arXiv:2601.06813v1 [cs.LG] 11 Jan 2026

分析预测准确率对分布鲁棒竞争比的影响

Toru Yoshinaga*[1] 和 Yasushi Kawase†[1]

[1]东京大学，日本

摘要

带预测的算法旨在将机器学习预测整合到算法设计中，以提升算法性能。该领域的核心问题之一是预测如何提升性能，而这一分析的关键方面在于预测准确率的作用。在此背景下，预测准确率被定义为一种有保证的概率，即从某个分布中抽取的实例属于预测集合的概率。作为一种纳入预测准确率的性能度量，我们关注 Sun 等人（ICML 2024）提出的分布鲁棒竞争比（DRCR）。DRCR 定义为算法成本与最优成本之比的期望，其中期望是针对满足给定预测和准确率要求的最坏情形实例分布计算的。一个已知的结构性质是，对于任意固定算法，DRCR 会随着预测准确率的提高而线性下降。在此基础上，我们证明最优分布鲁棒竞争比（即所有算法上的下确界）是预测准确率的单调凹函数。我们进一步将 DRCR 框架推广到多预测设置，并证明在该设置下单调性和凹性依然成立。最后，我们将结果应用于在线优化中的基准问题——滑雪租赁问题，以确定最优 DRCR 达到目标值所需的预测准确度条件。此外，我们还给出一种计算临界准确率的方法；临界准确率定义为使最优 DRCR 严格优于无准确率保证时可达到的性能所需的最低准确率。

1 引言

在线算法通常旨在保证最坏情况分析下的性能。评估算法性能的一项标准指标是竞争比，其定义为在线算法所产生的成本与最坏情况实例下最优离线成本之比。然而，一味追求良好的竞争比往往会导致过于保守，从而在实践中表现不佳。因此，我们可以利用预测来提升性能，同时保留最坏情况保证。近期的一系列研究通过带预测的算法 [Purohit 等人，2018年，Mitzenmacher 和 Vassilvitskii，2020年，Lykouris 和 Vassilvitskii， 2021]这一框架发展了这一思路。在该框架中，算法会获得有关问题实例的额外信息，这些信息通常来自机器学习预测。其目标是利用预测提升性能，同时维持最坏情况保证。

在以往大多数关于带预测的算法的研究中，算法通常通过两种性能指标进行评估：一致性，即满足预测的实例上的竞争比；以及鲁棒性，即所有实例上的竞争比。形式化地，这两项指标定义如下。设 $\mathcal{I}$ 为某个最小化问题的所有实例构成的集合。一个预测会导出一个实例

*yoshinaga-toru106@g.ecc.u-tokyo.ac.jp
†kawase@mist.i.u-tokyo.ac.jp

子集 $\Theta \subseteq \mathcal{I}$，其中的实例与该预测一致。我们将 Θ称为一致集。记算法 A 在实例 $I \in \mathcal{I}$ 上产生的期望成本为 ALG(I)，离线最优成本为 OPT(I)。算法 A的一致性和鲁棒性分别定义为

$$c(A) = \sup_{I \in \Theta} \frac{\mathrm{ALG}(I)}{\mathrm{OPT}(I)} \quad \text{and} \quad r(A) = \sup_{I \in \mathcal{I}} \frac{\mathrm{ALG}(I)}{\mathrm{OPT}(I)}.$$

这些指标仅刻画了两种极端情形，即预测完全正确和预测完全不可靠，因此忽略了有关预测准确率的信息。相比之下，机器学习预测通常附有量化的准确率保证。从算法设计中的风险管理角度来看，预测准确率至关重要：准确率高时，可以更大程度地依赖预测并采用更激进的设计；准确率低时，则需要基于最坏情况保证采取保守设计。因此，在实际应用中，有必要设计能够明确利用现有预测准确率的算法。

近年来，Sun 等人 [2024] 提出了一种纳入预测准确率的新性能指标，称为分布鲁棒竞争比（*DRCR*）。在他们的模型中，一个实例是从由一致集 Θ 和预测准确率$1-\delta$指定的分布集合中抽取的。实例分布集 $\mathcal{D}_{\Theta,\delta}$ 包含满足以下条件的所有分布：随机抽取的实例I 以至少 $1-\delta$ 的概率属于 Θ $\mathcal{I}$。采用这一记法，DRCR 定义如下：

定义 1. *(分布鲁棒竞争比 [Sun 等人, 2024])* 对于问题 $\mathcal{I}$，给定一个一致集 Θ 和预测准确率 $1-\delta$，算法 A 的 *DRCR* 定义为

$$\mathrm{DRCR}_{\Theta,\delta}(A) = \sup_{d \in \mathcal{D}_{\Theta,\delta}} \mathbb{E}_{I \sim d}\left[\frac{\mathrm{ALG}(I)}{\mathrm{OPT}(I)}\right].$$

本研究的主要目标是为在线算法在 DRCR 中达到理想性能水平所需的预测准确率奠定理论基础。进而，这些基准将为设计利用预测的算法时，构建预测所需的样本量或机器学习方法的选择提供依据。为此，我们研究预测准确率与最优 DRCR 之间的关系；最优 DRCR 是在给定预测和准确度的条件下，所有在线算法的 DRCR 的下确界。直观而言，较高的预测准确率应带来更好的性能。我们可以从 Sun 等人给出的 DRCR 凸组合表示中验证这一直觉 [2024]。

定理 1 ([Sun 等人, 2024])。 固定一个一致集 Θ。算法 A 的 *DRCR* 可以表示为其一致性与鲁棒性的凸组合，该凸组合由准确度参数化。也就是说，

$$\mathrm{DRCR}_{\Theta,\delta}(A) = (1-\delta)\cdot c(A) + \delta \cdot r(A).$$

然而，仅凭这一结果还不足以刻画预测准确率与最优 DRCR 之间的关系。预测准确率的变化可能会显著改变算法设计对预测的依赖程度，也可能大幅改变算法本身的构造。因此，预测准确率与最优 DRCR 之间的关系更加复杂，需要对这种关系进行详细分析，以考察在何种条件下预测有利于算法设计。

在对最优 DRCR 进行结构分析的同时，我们将 DRCR 框架扩展到多预测情形，以应对决策者能够获取多个预测来源的自然场景。在此设置下，我们假设多个预测及其准确度具有层次化 结构。从较高层面来看，层次结构意味着预测实例集及其相关的准确率保证是嵌套的，并且随着我们允许使用更大的实例集，准确率保证会变得更强。例如，假设我们想要预测某人一个月去健身房的次数。在这种情况下，假设我们得到以下两个预测：该人在一个月内去健身房 3–5 次的概率为 50%，以及访问次数处于（更宽的）2–10 范围内的概率为 80%。这两个预测具有层次关系。在这种层次结构下，最优 DRCR 的结构性质在多预测设置中得以保留。

1.1 我们的结果

在第3节中，我们将介绍 DRCR 关于预测准确率的若干结构性质。我们首先给出最优 DRCR 的形式化定义，并证明在单预测设置中，它是预测准确率的单调且凹函数。该性质源于定理 1 中给出的 DRCR 凸组合表示。接下来，我们证明，在存在层次化预测的情况下，DRCR 也可以表示为一致性与鲁棒性的凸组合，其参数由预测准确度决定。由这些结果可立即得到，在层次化预测设置中，最优 DRCR 关于预测准确度同样保持单调性和凹性。这些结构性质表明：如果存在一对包含足够信息、能够改善最优 DRCR 的一致集及其准确度，那么任何具有更高准确度的此类组合都能提供更加有效的信息，从而改善最优 DRCR。

在第4节中，我们提出了一种方法，用于确定所需的预测条件，以提升滑雪租赁问题的最优 DRCR；该问题是在线优化中的一个代表性基准。在我们的设置中，最后滑雪日通过层次化预测区间进行预测。首先，我们给出一个线性规划（LP）表述，用于计算该问题的最优随机化算法。然后，利用这一线性规划，我们以线性不等式组的形式描述所需的预测准确度条件，使最优 DRCR 达到期望值。最后，在单区间设置下，我们提出一种方法，用于计算最小预测准确率，使得最优 DRCR 严格优于最优鲁棒性值（即不提供任何准确率保证时的最优 DRCR）。

1.2 相关工作

经典竞争分析通过最坏情况实例来评估在线算法的性能，但这往往会导致过于悲观的保证。为了解决这一问题，研究人员探索了能够反映现实情况的模型，例如放宽对输入模型的假设，或允许利用关于未来输入的一些先验信息。Roughgarden 的教材[2021]综述了这方面的研究。具体而言，DRCR 框架可视为带预测的算法这一研究方向的一部分，因为它假设通过机器学习预测可以获得有关未来输入的部分信息。同时，它也与 Koutsoupias 和 Papadimitriou [2000], 提出的扩散对手模型 密切相关；在该模型中，对手从决策者已知的受限分布族中选取最坏情况分布。DRCR 框架可以看作是这一视角的进一步细化，其中，关于分布类的部分知识这一假设可由此类预测自然地加以解释。

滑雪租赁问题是在线优化中的一个基础问题，其众多变体已得到研究。在经典设定中，Karlin 等人 [1988, 1994] 提出了确定性算法和随机化算法，分别实现了最优竞争比。在带预测算法的研究脉络中，滑雪租赁问题的许多变体已得到研究 [Antoniadis 等人，2021年，Angelopoulos 等人，2024年，Anand 等人，2020年，Banerjee，2020年，Bhattacharya 和 Das，2022年，Bamas 等人，2020年，Benomar 和 Perchet，2023年，Diakonikolas 等人，2021年，Drygala 等人，2023年，Eichhorn 和 Krumke，2025年，Gollapudi 和 Panigrahi，2019年，Istrate 等人，2024年，Maghakian 等人，2023年，Purohit 等人，2018年，Sun 等人，2024年，Shin 等人，2023a,b，Shah 和 Rajkumar，2021年，Shen 等人，2025年，Wang 等人，2020年，Wei 和 Zhang，2020年，Zhao 等人， 2024]。在这些研究中，Shen 等人 [2025] 的工作与我们的研究（以及 Sun 等人 [2024]的研究）最为相关，因为该工作采用随机输入模型，并明确量化了预测准确率。他们将一个校准概率预测器纳入算法设计，用于预测滑雪时长是否超过购买阈值，并证明随着预测器准确度提高，竞争比会平滑改善。

2 预备知识

本节介绍用于分析具有层次化预测的 DRCR 的符号。全文中，我们使用符号 $[n] = \{1, 2, \ldots, n\}$，其中 n 是一个正整数。令 $\mathcal{I}$ 表示

一个在线最小化问题的实例集。实例 $I \in \mathcal{I}$ 从实例集 $\mathcal{I}$ 上的未知分布 d 中随机抽取。作为预测信息，我们给定一组 n 个预测（Θ_i，δ_i）$_{i\in[n]}$。每一对（Θ_i，δ_i）由一个一致集 $\Theta_i \subseteq \mathcal{I}$ 和相应的准确率参数 $\delta_i \in [0,1]$ 组成。对于每个 $i \in [n]$，预测（Θ_i，δ_i）保证未知分布 d 满足

$$\Pr_{I\sim d}[I \in \Theta_i] \geq 1 - \delta_i.$$

有时，我们希望只将一致集或只将准确率参数作为一个整体来处理。因此，我们分别用 $\boldsymbol{\Theta} := (\Theta_i)_{i\in[n]}$ 和 $\boldsymbol{\delta} := (\delta_i)_{i\in[n]}$ 表示它们。使用这一记法，一组 n预测可简写为 $(\boldsymbol{\Theta}, \boldsymbol{\delta})$，并称之为预测配置。令 $\mathcal{D}_{\boldsymbol{\Theta},\boldsymbol{\delta}}$ 表示满足上述条件的所有分布的集合。也就是说，

$$\mathcal{D}_{\boldsymbol{\Theta},\boldsymbol{\delta}} = \{d \mid \Pr_{I\sim d}[I \in \Theta_i] \geq 1 - \delta_i \text{ for all } i \in [n]\}.$$

采用这一记法，具有预测配置 $(\boldsymbol{\Theta}, \boldsymbol{\delta})$ 的算法 A 的 DRCR 定义为

$$\mathrm{DRCR}_{\boldsymbol{\Theta},\boldsymbol{\delta}}(A) = \sup_{d\in\mathcal{D}_{\boldsymbol{\Theta},\boldsymbol{\delta}}} \mathbb{E}_{I\sim d}\left[\frac{\mathrm{ALG}(I)}{\mathrm{OPT}(I)}\right].$$

在我们的设定中，我们为每个预测分别赋予一个一致性。对于每个一致集 Θ_i，我们将一致性 c_i 定义为

$$c_i(A) = \sup_{I\in\Theta_i} \frac{\mathrm{ALG}(I)}{\mathrm{OPT}(I)}.$$

此外，正如引言中所述，鲁棒性 $r(A)$ 定义为所有实例上的最坏情况性能。这里，为方便起见，我们令 $\Theta_{n+1} := \mathcal{I}$，并将 $r(A)$ 记作 $= c_{n+1}(A)$。根据定义，对于所有 $i \in [n]$，我们有 $c_i(A) \geq 1$，且 $r(A) \geq 1$。

如果预测配置 $(\boldsymbol{\Theta}, \boldsymbol{\delta})$ 满足以下条件，则称其具有层次结构：

$$\Theta_1 \subseteq \Theta_2 \subseteq \cdots \subseteq \Theta_n \quad \text{and} \quad \delta_1 \geq \delta_2 \geq \cdots \geq \delta_n.$$

如果某些连续的预测具有相同的一致集和准确率参数，即 $\Theta_i = \Theta_{i+1}$和 $\delta_i = \delta_{i+1}$，我们可以将它们合并为一个单一预测，而不影响任何性能指标。

除了前面的定义 $\Theta_{n+1} := \mathcal{I}$ 外，我们还进一步设定 $\Theta_0 = \emptyset$。相应地，集合 $\delta_0 = 1$（因为对于所有 d，$\Pr_{I\sim d}[I \in \emptyset] = 0$）以及 $\delta_{n+1} := 0$（因为对于所有 d，$\Pr_{I\sim d}[I \in \mathcal{I}] = 1$）。添加此记法不会破坏层次结构。根据层次结构所给出的包含链 $\Theta_1 \subseteq \Theta_2 \subseteq \cdots \subseteq \Theta_n \subseteq \Theta_{n+1} = \mathcal{I}$，一致性和鲁棒性满足 $1 \leq c_1(A) \leq c_2(A) \leq \cdots \leq c_n(A) \leq r(A)$。我们在证明中经常使用这一性质。

为了分析预测准确率对最优 DRCR 的影响，我们研究滑雪租赁问题，这是在线优化中的标准基准问题。在这个最小化问题中，关于滑雪者何时停止滑雪的信息对于最小化总成本至关重要。因此，我们利用层次化预测来估计最后滑雪日。作为层次化预测的自然应用场景，我们考虑嵌套预测区间。形式化地说，该问题定义如下。

定义 2（带预测区间的滑雪租赁问题）。一位初学者滑雪者计划一直滑雪，直到感到厌倦，我们将这个未知的日子记为 τ。滑雪者没有自己的滑雪板，因此必须以每天 \$1 的价格租赁，或以\$$B$的价格购买。其中，$B$ 是一个大于1的整数。第一天，滑雪者会获知购买成本 B 以及预测配置$(\boldsymbol{\Theta}, \boldsymbol{\delta})$。令实例集 $\mathcal{I}$ 为正整数集，并假设最后滑雪日 τ 服从定义在 $\mathcal{I}$ 上、满足预测条件的某个未知概率分布。每个一致集 Θ_i 都是一个区间 $[\ell_i, u_i]$，其中 ℓ_i 是正整数，而 u_i 是正整数或 $+\infty$。每项准确率保证均满足 $\Pr[\tau \in [\ell_i, u_i]] \geq 1 - \delta_i$。此处我们假设 $\ell_n \leq \cdots \leq \ell_1 \leq u_1 \leq \cdots \leq u_n$。这确保预测配置具有层次结构。我们的目标是设计一种使总成本最小化的算法。

3 DRCR 的结构

本节通过确立 DRCR 的凸组合性质，以及最优 DRCR 关于预测准确度的凹性，来分析 DRCR 的结构。

3.1 最优 DRCR 的凹性

通过利用最优 DRCR 的凸组合表示，我们证明了它关于预测准确率参数的凹性和单调性。

我们先讨论单一预测的情形，再将论证推广到层次化预测。固定一个预测 $\boldsymbol{\Theta} = \{\Theta\}$，并将预测准确率 $\boldsymbol{\delta} = \{\delta\}$ 视为变量。如定理 1 所示，对于任意算法 A，单一预测下的 DRCR 是 δ 的线性函数。将最优 DRCR 定义为关于 δ 的函数如下：

$$\mathrm{DRCR}^*(\delta) \coloneqq \inf_A \mathrm{DRCR}_{\boldsymbol{\Theta},\boldsymbol{\delta}}(A) = \inf_A \{(1-\delta)\cdot c(A) + \delta \cdot r(A)\}.$$

根据凸分析中的一个著名结论，$\mathrm{DRCR}^*(\delta)$ 是线性函数的逐点最小值，因此是凹函数（例如参见 Boyd and Vandenberghe [2004]）。此外，由于对于任意算法，$c(A) \le r(A)$ 都成立，因此 δ 的系数非负。于是，DRCR^* 是关于 δ 的非递减函数。

定理 2. 给定一个固定的一致集 Θ，最优 *DRCR* 关于 δ 是凹且单调非递减的。

3.2 层次化预测的凸组合表示

Sun 等人 [2024]证明，当给定单一预测时，DRCR 可以表示为一致性与鲁棒性的凸组合（见定理 1）。我们通过证明在给定层次化预测时也存在类似表示，拓展了这一结果。我们证明中的关键工具是 Abel 关于分部求和的引理，陈述如下。

引理 1 (Abel [1826])。对于任意序列 $\{a_i\}_{i=0}^{n+1}$ 和 $\{b_i\}_{i=0}^{n+1}$，

$$\sum_{i=1}^{n+1} a_i b_i = \sum_{i=1}^{n+1} a_i \cdot b_{n+1} - \sum_{i=1}^{n} \left(\sum_{j=1}^{i} a_j \right) (b_{i+1} - b_i).$$

利用这一引理，我们得到如下凸组合表示。

定理 3. 给定一个层次化预测剖面 $(\boldsymbol{\Theta}, \boldsymbol{\delta})$，算法 A 的 *DRCR* 可表示为

$$\mathrm{DRCR}_{\boldsymbol{\Theta},\boldsymbol{\delta}}(A) \coloneqq \sup_{d \in \mathcal{D}_{\boldsymbol{\Theta},\boldsymbol{\delta}}} \mathbb{E}_{I\sim d}\left[\frac{\mathrm{ALG}(I)}{\mathrm{OPT}(I)}\right] = \sum_{i=1}^{n} (\delta_{i-1} - \delta_i)\cdot c_i(A) + \delta_n \cdot r(A).$$

证明。对于每个 $i \in [n+1]$，令 $\mathbb{1}_i$ 表示子集 $\Theta_i \setminus \Theta_{i-1} \subseteq \mathcal{I}$的示性函数。此外，定义 $p_i^{(d)} = \mathbb{E}_{I\sim d}[\mathbb{1}_i(I)]$ 为从未知分布 d 中抽取的实例 I 属于 $\Theta_i \setminus \Theta_{i-1}$的概率。那么，对于每个 $i \in [n+1]$，总和 $\sum_{j=1}^{i} p_j^{(d)}$ 表示从d 中抽取的实例 I 属于 Θ_i的概率；当 $d \in \mathcal{D}_{\boldsymbol{\Theta},\ \boldsymbol{\delta}}$时，该概率至少为 $1-\delta_i$。

我们首先给出 DRCR 的上界。DRCR 可表示如下 :

$$\begin{aligned}
\mathrm{DRCR}_{\boldsymbol{\Theta},\boldsymbol{\delta}}(A) &= \sup_{d\in\mathcal{D}_{\boldsymbol{\Theta},\boldsymbol{\delta}}} \mathbb{E}_{I\sim d}\left[\frac{\mathrm{ALG}(I)}{\mathrm{OPT}(I)}\right] \\
&= \sup_{d\in\mathcal{D}_{\boldsymbol{\Theta},\boldsymbol{\delta}}} \mathbb{E}_{I\sim d}\left[\sum_{i=1}^{n+1}\mathbb{1}_i(I)\cdot\frac{\mathrm{ALG}(I)}{\mathrm{OPT}(I)}\right] \\
&\le \sup_{d\in\mathcal{D}_{\boldsymbol{\Theta},\boldsymbol{\delta}}} \mathbb{E}_{I\sim d}\left[\sum_{i=1}^{n+1}\mathbb{1}_i(I)\cdot c_i(A)\right] \\
&= \sup_{d\in\mathcal{D}_{\boldsymbol{\Theta},\boldsymbol{\delta}}} \left[\sum_{i=1}^{n+1}\mathbb{E}_{I\sim d}[\mathbb{1}_i(I)]\cdot c_i(A)\right] \\
&= \sup_{d\in\mathcal{D}_{\boldsymbol{\Theta},\boldsymbol{\delta}}} \left[\sum_{i=1}^{n+1}p_i^{(d)}\cdot c_i(A)\right].
\end{aligned}$$

应用引理 1，得到

$$\begin{aligned}
\sup_{d\in\mathcal{D}_{\boldsymbol{\Theta},\boldsymbol{\delta}}}\left[\sum_{i=1}^{n+1}p_i^{(d)}\cdot c_i(A)\right] &= \sup_{d\in\mathcal{D}_{\boldsymbol{\Theta},\boldsymbol{\delta}}}\left[\sum_{j=1}^{n+1}p_j^{(d)}c_{n+1}(A)-\sum_{i=1}^{n}\left(\sum_{j=1}^{i}p_j^{(d)}\right)(c_{i+1}(A)-c_i(A))\right] \\
&\le c_{n+1}(A)-\sum_{i=1}^{n}(1-\delta_i)\cdot(c_{i+1}(A)-c_i(A)) \\
&= c_{n+1}(A)-\sum_{i=2}^{n+1}(1-\delta_{i-1})\cdot c_i(A)+\sum_{i=1}^{n}(1-\delta_i)\cdot c_i(A) \\
&= \sum_{i=1}^{n}(\delta_{i-1}-\delta_i)\cdot c_i(A)+\delta_n\cdot r(A),
\end{aligned}$$

其中，不等式是根据以下事实得出的： $\sum_{j=1}^{i}p_j^{(d)}\ge 1-\delta_i$ 和 $c_i\le c_{i+1}$。

接下来，我们给出 DRCR 的下界。设 ε为正实数。对于每个 $i\in[n+1]$，令 $I_i^{(\varepsilon)}\in\Theta_i$为一个满足以下条件的实例

$$\mathrm{ALG}(I_i^{(\varepsilon)})/\mathrm{OPT}(I_i^{(\varepsilon)})\ge(1-\varepsilon)\cdot c_i(A).$$

这样的 $I_i^{(\varepsilon)}$ 总是存在，因为 $c_i(A)=\sup_{I\in\Theta_i}\mathrm{ALG}(I)/\mathrm{OPT}(I)$。令 $d^{(\varepsilon)}$ 表示如下分布：对于每个 $i\in[n+1]$，以概率 $\delta_{i-1}-\delta_i$ 选择$I^{(\varepsilon)}$。不难验证，$d^{(\varepsilon)}\in\mathcal{D}_{\boldsymbol{\Theta},\ \boldsymbol{\delta}}$。于是，我们有

$$\begin{aligned}
\mathrm{DRCR}_{\boldsymbol{\Theta},\boldsymbol{\delta}}(A) &= \sup_{d\in\mathcal{D}_{\boldsymbol{\Theta},\boldsymbol{\delta}}}\mathbb{E}_{I\sim d}\left[\frac{\mathrm{ALG}(I)}{\mathrm{OPT}(I)}\right] \\
&\ge \mathbb{E}_{I\sim d^{(\varepsilon)}}\left[\frac{\mathrm{ALG}(I)}{\mathrm{OPT}(I)}\right] \\
&= \sum_{i=1}^{n+1}\frac{\mathrm{ALG}(I_i^{(\varepsilon)})}{\mathrm{OPT}(I_i^{(\varepsilon)})}\cdot(\delta_{i-1}-\delta_i) \\
&\ge \sum_{i=1}^{n+1}(1-\varepsilon)\cdot c_i(A)\cdot(\delta_{i-1}-\delta_i) \\
&= (1-\varepsilon)\cdot\left(\sum_{i=1}^{n}(\delta_{i-1}-\delta_i)\cdot c_i(A)+\delta_n\cdot r(A)\right).
\end{aligned}$$

取极限 $\varepsilon \to +0$ 即可得到所需的下界。 □

由此结果可知，最优 DRCR 关于 $\boldsymbol{\delta}$是非递增且凹的。这是因为对于任意固定的算法 A，DRCR 关于 δ_i 是线性的，而 δ_i 的系数为 $c_{i-1}(A) - c_i(A)$，该系数非负。

4 应用：滑雪租赁问题

本节将第3节中得到的结构性质应用于具有层次化预测区间的滑雪租赁问题。回顾一下，在该问题中，每天的租赁成本为 1，购买成本为整数 B。实例集合 $\mathcal{I}$ 由所有正整数组成。预测配置 $(\boldsymbol{\Theta}, \boldsymbol{\delta})$ 为每个 $i \in [n]$指定区间 $\Theta_i = [\ell_i, u_i]$，且这些区间满足 $\ell_n \le \cdots \le \ell_1 \le u_1 \le \cdots \le u_n$。

在第4.1节中，我们介绍一种基于线性规划的方法，用于计算具有层次化预测区间的滑雪租赁问题的最优随机化算法。在第4.2节中，我们基于第4.1节中的线性规划，给出一种计算所需预测准确率的方法，使最优 DRCR 达到期望值。在第4.3节中，我们利用第4.2节给出的方法，推导出预测准确率必须满足的条件；在单区间情形下，只有满足该条件，最优 DRCR 才会严格优于鲁棒性值。

4.1 最优算法的线性规划表述

我们首先介绍一种方法，用于计算具有层次化预测区间的滑雪租赁问题的最优随机化算法。我们注意到，在只有一个预测区间的特殊情形下，Sun 等人 [2024]已经给出了求解该问题最优随机化算法的方法。我们将他们的框架扩展到层次化预测区间，并在这一更一般的设定下推导出最优随机化算法。该问题的任何随机化算法都可以表示为购买日分布，即在第 t 天进行购买的分布。我们用 f_t 表示在第 t 天购买滑雪板的概率。如果滑雪天数为 τ，那么在第 $t \le \tau$ 天购买的成本为（$t-1$）$+B$（截至第 $t-1$ 天租赁，然后在第 t 天购买）；而在第 $t > \tau$ 天购买则意味着截至第 τ 天都不会购买，因此成本为 τ。因此，如果滑雪天数为 τ，期望成本为

$$\sum_{t=1}^{\tau} (t - 1 + B) f_t + \tau \left(1 - \sum_{t=1}^{\tau} f_t\right).$$

我们的目标是设计一种期望成本较低的算法。为此，我们采用 DRCR 作为优化目标，并旨在针对给定的预测配置，计算出使 DRCR 最小化的算法。

在讨论主要结果之前，我们先说明，任何以正概率永远租滑雪板的随机化算法都可以排除在考虑范围之外。这类算法可以表示为设定 $f_\infty > 0$，其中 f_∞ 表示算法永远不购买滑雪板的概率。注意，$f_\infty = 1 - \sum_{t \in \mathcal{I}} f_t$。因此，当且仅当 $\sum_{t \in \mathcal{I}} f_t < 1$ 时，$f_\infty > 0$。

如果 $\delta_n = 0$，那么滑雪天数最多为 u_n。因此，任何可能永远租滑雪板的算法，都等价于以相同概率在第 $u_n + 1$ 天购买的算法。因此，在这种情况下，存在一个满足 $f_\infty = 0$ 的最优算法。

现在，假设 $\delta_n > 0$ 并固定任意一个满足 $f_\infty > 0$的算法。考虑一个实例分布：以概率 δ_n抽取 $\lceil 2B/(f_\infty \delta_n) \rceil \in \mathcal{I}$；该分布属于$\mathcal{D}_{\boldsymbol{\Theta},\, \boldsymbol{\delta}}$。在抽取到$\lceil 2B/(f_\infty \delta_n) \rceil$ 这一事件下，算法一直租用到最后的概率至少为 f_∞，因此其成本至少为 $\lceil 2B/(f_\infty \delta_n) \rceil$。因此，该算法的 DRCR 至少为

$$\delta_n \cdot \frac{f_\infty \cdot \lceil 2B/(f_\infty \delta_n) \rceil}{\min\{\lceil 2B/(f_\infty \delta_n) \rceil, B\}} \ge \delta_n \cdot \frac{f_\infty \cdot 2B/(f_\infty \delta_n)}{B} = 2.$$

另一方面，在第 B 天购买滑雪板的确定性算法具有 $2 - 1/B$ [Karlin 等人， 1988] 的鲁棒性。因此，没有任何最优随机化算法能够满足 $f_\infty > 0$。

我们现在引入一个无限维线性规划，然后证明它可以化简为一个具有有限个变量和约束的线性规划。在下面构造的无限维线性规划中，我们将一致性c_1，c_2，……，c_{n+1}（其中 c_{n+1} 表示鲁棒性）视为变量，并将 $(f_t)_{t\in\mathcal{I}}$ 一并纳入。对于每个 $\tau \in \mathcal{I}$，令 $i(\tau)$ 为满足 $\tau \in \Theta_i \setminus \Theta_{i-1}$ 的唯一索引 $i \in [n+1]$，其中 $\Theta_0 = \emptyset$ 且 $\Theta_{n+1} = \mathcal{I}$。回顾一下，若滑雪天数为 τ，则算法的期望成本和离线最优成本分别为$\sum_{t=1}^{\tau}(t-1+B)f_t + \tau(1-\sum_{t=1}^{\tau} f_t)$ 和 $\min\{\tau, B\}$。因此，该算法在一致集 $\Theta_i\,(i \in [n+1])$ 下的一致性为

$$\begin{aligned} c_i &= \max_{\tau\in\Theta_i} \frac{\sum_{t=1}^{\tau}(t-1+B)f_t + \tau(1-\sum_{t=1}^{\tau} f_t)}{\min\{\tau, B\}} \\ &= \max\left\{c_{i-1}, \max_{\tau\in\Theta_i\setminus\Theta_{i-1}} \frac{\sum_{t=1}^{\tau}(t-1+B)f_t + \tau(1-\sum_{t=1}^{\tau} f_t)}{\min\{\tau, B\}}\right\}, \end{aligned} \tag{1}$$

其中 c_0 定义为 1。因此，根据定理 3，DRCR 可表示为 $\sum_{i=1}^{n+1}(\delta_{i-1}-\delta_i)\cdot c_i$。采用这一记法，寻找使 DRCR 最小化的在线算法这一问题可化简为以下线性规划：

$$\begin{aligned} \min \quad & \sum_{i=1}^{n+1}(\delta_{i-1}-\delta_i)\cdot c_i && && (2)\\ \text{s.t.} \quad & \textstyle\sum_{t=1}^{\tau}(t-1+B)f_t + \tau\left(1-\sum_{t=1}^{\tau} f_t\right) \le \min\{\tau, B\}\cdot c_{i(\tau)} && (\tau\in\mathcal{I}), && (2a)\\ & c_i \le c_{i+1} && (i\in[n]), && (2b)\\ & \textstyle\sum_{t\in\mathcal{I}} f_t = 1, && && (2c)\\ & f_t \ge 0\ (t\in S), \quad c_i \ge 0\ (i\in[n+1]). \end{aligned}$$

约束 (2c) 与对所有 $t \in \mathcal{I}$ 均成立的 $f_t \ge 0$ 一起，确保 $(f_t)_{t\in\mathcal{I}}$ 构成关于购买日的有效概率分布。

现在证明，线性规划 (2) 的最优解确实给出了一个使 DRCR 最小化的在线算法。对于任意随机化算法 $(f_t)_{t\in\mathcal{I}}$ 及其在 (1) 中定义的一致性 $(c_i)_{i\in[n+1]}$，数对 $((f_t)_{t\in\mathcal{I}}, (c_i)_{i\in[n+1]})$ 满足 (2a) 和 (2b) 中的约束。目标函数为 $\sum_{i=1}^{n+1}(\delta_{i-1}-\delta_i)c_i$，根据定理 3，它等于该随机化算法的 DRCR。因此，线性规划 (2) 的最优值不大于最小 DRCR 值。反过来，对于任意可行解 $((f_t)_{t\in\mathcal{I}}, (c'_i)_{i\in[n+1]})$，相应的随机化算法 $(f_t)_{t\in\mathcal{I}}$ 可实现不超过 $\sum_{i=1}^{n+1}(\delta_{i-1}-\delta_i)c'_i$ 的 DRCR。为此，定义 $(c_i)_{i\in[n+1]}$ 为该算法所诱导的一致性。我们用归纳法证明，对所有 $i \in [n+1]$，$c'_i \ge c_i$ 均成立。令 $c'_0 = c_0 = 1$。假设对于某个 $i \in [n+1]$，$c'_{i-1} \ge c_{i-1}$ 成立，则

$$\begin{aligned} c'_i &\ge \max\left\{c'_{i-1}, \max_{\tau\in\Theta_i\setminus\Theta_{i-1}} \frac{\sum_{t=1}^{\tau}(t-1+B)f_t + \tau(1-\sum_{t=1}^{\tau} f_t)}{\min\{\tau, B\}}\right\} \\ &\ge \max\left\{c_{i-1}, \max_{\tau\in\Theta_i\setminus\Theta_{i-1}} \frac{\sum_{t=1}^{\tau}(t-1+B)f_t + \tau(1-\sum_{t=1}^{\tau} f_t)}{\min\{\tau, B\}}\right\} = c_i, \end{aligned}$$

其中，第一个不等式由约束 (2a) 和 (2b) 得出，第二个不等式根据归纳假设成立，最后一个等式则由 (1) 得出。因此，$c'_i \ge c_i$ 对所有 $i \in [n+1]$ 均成立，且 DRCR $\sum_{i=1}^{n+1}(\delta_{i-1}-\delta_i)c_i$ 至多为 $\sum_{i=1}^{n+1}(\delta_{i-1}-\delta_i)c'_i$。因此，线性规划 (2) 的最优值等于可达到的最小 DRCR 值。此外，对于线性规划的任意一个最优解，相应的在线算法都能达到最小 DRCR。

为使线性规划 (2) 在计算上可处理，我们利用问题的结构推导出一个等价的有限维线性规划。这是必要的，因为线性规划 (2) 具有无穷多个变量和约束，因此通常无法使用标准线性规划求解器。该约化依赖于以下结构性质：至少存在一个最优购买日分布，其支撑集（即购买概率为正的日期集合）可以限制在由预测 $\boldsymbol{\Theta}$ 和购买成本 B 确定的特定日期子集中，且该子集的大小至多为 $B+2n+1$。我们现在将这样一个受限日期子集记为 S。此外，

我们将 T 记为约化线性规划中需要考虑的实例子集。具体而言， S 和 T 定义如下：

- 若 $B < u_n$，令 $j = \min\{i \in [n] \mid u_i > B\}$。定义 $S := \{1, 2, \ldots, B\} \cup \{u_j + 1, u_{j+1} + 1, \ldots, u_n + 1\}$ 和 $T := \{1, 2, \ldots, B\} \cup \{\ell_i - 1,\, u_i \mid i \in [n]\} \cup \{u_n + 1\}$。
- 若 $u_n \le B$，定义 $S := [B]$ 和 $T := [B+1]$。

基于此记号，我们定义如下有限维线性规划：

$$
\begin{aligned}
\min \quad & \sum_{i=1}^{n+1} (\delta_{i-1} - \delta_i) \cdot c_i && && (3)\\
\text{s.t.} \quad & \textstyle\sum_{t \in S:\, t \le \tau} (t - 1 + B) \cdot f_t + \tau \left(1 - \sum_{t \in S:\, t \le \tau} f_t\right) \le \min\{\tau, B\} \cdot c_{i(\tau)} && (\tau \in T), && (x_\tau)\\
& c_i \le c_{i+1} && (i \in [n]), && (y_i)\\
& \textstyle\sum_{t \in S} f_t = 1, && && (z)\\
& f_t \ge 0 \ (t \in S), \quad c_i \ge 0 \ (i \in [n+1]). &&
\end{aligned}
$$

定理 **m 4.** 无限维线性规划(2) 可以化简为有限维线性规划(3)。

证明推迟至附录 A.1。因此，通过求解线性规划 (3)，可以得到具有层次化预测区间的滑雪租赁问题的最优随机化算法，该线性规划有 $O(B+n)$ 个变量和约束。此线性规划可以使用标准线性规划求解器求解（例如，基于内点法的求解器 [Karmarkar, 1984]）。

4.2 预测准确度的多面体刻画

在上一节中，我们构造了一个线性规划，用于计算在具有层次化预测区间的滑雪租赁问题中实现最优 DRCR 的随机化算法。接下来一个自然的问题是：哪些预测概况能够带来更优的最优 DRCR 值？作为回答这一问题的初步步骤，我们刻画了这样一组预测准确度：在该组预测准确度下，某个算法能够达到规定的性能水平。具体而言，通过固定层次化预测区间 $\boldsymbol{\Theta} = (\Theta_i)_{i \in [n]}$ 和目标最优 DRCR 值 v，我们通过一个线性不等式组刻画预测准确度 $\boldsymbol{\delta} = (\delta_i)_{i \in [n]}$ 的集合；对于该集合中的预测准确度，最优 DRCR 至少为 v。

我们基于线性规划 (3) 的对偶构造线性不等式组。采用这种基于对偶的方法有两个原因。首先，在最小化问题中，很难直接表达目标函数的最优值至少为 v 这一条件；而在对偶（最大化）问题中，只需要求目标值至少为 v，即可直接表达该条件。其次，如果基于原问题构造该不等式组，并将 $\boldsymbol{\delta}$ 作为决策变量，则项 $\sum_{i=1}^{n+1} (\delta_{i-1} - \delta_i) \cdot c_i$ 包含变量的乘积，从而使该系统变为非线性。相比之下，基于对偶问题的构造能够避免这一问题，并使不等式组保持线性。

我们将原问题约束的标签（即 $(x_\tau)_{\tau \in T}$、$(y_i)_{i \in [n]}$ 和 z）作为对偶变量。线性规划 (3) 的对偶问题如下：

$$
\begin{aligned}
\max \quad & z + \sum_{\tau \in T} \tau \cdot x_\tau && && (4)\\
\text{s.t.} \quad & z + \textstyle\sum_{\tau \in T:\, \tau \ge t} (\tau - (t - 1 + B)) \cdot x_\tau \le 0 && (t \in S), && (f_t)\\
& \textstyle\sum_{\tau \in T \cap (\Theta_i \setminus \Theta_{i-1})} \min\{\tau, B\} \cdot x_\tau + y_{i-1} - y_i \le \delta_{i-1} - \delta_i && (i \in [n+1]), && (c_i)\\
& x_\tau \ge 0 \ (\tau \in T), \quad y_i \ge 0 \ (i \in [n]). &&
\end{aligned}
$$

其中 y_0 和 $y_{n+1} = 0$ 设为 0。

然后，我们去掉线性规划 (4) 的目标函数，转而施加目标值至少为给定目标值（即 $v \le z + \sum_{\tau\in T} \tau \cdot x_\tau$）的约束，从而得到一个线性不等式组。在此操作下， z 实际上成为一个辅助变量，我们可以通过组合约束 (f_t) 将其从该系统中消去。此外，我们将 $(\delta_i)_{i\in[n]}$ 视为决策变量，并添加由层级预测结构所蕴含的约束 $1 \ge \delta_1 \ge \delta_2 \ge \cdots \ge \delta_n \ge 0$。由此得到一个线性不等式组，用于刻画最优 DRCR 至少为 v 时预测准确度的可行域 $(\delta_i)_{i\in[n]}$ 。形式上，该系统为

$$\begin{cases} v \le \sum_{\tau\in T:\,\tau<t} \tau x_\tau - (t-1+B)\sum_{\tau\in T:\,\tau\ge t} x_\tau & (t \in S), \\ \sum_{\tau\in T\cap(\Theta_i\setminus\Theta_{i-1})} \min\{\tau, B\}\cdot x_\tau + y_{i-1} - y_i \le \delta_{i-1} - \delta_i & (i \in [n+1]), \\ 1 = \delta_0 \ge \delta_1 \ge \delta_2 \ge \cdots \ge \delta_n \ge \delta_{n+1} = 0, \\ x_\tau \ge 0 \ (\tau \in T), \quad y_i \ge 0 \ (i \in [n]), \quad y_0 = y_{n+1} = 0. \end{cases} \tag{5}$$

接下来，我们得到以下定理。

定理 5. 固定层次化预测区间 $\boldsymbol{\Theta} = (\Theta_i)_{i\in[n]}$ 和目标 *DRCR* 值 v。那么，在预测配置 $(\boldsymbol{\Theta}, \boldsymbol{\delta})$ 下，使最优 *DRCR* 至少为 v 的预测准确度集合 $\boldsymbol{\delta} = (\delta_i)_{i\in[n]}$ 由线性系统 (5)刻画。

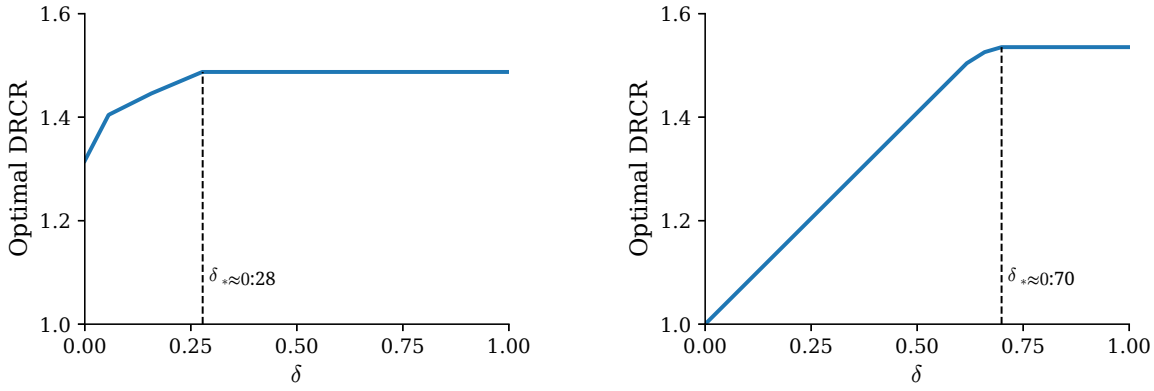
4.3 结构与刻画的应用

我们整合前几节得到的结果，推导出刻画滑雪租赁问题中一类理想预测配置的条件。为简便起见，我们聚焦于单预测区间设置，其中最后滑雪日 τ 以至少 $1-\delta$的概率属于区间 $[\ell, u]$ ，并记 $\boldsymbol{\Theta} = \{\Theta = [\ell, u]\}$, $\boldsymbol{\delta} = \{\delta\}$。

首先，我们通过实验观察最优 DRCR 值如何随预测准确率变化。图1展示了两个问题实例中 δ 与最优 DRCR 之间的关系。这些图可以通过求解线性规划 (3) 得到，即针对每个 δ求出最优 DRCR 值。选取这两个实例作为代表性示例；对于其他 B、 ℓ 和 u的取值，也可以观察到类似的定性行为。从图中可以确认，最优 DRCR 关于 δ具有单调性（非递减）和凹性，这与第3节的结果一致。此外，我们可以观察到 δ 的一个阈值，在该阈值处，最优 DRCR 值不再增加，并与最优鲁棒性值相同。这表明，即使提供了正的准确率保证，也存在一些无法改善最优 DRCR 的预测配置；由此自然引出一个问题：哪些预测配置能使最优 DRCR 严格优于鲁棒性值?

为回答这一问题，我们提出一种方法：给定 ℓ、 u 和 B，计算使最优 DRCR 严格小于最优鲁棒性值所需的最低预测准确率。我们将这一阈值称为临界准确率 ，并在下文中记为 δ^* 。形式化地，固定 ℓ、 u 和 B 后，临界准确率 δ^* 定义为使最优 DRCR 值与最优鲁棒性值相等的 δ 的最小值。我们的方法基于系统（5）；对于固定的预测集 $\boldsymbol{\Theta}$ 和目标值 v，该系统刻画了使最优 DRCR 至少为 v 的预测准确度 $\boldsymbol{\delta}$ 的集合。为计算 δ^*，我们将目标值 v 设为最优鲁棒性值，并在满足系统（5）的约束下最小化 δ 。为确定 v 的值，我们注意到鲁棒性和竞争比的定义相同。因此，最优鲁棒性值等于经典滑雪租赁问题的最优随机化算法所达到的竞争比，该值由 Karlin [1998] 给出，为

$$\frac{B^B}{B^B - (B-1)^B}.$$



(a) $\ell = 3$、 $u = 8$和 $B = 5$

(b) $\ell = 3$、 $u = 8$和 $B = 10$

图1：最优分布鲁棒竞争比与 δ之间关系的示例。

综上所述，用于计算临界准确率的线性规划如下：

$$
\begin{aligned}
\min \quad & \delta \\
\text{s.t.} \quad & \tfrac{B^B}{B^B-(B-1)^B} \le \textstyle\sum_{\tau\in T:\,\tau<t} \tau x_\tau + (t-1+B)\sum_{\tau\in T:\,\tau\ge t} x_\tau \quad (t\in S), \\
& \textstyle\sum_{\tau\in T\cap\Theta} \min\{\tau,B\}\cdot x_\tau - y \le 1-\delta, \\
& \textstyle\sum_{\tau\in T\cap(\mathcal{I}\setminus\Theta)} \min\{\tau,B\}\cdot x_\tau + y \le \delta, \\
& x_\tau \ge 0 \ (\tau\in T), \quad y\ge 0, \quad 0\le\delta\le 1.
\end{aligned}
\tag{6}
$$

这里，我们可以选择 $S = \{1, 2, \ldots, B, u+1\}$ 和 $T = \{1, 2, \ldots, B\} \cup \{\ell-1, u, u+1\}$。因此，该线性规划的规模是多项式级的，可以在多项式时间内求解。

5 结论

本文分析了 DRCR 与预测准确率之间关系的若干结构性质，并进一步考虑了向层次化预测的扩展。作为对所得结构性质的应用，我们研究了带预测区间的滑雪租赁问题，并推导出一个线性规划，用以给出最优 DRCR 严格小于最优鲁棒性值的条件。

未来值得探究的是，我们用于计算 δ^* 的方法能否推广到其他在线优化问题；这类问题的最优算法可通过线性规划形式计算。此外，还需要更深入地研究如何根据给定数据构建预测，才能有效地最小化 DRCR。

此外，还应考虑层次化预测在实践中的局限性。尽管层次化预测在理论上具有有利的结构，但现实中的预测器，尤其是学习模型，能否生成此类预测仍不确定。因此，仔细考察层级假设的有效性及其相关局限性十分重要。

参考文献

N.H. Abel. 关于级数的研究： $1 + \frac{m}{1}x + \frac{m_{(m-1)}}{1\cdot 2}\ x^2 + \cdots$. 纯粹与应用数学杂志 *¨ur die reine und angewandte Mathematik*, 1:311–339, 1826.

Keerti Anand、Rong Ge 和 Debmalya Panigrahi。为在线算法定制机器学习预测。载于国际机器学习会议，第 303–313 页。PMLR，2020年。

Spyros Angelopoulos、Christoph D¨urr、Shendan Jin、Shahin Kamali 和 Marc Renault。不可信建议下的在线计算。计算机与系统科学杂志，144:103545，2024年。Antonios Antoniadis、Christian Coester、Marek Eli´as、Adam Polak 和 Bertrand Simon。通过新的滑雪租赁界，利用学习增强的多状态动态功率管理。载于神经信息处理系统进展，第 34 卷，第 16714–16726 页，2021年。Etienne Bamas、Andreas Maggiori 和 Ola Svensson。用于学习增强算法的原始-对偶方法。载于神经信息处理系统进展，第 33 卷，第 20083–20094 页，2020年。Shom Banerjee。结合序贯决策和机器学习预测，改进在线租赁或购买算法。载于神经信息处理系统进展，第 21072–21080 页，2020年。Ziyad Benomar 和 Vianney Perchet。在线算法中预算约束下的建议查询。载于神经信息处理系统进展，第 36 卷，第 75026–75047 页，2023年。Arghya Bhattacharya 和 Rathish Das。带折扣的机器学习建议滑雪租赁问题。载于算法与计算国际会议及研讨会，第 213–224 页。施普林格，2022年。Stephen P Boyd 和 Lieven Vandenberghe。凸优化。剑桥大学出版社，2004。Ilias Diakonikolas、Vasilis Kontonis、Christos Tzamos、Ali Vakilian 和 Nikos Zarifis。利用分布式建议学习在线算法。载于国际机器学习会议，第 2687– 2696 页。PMLR，2021年。Marina Drygala、Sai Ganesh Nagarajan 和 Ola Svensson。具有高成本预测的在线算法。载于国际人工智能与统计会议，第 8078–8101 页。PMLR，2023年。Daniel Eichhorn 和 Sven O Krumke。利用预测在单条列车线路上进行在线延误管理。信息处理快报，190:106569，2025年。Sreenivas Gollapudi 和 Debmalya Panigrahi。结合专家建议的在线租赁或购买算法。载于国际机器学习会议，第 2319–2327 页。PMLR，2019年。Gabriel Istrate、Cosmin Bonchi¸s 和 Victor Bogdan。具有预-测的多智能体在线问题中的均衡。*arXiv* 预印本 *arXiv:2405.11873*，2024年。Anna R. Karlin。竞争算法在实践中的表现，第 373–384 页。施普林格·柏林·海德堡，1998年。Anna R Karlin、Mark S Manasse、Larry Rudolph 和 Daniel D Sleator。竞争性窥探缓存。*Algorithmica*，3(1):79–119，1988年。Anna R Karlin、Mark S Manasse、Lyle A McGeoch 和 Susan Owicki。非均匀问题的竞争性随机算法。*Algorithmica*，11(6):542–571，1994年。Narendra Karmarkar。线性规划的一种新型多项式时间算法。载于第十六届 *ACM* 计算理论研讨会论文集，第 302–311 页，1984年。Elias Koutsoupias 和 Christos H Papadimitriou。超越竞争分析。*SIAM* 计算杂志，30(1):300–317，2000年。Thodoris Lykouris 和 Sergei Vassilvitskii。结合机器学习建议的竞争性缓存。美国计算机学会期刊（*JACM*），68(4):1–25，2021年。Jessica Maghakian、Russell Lee、Mohammad Hajiesmaili、Jian Li、Ramesh Sitaraman 和 Zhenhua Liu。具有异构预测器的应用型在线算法。载于国际机器学习会议，第 23484–23497 页。PMLR，2023年。

Michael Mitzenmacher 和 Sergei Vassilvitskii。带预测的算法。载于 Tim Roughgarden（编），超越算法的最坏情况分析，第 646–662 页。剑桥大学出版社，2020年。Manish Purohit、Zoya Svitkina 和 Ravi Kumar。通过机器学习预测改进在线算法。载于神经信息处理系统进展，第 31 卷，2018年。Tim Roughgarden。超越算法的最坏情况分析。剑桥大学出版社，2021年。Anant Shah 和 Arun Rajkumar。序贯滑雪租赁问题。载于第 *20* 届国际自主代理与多代理系统会议论文集，第 1173–1181 页，2021年。Judy Hanwen Shen、Ellen Vitercik 和 Anders Wikum。具有校准机器学习预测的算法。*arXiv* 预印本 *arXiv:2502.02861*，2025年。Yongho Shin、Changyeol Lee 和 Hyung-Chan An。学习增强型多选项滑雪租赁的一致性-鲁棒性最优权衡。*arXiv* 预印本 *arXiv:2312.02547*，2023a。Yongho Shin、Changyeol Lee、Gukryeol Lee 和 Hyung-Chan An。通过最优可能竞争分析改进多选项滑雪租赁问题的学习增强算法。载于国际机器学习会议，第 31539–31561 页。PMLR，2023b。Bo Sun、Jerry Huang、Nicolas Christianson、Mohammad Hajiesmaili、Adam Wierman 和 Raouf Boutaba。具有不确定性量化预测的在线算法。载于第 *41* 届国际机器学习会议论文集，第 47056–47077 页，2024年。Shufan Wang、Jian Li 和 Shiqiang Wang。利用机器学习建议的多商店滑雪租赁在线算法。载于神经信息处理系统进展，第 33 卷，第 8150–8160 页，2020年。Alexander Wei 和 Fred Zhang。学习增强型在线算法的鲁棒性-一致性最优权衡。载于神经信息处理系统进展，第 33 卷，第 8042–8053 页，2020年。Hailiang Zhao、Xueyan Tang、Peng Chen 和 Shuiguang Deng。Bahncard问题的学习增强算法。神经信息处理系统进展，37:113234–113281，2024年。

A 推迟的证明

A.1 定理 4 的证明

证明。为从线性规划 (2) 中消去冗余变量和约束，我们首先对可行解引入一种变换。该操作将一个可行解 $(\boldsymbol{f},\ \boldsymbol{c}) = ((f_t)_{t\in\mathcal{I}},\ (c_i)^{i\in[n+1]})$ 映射为另一个可行解 $(\hat{}\boldsymbol{f},\ \boldsymbol{c}) = ((\hat{}f_t)_{t\in\mathcal{I}},\ (c_i)^{i\in[n+1]})$。由于一致性 $\boldsymbol{c}$ 保持不变，该操作保留了目标值。令 C_τ 表示 (2a) 中索引为 $\tau\in\mathcal{I}$ 的约束。对于可行解 $(\boldsymbol{f},\ \boldsymbol{c})$，我们将 $\mathrm{LHS}(\boldsymbol{f},\ \tau)$ 和 $\mathrm{RHS}(\boldsymbol{c},\ \tau)$ 分别记为 C_τ 左端项和右端项的值。

我们注意到，如果存在某一天 $\tau^*\in\mathcal{I}$，使得

$$\mathrm{RHS}(\boldsymbol{c},\tau^*)\geq\mathrm{RHS}(\boldsymbol{c},\tau^*+1), \tag{7}$$

那么可以构造另一个可行解，使得在第 τ^*+1 天购买的概率为零，且目标值不会变差。具体而言，我们考虑一个解 $(\boldsymbol{f},\ \boldsymbol{c})$，其中

$$\hat{f}_t=\begin{cases} f_{\tau^*}+f_{\tau^*+1} & \text{if } t=\tau^*,\\ 0 & \text{if } t=\tau^*+1,\\ f_t & \text{otherwise}\end{cases}\qquad (t\in\mathcal{I}). \tag{8}$$

现在我们验证 $(\hat{}\boldsymbol{f},\ \boldsymbol{c})$ 的可行性。除 (2a) 中的约束外，其他约束的可行性显然成立。因此，只需验证 (2a) 中约束的可行性。我们通过考虑以下三种情况来进行验证。

- 对于 $\tau\leq\tau^*-1$，对于所有 $\hat{t}\leq\tau$，我们有 $f_t=f_t$。因此，我们有

$$\mathrm{LHS}(\hat{\boldsymbol{f}},\tau)=\mathrm{LHS}(\boldsymbol{f},\tau)\leq\mathrm{RHS}(\boldsymbol{c},\tau).$$

- 对于 $\tau\geq\tau^*+1$，我们有

$$\begin{aligned}\mathrm{LHS}(\hat{\boldsymbol{f}},\tau)-\mathrm{LHS}(\boldsymbol{f_t},\tau)&=(\tau^*-1+B)(f_{\tau^*}+f_{\tau^*+1})-(\tau^*-1+B)f_{\tau^*}-(\tau^*+B)f_{\tau^*+1}\\&=-f_{\tau^*+1}\leq 0.\end{aligned}$$

  因此，我们得到

$$\mathrm{LHS}(\hat{\boldsymbol{f}},\tau)\leq\mathrm{LHS}(\boldsymbol{f},\tau)\leq\mathrm{RHS}(\boldsymbol{c},\tau).$$

- 对于 $\tau=\tau^*$，我们有

$$\mathrm{LHS}(\hat{\boldsymbol{f}},\tau)-\mathrm{LHS}(\hat{\boldsymbol{f}},\tau+1)=\tau^*\sum_{t\geq\tau^*+2}f_t-(\tau^*+1)\sum_{t\geq\tau^*+2}f_t=-\sum_{t\geq\tau^*+2}f_t\leq 0.$$

  因此，我们得到

$$\mathrm{LHS}(\hat{\boldsymbol{f}},\tau)\leq\mathrm{LHS}(\hat{\boldsymbol{f}},\tau+1)\leq\mathrm{RHS}(\boldsymbol{c},\tau+1)\leq\mathrm{RHS}(\boldsymbol{c},\tau),$$

  其中，第二个不等式由对情形 $\tau\geq\tau^*+1$ 的讨论可得，而最后一个不等式由假设 (7) 成立。

因此，新的解 $(\hat{}\boldsymbol{f},\boldsymbol{c})$ 是可行的，因为在所有情况下，(2a) 中的约束均满足：左端项$(\hat{}\boldsymbol{f},\tau)$ $\leq$右端项$(\boldsymbol{c},\tau)$。

通过这一变换，我们从线性规划 (2) 中消除了不必要的变量和约束。根据 B 与 u_n 之间的关系，我们通过分析以下两种情况来展开讨论：(i) $B<u_n$ 以及 (ii) $u_n\leq B$。为简化记号，设 $u_{n+1}=\infty$。

情况1： $B < u_n$。我们首先从线性规划 (2) 中消除冗余变量。为此，我们建立关于可行解的如下结论。令 $j = \min\{i \in [n] \mid u_i > B\}$。我们声称，对于任意可行解 $(\boldsymbol{f},\ \boldsymbol{c})$，都存在另一个具有相同目标值的可行解 $(\hat{\boldsymbol{f}},\ \boldsymbol{c})$，使得仅当 $t \in S$ 时 $\hat{f}_t \neq 0$

$$S \coloneqq \{1, 2, \ldots, B\} \cup \{u_j + 1, u_{j+1} + 1, \ldots, u_n + 1\}.$$

我们通过使用上述针对可行解的变换来证明这一结论。对于每个 $\tau \geq B$，我们有 $\mathrm{RHS}(\boldsymbol{c}, \tau) = B \cdot c_{i(\tau)}$。根据层次结构和 $B < u_j$，我们有 $i(B) \geq i(B+1) \geq \cdots \geq i(u_j)$。由于 $1 \leq c_1 \leq \cdots \leq c_{n+1}$，可得 $\mathrm{RHS}(\boldsymbol{c}, B) \geq \mathrm{RHS}(\boldsymbol{c}, B+1) \geq \cdots \geq \mathrm{RHS}(\boldsymbol{c}, u_j)$。通过对 τ^* 从 $u_j - 1$ 到 B 反复应用变换 (8)，我们可以将整个概率质量从区间 $\{B+1, B+2, \ldots, u_j\}$ 转移到第 B 天，同时不违反可行性。类似地，对于每个 $i \in \{j, j+1, \ldots, n\}$，我们可以将整个概率质量从区间 $\{u_i + 1, u_i + 2, \ldots, u_{i+1}\}$ 转移到第 $u_i + 1$ 天，因为对于所有 $\tau \in \{u_i + 1, u_i + 2, \ldots, u_{i+1}\}$，我们都有 $\mathrm{RHS}(\boldsymbol{c}, \tau) = B \cdot c_{i+1}$。因此，我们可以构造一个新的可行解 $(\boldsymbol{c}, \hat{\boldsymbol{f}})$，使其达到相同的目标值，其中 $\hat{\boldsymbol{f}} = (\hat{f}_t)_{t \in \mathcal{I}}$ 定义如下：

$$\hat{f}_t = \begin{cases} f_t & \text{if } t \leq B - 1, \\ \sum_{k=B}^{u_j} f_k & \text{if } t = B, \\ \sum_{k=u_i+1}^{u_{i+1}} f_k & \text{if } t = u_i + 1 \leq u_{i+1} \text{ for some } i \in \{j, j+1, \ldots, n\}, \\ 0 & \text{otherwise} \end{cases} \qquad (t \in \mathcal{I}).$$

这一变换证明，可以在满足 $f_t = 0$ for all $t \notin S$ 的子空间内找到一个最优解。因此，我们可以从线性规划 (2) 中消去变量 f_t for $t \notin S$，而不损失最优性。

接下来，利用消去这些变量后的线性规划，我们从 (2a) 中移除冗余约束。定义

$$T \coloneqq \{1, \ldots, B\} \cup \{\ell_i - 1,\ u_i \mid i \in [n]\} \cup \{u_n + 1\}.$$

我们证明，所有满足 C_τ with $\tau \notin T$ 的约束都是冗余约束。设 $(\boldsymbol{c}, (f_t)_{t \in S})$ 是约化线性规划的一个解。首先，由于对所有 $t \geq u_n + 2$ 都有 $f_t = 0$，所有约束 $(C_\tau)_{\tau \geq u_n + 1}$ 的形式都相同，即 $\sum_{t \in S} (t - 1 + B) f_t + \tau(1 - \sum_{t \in S} f_t) \leq B \cdot c_{n+1}$。因此，只需验证 $\tau = u_n + 1$ 处的约束，无需验证所有约束 $(C_\tau)_{\tau \geq u_n + 2}$。接着，对每个 $i \in \{j, j+1, \ldots, n\}$，考虑集合 $(\Theta_i \setminus \Theta_{i-1}) \cap \{B, B+1, \ldots\}$。对于该集合中的每个 τ，都有 $\mathrm{RHS}(\boldsymbol{c}, \tau) = B \cdot c_{i(\tau)}$，且 $\mathrm{LHS}(\hat{\boldsymbol{f}}, \tau)$ 关于 τ 单调不减。因此，如果存在 $\tau^* \in (\Theta_i \setminus \Theta_{i-1}) \cap \{B, B+1, \ldots\}$ 使得 $\mathrm{LHS}(\hat{\boldsymbol{f}}, \tau^*) \leq \mathrm{RHS}(\boldsymbol{c}, \tau^*)$ 成立，那么对于所有满足 $\tau \leq \tau^*$ 的 $\tau \in (\Theta_i \setminus \Theta_{i-1}) \cap \{B, B+1, \ldots\}$，$\mathrm{LHS}(\hat{\boldsymbol{f}}, \tau) \leq \mathrm{RHS}(\boldsymbol{c}, \tau)$ 都自动成立。因此，对每个 i，只需验证 $(\Theta_i \setminus \Theta_{i-1}) \cap \{B, B+1, \ldots\}$ 中最大元素处的约束是否成立，而该元素 τ^* 是以下元素之一：

$$\tau^* = u_i \text{ if } u_i > u_{i-1}, \qquad \tau^* = \ell_{i-1} - 1 \text{ if } u_i = u_{i-1} \text{ and } \ell_i < \ell_{i-1}.$$

因此，检查 $(C_\tau)_{\tau \in T}$ 中的约束就足够了。

情况2： $u_n \leq B$。与情况1一样，我们首先通过对解进行变换来消除冗余变量，然后移除冗余约束。我们首先证明，对于任意可行解 $(\boldsymbol{f},\ \boldsymbol{c})$，都存在另一个具有相同目标值的可行解 $(\hat{\boldsymbol{f}},\ \boldsymbol{c})$，使得仅当 t 时 $\hat{f}_t \neq 0$ 在

$$S \coloneqq [B].$$

对于所有 $\tau \geq B$，我们都有 $\mathrm{RHS}(\tau) = c_{n+1} B$。通过反复对 τ^* 应用从 ∞ 到 B 的变换 (8)，我们可以将所有日期 $t \geq B + 1$ 的全部概率质量转移到日期 B，且不违反可行性。因此，我们可以构造一个新的可行解 $(\hat{\boldsymbol{f}},\ \boldsymbol{c})$，使其达到相同的目标值，其中 $\hat{\boldsymbol{f}}$ 定义如下：

$$\hat{f}_t = \begin{cases} f_t & \text{if } t \leq B - 1, \\ \sum_{k=B}^{\infty} f_k \ (= 1 - \sum_{k=1}^{B-1} f_k) & \text{if } t = B, \\ 0 & \text{if } t > B \end{cases} \qquad (t \in \mathcal{I}).$$

根据上述相同的讨论，我们可以从线性规划 (2) 中消去变量 f_t （其中 $t \notin S$ ），且不损失最优性。

接下来，利用消去这些变量后的线性规划，我们从（2a）中删除冗余约束。我们证明所有满足 C_τ with $\tau \notin T$ 的约束都是冗余的，其中

$$T := [B+1].$$

ˆ由于对所有 $t \geq B+1$ 都有 $f_t = 0$，因此只需验证约束 C_{B+1}，即可确保满足所有约束（C_τ）$_{\tau \geq B+1}$。因此，只需检查约束（C_τ）$_{\tau \in T}$。

综合这些情况，我们得出结论：无限维线性规划（2）可以化简为有限维线性规划（3）。

□