2011 50th IEEE Conference on Decision and Control and
European Control Conference (CDC-ECC)
Orlando, FL, USA, December 12-15, 2011

Inverse polynomial optimization

Jean B. Lasserre

***Abstract*— We consider the inverse optimization problem associated with the polynomial program $f^* = \min\{f(\mathbf{x}) : \mathbf{x} \in \mathbf{K}\}$ and a given current feasible solution $\mathbf{y} \in \mathbf{K}$. We provide a systematic numerical scheme to compute an inverse optimal solution. That is, we compute a polynomial $\tilde{f}$ (which may be of same degree as f if desired) with the following properties: (a) $\mathbf{y}$ is a global minimizer of $\tilde{f}$ on $\mathbf{K}$ with a Putinar's certificate with an *a priori* degree bound d fixed, and (b), $\tilde{f}$ minimizes $\|f - \tilde{f}\|_1$ over all polynomials with such properties. The size of the semidefinite program can be adapted to the computational capabilities available. Moreover, $\tilde{f}$ takes a simple *canonical* form, and computing $\tilde{f}$ reduces to solving a semidefinite program whose optimal value also provides a bound on how far is $f(\mathbf{y})$ from the unknown optimal value f^*. Some variations are also discussed.**

I. Introduction

Let $\mathbf{P}$ be the optimization problem $f^* = \min\{f(\mathbf{x}) : \mathbf{x} \in \mathbf{K}\}$, where

$$\mathbf{K} := \{\mathbf{x} \in \mathbb{R}^n : g_j(\mathbf{x}) \geq 0,\ j = 1, \ldots, m\}, \qquad (1)$$

for some polynomials $f, (g_j) \subset \mathbb{R}[\mathbf{x}]$.

Problem $\mathbf{P}$ is in general NP-hard and one is often satisfied with a local minimum which can be obtained by running some local minimization algorithm among those available in the literature. Typically in such algorithms, at a current iterate (i.e. some feasible solution $\mathbf{y} \in \mathbf{K}$), one checks whether some optimality conditions (e.g. the Karush-Kuhn-Tucker (KKT) conditions) are satisfied within some ϵ-tolerance. However, as already mentioned those conditions are only valid for a local minimum, and in fact, even only for a stationary point of the Lagrangian. Moreover, in many practical situations the criterion f to minimize is subject to modeling errors or is questionable. In such a situation, the practical meaning of a local (or global) minimum f^* (and local (or global) minimizer) also becomes questionable. It could well be that the current solution $\mathbf{y}$ is in fact a global minimizer of an optimization problem $\mathbf{P}'$ with same feasible set as $\mathbf{P}$ but with a different criterion $\tilde{f}$. Therefore, if $\tilde{f}$ is *close enough* to f, one may not be willing to spend an enormous computing time and effort to find the local (or global) minimum f^* because one might be already satisfied with the current iterate $\mathbf{y}$ as an optimal solution of $\mathbf{P}'$.

Inverse Optimization is precisely concerned with the above issue of determining a criterion $\tilde{f}$ as close to f as possible, and for which the current solution $\mathbf{y}$ is an optimal solution of $\mathbf{P}'$ with this new criterion $\tilde{f}$. Pioneering work in Control dates back to Freeman and Kokotovic [6] for optimal stabilization. Whereas it was known that every value function of an optimal stabilization problem is also a Lyapunov function for the closed-loop system, in [6] the authors show the converse, that is, every Lyapunov function for every stable closed-loop system is also a *value function* for a meaningful optimal stabilization problem. In optimization, pioneering works in this direction date back to Burton and Toint [3] for shortest path problems, and Ahuja and Orlin [2] for linear programs. For integer programming, Schaefer [11] characterizes the feasible set of cost vectors $d \in \mathbb{R}^n$ candidates for inverse optimality. It is the projection on $\mathbb{R}^n$ of a (lifted) convex polytope obtained from the super-additive dual of integer programs. In Heuberger [5] the interested reader will find a nice survey on inverse optimization for combinatorial optimization problems. More recently, for linear programs Ahmed and Guan [1] have considered the variant called *inverse optimal value* problem in which one is now interested in finding a new linear criterion for which the optimal value is the closest to a desired specified value. This problem is shown to be NP-hard in [1].

As the reader may immediately guess, in inverse optimization the difficulty lies in having some characterization of global optimality for a given current point $\mathbf{y} \in \mathbf{K}$ and some candidate criterion $\tilde{f}$. And even more, another difficulty is to have a *tractable* characterization for practical computation. This is why most of all the above cited works address linear programs or combinatorial optimization problems for which some characterization of global optimality is available and can be effectively translated for practical computation. For instance, the characterization of global optimality for integer programs described in Schaefer [11] is exponential in the problem size, which prevents from its use in practice.

Contribution. We investigate the inverse optimization problem for polynomial optimization problems $\mathbf{P}$ as in (1), i.e., in a rather general context which includes nonlinear and nonconvex optimization problems and in particular, 0/1 and mixed integer nonlinear programs. Fortunately, in such a context, Putinar's Positivstellensatz [10] provides us with a very powerful certificate of global optimality that can be adapted to the actual computational capabilities for a given problem size. More precisely, and assuming that $\mathbf{y} = 0$ (possibly after the change of variable $\mathbf{x}' = \mathbf{x} - \mathbf{y}$), in the methodology that we propose, one computes the coefficients of a polynomial $\tilde{f}_d \in \mathbb{R}[\mathbf{x}]$ of same degree df as f (or possibly larger degree if desired and/or possibly with some additional constraints), such that:

- $\tilde{f}(0) = f(0)$ and 0 is a global optimum of the related problem $\min_{\mathbf{x}}\{\tilde{f}_d(\mathbf{x}) : \mathbf{x} \in \mathbf{K}\}$, with a Putinar's certificate of optimality with degree bound d (to be explained later).
- $\tilde{f}_d$ minimizes $\|\tilde{f} - f\|_1$ (where $\|\cdot\|_1$ is the ℓ_1-norm of

J.B. Lasserre is with LAAS-CNRS and the Institute of Mathematics, University of Toulouse, France. lasserre@laas.fr

the coefficient vector) over all polynomials $\tilde{f}$ of degree df, having the previous property. Of course other choices for the norm (and not discussed here) are possible, e.g. the ℓ_2- or ℓ_∞-norm.

It turns out that the optimal value $\rho_d := \|\tilde{f}_d - f\|_1$ also measures how close is $f(0)$ to the global optimum f^* of $\mathbf{P}$, as we also obtain that $f^* \leq f(0) \leq f^* + \rho_d$.

In addition, we prove that when df is even, $\tilde{f}_d$ has a *canonical form*, namely

$$\tilde{f}_d \,=\, f + \mathbf{b}'\mathbf{x} + \sum_{i=1}^n (\lambda_i\, x_i^2 + \gamma_i\, x_i^{df}),$$

for some vector $\mathbf{b} \in \mathbb{R}^n$, and nonnegative vectors $\lambda, \gamma \in \mathbb{R}^n$, optimal solutions of a semidefinite program. This canonical form is another example of the "sparsity" property of the ℓ_1-norm in optimization, already observed in other contexts (e.g., in some compressed sensing applications).

Importantly, to compute $\tilde{f}_d$ one has to solve a semidefinite program of size parametrized by d, where d is chosen so that the size of semidefinite program associated with Putinar's certificate (with degree bound d) is compatible with the capabilities of current semidefinite solvers available. Of course, even if d is relatively small, one is still restricted to problems of relatively modest size. However, if problem $\mathbf{P}$ exhibits some sparsity pattern (as is often the case in large scale problems) then one can use the specific Positivstellensatz developed for such problems [8], as a "sparse" certificate of global optimality for $\mathbf{y} = 0$.

Finally, one may also consider several options:

- Instead of looking for a polynomial $\tilde{f}$ of same degree as f, one might allow polynomials of higher degree, and/or restrict certain coefficients of $\tilde{f}$ to be the same as those of f (e.g. for structural modeling reasons).

- One may restrict $\tilde{f}$ to a certain class of functions, e.g., quadratic polynomials and even convex quadratic polynomials. In the latter case and if the g_j's that define $\mathbf{K}$ are concave, the resulting semidefinite program simplifies drastically as it reduces to (a) solving a linear program and (b) to computing the ℓ_1-projection of a matrix Q onto the cone of positive semidefinite matrices.

II. Notation and definitions

Let $\mathbb{R}[\mathbf{x}]$ (resp. $\mathbb{R}[\mathbf{x}]_d$) denote the ring of real polynomials in the variables $\mathbf{x} = (x_1, \ldots, x_n)$ (resp. polynomials of degree at most d), whereas $\Sigma[\mathbf{x}]$ (resp. $\Sigma[\mathbf{x}]_d$) denotes its subset of sums of squares (s.o.s.) polynomials (resp. of s.o.s. of degree at most $2d$). For every $\alpha \in \mathbb{N}^n$ the notation $\mathbf{x}^\alpha$ stands for the monomial $x_1^{\alpha_1} \cdots x_n^{\alpha_n}$ and for every $i \in \mathbb{N}$, let $\mathbb{N}^p_d := \{\beta \in \mathbb{N}^n : \sum_j \beta_j \leq d\}$ whose cardinal is $s(d) = \binom{n+d}{n}$. A polynomial $f \in \mathbb{R}[\mathbf{x}]$ is written

$$\mathbf{x} \mapsto f(\mathbf{x}) \,=\, \sum_{\alpha \in \mathbb{N}^n} f_\alpha\, \mathbf{x}^\alpha,$$

and f can be identified with its vector of coefficients $\mathbf{f} = (f_\alpha)$ in the canonical basis $(\mathbf{x}^\alpha)$, $\alpha \in \mathbb{N}^n$. Denote by $\mathcal{S}^t \subset \mathbb{R}^{t\times t}$ the space of real symmetric matrices, and for any $\mathbf{A} \in \mathcal{S}^t$ the notation $\mathbf{A} \succeq 0$ stands for $\mathbf{A}$ is positive semidefinite.

For $f \in \mathbb{R}[\mathbf{x}]_d$, let $\|f\|_1 := \sum_{\alpha \in \mathbb{N}^n_d} |f_\alpha|$. A real sequence $\mathbf{z} = (z_\alpha)$, $\alpha \in \mathbb{N}^n$, has a *representing measure* if there exists some finite Borel measure μ on $\mathbb{R}^n$ such that

$$z_\alpha \,=\, \int_{\mathbb{R}^n} \mathbf{x}^\alpha\, d\mu(\mathbf{x}), \qquad \forall\, \alpha \in \mathbb{N}^n.$$

Given a real sequence $\mathbf{z} = (z_\alpha)$ define the linear functional $L_\mathbf{z} : \mathbb{R}[\mathbf{x}] \to \mathbb{R}$ by:

$$f \,(= \sum_\alpha f_\alpha \mathbf{x}^\alpha) \quad \mapsto L_\mathbf{z}(f) \,=\, \sum_\alpha f_\alpha\, z_\alpha, \qquad f \in \mathbb{R}[\mathbf{x}].$$

Moment matrix

The *moment* matrix associated with a sequence $\mathbf{z} = (z_\alpha)$, $\alpha \in \mathbb{N}^n$, is the real symmetric matrix $\mathbf{M}_d(\mathbf{z})$ with rows and columns indexed by $\mathbb{N}^n_d$, and whose entry (α, β) is just $z_{\alpha+\beta}$, for every $\alpha, \beta \in \mathbb{N}^n_d$. Alternatively, let $\mathbf{v}_d(\mathbf{x}) \in \mathbb{R}^{s(d)}$ be the vector $(\mathbf{x}^\alpha)$, $\alpha \in \mathbb{N}^n_d$, and define the matrices $(\mathbf{B}_\alpha) \subset \mathcal{S}^{s(d)}$ by

$$\mathbf{v}_d(\mathbf{x})\, \mathbf{v}_d(\mathbf{x})^T \,=\, \sum_{\alpha \in \mathbb{N}^n_{2d}} \mathbf{B}_\alpha\, \mathbf{x}^\alpha, \qquad \forall \mathbf{x} \in \mathbb{R}^n. \tag{2}$$

Then $\mathbf{M}_d(\mathbf{z}) = \sum_{\alpha \in \mathbb{N}^n_{2d}} z_\alpha\, \mathbf{B}_\alpha$.

Localizing matrix

With $\mathbf{z}$ as above and $g \in \mathbb{R}[\mathbf{x}]$ (with $g(\mathbf{x}) = \sum_\gamma g_\gamma \mathbf{x}^\gamma$), the *localizing* matrix associated with $\mathbf{z}$ and g is the real symmetric matrix $\mathbf{M}_d(g\,\mathbf{z})$ with rows and columns indexed by $\mathbb{N}^n_d$, and whose entry (α, β) is just $\sum_\gamma g_\gamma z_{\alpha+\beta+\gamma}$, for every $\alpha, \beta \in \mathbb{N}^n_d$. Alternatively, let $\mathbf{C}_\alpha \in \mathcal{S}^{s(d)}$ be defined by:

$$g(\mathbf{x})\, \mathbf{v}_d(\mathbf{x})\, \mathbf{v}_d(\mathbf{x})^T \,=\, \sum_{\alpha \in \mathbb{N}^n_{2d+\deg g}} \mathbf{C}_\alpha\, \mathbf{x}^\alpha, \qquad \forall \mathbf{x} \in \mathbb{R}^n. \tag{3}$$

Then $\mathbf{M}_d(g\,\mathbf{z}) = \sum_{\alpha \in \mathbb{N}^n_{2d+\deg g}} z_\alpha\, \mathbf{C}_\alpha$.

With $\mathbf{K}$ as in (1), let $g_0 \in \mathbb{R}[\mathbf{x}]$ be the constant polynomial $\mathbf{x} \mapsto g_0(\mathbf{x}) = 1$, and for every $j = 0, 1, \ldots, m$, let $v_j := \lceil (\deg g_j)/2 \rceil$.

Definition 1: With $d, k \in \mathbb{N}$ and $\mathbf{K}$ as in (1), let $Q_k(g) \subset \mathbb{R}[\mathbf{x}]$ and $Q^d_k \subset \mathbb{R}[\mathbf{x}]_d$ be the convex cones:

$$Q(g) \;:=\; \left\{ \sum_{j=0}^m \sigma_j\, g_j \;:\; \sigma_j \in \Sigma[\mathbf{x}],\, \forall j \right\} \tag{4}$$

$$Q_k(g) \;=\; \left\{ \sum_{j=0}^m \sigma_j\, g_j \;:\; \sigma_j \in \Sigma[\mathbf{x}]_{k-v_j},\, \forall j \right\} \tag{5}$$

$$Q^d_k(g) \;=\; Q_k(g) \cap \mathbb{R}[\mathbf{x}]_d. \tag{6}$$

Every element $h \in Q_k(g)$ is said to have a *Putinar's certificate* of positivity on $\mathbf{K}$, with degree bound k.

The cone $Q(g)$ is called the quadratic module associated with the g_j's. Obviously, if $h \in Q(g)$ the associated s.o.s. polynomials σ_j's provide a certificate of nonnegativity of h on $\mathbf{K}$. The name "Putinar's certificate" is coming from Putinar's Positivstellensatz [10] which asserts that under some archimedean assumption on the g_j's, every polynomial h strictly positive on $\mathbf{K}$ belongs to $Q_k(g)$ for some k. Namely, let $\mathbf{C}_d(\mathbf{K}) \subset \mathbb{R}[\mathbf{x}]_d$ be the convex cone of polynomials of degree at most d that are nonnegative on $\mathbf{K}$.

Theorem 1 (Putinar): Let $\mathbf{K}$ be as in (1) and assume that there is some $M > 0$ such that the quadratic polynomial $\mathbf{x} \mapsto M - \|\mathbf{x}\|^2$ belongs to $Q(g)$. Then every polynomial $f \in \mathbb{R}[\mathbf{x}]$ strictly positive on $\mathbf{K}$ belongs to $Q(g)$. In addition,

$$\mathrm{cl}\left(\bigcup_{k=0}^{\infty} Q_k^d(g)\right) = \mathbf{C}_d(\mathbf{K}), \qquad \forall d \in \mathbb{N}. \tag{7}$$

The first statement is just Putinar's Positivstellensatz [10] whereas the second statement is an easy consequence. Indeed let $f \in \mathbf{C}_d(\mathbf{K})$. If $f > 0$ on $\mathbf{K}$ then $f \in Q_k^d(g)$ for some k. If $f(\mathbf{x}) = 0$ for some $\mathbf{x} \in \mathbf{K}$, let $f_n := f + 1/n$, so that $f_n > 0$ on $\mathbf{K}$ for every $n \in \mathbb{N}$. But then $f_n \in \cup_{k=0}^{\infty} Q_k^d(g)$ and the result follows because $\|f_n - f\|_1 \to 0$ as $n \to \infty$.

The ideal inverse problem

Let $\mathbf{P}$ be the global optimization problem $f^* = \min_{\mathbf{x}}\{f(\mathbf{x}) : \mathbf{x} \in \mathbf{K}\}$ with $\mathbf{K} \subset \mathbb{R}^n$ as in (1), and $f \in \mathbb{R}[\mathbf{x}]_{df}$ where $df := \deg f$.

Identifying a polynomial $f \in \mathbb{R}[\mathbf{x}]_{df}$ with its vector of coefficients $\mathbf{f} = (f_\alpha) \in \mathbb{R}^{s(df)}$, one may and will identify $\mathbb{R}[\mathbf{x}]_{df}$ with the vector space $\mathbb{R}^{s(df)}$, i.e., $\mathbb{R}[\mathbf{x}]_{df} \ni f \leftrightarrow \mathbf{f} \in \mathbb{R}^{s(df)}$. Similarly, the convex cone $\mathbf{C}_{df}(\mathbf{K}) \subset \mathbb{R}[\mathbf{x}]_{df}$ can be identified with the convex cone $\{\mathbf{h} \in \mathbb{R}^{s(df)} : \mathbf{h} \leftrightarrow h \in \mathbf{C}_{df}(\mathbf{K})\}$ of $\mathbb{R}^{s(df)}$. Next, let $\mathbf{y} \in \mathbf{K}$ be fixed. With no loss of generality (possibly after a change of variable $\mathbf{x}' = \mathbf{x} - \mathbf{y}$) we may and will assume that $\mathbf{y} = 0 \in \mathbf{K}$. Consider the following optimization problem $\mathcal{P}$:

$$\rho^* = \min_{\tilde f \in \mathbb{R}[\mathbf{x}]_{df}} \{\, \|\tilde f - f\|_1 \,:\, \tilde f - \tilde f(0) \in \mathbf{C}_{df}(\mathbf{K}) \,\}. \tag{8}$$

Theorem 2: Problem (8) has an optimal solution $\tilde f^* \in \mathbb{R}[\mathbf{x}]_{df}$. In addition, $\rho^* = 0$ if and only if 0 is an optimal solution of $\mathbf{P}$.

Proof: Obviously the constant polynomial $\tilde f(\mathbf{x}) := 1$ for all $\mathbf{x}$, is a feasible solution with associated value $\delta := \|\tilde f - f\|_1$. Moreover the optimal value of (8) is bounded below by 0. Consider a minimizing sequence $(\tilde f^j) \subset \mathbb{R}[\mathbf{x}]_{df}$, $j \in \mathbb{N}$, hence such that $\|\tilde f^j - f\|_k \to \rho^k$ as $j \to \infty$. As we have $\|\tilde f^j - f\|_1 \le \delta$ for every j, the sequence $(\tilde f^j)$ belongs to the ℓ_1-ball $\mathbf{B}(f) := \{\tilde f \in \mathbb{R}[\mathbf{x}]_{df} : \|\tilde f - f\|_1 \le \delta\} \subset \mathbb{R}^{s(df)}$, a compact set. Therefore, there is an element $\tilde f^* \in \mathbf{B}(f)$ and a subsequence (j_t), $t \in \mathbb{N}$, such that $\tilde f^{j_t} \to \tilde f^*$ as $t \to \infty$. Let $\mathbf{x} \in \mathbf{K}$ be fixed arbitrary. Obviously $(0 \le)\, \tilde f^{j_t}(\mathbf{x}) - \tilde f^{j_t}(0) \to \tilde f^*(\mathbf{x}) - \tilde f^*(0)$ as $t \to \infty$, which implies $\tilde f^*(\mathbf{x}) - \tilde f^*(0) \ge 0$, and so, as $\mathbf{x} \in \mathbf{K}$ was arbitrary, $\tilde f^* - \tilde f^*(0) \ge 0$ on $\mathbf{K}$, i.e., $\tilde f^* - \tilde f^*(0) \in \mathbf{C}_{df}(\mathbf{K})$. Finally, we also obtain the desired result

$$\rho^* = \lim_{j\to\infty} \|\tilde f^j - f\|_1 = \lim_{t\to\infty} \|\tilde f^{j_t} - f\|_1 = \|\tilde f^* - f\|_1.$$

Next, if 0 is an optimal solution of $\mathbf{P}$ then $\tilde f := f$ is an optimal solution of $\mathcal{P}$ with value $\rho^* = 0$. Conversely, if $\rho^* = 0$ then $\tilde f^* = f$, and so by feasibility of $\tilde f^* (= f)$ for (8), $f(\mathbf{x}) \ge f(0)$ for all $\mathbf{x} \in \mathbf{K}$, which shows that 0 is an optimal solution of $\mathbf{P}$. ∎

Theorem 2 states that the ideal inverse optimization problem is well-defined. However, even though $\mathbf{C}_{df}(\mathbf{K})$ is a finite-dimensional closed convex cone, it has no simple and tractable characterization to be used for practical computation. Therefore one needs an alternative and more tractable version of problem $\mathcal{P}$. Fortunately, we next show that in the polynomial context such a formulation exists, thanks to powerful positive certificates from real algebraic geometry.

III. MAIN RESULT

As the ideal inverse problem is intractable we here provide tractable formulations whose size depends on a parameter $d \in \mathbb{N}$. If $\tilde f^*$ in Theorem 2 belongs to $Q(g)$ then when d increases, $\tilde f^* \in \mathbb{R}[\mathbf{x}]_{df}$ can be obtained in finitely many steps, by solving finitely many semidefinite programs.

A. The practical inverse problem

With $d \in \mathbb{N}$ fixed, consider the following optimization problem $\mathbf{P}_d$ (where $g_0 = 1$):

$$\begin{aligned} \rho_d = \;& \min_{\tilde f, \sigma_j} \|f - \tilde f\|_1 \\ \text{s.t.} \;& \tilde f(\mathbf{x}) - \tilde f(0) = \sum_{j=0}^{m} \sigma_j(\mathbf{x})\, g_j(\mathbf{x}),\ \forall \mathbf{x} \in \mathbb{R}^n \\ & \tilde f \in \mathbb{R}[\mathbf{x}]_{df};\ \sigma_j \in \Sigma[\mathbf{x}]_{d-v_j},\ j = 0,\dots,m, \end{aligned} \tag{9}$$

where $df = \deg f$, and $v_j = \lceil (\deg g_j)/2 \rceil$, $j = 0,\dots,m$.

Remark 1: Observe that the constant coefficient $\tilde f_0$ plays no role in the constraints of (9), but since we minimize $\|\tilde f - f\|_k$ then it is always optimal to set $\tilde f_0 = f_0$. That is, $\tilde f(0) = \tilde f_0 = f_0 = f(0)$.

The parameter $d \in \mathbb{N}$ impacts (9) by the maximum degree allowed for the s.o.s. weights $(\sigma_j) \subset \Sigma[\mathbf{x}]$ in Putinar's certificate for the polynomial $\mathbf{x} \mapsto \tilde f(\mathbf{x}) - \tilde f(0)$.

For any feasible solution $\tilde f$ of (9), the constraints of (9) state that the polynomial $\mathbf{x} \mapsto \tilde f(\mathbf{x}) - \tilde f(0)$ is in $Q_d^{df}(g)$. Therefore, in particular, $\tilde f(\mathbf{x}) \ge \tilde f(0)$ for all $\mathbf{x} \in \mathbf{K}$, and so 0 is a global minimum of $\tilde f$ on $\mathbf{K}$. So $\mathbf{P}_d$ is a strengthening of $\mathcal{P}$ in that one has replaced the constraint $\tilde f - \tilde f(0) \in \mathbf{C}_{df}(\mathbf{K})$ witht the stronger condition $\tilde f - \tilde f(0) \in Q_d^{df}(g)$. And so $\rho^* \le \rho_d$ for all $d \in \mathbb{N}$. However, as we next see, $\mathbf{P}_d$ is a tractable optimization problem with nice properties. Indeed, $\mathbf{P}_d$ is a convex optimization problem and even a semidefinite program, as one may rewrite $\mathbf{P}_d$ as:

$$\begin{aligned} & \rho_d = \min_{\tilde f, \lambda_\alpha, \mathbf{Z}_j} \sum_{\alpha \in \mathbb{N}^n_{df} \setminus \{0\}} \lambda_\alpha \quad \text{subject to:} \\ & \lambda_\alpha + \tilde f_\alpha \ge f_\alpha, \quad \forall \alpha \in \mathbb{N}^n_{df} \setminus \{0\} \\ & \lambda_\alpha - \tilde f_\alpha \ge -f_\alpha, \quad \forall \alpha \in \mathbb{N}^n_{df} \setminus \{0\} \\ & \langle \mathbf{Z}_0, \mathbf{B}_\alpha \rangle + \sum_{j=1}^{m} \langle \mathbf{Z}_j, \mathbf{C}^j_\alpha \rangle = \begin{cases} \tilde f_\alpha, & \text{if } 0 < |\alpha| \le df \\ 0, & \text{if } \alpha = 0 \text{ or } |\alpha| > df \end{cases} \\ & \mathbf{Z}_j \succeq 0, \quad j = 0, 1, \dots, m, \end{aligned} \tag{10}$$

with $\mathbf{B}_\alpha$ as in (2) and $\mathbf{C}^j_\alpha$ as in (3) (with g_j in lieu of g).

The semidefinite program dual of (10) reads

$$
\begin{aligned}
&\max_{\mathbf{u},\mathbf{v},\mathbf{z}} \sum_{\alpha\in\mathbb{N}^n_{df}\setminus\{0\}} f_\alpha(u_\alpha - v_\alpha)\;(= L_{\mathbf{z}}(f(0)-f))\\
&\text{subject to:}\\
&u_\alpha + v_\alpha \;\leq\; 1, \quad \forall\alpha\in\mathbb{N}^n_{df}\setminus\{0\}\\
&u_\alpha - v_\alpha + z_\alpha \;=\; 0, \quad \forall\alpha\in\mathbb{N}^n_{df}\setminus\{0\}\\
&\mathbf{M}_d(\mathbf{z}),\, \mathbf{M}_{d-v_j}(g_j\,\mathbf{z}) \;\succeq\; 0, \quad j=1,\ldots,m.
\end{aligned}
\tag{11}
$$

Lemma 1: Assume that $\mathbf{K}\subset\mathbb{R}^n$ has nonempty interior. Then there is no duality gap between the semidefinite program (10) and its dual (11). Moreover, (10) has an optimal solution $\tilde{f}_d\in\mathbb{R}[\mathbf{x}]_{df}$.

Proof: Observe that $\rho_d\geq 0$ and the constant polynomial $\tilde{f}(\mathbf{x})=0$ for all $\mathbf{x}\in\mathbb{R}^n$, is an obvious feasible solution of (9) (hence of (10)). Therefore ρ_d being finite, it suffices to prove that Slater's condition[1] holds for the dual (11). Then the conclusion of Lemma 1 follows from a standard result of convex optimization. Let μ be the a finite Borel measure defined by

$$
\mu(B) \;:=\; \int_{B\cap\mathbf{K}} \mathrm{e}^{-\|\mathbf{x}\|^2} d\mathbf{x}, \qquad \forall B\in\mathcal{B}
$$

(with $\mathcal{B}$ the usual Borel σ-field), and let $\mathbf{z}=(z_\alpha)$, $\alpha\in\mathbb{N}^n_{2d}$, with

$$
z_\alpha \;:=\; \kappa\int_{\mathbf{K}} \mathbf{x}^\alpha\, d\mu(\mathbf{x}), \qquad \alpha\in\mathbb{N}^n_{2d},
$$

for some $\kappa>0$ sufficiently small to ensure that

$$
\kappa\left|\int \mathbf{x}^\alpha\, d\mu(\mathbf{x})\right| \;<\; 1, \qquad \forall\alpha\in\mathbb{N}^n_{2d}\setminus\{0\}. \tag{12}
$$

Define $u_\alpha=\max[0,-z_\alpha]$ and $v_\alpha=\max[0,z_\alpha]$, $\alpha\in\mathbb{N}^n_{df}$, so that $u_\alpha+v_\alpha<1$, $\alpha\in\mathbb{N}^n_{2d}$. Hence $(u_\alpha,v_\alpha,\mathbf{z})$ is a feasible solution of (11). In addition, $\mathbf{M}_d(\mathbf{z})\succ 0$ and $\mathbf{M}_{d-v_j}(g_j\,\mathbf{z})\succ 0$, $j=1,\ldots,m$, because $\mathbf{K}$ has nonempty interior, and so Slater's condition holds for (11), the desired result. ∎

As already mentioned, we could have chosen the ℓ_∞- or ℓ_2-norm rather than the ℓ_1-norm. For instance, with the ℓ_∞-norm the semidefinite program (10) becomes

$$
\begin{aligned}
&\rho_d = \min_{\tilde{f},\lambda,\mathbf{Z}_j} \lambda \quad \text{subject to:}\\
&\lambda + \tilde{f}_\alpha \geq f_\alpha, \quad \forall\alpha\in\mathbb{N}^n_{df}\setminus\{0\}\\
&\lambda - \tilde{f}_\alpha \geq -f_\alpha, \quad \forall\alpha\in\mathbb{N}^n_{df}\setminus\{0\}\\
&\langle\mathbf{Z}_0,\mathbf{B}_\alpha\rangle + \sum_{j=1}^m \langle\mathbf{Z}_j,\mathbf{C}^j_\alpha\rangle \;=\; \begin{cases} \tilde{f}_\alpha, & \text{if } 0<|\alpha|\leq df\\ 0, & \text{if } \alpha=0 \text{ or } |\alpha|>df\end{cases}\\
&\mathbf{Z}_j\succeq 0, \quad j=0,1,\ldots,m,
\end{aligned}
\tag{13}
$$

while its dual reads

$$
\begin{aligned}
&\max_{\mathbf{u},\mathbf{v},\mathbf{z}} \sum_{\alpha\in\mathbb{N}^n_{df}\setminus\{0\}} f_\alpha(u_\alpha - v_\alpha)\;(= L_{\mathbf{z}}(f(0)-f))\\
&\text{subject to:}\\
&\sum_{\alpha\in\mathbb{N}^n_{df}\setminus\{0\}} u_\alpha + v_\alpha \;\leq\; 1\\
&u_\alpha - v_\alpha + z_\alpha \;=\; 0, \quad \forall\alpha\in\mathbb{N}^n_{df}\setminus\{0\}\\
&\mathbf{M}_d(\mathbf{z}),\, \mathbf{M}_{d-v_j}(g_j\,\mathbf{z}) \;\succeq\; 0, \quad j=1,\ldots,m.
\end{aligned}
\tag{14}
$$

Theorem 3: Assume that $\mathbf{K}$ in (1) has nonempty interior and let $\mathbf{x}^*\in\mathbf{K}$ be a global minimizer of $\mathbf{P}$ with optimal value f^*, and let $\tilde{f}_d\in\mathbb{R}[\mathbf{x}]$ be an optimal solution of $\mathbf{P}_d$ in (9) with optimal value ρ_d. Then:

(a) $0\in\mathbf{K}$ is a global minimizer of the problem $\tilde{f}^*_d=\min_{\mathbf{x}}\{\tilde{f}_d(\mathbf{x})\,:\,\mathbf{x}\in\mathbf{K}\}$. In particular, if $\rho_d=0$ then $\tilde{f}_d=f$ and 0 is a global minimizer of $\mathbf{P}$.

(b) $f^*\;\leq\; f(0)\;\leq\; f^*+\rho_d\sup_{\alpha\in\mathbb{N}^n_{df}}|(\mathbf{x}^*)^\alpha|$. In particular if $\mathbf{K}\subseteq[-1,1]^n$ then $f^*\;\leq\; f(0)\;\leq\; f^*+\rho_d$.

Proof: (a) Existence of $\tilde{f}_d$ is guaranteed by Lemma 1. From the constraints of (9) we have: $\tilde{f}_d(\mathbf{x})-\tilde{f}(0)=\sum_{j=0}^m\sigma_j(\mathbf{x})\,g_j(\mathbf{x})$ which implies that $\tilde{f}_d(\mathbf{x})\geq\tilde{f}(0)$ for all $\mathbf{x}\in\mathbf{K}$, and so 0 is a global minimizer of the optimization problem $\mathbf{P}'\,:\,\min_{\mathbf{x}}\{\tilde{f}_d(\mathbf{x})\,:\,\mathbf{x}\in\mathbf{K}\}$.

(b) Let $\mathbf{x}^*\in\mathbf{K}$ be a global minimizer of $\mathbf{P}$. Observe that with $k=1$,

$$
\begin{aligned}
f^*=f(\mathbf{x}^*) \;&=\; \underbrace{f(\mathbf{x}^*)-\tilde{f}_d(\mathbf{x}^*)}+\underbrace{\tilde{f}_d(\mathbf{x}^*)-\tilde{f}_d(0)}_{\geq 0}+\tilde{f}_d(0)\\
&\geq\; \tilde{f}_d(0)-|\tilde{f}_d(\mathbf{x}^*)-f(\mathbf{x}^*)|\\
&\geq\; \tilde{f}_d(0)-\|\tilde{f}_d-f\|_1\times\sup_{\alpha\in\mathbb{N}^n_{df}}|(\mathbf{x}^*)^\alpha|\\
&=\; \tilde{f}_d(0)-\rho_d\sup_{\alpha\in\mathbb{N}^n_{df}}|(\mathbf{x}^*)^\alpha|,
\end{aligned}
\tag{15}
$$

and the result follows because $\tilde{f}_d(0)=\tilde{f}_{d0}=f_0=f(0)$; see Remark 1. ∎

So not only Theorem 3 states that 0 is the global optimum of the optimization problem $\min\{\tilde{f}_d(\mathbf{x})\,:\,\mathbf{x}\in\mathbf{K}\}$, but it also states that the optimal value ρ_d also measures how far is $f(0)$ from the optimal value f^* of the inital problem $\mathbf{P}$. Moreover, observe that Theorem 3(a) requires *no* assumption on the basic semi-algebraic set $\mathbf{K}$, and Theorem 3(b) only requires existence of a global minimizer of $\mathbf{P}$, and in particular, $\mathbf{K}$ may not be compact.

B. A canonical form

In fact when df is even, then the optimal solution $\tilde{f}_d\in\mathbb{R}[\mathbf{x}]_{df}$ in Theorem 3 (with $k=1$) takes a particularly simple *canonical* form, which is due to the use of the ℓ_1-norm.

Theorem 4: Assume that $\mathbf{K}$ has a nonempty interior and let $\mathbf{x}^*\in\mathbf{K}$ be a global minimizer of $\mathbf{P}$ with optimal value f^*. Let $\tilde{f}_d\in\mathbb{R}[\mathbf{x}]_{df}$ be an optimal solution of $\mathbf{P}_d$ in (9) with optimal value ρ^1_d for the ℓ_1-norm. Then $\tilde{f}_d$ is of the form:

$$
\tilde{f}_d(\mathbf{x}) \;=\; f(\mathbf{x})+\mathbf{b}'\mathbf{x}+\sum_{i=1}^n\left(\lambda^*_i\,x_i^2+\gamma^*_i\,x_i^{df}\right), \tag{16}
$$

[1] Slater's condition holds if there exists $\mathbf{x}_0\in\mathbf{K}$ such that $g_j(\mathbf{x}_0)>0$ for every $j=1,\ldots,m$.

for some vector $\mathbf{b} \in \mathbb{R}^n$ and some nonnegative vectors $\lambda^*, \gamma^* \in \mathbb{R}^n$, optimal solution of the semidefinite program:

$$\begin{aligned}\rho_d^1 := \min_{\lambda,\gamma\geq 0,\mathbf{b}} & \ \|\mathbf{b}\|_1 + \sum_{i=1}^n (\lambda_i + \gamma_i) \\ \text{s.t.} & \ f - f(0) + \mathbf{b}'\mathbf{x} + \sum_{i=1^n} \left(\lambda_i\, x_i^2 + \gamma_i\, x_i^{df}\right) \in Q_d(g)\end{aligned} \tag{17}$$

Proof: The dual (11) of (10) is equivalent to:

$$\begin{aligned}\max_{\mathbf{z}} \quad & L_{\mathbf{z}}(f(0) - f)) \\ \text{s.t.} \quad & \mathbf{M}_d(\mathbf{z}), \mathbf{M}_{d-v_j}(g_j\,\mathbf{z}) \succeq 0, \quad j = 1, \ldots, m, \\ & |z_\alpha| \leq 1, \quad \forall\, \alpha \in \mathbb{N}^n_{df} \setminus \{0\}.\end{aligned} \tag{18}$$

But $\mathbf{M}_d(\mathbf{z}) \succeq 0 \Rightarrow \mathbf{M}_{df/2}(\mathbf{z}) \succeq 0$. On the other hand, recalling that df is even and $\mathbf{M}_{df/2}(\mathbf{z}) \succeq 0$, one may use same arguments as those used in Lasserre and Netzer [9, Lemma 4.1, 4.2], to obtain $|z_\alpha| \leq \max_i\{\max[L_{\mathbf{z}}(x_i^2), L_{\mathbf{x}}(x_i^{df})]\}$, for very $\alpha \in \mathbb{N}^n_{df}$ with $1 < |\alpha| \leq df$. Hence in (18) one may replace the constraint $|z_\alpha| \leq 1$ for all $\alpha \in \mathbb{N}^n_{df} \setminus \{0\}$ with the $3n$ inequality constraints:

$$\pm L_{\mathbf{z}}(x_i) \leq 1, \ L_{\mathbf{z}}(x_i^2) \leq 1, \ L_{\mathbf{x}}(x_i^{df}) \leq 1, \tag{19}$$

for all $i = 1, \ldots, n$. And so the dual of the modified SDP (19) is now

$$\begin{aligned}\max_{\mathbf{b}^1,\mathbf{b}^2,\lambda,\gamma} \quad & \sum_{i=1}^n \left((b_i^1 + b_i^2) + \lambda_i + \gamma_i\right) \\ \text{s.t.} \quad & f - f(0) + (\mathbf{b}^1 - \mathbf{b}^2)'\mathbf{x} + \sum_{i=1}^n \left(\lambda_i x_i^2 + \gamma_i\, x_i^{df}\right) \\ & \in Q_d(g)\,;\ \mathbf{b}^1, \mathbf{b}^2, \lambda, \gamma \geq 0.\end{aligned}$$

which is equivalent to (17). ■

This special *canonical* form of $\tilde{f}_d$ is specific to the ℓ_1-norm, which yields the constraint $|z_\alpha| \leq 1$, $\alpha \in \mathbb{N}^n_{df} \setminus \{0\}$ in the dual (11) and allows its simplification (19) thanks to a property of the moment matrix described in [9].

C. Asymptotics when $d \to \infty$

We now relate $\mathbf{P}_d$, $d \in \mathbb{N}$, with the ideal inverse problem $\mathcal{P}$ in (8) when d increases. Recall that $\rho^* \leq \rho_d$ for every $d \in \mathbb{N}$.

Proposition 1: Let $\mathbf{K}$ in (1) be with nonempty interior, ρ_d be as in (9), and let $\tilde{f}^* \in \mathbb{R}[\mathbf{x}]_{df}$ be an optimal solution of (8) with associated optimal value ρ^*.

The sequence (ρ_d), $d \in \mathbb{N}$, is monotone nonincreasing and if the polynomial $\mathbf{x} \mapsto \tilde{f}^*(\mathbf{x}) - \tilde{f}^*(0)$ is in $Q(g)$, then $\rho_d = \rho^*$ for some d.

Proof: By definition if $\tilde{f}^* - \tilde{f}^*(0) \in Q(g)$ then $\tilde{f}^* - \tilde{f}^*(0) \in Q_d^{df}(g)$ for some $d = d_0$. Hence, $\tilde{f}^*$ is a feasible solution of (9) (when $d = d_0$) but with value $\rho^* \geq \rho_d$. Therefore, we conclude that $\tilde{f}^*$ is an optimal solution of (9) when $d = d_0$. ■

Proposition 1 relates ρ_d and ρ^* in a strong sense when $\tilde{f}^* - \tilde{f}^*(0) \in Q(g)$. However, we do know how restrictive is the assumption $\tilde{f}^* - \tilde{f}^*(0) \in Q(g)$ compared to $\tilde{f}^* - \tilde{f}^*(0) \in \mathbf{C}_{df}(\mathbf{K})$. Indeed, even though $\mathbf{C}_{df}(\mathbf{K}) = \mathrm{cl}\,(\cup_{k=0}^\infty Q_k^{df})$ when $\mathbf{K}$ satisfies the assumptions of Theorem 7, an approximating sequence $(f_n) \subset Q(g)$, $n \in \mathbb{N}$ (with $\|f_n - \tilde{f}^*\|_k \to 0$), may not satisfy $f_n(\mathbf{x}) - f_n(0) \geq 0$ for all $\mathbf{x}$ on $\mathbf{K}$.

D. Convexity

One may wish to restrict to search for *convex* polynomials $\tilde{f} \in \mathbb{R}[\mathbf{x}]_{df}$ (no matter if f itself is convex). For instance if the g_j's are concave (so that $\mathbf{K}$ is convex) but f is not, one may wish to find the convex optimization problem for which $\mathbf{y} = 0 \in \mathbf{K}$ is an optimal solution, and with convex polynomial criterion $\tilde{f} \in \mathbb{R}[\mathbf{x}]_{df}$ closest to f.

If $df > 2$ then it suffices to add to the semidefinite program (9) the additional Putinar's certificate

$$\begin{aligned}(\mathbf{x}, \mathbf{u}) \mapsto \ \mathbf{u}\,\nabla^2 \tilde{f}(\mathbf{x})\,\mathbf{u} \ &= \ \sum_{j=0}^m \psi_j(\mathbf{x}, \mathbf{u})\, g_j(\mathbf{x}) \\ & \quad + \psi_{m+1}(\mathbf{x}, \mathbf{u})(1 - \|\mathbf{u}\|^2),\end{aligned} \tag{20}$$

with $\psi_{m+1} \in \mathbb{R}[\mathbf{x}, \mathbf{u}]_{d-1}$ and $\psi_j \in \Sigma_{d-v_j}[\mathbf{x}, \mathbf{u}]$, for all $j = 0, 1, \ldots, m$. Indeed, (20) is a Putinar's certificate of convexity for $\tilde{f}$ on $\mathbf{K}$, with degree bound d, which also translates into additional semidefinite constraints.

If $df \leq 2$ (i.e. if $\tilde{f}(\mathbf{x}) = \mathbf{x}^T\mathbf{Q}\mathbf{x} + \mathbf{b}^T\mathbf{x} + c$) for some real symmetric matrix $\mathbf{Q} \in \mathbb{R}^{n\times n}$, some vector $\mathbf{b} \in \mathbb{R}^n$ and some scalar $c \in \mathbb{R}$) then in (9) it suffices to add constraint $\nabla^2 \tilde{f}(\mathbf{x}) = 2\tilde{\mathbf{Q}} \succeq 0$, which is just a Linear Matrix Inequality (LMI). And therefore, again, (9) can be rewritten as a semidefinite program, namely (10) with the additional LMI constraint $\tilde{\mathbf{Q}} \succeq 0$.

Notice that for $k = 1, 2$, it also makes sense to search for $\tilde{f} \in \mathbb{R}[\mathbf{x}]_2$ even if f has degree $df > 2$, i.e., if $f(\mathbf{x}) = c + \mathbf{b}^T\mathbf{x} + \frac{1}{2}\mathbf{x}^T\mathbf{Q}\mathbf{x} + h(\mathbf{x})$ where $h \in \mathbb{R}[\mathbf{x}]$ does not contains monomials of degree smaller than 3. This means that one searches for the convex program with quadratic cost closest to f.

So for instance, in the case where one searches for $\tilde{f} \in \mathbb{R}[\mathbf{x}]_2$, and given $0 \in \mathbf{K}$ let $J(0) := \{j \in \{1, \ldots, m\} : g_j(0) = 0\}$ be the set of constraints that are active at 0. If the g_j's that define $\mathbf{K}$ are concave and Slater's condition holds for $\mathbf{K}$, then one may simplify (9). Writing $\tilde{f} = \frac{1}{2}\mathbf{x}^T\tilde{\mathbf{Q}}\mathbf{x} + \tilde{\mathbf{b}}^T\mathbf{x} + \tilde{c}$, (9) now reads:

$$\begin{aligned}\rho := \min_{\tilde{\mathbf{Q}},\tilde{\mathbf{b}},\lambda} \quad & \|f - \tilde{f}\|_1 \\ \text{s.t.} \quad & \tilde{\mathbf{b}} = \sum_{j\in J(0)} \lambda_j\, \nabla g_j(0) \\ & \tilde{\mathbf{Q}} \succeq 0;\ \lambda_j \geq 0, \quad j \in J(0).\end{aligned} \tag{21}$$

which simplifies to:

$$\rho \ = \ \min_{\tilde{\mathbf{Q}}\succeq 0} \|\tilde{\mathbf{Q}} - \mathbf{Q}\|_1 + \min_{\lambda\geq 0} \|\mathbf{b} - \sum_{j\in J(0)} \lambda_j \nabla g_j(0)\|_1 \tag{22}$$

Problems (21) and (22) are much simpler than (9) because one has replaced Putinar's certificate of nonnegativity on $\mathbf{K}$ by the Karush-Kuhn-Tucker (KKT) optimality conditions at the point $\mathbf{x} = 0 \in \mathbf{K}$, and the convexity condition of $\tilde{f}$ reduces to the single LMI $\tilde{\mathbf{Q}} \succeq 0$. In particular, there is no index d! Moreover, problem (22) reduces to solving separately a linear program and a semidefinite program. The latter simply computes the ℓ_1-projection of $\mathbf{Q}$ onto the closed convex cone of semidefinite matrices.

Lemma 2: Let $\mathbf{K} \subset \mathbb{R}^n$ be as in (1) with g_j being concave for every $j = 1, \ldots, m$. Then (21) has an optimal solution $\tilde{f}^* \in \mathbb{R}[\mathbf{x}]_2$ and 0 is an optimal solution of the convex optimization problem $\mathbf{P}' : \min\{\tilde{f}^*(\mathbf{x}) \,:\, \mathbf{x} \in \mathbf{K}\}$.

Proof: Let $(\tilde{f}, \lambda)$ (with $\tilde{f} \in \mathbb{R}[\mathbf{x}]_2$) be any feasible solution of (21). The constraint in (21) states that $\nabla L(0) = 0$, where $L \in \mathbb{R}[\mathbf{x}]$ is the Lagrangian polynomial $\mathbf{x} \mapsto L(\mathbf{x}) := \tilde{f}(\mathbf{x}) - \sum_{j \in J(0)} \lambda_j \, g_j(\mathbf{x})$, which is convex on $\mathbf{K}$ because the g_j's are concave, the λ_j's are nonnegative, and $\tilde{f}$ is convex. Therefore $\nabla L(0) = 0$ implies that 0 is a global minimum of L on $\mathbb{R}^n$ and a global minimum of $\tilde{f}$ on $\mathbf{K}$ because

$$\tilde{f}(\mathbf{x}) \,\geq\, L(\mathbf{x}) \,\geq\, L(0) \,=\, \tilde{f}(0), \qquad \forall\, \mathbf{x} \in \mathbf{K}. \tag{23}$$

It remains to prove that (21) has an optimal solution $\tilde{f}^*$. But we have seen that (21) is equivalent to (22) for which an optimal solution can be found by solving a linear program and a semidefinite program. ∎

So in this case where the g_j's are concave (hence $\mathbf{K}$ is convex), one obtains the convex programming problem with quadratic cost whose criterion is the closest to f for the ℓ_1-norm.

E. Examples

Example 1: Let $n = 2$ and consider the optimization problem $\mathbf{P} \,:\, f^* = \min_{\mathbf{x}} \{f(\mathbf{x}) \,:\, \mathbf{x} \in \mathbf{K}\}$ with $\mathbf{x} \mapsto f(\mathbf{x}) = x_1 + x_2$, and

$$\mathbf{K} \,=\, \{\mathbf{x} \in \mathbb{R}^2 \,:\, (x_1+1)(x_2+1) \geq 1;\; -1/2 \leq \mathbf{x} \leq 1\,\}.$$

The polynomial f is convex and the set $\mathbf{K}$ is convex as well, but the polynomials that define $\mathbf{K}$ are not all concave. The point $\mathbf{y} = 0 \in \mathbf{K}$ is a global minimizer. With $d = 1$, one searches for an affine polynomial $\tilde{f}_1$ such that

$$\tilde{f}_1 - \tilde{f}_1(0) = \sigma_0 + ((x_1+1)(x_2+1) - 1)\sigma_1 + \sum_{i=1}^{2} (1 - x_i)\psi_i + (x_i + 1/2)\phi_i,$$

for some s.o.s. polynomials $\sigma_1, \psi_i\, \phi_i \in \Sigma[\mathbf{x}]_0$ and some s.o.s. polynomial $\sigma_0 \in \Sigma[\mathbf{x}]_1$. But then necessarily $\sigma_1 = 0$, which in turn implies that σ_0 is a constant polynomial. A straightforward calculation shows that $\tilde{f}_1(\mathbf{x}) = 0$ for all $\mathbf{x}$, and so $\rho_1 = 2$. On the other hand $\rho_3 = 0$. However, if now $\mathbf{K}$ has the representation:

$$\mathbf{K} = \begin{array}{ll} \{\mathbf{x} \,: & (x_1+1)(x_2+1) \,\geq\, 1 \\ & (x_i + 1/2)(1 - x_i) \,\geq\, 0,\; i = 1, 2\,\} \end{array},$$

then it turns out that

$$x_1 + x_2 \,=\, \tfrac{1}{5} + \tfrac{2}{5}(x_1 - x_2)^2 + \tfrac{4}{5}((x_1+1)(x_2+1) - 1) + \tfrac{2}{5}\sum_{i=1}^{2} (x_i + 1/2)(1 - x_i),$$

i.e., $f - f^* \in Q_1^1(g)$. Hence the test of inverse optimality yields $\rho_1 = 0$ with $\tilde{f}_1 = f$. This example shows that the representation of $\mathbf{K}$ may be important.

Example 2: Again consider Example 1 but now with $\mathbf{y} = (0.1, -0.091) \in \mathbf{K}$, which is not a global optimum of f on $\mathbf{K}$ any more. By solving (10) with $d = 1$ we still find $\rho_1 = 0$, and with $d = 2$ we find $\tilde{f}_2(\mathbf{x}) \approx 0.82782 x_1 + x_2$. And indeed by solving the new optimization problem with criterion $\tilde{f}_2$ (using GloptiPoly [4]) we find the global minimizer $(0.0991, -0.092) \approx \mathbf{y}$.

Example 3: Let $\mathbf{P}$ be the MAXCUT problem $\max\{\mathbf{x}'\mathbf{A}\mathbf{x} : \mathbf{x}_i^2 = 1, i = 1, \ldots, n\}$ where $\mathbf{A} = \mathbf{A}' \in \mathbb{R}^{n \times n}$ and $\mathbf{A}_{ij} = 1/2$ for all $i \neq j$. For n odd, the optimal solution is $\mathbf{y} = (y_j)$ with $y_j = 1$, $j = 1, \ldots \lceil n/2 \rceil$, and $y_j = -1$ otherwise. However, one cannot obtain

$$\mathbf{x}'\mathbf{A}\mathbf{x} - \mathbf{y}'\mathbf{A}\mathbf{y} = \sigma + \sum_{j=1}^{n} \gamma_i (x_i^2 - 1) \tag{24}$$

for some $\sigma \in \Sigma[\mathbf{x}]_1$ and $\lambda, \gamma \in \mathbb{R}$. Hence the inverse optimization problem reads: Find $\tilde{\mathbf{A}} \in \mathbb{R}^{n \times n}$ such that (24) holds and $\tilde{\mathbf{A}}$ minimizes the ℓ_1-norm $\|\mathbf{A} - \tilde{\mathbf{A}}\|_1$. Solving (10) for $n = 5$ with $\mathbf{y}$ as above, we find that $\tilde{\mathbf{A}} = \mathbf{A}$ except for the entries $(i, j) \in \{(1,2), (1,3), (2,3)\}$ now equal to $1/3$.

References

[1] S. Ahmed and Yongpei Guan. The inverse optimal value problem.

[2] R.K. Ahuja and J.B. Orlin. Inverse optimization, Oper. Res. **49**, 771–783 (2001)

[3] D. Burton and P. Toint. On an instance of the inverse shortest path problem, Math. Program. **63**, 45–61 (1992)

[4] D. Henrion, J.B. Lasserre, J. Lofberg. Gloptipoly 3: moments, optimization and semidefinite programming, Optim. Methods and Softw. **24**, 761–779 (2009)

[5] C. Heuberger. Inverse combinatorial oprtimization: A survey on problems, methods and results, Journal of Comb. Optim. **8**, 329–361 (2004).

[6] R.A. Freeman and P.V. Kokotovic. Inverse optimality in robust stabilization, SIAM J. Control Optim. **34**, 1365–1391 (1996)

[7] J.B. Lasserre. Global optimization with polynomials and the problem of moments, SIAM J. Optim. **11**, 796–817 (2001)

[8] J.B. Lasserre. Convergent SDP-relaxations in polynomial optimization with sparsity, SIAM J. Optim. **17**, 822–843 (2006)

[9] J.B. Lasserre, T. Netzer. SOS approximations of nonnegative polynomials via simpla high degree perturbation, Math. Z. **256**, 99–112 (2006)

[10] M. Putinar. Positive polynomials on compact semi-algebraic sets, Indiana Univ. Math. J. **42**, 969–984 (1993).

[11] A. Schaefer. Inverse integer programming, Optim. Letters. **3**, 483–489 (2004)