Operations Research Letters 46 (2018) 339–344

Contents lists available at ScienceDirect

Operations Research Letters

journal homepage: www.elsevier.com/locate/orl

Robust inverse optimization ☆

Kimia Ghobadi [a,*], Taewoo Lee [b], Houra Mahmoudzadeh [c], Daria Terekhov [d]

[a] Sloan School of Management, Massachusetts Institute of Technology, Cambridge, MA, USA
[b] Department of Industrial Engineering, University of Houston, Houston, TX, USA
[c] Department of Management Sciences, University of Waterloo, Waterloo, ON, Canada
[d] Department of Mechanical, Industrial and Aerospace Engineering, Concordia University, Montréal, QC, Canada

ARTICLE INFO

Article history:
Received 6 October 2017
Received in revised form 2 March 2018
Accepted 4 March 2018
Available online 28 March 2018

Keywords:
Inverse optimization
Robust optimization
Dietary recommendation

ABSTRACT

Given an observation of a decision-maker's uncertain behavior, we develop a robust inverse optimization model for imputing an objective function that is robust against mis-specifications of the behavior. We characterize the inversely optimized cost vectors for uncertainty sets that may or may not intersect the feasible region, and propose tractable solution methods for special cases. We demonstrate the proposed model in the context of diet recommendation.

1. Introduction

Given an observation as input, the inverse optimization problem determines objective function parameters of an (forward) optimization problem that make the observation an (often approximately) optimal solution for the forward problem. Inverse optimization has been applied in diverse areas, ranging from finance [6] and electricity markets [8] to medical decision-making [12,3]. It has been studied in various optimization frameworks including network and linear [1], combinatorial [15], conic [16], integer and mixed-integer [19,22,18], variational inequality [7], and countably infinite linear [14] problems. Recently, there has been a growing interest in inverse optimization with multiple solutions as input [7,17,21,20,11,2]. While it is unlikely that multiple solutions can be simultaneously optimal, the above studies aim to render the solutions nearly optimal by minimizing some function of residuals with respect to optimality of the solutions.

In this study, instead of considering multiple data points, we consider an uncertainty set that encapsulates all possible realizations of the input data. We adopt this idea from robust optimization, which has been widely used for solving (forward) optimization problems with uncertain parameters [4]. Thus, by bringing robust and inverse optimization together, we propose a *robust inverse* linear optimization model for uncertain input observations. We aim to find a cost vector for the underlying forward problem such that the associated error is minimized for the worst-case realization of the uncertainty in the observed solutions. That is, such a cost vector is *robust* in the sense that it protects against the worst mis-specification of a decision-maker's behavior.

As an example, consider a diet problem. Suppose an individual wants to find diets that follow specific nutritional constraints *and* satisfy his/her preferences. Our model aims to help infer the preferences of the person from past diet patterns, which might be inconsistent or vary over time, in order to generate personalized diets in the future. Assuming the person's behavior can be represented by an uncertainty set, it is important to infer his/her objective function that renders the worst-case behavior within the uncertainty set as close to optimal as possible. Under such an objective function, any non-worst-case diet will thus have a smaller deviation from optimal diets.

To the best of our knowledge, this study is the first to propose a robust inverse optimization framework. Chassein and Goerigk [10] consider robust optimization with variable-size uncertainty sets and determine how large an uncertainty set can be while the nominal solution remains optimal. That work can be seen as *inverse robust* optimization — it infers an uncertainty set that makes a given solution optimal, whereas we propose *robust inverse* optimization. More related to our work is the paper by Esfahani et al. [13], which considers pairs of signal-response data from an ambiguous distribution as input to the inverse problem and determines an objective function from a limited set of candidates that minimizes a prediction error. Our contributions are:

1. We develop an inverse optimization model that finds a cost vector that minimizes the worst-case fit error associated with a realization of an uncertain input data point. We show that solving the robust inverse model is equivalent to solving a finite number of smaller problems, which are tractable

☆ All authors contributed equally to this manuscript.
* Corresponding author.
E-mail address: kimiag@mit.edu (K. Ghobadi).

https://doi.org/10.1016/j.orl.2018.03.007
0167-6377/

in some cases, e.g., when the error function corresponds to p-norm distance where $p \in \{1, 2, \infty\}$ and the uncertainty set is polyhedral or ellipsoidal. Our model generalizes a previous single-observation nominal inverse model as the uncertainty set can be reduced to a singleton.

2. We derive robust inverse optimization models with uncertainty sets that may or may not intersect the feasible region, characterize the corresponding optimal solutions, and propose tractable reformulations for special cases. We do not make any assumptions on the structure of the input data.
3. We demonstrate the proposed methodology in the context of a diet problem and show how incorporating robust optimization into the inverse framework generates a cost vector that appears less sensitive to the uncertain behavior. We compare optimal cost vectors and diet plans from our robust inverse optimization model to those from a classical (non-robust) inverse optimization model.

2. Methodology

2.1. Preliminaries

Let $\mathbf{x} \in \mathbb{R}^n$, $\mathbf{c} \in \mathbb{R}^n$, $\mathbf{A} \in \mathbb{R}^{m\times n}$, and $\mathbf{b} \in \mathbb{R}^m$. We define our *forward* linear optimization (FO) problem as

$$\begin{aligned}\mathbf{FO}(\mathbf{c}):\quad &\underset{\mathbf{x}}{\text{minimize}} && \mathbf{c}'\mathbf{x}\\ &\text{subject to} && \mathbf{A}\mathbf{x} \geq \mathbf{b}.\end{aligned} \tag{1}$$

Let $I = \{1,\dots,m\}$ and $J = \{1,\dots,n\}$ index constraints and variables in (1), respectively, and $\mathbf{a}_i$ denote the i-th row of $\mathbf{A}$. Assume the feasible region of $\mathbf{x}$, denoted by $\mathcal{X}$, is nonempty, full-dimensional, and free of redundant constraints. Let $\mathbf{e}_i$ be the i-th unit vector, $\mathcal{X}^{\text{OPT}}(\mathbf{c})$ be the set of optimal solutions to $\mathbf{FO}(\mathbf{c})$, and $\mathcal{X}^{\text{OPT}} := \cup_{\mathbf{c}\neq\mathbf{0}}\mathcal{X}^{\text{OPT}}(\mathbf{c})$ assumed non-empty.

Given an observed solution $\mathbf{x}^0 \in \mathcal{X}^{\text{OPT}}$, an inverse optimization (IO) model finds a vector $\mathbf{c} \in \mathcal{C} \neq \emptyset$ that makes $\mathbf{x}^0$ optimal for $\mathbf{FO}(\mathbf{c})$. We consider $\mathcal{C} = \{\mathbf{c} \in \mathbb{R}^n \mid \|\mathbf{c}\|_1 = 1\}$, so as to normalize the cost vector and prevent a trivial $\mathbf{0}$ vector from being feasible. The IO formulation is written as

$$\begin{aligned}\mathbf{IO}(\mathbf{x}^0):\quad &\underset{\mathbf{c},\mathbf{y}}{\text{minimize}} && 0\\ &\text{subject to} && \mathbf{A}'\mathbf{y} = \mathbf{c},\\ &&& \mathbf{c}'\mathbf{x}^0 = \mathbf{b}'\mathbf{y},\\ &&& \|\mathbf{c}\|_1 = 1,\\ &&& \mathbf{y} \geq \mathbf{0}.\end{aligned} \tag{2}$$

The first and last constraints of $\mathbf{IO}(\mathbf{x}^0)$ enforce dual feasibility and the second constraint enforces strong duality. Showing that $\mathbf{IO}(\mathbf{x}^0)$ is not feasible for $\mathbf{x}^0 \in \mathcal{X} \setminus \mathcal{X}^{\text{OPT}}$ (e.g., an interior point), Chan et al. [9] propose a generalized inverse formulation, which finds a cost vector $\mathbf{c}$ that minimizes the "fit error" associated with $\mathbf{x}^0$, measured in some norm. Further generalizing their formulation, we write our *nominal* inverse optimization (NIO) model, which finds $\mathbf{c}$ that minimizes a distance function between the corresponding optimal solution $\mathbf{x} \in \mathcal{X}^{\text{OPT}}(\mathbf{c})$ and the given solution $\mathbf{x}^0$, denoted by $d(\mathbf{x}, \mathbf{x}^0)$, as

$$\begin{aligned}\mathbf{NIO}(\mathbf{x}^0):\quad &\underset{\mathbf{c},\mathbf{y},\mathbf{x}}{\text{minimize}} && d(\mathbf{x}, \mathbf{x}^0)\\ &\text{subject to} && \mathbf{A}'\mathbf{y} = \mathbf{c},\\ &&& \mathbf{c}'\mathbf{x} = \mathbf{b}'\mathbf{y},\\ &&& \mathbf{A}\mathbf{x} \geq \mathbf{b},\\ &&& \|\mathbf{c}\|_1 = 1,\\ &&& \mathbf{y} \geq \mathbf{0}.\end{aligned} \tag{3}$$

Note that, unlike Chan et al. [9], we allow $\mathbf{x}^0 \notin \mathcal{X}$, requiring the addition of primal feasibility constraints for $\mathbf{x}$. For the distance function d, we assume the following properties:

(i) $d(\mathbf{x}, \mathbf{x}^0) \geq 0$,
(ii) if $\mathbf{x} = \mathbf{x}^0$ then $d(\mathbf{x}, \mathbf{x}^0) = 0$,
(iii) $d(\mathbf{x}, \mathbf{x}^0) = d(\mathbf{x}^0, \mathbf{x})$,
(iv) $d(\mathbf{x}, \mathbf{x}^0) \leq d(\mathbf{x}, \tilde{\mathbf{x}}) + d(\tilde{\mathbf{x}}, \mathbf{x}^0)$,
(v) $d(\mathbf{x}, \mathbf{x}^0) = d(\mathbf{x} + \boldsymbol{\gamma}, \mathbf{x}^0 + \boldsymbol{\gamma})$, and
(vi) $|\lambda| d(\mathbf{x}, \mathbf{x}^0) = d(\lambda\mathbf{x}, \lambda\mathbf{x}^0)$.

These properties are general and apply to most distance functions (including general p-norm).

2.2. Robust inverse formulation

Consider a bounded, non-empty uncertainty set $\mathcal{U} \subset \mathbb{R}^n$. Our goal is to find a cost vector $\mathbf{c}$ that is "robust" against $\mathcal{U}$, i.e., minimizes the worst-case fit error over $\mathcal{U}$. We first formulate a general robust inverse optimization problem, and then consider three different cases for the uncertainty set: (a) $\mathcal{U}$ in the complement of $\mathcal{X}$ (denoted by $\bar{\mathcal{X}}$), (b) $\mathcal{U}$ in $\mathcal{X}$, and (c) $\mathcal{U} \cap \mathcal{X} \neq \emptyset$ and $\mathcal{U} \cap \bar{\mathcal{X}} \neq \emptyset$, for which we propose solution approaches under a certain choice of distance function. Formally, we say a cost vector $\mathbf{c}$ is *robust* if $\mathbf{c} \in \arg\min_{\mathbf{c}\in\mathcal{C}} \min_{\mathbf{x}\in\mathcal{X}^{\text{OPT}}(\mathbf{c})} \max_{\hat{\mathbf{x}}\in\mathcal{U}} d(\mathbf{x}, \hat{\mathbf{x}})$. We propose the robust inverse optimization (RIO) model:

$$\begin{aligned}\mathbf{RIO}(\mathcal{U}):\quad &\underset{\mathbf{c},\mathbf{y},\mathbf{x}}{\text{minimize}} && \max_{\hat{\mathbf{x}}\in\mathcal{U}} d(\mathbf{x}, \hat{\mathbf{x}})\\ &\text{subject to} && \mathbf{A}'\mathbf{y} = \mathbf{c},\\ &&& \mathbf{c}'\mathbf{x} = \mathbf{b}'\mathbf{y},\\ &&& \mathbf{A}\mathbf{x} \geq \mathbf{b},\\ &&& \|\mathbf{c}\|_1 = 1,\\ &&& \mathbf{y} \geq \mathbf{0}.\end{aligned} \tag{4}$$

Proposition 1. $\mathbf{RIO}(\mathcal{U})$ *has an optimal solution for any bounded non-empty* $\mathcal{U}$.

Proof. We first show that a feasible solution can always be constructed. Substituting $\mathbf{A}'\mathbf{y}$ for $\mathbf{c}$ using the first constraint, we rewrite constraints of problem (4) as

$$\begin{aligned}&\mathbf{y}'(\mathbf{A}\mathbf{x}) = \mathbf{b}'\mathbf{y},\\ &\mathbf{A}\mathbf{x} \geq \mathbf{b},\\ &\|\mathbf{A}'\mathbf{y}\|_1 = 1,\\ &\mathbf{y} \geq \mathbf{0}.\end{aligned} \tag{5}$$

We choose $\mathbf{x} = \tilde{\mathbf{x}}$ such that $\mathbf{a}_{\tilde{i}}'\tilde{\mathbf{x}} = b_{\tilde{i}}$ for some $\tilde{i}$ and $\mathbf{a}_i'\tilde{\mathbf{x}} > b_i$ for all $i \in I, i \neq \tilde{i}$. Then there exists $\mathbf{y} = \lambda\mathbf{e}_{\tilde{i}}$ for $\lambda > 0$ such that $\|\lambda\mathbf{a}_{\tilde{i}}\|_1 = 1$, i.e., $\lambda = 1/\|\mathbf{a}_{\tilde{i}}\|_1$, implying that the pair $\mathbf{x} = \tilde{\mathbf{x}}$ and $\mathbf{y} = \mathbf{e}_{\tilde{i}}/\mathbf{a}_{\tilde{i}}$ is feasible for (5). Thus $(\mathbf{c}, \mathbf{y}, \mathbf{x}) = (\lambda\mathbf{a}_{\tilde{i}},\ \lambda\mathbf{e}_{\tilde{i}},\ \tilde{\mathbf{x}})$ is feasible for problem (4). Since $\mathcal{U}$ is non-empty and bounded, and since any $\mathbf{x}$ feasible to (4) is in $\mathcal{X}^{\text{OPT}}$, $d(\mathbf{x}, \hat{\mathbf{x}})$ is finite for each $\hat{\mathbf{x}} \in \mathcal{U}$ and thus the optimal cost is bounded. □

Proposition 2. *A cost vector* $\mathbf{c}^*$ *optimal for* $\mathbf{RIO}(\mathcal{U})$ *is robust against* $\mathcal{U}$.

Proof. Let $(\mathbf{c}^*, \mathbf{y}^*, \mathbf{x}^*)$ be an optimal solution to $\mathbf{RIO}(\mathcal{U})$. Since the constraints of $\mathbf{RIO}(\mathcal{U})$ enforce primal and dual feasibility as well as strong duality for the given $(\mathbf{c}^*, \mathbf{y}^*, \mathbf{x}^*)$, it must be that $\mathbf{x}^* \in \mathcal{X}^{\text{OPT}}(\mathbf{c}^*)$, implying that $\mathbf{RIO}(\mathcal{U})$ is equivalent to

$\min_{\mathbf{c}\in\mathcal{C},\mathbf{x}\in\mathcal{X}^{OPT}(\mathbf{c})} \max_{\hat{\mathbf{x}}\in\mathcal{U}} d(\mathbf{x}, \hat{\mathbf{x}})$.

Thus,

$\mathbf{c}^* \in \arg\min_{\mathbf{c}\in\mathcal{C}} \min_{\mathbf{x}\in\mathcal{X}^{OPT}(\mathbf{c})} \max_{\hat{\mathbf{x}}\in\mathcal{U}} d(\mathbf{x}, \hat{\mathbf{x}})$,

which is the definition of a robust cost vector given earlier. □

Note that $\mathbf{RIO}(\mathcal{U})$ is non-convex due to the second constraint in (4). $\mathbf{RIO}(\mathcal{U})$ becomes a bilinear program when the distance function d can be linearized (e.g., p-norm distance where $p = 1, \infty$), $\mathbf{c} \geq \mathbf{0}$ and the substitution $\mathbf{c} = \mathbf{A}'\mathbf{y}$ is made as in (5), for which general bilinear programming algorithms may be used.

Next, we present theoretical results that lead to the development of solution approaches for a general distance function satisfying the above-stated properties (i)–(vi).

Proposition 3. *A solution* $(\mathbf{c}^*, \mathbf{y}^*, \mathbf{x}^*)$ *is optimal for* $\mathbf{RIO}(\mathcal{U})$ *if and only if* $\mathbf{x}^*$ *is optimal for*

$$\underset{\mathbf{x}\in\mathcal{X}^{OPT}}{\text{minimize}} \max_{\hat{\mathbf{x}}\in\mathcal{U}} \; d(\mathbf{x}, \hat{\mathbf{x}}), \tag{6}$$

and $(\mathbf{c}^*, \mathbf{y}^*)$ *is optimal to* $\mathbf{IO}(\mathbf{x}^*)$*.*

Proof. ($\Leftarrow$) Since $\mathbf{x}^* \in \mathcal{X}^{\text{OPT}}$ and $(\mathbf{c}^*, \mathbf{y}^*)$ is optimal for $\mathbf{IO}(\mathbf{x}^*)$, $(\mathbf{c}^*, \mathbf{y}^*, \mathbf{x}^*)$ is feasible for $\mathbf{RIO}(\mathcal{U})$. Since $\mathcal{X}^{\text{OPT}} = \cup_{\mathbf{c}\neq\mathbf{0}}\mathcal{X}^{\text{OPT}}(\mathbf{c}) = \cup_{\mathbf{c}\neq\mathbf{0}}\{\mathbf{x} \mid \mathbf{A}\mathbf{x} \geq \mathbf{b}, \mathbf{c}'\mathbf{x} = \mathbf{b}'\mathbf{y}, \mathbf{A}'\mathbf{y} = \mathbf{c} \text{ for some } \mathbf{y} \geq \mathbf{0}\} = \cup_{\mathbf{c}\in\mathcal{C}}\{\mathbf{x} \mid \mathbf{A}\mathbf{x} \geq \mathbf{b}, \mathbf{c}'\mathbf{x} = \mathbf{b}'\mathbf{y}, \mathbf{A}'\mathbf{y} = \mathbf{c} \text{ for some } \mathbf{y} \geq \mathbf{0}\}$ (recall $\mathcal{C} = \{\mathbf{c} \mid \|\mathbf{c}\|_1 = 1\}$), i.e., the choice of $\mathbf{x}^*$ in (6) is not restricted by $(\mathbf{c}, \mathbf{y})$, (6) provides a lower bound on the objective value of $\mathbf{RIO}(\mathcal{U})$. Substituting $(\mathbf{c}^*, \mathbf{y}^*, \mathbf{x}^*)$ into $\mathbf{RIO}(\mathcal{U})$, we achieve this lower bound, implying that $(\mathbf{c}^*, \mathbf{y}^*, \mathbf{x}^*)$ is optimal for $\mathbf{RIO}(\mathcal{U})$. ($\Rightarrow$) Let $(\mathbf{c}^*, \mathbf{y}^*, \mathbf{x}^*)$ be an optimal solution to $\mathbf{RIO}(\mathcal{U})$. Since the constraints of $\mathbf{RIO}(\mathcal{U})$ enforce primal and dual feasibility as well as strong duality, $\mathbf{x}^* \in \mathcal{X}^{\text{OPT}}(\mathbf{c}^*)$. Suppose that there exists $\tilde{\mathbf{x}} \in \mathcal{X}^{OPT}$ such that $\tilde{\mathbf{x}} \neq \mathbf{x}^*$ and $\max_{\hat{\mathbf{x}}\in\mathcal{U}} \; d(\tilde{\mathbf{x}}, \hat{\mathbf{x}}) < \max_{\hat{\mathbf{x}}\in\mathcal{U}} \; d(\mathbf{x}^*, \hat{\mathbf{x}})$. Since $\tilde{\mathbf{x}} \in \mathcal{X}^{OPT}$, there must exist a $\tilde{\mathbf{c}}$ such that $\tilde{\mathbf{x}} \in \mathcal{X}^{OPT}(\tilde{\mathbf{c}})$, which implies that $\tilde{\mathbf{x}} \in \mathcal{X}^{OPT}(\tilde{\mathbf{c}}/\|\tilde{\mathbf{c}}\|_1)$. Thus, $(\tilde{\mathbf{c}}/\|\tilde{\mathbf{c}}\|_1, \tilde{\mathbf{y}}/\|\tilde{\mathbf{c}}\|_1, \tilde{\mathbf{x}})$ where $\mathbf{A}'\tilde{\mathbf{y}} = \tilde{\mathbf{c}}$ and $\tilde{\mathbf{y}} \geq \mathbf{0}$ is a feasible solution to $\mathbf{RIO}(\mathcal{U})$. As a result $\mathbf{RIO}(\mathcal{U})$ now has the objective function value $\max_{\hat{\mathbf{x}}\in\mathcal{U}} d(\tilde{\mathbf{x}}, \hat{\mathbf{x}}) < \max_{\hat{\mathbf{x}}\in\mathcal{U}} d(\mathbf{x}^*, \hat{\mathbf{x}})$, which is a contradiction. Thus, $\mathbf{x}^* \in \arg\min_{\mathbf{x}\in\mathcal{X}^{\text{OPT}}} \max_{\hat{\mathbf{x}}\in\mathcal{U}} \; d(\mathbf{x}, \hat{\mathbf{x}})$ and $(\mathbf{c}^*, \mathbf{y}^*)$ is an optimal solution to $\mathbf{IO}(\mathbf{x}^*)$, as desired. $\square$

Proposition 3 elucidates the fact that once an $\mathbf{x}^* \in \mathcal{X}^{\text{OPT}}$ is identified that minimizes the maximum distance to $\mathcal{U}$, then $(\mathbf{c}, \mathbf{y}, \mathbf{x}^*)$ will be optimal to $\mathbf{RIO}(\mathcal{U})$ for any $(\mathbf{c}, \mathbf{y})$ that satisfy dual feasibility and strong duality with respect to $\mathbf{x}^*$. Importantly, Proposition 3 shows that formulation (4) can be decomposed into two problems: the upper-level optimization problem that finds $\mathbf{x}^*$ that minimizes the maximum distance to $\hat{\mathbf{x}} \in \mathcal{U}$, and the lower-level satisfaction problem where the corresponding $\mathbf{c}^*$ such that $\mathbf{x}^* \in \mathbf{FO}(\mathbf{c}^*)$ is determined. The next result exploits this structure and characterizes an optimal solution to $\mathbf{RIO}(\mathcal{U})$.

Corollary 1. *An optimal solution to* $\mathbf{RIO}(\mathcal{U})$ *is* $(\mathbf{c}^*, \mathbf{y}^*, \mathbf{x}^*) = (\mathbf{a}_i/\|\mathbf{a}_i\|_1, \mathbf{e}_i/\|\mathbf{a}_i\|_1, \mathbf{x}^*)$ *where* $\mathbf{x}^* = \arg\min_{\{\mathbf{x}|\mathbf{a}_i'\mathbf{x}=b_i, \mathbf{A}\mathbf{x}\geq\mathbf{b}\}} \max_{\hat{\mathbf{x}}\in\mathcal{U}} d(\mathbf{x}, \hat{\mathbf{x}})$ *for some* $i \in I$*.*

Proof. From Proposition 3, $\mathbf{c}^*$ is optimal for $\mathbf{RIO}(\mathcal{U})$ if and only if it is optimal for $\mathbf{IO}(\mathbf{x}^*)$ where $\mathbf{x}^* \in \arg\min_{\mathbf{x}\in\mathcal{X}^{\text{OPT}}} \max_{\hat{\mathbf{x}}\in\mathcal{U}} \; d(\mathbf{x}, \hat{\mathbf{x}})$. Let $\mathcal{I}(\mathbf{x}^*) = \{i \mid \mathbf{a}_i'\mathbf{x}^* = b_i\}$. Then $\mathbf{c}^* \in \mathcal{C}^* = \{\sum_{i\in\mathcal{I}(\mathbf{x}^*)}\mathbf{a}_i y_i \mid y_i \geq 0, \forall i \in \mathcal{I}(\mathbf{x}^*), \|\sum_{i\in\mathcal{I}(\mathbf{x}^*)}\mathbf{a}_i y_i\|_1 = 1\}$. By setting $y_{\tilde{i}} = 1/\|\mathbf{a}_{\tilde{i}}\|_1$ for some $\tilde{i} \in \mathcal{I}(\mathbf{x}^*)$ and $y_i = 0$ for $i \neq \tilde{i}$, we construct $\mathbf{c}^* = \mathbf{a}_{\tilde{i}}/\|\mathbf{a}_{\tilde{i}}\|_1 \in \mathcal{C}^*$, with the corresponding choice of $\mathbf{y}$ being $\mathbf{y}^*$, i.e., $\mathbf{y}^* = \mathbf{e}_{\tilde{i}}/\|\mathbf{a}_{\tilde{i}}\|_1$ as desired. $\square$

The above result implies that the search for an optimal $\mathbf{c}$ is reduced to searching over a set of finite alternatives, each corresponding to one of the constraints. Moreover, $\mathbf{x}^*$ is a point on some hyperplane $\mathbf{a}_i'\mathbf{x} = b_i$ that is a perturbation of the *best* worst-case point in $\mathcal{U}$, i.e., $\mathbf{x}^* = \mathbf{x}^w - \boldsymbol{\epsilon}^*$ where $\mathbf{x}^w \in \arg\max_{\hat{\mathbf{x}}\in\mathcal{U}} \; d(\mathbf{x}^*, \hat{\mathbf{x}})$ and $\boldsymbol{\epsilon}^*$ is an optimal perturbation vector. In the special case where $\mathbf{x}^w$ is known, e.g., $\mathcal{U} = \{\mathbf{x}^0\}$, $\mathbf{RIO}(\mathcal{U})$ reduces to $\mathbf{NIO}(\mathbf{x}^w)$, i.e., the single-point problem. For such a problem with a feasible single observation, Chan et al. [9] show the same structure for $\mathbf{c}^*$ and $\mathbf{y}^*$, and a closed-form expression for $\boldsymbol{\epsilon}^*$ (and hence $\mathbf{x}^*$) under an arbitrary norm objective. For a general $\mathcal{U}$ and a general distance function, a closed-form expression for $\mathbf{x}^*$ is unlikely to exist. However, below we leverage the solution structure described in Corollary 1 and show that $\mathbf{RIO}(\boldsymbol{\mathcal{U}})$ can be decomposed into m sub-problems.

Theorem 1. *For each* $i \in I$*, let* $(\tilde{\mathbf{x}}_i, \tilde{z}_i)$ *be an optimal solution to*

$$\begin{aligned} \underset{\mathbf{x}, z_i}{\text{minimize}} \quad & z_i \\ \textit{subject to} \quad & z_i \geq d(\mathbf{x}, \hat{\mathbf{x}}), \quad \hat{\mathbf{x}} \in \mathcal{U}, \\ & \mathbf{a}_i'\mathbf{x} = b_i, \\ & \mathbf{A}\mathbf{x} \geq \mathbf{b}. \end{aligned} \tag{7}$$

Then an optimal solution to $\mathbf{RIO}(\mathcal{U})$ *is* $(\mathbf{c}^*, \mathbf{y}^*, \mathbf{x}^*) = (\mathbf{a}_{i^*}/\|\mathbf{a}_{i^*}\|_1, \mathbf{e}_{i^*}/\|\mathbf{a}_{i^*}\|_1, \tilde{\mathbf{x}}_{i^*})$ *where* $i^* \in \arg\min_{i\in I} \tilde{z}_i$*.*

Proof. For each $i \in I$, $\tilde{z}_i = \text{minimize}_{\{\mathbf{x}|\mathbf{a}_i'\mathbf{x}=b_i, \mathbf{A}\mathbf{x}\geq\mathbf{b}\}} \max_{\hat{\mathbf{x}}\in\mathcal{U}} \; d(\mathbf{x}, \hat{\mathbf{x}})$, and $\tilde{\mathbf{x}}_i \in \arg\min_{\{\mathbf{x}|\mathbf{a}_i'\mathbf{x}=b_i, \mathbf{A}\mathbf{x}\geq\mathbf{b}\}} \max_{\hat{\mathbf{x}}\in\mathcal{U}} \; d(\mathbf{x}, \hat{\mathbf{x}})$. Thus, $(\mathbf{c}^*, \mathbf{y}^*, \mathbf{x}^*) = (\mathbf{a}_{i^*}/\|\mathbf{a}_{i^*}\|_1, \mathbf{e}_{i^*}/\|\mathbf{a}_{i^*}\|_1, \tilde{\mathbf{x}}_{i^*})$ where $i^* \in \arg\min_{i\in I} \; \tilde{z}_i$ satisfies the structure identified in Corollary 1. It remains to show that the optimal value of $\mathbf{RIO}(\mathcal{U})$ is achieved for $i^* \in \arg\min_{i\in I} \; \tilde{z}_i$. To see that this property is true, note that

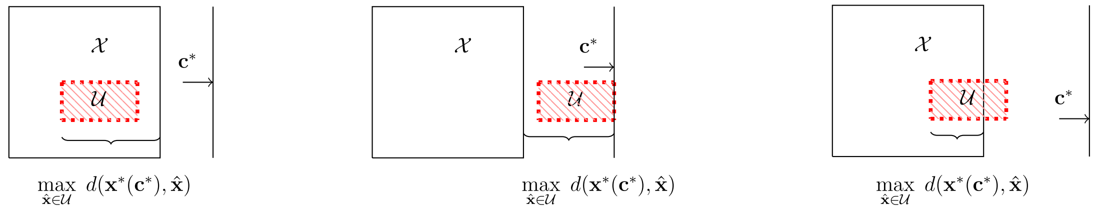
$$\tilde{\mathbf{x}}_{i^*} \in \underset{\bigcup_{i\in I}\{\mathbf{x}|\mathbf{a}_i'\mathbf{x}=b_i, \mathbf{A}\mathbf{x}\geq\mathbf{b}\}}{\arg\min} \left\{ \max_{\hat{\mathbf{x}}\in\mathcal{U}} d(\mathbf{x}, \hat{\mathbf{x}}) \right\}, \tag{8}$$

and $\bigcup_{i\in I}\{\mathbf{x} \mid \mathbf{a}_i'\mathbf{x} = b_i, \mathbf{A}\mathbf{x} \geq \mathbf{b}\}$ in (8) is equivalent to $\mathcal{X}^{OPT}$, implying that (8) is identical to (6). Thus, $(\mathbf{a}_{i^*}/\|\mathbf{a}_{i^*}\|_1, \mathbf{e}_{i^*}/\|\mathbf{a}_{i^*}\|_1, \tilde{\mathbf{x}}_{i^*})$, where $i^* \in \arg\min_{i\in I} \tilde{z}_i$, is optimal to $\mathbf{RIO}(\mathcal{U})$. $\square$

Remark 1. Formulation (7) for each $i \in I$ is a robust optimization problem with uncertainty in constraint parameters. Therefore, Theorem 1 suggests that $\mathbf{RIO}(\mathcal{U})$ can be solved by decomposing it into m robust optimization problems, one for each $i \in I$, and finding the minimum worst-case distance over all $i \in I$.

Note that there may exist multiple i^*'s that lead to the same minimum worst-case distance. In this case, one may use a secondary objective function to break the ties, e.g., choose i^* that corresponds to a cost vector closest to a target vector, as proposed by Ahuja and Orlin [1]. The complexity of (7) depends on the function $d(\mathbf{x}, \hat{\mathbf{x}})$ and the structure of $\mathcal{U}$. In general, if d can be rewritten as a convex function in (7), then a tractable reformulation of model (7) can be derived for some well-studied uncertainty sets [5]. For instance, for $d(\mathbf{x}, \hat{\mathbf{x}}) = \|\mathbf{x} - \hat{\mathbf{x}}\|_p$ where $p \in \{1, \infty\}$, which can be linearized, (7) can be reformulated as a linear or conic program if $\mathcal{U}$ is a polyhedron or an ellipsoid, respectively. As an example, a linear reformulation of (7) when $d(\mathbf{x}, \hat{\mathbf{x}}) = \|\mathbf{x} - \hat{\mathbf{x}}\|_\infty$ and $\mathcal{U}$ is a polyhedron is provided in the Appendix.

In the following subsections, we consider three types of uncertainty sets in geometric relation to $\mathcal{X}$: (a) $\mathcal{U} \subset \bar{\mathcal{X}}$, (b) $\mathcal{U} \subseteq \mathcal{X}$, and (c) $\mathcal{U} \cap \mathcal{X} \neq \emptyset$ and $\mathcal{U} \cap \bar{\mathcal{X}} \neq \emptyset$. Fig. 1 illustrates the geometrical intuition for the three cases when the 2-norm distance function is used (i.e., $d(\mathbf{x}, \hat{\mathbf{x}}) = \|\mathbf{x} - \hat{\mathbf{x}}\|_2$). In general, solving $\mathbf{RIO}(\mathcal{U})$ corresponds to finding and projecting $\mathbf{x}^w \in \mathcal{U}$ onto $\mathcal{X}^{\text{OPT}}$. When $\mathcal{U} \subset \bar{\mathcal{X}}$, doing so is equivalent to projecting $\mathbf{x}^w$ onto a convex set $\mathcal{X}$ whereas when $\mathcal{U} \subseteq \bar{\mathcal{X}}$ the projection is onto the non-convex $\bar{\mathcal{X}}$. When $\mathcal{U} \cap \mathcal{X} \neq \emptyset$, and $\mathcal{U} \cap \bar{\mathcal{X}} \neq \emptyset$, the problem requires partitioning $\mathcal{U}$ into $\mathcal{U} \cap \mathcal{X}$ and $\mathcal{U} \cap \bar{\mathcal{X}}$.



(a) $\mathcal{U} \subset \mathcal{X}$. (b) $\mathcal{U} \subset \bar{\mathcal{X}}$. (c) $\mathcal{U} \cap \mathcal{X} \neq \emptyset$ and $\mathcal{U} \cap \bar{\mathcal{X}} \neq \emptyset$.

Fig. 1. Illustration of the optimal cost vector $\mathbf{c}^*$ when $d(\mathbf{x}, \hat{\mathbf{x}}) = \|\mathbf{x} - \hat{\mathbf{x}}\|_2$.

2.2.1. $\mathcal{U} \subset \bar{\mathcal{X}}$

First, consider the uncertainty set $\mathcal{U}$ that is outside the feasible region of the forward problem. In this case, solving $\mathbf{RIO}(\mathcal{U})$ is easier than in the case of a general $\mathcal{U}$.

Proposition 4. *Let $\mathcal{U} \subset \bar{\mathcal{X}}$. If $\mathbf{x}^*$ is optimal for*

$$\underset{\mathbf{x}\in\mathcal{X}}{\text{minimize}} \max_{\hat{\mathbf{x}}\in\mathcal{U}} \quad d(\mathbf{x}, \hat{\mathbf{x}}), \tag{9}$$

then there exist $\mathbf{c}^* \in \mathbb{R}^n$ and $\mathbf{y}^* \in \mathbb{R}^m$ such that $(\mathbf{c}^*, \mathbf{y}^*, \mathbf{x}^*)$ is optimal for $\mathbf{RIO}(\mathcal{U})$.

Proof. Let $\mathbf{x}^*$ be an optimal solution for (9). All we need to show is that $\mathbf{x}^*$ optimal for (9) is feasible for (6) because the feasible region of (6) is a subset of that of (9). We use contradiction. Suppose that $\mathbf{x}^* \in \mathcal{X} \setminus \mathcal{X}^{OPT}$, i.e., $\max_{\hat{\mathbf{x}}\in\mathcal{U}} d(\mathbf{x}^*, \hat{\mathbf{x}}) < \max_{\hat{\mathbf{x}}\in\mathcal{U}} d(\mathbf{x}, \hat{\mathbf{x}})$ for all $\mathbf{x} \in \mathcal{X}^{OPT}$. For every $\hat{\mathbf{x}} \in \mathcal{U}$, since it is outside the feasible region $\mathcal{X}$ and $\mathcal{X}$ is convex, there must exist a point in $\mathcal{X}^{OPT}$, say $\tilde{\mathbf{x}} \in \mathcal{X}^{OPT}$, such that $\tilde{\mathbf{x}} = \lambda\mathbf{x}^* + (1-\lambda)\hat{\mathbf{x}}$, $\lambda \in (0, 1)$. Then we have

$$\begin{aligned} d(\mathbf{x}^*, \hat{\mathbf{x}}) - d(\tilde{\mathbf{x}}, \hat{\mathbf{x}}) &= d(\mathbf{x}^*, \hat{\mathbf{x}}) - d(\lambda\mathbf{x}^* + (1-\lambda)\hat{\mathbf{x}}, \hat{\mathbf{x}}) \\ &= d(\mathbf{x}^*, \hat{\mathbf{x}}) - d(\lambda\mathbf{x}^*, \hat{\mathbf{x}} - (1-\lambda)\hat{\mathbf{x}}) \\ &= d(\mathbf{x}^*, \hat{\mathbf{x}}) - d(\lambda\mathbf{x}^*, \lambda\hat{\mathbf{x}}) \\ &= d(\mathbf{x}^*, \hat{\mathbf{x}}) - |\lambda| d(\mathbf{x}^*, \hat{\mathbf{x}}) > 0, \end{aligned} \tag{10}$$

where the second and fourth equalities are due to distance function properties (v) and (vi), respectively. From (10), $\max_{\hat{\mathbf{x}}\in\mathcal{U}} d(\tilde{\mathbf{x}}, \hat{\mathbf{x}}) < \max_{\hat{\mathbf{x}}\in\mathcal{U}} d(\mathbf{x}^*, \hat{\mathbf{x}})$, which is a contradiction. Therefore, $\mathbf{x}^* \in \mathcal{X}^{OPT}$ and thus is feasible for (6). As a result, $\mathbf{x}^*$ is an optimal solution to (6). Since $\mathbf{x}^* \in \mathcal{X}^{OPT}$, there exists a solution $(\mathbf{c}^*, \mathbf{y}^*)$ feasible for $\mathbf{IO}(\mathbf{x}^*)$. The rest of the proof follows the proof of Proposition 3. □

Thus, when $\mathcal{U} \subset \bar{\mathcal{X}}$, Proposition 4 suggests that $\mathbf{x}^*$ can be found by solving a single optimization problem, instead of m problems as described in Theorem 1 for the general case. That is, formulation (7) over all i is replaced by

$$\begin{aligned} \underset{\mathbf{x}}{\text{minimize}} \quad & z \\ \text{subject to} \quad & z \geq d(\mathbf{x}, \hat{\mathbf{x}}), \quad \forall \hat{\mathbf{x}} \in \mathcal{U}, \\ & \mathbf{A}\mathbf{x} \geq \mathbf{b}. \end{aligned} \tag{11}$$

Corollary 2. *Given $\mathbf{x}^*$ from formulation (11), an optimal solution to $\mathbf{RIO}(\mathcal{U})$ is $(\mathbf{c}^*, \mathbf{y}^*, \mathbf{x}^*) = (\mathbf{a}_i/\|\mathbf{a}_i\|_1, \mathbf{e}_i/\|\mathbf{a}_i\|_1, \mathbf{x}^*)$ where i satisfies $\mathbf{a}_i'\mathbf{x}^* = b_i$.*

2.2.2. $\mathcal{U} \subseteq \mathcal{X}$

If $\mathcal{U}$ is contained in the feasible region of the forward problem, formulation (6) is equivalent to finding the distance of the worst-case point to the complement of a convex set. In general, the problem can be solved by decomposition into m sub-problems, as shown in Theorem 1. One important special case that can be solved efficiently is when the distance function corresponds to the duality gap.

Proposition 5. *Let $d(\mathbf{x}, \hat{\mathbf{x}}; \mathbf{c}) = |\mathbf{c}'(\hat{\mathbf{x}} - \mathbf{x})|$ for a given $\mathbf{c}$. Then $\mathbf{RIO}(\mathcal{U})$ can be written as*

$$\begin{aligned} \underset{\mathbf{c},\mathbf{y},\epsilon}{\text{minimize}} \quad & \epsilon \\ \textit{subject to} \quad & \mathbf{A}'\mathbf{y} = \mathbf{c}, \\ & \mathbf{c}'\hat{\mathbf{x}} - \mathbf{b}'\mathbf{y} \leq \epsilon, \quad \forall \hat{\mathbf{x}} \in \mathcal{U}, \\ & \|\mathbf{c}\|_1 = 1, \\ & \mathbf{y} \geq \mathbf{0}. \end{aligned} \tag{12}$$

Proof. In $\mathbf{RIO}(\mathcal{U})$, replace $d(\mathbf{x}, \hat{\mathbf{x}})$ with $|\mathbf{c}'(\hat{\mathbf{x}} - \mathbf{x})|$ and remove the primal feasibility constraint since $\hat{\mathbf{x}} \in \mathcal{X}$. Additionally, since any optimal $\mathbf{x}$ for $\mathbf{RIO}(\mathcal{U})$ is in $\mathcal{X}^{\text{OPT}}$ and $\hat{\mathbf{x}} \in \mathcal{X}$, due to weak duality we have $\mathbf{c}'(\hat{\mathbf{x}} - \mathbf{x}) \geq 0$, implying that the absolute value operator can be removed. Further, we substitute $\max_{\hat{\mathbf{x}}\in\mathcal{U}} d(\mathbf{x}, \hat{\mathbf{x}})$ by ϵ where $\epsilon \geq \mathbf{c}'(\hat{\mathbf{x}} - \mathbf{x})$ for all $\hat{\mathbf{x}} \in \mathcal{U}$. Substituting $\mathbf{c}'\mathbf{x}$ by $\mathbf{b}'\mathbf{y}$ using the second constraint of $\mathbf{RIO}(\mathcal{U})$ (strong duality), we obtain formulation (12). □

In formulation (12), the second constraint relaxes strong duality and ensures that the duality gap is minimized for the worst-case realization of $\hat{\mathbf{x}} \in \mathcal{U}$. In practice, we often *a priori* know that all cost coefficients should be non-negative, in which case we impose an additional constraint $\mathbf{c} \geq \mathbf{0}$ to (12), which allows the constraint $\|\mathbf{c}\|_1 = 1$ to be replaced by $\sum_{j\in J} c_j = 1$, and thus the modified formulation becomes a linear or conic program when $\mathcal{U}$ is a polyhedron or an ellipsoid, respectively [5].

2.2.3. $\mathcal{U} \cap \mathcal{X} \neq \emptyset$ *and* $\mathcal{U} \cap \bar{\mathcal{X}} \neq \emptyset$

Finally we consider $\mathcal{U}$ that partially overlaps $\mathcal{X}$. We propose to divide the robust inverse problem into two problems, $\mathbf{RIO}(\mathcal{U} \cap \mathcal{X})$ and $\mathbf{RIO}(\mathcal{U} \cap \bar{\mathcal{X}})$, which we solve separately as described in Sections 2.2.1 and 2.2.2, and find $\mathbf{c}^*$ that corresponds to the worst-case distance between the two.

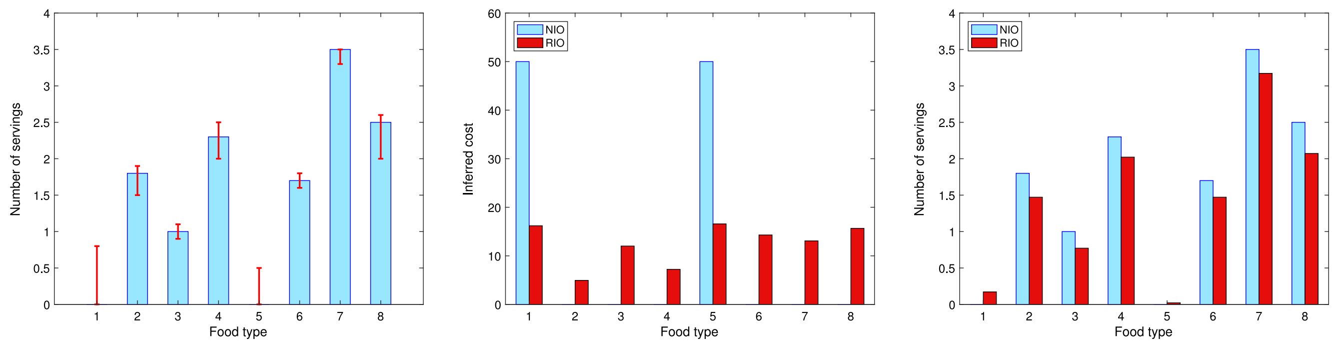
Proposition 6. $\mathbf{RIO}(\mathcal{U})$ *is equivalent to* $\max\{\mathbf{RIO}(\mathcal{U} \cap \mathcal{X}), \mathbf{RIO}(\mathcal{U} \cap \bar{\mathcal{X}})\}$.

Proof. Assume $\mathcal{U} \cap \mathcal{X} \neq \emptyset$ and $\mathcal{U} \cap \bar{\mathcal{X}} \neq \emptyset$. Recall the objective function of $\mathbf{RIO}(\mathcal{U})$ is $\min_{\mathbf{c},\mathbf{y},\mathbf{x}} \left\{\max_{\hat{\mathbf{x}}\in\mathcal{U}} d(\mathbf{x}, \hat{\mathbf{x}})\right\}$. Then we have

$$\begin{aligned} & \min_{\mathbf{c},\mathbf{y},\mathbf{x}} \left\{ \max_{\hat{\mathbf{x}}\in\mathcal{U}} d(\mathbf{x}, \hat{\mathbf{x}}) \right\} \\ = & \min_{\mathbf{c},\mathbf{y},\mathbf{x}} \left\{ \max \left\{ \max_{\hat{\mathbf{x}}\in\mathcal{U}\cap\mathcal{X}} d(\mathbf{x}, \hat{\mathbf{x}}), \max_{\hat{\mathbf{x}}\in\mathcal{U}\cap\bar{\mathcal{X}}} d(\mathbf{x}, \hat{\mathbf{x}}) \right\} \right\} \\ = & \max \left\{ \min_{\mathbf{c},\mathbf{y},\mathbf{x}} \left\{ \max_{\hat{\mathbf{x}}\in\mathcal{U}\cap\mathcal{X}} d(\mathbf{x}, \hat{\mathbf{x}}) \right\}, \min_{\mathbf{c},\mathbf{y},\mathbf{x}} \left\{ \max_{\hat{\mathbf{x}}\in\mathcal{U}\cap\bar{\mathcal{X}}} d(\mathbf{x}, \hat{\mathbf{x}}) \right\} \right\} \\ = & \max\{\mathbf{RIO}(\mathcal{U} \cap \mathcal{X}), \mathbf{RIO}(\mathcal{U} \cap \bar{\mathcal{X}})\}. \quad \square \end{aligned}$$

3. Numerical example

In this section, we apply the proposed robust inverse optimization methodology to a small-scale example of the diet problem



(a) Observation of varying food intake.

(b) Inferred costs.

(c) Optimal diet recommendations.

Fig. 2. Food intake observation and nominal and robust costs. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

briefly described in Section 1. The forward optimization problem is to determine optimal intake of each food type to minimize the overall cost while satisfying pre-specified nutritional constraints. What we refer to as cost is not necessarily monetary but represents the user's general perception of a specific food type: higher cost implies lower preference. Inverse optimization determines the cost function of the forward diet problem such that the preferences encoded in the observed diets are preserved.

The nominal inverse problem takes a single observation of the user's food intake, $\mathbf{x}^0$, as input (e.g., a one-day observation of the user behavior); that is, there is no uncertainty. We assume that the observed diet may or may not be feasible for the forward problem and may be represented as an uncertainty set. Data used for the forward and inverse problems is shown in Table B.1 in the Appendix. For the robust inverse problem, we assume a polyhedral uncertainty set surrounding the nominal, one-day observation, as outlined in Table B.2 in the Appendix. We test all three cases: the uncertainty set being inside, outside, and partially inside the feasible region. For the inside case, both the absolute duality gap and the ℓ_∞ norm are used for the distance function, while for the other cases only the ℓ_∞ norm is used. In our experiments, the inside duality gap case is solved by solving (12) with an additional constraint $\mathbf{c} \geq \mathbf{0}$, and other instances are solved using Theorem 1 (i.e., formulation (7)). The linear programming reformulation of (7) for the ℓ_∞ norm case is shown in (A.1) in the Appendix; the reformulations of (12) for the inside duality gap case can be derived similarly. While we only discuss the results for the inside duality gap instance in this section, implications are similar for the remaining instances. Results for all of the instances are provided in the Appendix.

3.1. Nominal and robust cost vectors

We solve $\mathbf{NIO}(\mathbf{x}^0)$ and $\mathbf{RIO}(\mathcal{U})$ with $\mathbf{x}^0$ and $\mathcal{U}$ as input, respectively. If there are multiple optimal solutions, we pick the cost vector that minimizes the maximum cost of any food item. The cost values were normalized so they add up to 100. Fig. 2(a) shows an observed one-time food intake (i.e., $\mathbf{x}^0$) in blue (light color) bars and the uncertainty set around the initial observation (i.e., $\mathcal{U}$) in red (dark color). As $\mathbf{x}^0$ does not contain any meat and yogurt (food type 1 and 5, respectively), the nominal inverse model returns $\mathbf{c}^* \in \{\mathbf{c} \mid c_1 + c_5 = 100, c_j = 0 \text{ for } j \neq 1, 5, j \in \mathcal{J}\}$ (see Fig. 2(b) for example), i.e., costs are non-zero only for these two food types so that avoiding them is optimal for this specific person. On the other hand, the red bars in Fig. 2(b) show that the costs obtained by the robust model are non-zero for *all* food types, with decreased values for types 1 and 5, which reflects the increase in food types 1 and 5 in the uncertainty set.

3.2. Nominal and robust diet recommendation

To compare diet plans generated by the nominal and robust cost vectors, we solve the forward problem with the respective costs. If there exist multiple optimal solutions, we use a secondary objective for the forward problem by which to select the solution that is closest to the original observation(s): when we use the nominal cost vector, we find the solution closest to $\hat{\mathbf{x}}$ in ℓ_∞ norm; for the robust cost vector, we find the solution closest to the centroid of the uncertainty set in ℓ_∞ norm. As shown in Fig. 2(c), food types 1 and 5 are not included in the nominal diet plan but are present in the robust one, reflecting the observed increase in their intake in the uncertainty set.

3.3. The effect of uncertainty on cost values

Next, we compare the robustness of the nominal and robust cost vectors by comparing variation of the achieved objective function values. Let $\mathbf{c}^N$ and $\mathbf{c}^R$ be the robust and nominal cost vectors, respectively. For randomly generated 100 diets $\tilde{\mathbf{X}} = [\mathbf{x}^1, \ldots, \mathbf{x}^{100}]$, each in $\mathcal{U}$, the variance of the resulting objective function values associated with $\mathbf{c}^N$ (i.e., var($\mathbf{c}^{N'}\tilde{\mathbf{X}}$)) was 175.02, whereas var($\mathbf{c}^{R'}\tilde{\mathbf{X}}$) was 27.47. That is, the robust cost vector gives a more consistent cost evaluation of the person's changing behavior and thus renders the varying diets closer to optimality. Results for other instances can be found in Table B.4 in the Appendix.

Appendix A. Supplementary data

Supplementary material related to this article can be found online at https://doi.org/10.1016/j.orl.2018.03.007.

References

[1] R.K. Ahuja, J.B. Orlin, Inverse optimization, Oper. Res. 49 (5) (2001) 771–783.
[2] A. Aswani, Z.-J.M. Shen, A. Siddiq, Inverse optimization with noisy data, Oper. Res. (2018) forthcoming.
[3] T. Ayer, Inverse optimization for assessing emerging technologies in breast cancer screening, Ann. Oper. Res. 230 (2015) 57–85.
[4] A. Ben-Tal, L. Ghaoui, A. Nemirovski, Robust optimization, Princeton University Press, 2009.
[5] D. Bertsimas, D.B. Brown, C. Caramanis, Theory and applications of robust optimization, SIAM Rev. 53 (3) (2011) 464–501.
[6] D. Bertsimas, V. Gupta, I.Ch. Paschalidis, Inverse optimization: A new perspective on the Black-Litterman model, Oper. Res. 60 (6) (2012) 1389–1403.
[7] D. Bertsimas, V. Gupta, I.Ch. Paschalidis, Data-driven estimation in equilibrium using inverse optimization, Math. Program. 153 (2) (2015) 595–633.
[8] J. Birge, A. Hortaçsu, M. Pavlin, Inverse optimization for the recovery of market structure from market outcomes: An application to the MISO electricity market, Oper. Res. 65 (4) (2017) 837–855.

[9] T.C.Y. Chan, T. Lee, D. Terekhov, Inverse optimization: Closed-form solutions, geometry and goodness of fit, Manage. Sci. (2018) forthcoming.
[10] A. Chassein, M. Goerigk, Variable-sized uncertainty and inverse problems in robust optimization, European J. Oper. Res. 264 (1) (2018) 17–28.
[11] J.Y.J. Chow, W.W. Recker, Inverse optimization with endogenous arrival time constraints to calibrate the household activity pattern problem, Transp. Res. B 46 (3) (2012) 463–479.
[12] Z. Erkin, M.D. Bailey, L.M. Maillart, A.J. Schaefer, M.S. Roberts, Eliciting patients' revealed preferences: An inverse Markov decision process approach, Decis. Anal. 7 (4) (2010) 358–365.
[13] P.M. Esfahani, S. Shafieezadeh-Abadeh, G.Adiwena Hanasusanto, D. Kuhn, Data-driven inverse optimization with incomplete information, Math. Program. 167 (1) (2018) 191–234.
[14] A. Ghate, Inverse optimization in countably infinite linear programs, Oper. Res. Lett. 43 (3) (2015) 231–235.
[15] C. Heuberger, Inverse combinatorial optimization: A survey on problems, methods, and results, J. Comb. Optim. 8 (3) (2004) 329–361.
[16] G. Iyengar, W. Kang, Inverse conic programming with applications, Oper. Res. Lett. 33 (3) (2005) 319–330.
[17] A. Keshavarz, Y. Wang, S. Boyd, Imputing a convex objective function, in: 2011 IEEE International Symposium on Intelligent Control, (ISIC), IEEE, 2011, pp. 613–619.
[18] J.B. Lamperski, A.J. Schaefer, A polyhedral characterization of the inverse-feasible region of a mixed-integer program, Oper. Res. Lett. 43 (6) (2015) 575–578.
[19] A.J. Schaefer, Inverse integer programming, Optim. Lett. 3 (4) (2009) 483–489.
[20] M.D. Troutt, A.A. Brandyberry, C. Sohn, S.K. Tadisina, Linear programming system identification: The general nonnegative parameters case, European J. Oper. Res. 185 (1) (2008) 63–75.
[21] M.D. Troutt, W.-K. Pang, S.-H. Hou, Behavioral estimation of mathematical programming objective function coefficients, Manage. Sci. 52 (3) (2006) 422–434.
[22] L. Wang, Cutting plane algorithms for the inverse mixed integer linear programming problem, Oper. Res. Lett. 37 (2) (2009) 114–116.