Inverse Optimization with Prediction Market: A Characterization of Scoring Rules for Elciting System States

Han Bao
Kyoto University

Shinsaku Sakaue
The University of Tokyo & RIKEN AIP

Abstract

Inverse optimization aims to recover the unknown state in forward optimization after observing a state-outcome pair. This is relevant when we want to identify the underlying state of a system or to design a system with desirable outcomes. Whereas inverse optimization has been investigated in the algorithmic perspective during past two decades, its formulation is intimately tied with the principal's subjective choice of a desirable state—indeed, this is crucial to make the inverse problem well-posed. We propose an alternative model to the conventional inverse optimization by building upon prediction market, where multiple agents submit their beliefs until converging to market equilibria. The market equilibria express the crowd consensus on a desirable state, effectively eschewing the subjective design by the principal. To this end, we derive a proper scoring rule for prediction market design in the context of inverse optimization. We can benefit from this model of inverse optimization to represent crowd knowledge in optimization problems.

1 INTRODUCTION

While many decision-making problems are formulated through *forward* optimization problems, their objective functions may depend on unknown states. The goal of *inverse* optimization is to identify the unknown state involved in a forward optimization problem. Consider a forward optimization problem

$$\underset{\boldsymbol{\omega}\in\Omega}{\text{maximize}}\, f(\boldsymbol{\omega};\boldsymbol{\theta}), \tag{1}$$

where $\Omega \subseteq \mathbb{R}^d$ is the feasible region and $\boldsymbol{\theta}$ is the state determining the objective function. In inverse optimization, we aim to identify the underlying state $\boldsymbol{\theta}$ from an incoming outcome $\boldsymbol{\omega} \in \Omega$. This problem is motivated by the following three (non-exhaustive) scenarios, among which some examples have been previously discussed (Iyengar and Kang, 2005).

Example 1 (System identification)**.** The first scenario is system identification (Burton and Toint, 1992): suppose that we aim to identify the transmission time of seismic waves between two geological locations supported on a directed graph. Once an earthquake is observed, we regard the actual transmission time as the unknown state $\boldsymbol{\theta}$ of the shortest path problem. Then, we are able to estimate the transmission time between two geological locations after observing an earthquake.

Example 2 (Outcome-aligned design)**.** The second scenario is enforcement of a desired outcome $\boldsymbol{\omega}$: suppose that the traffic flow is a solution of the shortest path problem defined by the toll as a state $\boldsymbol{\theta}$ and we seek $\boldsymbol{\theta}$ to realize a specific flow $\boldsymbol{\omega}$ (Dial, 1999, 2000).

Example 3 (Inverse optimal transport)**.** The last scenario is inverse optimal transport, where we want to recover the ground metric of transport from the actual transport plan. Inverse optimal transport is particularly interesting because its connection to contrastive self-supervised learning has been pointed out recently (Shi et al., 2023), where the transport plan corresponds to sample alignment between semantically close data points and we seek a better ground metric implemented by a learnable data representation.

Despite the popularity of inverse optimization, its formulation is not unanimous because of the ill-posed nature. Actually, there can be multiple states $\boldsymbol{\theta}$ yielding the same outcome $\boldsymbol{\omega}$ in forward optimization (1). To make the problem well-posed, we need inductive bias on the underlying state $\boldsymbol{\theta}$. Iyengar and Kang (2005) formulated inverse optimization by choosing a state as close to a reference state as possible while satisfying the optimality of the forward optimization (1) as follows:

$$\begin{aligned}&\underset{\boldsymbol{\theta}\in\mathcal{S}}{\text{minimize}}\, \|\boldsymbol{\theta}-\boldsymbol{\theta}_{\text{ref}}\| \\ &\text{subject to} \quad \widehat{\boldsymbol{\omega}} \in \underset{\boldsymbol{\omega}\in\Omega}{\arg\min}\, f(\boldsymbol{\omega};\boldsymbol{\theta}),\end{aligned} \tag{2}$$

Proceedings of the 28[th] International Conference on Artificial Intelligence and Statistics (AISTATS) 2025, Mai Khao, Thailand. PMLR: Volume 258.

where $\boldsymbol{\theta}_{\text{ref}}$ is a given reference point, $\mathcal{S} \subseteq \mathbb{R}^d$ is a set of admissible states, and $\widehat{\boldsymbol{\omega}}$ is the observed outcome. More recently, Bärmann et al. (2017) considered inverse online optimization by measuring the similarity between the estimated and underlying states with their associated outcomes. Sun et al. (2023) and Besbes et al. (2023) focused on the margin of the estimated state from the boundary of the polar cone associated with an outcome. Though all of these criteria are reasonable to some extent, the choice of the evaluation criterion (such as $\|\cdot - \boldsymbol{\theta}_{\text{ref}}\|$ in Eq. (2)) remains subjective to the principal, who wants to identify the state. Robust inverse optimization (Ghobadi et al., 2018; Mohajerin Esfahani et al., 2018) formulates the uncertainty in the reference point $\boldsymbol{\theta}_{\text{ref}}$, which still requires the subjective design of the uncertainty set. In the motivating scenarios such as traffic network design, the decision making eventually affects the utility of general public, who should be involved in the process of the evaluation criterion design for enhancing informed decision making. Can we incorporate the crowd knowledge to formulate inverse optimization?

We borrow the idea from information elicitation and decision markets, where agents have their own beliefs about outcomes conditioned on principal's action, and the principal is interested in eliciting agents' truthful beliefs (Othman and Sandholm, 2010; Chen and Kash, 2011). Consider asking "What is the conversion probability when we make a promotion X (an action)? What if a promotion Y (another action)?" By aligning with *forward* optimization (1), principal's action corresponds to choosing a state $\boldsymbol{\theta}$, and an agent makes a guess about the outcome $\boldsymbol{\omega}$. Proper scoring rules (Savage, 1971; Buja et al., 2005; Gneiting and Raftery, 2007; Reid and Williamson, 2010; Frongillo and Kash, 2021; Bao, 2023) are the basis of designing prediction markets—such a scoring rule incentivizes an agent to truthfully report her belief about the event. For decision problems, Othman and Sandholm (2010) gave a seminal formulation: the agent first reveals her belief about an outcome $\boldsymbol{\omega}$ *conditioned on* a played action $\boldsymbol{\theta}$, the principal plays an action by choosing a state $\boldsymbol{\theta}$, and then an outcome $\boldsymbol{\omega}$ is realized based on the principal's action. The question is whether we can incentivize the agent to report her conditional belief truthfully. Chen and Kash (2011) showed that there exists a strictly proper scoring rule to achieve this. Based on the correspondence between a scoring rule and prediction market, we can design a prediction market, which aggregates beliefs of market participants through the equilibrium market price (Hanson, 2003; Lambert et al., 2008).

The goal of this paper is to formulate inverse optimization to take multiple agents' belief about the state. We incentivize an agent to truthfully report her belief about the state given an outcome by a proper scoring rule. By turning such a scoring rule into prediction market, we can aggregate multiple agents' belief about reasonable states. While the traditional formulations of inverse optimization, represented by Eq. (2), have been solely governed by the principal's inductive bias, our formulation of inverse optimization relies on the agents' belief on the state, avoiding the principal's subjective design. Eventually, our formulation falls into a Stackelberg model between the principal and agents.

Summary of contributions. Focusing on inverse linear optimization (ILP), we contribute as follows: (i) to derive a strictly proper scoring rule for ILP and turn it into prediction market (Sections 3.2 and 3.3); (ii) to formulate ILP as a Stackelberg game between the agents and principal, who correspond to the leader maximizing the market scoring rule and the follower solving forward optimization (Section 3.4); (iii) to discuss the dual relationship between the agent's belief and the uncertainty in forward optimization, shedding light on a rich structure of ILP and prediction market (Section 3.5). Therefore, we propose a non-subjective model of inverse optimization.

2 BACKGROUND

We use bold-face $\mathbf{p}$ to denote a vector with each element p_ω, and capital P to denote a matrix with each element P_{ij}. The dual pair $\langle \cdot, \cdot \rangle$ recovers the standard inner product $\langle \mathbf{r}, \mathbf{p} \rangle = \sum_\omega r_\omega p_\omega$ for two vectors $\mathbf{r}, \mathbf{p}$ and the Frobenius inner product $\langle R, P \rangle = \sum_{i,j} R_{ij} P_{ij}$ for two matrices R, P.

2.1 Proper Scoring Rules

Proper scoring rules have been used to define a supervised loss for machine learning (Buja et al., 2005; Reid and Williamson, 2010; Bao, 2023) and incentivize a risk-neutral agent to truthfully report her probability assessment for an uncertain event (Lambert et al., 2008; Abernethy and Frongillo, 2012; Frongillo and Kash, 2021). We introduce scoring rules in the context of information elicitation. Let $\Omega \coloneqq \{\omega_1, \ldots, \omega_K\}$ be a size-K discrete outcome space for the event and $\triangle(\Omega)$ be the probability space over outcomes. Denote $\underline{\mathbb{R}} \coloneqq \mathbb{R} \cup \{-\infty\}$. A scoring rule $\mathbf{s} \colon \triangle(\Omega) \to \underline{\mathbb{R}}^K$ gives a score $s_\omega(\mathbf{p})$ to the agent when her report is $\mathbf{p} \in \triangle(\Omega)$ and the outcome is $\omega \in \Omega$. A scoring rule is said *regular* if $\mathbf{s}(\mathbf{p}) \in \mathbb{R}^K$, except $s_\omega(\mathbf{p}) = -\infty$ being allowed for $p_\omega = 0$. Let us write the expected score as follows:

$$V(\mathbf{p}, \mathbf{q}) \coloneqq \sum_{\omega \in \Omega} p_\omega s_\omega(\mathbf{q}),$$

where $\mathbf{p}, \mathbf{q} \in \triangle(\Omega)$ is agent's truthful and reported beliefs. Note that the agent can misreport her truthful belief. A regular scoring rule is *proper* if her truthful report $\mathbf{p}$ maximizes the expected score $V(\mathbf{p}, \cdot)$, and *strictly proper* if the truthful report $\mathbf{p}$ is the unique maximizer. For example, the log score $s_\omega(\mathbf{p}) := a_\omega + b \log p_\omega$ and Brier score $s_\omega(\mathbf{p}) := a_\omega + b(2p_\omega - \sum_i p_i^2)$ are common strictly proper scoring rules for arbitrary $b > 0$ and $a_\omega \in \mathbb{R}$.

A proper scoring rule is tightly connected to a convex function. The following relationship dates back to Savage (1971), and has been rediscovered in many fields, including statistics (Gneiting and Raftery, 2007), machine learning (Reid and Williamson, 2010), and mechanism design (Frongillo and Kash, 2021).

Theorem 1 (Savage (1971))**.** *A regular scoring rule is (strictly) proper if and only if there exists a proper*[1] *and (strictly) convex function* $\Lambda \colon \triangle(\Omega) \to \mathbb{R}$ *such that for all* $\mathbf{p} \in \triangle(\Omega)$*, there exists a subgradient* $\mathbf{g} \in \partial\Lambda(\mathbf{p})$ *satisfying*

$$s_\omega(\mathbf{p}) = \Lambda(\mathbf{p}) - \langle \mathbf{g}, \mathbf{p} \rangle + g_\omega \quad \textit{for } \mathbf{p} \in \triangle(\Omega),$$

where $\partial\Lambda$ *is the subdifferential of* Λ*.*

Theorem 1 has two implications. First, the regret of the expected score of a regular proper scoring rule is a Bregman divergence. Indeed, we have

$$\begin{aligned} V(\mathbf{p}, \mathbf{p}) - V(\mathbf{p}, \mathbf{q}) &= \Lambda(\mathbf{p}) - \Lambda(\mathbf{q}) - \langle \mathbf{g}, \mathbf{p} - \mathbf{q} \rangle \\ &=: B_\Lambda(\mathbf{p} \| \mathbf{q}), \end{aligned}$$

where $B_\Lambda(\mathbf{p}\|\mathbf{q})$ is the Bregman divergence generated by Λ. Hence, the expected score maximization with respect to $\mathbf{q}$ can be viewed as the Bregman projection. Second, $\mathbf{s}(\mathbf{p})$ is a subgradient of Λ at the point $\mathbf{p}$, namely, $\mathbf{s}(\mathbf{p}) \in \partial\Lambda(\mathbf{p})$. This is utilized by Frongillo and Kash (2021) to turn a scoring rule into a mechanism and vice versa. Confer Bao and Takatsu (2024) for more details of convex analysis such as the subdifferential and Bregman divergence.

Example 4. To see an example of Theorem 1, one can consider the log score $s_\omega(\mathbf{p}) = a_\omega + b \log p_\omega$ again. Then, the convex function Λ is nothing else but the negative Shannon entropy $\Lambda(\mathbf{p}) = \langle \log \mathbf{p}, \mathbf{p} \rangle$.

2.2 Information Elicitation with Decision Rules

We review information elicitation with decision rules by following the convention of Othman and Sandholm (2010) and Chen and Kash (2011). We are interested in agent's probabilistic assessment about uncertain outcomes $\omega \in \Omega$, but now outcomes may depend on *states* $\mathcal{S} := \{\theta_1, \dots, \theta_M\}$. Each state affects the realization probability of outcomes. In this scenario, we assume that the principal may select a state, which is principal's action. We write the belief space $\mathcal{P} := \mathcal{S} \times \triangle(\Omega)$ in this context. In information elicitation with decision rules, we work on *conditional* beliefs on outcomes given an action, unlike (marginal) beliefs on outcomes in Section 2.1. The report procedure, termed as decision game for the convenience, goes as follows:

Decision game (Chen and Kash, 2011)

1. The principal asks an agent to **report** her conditional belief $Q_{\theta\omega} = \Pr[\omega \mid \theta]$.
2. The principal **plays an action** $\theta \in \mathcal{S}$ with a fixed decision rule $D \colon \mathcal{P} \to \triangle(\mathcal{S})$, where $D_\theta(Q)$ is the probability selecting action θ given the reported belief Q.
3. Nature **samples an outcome** $\omega \in \Omega$ according to agent's truthful belief P. The agent is expected to assess likely outcomes.
4. The principal **scores agent's report** based on a scoring rule $S \colon \mathcal{P} \to \underline{\mathbb{R}}^{M \times K}$, where $S_{\theta\omega}(Q)$ denotes the score of the reported belief Q when the realized outcome is ω given the selected action θ.

The decision rule D and scoring rule S are revealed before agent's report. Our goal is to design a scoring rule encouraging the agent to truthfully report her belief. To do this, we introduce proper scoring rules *for decision rules*. Hereafter, we extend the expected score introduced in Section 2.1 to matrix beliefs conditioned on actions: for $P, Q \in \mathcal{P}$,

$$V(P, Q) := \sum_{\theta \in \mathcal{S}, \omega \in \Omega} D_\theta(Q) P_{\theta\omega} S_{\theta\omega}(Q). \tag{3}$$

Definition 2. *A scoring rule* S *is* regular *for a decision rule* D *if* $S_{\theta\omega}(P) \in \mathbb{R}$ *unless* $P_{\theta\omega} = 0$*.*

Definition 3. *A regular scoring rule* S *is* proper *for a decision rule* D *if*

$$V(P, P) \geq V(P, Q) \quad \textit{for all } P, Q \in \mathcal{P} \textit{ with } Q \neq P.$$

It is strictly proper *for a decision rule* D *when the inequality is strict.*

A strictly proper scoring rule (if it exists) elicits agent's truthful belief. However, Othman and Sandholm

[1] A convex function is *proper* if its effective domain is non-empty. This is irrelevant to the properness of a scoring rule.

(2010) showed that no *deterministic* decision rules have a *strictly* proper scoring rule—nonetheless, they formulate a weaker version of the strict properness, *quasi-strictly* proper scoring rules, to characterize admissible decision rules. Refer to Othman and Sandholm (2010) for the further details.

In parallel with Theorem 1, Chen and Kash (2011) characterized strictly proper scoring rules for arbitrary decision rules, including randomized ones.

Theorem 4 (Chen and Kash (2011))**.** *A regular scoring rule is (strictly) proper for a decision rule D if and only if there exists a proper and (strictly) convex function $\Lambda\colon \mathcal{P} \to \mathbb{R}$ such that for all $P \in \mathcal{P}$, there exists a subgradient $G \in \partial\Lambda(P)$ with $G_{\theta\omega} = 0$ when $D_\theta(P) = 0$ and* <u>*(i) when $D_\theta(P) > 0$*</u>*,*

$$S_{\theta\omega}(P) = \Lambda(P) - \langle G, P \rangle + \frac{G_{\theta\omega}}{D_\theta(P)} \quad \text{for } P \in \mathcal{P},$$

and <u>*(ii) when $D_\theta(P) = 0$*</u>*, $S_{\theta\omega}(P)$ is arbitrary and allowed to be $-\infty$ only when $P_{\theta\omega} = 0$.*

This theorem gives a way to test whether a given scoring rule is strictly proper for a decision rule D. We make two remarks. First, a scoring rule form given by Theorem 4 is similar to the Bregman score form by Theorem 1 but the last subgradient term is inversely weighted by $D_\theta(P)$. This indicates that one should upweight scores for action θ that the principal selects with low probability. Intuitively, the upweight is necessary for "debiasing" similarly to inverse propensity weighting. Second, Theorem 4 requires an additional condition on the convex function: its subgradient satisfies $G_{\theta\omega} = 0$ when $D_\theta(P) = 0$. This constraint is necessary to ignore agent's report conditioned on actions never taken, but makes the design of Λ challenging when the decision rule D is not fully supported (Othman and Sandholm, 2010).

Example 5. Othman and Sandholm (2010) considers the following scenario: the outcome space is binary $\Omega = \{\top, \bot\}$, the state space is $\mathcal{S} = \{\theta_1, \ldots, \theta_M\}$, and the MAX decision rule is $D_\theta(P) = \mathbb{1}_{\{\theta = \arg\max_{\vartheta\in\mathcal{S}} P_{\vartheta\top}\}}$ for $P \in \mathcal{S} \times \triangle(\Omega)$. They showed that no strictly proper scoring rules exist just because the MAX decision rule is deterministic.

For randomized decision rules, let us mention another example provided by Chen and Kash (2011). Consider the binary outcome case again and the decision rule $D_\theta(P) = P_{\theta\top} / \sum_{\vartheta\in\mathcal{S}} P_{\vartheta\top}$. Then, $\Lambda(P) = \sum_{\theta\in\mathcal{S}} P_{\theta\top}^2$ (with Theorem 4) yields a strictly proper scoring rule:

$$S_{\theta\omega} = \begin{cases} \sum_{\vartheta\in\mathcal{S}} (2P_{\vartheta\top} - P_{\vartheta\top}^2) & \text{if } \omega = \top, \\ -\sum_{\vartheta\in\mathcal{S}} P_{\vartheta\top}^2 & \text{if } \omega = \bot. \end{cases}$$

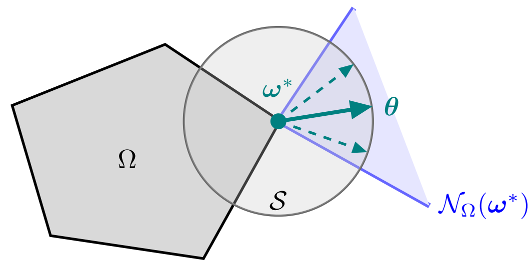

Figure 1: The relationship between an LP cost vector $\boldsymbol{\theta} \in \mathcal{S}$ and the LP solution $\boldsymbol{\omega}^*(\boldsymbol{\theta}) \in \Omega$. The LP cost vector $\boldsymbol{\theta}$ must lie in the normal cone $\mathcal{N}_\Omega(\boldsymbol{\omega}^*)$. Conversely, any cost vector lying in the normal cone $\boldsymbol{\theta} \in \mathcal{N}_\Omega(\boldsymbol{\omega}^*)$ yields the same LP solution $\boldsymbol{\omega}^*$.

3 INVERSE LINEAR PROGRAM BASED ON PREDICTION MARKET

In this section, we formulate inverse optimization via information elicitation and prediction market. We herein focus on inverse linear program (ILP) with the following standard form of an LP:

$$\boldsymbol{\omega}^*(\boldsymbol{\theta}) \coloneqq \operatorname*{arg\,min}_{\boldsymbol{\omega}\in\Omega} \left\{ f(\boldsymbol{\omega}; \boldsymbol{\theta}) \coloneqq \langle \boldsymbol{\theta}, \boldsymbol{\omega} \rangle \right\}, \tag{4}$$

where $\Omega \coloneqq \{\boldsymbol{\omega} \mid A\boldsymbol{\omega} \le \mathbf{b}\} \subseteq \mathbb{R}^d$ is the feasible region defined by $A \in \mathbb{R}^{n\times d}$ and $\mathbf{b} \in \mathbb{R}^n$. Here, the feasible region Ω is supposed to be known in advance. In ILP, we are given an incoming fresh outcome $\boldsymbol{\omega}$ and infer the LP cost $\boldsymbol{\theta} \in \mathcal{S}$, where $\mathcal{S} \subseteq \mathbb{R}^d$ is a state space.

It is not straightforward at all to infer an LP state $\boldsymbol{\theta}$ from an outcome $\boldsymbol{\omega}$ because of the following reasons: in forward optimization, the principal is given an LP state $\boldsymbol{\theta}$ and asked to solve Eq. (4), whose outcome $\boldsymbol{\omega}^*(\boldsymbol{\theta})$ is uniquely determined (unless $\boldsymbol{\theta}$ is normal to any face of the convex polytope Ω). However, inverse optimization is underdetermined as seen in Fig. 1, where different LP states $\boldsymbol{\theta}$ can yield the same outcome $\boldsymbol{\omega}$ as long as $\boldsymbol{\theta}$ lies in the normal cone $\mathcal{N}_\Omega(\boldsymbol{\omega}) \coloneqq \{\boldsymbol{\theta} \mid \langle \boldsymbol{\omega}' - \boldsymbol{\omega}, \boldsymbol{\theta} \rangle \le 0 \ \forall \boldsymbol{\omega}' \in \Omega\}$. Even worse, a cost vector $\boldsymbol{\theta}$ could be very brittle against a tiny perturbation if $\boldsymbol{\theta}$ is close to the boundary of the normal cone $\mathcal{N}_\Omega(\boldsymbol{\omega})$. The principal should ask for agent's preference through information elicitation to reliably infer an LP state from a given LP outcome, ideally.

3.1 Eliciting State from Agent

We suppose that an LP cost vector $\boldsymbol{\theta} \in \mathcal{S}$ shall never be normal to any face of Ω so that the LP solution $\boldsymbol{\omega}^*(\boldsymbol{\theta})$ does not have a tie. This is a common assumption akin to the LP non-degeneracy (Sun et al., 2023). Let $\overline{\Omega} \coloneqq \{\boldsymbol{\omega}_1, \ldots, \boldsymbol{\omega}_K\}$ be the finite vertex set of the convex polytope Ω containing the basic feasible solu-

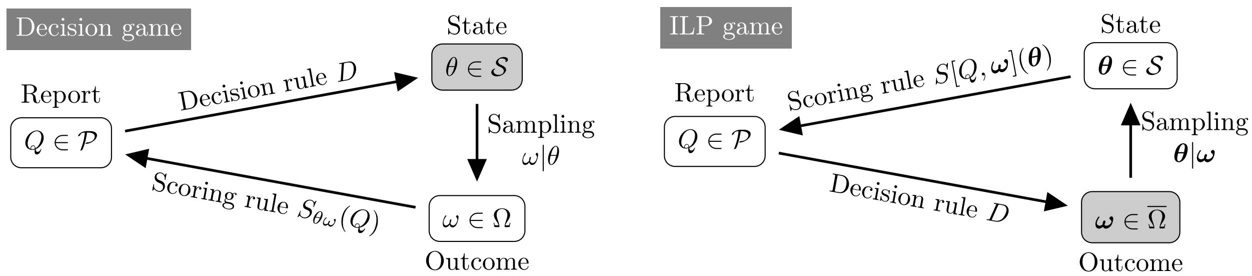

Figure 2: Illustrative comparison between the decision game (Section 2.2) and ILP game (Section 3.4). In the decision game, the principal acts on a state $\theta \in \mathcal{S}$ according to agent's report Q, while in the ILP game he acts on an outcome $\boldsymbol{\omega} \in \overline{\Omega}$. Both the action targets are highlighted in the figure. In both games, a lifecycle in each round proceeds with the decision rule D, sampling (based on agent's truthful conditional belief), and evaluating the report by the scoring rule S; nonetheless, the decision game goes with "$Q \to \theta \to \omega \to Q$", while the ILP game goes with "$Q \to \boldsymbol{\omega} \to \boldsymbol{\theta} \to Q$".

tions. Departing from the decision game (Section 2.2), we identify $\mathcal{S}$ and $\overline{\Omega}$ as the state and outcome spaces, respectively, in information elicitation with decision rules. Note that $\mathcal{S}$ is infinitely-sized in general.

Moreover, we assume that the state space $\mathcal{S} = \mathbb{B}^d(1) := \left\{\boldsymbol{\theta} \in \mathbb{R}^d \;\middle|\; \|\boldsymbol{\theta}\|_2 \le 1\right\}$ because the information about the norm $\|\boldsymbol{\theta}\|_2$ cannot be recovered with the observation of an LP solution $\boldsymbol{\omega}$ solely—all we know is that the cost vector lies in the normal cone of Ω at $\boldsymbol{\omega}$ such that $\boldsymbol{\theta} \in \mathcal{N}_\Omega(\boldsymbol{\omega})$ (Fig. 1). Similar assumptions are common in the literature of contextual inverse optimization (Bärmann et al., 2017; Besbes et al., 2023).

Assumption 1. *The state space is* $\mathcal{S} = \mathbb{B}^d(1)$.

We formulate ILP as the report game with several tweaks to the decision game (Section 2.2). The conditional belief space is extended to $\mathcal{P} := \overline{\Omega} \to \triangle(\mathcal{S})$ to deal with infinitely-sized $\mathcal{S}$, and a decision rule makes an output on the outcome space $\overline{\Omega}$ instead of the state space $\mathcal{S}$. We consider the following report game, called the ILP game for the convenience:

ILP game

1. The principal asks an agent to **report** her conditional belief $Q[\boldsymbol{\omega}](\boldsymbol{\theta}) = \Pr[\boldsymbol{\theta} \,|\, \boldsymbol{\omega}]$, where $\boldsymbol{\theta} \in \mathcal{S}$ and $\boldsymbol{\omega} \in \overline{\Omega}$ are an LP cost and the associated solution, respectively.

2. The principal **plays an action** $\boldsymbol{\omega} \in \overline{\Omega}$ with a fixed decision rule $D\colon \mathcal{P} \to \triangle(\overline{\Omega})$, where $D[Q](\boldsymbol{\omega})$ is the probability selecting action $\boldsymbol{\omega}$ given the reported belief Q.

3. Nature **samples a state** $\boldsymbol{\theta} \in \mathcal{S}$ given the played decision $\boldsymbol{\omega}$, according to agent's conditional belief $\boldsymbol{\theta} \sim P[\boldsymbol{\omega}]$. Agent's task is to assess likely states.

4. The principal **scores agent's report** based on a scoring rule $S\colon \mathcal{P} \times \overline{\Omega} \to (\mathcal{S} \to \mathbb{R})$, where $S[Q, \boldsymbol{\omega}](\boldsymbol{\theta})$ denotes the score of the reported belief Q when the sampled state $\boldsymbol{\theta}$ is given the played action $\boldsymbol{\omega}$.

The primary goal of the ILP game is to elicit agent's truthful belief on states. A reasonable conditional belief is supposed to make a good guess on an LP state $\boldsymbol{\theta}$ from an outcome $\boldsymbol{\omega}$ such that a reliable decision process yields the LP solution $\boldsymbol{\omega}$. This decision process is modeled as principal's decision rule D. Agent's conditional belief is refined by maximizing a proper scoring rule, which incentivizes a belief closely aligned with the observed state-outcome pair. After eliciting the agent's truthful belief, we can turn it into a criterion to identify the underlying state in ILP, as we will see in Section 3.4. Therein, the principal's decision rule is subjective to agents' belief, which effectively eliminates the principal's selfhood. We discuss a possible implementation of agent's report in Section 3.5.

Now compare the decision game (Section 2.2) with the ILP game in Fig. 2. Whereas the principal acts in the state space $\mathcal{S}$ in the former, he acts in the outcome space $\overline{\Omega}$ in the latter. Accordingly, the agent infers states in the ILP game, unlike inferring outcomes in the decision game. The source of these differences is rooted in what one wants to elicit: the decision game cares about outcomes, and inverse optimization cares about likely states of the optimization problem.

3.2 Proper Scoring Rules

We first derive an admissible form of proper scoring rules for the ILP game without assuming a specific form of the ILP decision rule. For a scoring rule $S\colon \mathcal{P} \times \overline{\Omega} \to (\mathcal{S} \to \mathbb{R})$, the expected score is extended from Eq. (3) as follows:

$$V(P, Q) := \int_{\mathcal{S}} \mathrm{d}\boldsymbol{\theta} \sum_{\boldsymbol{\omega} \in \overline{\Omega}} D[Q](\boldsymbol{\omega}) P[\boldsymbol{\omega}](\boldsymbol{\theta}) S[Q, \boldsymbol{\omega}](\boldsymbol{\theta}) \quad (5)$$

for $P, Q \in \mathcal{P}$.

Theorem 5. *A regular scoring rule is (strictly) proper for an ILP decision rule $D\colon \mathcal{P} \to \triangle(\overline{\Omega})$ if and only if there exists a proper and (strictly) convex function $\Lambda\colon \mathcal{P} \to \overline{\mathbb{R}}$ such that for all $P \in \mathcal{P}$, there exists a subgradient $G_P \in \partial\Lambda(P) : \overline{\Omega} \to (\mathcal{S} \to \overline{\mathbb{R}})$ with $G_P[\boldsymbol{\omega}](\boldsymbol{\theta}) = 0$ when $D[P](\boldsymbol{\omega}) = 0$ and (i) when $D[P](\boldsymbol{\omega}) > 0$,*

$$S[P, \boldsymbol{\omega}](\boldsymbol{\theta}) = \Lambda(P) - \langle G_P, P\rangle + \frac{G_P[\boldsymbol{\omega}](\boldsymbol{\theta})}{D[P](\boldsymbol{\omega})} \quad \text{for } P \in \mathcal{P}, \tag{6}$$

and (ii) when $D[P](\boldsymbol{\omega}) = 0$, $S[P, \boldsymbol{\omega}](\boldsymbol{\theta})$ is arbitrary and allowed to be $-\infty$ only when $P[\boldsymbol{\omega}](\boldsymbol{\theta}) = 0$.

The proof is given in Appendix A. Once we derive a strictly proper scoring rule in the form of Eq. (6), the principal can elicit agent's truthful conditional belief on state $\boldsymbol{\theta}$. Thus, for an incoming fresh outcome $\boldsymbol{\omega}$, an LP state can be inferred based on the elicited belief.

While it is challenging to design a strictly proper scoring rule for a general ILP decision rule, it is rather easy for the full-support decision rule. We suppose this for the time being and give two examples of proper scoring rules for the ILP game. The first example of strictly proper scoring rules is the Brier score, generated by

$$\Lambda(P) = \sum_{\boldsymbol{\omega} \in \overline{\Omega}} \int \mathrm{d}\boldsymbol{\theta} \{P[\boldsymbol{\omega}](\boldsymbol{\theta})\}^2.$$

This is strictly convex, and by noting that the subgradient $G_P[\boldsymbol{\omega}](\boldsymbol{\theta}) = 2P[\boldsymbol{\omega}](\boldsymbol{\theta})$, the corresponding Brier score is

$$S[P, \boldsymbol{\omega}](\boldsymbol{\theta}) = -\sum_{\boldsymbol{\omega} \in \overline{\Omega}} \int \mathrm{d}\boldsymbol{\theta} \{P[\boldsymbol{\omega}](\boldsymbol{\theta})\}^2 + \frac{2P[\boldsymbol{\omega}](\boldsymbol{\theta})}{D[P](\boldsymbol{\omega})}.$$

The second example is the log score, generated by

$$\Lambda(P) = \sum_{\boldsymbol{\omega} \in \overline{\Omega}} \int \mathrm{d}\boldsymbol{\theta} P[\boldsymbol{\omega}](\boldsymbol{\theta}) \{\log P[\boldsymbol{\omega}](\boldsymbol{\theta}) - 1\},$$

which is strictly convex. By noting that the subgradient $G_P[\boldsymbol{\omega}](\boldsymbol{\theta}) = \log P[\boldsymbol{\omega}](\boldsymbol{\theta})$, the log score is

$$S[P, \boldsymbol{\omega}](\boldsymbol{\theta}) = -1 + \frac{\log P[\boldsymbol{\omega}](\boldsymbol{\theta})}{D[P](\boldsymbol{\omega})}.$$

This is essentially the standard log score inversely weighted by the decision rule $D[P](\boldsymbol{\omega})$. The log score is a local scoring rule (Party et al., 2012) such that we are free from integration over $\mathcal{S}$, which is practically convenient.

Remark 1. The full-support assumption of an ILP decision rule D is not restrictive because the decision rule in the ILP game is our design choice. We can choose any decision rule that is convenient for us. This is in contrast to the decision game, where the decision rule is given ex-ante. As noted in Section 3.5, however, full-support decision rules usually face an intractability issue in practice.

3.3 Prediction Market

Building on the above framework, we formulate inverse optimization without the principal's subjective inductive bias, such as the choice of the proximity measure and reference point in Eq. (2). To this end, we utilize the idea of prediction market to reflect agent's preference on the recovered LP state. The seminal work, Hanson (2003), formulated a market, where agents are allowed to place bets on an event of our interest. In prediction market, market equilibria lead to a consensus of beliefs, which we can hypothetically regard as the belief of a single agent. Thus, prediction market can be implemented with a scoring rule by eliciting the truthful belief from this hypothetical agent. This process to reach market equilibria can be understood as the Follow-The-Regularized-Leader (Chen and Vaughan, 2010) and online mirror descent (Frongillo et al., 2012).[2]

We describe the implementation of a prediction market in the context of ILP herein. First, the principal discloses strictly proper scoring rule S for an ILP decision rule D. The market initializes a crowd belief with an arbitrary estimate $Q_0 \in \mathcal{P}$. On each time t, the market maintain a crowd belief $Q_{t+1} \in \mathcal{P}$, and any agent is allowed to submit a new belief by updating the belief from Q_t to Q_{t+1}. Eventually, the market reaches an equilibrium Q_T for sufficiently large number of rounds T. This equilibrium Q_T corresponds to agent's report in the first line of the ILP game. At the close of the market, the principal plays an action $\boldsymbol{\omega}$ with the ILP decision rule D given the equilibrium belief Q_T, and nature samples a state $\boldsymbol{\theta}$ according to the crowd belief $P \in \mathcal{P}$—corresponding to the second and third lines of the ILP game. Then, each agent (who updates the belief from Q_t to Q_{t+1}) receives a (potentially negative) payoff of

$$S[Q_{t+1}, \boldsymbol{\omega}](\boldsymbol{\theta}) - S[Q_t, \boldsymbol{\omega}](\boldsymbol{\theta}).$$

The total payout for which the principal is liable is $S[Q_T, \boldsymbol{\omega}](\boldsymbol{\theta}) - S[Q_0, \boldsymbol{\omega}](\boldsymbol{\theta})$ due to the telescoping sum, which is always bounded from above by $S[P, \boldsymbol{\omega}](\boldsymbol{\theta}) -$

[2] If we suppose the agent's rationality (to perform these online learning algorithms), we can immediately guarantee the sublinear regret for reaching market equilibria. However, such rationality cannot be necessarily assumed.

$S[Q_0, \boldsymbol{\omega}](\boldsymbol{\theta})$ thanks to the properness of S. The principal never needs to worry if his payout is unbounded.

Therefore, multiple agents (selfishly) maximize their own profits, which leads to a consensus on a plausible LP state, under market liquidity. As we will see in Section 3.4, this framework enables us to get rid of the principal's subjectivity when identifying a state.

Remark 2. If a malicious agent bids adversely to the majority such that the reported belief moves the current belief away from the equilibrium, she would incur a negative payoff at the end of the market. This would not happen under the assumption of rationality. However, if malicious agents dominated the population, they would govern the equilibrium. In this case, we may regard malicious agents as the majority, and the model would fall into ochlocracy. This is a limitation of the "agent-centric" model.

3.4 Stackelberg Model for ILP

Now we discuss how to formulate the decision rule and ILP with the prediction market. In the ILP game, the principal's action corresponds to solving the forward optimization (4) but with an LP state $\boldsymbol{\theta}$ aligning with the agent's belief. Here, the agent's report in prediction market is regarded as the leader's action, and the principal as the follower. By casting the principal as the follower, we can prevent the principal from arbitrarily designing inverse optimization. Then, the leader's problem, or the prediction market, can be written as the following Stackelberg model:

$$\begin{aligned} &\underset{Q\in\mathcal{P}}{\text{minimize}}\; S[Q,\boldsymbol{\omega}](\boldsymbol{\theta}) \\ &\text{subject to} \begin{cases} D[Q] \in \arg\min_{D'\in\triangle(\overline{\Omega})} \langle \mathbf{f}_Q, D'\rangle \\ \boldsymbol{\omega} \sim D[Q] \\ \boldsymbol{\theta} \sim P[\boldsymbol{\omega}] \end{cases} \end{aligned} \tag{7}$$

where

$$\mathbf{f}_Q \coloneqq \begin{bmatrix} \vdots \\ f(\boldsymbol{\omega}_k; \boldsymbol{\theta}^*(\boldsymbol{\omega}_k)) \\ \vdots \end{bmatrix}_{k\in[K]} \in \mathbb{R}^K,$$

$$\boldsymbol{\theta}^*(\boldsymbol{\omega}) \in \underset{\boldsymbol{\theta}\in\mathcal{S}}{\arg\max}\, Q[\boldsymbol{\omega}](\boldsymbol{\theta}),$$

and $P \in \mathcal{P}$ is the truthful market belief. The leader takes an action (i.e., the agent reports an belief) Q with ex ante knowledge of the follower's policy. The follower submits a decision $D \in \triangle(\overline{\Omega})$ by seeking the best response $\boldsymbol{\omega} \in \overline{\Omega}$ in terms of the LP value $f(\boldsymbol{\omega}; \boldsymbol{\theta}^*(\boldsymbol{\omega})) = \langle \boldsymbol{\theta}^*(\boldsymbol{\omega}), \boldsymbol{\omega}\rangle$, where the LP value is inferred based on the market result. Even though the scoring rule S is arbitrarily chosen, the market reveals the agent's truthful belief as long as S is strictly proper. In this way, an alternative model to ILP is proposed based on the agents' crowd knowledge, rather than the principal's subjective belief.

We briefly discuss sampling of $\boldsymbol{\omega}$ and $\boldsymbol{\theta}$ in Eq. (7). Sampling $\boldsymbol{\omega}$ is not challenging: the principal wait until the market reaches the equilibrium Q_T and roll the dice with the decision rule $D[Q_T]$. To sample $\boldsymbol{\theta}$, one needs access to the truthful market belief P, which we can hardly hope. One possible remedy is to use the market equilibrium Q_T as a proxy when sampling $\boldsymbol{\theta}$. This issue has been known as a notorious challenge in decision market design (Chen and Kash, 2011). Though not fully satisfactory, this gives a reasonable approximation under the sufficient market liquidity if the scoring rule is strictly proper.

Now we derive an ILP decision rule from the Stackelberg model (7). To ensure that the decision rule is fully supported, we consider the relaxed version of the follower's action in Eq. (7):

$$D^\beta[Q] = \underset{D\in\triangle(\overline{\Omega})}{\arg\min}\, \langle \mathbf{f}_Q, D\rangle + \frac{1}{\beta} H(D), \tag{8}$$

where $H(D) \coloneqq -\langle \log D, D\rangle$ is the Shannon entropy and $\beta > 0$ is an inverse temperature. While D in Eq. (7) is always supported on a single $\boldsymbol{\omega}$ and deterministically chooses outcomes, the minimizer D^β in this relaxed version gives a randomized policy and is fully supported, taking the following Gibbs form:

$$D^\beta[Q](\boldsymbol{\omega}) = \frac{\exp(-\beta f(\boldsymbol{\omega}; \boldsymbol{\theta}^*(\boldsymbol{\omega})))}{Z_Q} \tag{9}$$

where $Z_Q \coloneqq \sum_{\boldsymbol{\omega}\in\overline{\Omega}} \exp(-\beta f(\boldsymbol{\omega}; \boldsymbol{\theta}^*(\boldsymbol{\omega})))$ denotes the normalizer. Here, Assumption 1 matters because the LP value $f(\boldsymbol{\omega}; \boldsymbol{\theta}^*(\boldsymbol{\omega})) = \langle \boldsymbol{\theta}^*(\boldsymbol{\omega}), \boldsymbol{\omega}\rangle$ can be made arbitrarily small otherwise. The Gibbs decision rule D^β operates by enumerating the most likely state $\boldsymbol{\theta}^*(\boldsymbol{\omega})$ (MAP state) for every outcome $\boldsymbol{\omega}$ and then choosing an outcome $\boldsymbol{\omega}$ with probability inversely proportional to the LP value $\langle \boldsymbol{\theta}^*(\boldsymbol{\omega}), \boldsymbol{\omega}\rangle$. With the low-temperature limit $\beta \to \infty$, D^β recovers the MAP decision rule in Eq. (7) choosing an outcome with the least LP value.

Remark 3. Our formulation of the ILP game in (7) supposes that the market equilibrium is achieved. This is a practically challenging point. The classical prediction market faces the same issue, and Chen and Pennock (2007, Theorem 8) showed a trade-off between the principal's total payout and the market liquidity. Hence, the principal needs more subsidy to gain higher liquidity. The same trade-off exists in our ILP game.

Remark 4. To compute the Gibbs decision rule (9), $K = |\overline{\Omega}|$ must not be too large to enumerate otherwise the normalizer is intractable in general.

Algorithm 1 Principal's play of an action

Input: agent's convex potential π

for $\boldsymbol{\omega} \in \overline{\Omega}$ **do**

  $\boldsymbol{\theta}^*(\boldsymbol{\omega}) \leftarrow \arg\min_{\boldsymbol{\theta}\in\mathcal{S}} \pi(\boldsymbol{\theta};\boldsymbol{\omega})$ ▷ Frank–Wolfe

end for

$\boldsymbol{\omega} \sim \exp(-\beta \langle \boldsymbol{\theta}^*(\boldsymbol{\omega}), \boldsymbol{\omega}\rangle)$ ▷ Rejection sampling

3.5 Computing Decision Rule

Consider the computation of the Gibbs decision rule. The decision rule is invoked in two situations: when the principal plays an action $\boldsymbol{\omega} \sim D^\beta[Q]$ and when the agent's report is scored by $S[Q, \boldsymbol{\omega}]$. Since it is intractable to deal with general belief forms, we hereafter suppose that the agent have a log-concave belief:

$$Q[\boldsymbol{\omega}](\boldsymbol{\theta}) \propto \exp(-\pi(\boldsymbol{\theta};\boldsymbol{\omega})), \tag{10}$$

where $\pi(\cdot;\boldsymbol{\omega})\colon \mathcal{S} \to \mathbb{R}$ is a convex L-smooth potential indexed by $\boldsymbol{\omega}$. In the case where we implement a log-concave belief by an ML model, we may implement $\pi(\cdot;\boldsymbol{\omega})$ by an input convex neural network (ICNN) (Amos et al., 2017), a neural network that is convex in its input by design. One can check ICNNs are smooth if the weight matrices are bounded. Implementing each agent with an ML model corresponds to machine learning markets (Storkey, 2011).

To play an action, principal's simple strategy is to invoke the rejection sampling for $D^\beta[Q]$. To this end, we need to have access to the unnormalized density value of $D^\beta[Q]$ (9), which requires the MAP state $\boldsymbol{\theta}^*(\boldsymbol{\omega})$ by solving the following optimization for every $\boldsymbol{\omega} \in \overline{\Omega}$:

$$\boldsymbol{\theta}^*(\boldsymbol{\omega}) = \operatorname*{arg\,min}_{\boldsymbol{\theta}\in\mathcal{S}} \pi(\boldsymbol{\theta};\boldsymbol{\omega}). \tag{11}$$

When adopted with the Frank–Wolfe algorithm, the iteration complexity to solve Eq. (11) for a fixed $\boldsymbol{\omega}$ is $O(LR^2\varepsilon^{-1})$, where $R := \operatorname{diam}(\mathcal{S})$ and $\varepsilon > 0$ is the tolerance parameter (Bubeck, 2015). Having computed $\boldsymbol{\theta}^*(\boldsymbol{\omega})$ for every $\boldsymbol{\omega} \in \overline{\Omega}$, now we can invoke the rejection sampling. The entire procedure for the principal to play a single action is listed in Algorithm 1. The whole computational complexity is $O(KLR^2\varepsilon^{-1})$.

To score agent's report, we need to evaluate $D^\beta[Q](\boldsymbol{\omega})$ for the played action $\boldsymbol{\omega}$. This requires the marginal inference of the decision rule (9). After the principal plays an action for the report Q by Algorithm 1, we have already computed $\boldsymbol{\theta}^*(\boldsymbol{\omega})$ for every $\boldsymbol{\omega} \in \overline{\Omega}$. Thus, we only compute the normalizer $Z_Q = \sum_{\boldsymbol{\omega}\in\overline{\Omega}} \exp(-\beta f(\boldsymbol{\omega};\boldsymbol{\theta}^*(\boldsymbol{\omega})))$. Its computational complexity is $O(dK)$, which is affordable unless K is so gigantic that enumeration of $\overline{\Omega}$ is intractable.

Dual perspective. The log-concave belief (10) has a nice dual interpretation in the ILP pipeline. First,

$$\begin{tikzcd}
\boxed{\text{LP potential } \psi(\boldsymbol{\omega})} \arrow[r, leftrightarrow, "\text{conjugate}"] \arrow[d, "\nabla_{\boldsymbol{\omega}}"] & \boxed{\text{Agent's belief } \psi^\star(\boldsymbol{\theta})} \arrow[d, "\nabla_{\boldsymbol{\theta}}"] \\
\boxed{\text{MAP state } \nabla\psi(\boldsymbol{\omega})} \arrow[r, leftrightarrow, "\text{inverse}"] & \boxed{\text{Reg. outcome } \nabla\psi^\star(\boldsymbol{\theta})}
\end{tikzcd}$$

Figure 3: Through the lens of the relaxed LP (12), agent's belief $\pi(\equiv \psi^\star(\boldsymbol{\theta}))$ is dual to the LP potential ψ. This commutative diagram shares the same structure as Bao and Sugiyama (2021, Figure 1), which characterizes the duality structure of canonical proper scoring rules.

we consider a *relaxed* LP of Eq. (4) with a convex potential. For a proper convex function $\psi\colon \Omega \to \mathbb{R}$,

$$\boldsymbol{\omega}^*_\psi(\boldsymbol{\theta}) := \operatorname*{arg\,min}_{\boldsymbol{\omega}\in\Omega} \langle \boldsymbol{\theta}, \boldsymbol{\omega}\rangle + \psi(\boldsymbol{\omega}). \tag{12}$$

The convex potential ψ represents *uncertainty* of an outcome $\boldsymbol{\omega}$—smaller potential value $\psi(\boldsymbol{\omega})$ indicates that $\boldsymbol{\omega}$ is less certain, or far from vertices of Ω. By Danskin's theorem (Blondel and Roulet, 2024, Section 11.2), we identify (12) with differentiable ψ as

$$\boldsymbol{\omega}^*_\psi(\boldsymbol{\theta}) = \operatorname*{arg\,max}_{\boldsymbol{\omega}\in\boldsymbol{\Omega}} \langle -\boldsymbol{\theta}, \boldsymbol{\omega}\rangle - \psi(\boldsymbol{\omega}) = \nabla\psi^\star(-\boldsymbol{\theta}),$$

where $\psi^\star(\boldsymbol{\theta}) := \max_{\boldsymbol{\omega}\in\Omega} \langle \boldsymbol{\theta}, \boldsymbol{\omega}\rangle - \psi(\boldsymbol{\omega})$ is the Fenchel–Legendre conjugate. Moreover, if ψ is of Legendre-type, $\nabla\psi^\star = [\nabla\psi]^{-1}$ holds (Rockafellar, 1970, Theorem 26.5). Therefore, the mirror map $\nabla\psi\colon \Omega \to \mathcal{S}$ induces the primal-dual relationship between an outcome $\boldsymbol{\omega}$ and state $\boldsymbol{\theta}$.

Next, we focus on the MAP state (11), where $\boldsymbol{\theta}^*(\boldsymbol{\omega})$ can be identified as the mirror map $\nabla\psi\colon \Omega \to \mathcal{S}$ with

$$\pi(\boldsymbol{\theta};\boldsymbol{\omega}) = \langle \boldsymbol{\theta}, \boldsymbol{\omega}\rangle + \psi^\star(\boldsymbol{\theta}).$$

Thus, π associated with agent's belief is interpreted as the dual potential $\psi^\star$. In this view, the dual potential $\psi^\star$ in agent's belief π penetrates into outcome uncertainty via the primal potential ψ (12). Figure 3 summarizes the duality between $\boldsymbol{\omega}$ and $\boldsymbol{\theta}$.

4 SIMULATION

We report a simulation result to minimally demonstrate that our Stackelberg model practically works.

Setup overview. In the ILP game, we used the von Mises–Fisher distribution to simulate the belief because of its simplicity. For each outcome ω, we prepare a direction parameter $\boldsymbol{\mu}_\omega \in \mathbb{R}^d$ to model the conditional belief on states by $Q[\omega](\boldsymbol{\theta}) \propto \exp(\boldsymbol{\mu}_\omega^\top \boldsymbol{\theta})$. Thus, in each round of the ILP game, the agent's report is

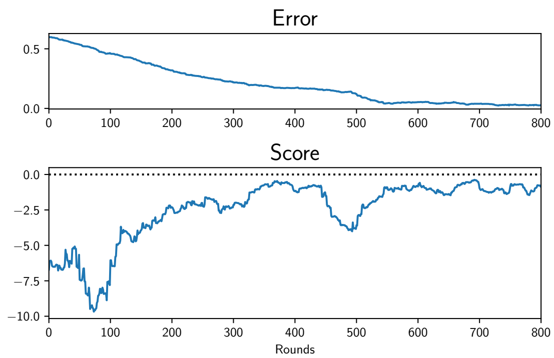

Figure 4: The simulation result. (Top) The mean-squared error of the agent's belief. (Bottom) The log score.

equivalent to the update of $K \times d$ matrix with each row being $\boldsymbol{\mu}_\omega$. For simplicity, we model the agent's report by one step of the gradient descent on the scoring rule.

We used the log score in Section 3.2. By assuming the Gibbs decision rule, the gradient of the log score (with respect to $\boldsymbol{\mu}_\omega$) can be calculated analytically. During the gradient calculation, we have to note $\nabla_{\boldsymbol{\mu}_\omega} \boldsymbol{\theta}^*(\omega) = \nabla_{\boldsymbol{\mu}_\omega}(\boldsymbol{\mu}_\omega / \|\boldsymbol{\mu}_\omega\|)$, which can be obtained via

$$\arg\max_{\boldsymbol{\theta}} Q[\omega](\boldsymbol{\theta}) = \arg\max_{\boldsymbol{\theta}} \boldsymbol{\mu}_\omega^\top \boldsymbol{\theta} = \frac{\boldsymbol{\mu}_\omega}{\|\boldsymbol{\mu}_\omega\|}.$$

All the other calculations follow from the chain rule.

Detailed setup. To sample outcomes from the decision rule, we randomly generated from the multinomial distribution. To sample states from the truthful agent's belief, we used the rejection sampling. For agent's belief, we used $\kappa = 10.0$ for the concentration parameter of von Mises–Fisher distributions. The inverse temperature of the ILP decision rule is set $\beta = 1.0$. To simulate the agent's report, we run the gradient descent with the step size 0.0005 and 800 updates. For outcomes $(\boldsymbol{\omega}_i)_i$ and the truthful agent's belief $(\boldsymbol{\mu}_\omega)_\omega$, we used

$$(\boldsymbol{\omega}_1, \boldsymbol{\omega}_2, \boldsymbol{\omega}_3, \boldsymbol{\omega}_4) = \left(\begin{bmatrix} 1.0 \\ 0.3 \end{bmatrix}, \begin{bmatrix} 0.5 \\ 0.1 \end{bmatrix}, \begin{bmatrix} 0.02 \\ 0.2 \end{bmatrix}, \begin{bmatrix} -0.3 \\ -0.8 \end{bmatrix} \right),$$

$$(\boldsymbol{\mu}_1, \boldsymbol{\mu}_2, \boldsymbol{\mu}_3, \boldsymbol{\mu}_4) = \left(\begin{bmatrix} 1.0 \\ 0.0 \end{bmatrix}, \begin{bmatrix} 0.0 \\ 1.0 \end{bmatrix}, \begin{bmatrix} 0.0 \\ 1.0 \end{bmatrix}, \begin{bmatrix} 0.6 \\ 0.8 \end{bmatrix} \right).$$

The agent's belief is initialized by the direction parameter $\boldsymbol{\mu}_\omega = [1/\sqrt{2}, 1/\sqrt{2}]^\top$ for all outcomes ω.

Result. We report the log score values and the relative errors of the elicited agent's beliefs in Fig. 4. As seen, the log score (maximum value: 0) gradually increases while the relative error decreases throughout the rounds. This indicates that the agents gradually converge to the truthful market belief, leading to the ILP decision rule in (7) based on crowd knowledge.

5 RELATED WORK

A relevant (yet different) field to inverse optimization is decision-focused learning (DFL), where we predict the unknown state of a downstream optimization problem (Wilder et al., 2019; Mandi et al., 2024). In its common formulation, we suppose that the form of the optimization problem of our interest $f(\boldsymbol{\omega}; \boldsymbol{\theta})$ is known—including the feasible region $\Omega(\ni \boldsymbol{\omega})$. The state $\boldsymbol{\theta}$ is given with a contextual information $\mathbf{x}$. Then, a predictive model takes context $\mathbf{x}$ as input features to output the state $\widehat{\boldsymbol{\theta}}$ in the first step, and the decision $\boldsymbol{\omega}$ is made by minimizing the objective function $f(\boldsymbol{\omega}; \widehat{\boldsymbol{\theta}})$ with the predicted $\widehat{\boldsymbol{\theta}}$ in the second step. *Smart predict-then-optimize* (SPO) is a standard approach to this problem (Donti et al., 2017; Elmachtoub and Grigas, 2022), which specifically minimizes the suboptimality $\ell(\widehat{\boldsymbol{\theta}}, \boldsymbol{\theta}) := f(\boldsymbol{\omega}^*(\widehat{\boldsymbol{\theta}}); \boldsymbol{\theta}) - f(\boldsymbol{\omega}^*(\boldsymbol{\theta}); \boldsymbol{\theta})$ defined via the optimal outcome $\boldsymbol{\omega}^*(\boldsymbol{\theta})$ of $f(\boldsymbol{\omega}; \boldsymbol{\theta})$ to make the predicted state $\widehat{\boldsymbol{\theta}}$ as close to the ground-truth state $\boldsymbol{\theta}$ as possible. Technically, implicit differentiation through the KKT conditions (Amos and Kolter, 2017; Donti et al., 2017; Wilder et al., 2019) or the suboptimality loss with a convex surrogate loss (Elmachtoub and Grigas, 2022) have been used. Theoretically, the generalization ability has been studied (El Balghiti et al., 2019; Liu and Grigas, 2021). Thus, DFL primarily focuses on a context-aligned predictive model of a state, while we focus on identifying outcome-aligned states.

Contextual inverse optimization is related to DFL, where historical observations $\{(\mathbf{x}_i, \boldsymbol{\omega}_i)\}_{i \in [N]}$, consisting of a context $\mathbf{x}_i$ and outcome $\boldsymbol{\omega}_i$, are given, and we aim to predict the unknown state from a context with high fidelity to the underlying context-outcome relationship (Chan et al., 2023). Unlike inverse optimization and DFL, one typically assumes no access to a state $\boldsymbol{\theta}$ but optimizes the worst-case suboptimality loss with respect to a state (Besbes et al., 2023). That being said, the identification of the underlying state is usually out of scope. This is a primary difference of contextual inverse optimization from our setup.

6 CONCLUSION

We proposed a new model to inverse linear optimization without relying on the principal's subjective criteria to recover the state. The proposed model is based on the Stackelberg game, where the agents act as the leader through prediction market and the principal plays the follower's role to solve forward optimization. Proper scoring rules for this prediction market can be derived based on Chen and Kash (2011). More efficient implementation of the decision rule is left open.

Acknowledgment

HB is supported by JST PRESTO (Grant No. JPMJPR24K6). SS is supported by JST ERATO (Grant No. JPMJER1903).

References

Jacob D Abernethy and Rafael M Frongillo. A characterization of scoring rules for linear properties. In *Proceedings on the 25th Annual Conference on Learning Theory*, pages 27.1–27.13, 2012.

Brandon Amos and J Zico Kolter. OptNet: Differentiable optimization as a layer in neural networks. In *Proceedings of the 34th International Conference on Machine Learning*, pages 136–145, 2017.

Brandon Amos, Lei Xu, and J Zico Kolter. Input convex neural networks. In *Proceedings of the 34th International Conference on Machine Learning*, pages 146–155, 2017.

Han Bao. Proper losses, moduli of convexity, and surrogate regret bounds. In *Proceedings of the 36th Conference on Learning Theory*, pages 525–547, 2023.

Han Bao and Masashi Sugiyama. Fenchel-Young losses with skewed entropies for class-posterior probability estimation. In *Proceedings of the 24th International Conference on Artificial Intelligence and Statistics*, pages 1648–1656, 2021.

Han Bao and Asuka Takatsu. Proper losses regret at least 1/2-order. *arXiv preprint arXiv:2407.10417*, 2024.

Andreas Bärmann, Sebastian Pokutta, and Oskar Schneider. Emulating the expert: Inverse optimization through online learning. In *Proceedings of the 34th International Conference on Machine Learning*, pages 400–410, 2017.

Omar Besbes, Yuri Fonseca, and Ilan Lobel. Contextual inverse optimization: Offline and online learning. *Operations Research*, 73:424–443, 2023.

Mathieu Blondel and Vincent Roulet. The elements of differentiable programming. *arXiv preprint arXiv:2403.14606*, 2024.

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

Andreas Buja, Werner Stuetzle, and Yi Shen. Loss functions for binary class probability estimation and classification: Structure and applications. *Technical Report*, 2005.

Didier Burton and Ph L Toint. On an instance of the inverse shortest paths problem. *Mathematical Programming*, 53:45–61, 1992.

Timothy CY Chan, Rafid Mahmood, and Ian Yihang Zhu. Inverse optimization: Theory and applications. *Operations Research*, 2023.

Yiling Chen and Ian A Kash. Information elicitation for decision making. In *Proceedings of the 10th International Conference on Autonomous Agents and Multiagent Systems*, pages 175–182, 2011.

Yiling Chen and David M Pennock. A utility framework for bounded-loss market makers. In *Proceedings of the 23rd Conference on Uncertainty in Artificial Intelligence*, pages 49–56, 2007.

Yiling Chen and Jennifer Wortman Vaughan. A new understanding of prediction markets via no-regret learning. In *Proceedings of the 11th ACM Conference on Electronic Commerce*, pages 189–198, 2010.

Robert B Dial. Minimal-revenue congestion pricing part I: A fast algorithm for the single-origin case. *Transportation Research Part B: Methodological*, 33 (3):189–202, 1999.

Robert B Dial. Minimal-revenue congestion pricing part II: An efficient algorithm for the general case. *Transportation Research Part B: Methodological*, 34 (8):645–665, 2000.

Priya Donti, Brandon Amos, and J Zico Kolter. Task-based end-to-end model learning in stochastic optimization. *Advances in Neural Information Processing Systems*, 30:5484–5494, 2017.

Othman El Balghiti, Adam N Elmachtoub, Paul Grigas, and Ambuj Tewari. Generalization bounds in the predict-then-optimize framework. *Advances in Neural Information Processing Systems*, 32:14412–14421, 2019.

Adam N Elmachtoub and Paul Grigas. Smart "predict, then optimize". *Management Science*, 68(1):9–26, 2022.

Rafael Frongillo, Nicholás Della Penna, and Mark D Reid. Interpreting prediction markets: a stochastic approach. *Advances in Neural Information Processing Systems*, 25:3266–3274, 2012.

Rafael M Frongillo and Ian A Kash. General truthfulness characterizations via convex analysis. *Games and Economic Behavior*, 130:636–662, 2021.

Kimia Ghobadi, Taewoo Lee, Houra Mahmoudzadeh, and Daria Terekhov. Robust inverse optimization. *Operations Research Letters*, 46(3):339–344, 2018.

Tilmann Gneiting and Adrian E Raftery. Strictly proper scoring rules, prediction, and estimation. *Journal of the American Statistical Association*, 102 (477):359–378, 2007.

Robin Hanson. Combinatorial information market design. *Information Systems Frontiers*, 5:107–119, 2003.

Garud Iyengar and Wanmo Kang. Inverse conic programming with applications. *Operations Research Letters*, 33(3):319–330, 2005.

Nicolas S Lambert, David M Pennock, and Yoav Shoham. Eliciting properties of probability distributions. In *Proceedings of the 9th ACM Conference on Electronic Commerce*, pages 129–138, 2008.

Heyuan Liu and Paul Grigas. Risk bounds and calibration for a smart predict-then-optimize method. *Advances in Neural Information Processing Systems*, 34:22083–22094, 2021.

Jayanta Mandi, James Kotary, Senne Berden, Maxime Mulamba, Victor Bucarey, Tias Guns, and Ferdinando Fioretto. Decision-focused learning: Foundations, state of the art, benchmark and future opportunities. *Journal of Artificial Intelligence Research*, 80:1623–1701, 2024.

Peyman Mohajerin Esfahani, Soroosh Shafieezadeh-Abadeh, Grani A Hanasusanto, and Daniel Kuhn. Data-driven inverse optimization with imperfect information. *Mathematical Programming*, 167:191–234, 2018.

Abraham Othman and Tuomas Sandholm. Decision rules and decision markets. In *Proceedings of the 9th International Conference on Autonomous Agents and Multiagent Systems*, pages 625–632, 2010.

Mattew Party, A Philip Dawid, and Steffen Lauritzen. Proper local scoring rules. *The Annals of Statistics*, 40(1):561–592, 2012.

Mark D Reid and Robert C Williamson. Composite binary losses. *Journal of Machine Learning Research*, 11:2387–2422, 2010.

Ralph Tyrrell Rockafellar. *Convex Analysis*, volume 28. Princeton University Press, 1970.

Leonard J Savage. Elicitation of personal probabilities and expectations. *Journal of the American Statistical Association*, 66(336):783–801, 1971.

Liangliang Shi, Gu Zhang, Haoyu Zhen, Jintao Fan, and Junchi Yan. Understanding and generalizing contrastive learning from the inverse optimal transport perspective. In *Proceedings of the 40th International Conference on Machine Learning*, pages 31408–31421, 2023.

Amos Storkey. Machine learning markets. In *Proceedings of the 14th International Conference on Artificial Intelligence and Statistics*, pages 716–724, 2011.

Chunlin Sun, Shang Liu, and Xiaocheng Li. Maximum optimality margin: A unified approach for contextual linear programming and inverse linear programming. In *Proceedings of the 40th International Conference on Machine Learning*, pages 32886–32912, 2023.

Bryan Wilder, Bistra Dilkina, and Milind Tambe. Melding the data-decisions pipeline: Decision-focused learning for combinatorial optimization. In *Proceedings of the AAAI Conference on Artificial Intelligence*, volume 33, pages 1658–1665, 2019.

Checklist

1. For all models and algorithms presented, check if you include:
   (a) A clear description of the mathematical setting, assumptions, algorithm, and/or model. [Yes] See Section 3.
   (b) An analysis of the properties and complexity (time, space, sample size) of any algorithm. [Yes] See Section 3.2 for the properness of scoring rules and Section 3.5 for the computational complexity of the decision rule.
   (c) (Optional) Anonymized source code, with specification of all dependencies, including external libraries. [Not Applicable]

2. For any theoretical claim, check if you include:
   (a) Statements of the full set of assumptions of all theoretical results. [Yes] See Section 3.2 for the statement to derive proper scoring rules.
   (b) Complete proofs of all theoretical results. [Yes] See Appendix A.
   (c) Clear explanations of any assumptions. [Yes] See Section 3.2 and Theorem 5 for the assumptions on the decision rule.

3. For all figures and tables that present empirical results, check if you include:
   (a) The code, data, and instructions needed to reproduce the main experimental results (either in the supplemental material or as a URL). [Not Applicable]
   (b) All the training details (e.g., data splits, hyperparameters, how they were chosen). [Yes] See Section 4.
   (c) A clear definition of the specific measure or statistics and error bars (e.g., with respect to the random seed after running experiments multiple times). [Not Applicable]
   (d) A description of the computing infrastructure used. (e.g., type of GPUs, internal cluster, or cloud provider). [Not Applicable]

4. If you are using existing assets (e.g., code, data, models) or curating/releasing new assets, check if you include:

(a) Citations of the creator If your work uses existing assets. [Not Applicable]

(b) The license information of the assets, if applicable. [Not Applicable]

(c) New assets either in the supplemental material or as a URL, if applicable. [Not Applicable]

(d) Information about consent from data providers/curators. [Not Applicable]

(e) Discussion of sensible content if applicable, e.g., personally identifiable information or offensive content. [Not Applicable]

5. If you used crowdsourcing or conducted research with human subjects, check if you include:

(a) The full text of instructions given to participants and screenshots. [Not Applicable]

(b) Descriptions of potential participant risks, with links to Institutional Review Board (IRB) approvals if applicable. [Not Applicable]

(c) The estimated hourly wage paid to participants and the total amount spent on participant compensation. [Not Applicable]

A Proof

Proof of Theorem 5. First, we show S is (strictly) proper with the given construction of S. Fix an arbitrary $P, Q \in \mathcal{P}$ with $P \neq Q$. Then,

$$
\begin{aligned}
V(P,Q) &= \int_{\mathcal{S}} \mathrm{d}\boldsymbol{\theta} \sum_{\boldsymbol{\omega}} D[Q](\boldsymbol{\omega}) P[\boldsymbol{\omega}](\boldsymbol{\theta}) S[Q,\boldsymbol{\omega}](\boldsymbol{\theta}) \\
&= \Lambda(Q) - \langle G_Q, Q \rangle + \int_{\mathcal{S}} \mathrm{d}\boldsymbol{\theta} \sum_{\boldsymbol{\omega}: D[Q](\boldsymbol{\omega})>0} G_Q[\boldsymbol{\omega}](\boldsymbol{\theta}) P[\boldsymbol{\omega}](\boldsymbol{\theta}) \\
&= \Lambda(Q) - \langle G_Q, Q \rangle + \langle G_Q, P \rangle \\
&= \Lambda(Q) + \langle G_Q, P - Q \rangle,
\end{aligned}
$$

where the third line is attributed to the condition $G_Q[\boldsymbol{\omega}](\boldsymbol{\theta}) = 0$ for $D[Q](\boldsymbol{\omega}) = 0$. Similarly, we have $V(P,P) = \Lambda(P)$. By the convexity of Λ, $V(P,P) \geq V(P,Q)$. When Λ is strictly convex, the inequality becomes strict.

Now we consider a regular (strictly) proper scoring rule S for D and show that S must be of the form of (6). Define

$$
\Lambda(P) \coloneqq V(P,P) \quad \text{and} \quad G_P[\boldsymbol{\omega}](\boldsymbol{\theta}) \coloneqq D[P](\boldsymbol{\omega}) S[P,\boldsymbol{\omega}](\boldsymbol{\theta}).
$$

Then, $\Lambda(P) = \sup_Q V(P,Q)$ by the properness of S, meaning that Λ is a pointwise supremum of convex functions $V(\cdot, Q)$, and thus Λ is convex. If $D[P](\boldsymbol{\omega}) = 0$, $G_P[\boldsymbol{\omega}](\boldsymbol{\theta}) = 0$ is satisfied by definition. To see S is of the form of (6), we first see for $Q \neq P$,

$$
\begin{aligned}
\Lambda(P) + \langle Q - P, G_P \rangle &= \int \mathrm{d}\boldsymbol{\theta} \sum_{\boldsymbol{\omega}} D[P](\boldsymbol{\omega}) Q[\boldsymbol{\omega}](\boldsymbol{\theta}) S[P,\boldsymbol{\omega}](\boldsymbol{\theta}) \\
&= V(Q,P) \\
&\leq \Lambda(Q),
\end{aligned}
$$

where the last inequality is due to the properness of S and it becomes strict when S is strictly proper. This inequality implies that $G_P \in \partial\Lambda(P)$, and that Λ is strictly convex if S is strictly proper. Therefore, for P and $\boldsymbol{\omega}$ such that $D[P](\boldsymbol{\omega}) > 0$,

$$
\Lambda(P) - \langle G_P, P \rangle + \frac{G_P[\boldsymbol{\omega}](\boldsymbol{\theta})}{D[P](\boldsymbol{\omega})} = \frac{G_P[\boldsymbol{\omega}](\boldsymbol{\theta})}{D[P](\boldsymbol{\omega})} = S[P,\boldsymbol{\omega}](\boldsymbol{\theta}).
$$

Hence, the converse is proven. $\square$