On the Complexity of Inverse Mixed Integer Linear Optimization

AYKUT BULUT AND TED K. RALPHS
Department of Industrial and Systems Engineering, Lehigh University, USA

COR@L Technical Report 15T-001-R1
On the Complexity of Inverse Mixed Integer Linear Optimization

AYKUT BULUT*† AND TEO K. RALPHS††

1Department of Industrial and Systems Engineering, Lehigh University, USA

Abstract

Inverse optimization is the problem of determining the values of missing input parameters that are closest to given estimates and that will make a given target solution optimal. This study is concerned with inverse mixed integer linear optimization problems (MILPs) in which the missing parameters are objective function coefficients. This class generalizes the class studied by Ahuja and Orlin [2001], who showed that inverse linear optimization problems can be solved in polynomial time under mild conditions. We extend their result to the discrete case and show that the decision version of the inverse MILP is \textsc{coNP}-complete, while the optimal value verification problem is \textsc{D^p}-complete. We derive a cutting plane algorithm for solving inverse MILPs and show that there is a close relationship between the inverse problem and the well-known separation problem in both a complexity and an algorithmic sense.

Keywords: Inverse optimization, mixed integer linear optimization, computational complexity, polynomial hierarchy

1 Introduction

Optimization problems arise in many fields and the literature abounds with techniques for solving various classes of such problems. In general the goal of optimization is to determine the member of a given feasible set (the solution) that minimizes the value of a given objective function. The set is typically described as the points in a vector space satisfying a given set of equations, inequalities, and disjunctions (the last are usually in the form of a requirement that the value of a certain element of the solution take on an integral value).

An inverse optimization problem, in contrast, is a related problem in which the description of the original optimization problem, which we refer to as the forward problem, is not complete (some parameters are missing or cannot be observed), but a full or partial solution can be observed. The goal is to determine values for the missing parameters for which the given solution would be optimal. Estimates for the missing parameters may be given, in which case the goal is to produce a set of parameters that is as “close” to the given estimate as possible.

Formally, the optimization problem of primary interest in this paper is the mixed integer linear optimization problem (MILP)

$$\min_{x \in P} d^\top x$$ (1)
where \( d \in \mathbb{R}^n \) and 
\[
\mathcal{P} = \{ x \in \mathbb{R}^n \mid Ax = b, x \geq 0 \} \cap (\mathbb{Z}^r \times \mathbb{R}^{n-r}).
\]
for \( A \in \mathbb{R}^{m \times n}, \ b \in \mathbb{R}^m \), where the first \( r \) components of the solution \( x \) are restricted to integer values.

One can define a number of different associated inverse problems, depending on what parts of the description \((A, b, d)\) are unknown. Here, we study the case in which the objective function \( d \) is unknown. We are given \( A, b, c, \) and \( x^0 \in \mathbb{R}^n \). To formalize our statement of the problem, we first present several related mathematical optimization problems whose relationship to the inverse will be made clear. First, consider the following candidate for formulating the inverse problem,
\[
\min \|c - d\| \\
\text{s.t.} \quad d^\top x^0 \leq d^\top x \quad \forall x \in \mathcal{P}
\]
where \( \| \cdot \| \) can be any norm in general. In the above, \( d \) is the unspecified vector to be determined, while \( c \in \mathbb{R}^n \) is the estimate or target value. As we will see, this is not precisely the inverse problem we have described, but roughly speaking, the constraints require that \( d \) be an objective vector for which \( x^0 \) is optimal.

Note that if we fix \( d \) and consider \( x^0 \) to be the unknown quantity in (2), while replacing the objective function with that of the forward problem, we get a formulation of the forward problem itself.

Problem (2) can be re-formulated as a conic problem. For this we define two conic sets, \( K(y) \) and \( D \), as
\[
K(y) = \{ \alpha d \in \mathbb{R}^n : \|c - d\| \leq y, \alpha > 0, \alpha \in \mathbb{R} \}
\]
\[
D = \{ d \in \mathbb{R}^n : d^\top (x^0 - x) \leq 0 \ \forall x \in \mathcal{P} \}.
\]
In terms of these two sets, (2) can be reformulated as
\[
\min_{d \in K(y) \cap D} y.
\]
The set \( D \) can be interpreted either as the set of objective function vectors that prefer \( x^0 \) over all points in \( \mathcal{P} \) or, alternatively, as the set of all hyperplanes containing \( x^0 \) and that define inequalities valid for \( \mathcal{P} \). The latter interpretation leads to a third formulation in terms of the so-called 1-polar. Assuming that \( \text{conv}(\mathcal{P}) \) is a polytope, we define the 1-polar as
\[
\mathcal{P}^1 = \{ \pi \in \mathbb{R}^n \mid \pi^\top x \leq 1, \forall x \in \mathcal{P} \}.
\]
When \( \text{conv}(\mathcal{P}) \) is full-dimensional and 0 is in \( \text{conv}(\mathcal{P}) \) (this latter requirement is without loss of generality by translating \( \text{conv}(\mathcal{P}) \)), the 1-polar is the (normalized) set of all inequalities valid for \( \text{conv}(\mathcal{P}) \) (see [Schrijver, 1986] for a more formal definition). Under these assumptions, we then have this second re-formulation of (2).
\[
\min \|c - d\| \\
\text{s.t.} \quad \pi \in \mathcal{P}^1 \\
\pi^\top x^0 \geq 1 \\
d = \alpha \pi \\
\alpha \in \mathbb{R}_-
\]
In (4), the constraint \( d = \alpha \pi \) allows \( d \) to be scaled in order to improve the objective function value. The scale factor \( \alpha \) must be negative to reverse the sense of the inequality, since we are minimizing. We might also require \( \|c\| = 1 \) or normalize in some other way to avoid this scaling. The constraint \( \pi^\top x^0 \leq 1 \) ensures that \( d \) is feasible to (2). Note that in (2), the quantity \( d^\top x^0 \) appears explicitly as the right-side of the valid inequality generated, whereas in (4), it is not necessary to require \( \pi^\top x^0 = 1 \) in order to ensure that \( d^\top x \geq d^\top x^0 \) is valid, which is our requirement. Observe also that relaxing the constraint \( \pi^\top x^0 \leq 1 \) yields a problem something like the classical separation problem, but with a different objective function. We revisit this idea in Section 3.
We have so far avoided an important point and that is to be explicit about what assumptions we make about the point \( x^0 \). On the one hand, the problem, as informally stated, can only have a solution if \( x^0 \in \mathcal{P} \), since otherwise, \( x^0 \) cannot be optimal for any objective function. On the other hand, the formulations above can be interpreted whether or not \( x^0 \in \mathcal{P} \). As a practical matter, this subtle point is not very important, since membership in \( \mathcal{P} \) can be verified in a pre-processing step if this is important in a given setting. However, in the context of formal complexity analysis, this point is important and we will return to it. For now, we do not assume \( x^0 \in \mathcal{P} \), in which case \( d \) can be more accurately interpreted as specifying a valid inequality which is satisfied at equality by \( x^0 \).

In order to paint a complete picture, there is one other pathological case to be considered and that is when \( x^0 \) is in the interior of \( \text{conv}(\mathcal{P}) \). In this case, the optimal value of (2) is \( \|c\| \). For the objective value be anything other than \( \|c\| \), \( x^0 \) must either be on the boundary of \( \text{conv}(\mathcal{P}) \) or outside of \( \text{conv}(\mathcal{P}) \). Hence, (2) implicitly solves the decision problem of whether \( x^0 \) is in the interior of \( \text{conv}(\mathcal{P}) \). This decision problem is in turn equivalent to determining the feasibility of (2) (i.e., resolving the question of whether there exists an inequality valid for \( \mathcal{P} \) that is satisfied at equality by \( x^0 \)). It can thus further be seen as a sort of complement of the membership problem.

Figure 1 demonstrates the inverse MILP geometrically. \( \mathcal{P} \) is a pure integer set indicated by the black dots. The vector \( c \) is \((0, -2)\) and \( x^0 \) is \((3, 1)\). The convex hull of \( \mathcal{P} \) and the cone \( \mathcal{D} \) (translated to \( x^0 \)) are shaded. The ellipsoids show the sets of points with a fixed distance to \( x^0 + c \) for some given norm. The optimal objective function in this example is vector \( d^* \), and the point indicated in the figure is \( x^0 + d^* \).

In this study, we address the computational complexity of (2). As written, this is a semi-infinite program, but it is easy to see that we can replace the infinite set of constraints with a finite set corresponding to the extreme points of \( \text{conv}(\mathcal{P}) \). This still leaves us with a potentially non-linear objective function. We show in Section 3 that for the \( \ell_\infty \) and \( \ell_1 \) norms, this problem can in be expressed as a standard linear optimization problem (LP), albeit one with an exponential number of constraints. The reformulation can be readily solved in practice using a standard cutting plane approach. On the other hand, we show in Section 4 that the formal complexity does not depend on the norm.

In what follows, we discuss these problems in the traditional framework of computational complexity theory. The classical NP-completeness theory of Garey and Johnson [1979] addresses decision problems and as such, it will be convenient to refer in the sequel to several decision problem associated with (1). The most commonly associated decision version of (1) is a feasibility problem involving an extra scalar parameter \( \gamma \), as follows.
Definition 1 **MILP decision problem (MILPD):** Given \( \gamma \in \mathbb{Q}, \ d \in \mathbb{R}^n \), and an MILP with feasible region \( \mathcal{P} \), does there exist \( x \in \mathcal{P} \) such that \( d^T x \leq \gamma \)?

It is well-known that this problem is in the complexity class \( \text{NP} \)-complete and the optimal solution of problem (1) can be determined using bisection search in a polynomial number of calls to an oracle for MILPD. In the terminology of complexity theory, the *input* to this decision problem is the quintuplet \((A, b, d, r, \gamma)\) and the set of such inputs that yields the answer YES is the *language* recognized by an algorithm for solving this problem (formally specified as a Turing machine).

On the other hand, the problem of determining whether a given \( \gamma \) is a lower bound on the value of an optimal solution is also a decision problem.

Definition 2 **MILP lower-bounding problem (MILPL):** Given \( \gamma \in \mathbb{Q}, \ d \in \mathbb{R}^n \), and an MILP with feasible region \( \mathcal{P} \), is \( \min_{x \in \mathcal{P}} d^T x \geq \gamma \)?

This decision problem is in the class \( \text{coNP} \), i.e., there is a short certificate for the negative answer. When answer is NO then a feasible solution in \( \mathcal{P} \) with an objective value strictly less than \( \gamma \) is a short certificate. Furthermore, we can define a third decision problem, which is that of determining whether the optimal solution value is *exactly* \( \gamma \).

Definition 3 **MILP optimal value verification problem (MILPO):** Given \( \gamma \in \mathbb{Q}, \ d \in \mathbb{R}^n \), and an MILP with feasible region \( \mathcal{P} \), is \( \min_{x \in \mathcal{P}} d^T x = \gamma \)?

This problem is in the class \( \text{D}^P \) of problems defined by Papadimitriou and Yannakakis [1982] as those for which the language to be recognized is the intersection of languages recognized by two algorithms defined over the same set of inputs, one of which is known to be in the class \( \text{NP} \) and one of which is known to be in the class \( \text{coNP} \). Papadimitriou and Yannakakis [1982] showed that MILPO is complete (and 1 is hard) for the class \( \text{D}^P \).

In the remainder of the paper, we provide a literature review, formulate the inverse MILP, and discuss its computational complexity. We refer interested readers to Ahuja and Orlin [2001] for the history and development of inverse problems.

2 Literature Review

In this section, we first give a brief survey of the inverse optimization literature and then also briefly review relevant concepts from the literature on computational complexity.

**Inverse Problems.** There are a range of different flavors of inverse optimization problem in the literature. The inverse problem we investigate is to determine objective function coefficient that make a given solution optimal, but other flavors of inverse optimization include constructing a missing part of either the coefficient matrix or the right-hand side that makes a given solution optimal. Heuberger [2004] gives a detailed survey about inverse combinatorial optimization. He examines different types of inverse problems, including types for which the inverse problem seeks parameters other than objective function coefficients. He surveys solution procedures for specific combinatorial problems. He provides a classification of the inverse problems that are common in the literature. According to his classification, the inverse problem we study in this paper is *unconstrained*, *single feasible object*, and *unit weight norm* inverse problem. All the forward problems he considers are linear and some of them include discrete variables. We thus investigate the same class of problems, but only for the case of missing objective function and for one given solution. Our results can be easily extended to the case of multiple given solutions.

Cai et al. [1999] examines the inverse center location problem. This problem tries to retrieve the distances between nodes from a given optimal solution. They show that even though the center location problem is
polynomially solvable, its inverse is NP-hard. They reach this result by presenting a polynomial transformation of the satisfiability problem to the decision version of the inverse center location problem. The inverse problem they consider retrieves the coefficient matrix, which indicates that constructing coefficient matrix is harder than the forward version of the problem.

Huang [2005] examines the inverse knapsack problem and inverse integer optimization problems. He presents a pseudo-polynomial algorithm for the inverse knapsack problem. He also shows that inverse integer optimization with a fixed number of constraints is pseudo-polynomial. He reaches this conclusion by transforming the inverse problem to a shortest path problem on a directed graph. When the number of constraints are fixed, this results a pseudo-polynomial algorithm for inverse integer optimization.

Schaefer [2009] studies general inverse integer optimization problems. Using super-additive duality, he derives a polyhedral description of the set of all feasible objective functions. This description has only continuous variables but an exponential number of constraints. He then proposes solution methods that use this polyhedral description.

The case when the feasible set is an explicitly described polyhedron is well-studied by Ahuja and Orlin [2001]. In their study, they analyze the shortest path, assignment, minimum cut, and minimum cost flow problems under the $l_1$ and $l_\infty$ norms in detail. They also conclude that inverse optimization problem is polynomially solvable when the forward problem is polynomially solvable. This study aims to generalize the result of Ahuja and Orlin [2001] to the case when the forward problem is not necessarily polynomially solvable.

**Complexity Theory.** We have already mentioned the well-known NP-completeness theory of Garey and Johnson [1979], but for the results we present here, we need to refer to some additional complexity classes. In their contribution to this theory, Papadimitriou and Yannakakis [1982] define the class $\text{D}^p$ to be the class of languages that are the intersection of two languages, the first of which is in $\text{NP}$ and the second of which is in $\text{coNP}$. $\text{D}^p$ is a broader class that includes $\text{NP}$ and $\text{coNP}$. They give examples of problems in $\text{D}^p$. Moreover they prove completeness of some of these problems for $\text{D}^p$. One of the complete problems for $\text{D}^p$ is the optimal value problem. It is the problem of deciding whether a given value is the exact optimal one for a given MILP. We will show that optimal value problem for inverse optimization over MILP feasible sets is also $\text{D}^p$-complete.

$\Delta^p_2$ is the class of decision problems that can be solved in polynomial time given an $\text{NP}$ oracle. This class is defined as one member of the so-called *polynomial-time hierarchy* in the seminal work of Stockmeyer [1976]. It is a broader class that includes $\text{NP}$, $\text{coNP}$ and $\text{D}^p$. Every problem in $\text{NP}$, $\text{coNP}$ or $\text{D}^p$ is also in $\Delta^p_2$.

Figure 2 illustrates class $\Delta^p_2$ relative to $\text{D}^p$, $\text{NP}$, $\text{coNP}$ and $\text{P}$, assuming $\text{P} \neq \text{NP}$. If $\text{P} = \text{NP}$, we conclude that all classes are equivalent, i.e., $\Delta^p_2 = \text{D}^p = \text{NP} = \text{coNP} = \text{P}$. This theoretical possibility is known as the collapse of polynomial hierarchy to its first level, Papadimitriou [1994], but it is thought to be highly unlikely.

### 3 An Algorithmic Framework for Inverse MILP

We now show how to formulate (2) explicitly for two common norms using standard techniques for linearization. The objective function of an inverse MILP under the $l_1$ norm can be linearized by the introduction of
variable vector $\theta$, and associated constraints as follows,

$$
\begin{align*}
    z^1_{IP} &= \min_y \quad & s.t. \quad y = \sum_{i=1}^{n} \theta_i \\
    & & c_i - d_i \leq \theta_i & \forall i \in \{1, 2, \ldots, n\} \\
    & & d_i - c_i \leq \theta_i & \forall i \in \{1, 2, \ldots, n\} \\
    & & d^T x^0 \leq d^T x & \forall x \in P.
\end{align*}
$$

(5)

The objective function of inverse MILP under $l_\infty$ norm can be linearized by the introduction of variable $y$ and two sets of constraint sets as follows,

$$
\begin{align*}
    z^\infty_{IP} &= \min_y \quad & s.t. \quad c_i - d_i \leq y & \forall i \in \{1, 2, \ldots, n\} \\
    & & d_i - c_i \leq y & \forall i \in \{1, 2, \ldots, n\} \\
    & & d^T x^0 \leq d^T x & \forall x \in P.
\end{align*}
$$

(6)

This formulation is a continuous problem, but is a semi-infinite program when written in form above, since $P$ may contain infinitely many points.

It is enough to write this constraint set over extreme points and rays of $\text{conv}(P)$. We know that $\text{conv}(P)$ is a polyhedron and hence has finitely many extreme points and rays. Nevertheless, the number of extreme points and rays may be still very large and we cannot hope to write the inverse problem by enumerating all extreme points and rays of $\text{conv}(P)$ a priori. A better approach is using a separation–optimization procedure and generating these inequalities on the fly as we need them. This is a natural application of separation–optimization procedure of Grötschel et al. [1993]. A similar algorithm was described by Wang [2009], although this work was unknown to us during the development of these results. The following section describes this algorithm for inverse problems.

We propose a cutting plane algorithm for solving inverse MILPs under $l_\infty$ norm. We describe this method with respect to $l_\infty$ norm model (6), but it works similarly for the $l_1$ norm model (5). First, we define two parametric problems named $P_k$ and $InvP_k$ as follows,

$$
\min_{x \in P} d^{k\top} x
$$

($P_k$)
\[
\begin{align*}
\min & \quad y \\
\text{s.t.} & \quad c_i - d_i \leq y \quad \forall i \in \{1, 2, \ldots, n\} \\
& \quad d_i - c_i \leq y \quad \forall i \in \{1, 2, \ldots, n\} \\
& \quad d^\top x^0 \leq d^\top x \quad \forall x \in \mathcal{P}.
\end{align*}
\]

where \( \mathcal{E}^k \) is the set of solutions found by solving \( P_1, \ldots, P_{k-1} \). Note that \( P_k \) is the forward MILP problem with \( d^k \) as the objective coefficient and is precisely the problem of separating \( d^k \) from the feasible region of (2). \( \text{InvP}_k \) is the relaxation of MILP (6) considering only valid inequalities that correspond to forward problem solutions found so far.

The overall procedure is given in Algorithm 1. In this algorithm, we solve an instance of the forward problem in each iteration in order to generate a cut. The algorithm stops when the current \( d^k \) is feasible.

**Algorithm 1** Cutting plane for inverse MILP under \( l_\infty \) norm

\[
k \leftarrow 0, \mathcal{E}^1 \leftarrow \emptyset.
\]

**do**

\[
k \leftarrow k + 1.
\]

Solve \( \text{InvP}_k \), \( d^k \leftarrow d^* \).

Solve \( P_k \).

if \( P_k \) unbounded then

\[y^* \leftarrow \|c\|_\infty, \ d^* \leftarrow 0, \ \text{STOP}.
\]

else

\[x^k \leftarrow x^*.
\]

end if

\[\mathcal{E}^{k+1} \leftarrow \mathcal{E}^k \cup \{x^k\}.
\]

**while** \( d^k \top (x^0 - x^k) > 0 \)

\[y^* \leftarrow \|c - d^k\|_\infty, \ d^* \leftarrow d^k, \ \text{STOP}.
\]

When \( P_k \) is unbounded, then \( d = 0 \) is an optimal solution, since this shows that only \( d = 0 \) satisfies \( d^\top (x^0 - x) \leq 0 \) for all \( x \) in \( \mathcal{P} \).

Before illustrating with a small example, we would like to again point out the close relationship of the inverse problem and the separation problem. Algorithm 1 can be easily modified to solve the generic separation problem for \( \text{conv}(\mathcal{P}) \) by interpreting \( x^0 \) as the point to be separated and replacing the objective function (and associated auxiliary constraints) of \( \text{InvP}_k \) with one measuring the degree of violation of \( x^0 \). The generated valid inequalities are sometimes called Fenchel cuts, Boyd [1994].

A Small Example: Let \( c = (-2, 1) \), \( x^0 = (0, 3) \) and \( \mathcal{P} \) given as in Figure 3 where both \( x_1 \) and \( x_2 \) are integer and convex hull of \( \mathcal{P} \) is given. \( k, d^k \) and \( x^k \) values through iterations are given in Table 1.

| Table 1: \( k \), \( d^k \), \( x^k \) and \( \mathcal{E}^k \) values through iterations |
|---|---|---|---|
| \( k \) | \( \mathcal{E}^k \) | \( d^k \) | \( x^k \) | \( \|d^k - c\|_\infty \) |
| initialization | 1 | \( \emptyset \) | \((-2, 1)\) | \( (3, 0) \) | 0 |
| iteration 1 | 2 | \( \{(3, 0)\} \) | \((-0.5, -0.5)\) | \( (3, 1) \) | 1.5 |
| iteration 2 | 3 | \( \{(3, 0), (3, 1)\} \) | \((-0.4, -0.6)\) | \( (3, 1) \) | 1.6 |

Inverse MILP optimal value is \( y^* = \|c - d^3\|_\infty = 1.6 \). Inverse MILP optimal solution is \( d^3 = (0.4, 0.6) \).
4 Complexity of Inverse MILP

Ahuja and Orlin [2001] show that the inverse problem can be solved in polynomial time when the forward problem is polynomially solvable.

**Theorem 1 (Ahuja and Orlin [2001])** If a forward problem is polynomially solvable for each linear cost function, then the corresponding inverse problems under $l_1$ and $l_\infty$ norms are polynomially solvable.

They use the well-known result of Grötschel et al. [1993] to conclude that inverse LP, in particular, is polynomially solvable. Note that this result already indicates that if a given MILP is polynomially solvable, then the associated inverse problem is also polynomially solvable. To obtain a formal complexity result for the more general case, we first consider the decision version of the inverse problem. The decision version is derived in a fashion similar to that of MILPD. It asks whether a solution with objective value less than some given threshold exists.

**Definition 4 Inverse MILP decision problem (INVD):** Given $\gamma \in \mathbb{Q}$, $c \in \mathbb{R}^n$, $x^0 \in \mathbb{R}^n$, and polyhedron $P \subseteq \mathbb{R}^n$, is the set $K(\gamma) \cap D$ non-empty?

The result of Grötschel et al. [1993] bounds the running time for optimizing a linear objective function over an implicitly defined polyhedron in terms of call to a separation oracle. Their result can be stated as follows.

**Theorem 2 (Grötschel et al. [1993])** Given an oracle for the separation problem, the optimization problem over a given polyhedron with linear objective can be solved in time polynomial in $\varphi$, $n$ and the encoding length of objective coefficient vector, where $\varphi$ is the facet complexity of the given polyhedron.

A polyhedron has facet-complexity at most $\varphi$ if there exists a rational system of inequalities describing the polyhedron in which the encoding length of each inequality is at most $\varphi$. The facet complexity thus measures the complexity of a polyhedron independent of its representation. Theorem 2 indicates that, given an oracle for inverse MILP separation, the inverse MILP optimization problem can be solved in time polynomial in $\varphi$ and $n$, where the feasible set of (6) has facet-complexity at most $\varphi$, since the objective function of 6 has an encoding length polynomial in $n$.

To find a bound on $\varphi$, consider the third set of constraints of the formulation (6). The encoding length of the first two sets of constraints depends on the maximum encoding length of $c_i$, $i \in 1, \ldots, n$. The encoding length of the third set of constraints depends on the encoding length of $x^0$ and the largest encoding length of
any extreme point of the convex hull of $\mathcal{P}$. This latter quantity is known as the vertex complexity of $\mathcal{P}$ and is a related measure of the complexity of a polyhedron that is bounded by a polynomial function of the facet complexity. Thus, we can say that the running time of the separation–optimization algorithm is polynomial in the encoding length of $c_i$ for $i = 1, \ldots, n, x^0$ and the vertex complexity of the convex hull of $\mathcal{P}$. Note that in the case of binary integer optimization problems, the vertex complexity of $\text{conv}(\mathcal{P})$ is always polynomial in $n$.

The above conclusions can be interpreted as stating that the inverse MILP separation problem is equivalent to the MILP optimization problem, but it is important to note that this equivalence is only a polynomial equivalence, not a complexity-wise equivalence. The MILP optimization problem can be solved in polynomial time, given an oracle for the MILP decision problem. Similarly, we conclude that the inverse MILP optimization problem can be solved in polynomial time, given an oracle for the MILP decision problem, which we know to be $\text{NP}$–complete. The following theorem summarizes this result.

**Theorem 3** The inverse MILP optimization problem under the $l_\infty/l_1$ norms is solvable in time polynomial in $\varphi$ and $n$, given an oracle for the MILP decision problem.

This theorem hints at the complexity of inverse optimization problem. We now know that Algorithm 1 solves inverse MILP in polynomial time, given an $\text{NP}$ oracle. This algorithm can be used to solve the decision version. In complexity terms, this shows that the inverse problem is in $\Delta^P_2$. The following is the restatement of Theorem 3 in complexity terms.

**Theorem 4** $\text{INVD}$ under $l_1$ and $l_\infty$ norms is in $\Delta^P_2$.

The next natural question that comes to mind is whether $\text{INVD}$ is complete for this class. Somewhat surprisingly (though not in hindsight), the answer is no. This indicates that GLS result does not yield the tightest complexity class. The first main result of this paper is the following theorem that shows $\text{INVD}$ with an arbitrary norm ($l_1$, $l_\infty$, or any other $p$-norm) is in $\text{coNP}$.

**Theorem 5** $\text{INVD}$ is in $\text{coNP}$.

**Proof** We show existence of a short certificate when the answer to $\text{INVD}$ problem is NO. Note that when answer is NO then $\gamma < \|c\|$, since $d = 0$ is a valid solution otherwise. Furthermore, when $\gamma = 0$ the problem reduces to MILP (is $c^\top x^0$ a lower bound for minimization along $c$ over $\mathcal{P}$), which is already known to be in $\text{coNP}$. Therefore it is enough to consider the case where $0 < \gamma < \|c\|$.

When the answer to $\text{INVD}$ is NO, then, for each $d$ in $K(\gamma)$, there exists an $x \in \mathcal{P}$ such that $d^\top (x - x^0) < 0$. Hence, the NO answer can be validated by enumeration over $\mathcal{P}$ in principle. What we will show is that we do not need to check the inequality for all $x \in \mathcal{P}$, but only for a subset of polynomial size. We define the set of all such $x$ values as follows,

$$\mathcal{X}(\gamma) = \{x \in \mathcal{P} \mid \exists d \in K(\gamma) \text{ s.t. } d^\top (x - x^0) < 0\}.$$  

$\mathcal{X}(\gamma)$ is the set of points in $\mathcal{P}$ that are better than $x^0$ for at least one direction $d$ in $K(\gamma)$. Note that set $\mathcal{X}(\gamma)$ is not an empty since answer to problem is NO. Moreover since $\mathcal{X}(\gamma)$ is a subset of $\mathcal{P}$ it is a discrete set. We define another set $K^*(\gamma)$ as follows,

$$K^*(\gamma) = \{x \in \mathbb{R}^n \mid d^\top (x - x^0) \leq 0 \forall d \in K(\gamma)\}.$$  

$K^*(\gamma)$ is the set of points that are better than $x^0$ for all the directions in $K(\gamma)$. Note that $K^*(\gamma)$ is nothing but the dual cone of $K(\gamma)$ moved along $x^0$. Both $K(\gamma)$ and $K^*(\gamma)$ are full dimensional pointed cones, since $0 < \gamma < \|c\|$.

Cone $K^*(\gamma)$, set $\mathcal{X}(\gamma)$ and set $\mathcal{P}$ can be considered to be in the primal space, i.e., space of primal solution values and cones $D$ and $K(\gamma)$ can be considered in the dual space, i.e., space of directions.
We claim the following holds for the negative answer and continue to construct our short certificate. We prove our claim after we construct the short certificate.

**Claim 1** \( \text{conv}(\mathcal{X}(\gamma)) \cap \text{int}(K^*(\gamma)) \neq \emptyset \).

Let \( \bar{x} \in \text{conv}(\mathcal{X}(\gamma)) \cap \text{int}(K^*(\gamma)) \), then a subset of \( \mathcal{X}(\gamma) \) that can give \( \bar{x} \) as a convex combination is a certificate. Moreover, it is a short certificate since we need \( n+1 \) elements from \( \mathcal{X}(\gamma) \) at most. Let \( \{x^1, x^2, \ldots, x^k\} \subset \mathcal{X}(\gamma) \) be such a subset and \( \{\lambda_1, \ldots, \lambda_k\} \) be the corresponding values such that \( \bar{x} = \sum_{i=1}^{k} \lambda_i x^i, \sum_{i=1}^{k} \lambda_i = 1 \) and \( \lambda_i \geq 0 \) for \( i = 1, \ldots, k \) and \( k \leq n + 1 \). Next we show how sets \( \{x^1, x^2, \ldots, x^k\} \) and \( \{\lambda_1, \ldots, \lambda_k\} \) can be used to validate the NO answer.

For any given \( d \in D, \bar{x} \) being and element of \( \text{int}(K^*(\gamma)) \) gives us the following,

\[
d^T(\bar{x} - x^0) < 0.\]

We can write \( \bar{x} \) as a convex combination of \( x^i \) values. When we replace \( \bar{x} \) using this we get the following inequality,

\[
d^T \left( \sum_{i=1}^{k} \lambda_i x^i - x^0 \right) < 0.\]

We can manipulate this inequality to get the following inequalities,

\[
d^T \left( \sum_{i=1}^{k} \lambda_i x^i - \sum_{i=1}^{k} \lambda_i x^0 \right) < 0,\]

\[
\sum_{i=1}^{k} \lambda_i d^T(x^i - x^0) < 0.\]

Then, there exists at least one index \( j \in \{1, \ldots, k\} \) such that \( d^T(x^j - x^0) < 0 \). Then \( x^j \in P \) and being a better solution for direction \( d \) means \( x^0 \) can not be optimal. Direction \( d \) is arbitrary, meaning this result holds for all \( d \) in set \( K(\gamma) \). Using sets \( \{x^1, \ldots, x^k\} \) and \( \{\lambda_1, \ldots, \lambda_k\} \) we validated the NO answer for an arbitrary \( d \) in set \( K(\gamma) \). This shows that sets \( \{x^1, \ldots, x^k\} \) and \( \{\lambda_1, \ldots, \lambda_k\} \) together is a short certificate for the inverse problem defined. \( \square \)

**Proof of Claim** Assume \( \text{conv}(\mathcal{X}(\gamma)) \cap \text{int}(K^*(\gamma)) = \emptyset \) for a contradiction. \( \text{conv}(\mathcal{X}(\gamma)) \) and \( K^*(\gamma) \) are both convex sets. Then there exists a hyperplane that separates these two sets. Let \( \{x \in \mathbb{R}^n : a^T x = \alpha, \alpha \in \mathbb{R}, a \in \mathbb{R}^n \} \) be such a hyperplane that separates \( \text{conv}(\mathcal{X}(\gamma)) \) and \( K^*(\gamma) \) as follows,

\[
a^T x \geq \alpha \ \forall x \in \text{conv}(\mathcal{X}(\gamma)),
\]

\[
a^T x \leq \alpha \ \forall x \in K^*(\gamma).
\]

Then we can write the following inequality,

\[
\min_{x \in \text{conv}(\mathcal{X}(\gamma))} a^T x \geq \max_{x \in K^*(\gamma)} a^T x. \tag{7}
\]

Note that problem on the right-hand side is unbounded when \( a \) is not in \( K(\gamma) \). Then we can conclude that \( a \in K(\gamma) \) for a valid separating hyperplane. This indicates that \( x^0 \) maximizes \( a^T x \) over cone \( K^*(\gamma) \). Then we have the following inequality,
Figure 4: A small example demonstrates $\text{conv}(P), \mathcal{K}(\gamma), \mathcal{K}^*(\gamma), \text{conv}(\mathcal{X}(\gamma))$

\[
\min_{x \in \text{conv}(\mathcal{X}(\gamma))} a^\top x \geq a^\top x^0.
\]

Since direction $a$ is in $\mathcal{K}(\gamma)$ and answer to our problem is NO, there exists an $\bar{x}$ in $\mathcal{X}(\gamma)$ such that $a^\top (\bar{x} - x^0) < 0$. Since $\bar{x}$ is a feasible solution for the optimization problem over $\text{conv}(\mathcal{X}(\gamma))$ we have the following inequality,

\[
a^\top \bar{x} \geq \min_{x \in \text{conv}(\mathcal{X}(\gamma))} a^\top x \geq a^\top x^0.
\]

Using $a^\top (\bar{x} - x^0) > 0$, we can rewrite the inequality as follows,

\[
a^\top x^0 < a^\top \bar{x} \geq \min_{x \in \text{conv}(\mathcal{X}(\gamma))} a^\top x \geq a^\top x^0.
\]

Which is a contradiction. This indicates that the contradiction assumption, existence of a separating hyperplane, is wrong. This proves that $\text{conv}(\mathcal{X}) \cap \text{int}(\mathcal{K}^*) \neq \emptyset$.

Figure 4 shows sets $\text{conv}(P), \mathcal{K}(\gamma), \mathcal{K}^*(\gamma)$ and $\text{conv}(\mathcal{X}(\gamma))$ for the example introduced in Section 3, where $c$ and $x^0$ are redefined. In this case $c = (-1, -2)$, $x^0 = (2, 1)$ and $\gamma = 1$.

All this theory indicates that complexity of INVD is same as MILPL. The difference is the certificate for lower bound problem is just a feasible point where the certificate for INVD problem is at most $n + 1$ points with corresponding weights. Certificate for INVD problem is a little more complicated than certificate of MILPL.

**Theorem 6** \text{INVD is coNP–complete.}

**Proof** MILPL can be reduced to INVD. Let inputs of MILPL be $(c, \gamma, P)$ then MILPL can be decided by deciding INVD with inputs $(c^2 \leftarrow c, \gamma^2 \leftarrow 0, P^2 \leftarrow P, x^0 \leftarrow \frac{\gamma c}{\|c\|})$. INVD asks whether some $d$ in cone
\{ d \in \mathbb{R}^n | d^\top \left( \frac{\gamma c}{\|c\|^2} - x \right) \leq 0 \ \forall x \in \mathcal{P} \} \text{ satisfies } \|c - d\| \leq 0. \text{ Only } d \text{ that satisfies } \|c - d\| \leq 0 \text{ is } d = c. \text{ For answer to be positive } c \text{ must be in this cone. } c \text{ is in this cone if and only if }
\begin{align*}
c^\top \left( \frac{\gamma c}{\|c\|^2} - x \right) & \leq 0 \quad \forall x \in \mathcal{P}, \\
\gamma - c^\top x & \leq 0 \quad \forall x \in \mathcal{P}, \\
\gamma & \leq c^\top x \quad \forall x \in \mathcal{P},
\end{align*}

which means answer to MILPL is positive. This indicates answer to INVD is positive if and only if answer to MILPL is positive.

Figure 5 shows \( x^0 \) for various \( \gamma \) values. Answer for \( \gamma_1 \) is negative and for \( \gamma_2 \) and \( \gamma_3 \) is positive. Position of \( x^0 \) is just for presentation. For \( \gamma_1 \) case \( x^0 \) is displayed to be outside of \( \text{conv}(\mathcal{P}) \). This is just for display and the result is independent of \( x^0 \) being in \( \text{conv}(\mathcal{P}) \) or not. The answer is negative for both of the cases.

\[ \square \]

Lower bound problem for inverse MILP can be defined as follows,

**Definition 5** Inverse MILP lower-bounding problem (INVL): Given \( \gamma \in \mathbb{R}, c \in \mathbb{R}^n, x^0 \in \mathbb{R}^n \) and an MILP with feasible set \( \mathcal{P} \), is \( \min_{d \in \mathcal{K}(\gamma) \cap \mathbb{R}^n} y \geq \gamma ? \)

**Theorem 7** INVL problem is in \( \text{NP} \).

**Proof** We need to show existence of a short certificate that can validate YES answer. When answer is YES, optimal value of the inverse problem is greater than equal to \( \gamma \). We show existence of a short certificate that validates optimal value can not be less than \( \gamma \), i.e. no feasible direction \( d \) that optimizes \( x^0 \) and its distance to \( c \) is less than \( \gamma \). This is same as validating NO answer for INVD. The only difference is now the directions are strictly less than \( \gamma \)-distance to \( c \). Remember the claim we proved,

\( \text{conv}(\mathcal{X}(\gamma)) \cap \text{int}(\mathcal{K}^+(\gamma)) \neq \emptyset. \)

Note that \( \mathcal{K}^+(\gamma) \) is the set of points that are at least as good as \( x^0 \) for all directions \( d \) at most \( \gamma \)-distant to \( c \). \( \text{int}(\mathcal{K}^+(\gamma)) \) is the set of points that are strictly better than \( x^0 \) for all directions \( d \) that are strictly less than \( \gamma \)-distant to \( c \).

After this point the proof goes on same as proof of INVD being in \( \text{coNP} \). The short certificate is the same and it can be used to show that the optimal value of inverse problem can not be less than \( \gamma \).

**Theorem 8** INVL is \( \text{NP-complete} \).
Proof MILPD can be reduced to INVL. Let inputs of MILPD be \((c, \gamma, \mathcal{P})\) then MILPD can be resolved by deciding INVL with inputs \((c^2 \leftarrow c, \gamma^2 \leftarrow \delta, \mathcal{P}^2 \leftarrow \mathcal{P}, x^0 \leftarrow \frac{(\gamma + \epsilon)c}{\|c\|^2})\). \(\epsilon\) and \(\delta\) are small positive rationals computed from inputs of MILPD. \(\epsilon\) is such that when \(c^\top x > \gamma \forall x \in \mathcal{P}\) then \(c^\top x > \gamma + \epsilon \gamma \in \mathcal{P}\). Such an \(\epsilon\) exists since optimal value of MILP is bounded by a polynomial function of the vertex complexity of \(\mathcal{P}\).

We can compute this epsilon using the vertex complexity, encoding of \(\gamma\) and \(c\). When given \(\gamma\) is optimal for MILP then \(c\) does not optimize \(x^0\) since

\[
c^\top x^0 = c^\top \left(\frac{\gamma + \epsilon}{\|c\|^2}\right)c = \gamma + \epsilon.
\]

(8)

Optimal value of inverse problem is 0 when \(c^\top x > \gamma \forall x \in \mathcal{P}\). When optimal value of MILP is \(\gamma\) then inverse optimal value is not 0 but a lower bound \(\delta\) can be computed from vertex complexity of \(\mathcal{P}\) and encoding length of \(\gamma\) and \(\epsilon\). Any nonzero objective value of inverse problem is larger than this \(\delta\) value.

INVL asks whether \(\|c - d\| \geq \delta\) for all \(d \in \mathbb{R}^n\) \(d^\top \left(\frac{(\gamma + \epsilon)c}{\|c\|^2}\right) - x \forall x \in \mathcal{P}\).

Deciding INVL with described inputs resolves MILPD. If answer to INVL is positive than \(c\) does not optimize \(x^0\) over \(\mathcal{P}\). There exists \(\pi\) in \(\mathcal{P}\) such that

\[
c^\top \pi < c^\top x^0 = \gamma + \epsilon.
\]

This indicates \(c^\top \pi \leq \gamma\) by our design of \(\epsilon\). This means answer to MILPD is positive.

When answer to INVL is negative than optimal value of inverse problem is 0 by our design of \(\delta\). This indicates \(c\) optimizes \(x^0\),

\[
c^\top x^0 = \gamma + \epsilon < \gamma < c^\top x \forall x \in \mathcal{P}.
\]

This means answer to MILPD is negative. □

Figure 6 illustrates the case where \(c^\top x > \gamma\) for all \(x \in \mathcal{P}\). Note that for this case \(c^\top x^0 < \gamma\) for all \(x \in \mathcal{P}\) and answer to both of the problems is negative.
Figure 7 illustrates the case where optimal value of MILP is exactly $\gamma$. In this figure inverse optimal value is $\delta$ which is a small positive number. For this figure we can not compute the optimal value of inverse problem since we do not know optimal value of MILP but we can compute $\delta$ which is a lower bound for inverse problem. For this case answer to both of the problems is positive.

**Definition 6** Inverse MILP optimal value verification problem (INVO): Given $\gamma \in \mathbb{Q}$, $x^0 \in \mathbb{R}^n$, $c \in \mathbb{R}^n$, and an MILP with feasible region $\mathcal{P}$, is $\min_{d \in \mathcal{K}(y) \cap \mathcal{D}} y = \gamma$?

**Theorem 9** INVO problem is in class $\text{DP}$.

**Proof** Language of INVO can be written as an intersection of languages of INVD and INVL which are in coNP and NP respectively. □

Note that verifying exact optimal value of both inverse and forward problems are in the same complexity class. Deciding whether a given point is on the boundary of $\text{conv}(\mathcal{P})$ is also in this class since it is equivalent to verifying optimal value.

Note that asking if a given point is in $\text{conv}(\mathcal{P})$ (membership problem) is equivalent to MILPD problem. Similarly asking whether a given inequality is valid (validity problem) for $\text{conv}(\mathcal{P})$ is equivalent to MILPL. Note that validity problem itself is a membership problem over polar set of $\text{conv}(\mathcal{P})$.

### 5 Conclusion and Future Directions

In this paper, we formally defined various problems related to the inverse MILP problem in which we try to derive an objective function $d$ closest to a given estimate $c$ that make a given solution $x^0$ optimal over the feasible region $\mathcal{P}$ to an MILP. This problem can be seen as an optimization problem over the set of all inequalities valid for $\mathcal{P}$ and satisfied at equality by $x^0$. Alternatively, it can also be seen as optimization
over the 1-polar with some additional constraints. Both these characterization make the connection the separation problem associated with $P$ evident.

After defining the problem formally, we gave a cutting plane algorithm for solving it under the $l_1$ and $l_\infty$ norms and observed that the separation problem for the feasible region is equivalent to the original forward problem, enabling us to conclude by the framework of Grötschel et al. [1993] that the problem can be solved with a polynomial number of calls to an oracle for solving the forward problem.

This algorithm places the decision version of inverse MILP in the complexity class $\Delta_2^P$, but it is possible to prove a stronger result. The main contribution of this study is to show that this decision problem is complete for the class $\text{coNP}$, which is on the same level of the polynomial-time hierarchy of that of the forward problem. We proved the problem is in $\text{coNP}$ by giving a short certificate for the negative answer and then show it is complete for $\text{coNP}$ by reducing the MILP lower bound problem (MILPL) to inverse MILP decision problem. We also provide a reduction for the inverse lower bound problem. Finally, we show that the inverse optimal value verification problem to class $\text{D}^P$, which is precisely the same class containing the MILP optimal value verification problem.

Theorem 2 states that an optimization problem (over a convex set) can be solved in polynomial time given an oracle for the separation problem. Technically, this does not allow us to place the optimization and separation problems on precisely the same level of the polynomial hierarchy. It is likely that the GLS can be modified slightly in order to show that optimization and separation are indeed on the same level of the hierarchy. There are also some interesting open questions remaining to be explored with respect to complexity.

Finally, we have implemented the algorithm and a computationally oriented study is left as future work. Such a study would reveal the practical performance of the separation-optimization procedure and investigate the possible relationship between the number of iterations (oracle calls) and the polyhedral complexity (vertex/facet complexity), among other things. This may provide practical estimates for the number of iterations required to solve certain classes of problems.

References


