Inverse optimization in countably infinite linear programs

Archis Ghate (archis@uw.edu)

Abstract

Given the costs and a feasible solution for a linear program, inverse optimization involves finding new costs that are close to the original ones and make the given solution optimal. We develop an inverse optimization framework for countably infinite linear programs using the weighted absolute sum metric. We reformulate this as an infinite-dimensional mathematical program using duality. We propose a convergent algorithm that solves a sequence of finite-dimensional LPs to tackle it. We apply this to non-stationary Markov decision processes.

Keywords: infinite-dimensional optimization; duality theory; non-stationary Markov decision processes

1 Introduction

Countably infinite linear programs (CILPs) are infinite-dimensional linear programs (LPs) that include a countable number of variables and a countable number of constraints [3, 10, 11]. CILPs arise in infinite-horizon planning applications such as production planning, equipment replacement, and capacity expansion. Special cases of CILPs include minimum cost flow problems on infinite networks [17, 21]; infinite horizon stochastic programs [12]; LP formulations of countable-state Markov decision processes (MDPs) [11, 14, 16, 20]; and problems in robust optimization [9].

Inverse optimization in n-dimensional LPs refers to the following problem: given a feasible solution $x^* \in \Re^n$ to an LP with cost vector $c^* \in \Re^n$, find a $d \in \Re^n$ that (i) is as close (in an appropriate distance metric) as possible to c^*, and (ii) makes x^* optimal to a new LP where the cost vector is d. Ahuja and Orlin [1] showed that if we used the weighted absolute sum metric in $\Re^n$, then this problem reduces to a finite-dimensional LP. Chan et al. [6] stated that inverse optimization has been studied in the context of shortest path problems; network and combinatorial optimization; integer programming; mixed integer programming; and convex optimization. Despite the recent surge of interest in CILPs, inverse optimization has not yet been studied in the CILP context. The goal in this paper is to develop an inverse optimization framework for CILPs.

We pursue a duality-based as in Ahuja and Orlin. However, the difficulty is that unlike finite-dimensional LPs, weak duality, complementary slackness, and strong duality may not hold in general in primal-dual pairs of CILPs [3, 9, 15, 18, 19]. It is essential to embed the primal and the dual CILPs in appropriately chosen sequence spaces to ensure that weak duality and complementary slackness hold, and then to impose additional restrictions for strong duality to hold. This task is rendered difficult owing to mathematical pathologies in sequence spaces and has been called the "Slater conundrum" [15]. The duality approach in Ghate [9] allows for the broadest class of CILPs. That framework is therefore utilized to cast CILP formulations here.

Following Ahuja and Orlin, we use the weighted absolute sum metric. The weights are embedded in an appropriate sequence space so that the corresponding metric is finite. We show that the constraints in our inverse optimization problem can be equivalently reformulated as a countably infinite set of linear constraints. We propose to solve a sequence of finite-dimensional LPs to tackle the resulting infinite-dimensional inverse optimization problem. We prove that accumulation points of any sequence of optimal solutions to these finite-dimensional LPs are optimal to the inverse optimization problem. We also prove that the sequence of optimal values of these finite-dimensional LPs converges to the optimal value of the inverse optimization problem. These results are applied to infinite-horizon *non-stationary* MDPs, thus extending recent work on inverse optimization in infinite-horizon *stationary* MDPs [8]. Proofs are provided in the supplementary material available via the author's website at http://faculty.washington.edu/archis/orl-inverse-opt-suppl.pdf.

2 Review of duality in CILPs

We first review CILP duality results from Ghate [9]. The symbol $\triangleq$ means "defined as." We use $\mathbb{N} \triangleq \{1, 2, \ldots\}$ to denote the set of all natural numbers and let $\Re^{\mathbb{N}}$ denote the space of all real-valued sequences. Let $Z \subseteq \Re^{\mathbb{N}}$ be a sequence space. Generic sequences in Z will often be denoted by c and will form the objective function coefficients in our CILPs. Let $b \in \Re^{\mathbb{N}}$ be a given sequence; and, for $i = 1, 2, \ldots$, let $A_{i\cdot} \triangleq (a_{i1}, a_{i2}, \ldots) \in \Re^{\mathbb{N}}$ be the ith row of a given doubly-infinite matrix A. Similarly, let $A_{\cdot j} \triangleq (a_{1j}, a_{2j}, \ldots) \in \Re^{\mathbb{N}}$ be the jth column of this matrix. As is common in the existing literature, we assume that, for each i, only a finite number of entries in $A_{i\cdot}$ is non-zero; similarly, we assume that, for each j, only a finite number of entries in $A_{\cdot j}$ is non-zero. This structure is ubiquitous (see [18, 19]) in Operations Research such as in shortest path formulations of infinite-horizon dynamic programs with finite states and actions [10]; minimum cost flow problems in infinite-staged networks with finite node degrees [17, 21]; and CILP formulations of infinite-horizon non-stationary MDPs with finite states and actions [11, 13].

Now let $X \subseteq \Re^{\mathbb{N}}$ be the subspace of all sequences $x \in \Re^{\mathbb{N}}$ for which

C1. the series $C(x) \triangleq \sum_{j=1}^{\infty} c_j x_j$ converges for any $c \in Z$.

Let Y be the subspace of all $y \in \Re^{\mathbb{N}}$ for which

C2. the series $B(y) \triangleq \sum_{i=1}^{\infty} b_i y_i$ converges; and

C3. for every $x \in X$, we have $\sum_{i=1}^{\infty} L_i(x, y_i) < \infty$, where $L_i(x, y_i) \triangleq \sum_{j=1}^{\infty} |a_{ij} x_j y_i|$ for each $i = 1, 2, \ldots$.

Consider the following pair of primal-dual CILPs

$$(P) \quad V(P) = \inf \sum_{j=1}^{\infty} c_j x_j \tag{1}$$

$$\sum_{j=1}^{\infty} a_{ij} x_j = b_i, \quad i = 1, 2, \ldots, \tag{2}$$

$$x_j \geq 0, \quad j = 1, 2, \ldots, \tag{3}$$

$$x \in X, \tag{4}$$

and

$$(D) \quad V(D) = \sup \sum_{i=1}^{\infty} b_i y_i \tag{5}$$

$$\sum_{i=1}^{\infty} a_{ij} y_i \leq c_j, \ j = 1, 2, \ldots, \tag{6}$$

$$y \in Y. \tag{7}$$

Let F and G denote the feasible regions of these two problem, respectively.

Theorem 2.1. Weak duality: *(Ghate [9]) For any $x \in F$ and any $y \in G$, $\sum_{j=1}^{\infty} c_j x_j \geq \sum_{i=1}^{\infty} b_i y_i$. Hence $\infty \geq V(P) \geq V(D) \geq -\infty$ (here, the infimum over an empty set is interpreted as $+\infty$ and the supremum over an empty set is interpreted as $-\infty$). Also, if $x \in F$ and $y \in G$ are such that $\sum_{j=1}^{\infty} c_j x_j = \sum_{i=1}^{\infty} b_i y_i$, then x is optimal to (P) and y is optimal to (D), and thus strong duality holds.*

Vectors $x \in X$ and $y \in Y$ are called complementary if $x_j \Big(c_j - \sum_{i=1}^{\infty} a_{ij} y_i \Big) = 0$ for each $j = 1, 2, \ldots$.

Theorem 2.2. Complementary slackness: *(Ghate [9])*

1. *Suppose $x \in F$ and $y \in G$, and suppose x and y are complementary. Then x is optimal to (P), y is optimal to (D), and $V(P) = V(D)$. Thus, strong duality holds in this case.*

2. *Suppose x is optimal to (P), y is optimal to (D), and $V(P) = V(D)$ (that is, strong duality holds). Then x and y are complementary.*

For any increasing sequences of positive integers N_n and M_n, consider the truncation given by

$$P(n) \quad V(P(n)) = \inf \sum_{j=1}^{M_n} c_j x_j \tag{8}$$

$$\sum_{j=1}^{M_n} a_{ij} x_j \ = \ b_i, \ \ i = 1, 2, \ldots, N_n, \tag{9}$$

$$x_j \ \geq \ 0, \ \ j = 1, 2, \ldots, M_n. \tag{10}$$

Let $X^*(n) \subseteq X$ denote the (possibly empty) set of optimal solutions to $P(n)$. The dual of $P(n)$ is

$$D(n) \quad V(D(n)) = \sup \sum_{i=1}^{N_n} b_i y_i \tag{11}$$

$$\sum_{i=1}^{N_n} a_{ij} y_i \leq c_j, \ j = 1, 2, \ldots, M_n. \tag{12}$$

Let $Y^*(n) \subseteq Y$ denote the (possibly empty) set of optimal solutions to $D(n)$.

We use the product topology on sequence spaces in $\Re^{\mathbb{N}}$.

Theorem 2.3. Strong duality: *(Ghate [9]) Suppose there exist the aforementioned sequences $P(n)$ and $D(n)$ of finite-dimensional primal-dual problems and sets $\mathcal{C} \subseteq X$ and $\mathcal{K} \subseteq Y$ such that*

C4. *for each n, $X_{\mathcal{C}}(n) \triangleq \Big(X^*(n) \bigcap \mathcal{C} \Big) \neq \emptyset$;*

C5. *for each n, $Y_{\mathcal{K}}(n) \triangleq \Big(Y^*(n) \bigcap \mathcal{K} \Big) \neq \emptyset$;*

C6. $\exists$ *a sequence in* $X_{\mathcal{C}}(n) \times Y_{\mathcal{K}}(n)$ *with a convergent subsequence with a limit in* $\mathcal{C} \times \mathcal{K}$.

Then (P) *and* (D) *have optimal solutions in* $\mathcal{C}$ *and* $\mathcal{K}$*, and* $V(P) = V(D)$.

The product topologies on X and Y are countable products of the usual metrizable topology on $\Re$; hence they are metrizable (see Theorem 3.36 on page 89 of [2]). Every sequence in a compact set in the product topology on X thus has a convergent subsequence; similarly for Y (see Theorem 3.28 on page 86 of [2]). Condition C6 thus holds when $\mathcal{C}$ and $\mathcal{K}$ are compact. By the Tychonoff product theorem (see Theorem 2.61 on page 52 of [2]), this compactness holds if each component of optimal solutions to $P(n)$ and each component of optimal solutions to $D(n)$ can be bounded independently of n. See [9, 18, 19] for several applications where these conditions are met.

3 Inverse optimization formulation

Now suppose that a fixed $c^* \in Z$ is given. Also suppose that an $x^* \in X$ that is feasible to (P) is given. For each $d \in Z$, we will use (P_d), for emhpasis, to denote the CILP that is identical in form to (P) but now with cost coefficients d. We will use (D_d) to denote the dual of (P_d). Note that (P_d) and (D_d) satisfy C1-C3 and hence weak duality as in Theorem 2.1 and complementary slackness as in Theorem 2.2 hold for this primal-dual pair.

Following Ahuja and Orlin, we say that $d \in Z$ is *inverse feasible* with respect to x^* if x^* is an optimal solution to (P_d). Similarly, we will use $\mathcal{D}(x^*) \subseteq Z$ to denote the set of all $d \in Z$ that are inverse feasible with respect to x^*. Our inverse optimization problem involves finding a $d \in \mathcal{D}(x^*) \cap \mathcal{C}$ that is as close as possible to c in the weighted absolute sum metric, where $\mathcal{C} \triangleq \{d \in Z : l_j \leq d_j \leq u_j, \ \forall j\}$ is some compact set in the product topology on Z. Here, the given lower and upper bound vectors l, u belong to Z. Let $w \in \Re^{\mathbb{N}}$ be a fixed sequence of non-negative weights chosen by the decision-maker such that

$$\sum_{j=1}^{\infty} |w_j z_j| < \infty \text{ for every } z \in Z. \tag{13}$$

We then formulate the inverse optimization problem as

$$(\text{INV}(x^*)) \ \inf \ \sum_{j=1}^{\infty} w_j |c_j^* - d_j| \tag{14}$$

$$d \in \mathcal{D}(x^*), \tag{15}$$

$$l_j \leq d_j \leq u_j, j = 1, 2, \ldots. \tag{16}$$

We will now use complementary slackness to characterize $\mathcal{D}(x^*)$. Let $\mathcal{T}(x^*)$ denote the support of x^*; that is, $\mathcal{T}(x^*) \triangleq \{j : x_j^* > 0\}$. Let $\mathcal{U}(x^*)$ denote the subset of all $(y, d) \in Y \times Z$ that satisfy

$$\sum_{i=1}^{\infty} a_{ij} y_i = d_j, \ j \in \mathcal{T}(x^*), \tag{17}$$

$$\sum_{i=1}^{\infty} a_{ij} y_i \leq d_j, \ j \notin \mathcal{T}(x^*). \tag{18}$$

Lemma 3.1. *If there exists a pair* $(y, d) \in \mathcal{U}(x^*)$*, then* $d \in \mathcal{D}(x^*)$*,* $V(P_d) = V(D_d)$*, and* y *is an optimal solution to* (D_d)*. If there exists a* $d \in \mathcal{D}(x^*)$ *and if* $V(P_d) = V(D_d)$ *with* y *as an optimal solution to* (D_d)*, then* $(y, d) \in \mathcal{U}(x^*)$.

The second claim in Lemma 3.1 shows that the sets $\mathcal{D}(x^*)$ and $\mathcal{U}(x^*)$ are not quite "equivalent"; such equivalence needs strong duality. This complication does not arise in inverse optimization in finite-dimensional LPs. We therefore make the following overarching assumption.

Assumption 3.2. *For each $d \in Z$, conditions C4-C6 hold for the primal-dual pair (P_d) and (D_d).*

This assumption assures us that strong duality holds between (P_d) and (D_d) for each $d \in Z$, and hence $\mathcal{D}(x^*)$ and $\mathcal{U}(x^*)$ are "equivalent." We thus rewrite the inverse optimization problem, after expanding the set $\mathcal{U}(x^*)$ as defined in (17)-(18), as

$$(\text{INV}(x^*)) \ \inf \ \sum_{j=1}^{\infty} w_j |c_j^* - d_j| \tag{19}$$

$$\sum_{i=1}^{\infty} a_{ij} y_i = d_j, \ j \in \mathcal{T}(x^*), \tag{20}$$

$$\sum_{i=1}^{\infty} a_{ij} y_i \leq d_j, \ j \notin \mathcal{T}(x^*), \tag{21}$$

$$l_j \leq d_j \leq u_j, j = 1, 2, \ldots, \tag{22}$$

$$d \in Z, \tag{23}$$

$$y \in Y. \tag{24}$$

We use $\mathcal{F}(x^*) \subseteq Y \times Z$ to denote the (possibly empty) set of feasible solutions to this problem. We next present an algorithm that tackles this problem by solving a sequence of finite-dimensional LPs. We investigate the asymptotic convergence behavior of this algorithm.

4 A convergent solution algorithm

We first make a key observation that relies on the property of our doubly infinite constraint matrix A that each row and each column include a finite number of non-zero entries. Specifically, there exist increasing sequences M_n and N_n of positive integers indexed by $n = 1, 2, \ldots$ such that variables $y_{N_n+1}, y_{N_n+2}, \ldots$ do not appear in constraints (20)-(21) that correspond to $j = 1, 2, \ldots, M_n$. Now let $\mathcal{T}_n(x^*)$ denote a truncation of $\mathcal{T}(x^*)$ that only considers the first M_n components of x^*. That is, $\mathcal{T}_n \triangleq \{1 \leq j \leq M_n : x_j^* > 0\}$. We then consider a truncation of $(\text{INV}(x^*))$ that is given by

$$(\text{INV}^n(x^*)) \ \inf \ \sum_{j=1}^{M_n} w_j |c_j^* - d_j| \tag{25}$$

$$\sum_{i=1}^{N_n} a_{ij} y_i = d_j, \ j \in \mathcal{T}_n(x^*), \tag{26}$$

$$\sum_{i=1}^{N_n} a_{ij} y_i \leq d_j, \ j \notin \mathcal{T}_n(x^*), \tag{27}$$

$$l_j \leq d_j \leq u_j, \ j = 1, 2, \ldots, M_n. \tag{28}$$

We use $\mathcal{F}_n(x^*) \in \Re^{N_n} \times \Re^{M_n}$ to denote the (possibly empty) set of feasible solutions to $(\text{INV}^n(x^*))$. It is possible to view $\mathcal{F}_n(x^*)$ as a subset of $Y \times Z$ by appropriately appending values for variables $y_{N_n+1}, y_{N_n+2}, \ldots$ and $d_{M_n+1}, d_{M_n+2}, \ldots$.

Lemma 4.1. *The truncation to $\Re^{N_n} \times \Re^{M_n}$ of any feasible solution to $(INV(x^*))$ is feasible to $(INV^n(x^*))$. Similarly, the truncation to $\Re^{N_n} \times \Re^{M_n}$ of any feasible solution to $(INV^{n+1}(x^*))$ is feasible to $(INV^n(x^*))$. In particular, $\mathcal{F}_{n+1}(x^*) \subseteq \mathcal{F}_n(x^*)$ for $n = 1, 2, \ldots$.*

C7. The feasible sets $\mathcal{F}_n(x^*)$ are non-empty for each n, and there exist sequences $\gamma \in Y$, $\delta \in Z$ independently of n such that the compact set $\mathcal{E} \triangleq \{(y, d) \in Y \times Z : |y_i| \leq \gamma_i, \forall i; \ |d_j| \leq \delta_j, \ \forall j\}$ satisfies $\mathcal{F}_n(x^*) \subseteq \mathcal{E}$ for all n.

Lemma 4.2. *Suppose C7 holds. Then problem $(INV^n(x^*))$ has an optimal solution; we denote the optimal value of $(INV^n(x^*))$ by $R_n(x^*)$.*

Lemma 4.3. *Suppose C7 holds. Then the feasible region of $(INV(x^*))$ is non-empty and compact.*

Lemma 4.4. *The objective function in $(INV(x^*))$ is continuous over the compact set $\mathcal{E}$.*

Corollary 4.5. *Suppose C7 holds. Then $(INV(x^*))$ has an optimal solution in $\mathcal{E}$; we denote the optimal value of $(INV(x^*))$ by $R(x^*)$.*

Theorem 4.6. Solution and value convergence: *Suppose C7 holds. Consider any sequence of optimal solutions $(\hat{y}(n), \hat{d}(n)) \in \mathcal{E}$ to problems $(INV^n(x^*))$; this sequence has a convergent subsequence that converges to an optimal solution of $(INV(x^*))$ in $\mathcal{E}$. In fact, every convergent subsequence of $(\hat{y}(n), \hat{d}(n))$ exhibits this property. Furthermore, the sequence of optimal objective values $R_n(x^*)$ of problems $(INV^n(x^*))$ converges to the optimal objective value $R(x^*)$ of problem $(INV(x^*))$.*

Our proof of this result employs a modification of the argument used in the proof of Berge's maximum principle [5]. The overall idea is to show that a sequence of optimal solutions in $\mathcal{E}$ to problems (INV$^n(x^*)$) has a convergent subsequence. Such optimal solutions exist by Lemma 4.2 and such a convergent subsequence exists because $\mathcal{E}$ is compact. Then we show by contradiction that the limit of this convergent subsequence is optimal to (INV(x^*)). The existence of an optimal solution to (INV(x^*)) that is exploited in this proof by contradiction is assured by Corollary 4.5, which is in turn derived from Lemmas 4.3 and 4.4. The proof by contradiction also uses the result in Lemma 4.1 that truncations of feasible solutions to (INV(x^*)) are feasible to (INV$^n(x^*)$).

As in Ahuja and Orlin, we simplify (INV$^n(x^*)$) using a standard technique. We express $c_j^* - d_j$ as $\alpha_j - \beta_j$, for $\alpha_j, \beta_j \geq 0$, and instead of minimizing $\sum_{j=1}^{M_n} w_j |c_j^* - d_j|$, we minimize $\sum_{j=1}^{\infty} w_j(\alpha_j + \beta_j)$. The idea is that in any optimal solution to the resulting problem, at most one of α_j, β_j would be positive for each j. This holds because otherwise we can reduce the objective value by decreasing both α_j and β_j without sacrificing feasibility. This reduces (INV$^n(x^*)$) to the LP

$$(\text{INV}^n(x^*)) \ \inf \ \sum_{j=1}^{M_n} w_j \alpha_j + \sum_{j=1}^{M_n} w_j \beta_j \tag{29}$$

$$\sum_{i=1}^{N_n} a_{ij} y_i + \alpha_j - \beta_j = c_j^*, \ j \in \mathcal{T}_n(x^*), \tag{30}$$

$$\sum_{i=1}^{N_n} a_{ij} y_i + \alpha_j - \beta_j \leq c_j^*, \ j \notin \mathcal{T}_n(x^*), \tag{31}$$

$$l_j \leq d_j \leq u_j, \ j = 1, 2, \ldots, M_n, \tag{32}$$

$$\alpha, \beta \geq 0. \tag{33}$$

5 Application to non-stationary Markov decision processes

Infinite-horizon non-stationary Markov decision processes (MDPs) relax the restrictive assumption in stationary MDPs that problem data do not change over time [7, 11, 13]. In a non-stationary MDP, a decision-maker observes a dynamic system at the beginning of periods $n = 1, 2, \ldots$ to be in some states $s \in \mathcal{S}$, where $\mathcal{S} \triangleq \{1, 2, \ldots, S\}$ is a finite set. The decision-maker then chooses an action $a \in \mathcal{A}$; here, $\mathcal{A} \triangleq \{1, 2, \ldots, A\}$ is also a finite set. Given that action a was chosen in state s in period n, the system makes a transition to state s' with probability $p_n(s'|s, a)$, incurring expected cost $c_n(s, a)$ such that $|c_n(s, a)| \leq \bar{c} < \infty$. The goal is to minimize the total infinite-horizon discounted expected cost when the discount factor is $0 < \lambda < 1$.

It is known that a deterministic Markovian decision rule is optimal to the non-stationary MDP [20, 11]. Such a decision rule prescribes an action $a_n(s) \in \mathcal{A}$ whenever the system occupies state $s \in \mathcal{S}$ in period n, irrespective of the history of the process. Suppose $q \triangleq \{q_n\}$ is a sequence of *positive* vectors in $\Re^S$ such that $\sum\limits_{n \in \mathbb{N}} \sum\limits_{s \in \mathcal{S}} q_n(s) = 1$. A deterministic Markovian policy can be found (see [11, 20]) by solving the CILP

$$(P0) \;\; \min \sum_{n \in \mathbb{N}} \sum_{s \in \mathcal{S}} \sum_{a \in \mathcal{A}} c_n(s, a) x_n(s, a) \tag{34}$$

$$\sum_{a \in \mathcal{A}} x_1(s, a) = q_1(s), \;\; \text{for } s \in \mathcal{S}, \tag{35}$$

$$\sum_{a \in \mathcal{A}} x_n(s, a) - \lambda \sum_{s' \in \mathcal{S}} \sum_{a \in \mathcal{A}} p_{n-1}(s|s', a) x_{n-1}(s', a) = q_n(s), \;\; \text{for } s \in \mathcal{S}, \;\; n \in \mathbb{N} \setminus \{1\}, \tag{36}$$

$$x_n(s, a) \geq 0, \;\; \text{for } s \in \mathcal{S}, \;\; a \in \mathcal{A}, \;\; n \in \mathbb{N}, \tag{37}$$

$$x \in l_1. \tag{38}$$

Note here that the costs are bounded and hence reside in l_∞. That is, $Z = l_\infty$. Moreover, it is easy to show, with a proof by induction similar to Lemma 2.1 in Ghate [11], that, for any feasible solution to constraints (35)-(37), we have, $\sum\limits_{n \in \mathbb{N}} \sum\limits_{s \in \mathcal{S}} \sum\limits_{a \in \mathcal{A}} |x_n(s, a)| < \infty$. Thus, the restriction $x \in l_1$ in the above CILP is in fact without loss of feasibility with respect to the rest of the constraints.

The CILP $(P0)$ has an extreme point optimal solution (see [11]). Its extreme points are characterized by the property that, for every $(n, s) \in \mathbb{N} \times \mathcal{S}$, there is a single action $a_n(s) \in \mathcal{A}$ (called a *basic* action) for which $x_n(s, a) > 0$, and $x_n(s, a) = 0$ for all other (non-basic) actions in $\mathcal{A}$. Thus, an optimal extreme point solution induces an optimal deterministic Markovian policy that prescribes the basic actions in each state in each period. We also mention that, as explained in Ghate [11], problem $(P0)$ can be viewed as a minimum cost flow problem in a staged *hypernetwork*. The stages in this hypernetwork are indexed by n; nodes are indexed by (n, s); hyperarcs are indexed by (n, s, a); hyperarc flows are denoted by $x_n(s, a)$; and hyperarc unit flow costs are given by $c_n(s, a)$.

The dual of $(P0)$ is given by

$$(D0) \;\; \max \sum_{n \in \mathbb{N}} \sum_{s \in \mathcal{S}} q_n(s) y_n(s) \tag{39}$$

$$y_n(s) - \lambda \sum_{s' \in \mathcal{S}} p_n(s'|s, a) y_{n+1}(s') \leq c_n(s, a), \;\; \text{for } s \in \mathcal{S}, \;\; a \in \mathcal{A}, \;\; n \in \mathbb{N}, \tag{40}$$

$$y \in l_\infty. \tag{41}$$

It is shown in [11] based on a result in [20] that the (unique) optimal values of variables $y_n(s)$ in $(D0)$ equal the optimal-costs-to-go starting from state $s \in \mathcal{S}$ in period n. Owing to this interpretation, we have that $|y_n(s)| \leq \frac{\bar{c}}{1-\lambda}$ for all (n, s) without loss of optimality in $(D0)$. It is shown in Ghate [11] that conditions C1-C6 hold for the primal-dual pair $(P0) - (D0)$. This conclusion does not depend on the specific values of costs $c_n(s, a)$. That is, Assumption 3.2 holds, and thus, weak duality, complementary slackness, and strong duality hold.

Now suppose that costs $c^*_n(s, a)$, for $n \in \mathbb{N}$, $s \in \mathcal{S}$, and $a \in \mathcal{A}$ are given such that $|c^*_n(s, a)| \leq \bar{c}$. Suppose a deterministic Markovian policy is also given. Equivalently, a feasible (extreme point) solution x^* to $(P0)$ is given. Inverse optimization then involves finding bounded costs $|d_n(s, a)| \leq \bar{c}$, for $n \in \mathbb{N}$, $s \in \mathcal{S}$, $a \in \mathcal{A}$, such that this deterministic Markovian policy is optimal; in other words, we want the solution x^* to be optimal to $(P0_d)$. Moreover, we need to do this such that d is as close as possible to c in the weighted absolute sum distance. We embed the costs d in l_∞ and embed the non-negative weight vector $w \in l_1$ so that (13) holds. Here, we use $\mathcal{T}(x^*)$ to denote the set of basic hyperarcs; that is, the set of $(n, s, a) \in \mathbb{N} \times \mathcal{S} \times \mathcal{A}$ such that $x^*_n(s, a) > 0$. The inverse non-stationary MDP problem is then given by

$$(\text{P0INV}(x^*)) \;\; \inf \sum_{n\in\mathbb{N}} \sum_{s\in\mathcal{S}} \sum_{a\in\mathcal{A}} w_n(s, a)|c^*_n(s, a) - d_n(s, a)| \tag{42}$$

$$y_n(s) - \lambda \sum_{s'\in\mathcal{S}} p_n(s'|s, a)y_{n+1}(s') = d_n(s, a),\;\; (n, s, a) \in \mathcal{T}(x^*), \tag{43}$$

$$y_n(s) - \lambda \sum_{s'\in\mathcal{S}} p_n(s'|s, a)y_{n+1}(s') \leq d_n(s, a),\;\; (n, s, a) \notin \mathcal{T}(x^*), \tag{44}$$

$$-\bar{c} \leq d_n(s, a) \leq \bar{c},\;\; s \in \mathcal{S},\;\; a \in \mathcal{A},\;\; n \in \mathbb{N}, \tag{45}$$

$$-\frac{\bar{c}}{1-\lambda} \leq y_n(s) \leq \frac{\bar{c}}{1-\lambda},\;\; s \in \mathcal{S},\;\; n \in \mathbb{N}, \tag{46}$$

$$d \in l_\infty, \tag{47}$$

$$y \in l_\infty. \tag{48}$$

Constraint (46) is actually not present in the original inverse optimization formulation. We added it here without loss of feasibility with respect to the other constraints in $(\text{P0INV}(x^*))$ because of the aforementioned cost-to-go interpretation of (and the corresponding bounds on) the y variables.

Constraints (43)-(44) corresponding to $n = 1, 2, \ldots, N$ do not include variables $y_{n+2}(s), y_{n+3}(s), \ldots$ for $s \in \mathcal{S}$. For $N = 1, 2, 3, \ldots$, the N-horizon truncation of $(\text{P0INV}(x^*))$ is thus

$$(\text{P0INV}^N(x^*)) \;\; \min \sum_{n=1}^{N} \sum_{s\in\mathcal{S}} \sum_{a\in\mathcal{A}} w_n(s, a)|c^*_n(s, a) - d_n(s, a)| \tag{49}$$

$$y_n(s) - \lambda \sum_{s'\in\mathcal{S}} p_n(s'|s, a)y_{n+1}(s') = d_n(s, a),\;\; (n, s, a) \in \mathcal{T}_N(x^*), \tag{50}$$

$$y_n(s) - \lambda \sum_{s'\in\mathcal{S}} p_n(s'|s, a)y_{n+1}(s') \leq d_n(s, a),\;\; (n, s, a) \notin \mathcal{T}_N(x^*), \tag{51}$$

$$-\bar{c} \leq d_n(s, a) \leq \bar{c},\;\; s \in \mathcal{S},\;\; a \in \mathcal{A},\;\; n = 1, \ldots, N, \tag{52}$$

$$-\frac{\bar{c}}{1-\lambda} \leq y_n(s) \leq \frac{\bar{c}}{1-\lambda},\;\; s \in \mathcal{S},\;\; n = 1, 2, \ldots, N+1. \tag{53}$$

For each $N = 1, 2, \ldots$, the feasible region of problem $(\text{P0INV}^N(x^*))$ is non-empty because the solution obtained by setting $y_{N+1}(s) = 0$ for all $s \in \mathcal{S}$; $d_n(s, a) = -\bar{c}$ for $(n, s, a) \in \mathcal{T}_N(x^*)$; $d_n(s, a) = \bar{c}$ for $(n, s, a) \notin \mathcal{T}_N(x^*)$; and finally, $y_n(s)$, for $n = 1, 2, \ldots, N$ and $s \in \mathcal{S}$, as the optimal costs-to-go for the resulting N-horizon MDP, is feasible to $(\text{P0INV}^N(x^*))$. The boundedness requirement in condition C7 clearly holds for problems (P0INV$^N(x^*)$) for all N owing to the explicit bounds on variables y and d in these problems. Thus, our solution and value convergence Theorem 4.6 applies. Finally, (P0INV$^N(x^*)$) can be reduced to a finite-dimensional LP as explained earlier in Section 4.

6 Acknowledgments

Funded in part by the National Science Foundation through grant #CMMI 1333260.

References

[1] R K Ahuja and J B Orlin. Inverse optimization part I: Linear programming and general problem. *Operations Research*, 49(5):771–783, 2001.

[2] C D Aliprantis and K C Border. *Infinite-dimensional analysis: a hitchhiker's guide*. Springer-Verlag, Berlin, Germany, 1994.

[3] E J Anderson and P Nash. *Linear programming in infinite-dimensional spaces: theory and applications*. John Wiley and Sons, Chichester, UK, 1987.

[4] T M Apostol. *Mathematical Analysis*. Addison Wesley, Reading, MA, USA, 1974.

[5] C Berge. *Topological Spaces*. McMillan, New York, NY, USA, 1963.

[6] T C Y Chan, T Craig, T Lee, and M B Sharpe. Generalized inverse multi-objective optimization with application to cancer therapy. *Operations Research*, 62(3):680–695, 2014.

[7] T Cheevaprawatdomrong, I E Schochetman, R L Smith, and A Garcia. Solution and forecast horizons for infinite-horizon non-homogeneous Markov decision processes. *Mathematics of Operations Research*, 32(1):51–72, 2007.

[8] Z Erkin, M D Bailey, L M Maillart, A J Schaefer, and M S Roberts. Eliciting patient's revealed preferences: an inverse Markov decision process approach. *Decision Analysis*, 7(4):358–365, 2010.

[9] A Ghate. Duality in countably infinite linear programs. http://papers.ssrn.com/sol3/papers.cfm?abstract_id=2497275, September 2014.

[10] A Ghate, D Sharma, and R L Smith. A shadow simplex method for infinite linear programs, forthcoming. *Operations Research*, 58(4):865–877, 2010.

[11] A Ghate and R L Smith. A linear programming approach to nonstationary infinite horizon markov decision processes. *Operations Research*, 61(2):413–425, 2013.

[12] R C Grinold. Infinite horizon stochastic programs. *SIAM Journal on Control and Optimization*, 24(6):1246–1260, 1986.

[13] I Lee, M A Epelman, H E Romeijn, and R L Smith. A linear programming approach to constrained nonstationary infinite-horizon Markov decision processes. Technical Report TR13-01, University of Michigan, Ann Arbor, MI, USA, 2013.

[14] I Lee, M A Epelman, H E Romeijn, and R L Smith. Simplex algorithm for countable-state discounted markov decision processes. submitted, 2014.

[15] K Martin, C T Ryan, and M Stern. The Slater Conundrum: Duality and Pricing in Infinite Dimensional Optimization. http://faculty.chicagobooth.edu/kipp.martin/root/Slater_Conundrum.pdf, April 2014.

[16] M L Puterman. *Markov decision processes : Discrete stochastic dynamic programming*. John Wiley and Sons, New York, NY, USA, 1994.

[17] H E Romeijn, D Sharma, and R L Smith. Extreme point solutions for infinite network flow problems. *Networks*, 48(4):209–222, 2006.

[18] H E Romeijn and R L Smith. Shadow prices in infinite dimensional linear programming. *Mathematics of Operations Research*, 23(1):239–256, 1998.

[19] H E Romeijn, R L Smith, and J C Bean. Duality in infinite dimensional linear programming. *Mathematical Programming*, 53(1-3):79–97, 1992.

[20] S M Ross. *Introduction to stochastic dynamic programming*. Academic Press, New York, NY, USA, 1983.

[21] T C Sharkey and H E Romeijn. A simplex algorithm for minimum cost network flow problems in infinite networks. *Networks*, 52(1):14–31, 2008.