



INFORMS TutORials in Operations Research

Publication details, including instructions for authors and subscription information:
<http://pubsonline.informs.org>

Integrating Prediction/Estimation and Optimization with Applications in Operations Management

<https://orcid.org/0000-0002-0984-4846>Meng Qi, <https://orcid.org/0000-0003-4538-8312>Zuo-Jun (Max) Shen

To cite this entry: <https://orcid.org/0000-0002-0984-4846>Meng Qi, <https://orcid.org/0000-0003-4538-8312>Zuo-Jun (Max) Shen. Integrating Prediction/Estimation and Optimization with Applications in Operations Management. *In* INFORMS TutORials in Operations Research. Published online: 14 Oct 2022; 36–58.
<https://doi.org/10.1287/educ.2022.0249>

Full terms and conditions of use: <https://pubsonline.informs.org/Publications/Librarians-Portal/PubsOnLine-Terms-and-Conditions>

This article may be used only for the purposes of research, teaching, and/or private study. Commercial use or systematic downloading (by robots or other automatic processes) is prohibited without explicit Publisher approval, unless otherwise noted. For more information, contact permissions@informs.org.

The Publisher does not warrant or guarantee the article's accuracy, completeness, merchantability, fitness for a particular purpose, or non-infringement. Descriptions of, or references to, products or publications, or inclusion of an advertisement in this article, neither constitutes nor implies a guarantee, endorsement, or support of claims made of that product, publication, or service.

Copyright © 2022, INFORMS

Please scroll down for article—it is on subsequent pages



With 12,500 members from nearly 90 countries, INFORMS is the largest international association of operations research (O.R.) and analytics professionals and students. INFORMS provides unique networking and learning opportunities for individual professionals, and organizations of all types and sizes, to better understand and use O.R. and analytics tools and methods to transform strategic visions and achieve better outcomes. For more information on INFORMS, its publications, membership, or meetings visit <http://www.informs.org>

Integrating Prediction/Estimation and Optimization with Applications in Operations Management

Meng Qi,^{a,*} Zuo-Jun (Max) Shen^{b,c}

^aCornell University, Ithaca, New York 14853; ^bUniversity of Hong Kong, Pokfulam, Hong Kong;

^cUniversity of California, Berkeley, Berkeley, California 94720

*Corresponding author

Contact: mq56@cornell.edu,  <https://orcid.org/0000-0002-0984-4846> (MQ);

maxshen@hku.hk,  <https://orcid.org/0000-0003-4538-8312> (Z-J(M)S)

Abstract Big data provide new opportunities to tackle one of the main difficulties in decision-making systems—the uncertain behavior that follows unknown probability distribution. Standard data-driven approaches usually consist of two steps. The first step involves predicting or estimating the uncertainty behavior using data. Then the second step requires finding decisions that optimize an objective function that depends on the output of the first step. Instead of the classical two-step predict-then-optimize/estimate-then-optimize procedure, this tutorial examines data-driven solutions that integrate these two steps. We first introduce the problem formulation as a contextual stochastic optimization problem. In this formulation, the objective function depends on the unknown uncertainty, and the distribution of the uncertainty is associated with some contextual information. Massive data are often available to solve this problem, including historical observations of the uncertainty and contextual information. Therefore, machine learning tools have become an important technique for achieving integrated data-driven solutions. Yet it is noteworthy that the goal of the integrated data-driven solution is very different from traditional predictive tasks for machine learning. Moreover, different integrated data-driven methods have shown applicability in many real-world decision-making situations, including supply chain management, portfolio optimization, and power system operations. To demonstrate this, we review current applications of integrated methods in application areas in operations management.

Funding: The authors are grateful for the support from the National Key R&D Program of China [Grant 2018YFB1700600].

Keywords data-driven decision-making • machine learning • operations management

1. Introduction

One of the main difficulties in decision-making systems is the uncertain behavior of some parameters, which are often viewed as random variables following some unknown ground-truth probability distributions. Traditional models often require the knowledge or an assumption of the underlying distribution of the random variable of interest and apply *stochastic optimization* (SO) frameworks to solve the optimal decision (see Birge and Louveaux [8] and Dantzig [13] for more details). However, it is often not realistic to assume that the decision maker knows the ground-truth distribution. Instead, big data provide opportunities to tackle the uncertainty in decision making when there has been tremendous growth in the availability of voluminous high-quality data in the past decade.

Besides the increasing availability of data, the presence of *contextual information* (also called features, covariates, or side information) is another emergent trend in the modern decision-making system. In this case, the ground-truth distribution of the random parameter of interest depends on the contextual information, and the decision maker seeks personalized solutions associated with features. With the presence of features, the available data often include historical observations of the random parameter and features. In this tutorial, we consider the case where data samples are collected from an offline source prior to the decision-making process.

Standard data-driven approaches with offline data usually consist of two steps. The first step involves predicting or estimating the uncertain parameter using data. The prediction/estimation obtained in this stage can be either probabilistic or a point estimation. In the former case, the distributional estimation can be parametric or nonparametric. In the latter case, various regression models often come into play and output a point estimator. In the second step, the output estimation or prediction is incorporated into an optimization problem, which is then solved for the ultimate decision. Instead of the classical two-step predict-then-optimize (PTO) procedure, this tutorial examines a current emerging trend of methods that integrate these two steps.

This tutorial reviews recent progress in this direction. Section 2 presents the problem setting of contextual stochastic optimization and describes the available data sets. Section 3 then reviews existing methodologies that integrate these two steps. Section 4 demonstrates current applications in operations management.

2. Problem Setting

In this part, we specify notation and a general problem formulation. The static decision-making problem with uncertainty and context can be formulated as a contextual stochastic optimization problem.

2.1. Constrained Contextual Stochastic Optimization

Let $\mathbf{z} \in \mathcal{Z} \subseteq \mathbb{R}^k$ denote the random variable that characterizes the random parameter in the problem setting, and let $\mathbf{x} \in \mathcal{X} \subseteq \mathbb{R}^p$ denote the associated contextual information that affects the distribution of \mathbf{z} . Suppose that the ground-truth distribution of \mathbf{z} conditioned on \mathbf{x} is denoted by $\mathbb{P}_{\mathbf{z}|\mathbf{x}}$. We also let $\mathbb{P}_{(\mathbf{x}, \mathbf{z})}$ denote the joint distribution of (\mathbf{x}, \mathbf{z}) , and let $\mathbb{P}_{\mathbf{x}}$ denote the marginal distribution of \mathbf{x} .

The decision maker aims to make a decision $\mathbf{w} \in S \subseteq \mathbb{R}^d$, where S is a feasible set of decisions characterized by a group of constraints. We assume that S is a convex set. Suppose there is a cost function $c(\cdot, \cdot) : S \times \mathcal{Z} \rightarrow \mathbb{R}$. We assume that the cost function $c(\cdot, \mathbf{z})$ is convex with respect to the decision \mathbf{w} given any value of \mathbf{z} . The goal of the decision maker is to minimize the expected cost conditioned on a given feature vector, which can be formulated as the following stochastic optimization problem:

$$\min_{\mathbf{w} \in S} \mathbb{E}[c(\mathbf{w}, \mathbf{z})|\mathbf{x}], \tag{1}$$

where the expectation is taken with respect to the true conditional distribution $\mathbb{P}_{\mathbf{z}|\mathbf{x}}$. We denote the optimal objective value of (1) by $v^*(\mathbf{x}) := \min_{\mathbf{w} \in S} \mathbb{E}[c(\mathbf{w}, \mathbf{z})|\mathbf{x}]$ and represent the set of corresponding optimal solutions as

$$W^*(\mathbf{x}) = \arg \min_{\mathbf{w} \in S} \mathbb{E}[c(\mathbf{w}, \mathbf{z})|\mathbf{x}].$$

Note that, without further assumption, there might be multiple optimal solutions, so that $W^*(\mathbf{x})$ is an optimal solution set. The decision maker aims to identify an optimal decision policy, which gives personalized solutions corresponding to input features. This can be achieved by calling an optimal solution mapping $w^*(\cdot)$, which outputs an optimal solution $w^*(\mathbf{x}) \in W^*(\mathbf{x})$ given any input feature \mathbf{x} . Note that if $W^*(\mathbf{x})$ contains multiple optimal solutions, the optimal solution mapping $w^*(\cdot)$ is not unique and needs to be specified by a selection rule.

However, without knowing the conditional distribution $\mathbb{P}_{\mathbf{z}|\mathbf{x}}$, (1) cannot be solved. Therefore, neither $v^*(\cdot)$, $W^*(\cdot)$, nor $w^*(\cdot)$ is obtainable. To overcome this issue, data-driven methods are widely adopted to approximate the desired personalized optimal decision.

2.2. Data

One of the most common settings for data-driven decision making is that there is an available offline data set that contains historical observations of the random parameter and features. We let $\{\mathbf{x}_i, \mathbf{z}_i\}_{i=1}^n$ denote a data set with n samples, where $\mathbf{x}_i, i = 1, \dots, n$ denote the samples of \mathbf{x} and $\mathbf{z}_i, i = 1, \dots, n$ denote the samples of \mathbf{z} . If not stated otherwise, we assume that the data samples are *independent and identically distributed* (i.i.d.); that is, $\{\mathbf{x}_i, \mathbf{z}_i\}_{i=1}^n$ are generated independently from the joint distribution $\mathbb{P}_{(\mathbf{x}, \mathbf{z})}$.

Conventional data-driven approaches to solving (1) adopt a two-step framework, where the first step involves estimations or predictions of the unknown parameter \mathbf{z} based on data. In the second step, those estimations/predictions are considered as inputs to an optimization problem. As discussed in Grigas et al. [18], sometimes the phrases “prediction” and “estimation” are not distinguished in the literature. In this tutorial, we follow the convention that specifically lets *prediction* represent point estimates of random parameters \mathbf{z} conditioned on features \mathbf{x} and lets *estimation* denote distributional estimations of the conditional distribution estimating $\mathbb{P}_{\mathbf{z}|\mathbf{x}}$.

2.3. Motivating Examples

In this part, we demonstrate some examples that share the general formulation stated in Section 2.1.

Example 1 (Single-Product Newsvendor Problem). The single-item, unconstrained newsvendor problem aims to find the optimal replenishment quantity for a product regarding random demand, which is denoted by \mathbf{z} . The decision, $\mathbf{w} \in \mathbb{R}$, represents the order quantity. The distribution of demand may depend on associated contextual information, denoted as \mathbf{x} . Some examples of \mathbf{x} can be promotions, holiday seasons, and product features, to name just a few. The objective is to minimize the total inventory cost, including the holding cost, if there is overstock and stockout cost if there is understock. The holding cost is proportional to the overstocked items, if there is any, and the unit holding cost is denoted by h . Similarly, the unit stockout cost is denoted by b . Then, the single-product, unconstrained newsvendor problem can be represented as

$$\min_{\mathbf{w} \geq 0} \mathbb{E}[h(\mathbf{w} - \mathbf{z})^+ + b(\mathbf{z} - \mathbf{w})^+ | \mathbf{x}]. \quad (2)$$

A natural extension of the single-product, unconstrained newsvendor problem is the multi-product version. When there are multiple products, there often exist productwise joint constraints. In the following example, we review the multiproduct newsvendor problem with a budget constraint.

Example 2 (Multiproduct Newsvendor Problem). The newsvendor aims to decide optimal replenishment quantities for d different products, denoted by $\mathbf{w} \in \mathbb{R}^d$. The random demand of all products is characterized by the random vector $\mathbf{z} \in \mathbb{R}^d$. If there is a total budget capacity $C > 0$ on the total order quantities, the feasible set can be formulated as $S := \{\mathbf{w} \in \mathbb{R}^d : \sum_{l=1}^d w_l \leq C, \mathbf{w} \geq 0\}$. Similar to (2), the objective is to minimize the total inventory cost of all products. Then the multiproduct, constrained newsvendor problem can be formulated as

$$\begin{aligned} \min_{\mathbf{w}} \quad & \mathbb{E} \left[\sum_{l=1}^d h_l (w_l - z_l)^+ + b_l (z_l - w_l)^+ \mid \mathbf{x} \right] \\ \text{s.t.} \quad & \sum_{l=1}^d w_l \leq C, \\ & \mathbf{w} \geq 0, \end{aligned} \quad (3)$$

where the unit holding and stockout costs, h_l and b_l , can be different for different products.

Portfolio optimization is another classic example in operations management. In the following example, we state the portfolio optimization problem with mean-variance quadratic objective.

Example 3 (Mean-Variance Portfolio Optimization). The mean-variance portfolio optimization problem aims to maximize the expected return while minimizing the variance of return. Suppose there are d possible assets to invest in, and let $\mathbf{z} \in \mathbb{R}^d$ denote the random return of the assets. The portfolio $\mathbf{w} \in \mathbb{R}^d$ represents the percentage of the total capital investment on a set of assets. The mean-variance objective is formulated as

$$c((\mathbf{w}, w_0), \mathbf{z}) := -\rho \mathbf{z}^T \mathbf{w} + \frac{\delta}{2} (\mathbf{z}^T \mathbf{w} - w_0)^2,$$

where $w_0 \in \mathbb{R}$ is an auxiliary decision variable such that minimizing $\mathbb{E}[c((\mathbf{w}, w_0), \mathbf{z})]$ is equivalent to minimizing the weighted sum of variance of return and negative averaged return. There are linear equality and inequality constraints, denoted by $A\mathbf{w} = \mathbf{b}$ and $G\mathbf{w} \leq \mathbf{h}$, respectively. Therefore, the decision maker aims to solve the following constrained contextual stochastic optimization problem:

$$\begin{aligned} \min_{\mathbf{w}, w_0} \quad & \mathbb{E}[c((\mathbf{w}, w_0), \mathbf{z})|\mathbf{x}] \\ \text{s.t.} \quad & A\mathbf{w} = \mathbf{b}, \\ & G\mathbf{w} \leq \mathbf{h}. \end{aligned}$$

Another important application is the scheduling problem for grid generation in power systems.

Example 4 (Grid Generation Scheduling). The grid generation scheduling problem requires the operator to decide the electricity generated for an upcoming time period $1, \dots, T$. The grid operator aims to satisfy random electricity demands z_1, \dots, z_T over the grid by deciding on the power generation w_1, \dots, w_T . There are penalties for unit generation excess (more power generated than needed) as well as unit generation shortage (power generated cannot meet the demand). The problem can be formulated as

$$\begin{aligned} \min_{\mathbf{w} \in \mathbb{R}^T} \quad & \mathbb{E} \left[\sum_{t=1}^T c_h(w_t - z_t)^+ + c_b(z_t - w_t)^+ + \rho(w_t - z_t)^2 \mid \mathbf{x} \right] \\ \text{s.t.} \quad & |w_t - w_{t-1}| \leq \delta, \quad \forall t = 2, \dots, T, \end{aligned}$$

where c_h and c_b denote unit excess cost and unit shortage cost, respectively. The coefficient of a quadratic regularization term is denoted by ρ , and δ is the restriction on the change between consecutive time periods.

Note that the formulation of the grid generator scheduling problem is similar to that of the multiproduct newsvendor problem stated in Example 2. Rather than a total budget constraint, the grid generation problem restricts the difference in electricity output between time periods $t - 1$ and t as a result of physical limits in the grid.

3. Methodologies and Performance Guarantees

The remainder of this section summarizes classical two-step data-driven methods and corresponding methodologies that integrate the two steps. We start with point predictions.

3.1. Learning Point Prediction with Regression Models

3.1.1. PTO. Classical PTO approaches include two steps. The first step focuses on providing point predictions for the random parameter, whereas the second stage is devoted to finding optimal decisions based on these predictions. The state-of-art regression models are frequently applied for the prediction stage in PTO methods. Herein, we summarize a list of critical ingredients of a regression model.

1. A *hypothesis class* \mathcal{H} : A hypothesis class is often used to capture the mapping from features to the predicted random parameters. Here, \mathcal{H} denotes the hypothesis class that consists of a group of functions $h: \mathcal{X} \rightarrow \mathcal{Z}$. Then a prediction of \mathbf{z} , denoted by $\hat{\mathbf{z}} := h(\mathbf{x})$, is interpreted as the predicted random parameter given feature \mathbf{x} .

2. A *loss function* l : A loss function $l(\cdot, \cdot): \mathcal{Z} \times \mathcal{Z} \rightarrow \mathbb{R}_+$ takes a realized random parameter and a predicted random parameter as inputs. The loss function is supposed to quantify the prediction error.

In PTO methods, the loss function is purely out of statistical concern and completely independent of the optimization problem (1). A common example of l is the least square (LS) loss,

$$l_{\text{LS}}(\hat{\mathbf{z}}, \mathbf{z}) = \|\mathbf{z} - \hat{\mathbf{z}}\|^2. \quad (4)$$

Being a critical component of modern machine learning techniques, regression models now can obtain accurate predictions and have been widely used in a variety of application fields. We refer to Sarker [29] for a thorough and detailed review of the machine learning models for regression and their applications in prediction.

In PTO methods, the prediction model is determined by solving an empirical risk minimization (ERM) problem in the training process

$$\min_{h \in \mathcal{H}} \frac{1}{n} \sum_{i=1}^n l(h(\mathbf{x}_i), \mathbf{z}_i), \quad (5)$$

and we let \hat{h} denote the prediction model obtained by solving (5). It is worth highlighting that in the objective function of (5) there may exist an additional regularization term $\varphi(\cdot): \mathcal{H} \rightarrow \mathbb{R}$. That is to say, instead of minimizing $\frac{1}{n} \sum_{i=1}^n l(h(\mathbf{x}_i), \mathbf{z}_i)$, the training process minimizes the regularized empirical risk $\frac{1}{n} \sum_{i=1}^n l(h(\mathbf{x}_i), \mathbf{z}_i) + \lambda \varphi(h)$, with λ denoting the coefficient of regularization. There are various options for the regularization term. For example, if the hypothesis class \mathcal{H} is a group of linear functions $h_{\beta}(\mathbf{x}) = \beta^T \mathbf{x}$, then two commonly used regularization terms are LASSO (least absolute shrinkage and selection operator) regularization (i.e., $\varphi^{\text{LASSO}}(h_{\beta}) := \|\beta\|_1$) and Ridge regularization (i.e., $\varphi^{\text{Ridge}}(h_{\beta}) := \|\beta\|_2$). In practice, adding regularization terms often improves (finite-sample) accuracy while causing biased prediction.

In the optimization stage, instead of solving (1), the decision maker plugs in the prediction $\hat{h}(\mathbf{x})$ given the value of \mathbf{x} and solves

$$\min_{\mathbf{w} \in S} c(\mathbf{w}, \hat{h}(\mathbf{x})). \quad (6)$$

Note that (6) is an approximation of the original stochastic optimization (1) with point predictions. Without stochasticity, (6) is generally easier to solve by minimizing a deterministic version of (1). Unfortunately, in general, there is no guarantee on the decision obtained by solving (6). For the special case where the optimization stage has a linear objective function (i.e., $c(\mathbf{w}, \mathbf{z}) := \mathbf{z}^T \mathbf{w}$), the decision would be optimal in an ideal world such that $\hat{h}(\mathbf{x})$ recovers the true value of $\mathbb{E}[\mathbf{z}|\mathbf{x}]$ for any $\mathbf{x} \in \mathcal{X}$. However, it is most likely that $\hat{h}(\mathbf{x})$ does not equal $\mathbb{E}[\mathbf{z}|\mathbf{x}]$. In such cases, the prediction accuracy of $\hat{h}(\mathbf{x})$ does not lead to decision optimality obtained by solving (6). Loke et al. [23] demonstrated that there is a bias-variance trade-off in

the estimation of objective function $\mathbb{E}_{z|x}[\mathbf{z}]^T \mathbf{w}$. In fact, the decision \mathbf{w} nonlinearly transforms the bias-variance curve of $\hat{\mathbf{z}}$ to the performance curve $\hat{\mathbf{z}}^T \mathbf{w}$. Thus, the prediction accuracy does not align with the performance optimality. This explains why improving the prediction accuracy might be counterproductive (we refer to proposition 1 in Loke et al. [23] for details).

3.1.2. Smart “Predict, Then Optimize.” As PTO methods ignore the optimization goal in the prediction stage, it is a natural idea to produce “smarter” predictions by integrating the optimization goal into point predictions. One way to achieve this is to twist the loss function l according to the optimization problem.

Elmachtoub and Grigas [17] recently proposed the smart “predict, then optimize” (SPO) framework that trains prediction models to aim at minimizing decision error rather than purely statistical prediction error. The SPO framework focuses on the special case where the objective function is a linear combination of the random parameter and the decision—that is, $c(\mathbf{w}, \mathbf{z}) := \mathbf{z}^T \mathbf{w}$, with $\mathbf{w} \in \mathbb{R}^d$ and $\mathbf{z} \in \mathbb{R}^d$. This special case with a linear objective function can be found in applications such as the shortest path problem. Moreover, the authors also mentioned that their method can accommodate piecewise linear objective functions as well.

In this section, we review the main results of the SPO framework and slightly abuse the notation by letting $W^*(\hat{\mathbf{z}})$ denote the optimal solution set based on input prediction $\hat{\mathbf{z}}$ (i.e., $W^*(\hat{\mathbf{z}}) \in \arg \min_{\mathbf{w} \in S} \hat{\mathbf{z}}^T \mathbf{w}$). Similarly, we let $v^*(\hat{\mathbf{z}}) = \min_{\mathbf{w} \in S} \hat{\mathbf{z}}^T \mathbf{w}$ denote the corresponding optimal objective value. In this setting, we restate the formulation of the stochastic optimization (1) as

$$\min_{\mathbf{w} \in S} \mathbb{E}[\mathbf{z}^T \mathbf{w} | \mathbf{x}] = \min_{\mathbf{w} \in S} \mathbb{E}[\mathbf{z} | \mathbf{x}]^T \mathbf{w}.$$

With a prediction of \mathbf{z} , denoted by $\hat{\mathbf{z}}$, the SPO framework then solves the deterministic optimization problem:

$$\min_{\mathbf{w} \in S} \hat{\mathbf{z}}^T \mathbf{w}.$$

In contrast to the two-step PTO approaches where the loss function l is purely out of statistical concerns and independent of the optimization problem, the SPO framework adopts a loss function that quantifies the excess cost incurred in the downstream optimization stage because of the imprecise prediction. The SPO loss is stated as in the following definition.

Definition 1 (Unambiguous SPO Loss (Elmachtoub and Grigas [17])). Given the prediction of random parameter $\hat{\mathbf{z}}$ and a realized random parameter \mathbf{z} , the (unambiguous) true SPO loss $l_{\text{SPO}}(\hat{\mathbf{z}}, \mathbf{z})$ is defined as

$$l_{\text{SPO}}(\hat{\mathbf{z}}, \mathbf{z}) := \max_{\mathbf{w} \in W^*(\hat{\mathbf{z}})} \{\mathbf{z}^T \mathbf{w}\} - v^*(\hat{\mathbf{z}}).$$

The SPO loss l_{SPO} characterizes the maximum value of the excess cost given an inaccurate prediction $\hat{\mathbf{z}}$ among the optimal solution set $W^*(\hat{\mathbf{z}})$. Similar to the PTO approach, the prediction model can be determined by minimizing the empirical SPO risk:

$$\min_{h \in \mathcal{H}} \frac{1}{n} \sum_{i=1}^n l_{\text{SPO}}(h(\mathbf{x}_i), \mathbf{z}_i), \tag{7}$$

and we define the empirical SPO risk as

$$\hat{R}_{\text{SPO}}(h) := \frac{1}{n} \sum_{i=1}^n l_{\text{SPO}}(h(\mathbf{x}_i), \mathbf{z}_i).$$

By adopting the SPO loss function, the training process of the prediction model integrates the information of the subsequent optimization problem. However, the formulation (7) is difficult to solve and is NP-hard in many cases. Therefore, the authors derived a tractable

surrogate loss $l_{\text{SPO}+}(\cdot, \cdot)$ based on the dual representation of the SPO loss l_{SPO} . We state the definition of $l_{\text{SPO}+}$ for completeness and refer readers to Elmachtoub and Grigas [17] for the details of derivation.

Definition 2 (SPO+ Loss (Elmachtoub and Grigas [17])). Given a prediction $\hat{\mathbf{z}}$ and a realization \mathbf{z} of the random parameter, the SPO+ loss is defined as

$$l_{\text{SPO}+}(\hat{\mathbf{z}}, \mathbf{z}) := \max_{\mathbf{w} \in S} \{ \mathbf{z}^T \mathbf{w} - 2\hat{\mathbf{z}}^T \mathbf{w} \} + 2\hat{\mathbf{z}}^T \mathbf{w}^*(\mathbf{z}) - v^*(\mathbf{z}).$$

Replacing the SPO loss function with the SPO+ loss function, (7) becomes

$$\min_{h \in \mathcal{H}} \frac{1}{n} \sum_{i=1}^n l_{\text{SPO}+}(h(\mathbf{x}_i), \mathbf{z}_i), \quad (8)$$

and we denote the SPO+ empirical risk as

$$\hat{R}_{\text{SPO}+}(h) := \frac{1}{n} \sum_{i=1}^n l_{\text{SPO}+}(h(\mathbf{x}_i), \mathbf{z}_i).$$

Elmachtoub and Grigas [17] demonstrated that $l_{\text{SPO}+}$ is a convex function of the prediction $\hat{\mathbf{z}}$; thus the objective of (8) is also convex in $\hat{\mathbf{z}}$. To address the issue that $l_{\text{SPO}+}$ in general is nondifferentiable, the authors provided a subgradient of $l_{\text{SPO}+}$. It is worth pointing out that because the SPO framework focuses on linear optimization problems, it can be easily applied to combinatorial or mixed-integer feasible regions. Suppose the feasible set of the optimization problem is a bounded but possibly nonconvex or nonclosed set \tilde{S} ; then the SPO framework can be applied by replacing \tilde{S} with its convex hull S .

As the size of data samples grows to infinity, the empirical risk minimization problem given a loss function l becomes the associated true risk minimization problem. If the loss function is the LS loss l_{LS} as defined in (4), it is straightforward that the minimizer of the true risk,

$$R_{\text{LS}}(h) := \mathbb{E}_{(\mathbf{x}, \mathbf{z}) \sim \mathbb{P}_{(\mathbf{x}, \mathbf{z})}} [l_{\text{LS}}(h(\mathbf{x}), \mathbf{z})],$$

is almost surely equal to $\mathbb{E}[\mathbf{z}|\mathbf{x}]$ if $\mathbb{E}[\mathbf{z}|\mathbf{x}]$, viewed as a function of \mathbf{x} , is in the hypothesis class \mathcal{H} . Moreover, it can be shown that any minimizer of the SPO true risk

$$R_{\text{SPO}}(h) := \mathbb{E}_{(\mathbf{x}, \mathbf{z}) \sim \mathbb{P}_{(\mathbf{x}, \mathbf{z})}} [l_{\text{SPO}}(h(\mathbf{x}), \mathbf{z})] \quad (9)$$

and the SPO+ true risk

$$R_{\text{SPO}+}(h) := \mathbb{E}_{(\mathbf{x}, \mathbf{z}) \sim \mathbb{P}_{(\mathbf{x}, \mathbf{z})}} [l_{\text{SPO}+}(h(\mathbf{x}), \mathbf{z})] \quad (10)$$

is also almost surely equal to $\mathbb{E}[\mathbf{z}|\mathbf{x}]$ under certain conditions. Herein, we restate the assumptions and results in Elmachtoub and Grigas [17] for completeness.

Assumption 1. Suppose there is no model misspecification of the hypothesis class \mathcal{H} ; that is, $\mathbb{E}[\mathbf{z}|\mathbf{x}]$, viewed as a function of \mathbf{x} , is in the hypothesis class \mathcal{H} . Then the following holds.

- (A) Almost surely, $W^*(\mathbb{E}[\mathbf{z}|\mathbf{x}])$ is a singleton.
- (B) For all $\mathbf{x} \in \mathcal{X}$, the distribution of $\mathbf{z}|\mathbf{x}$ is continuous and centrally symmetric about its mean $\mathbb{E}[\mathbf{z}|\mathbf{x}]$.
- (C) The interior of the feasible region S is nonempty.

Theorem 1 (Fisher Consistency of the SPO Framework). Suppose parts (A)–(C) of Assumption 1 hold. Then, any minimizer of the SPO+ true risk (10) is almost surely (over the distribution of \mathbf{x}) equal to $\mathbb{E}[\mathbf{z}|\mathbf{x}]$ and is also a minimizer of the SPO true risk (9). Thus, we say that the SPO+ loss function is Fisher consistent with respect to the SPO loss.

In essence, this theorem states that, under certain conditions, the hypothesis that minimizes the SPO+ true risk is the regression function $\mathbb{E}[\mathbf{z}|\mathbf{x}]$, and it also minimizes the SPO true risk. As a surrogate loss of the SPO loss, the SPO+ loss introduces the convexity of $\hat{\mathbf{z}}$ and thus makes it easier to develop computational approaches. In the meantime, SPO+ loss is Fisher consistent with respect to the SPO loss. The consistency result of SPO+ loss means that minimizing the SPO+ loss becomes equivalent to minimizing SPO loss, as the sample size grows large. As mentioned earlier, $\mathbb{E}[\mathbf{z}|\mathbf{x}]$ also minimizes the true risk with LS loss. Thus, the SPO estimators (with either SPO+ loss or the SPO loss) share the same asymptotic validity as the prediction with LS loss by maintaining the same fundamental consistency property.

Beyond asymptotic consistency, El Balghiti et al. [16] provided finite sample performance guarantees in the form of generalization bounds for the SPO loss function. A generalization bound is a high probability upper bound on the true risk $R_{\text{SPO}}(h)$, given the empirical risk $\hat{R}_{\text{SPO}}(h)$ and the sample size n , for any hypothesis h in the hypothesis class \mathcal{H} . The generalization bound is constructed based on the (multivariate) Rademacher complexity of the function class $l_{\text{SPO}} \circ \mathcal{H}$. Intuitively speaking, Rademacher complexity quantifies the richness and representativeness of a class of function. We refer to Shalev-Shwartz and Ben-David [30] and Bartlett and Mendelson [5] for more details of Rademacher complexity and the corresponding generalization bounds. The key to obtain an insightful generalization bound is to analyze the Rademacher complexity of $l_{\text{SPO}} \circ \mathcal{H}$ by exploiting the structure of feasible region S or function class \mathcal{H} . El Balghiti et al. [16] provided generalization bounds for two cases. If the feasible region S is a polyhedron, the generalization bound can be constructed based on the Natarajan dimension (Natarajan [24]). If S is strongly convex, the Rademacher complexity of $l_{\text{SPO}} \circ \mathcal{H}$ can be bounded by the Rademacher complexity of \mathcal{H} after demonstrating the Lipschitz property of the SPO loss. In both cases, the generalization bounds achieve a dependency of $\frac{1}{\sqrt{n}}$ up to logarithmic terms.

Remark 1 (Regret Minimization as a Regularization). The SPO loss and SPO+ loss are formulated in order to minimize the regret caused by a prediction. Recently, Loke et al. [23] proposed the idea of a blended model that targets both prediction accuracy and decision optimality. To be more specific, the authors proposed a decision-driven regularization model that considers the loss function

$$l_{\text{DDR}}(\hat{\mathbf{z}}, \mathbf{z}) := l(\hat{\mathbf{z}}, \mathbf{z}) - \lambda \min_{\mathbf{w} \in S} \{ \mu \mathbf{z}^T \mathbf{w} + (1 - \mu) \hat{\mathbf{z}}^T \mathbf{w} \},$$

where λ and μ are two hyperparameters, and l can be any loss function—for example, l_{LS} . The second term, $\min_{\mathbf{w} \in S} \{ \mu \mathbf{z}^T \mathbf{w} + (1 - \mu) \hat{\mathbf{z}}^T \mathbf{w} \}$, can be interpreted as the lowest possible value of a weighted summation of estimated and empirical cost, among all feasible decisions. It is proved to be a scaled upper bound of the optimality regret and can be viewed as a regularization on the loss function l . In fact, the SPO+ loss $l_{\text{SPO}+}$ is a special case when $\mu = -1$, $\lambda = 1$, and $l(\hat{\mathbf{z}}, \mathbf{z}) := 2\hat{\mathbf{z}}^T \mathbf{w}^*(\mathbf{z})$.

3.2. Learning Distributional Estimation by Reweighting the Empirical Distribution

Besides point predictions, distributional estimations are also widely adopted in two-step methods. In estimate-then-optimize (ETO) methods, the first step focuses on an estimation of the conditional distribution of the random parameters.

When the objective function in the optimization stage is more complicated than a linear function, ETO methods often achieve better results compared with PTO methods. The advantage of distributional estimation is that, if the decision-maker can obtain a perfect estimation of the underlying conditional distribution, then all information that is needed to solve the stochastic optimization problem has been revealed.

A widely adopted nonparametric method is to estimate the conditional distribution by assigning weights to the empirical distribution, and the weights depend on the input feature.

3.2.1. ETO Approach. Bertsimas and Kallus [6] and Ban and Rudin [4] investigated a nonparametric estimation of $\mathbb{P}_{z|\mathbf{x}}$ by adding weights $w_{n,i}(\mathbf{x})$ that depend on the feature \mathbf{x} . The weights can be constructed based on various machine learning models. In the following, we demonstrate several examples proposed by Bertsimas and Kallus [6] that are motivated by different regression methods:

1. *k*-nearest-neighbor (KNN) regression:¹

$$\omega_{n,i}^{\text{KNN}}(\mathbf{x}) = \frac{1}{k} \mathbb{1}[\mathbf{x}_i \text{ is a KNN of } \mathbf{x}].$$

2. Nadaraya–Watson kernel regression:

$$\omega_{n,i}^{\text{KR}}(\mathbf{x}) = \frac{K((\mathbf{x}_i - \mathbf{x})/h_n)}{\sum_{j=1}^n K((\mathbf{x}_j - \mathbf{x})/h_n)},$$

where h_n is a bandwidth parameter that may depend on the sample size n . Examples of the following kernels are $K(\mathbf{x}) = \mathbb{1}[\|\mathbf{x}\| \leq 1]$ (Naive), $K(\mathbf{x}) = (1 - \|\mathbf{x}\|^2)\mathbb{1}[\|\mathbf{x}\| \leq 1]$ (Epanechnikov), and $K(\mathbf{x}) = (1 - \|\mathbf{x}\|^3)^3\mathbb{1}[\|\mathbf{x}\| \leq 1]$ (Tricubic). A slight modification to $\omega_{n,i}^{\text{KR}}(\mathbf{x})$ is to have the bandwidths be selected per data point:

$$\omega_{n,i}^{\text{recursive-KR}}(\mathbf{x}) = \frac{K((\mathbf{x}_i - \mathbf{x})/h_i)}{\sum_{j=1}^n K((\mathbf{x}_j - \mathbf{x})/h_j)},$$

where h_i varies across data points and does not depend on the sample size.

3. Local linear methods:

$$\omega_{n,i}^{\text{LOESS}}(\mathbf{x}) = \frac{\tilde{\omega}_{n,i}(\mathbf{x})}{\sum_{j=1}^n \tilde{\omega}_{n,j}(\mathbf{x})},$$

where $\tilde{\omega}_{n,i}(\mathbf{x})$ is defined as

$$\tilde{\omega}_{n,i}(\mathbf{x}) = k_i(\mathbf{x}) \left(1 - \sum_{j=1}^n k_j(\mathbf{x})(\mathbf{x}_j - \mathbf{x})^T \Xi(\mathbf{x})^{-1}(\mathbf{x}_i - \mathbf{x}) \right),$$

with $\Xi(\mathbf{x}) := \sum_{i=1}^n k_i(\mathbf{x})(\mathbf{x}_i - \mathbf{x})(\mathbf{x}_i - \mathbf{x})^T$ and $k_i(\mathbf{x}) = K((\mathbf{x}_i - \mathbf{x})/h_n(\mathbf{x}))$. Note that the weights $\omega_{n,i}^{\text{LOESS}}(\mathbf{x})$ may be negative. In such scenarios, the interpretation of the reweighted empirical distribution as an estimation of the conditional distribution breaks down. Therefore, a modification of $\omega_{n,i}^{\text{LOESS}}(\mathbf{x})$ to guarantee nonnegative weights can be

$$\omega_{n,i}^{\text{LOESS}^*}(\mathbf{x}) = \frac{\tilde{\omega}_{n,i}(\mathbf{x})}{\sum_{j=1}^n \tilde{\omega}_{n,j}(\mathbf{x})},$$

with $\tilde{\omega}_{n,i}(\mathbf{x})$ being modified as

$$\tilde{\omega}_{n,i}(\mathbf{x}) = k_i(\mathbf{x}) \max \left\{ 1 - \sum_{j=1}^n k_j(\mathbf{x})(\mathbf{x}_j - \mathbf{x})^T \Xi(\mathbf{x})^{-1}(\mathbf{x}_i - \mathbf{x}), 0 \right\}.$$

With weights $\omega_{n,i}(\mathbf{x})$ specified, the decision maker may consider the weighted empirical distribution with samples $\{\mathbf{x}_i, \mathbf{z}_i\}_{i=1}^n$ as an estimation of the conditional distribution $\mathbb{P}_{z|\mathbf{x}}$. Then the corresponding optimal decision is obtained by solving the following optimization problem:

$$\min_{\mathbf{w} \in S} \sum_{i=1}^n [\omega_{n,i}(\mathbf{x}) c(\mathbf{w}, \mathbf{z}_i)]. \quad (11)$$

We let

$$\hat{\mathbf{w}}_n(\mathbf{x}) \in \hat{W}_n(\mathbf{x}) := \arg \min_{\mathbf{w} \in S} \sum_{i=1}^n [\omega_{n,i}(\mathbf{x}) c(\mathbf{w}, \mathbf{z}_i)]$$

denote the optimal solution of (11) and

$$\hat{v}_n(\mathbf{x}) := \min_{\mathbf{w} \in S} \sum_{i=1}^n [\omega_{n,i}(\mathbf{x}) c(\mathbf{w}, \mathbf{z}_i)].$$

It is noteworthy to point out that, as an approximation of (1), (11) can accommodate and keep all the constraints on decision \mathbf{w} , described with the feasible set S . Another important question is the computational tractability of (11). Bertsimas and Kallus [6] demonstrated the following sufficient conditions for (11) to be solvable in polynomial time. We restate the result in the following theorem.

Theorem 2 (Theorem 2 in Bertsimas and Kallus [6]). *Fix \mathbf{x} and weights $\omega_{n,i} \geq 0$. Suppose S is a closed convex set, and let a separation oracle for it be given. Suppose also that $c(\mathbf{w}, \mathbf{z})$ is convex in \mathbf{w} for every fixed \mathbf{z} , and let oracles be given for evaluation and subgradient in \mathbf{w} . Then for any \mathbf{x} , we can find an ϵ -optimal solution to (11) in time and oracle calls polynomial in the effective sample size N_0 , dimension d , and $\log(1/\epsilon)$, where $N_0 = \sum_{i=1}^n \mathbb{1}[w_{n,i}(\mathbf{x}) > 0]$ is the effective sample size.*

Bertsimas and Kallus [6] also demonstrated the asymptotic optimality of this approach—that is, whether the solution of (11) converges to the full-information optimal solution as the sample size grows to infinity. Note that the asymptotic optimal may have various formats. Here, we restate the mathematical definition of the asymptotic optimality in Definition 3 and the associated results in Theorem 3.

Definition 3 (Asymptotic Optimality and Consistency (Bertsimas and Kallus [6])). We say that $\hat{\mathbf{w}}_n(\mathbf{x})$ is *asymptotically optimal* if, with probability 1, we have that for $\mathbb{P}_{\mathbf{x}}$ almost everywhere $\mathbf{x} \in \mathcal{X}$:

$$\lim_{n \rightarrow \infty} \mathbb{E}[c(\hat{\mathbf{w}}_n(\mathbf{x}), \mathbf{z}) | \mathbf{x}] = v^*(\mathbf{x}).$$

We say that $\hat{\mathbf{w}}_n(\mathbf{x})$ is *consistent* if we have that for $\mathbb{P}_{\mathbf{x}}$ almost everywhere $\mathbf{x} \in \mathcal{X}$:

$$\lim_{n \rightarrow \infty} \text{dist}(\hat{\mathbf{w}}_n(\mathbf{x}), W^*(\mathbf{x})) = 0,$$

where $\text{dist}(\mathbf{w}, W)$ denotes the distance from a point $\mathbf{w} \in \mathbb{R}^d$ to a set $W \subseteq \mathbb{R}^d$ and is defined as $\text{dist}(\mathbf{w}, W) := \inf_{\mathbf{u} \in W} \|\mathbf{w} - \mathbf{u}\|$.

Bertsimas and Kallus [6] proved the asymptotic optimality and consistency given the assumptions of existence, continuity, and regularity. For the purpose of completeness, we summarize the assumptions as the following.

Assumption 2. *Suppose the data samples $\{(\mathbf{x}_i, \mathbf{z}_i)\}_{i=1}^n$ are generated by i.i.d. sampling following the joint distribution $\mathbb{P}_{(\mathbf{x}, \mathbf{z})}$; then the following holds:*

- (A) Existence: (1) is well defined: We have $\mathbb{E}[|c(\mathbf{w}, \mathbf{z})|] < \infty$ for every $\mathbf{w} \in S$ and $W^*(\mathbf{x}) \neq \emptyset$.
- (B) Continuity: Cost function $c(\mathbf{w}, \mathbf{z})$ is equicontinuous in w .
- (C) Regularity: Set S is closed and nonempty. In addition, either S is bounded or the following holds: $\liminf_{\|\mathbf{w}\| \rightarrow \infty} \inf_{\mathbf{z} \in \Xi} c(\mathbf{w}, \mathbf{z}) > -\infty$, and for every $\mathbf{x} \in \mathcal{X}$, there exists a subset $D_{\mathbf{x}} \in \Xi$ and $\mathbb{P}(\mathbf{z} \in D_{\mathbf{x}} | \mathbf{X} = \mathbf{x}) > 0$.

We restate the asymptotic properties of the aforementioned examples in the following theorem.

Theorem 3. Suppose parts (A)–(C) of Assumption 2 hold. Then $\hat{w}_n(\mathbf{x})$ is asymptotically optimal and consistent under each of the following conditions:

(1) If the weights $\omega_{n,i}^{\text{KNN}}(\mathbf{x})$ are adopted with $k = \min\{\lceil CN^\delta \rceil, n-1\}$ for some $C > 0$ and $0 < \delta < 1$.

(2) If the weights $\omega_{n,i}^{\text{KR}}(\mathbf{x})$ are adopted and K can be any of the three kernels (Naive, Epanechnikov, and Tricubic). Suppose that $\mathbb{E}[|c(\mathbf{w}, \mathbf{z}) \max\{\log|c(\mathbf{w}, \mathbf{z})|, 0\}|] < \infty$ for each \mathbf{w} and that $h_N = cN^{-\delta}$ for $C > 0$, $0 < \delta < 1/d_x$.

(3) If the weights $\omega_{n,i}^{\text{recursive-KR}}(\mathbf{x})$ are adopted and K is the Naive kernel. Suppose that $h_i = Ci^{-\delta}$ for $C > 0$, $0 < \delta < 1/(2d_x)$.

(4) If the weights $\omega_{n,i}^{\text{LOESS}}(\mathbf{x})$ or $\omega_{n,i}^{\text{LOESS}^*}(\mathbf{x})$ are adopted with $h_N = cN^{-\delta}$ for $C > 0$, $0 < \delta < 1/d_x$ and with K being any of the three kernels (Naive, Epanechnikov, or Tricubic). Suppose \mathbb{P}_x is absolutely continuous and has density bounded away from 0 and infinity on \mathcal{X} and twice continuously differentiable. Moreover, suppose that costs are bounded over \mathbf{z} for each \mathbf{w} and twice continuously differentiable.

3.2.2. Stochastic Optimization Forest. Kallus and Mao [20] investigated the nonparametric approach that estimates conditional distribution by reweighting the empirical distribution based on a forest. Consider a forest $\Gamma = \{\tau_1, \dots, \tau_T\}$ consisting of T trees. The j th tree, $\tau_j: \mathcal{X} \rightarrow \{1, \dots, L_j\}$, forms a partition of $\mathcal{X} \subset \mathbb{R}^p$ into L_j regions. With the forest Γ , the weights are constructed as

$$\omega_{n,i}^{\text{Forest}} = \frac{1}{T} \sum_{j=1}^T \frac{\mathbb{1}[\tau_j(\mathbf{x}_i) = \tau_j(\mathbf{x})]}{\sum_{r=1}^n \mathbb{1}[\tau_j(\mathbf{x}_r) = \tau_j(\mathbf{x})]}.$$

In the ETO methods, the forest Γ can be trained by running the random forest algorithm that focuses on prediction error. That is, each split that divides a region $R_0 \subseteq \mathcal{X}$ into two subregions R_1 and R_2 minimizes the sum of squared distance to the average in each subregion. We denote this split criterion by

$$\hat{C}^{\text{sos}}(R_1, R_2) = \sum_{j=1,2} \min_z \sum_{i: \mathbf{x}_i \in R_j} \|\mathbf{z} - \mathbf{z}_i\|_2^2.$$

In Kallus and Mao [20], the forest is trained by targeting the downstream optimization problem. In contrast to minimizing the sum of squared prediction errors, the authors proposed a new criterion for the recursive partitions that characterizes the optimization goal:

$$C^{\text{oracle}}(R_1, R_2) = \sum_{j=1,2} \min_{\mathbf{w} \in S} \mathbb{E}[c(\mathbf{w}, \mathbf{z}) \mathbb{1}[\mathbf{x} \in R_j]].$$

However, directly implementing this criterion to evaluate every candidate split of each tree in the forest Γ would be too computationally difficult. Instead, Kallus and Mao [20] proposed a second-order perturbation analysis that allows one to solve a single problem corresponding to the root node and extrapolate from there for all candidate splits. This approximation splitting criterion provides a computationally efficient manner for a large-scale forest.

Similar to the original prescriptive methods proposed in Bertsimas and Kallus [6], asymptotic optimality persists in stochastic optimization forests. The authors proved that, under certain conditions, the forest policy attains the optimal risk $v^*(\mathbf{x})$ as the sample sizes go to infinity, pointwisely for any $\mathbf{x} \in \mathcal{X}$ (this convergence result can straightforwardly be extended to be uniform in \mathbf{x} with slightly different conditions). That is, the asymptotic optimality defined in Definition 3 automatically holds for any measure \mathbb{P}_x on \mathcal{X} .

3.3. Learning the Conditional Probability Vector

Grigas et al. [18] investigated the case that the random parameter \mathbf{z} is a discrete random variable with a finite support. That is, there are K possible values of \mathbf{z} (i.e., $\mathcal{Z} := \{\tilde{\mathbf{z}}_1, \dots, \tilde{\mathbf{z}}_K\}$). In this setting, the conditional distribution $\mathbb{P}_{\mathbf{z}|\mathbf{x}}$ can be characterized by a probability vector $p^*(\mathbf{x}) \in \Delta_K := \{\mathbf{p} \in \mathbb{R}^K : \sum_{k=1}^K p_k = 1, \mathbf{p} \geq 0\}$. The k th component of $p^*(\mathbf{x})$, $p_k^*(\mathbf{x})$ represents the probability of $\mathbf{z} = \tilde{\mathbf{z}}_k$ given the value of \mathbf{x} . Therefore, although focusing on distributional estimation, it is still possible to represent the conditional probability estimators using a hypothesis class \mathcal{F} .

Remark 2 (Hypothesis Class of Estimators). Suppose \mathcal{F} is a compact set of functions $f : \mathcal{X} \rightarrow \Delta_K$. The major difference between \mathcal{F} and the hypothesis class used for point predictions \mathcal{H} is that there is a constraint on the output of $f \in \mathcal{F}$ such that $f(\mathbf{x}) \in \Delta_K$ for all $\mathbf{x} \in \mathcal{X}$. One example to accommodate this constraint is to consider the softmax operator $\text{soft} : \mathbb{R}^K \rightarrow \mathbb{R}^K$, defined as $\text{soft}_k(\mathbf{a}) = \frac{\exp(a_k)}{\sum_{j=1}^K \exp(a_j)}$. With the softmax operator, the user can adopt any unconstrained hypothesis class $\tilde{\mathcal{F}}$ consisting of $\tilde{f} : \mathcal{X} \rightarrow \mathbb{R}^K$ and use $\tilde{\mathcal{F}} := \text{soft} \circ \tilde{\mathcal{F}}$. This example demonstrates the flexibility of the choice of hypothesis class.

3.3.1. An ETO Approach for Learning the Conditional Probability Vector. To learn the model \hat{f}_{ETO} in an ETO approach, it is natural to adopt the cross-entropy loss, defined as

$$l_{\text{ce}}(f(\mathbf{x}), \mathbf{z}) := -\sum_{k=1}^K \log(f_k(\mathbf{x})) \mathbb{1}[z = \tilde{\mathbf{z}}_k],$$

where $f_k(\mathbf{x})$ denotes the k th component of $f(\mathbf{x})$. Thus, learning the estimation model requires solving the corresponding ERM problem

$$\min_{f \in \mathcal{F}} \frac{1}{n} \sum_{i=1}^n l_{\text{ce}}(f(\mathbf{x}_i), \mathbf{z}_i).$$

Note that the cross-entropy loss completely ignores the downstream optimization goal. To obtain a decision, the decision maker may plug in the learned estimator \hat{f}_{ETO} and, given feature \mathbf{x} , solve

$$\min_{\mathbf{w} \in S} \sum_{k=1}^K \hat{f}_{\text{ETO},k}(\mathbf{x}) c_k(\mathbf{w}),$$

where $c_k(\mathbf{w}) := c(\mathbf{w}, \tilde{\mathbf{z}}_k)$.

3.3.2. Integrated Conditional Estimation-and-Optimization. In contrast to ETO approaches, Grigas et al. [18] recently proposed an integrated conditional estimation-optimization (ICEO) method that integrates the optimization goal in the estimation process.

The ICEO approach is based on the access to an oracle associated with the downstream optimization problem. Given any input probability vector, the oracle outputs the associated optimal decision. From the assumption that $c(\mathbf{w}, \mathbf{z})$ is convex on \mathbf{w} for any $\mathbf{z} \in \mathcal{Z}$ and that S is a convex and compact feasible region, it is computationally tractable to have an oracle that outputs the optimal solution given any input probability vector $\mathbf{p} \in \Delta_K$. The authors introduced a decision regularization function $\phi(\cdot) : S \rightarrow \mathbb{R}$, where $\phi(\cdot)$ can be any nonnegative and strongly convex function such as the squared l_2 -norm $\|\cdot\|_2^2$.

Then we defined *regularized optimal solution mapping* as

$$w_\rho(\mathbf{p}) := \arg \min_{\mathbf{w} \in S} \mathbb{E}_{\mathbf{z}}[c(\mathbf{w}, \mathbf{z})] + \rho\phi(\mathbf{w}) = \arg \min_{\mathbf{w} \in S} \sum_{k=1}^K p_k c_k(\mathbf{w}) + \rho\phi(\mathbf{w}),$$

where ρ denotes the regularization coefficient. The regularized oracle outputs $w_\rho(\mathbf{p})$ given any feature vector $\mathbf{p} \in \Delta_K$. The strongly convexity of ϕ ensures the Lipschitz property of the regularized optimal solution mapping, which leads to the finite-sample performance guarantee of the ICEO approach. We also slightly abuse the notation of W^* and define

$$W^*(\mathbf{p}) := \arg \min_{\mathbf{w} \in S} \mathbb{E}_z[c(\mathbf{w}, \mathbf{z})] = \arg \min_{\mathbf{w} \in S} \sum_{k=1}^K p_k c_k(\mathbf{w})$$

as the optimal solution(s) of the stochastic optimization problem induced by the input probability vector $\mathbf{p} \in \Delta_K$.

Instead of the cross-entropy loss, the ICEO approach directly targets the objective of the downstream optimization problem. Learning the estimator \hat{f}_{ICEO} is to solve the ERM problem with respect to the regularized in-sample cost induced by the estimator and the regularized oracle:

$$\begin{aligned} \min_{f \in \mathcal{F}; \mathbf{w}_1, \dots, \mathbf{w}_n \in S} \quad & \frac{1}{n} \sum_{i=1}^n c(\mathbf{w}_i, \mathbf{z}_i) + \rho \phi(\mathbf{w}_i) \\ \text{s.t.} \quad & \mathbf{w}_i = w_\rho(f(\mathbf{x}_i)). \end{aligned} \quad (12)$$

Then, the policy associated with \hat{f}_{ICEO} is achieved by solving

$$\min_{\mathbf{w} \in S} \sum_{k=1}^K \hat{f}_{\text{ICEO},k}(\mathbf{x}) c_k(\mathbf{w}), \quad (13)$$

given any feature \mathbf{x} .

3.3.2.1. Nondifferentiability. Unfortunately, solving the training problem (12) is not easy. In general, the formulation (12) is nonconvex and even nondifferentiable. The nondifferentiability of the training objective arises from the nondifferentiability of the optimal decision mapping $w^*(\cdot)$. As a mapping from input probability vector \mathbf{p} to associated optimal decision $w^*(\mathbf{p})$, $w^*(\cdot)$ may have a piecewise constant shape because of the presence of the constraints from the feasible region S .

Note that the existence of poorly shaped nondifferentiability prevents a direct application of gradient-based methods to solve (12). To overcome this issue, the authors proposed a practical approach that approximates the regularized optimal solution mapping $w_\rho(\cdot)$ with smooth functions (this can be achieved via a separate learning approach). The approximated oracle $\tilde{w}_\rho(\cdot)$ can be constructed based on a randomly generated data set $\{(\mathbf{p}_i, \mathbf{w}_i)\}_{i=1}^m$, where $\mathbf{p}_i, i = 1, \dots, m$ are randomly sampled from the simplex Δ_K , and $\mathbf{w}_i = w_\rho(\mathbf{p}_i), i = 1, \dots, m$ are calculated by calling the original oracle. Regression or interpolation methods can be applied to fit the oracle, resulting in different approximation error bounds.

Given the approximated oracle $\tilde{w}_\rho(\cdot)$, (12) becomes

$$\begin{aligned} \min_{f \in \mathcal{F}; \mathbf{w}_1, \dots, \mathbf{w}_n \in S} \quad & \frac{1}{n} \sum_{i=1}^n c(\mathbf{w}_i, \mathbf{z}_i) + \rho \phi(\mathbf{w}_i) \\ \text{s.t.} \quad & \mathbf{w}_i = \tilde{w}_\rho(f(\mathbf{x}_i)). \end{aligned} \quad (14)$$

Note that (14) is differentiable, and any gradient-based algorithm can be directly applied to find \hat{f}_{ICEO} . For a special case where the downstream problem is semialgebraic (Lasserre [21]) and, in the meantime, when the approximated optimal solution mapping is constructed based on polynomial functions, the authors provided a polynomial optimization approach to solve (12).

To analyze the performance of the ICEO approach, we first introduce the following definitions of risk functions. The empirical regularized risk of $f \in \mathcal{F}$ is defined as

$$\hat{R}_n^{\text{ICEO}}(f; \rho) := \frac{1}{n} \sum_{i=1}^n c(w_\rho(f(\mathbf{x}_i)), \mathbf{z}_i) + \rho \phi(w_\rho(f(\mathbf{x}_i))),$$

whereas its unregularized version is

$$\hat{R}_n^{\text{ICEO}}(f) := \frac{1}{n} \sum_{i=1}^n c(w_\rho(f(\mathbf{x}_i)), \mathbf{z}_i).$$

Similarly, the expected true risk is defined as

$$R^{\text{ICEO}}(f) := \mathbb{E}_{\mathbf{x}} \left[\sum_{k=1}^K f_k^*(\mathbf{x}) c_k(w_\rho(f(\mathbf{x}))) \right].$$

Grigas et al. [18] demonstrated a threefold asymptotic consistency, in terms of risks, decisions, and hypotheses, under certain conditions. The asymptotic consistency guarantees the performance of the ICEO approach as the sample size increases by showing that the estimated models and associated solutions converge to the optimal values given full information of the conditional distribution. To obtain the asymptotic consistency of the ICEO approach, model specification and the uniqueness of true hypothesis is required. Herein, we restate assumptions and results from Grigas et al. [18].

Assumption 3 (Grigas et al. [18]). *For the compact hypothesis class \mathcal{H} , we have the following:*

(A) Model specification: *The hypothesis class \mathcal{F} includes the true probability vector f^* .*

(B) Unique true hypothesis: *There does not exist a hypothesis $f \neq f^*$ in \mathcal{F} such that $W(f(\mathbf{x})) \cap W(f^*(\mathbf{x})) \neq \emptyset$, $\mathbb{P}_{\mathbf{x}}$ almost surely for all $\mathbf{x} \in \mathcal{X}$.*

The first part of the assumption is always a must-have to guarantee the consistency of the estimator. The uniqueness assumption in the second part helps to identify the true hypothesis within the hypothesis class.

Note that there is a regularization term on the oracle and the training objective. To guarantee asymptotic consistency, the coefficient ρ should decrease as the sample size n increases and converge to 0 as n grows to infinity. Here, we restate the three levels of asymptotic consistency.

Theorem 4 (Asymptotic Consistency of ICEO (Grigas et al. [18])). *Suppose that the training data $(\mathbf{x}_i, \mathbf{z}_i)$ are an i.i.d. sequence from the distribution \mathcal{D} and that the sequence of regularization parameters ρ_n satisfies $\lim_{n \rightarrow \infty} \rho_n = 0$. Then, under Assumption 1, we have the following:*

(i) *The optimal empirical regularized risk converges to the optimal expected risk; that is,*

$$\min_{f \in \mathcal{F}} \hat{R}_n^{\text{ICEO}}(f, \rho_n) \rightarrow \min_{f \in \mathcal{F}} R^{\text{ICEO}}(f),$$

as the sample size n goes to infinity.

(ii) *With $\mathbb{P}_{\mathbf{x}}$ almost surely for all $\mathbf{x} \in \mathcal{X}$, the sequence of ICEO decisions $w_{\rho_n}(\hat{f}_{\rho_n}^{\text{ICEO}}(\mathbf{x}))$ converges to the true set of optimal decisions $W(f^*(\mathbf{x}))$:*

$$\text{dist}\left(w_{\rho_n}(\hat{f}_{\rho_n}^{\text{ICEO}}(\mathbf{x})), W(f^*(\mathbf{x}))\right) \rightarrow 0$$

with probability 1.

(iii) Additionally, with Assumption 2, the sequence of ICEO hypotheses converges to the true hypothesis:

$$\hat{f}_{\rho_n}^{\text{ICEO}} \rightarrow f^*$$

with probability 1.

It is noteworthy that the asymptotic consistency stated in Theorem 4 is not the same as the asymptotic optimality in Definition 3. In Definition 3, the asymptotic optimality refers to the decisions obtained with an estimator reaching the best objective function with full information. With the continuity of the cost function c , the convergence in terms of decisions stated in Theorem 4 implies the asymptotic optimality defined by Definition 3.

Beyond the asymptotic performances, the authors also provided a finite sample performance of the ICEO estimator in the form of generalization bounds. The generalization bounds are high-probability upper bounds of true risk $R^{\text{ICEO}}(f)$ based on empirical risks $\hat{R}^{\text{ICEO}}(f)$ or $\hat{R}^{\text{ICEO}}(f; \rho)$. Note that $R^{\text{ICEO}}(f)$ is the out-of-sample performance, to the interest of decision makers, and both $\hat{R}^{\text{ICEO}}(f)$ and $\hat{R}^{\text{ICEO}}(f; \rho)$ are accessible via training data. Similar to El Balghiti et al. [16], the generalization bounds are constructed based on (multivariate) Rademacher complexity. The strongly convex regularization on decisions guarantees the Lipschitz property of the regularized optimal solution mapping. Therefore, the multivariate Rademacher complexity can be bounded, which leads to the desired generalization bounds. The generalization results can also be extended to the ICEO solution with the approximated oracle.

3.4. Learning Feature-to-Decision Mapping

So far, in Section 2.2, we introduce two-step approaches and corresponding methods that integrate the optimization goal with the predictions/estimations. Besides the aforementioned stream of research, there is another vein of methods that seeks to learn the optimal feature-to-decision mapping $w^*(\cdot)$.

Bertsimas and Koduri [7] proposed an approach to learn the feature-to-decision mapping based on reproducing kernel Hilbert spaces (RKHS). We restate the definition of RKHS from Bertsimas and Koduri [7].

Definition 4 (RKHS Generated by Kernel (Bertsimas and Koduri [7], p. 457)). Consider a positive definite kernel $K(\cdot, \cdot) : \mathcal{X} \times \mathcal{X} \rightarrow \mathbb{R}$ such that

$$\sum_{i=1}^n \sum_{j=1}^n a_i a_j K(\mathbf{x}_i, \mathbf{x}_j) \geq 0, \quad \forall n \in \mathbb{N}, \mathbf{x}_1, \dots, \mathbf{x}_n \in \mathcal{X}, a_1, \dots, a_n \in \mathbb{R}.$$

Then a reproducing kernel Hilbert space \mathcal{H} generated by K is the closure of the set of functions

$$h : \mathcal{X} \rightarrow \mathbb{R} | h(\mathbf{x}) = \sum_{l=1}^L a_l K(\mathbf{v}_l, \mathbf{x}) \text{ for } \mathbf{v}_1, \dots, \mathbf{v}_L \in \mathcal{X}, L \in \mathbb{N}.$$

The authors considered the case when the feasible region is convex and characterized by a group of convex constraints:

$$g_j(\mathbf{w}) \leq 0, j \in J.$$

To construct $\hat{w}(\mathbf{x}) : \mathcal{X} \rightarrow \mathbb{R}^d$ as an approximation of $w^*(\mathbf{x})$, we minimize the following empirical regularized risk, with an additional penalty term corresponding to the constraints:

$$\min_{\mathbf{w}^{(\cdot)} \in \mathcal{H}, \forall i=1, \dots, d} \frac{1}{n} \sum_{i=1}^n \left[c(w(\mathbf{x}_i), \mathbf{z}_i) + \gamma \sum_{j=1}^J \max\{0, g_j(w(\mathbf{x}_i))\} \right] + \lambda \sum_{l=1}^d \|w^l\|_{\mathcal{H}}^2, \quad (15)$$

where γ is the regularization parameter and λ is the penalty coefficient for infeasibility. The authors provided a reformulation of (15). When the penalty coefficient grows infinity (i.e., $\lambda \rightarrow \infty$), (15) ensures full feasibility of decisions. The authors proved that, under certain conditions, the minimizer of (15), denoted as \hat{w}^{RKHS} , is asymptotically optimal as defined in Definition 3. In addition, the authors also provided a finite-sample performance bound that gives the high probability upper bound of the expected risk with penalty:

$$R^\gamma(\hat{w}^{\text{RKHS}}) := \mathbb{E} \left[c(\hat{w}^{\text{RKHS}}(\mathbf{x}), \mathbf{z}) + \gamma \sum_{j=1}^J \max\{0, g_j(\hat{w}^{\text{RKHS}}(\mathbf{x}_i))\} \right]$$

using the corresponding empirical risk with penalty

$$\hat{R}^\gamma(\hat{w}^{\text{RKHS}}) := \frac{1}{n} \sum_{i=1}^n \left[c(\hat{w}^{\text{RKHS}}(\mathbf{x}_i), \mathbf{z}_i) + \gamma \sum_{j=1}^J \max\{0, g_j(\hat{w}^{\text{RKHS}}(\mathbf{x}_i))\} \right].$$

Note that R^γ and $\hat{R}^\gamma(\hat{w}^{\text{RKHS}})$ are not the risks that concern the decision maker, especially when γ must take a large value to ensure feasibility.

Remark 3 (Special Case: Unconstrained Decisions). When there are no constraints (i.e., $S = \mathbb{R}^d$), it is less challenging to identify the optimal solution mapping because there is no need to address the feasibility of output decisions. The aforementioned RKHS-based approach can be directly applied to the unconstrained case. For the unconstrained case, (15) (with $\gamma = 0$) is numerically tractable, and the asymptotic optimality holds. Moreover, in the unconstrained special case, the finite-sample performance bound when $\gamma = 0$ characterizes the objective concerned by decision makers.

An example of the unconstrained setting is the single-product newsvendor problem stated in Example 1. For the newsvendor setting, there is no effective constraint on decision \mathbf{w} , except for the nonnegativity constraint, which is often trivial. Moreover, (2) shares the same formulation as the contextual quantile regression problem, and finding the optimal decision is equivalent to predicting the $\frac{b}{b+h}$ th quantile of demand. Thus, given any hypothesis class \mathcal{H} , the prediction model $\hat{\mathbf{z}} := h(\mathbf{x})$ can be achieved by solving (5) with the quantile loss function $l_{\frac{b}{b+h}}(\hat{\mathbf{z}}, \mathbf{z}) := \frac{b}{b+h}(\mathbf{w} - \mathbf{z})^+ + \frac{b}{b+h}(\mathbf{z} - \mathbf{w})^+$. There is a handful of literature investigating learning methods for this problem: some representative studies include Ban and Rudin [4], Cao and Shen [10], Lin et al. [22], Oroojlooyjadid et al. [26], and Qi et al. [27].

3.5. Practical End-to-End Learning with Neural Networks

All methods covered so far possess nice theoretical guarantees, such as asymptotic optimality and finite-sample performance guarantees. Beyond the aforementioned methods, there are other existing studies that provide practical, easy-to-implement models, without strict theoretical analysis of performance. These methods are often built with the help of neural networks.

Donti et al. [15] considered the setting where the optimization problem with constraints could be either deterministic or stochastic, which is a more general version of (1):

$$\begin{aligned} \min_{\mathbf{w}} \quad & \mathbb{E}[c(\mathbf{w}, \mathbf{z})|\mathbf{x}] \\ \text{s.t.} \quad & \mathbb{E}[g_i(\mathbf{w}, \mathbf{z}, \mathbf{x})] \leq 0, \quad i = 1, \dots, n_{\text{ineq}}, \\ & h_i(\mathbf{w}) = 0, \quad i = 1, \dots, n_{\text{eq}}, \end{aligned}$$

where the inequality constraints may depend on the random parameters \mathbf{z} and \mathbf{x} but the equality constraints do not. Both objective function and constraints are supposed to be strongly convex in this setting.

By modeling the conditional distribution $\mathbb{P}_{\mathbf{z}|\mathbf{x}}$ using some parametric model $p(\mathbf{z}|\mathbf{x}; \theta)$, the corresponding optimal decision can be viewed as a function of parameter θ given input

features \mathbf{x} (i.e., $w^*(\mathbf{x}; \theta)$). The parameter θ is learned from data by minimizing the optimization objective and penalties of constraint violation:

$$l(\theta) = \mathbb{E}[c(w^*(\mathbf{x}; \theta), \mathbf{z})] + \sum_{i=1}^{n_{\text{ineq}}} \mathbb{I}\{\mathbb{E}[g_i(w^*(\mathbf{x}; \theta), \mathbf{z}, \mathbf{x})] \leq 0\} + \sum_{i=1}^{n_{\text{eq}}} \mathbb{I}\{\mathbb{E}[h_i(w^*(\mathbf{x}; \theta))] = 0\},$$

where the indicator functions lead to infinitely large penalties if any constraint is violated. The main difficulty of learning with the aforementioned loss function is computing the gradient of $l(\theta)$. By the chain rule, $\frac{\partial l}{\partial \theta} = \frac{\partial l}{\partial w^*} \frac{\partial w^*}{\partial \theta}$. Usually, $\frac{\partial l}{\partial w^*}$ is easy to calculate for most choices of the loss function l . However, the optimal solution mapping $w^*(\mathbf{x}; \theta)$ is, in general, nondifferentiable with respect to θ (as elaborated in Section 3.3.2), which makes it challenging to calculate $\frac{\partial w^*}{\partial \theta}$. Donti et al. [15] proposed a method to compute $\frac{\partial w^*}{\partial \theta}$ by solving a group of linear equations based on the Karush-Kuhn-Tucker conditions. Amos and Kolter [3] presented a neural network structure with practical differentiable optimization layers that integrate the optimization problem into traditional fully connected layers. Agrawal et al. [2] continued the study of Amos and Kolter [3] and converted convex programs to the canonical forms and implemented their grammar in CVXPY, a domain-specific, easy-to-implement language for convex optimization embedded in Python (Diamond and Boyd [14]). When the cost function is quadratic, this stream of methods is easy to use and nicely integrated with existing Python libraries for convex optimization.

Cristian et al. [12] focused on the setting where the optimization task has a linear objective (i.e., $c(\mathbf{w}, \mathbf{z}) := \mathbf{z}^T \mathbf{w}$), and the feasible region S is defined by linear constraints $S := \{\mathbf{w} \in \mathbb{R}^d : A\mathbf{w} = \mathbf{b}, \mathbf{w} \geq 0\}$. A parametric prediction model $h_\theta(\mathbf{x})$ (such as a neural network) for the random parameter \mathbf{z} is considered, and the corresponding optimal solution given the input predictions is denoted by $w^*(h_\theta(\mathbf{x}))$. To overcome the nondifferentiability, the authors adopted a similar idea as Grigas et al. [18] and approximated the optimal solution mapping by training a separate neural network $\tilde{w}(\cdot)$ to approximate $w^*(\cdot)$. The prediction model is trained by minimizing the empirical risk, expressed as

$$\min_{\theta} \sum_{i=1}^n \mathbf{z}_i^T w^*(h_\theta(\mathbf{x}_i)). \quad (16)$$

Thus, in each gradient descent step of the proposed ProjectNet model, the direction to move is calculated based on (16), and the updated decision is projected alternatively on $S_1 := \{\mathbf{w} \in \mathbb{R}^d : A\mathbf{w} = \mathbf{b}\}$ and $S_2 := \{\mathbf{w} \in \mathbb{R}^d : \mathbf{w} \geq 0\}$ until it satisfies the feasibility to a desired accuracy. Note that the projections are differentiable and easy to implement because of the linearity of the constraints. The projection of $\mathbf{w} \in \mathbb{R}^d$ on S_1 has a closed-form solution $\mathbf{w} - A^T(AA^T)^{-1}(A\mathbf{w} - \mathbf{b})$, and the projection on S_2 can be implemented easily by a ReLU layer (Agarap [1]).

Zhang et al. [33] proposed an end-to-end framework for portfolio optimization based on deep neural networks. Similar to the aforementioned methods, the loss function used to train the neural network is the objective cost of the portfolio management. The proposed network transforms the input features into fitness scores for each asset, where a higher fitness score means there is a higher probability of an asset to be emphasized (to be selected or assigned higher weights). Then the fitness scores are converted to portfolio weights satisfying constraints that are required in the problem setting. The proposed approach is flexible enough to incorporate various formulations of portfolio management.

4. Applications in Operations Management

In this section, we introduce representative applications of previously reviewed methods in different fields of operations management (OM). We cover three representative areas in OM: supply chain management, portfolio management, and power system operations.

4.1. Supply Chain Management

Many of the previously reviewed methods can be applied to the multiproduct newsvendor problem with budget constraints, as stated in Example 2. Whereas the nonnegativity constraint may often be addressed by rounding the unconstrained solution, the joint capacity constraint causes the nondifferentiability and the difficulty of ensuring feasibility. The ICEO method proposed by Grigas et al. [18] can be directly applied when demand is a discrete random variable. Bertsimas and Kallus [6] and Kallus and Mao [20] also applied their proposed nonparametric prescriptive methods that estimate the conditional distribution to this problem setting for both continuous and discrete demand distribution.

Besides the newsvendor problem, the multiperiod replenishment problem is another popular application in inventory management. In the multiperiod setting, vendor lead time (VLT), defined as the time period from the order placing time to the order arrival time, is another source of uncertainty. In Qi et al. [28], the authors investigated a single-location, single-product, multiperiod replenishment problem with finite horizon $1, \dots, T$, where T denotes the end of the horizon. We let $z_t, \forall t = 1, \dots, T$ denote the sequence of random demands. Suppose the review periods are given as a sequence of dates, and there are in total M orders from period 1 to T that are placed at $t_m, \forall m = 1, \dots, M$. Then, l_m denotes the VLT of the m th order placed at t_m , and the arrival time of this order is denoted by $v_m := t_m + l_m$. The system dynamic can be described as the following: at each period, the system first updates the inventory level, I_t , by checking whether there is any order arrival and demand realization. At the end of each period, either a holding cost or a backorder cost occurs. The unsatisfied demand is fully backordered. Let w_m denote the order quantity of the m th order; then the aforementioned problem can be expressed as

$$\begin{aligned}
 \min_{w_1, \dots, w_M} \quad & \mathbb{E} \left[\sum_{t=1}^T h \left[I_t - z_t + \sum_{m=1}^M w_m \mathbb{1}\{t = t_m + l_m\} \right]^+ + b \left[-I_t + z_t - \sum_{m=1}^M w_m \mathbb{1}\{t = t_m + l_m\} \right]^+ \middle| x \right] \\
 \text{s.t.} \quad & I_{t+1} = I_t - z_t + \sum_{m=1}^M w_m \mathbb{1}\{t = t_m + l_m\}.
 \end{aligned} \tag{17}$$

The multiperiod setting and the two sources of uncertainty (demand and VLT) make fundamental challenges for developing automatic solutions to identify the optimal or close-to-optimal solutions. Instead of integrating the optimization objective into the training loss function, Qi et al. [28] constructed a training data set $\{(\mathbf{x}_i, \tilde{w}_i^*)\}_{i=1}^n$ based on the original data set $\{(\mathbf{x}_i, z_i)\}$. Labels \tilde{w}_i^* in the training set are calculated by solving (17) with a realized trajectory of demand and VLT, denoted as $\mathbf{z}_i := (\mathbf{d}_{[1, \dots, T]}, \mathbf{l}_{[1, \dots, n]})$. The authors demonstrated that the labeling process is computationally efficient for a large training data set. After the labeling process, the authors trained a neural network, noted by $f_\theta(\mathbf{x})$, to minimize the empirical risk

$$\min_{\theta} \sum_{i=1}^n L(f_\theta(\mathbf{x}_i), \tilde{w}_i^*).$$

Qi et al. [28] conducted a field experiment of their proposed method at an online retailer. According to the results, the proposed approach outperforms multiple baseline models, including the original practice implemented at the retailer. The empirical success implies the applicability of this model to real-world inventory management problems.

Another variant of the classical newsvendor problem is the multiwarehouse setting investigated in Cristian et al. [12]. In this setting, there is a network with multiple warehouses, each facing a random local demand. Fulfillment across the warehouses is allowed, and different purchase costs occur if fulfilled from different warehouses. The decision maker needs to decide the amount of inventory that is allocated to the warehouse as well as auxiliary decisions that specify the quantities of demand that are fulfilled from one warehouse to another. Fulfillment

rules can be accommodated as well. In numerical experiments, the authors demonstrated the advantage of their proposed network architecture.

Donti et al. [15] considered a quadratic inventory management problem. The inventory costs are characterized as

$$c_{inv} = c_0 w + c_1 w^2 + b_0(z - w)^+ + b_1((z - w)^+)^2 + h_0(w - z)^+ + h_1((w - z)^+)^2.$$

Although not commonly adopted in practice, the above-mentioned inventory cost can be viewed as a general formulation of inventory cost with quadratic terms on the amount of purchase, overordering quantities, and inventory shortage. The authors showed that the proposed end-to-end model outperforms two-step benchmarks with numerical experiments.

Beyond inventory management, Notz and Pibernik [25] developed a novel approach based on RKHS. This method learns the optimal feature-to-decision mapping for a practical two-stage capacity planning problem. For this particular capacity planning, the authors extended the theoretical results in Bertsimas and Kallus [6] for various kernels.

Chu et al. [11] studied a last-mile delivery problem of online platforms. In the problem setting, there is a service provider for a given area. It can be the central depot that prepares food or parcels. The provider needs to assign the delivery tasks to vehicles and determine the sequence of the tasks. The drivers deliver the items to customers' destinations and then return to the central depot. The main challenge is the uncertainty of the traveling time on each service arc, as a result of the heterogeneity of drivers. The authors formulated the last-mile problem as a capacitated vehicle routing problem and applied the SPO method originally proposed by Elmachtoub and Grigas [17] to solve the last-mile delivery problem.

4.2. Portfolio Optimization

Portfolio optimization is another area of OM that involves decision-making with uncertainty. The goal of portfolio optimization is to find the most profitable assortment of assets facing a random market. Kallus and Mao [20] considered the mean-variance portfolio optimization problem described in Example 3. The authors applied the stochastic optimization forest method on the mean-variance and variance-only (when $\rho = 0$) portfolio optimization problems with linear constraints $\sum_{l=1}^d w_l = 1, w_l \geq 0$ to ensure that the portfolio is in the simplex. The authors implemented the stochastic optimization forest algorithm on these examples and evaluated the numerical performances. Butler and Kwon [9] considered a parametric assumption of the stochastic return

$$\mathbf{z} = P \text{diag}(\mathbf{x})\boldsymbol{\theta} + \boldsymbol{\epsilon},$$

where P is a known binary matrix that specifies the regression design, and $\boldsymbol{\epsilon} \sim N(0, \sigma)$ is the random noise. The authors provided a closed-form solution when there are no constraints or only with the equality constraints $A\mathbf{z} = \mathbf{b}$. With both equality and inequality constraints, the differentiable neural network proposed in Amos and Kolter [3] can be applied, and the authors demonstrated the advantage of integrating prediction and optimization in numerical experiments.

Besides the classical mean-variance portfolio optimization, there are other variants of portfolio optimization with different criteria on return $\mathbf{z}^T \mathbf{w}$ and constraints regarding selling rules and positions. In the following, we summarize some widely adopted objectives.

1. Minimizing the conditional value-at-risk (CVaR):

$$\min_{\mathbf{w}} \mathbb{E} \left[\min_{y \in \mathbb{R}} \mathbb{E} \left[\frac{1}{\alpha} \max\{y - \mathbf{z}^T \mathbf{w}, 0\} - y \right] \right].$$

2. Minimizing the Sharpe ratio:

$$\min_{\mathbf{w}} \frac{\mathbb{E}[\mathbf{z}^T \mathbf{w}]}{\text{Var}(\mathbf{z}^T \mathbf{w})}.$$

In terms of optional constraints, we summarize the following options:

1. *Long selling only*: No short selling allowed (i.e., $\mathbf{w} \in \Delta_d := \{\mathbf{w} \in \mathbb{R}^d \mid \sum_{l=1}^d w_l = 1, \mathbf{w} \geq 0\}$).
2. *Limited maximum position*: An upper bound on weights allocated to a single asset to ensure diversification (i.e., $\|\mathbf{w}\|_\infty \leq u$).
3. *Cardinality constraint*: Maximum number of total number of assets in a portfolio (i.e., $\|\mathbf{w}\|_0 = K$).
4. *Leveraged portfolio*: Allow leverage to increase the investor's exposure to the market (i.e., $\|\mathbf{w}\|_1 = L$ with $L \geq 1$).

Kallus and Mao [20] applied the stochastic optimization forest method to the portfolio optimization with the CVaR criteria and demonstrated the effectiveness of their proposed method compared with the nonparametric prescriptive solution with a regular random forest. Zhang et al. [33] investigated multiple combinations of the aforementioned criteria and constraints. The authors adopted their proposed end-to-end framework. Because of the structure of the aforementioned constraints, it is possible to use differential functions that can be easily implemented in neural network layers to ensure the feasibility of outputs. For example, the softmax function can be adopted to ensure the output decisions satisfy the long-selling-only constraint. The authors showed superior numerical performance of the proposed end-to-end framework on a real-world data set from the Wharton Research Data Services, including hundreds of stocks from the Russell 3000 Index from an almost 40-year range.

4.3. Power System Operations

Power systems have become an important component of modern life and industry. One important problem is the scheduling of the operators in the power system. The key challenge is that the demand of electricity is random. Donti et al. [15] investigated the grid generation scheduling problem stated in Example 4 and assumed that the electricity demand follows a Gaussian distribution. With this assumption, closed-form expressions are available to calculate the gradients while performing gradient-based algorithms. The authors then demonstrated the effectiveness of their proposed end-to-end learning model. Following up on Donti et al. [15], Han et al. [19] studied a similar generation scheduling problem and proposed a task-based neural network that integrates load forecasting and economic generator dispatch.

Wu et al. [32] considered a more practical network-constrained unit commitment (NCUC) problem that also aims to determine unit commitment, generation dispatch, and reserve schedules to minimize the total operation cost. The operation cost includes start-up, shut-down, generation, and load shedding costs; the operator's decision is restricted by complicated physical and network constraints such as ramping limits, generation limits, and minimum on and off time requirements. This problem can be formulated as a mixed-integer linear program. Therefore, the authors applied the SPO framework proposed by Elmachtoub and Grigas [17] for cost-oriented predictions, which enable noticeable economic improvement compared with classical predictions without considering the downstream NCUC problem.

Donti et al. [15] also studied the battery storage problem in the power system, where the operator needs to determine the charge and discharge of a grid-scale battery based on unknown electricity prices in the future. The authors evaluated their proposed method in this setting and demonstrated that learning based on task-based loss outperforms learning based on the traditional root mean square error loss in numerical experiments. Learning with task-based loss provides more reliable solutions with less variability.

Stratigakos et al. [31] applied the nonparametric prescriptive solutions to renewable energy trading problems. The producer needs to decide on an energy offer (an offered price of renewable energy) facing an uncertain day-ahead market. On the basis of certain market balancing, the producer then buys (sells) the amount of energy shortage (surplus). The optimal offer decision w depends on the random status of the market z , and the cost function $c(w, z)$ includes trading cost and prediction error. The authors conducted numerical experiments on different market designs and demonstrated improved performance by adopting the nonparametric prescriptive methods.

5. Conclusion

One of the main difficulties in decision-making systems is the uncertain behavior of some parameters, which are often viewed as random variables following some unknown ground-truth probability distributions. Decision making in various application areas faces the same challenge of uncertainty. For example, supply chain managers face random demand and order arrival times, portfolio managers invest in assets when the market return is uncertain, and power system operators make decisions regarding random electricity demand and market status.

In the era of big data, there are new opportunities to tackle the uncertainty in decision making based on the availability of high-quality data. Besides the increasing availability of data, the presence of contextual information brings the possibility of developing personalized data-driven solutions. Standard data-driven approaches usually consist of two steps. The first step involves using data to build point predictions or distributional estimations for the uncertainty behavior. In the second step, an optimization problem is solved given input predictions/estimations. By contrast, this tutorial examines data-driven solutions that integrate these two steps.

This tutorial reviews recently developed methodologies that integrate prediction/estimation and optimization. In this tutorial, we introduce two-step predict-then-optimize and estimate-then-optimize procedures and describe how to integrate two stages. We also summarize representative examples of applications in multiple areas of operations management, including supply chain management, portfolio optimization, and power system operations. These applications demonstrate the (potential) practical impact of these integrated methods in the real world.

There is plentiful room for future development regarding this topic. From the results presented in this tutorial, we know that integrating prediction/estimation with optimization cannot be achieved by directly applying standard machine learning methods. In fact, it requires new methodologies based on tools and concepts from both machine learning and optimization. With recent development in statistical learning theory, we see great potential in this topic. Another direction for future research is to investigate the applicability and practical impact of integrating prediction/estimation with downstream optimization. Although multiple works have investigated applications of integrated methods, there are still exciting opportunities to investigate when and how integrated methods improve decision making in practice with more realistic data sets.

Acknowledgments

The authors thank Mabel Chou and anonymous reviewers for their comments and suggestions on this tutorial.

Endnote

¹Equidistant ties are broken down either randomly or by a lower-index-first rule.

References

- [1] A. F. Agarap. Deep learning using rectified linear units (ReLU). Preprint, submitted March 22, <https://doi.org/10.48550/arXiv.1803.08375>, 2018.

- [2] A. Agrawal, B. Amos, S. Barratt, S. Boyd, S. Diamond, and Z. J. Kolter. Differentiable convex optimization layers. H. M. Wallach, H. Larochelle, A. Beygelzimer, F. d'Alché-Buc, E. A. Fox, and R. Garnett, eds. *Advances in Neural Information Processing Systems*, Vol. 32. Curran Associates, Red Hook, NY, 9558–9570, 2019.
- [3] B. Amos and Z. J. Kolter. OptNet: Differentiable optimization as a layer in neural networks. D. Precup and Y. W. Teh, eds. *Proceedings of the 34th International Conference on Machine Learning*. PMLR, 136–145, 2017.
- [4] G. Y. Ban and C. Rudin. The big data newsvendor: Practical insights from machine learning. *Operations Research* 67(1):90–108, 2019.
- [5] P. L. Bartlett and S. Mendelson. Rademacher and Gaussian complexities: Risk bounds and structural results. *Journal of Machine Learning Research* 3(November):463–482, 2002.
- [6] D. Bertsimas and N. Kallus. From predictive to prescriptive analytics. *Management Science* 66(3):1025–1044, 2020.
- [7] D. Bertsimas and N. Koduri. Data-driven optimization: A reproducing kernel Hilbert space approach. *Operations Research* 70(1):454–471, 2022.
- [8] J. R. Birge and F. Louveaux. *Introduction to Stochastic Programming*. Springer Science & Business Media, New York, 2011.
- [9] A. Butler and R. Kwon. Integrating prediction in mean-variance portfolio optimization. Preprint, submitted February 18, <https://doi.org/10.48550/arXiv.2102.09287>, 2021.
- [10] Y. Cao and Z. M. Shen. Quantile forecasting and data-driven inventory management under non-stationary demand. *Operations Research Letters* 47(6):465–472, 2019.
- [11] H. Chu, W. Zhang, P. Bai, and Y. Chen. Data-driven optimization for last-mile delivery. *Complex & Intelligent Systems*, ePub ahead of print February 22, <https://doi.org/10.1007/s40747-021-00293-1>, 2021.
- [12] R. Cristian, P. Harsha, G. Perakis, B. L. Quanz, and I. Spantidakis. End-to-end learning via constraint-enforcing approximators for linear programs with applications to supply chains. Presentation, AAAI-22 Workshop on AI for Decision Optimization, February 28, Association for the Advancement of Artificial Intelligence, Menlo Park, CA, 2022.
- [13] G. B. Dantzig. Linear programming under uncertainty. *Management Science* 1(3–4):197–206, 1955.
- [14] S. Diamond and S. Boyd. CVXPY: A Python-embedded modeling language for convex optimization. *Journal of Machine Learning Research* 17(1):2909–2913, 2016.
- [15] P. Donti, B. Amos, and Z. J. Kolter. Task-based end-to-end model learning in stochastic optimization. I. Guyon, U. von Luxburg, S. Bengio, H. M. Wallach, R. Fergus, S. V. N. Vishwanathan, and R. Garnett, eds. *Advances in Neural Information Processing Systems*, Vol. 30. Curran Associates, Red Hook, NY, 5484–5494, 2017.
- [16] O. El Balghiti, A. N. Elmachtoub, P. Grigas, and A. Tewari. Generalization bounds in the predict-then-optimize framework. H. M. Wallach, H. Larochelle, A. Beygelzimer, F. d'Alché-Buc, E. A. Fox, and R. Garnett, eds. *Advances in Neural Information Processing Systems*, Vol. 32. Curran Associates, Red Hook, NY, 14389–14398, 2019.
- [17] A. N. Elmachtoub and P. Grigas. Smart “predict, then optimize.” *Management Science* 68(1):9–26, 2022.
- [18] P. Grigas, M. Qi, and Z. M. Shen. Integrated conditional estimation-optimization. Preprint, submitted October 24, <https://doi.org/10.48550/arXiv.2110.12351>, 2021.
- [19] J. Han, L. Yan, and Z. Li. A task-based day-ahead load forecasting model for stochastic economic dispatch. *IEEE Transactions on Power Systems* 36(6):5294–5304, 2021.
- [20] N. Kallus and X. Mao. Stochastic optimization forests. *Management Sci.*, ePub ahead of print, <https://doi.org/10.1287/mnsc.2022.4458>, 2022.
- [21] J. B. Lasserre. *An Introduction to Polynomial and Semi-Algebraic Optimization*, Vol. 52. Cambridge University Press, Cambridge, UK, 2015.

- [22] S. Lin, Y. Chen, Y. Li, and Z.-J. M. Shen. Data-driven newsvendor problems regularized by a profit risk constraint. *Production and Operations Management* 31(4):1630–1644, 2022.
- [23] G. G. Loke, Q. Tang, and Y. Xiao. Decision-driven regularization: A blended model for predict-then-optimize. Preprint, submitted February 1, <http://dx.doi.org/10.2139/ssrn.3623006>, 2022.
- [24] B. K. Natarajan. On learning sets and functions. *Machine Learning* 4(1):67–97, 1989.
- [25] P. M. Notz and R. Pibernik. Prescriptive analytics for flexible capacity management. *Management Science* 68(3):1756–1775, 2022.
- [26] A. Oroojlooyjadid, L. V. Snyder, and M. Takáč. Applying deep learning to the newsvendor problem. *IIEE Transactions* 52(4):444–463, 2020.
- [27] M. Qi, Z. M. Shen, and Z. Zheng. Learning newsvendor problem with intertemporal dependence and moderate non-stationarities. Preprint, submitted July 10, <http://dx.doi.org/10.2139/ssrn.3648615>, 2020.
- [28] M. Qi, Y. Shi, Y. Qi, C. Ma, R. Yuan, D. Wu, and Z. M. Shen. A practical end-to-end inventory management model with deep learning. Preprint, submitted November 25, <http://dx.doi.org/10.2139/ssrn.3737780>, 2020.
- [29] I. H. Sarker. Machine learning: Algorithms, real-world applications and research directions. *SN Computer Science* 2(3):1–21, 2021.
- [30] S. Shalev-Shwartz and S. Ben-David. *Understanding Machine Learning: From Theory to Algorithms*. Cambridge University Press, New York, 2014.
- [31] A. C. Stratigakos, S. Camal, A. Michiorri, and G. Kariniotakis. Prescriptive trees for integrated forecasting and optimization applied in trading of renewable energy. *IEEE Transactions on Power Systems*, ePub ahead of print February 18, <https://doi.org/10.1109/TPWRS.2022.3152667>, 2022.
- [32] L. Wu, X. Chen, Y. Yang, and Y. Liu. Feature-driven economic improvement for network-constrained unit commitment: A closed-loop predict-and-optimize framework. *IEEE Transactions on Power Systems* 37(4):3104–3118, 2022.
- [33] C. Zhang, Z. Zhang, M. Cucuringu, and S. Zohren. A universal end-to-end approach to portfolio optimization via deep learning. Preprint, submitted November 17, <https://doi.org/10.48550/arXiv.2111.09170>, 2021.