)\frac{\sqrt{(D^2G^2+\ln m)n_1}}{n_i}. \end{aligned} \tag{64}$$

至此，定理4证明完毕。

6 总结与展望

6.1 总结

GDRO 的目的是学习一个在 m 个不同分布表现良好的模型。其可以形式化为随机凸凹鞍点问题，进而利用随机镜像下降（SMD）算法解决。本文主要介绍了近年来 GDRO 领域的三篇论文，第一篇论文提出 SMD(1) 算法，从实验方面表明了组分布鲁棒优化可以提高过参数化模型的鲁棒性，但由于构建的随机梯度方差上界较大，仅能获得次优的样本复杂度 $O(m^2(\log m)/\epsilon^2)$；第二篇论文将组分布鲁棒优化看作是一个双人零和博弈问题，并提出了 GDRO-EXP3 算法，使得样本复杂度降低到 $O(m(\log m)/\epsilon^2)$，但由于没有考虑在线学习的非遗忘性质，其理论证明存在依赖性问题；第三篇论文从减小随机梯度的对偶范数角度出发，通过每轮使用 m 个样本，提出了 SMD(m) 算法，将样本复杂度降低到 $O(m(\log m)/\epsilon^2)$，这几乎与该问题的下界 $\Omega(m/\epsilon^2)$ 相匹配。然后，又针对从每个分布中采样的数量不同，提出了一种新的加权组分布鲁棒优化（weighted DRO）情景，即用 n_i 表示第 i 个分布的样本预算，并假设 $n_1 \geqslant n_2 \geqslant \cdots \geqslant n_m$。为了解决这个问题，其在 SMD 中引入非均匀抽样，使得样本预算满足期望，并证明了第 i 个分布的超额风险以 $O(\sqrt{n_1\log m}/n_i)$ 的速率收敛。

6.2 心得

GDRO 可以形式化为随机凸凹鞍点问题，进而可以使用随机优化的一些技术，例如 SMD 算法来解决，而要想取得最优的样本复杂度，关键在于随机梯度的构建方式，三篇论文整体框架比较相似，不同点也基本都在随机梯度的构建方式上，如何平衡定义域直径大小与随机梯度对偶范数大小，

如何平衡无偏性与方差大小，都是需要考虑的问题。此外，理论研究的证明也需要注重细节，比如第二篇论文就没有考虑到条件期望的依赖性问题，导致证明存在问题。最后，提出一些新的有实际意义的设定，并推导其理论保证也是一种较好的科研方式。

6.3 展望

关于未来展望，虽然本文主要关注 GDRO，但可能的未来工作是将 SA 算法扩展到具有类似形式的问题，例如 PAC 协作学习 (Blum et al., 2017)、联邦学习 (Mohri et al., 2019) 和极小极大遗憾优化 (Agarwal and Zhang, 2022) 问题。

参考文献

Aharon Ben-Tal, Dick den Hertog, Anja De Waegenaere, Bertrand Melenberg, and Gijs Rennen. Robust solutions of optimization problems affected by uncertain probabilities. *Management Science*, 59(2):341–357, 2013.

Shiori Sagawa, Pang Wei Koh, Tatsunori B. Hashimoto, and Percy Liang. Distributionally robust neural networks for group shifts: On the importance of regularization for worst-case generalization. In *International Conference on Learning Representations*, 2020.

A. Nemirovski, A. Juditsky, G. Lan, and A. Shapiro. Robust stochastic approximation approach to stochastic programming. *SIAM Journal on Optimization*, 19(4):1574–1609, 2009.

Herbert Scarf. A min-max solution of an inventory problem. *Studies in the Mathematical Theory of Inventory and Production*, pages 201–209, 1958.

Aman Sinha, Hongseok Namkoong, and John Duchi. Certifying some distributional robustness with principled adversarial training. In *International Conference on Learning Representations*, 2018.

Tatsunori Hashimoto, Megha Srivastava, Hongseok Namkoong, and Percy Liang. Fairness without demographics in repeated loss minimization. In *Proceedings of the 35th International Conference on Machine Learning*, pages 1929–1938, 2018.

Ziyu Xu, Chen Dan, Justin Khim, and Pradeep Ravikumar. Class-weighted classification: Trade-offs and robust approaches. In *Proceedings of the 37th International Conference on Machine Learning*, pages 10544–10554, 2020.

Dvir Samuel and Gal Chechik. Distributional robustness loss for long-tail learning. In *Proceedings of the 2021 IEEE/CVF International Conference on Computer Vision*, pages 9475–9484, 2021.

Jingzhao Zhang, Aditya Krishna Menon, Andreas Veit, Srinadh Bhojanapalli, Sanjiv Kumar, and Suvrit Sra. Coping with label shift via distributionally robust optimisation. In *International Conference on Learning Representations*, 2021.

Nika Haghtalab, Michael I. Jordan, and Eric Zhao. On-demand sampling: Learning optimally from multiple distributions. In *Advances in Neural Information Processing Systems 35*, pages 406–419, 2022.

Tasuku Soma, Khashayar Gatmiry, and Stefanie Jegelka. Optimal algorithms for group distributionally robust optimization and beyond. *ArXiv e-prints*, arXiv:2212.13669, 2022.

Alekh Agarwal and Tong Zhang. Minimax regret optimization for robust machine learning under distribution shift. In *Proceedings of 35th Conference on Learning Theory*, pages 2704–2729, 2022.

Agnieszka Słowik and Léon Bottou. On distributionally robust optimization and data rebalancing. In *Proceedings of the 25th International Conference on Artificial Intelligence and Statistics*, pages 1283–1297, 2022.

Yoav Freund and Robert E Schapire. A decision-theoretic generalization of on-line learning and an application to boosting. *Journal of Computer and System Sciences*, 55(1):119–139, 1997.

Predrag Radivojac, Nitesh V. Chawla, A. Keith Dunker, and Zoran Obradovic. Classification and knowledge discovery in protein databases. *Journal of Biomedical Informatics*, 37(4):224–239, 2004.

Alekh Agarwal and Tong Zhang. Minimax regret optimization for robust machine learning under distribution shift. In *Proceedings of 35th Conference on Learning Theory*, pages 2704–2729, 2022.

L. Zhang, P. Zhao, T. Yang, and Z.-H. Zhou. Stochastic approximation approaches to group distributionally robust optimization, 2023.

Avrim Blum, Nika Haghtalab, Ariel D. Procaccia, and Mingda Qiao. Collaborative PAC learning. In *Advances in Neural Information Processing Systems 30*, pages 2389–2398, 2017.

Mehryar Mohri, Gary Sivek, and Ananda Theertha Suresh. Agnostic federated learning. In *Proceedings of the 36th International Conference on Machine Learning*, pages 4615–4625, 2019.

Tasuku Soma, Khashayar Gatmiry, and Stefanie Jegelka. Optimal algorithms for group distributionally robust optimization and beyond. *arXiv preprint arXiv:2212.13669*, 2022.

Peter Auer, Nicolò Cesa-Bianchi, Yoav Freund, and Robert E. Schapire. The Nonstochastic Multiarmed Bandit Problem. *SIAM Journal on Computing*, 32(1):48–77, 2002.

Francesco Orabona. A modern introduction to online learning. *arXiv preprint arXiv:1912.13213*, 2019.