IBM Research Report

Intractability of Approximate Multi-dimensional Nonlinear Optimization on Independence Systems

Jon Lee
IBM Research Division
Thomas J. Watson Research Center
P.O. Box 218
Yorktown Heights, NY 10598, USA

Shmuel Onn
Technion - Israel Institute of Technology
Haifa, Israel

Robert Weismantel
ETH
Zürich, Switzerland
Intractability of approximate multi-dimensional nonlinear optimization on independence systems

Jon Lee  Shmuel Onn  Robert Weismantel

Abstract

We consider optimization of nonlinear objective functions that balance $d$ linear criteria over $n$-element independence systems presented by linear-optimization oracles. For $d = 1$, we have previously shown that an $r$-best approximate solution can be found in polynomial time. Here, using an extended Erdős-Ko-Rado theorem of Frankl, we show that for $d = 2$, finding a $\rho n$-best solution requires exponential time.

1 Introduction

Given system $S \subseteq \{0, 1\}^n$, integer $d \times n$ matrix $W$, and function $f : \mathbb{Z}^d \rightarrow \mathbb{Z}$, consider the problem of minimizing the nonlinear composite function $f(Wx)$ over $S$, that is,

$$\min \{ f(Wx) : x \in S \}.$$  

(1)

This can be interpreted as multi-criteria optimization, where row $W_i$ of $W$ gives a linear function $W_ix$ representing the value of feasible point $x \in S$ under criterion $i$, and the objective-function value $f(Wx) = f(W_1x, \ldots, W_dx)$ is a balancing of these $d$ criteria.

We assume that we can do linear optimization over $S$ to begin with; namely $S$ is presented by a linear-optimization oracle, which when queried on $w \in \mathbb{Z}^n$, solves $\max \{ wx : x \in S \}$. For restricted systems $S$, such as matroids and matroid intersections, or restricted functions $f$, such as concave functions, problem (1) can be solved in polynomial time (see [1, 2]). A comprehensive description of the state of the art on this area can be found in [5].

Here, we continue our investigation from [4] of problem (1), where $S$ is an arbitrary independence system, that is, $S$ is nonempty, and $x \leq y \in S$ with $x \in \{0, 1\}^n$ imply $x \in S$.

A feasible point $x^* \in S$ is called an $r$-best solution of problem (1) if there are at most $r$ better objective-function values attainable by other feasible points, that is,

$$|\{ f(Wx) : f(Wx) < f(Wx^*), x \in S \}| \leq r.$$ 

So it provides a suitable approximation to (1). In particular, a 0-best solution is optimal.

In [4], we considered the case of $d = 1$, that is, the problem $\min \{ f(wx) : x \in S \}$ with $w \in \mathbb{Z}^n$. We showed that for any fixed positive integers $a_1, \ldots, a_p$ there is a polynomial-time algorithm that, given any $w \in \{a_1, \ldots, a_p\}^n$, provides an $r(a_1, \ldots, a_p)$-best solution to the problem, where $r(a_1, \ldots, a_p)$ is a constant related to Frobenius numbers of some of the $a_i$. In particular, for any $p = 2$ integers, $r(a_1, a_2) = F(a)$ is the Frobenius number.
In this note, we consider the problem in dimension $d = 2$. We restrict attention to $2 \times n$ matrices $W$ that are $\{0,1\}$-valued. Then the image of $S$ under $W$ satisfies

$$WS := \{Wx : x \in S\} \subseteq \{0,1,\ldots,n\}^2.$$  

Therefore, the problem of computing the optimal objective-function value of (1) is seemingly reducible to computing the image $WS$ by checking if $y \in WS$ for each of the $(n + 1)^2$ points $y$ in $\{0,1,\ldots,n\}^2$, and determining the minimum value of $f$ over $WS$. Unfortunately, this so called fiber problem, of checking if $y \in WS$, is computationally hard. In particular, already for $S$ the set of (indicators of) matchings in a bipartite graph, over which linear optimization is easy, this problem includes as a special case the notorious exact matching problem whose complexity is long open (see [6]).

Here we show that there is a universal positive constant $\rho$ such that, already for $d = 2$, matrix $W$ each column of which is one of the two unit vectors in $\mathbb{Z}^2$, and a very simple explicit function $f$ supported on $\{0,1,\ldots,n\}^2$, there is no polynomial-time algorithm that can produce even a $\rho n$-best solution of problem (1) for every independence system $S \subseteq \{0,1\}^n$, let alone find a constant $\tau$-best or optimal solution. Our construction makes use of a beautiful extension of the classical Erdős-Ko-Rado theorem due to Frankl (see [3]).

It would be interesting to know whether our construction can be refined to shed some light on the exact matching and related open problems of [6], and whether other natural oracles for $S$ could lead to polynomial-time solution of problem (1) in dimensions $d = 2$ and higher.

2 A $\rho n$-best solution cannot be found in polynomial time

Theorem 2.1. There exists a universal positive constant $\rho$ such that no polynomial-time algorithm can compute a $\rho n$-best solution of the 2-dimensional nonlinear optimization problem $\min \{f(Wx) : x \in S\}$ over every independence system $S \subseteq \{0,1\}^n$ presented by a linear-optimization oracle, with $W$ an integer $2 \times n$ weight matrix each column of which is one of the unit vectors in $\mathbb{Z}^2$, and $f$ an explicit function supported on $\{0,1,\ldots,n\}^2$.

In fact, the following explicit statement holds. Let $l$ be any positive integer with $l \geq 2^{10}$, $k := 7l$, $m := 8l^2$, $n := 2m$, and $\rho := \frac{1}{17}$. Let $W$ be the $2 \times n$ matrix with first $m$ columns the unit vector $1_1$ and last $m$ columns the unit vector $1_2$. Define $f$ on $\mathbb{Z}^2$ explicitly by

$$f(y) = f(y_1,y_2) := \begin{cases} (y_1 - k) - l(y_2 - k) - 1 & \text{if } k + 1 \leq y_1, y_2 \leq k + l, \\ 0 & \text{otherwise.} \end{cases}$$

Then at least $2^{\frac{7}{17}n}$ queries to the oracle of $S$ are needed to compute a $\frac{1}{17}n$-best solution.

Proof. Let $l \geq 2^{10}$ be a positive integer, $k, m, n, \rho$ and $W$ as above, and $f$ as in (3) above. It is more convenient here to work with set systems over ground set $N := \{1,\ldots,n\}$ rather than sets of vectors in $\{0,1\}^n$. As usual, vectors $x \in \{0,1\}^n$ are in bijection with subsets $X \subseteq N$, with corresponding elements satisfying $X = \text{supp}(x)$ the support of $x$, and $x = 1_X$ being the indicator of $X$. So we replace each $S \subseteq \{0,1\}^n$ by the set system $\mathcal{S} := \{X = \text{supp}(x) : x \in S\}$. Also, for $c \in \mathbb{Z}^n$ and $X \subseteq N$, we write
Claim: Moreover, the objective-function value of every \( S \) implies that either there exists a \( \rho_n \)-best solution of the minimization problem over \( S \). For each \( y_1, y_2 \leq m \), let

\[
S_{y_1,y_2} := \{ X : X = X_1 \cup X_2 : |X_1| = y_1, |X_2| = y_2 \}.
\]

Next, let

\[
S^* := \{ X : (|X_1|, |X_2|) \leq (m, k) \text{ or } (|X_1|, |X_2|) \leq (k, m) \}.
\]

Then \( S^* \) is an independence system whose image is given by

\[
WS^* = \{(y_1, y_2) : (y_1, y_2) \leq (m, k) \text{ or } (y_1, y_2) \leq (k, m) \}.
\]

Moreover, the objective-function value of every \( X \in S^* \), and hence in particular of every \( \rho_n \)-best solution of the minimization problem over \( S^* \), satisfies \( f(WX) = 0 \).

Next, for each \( Y \in S_{k+l,k+l} \), let

\[
S_Y := S^* \cup \{ X : X \subseteq Y \}.
\]

Then \( S_Y \) is also an independence system, with image

\[
WS_Y = WS^* \cup \{(y_1, y_2) : (k+1, k+1) \leq (y_1, y_2) \leq (k+l, k+l) \}.
\]

Moreover, the objective-function values of the points in \( S_Y \setminus S^* \), whose images lie in \( WS_Y \setminus WS^* \), attain exactly all \( l^2 = \frac{16}{16}n > \rho n \) values \(-1, -2, \ldots, -l^2 \), and so the value of every \( \rho n \)-best solution of the minimization problem over \( S_Y \) satisfies \( f(WX) \leq -1 \).

For each vector \( c \in \mathbb{Z}^n \) and each pair \( 1 \leq i_1, i_2 \leq l \), let

\[
T_{i_1,i_2}(c) := \{ Z \in S_{k+i_1,k+i_2} : cZ > \max\{cX : X \in S^*\} \}.
\]

Claim: For every \( c \in \mathbb{Z}^n \) and every pair \( 1 \leq i_1, i_2 \leq l \), we have

\[
|T_{i_1,i_2}(c)| \leq \binom{m}{l} \binom{m}{k+l}.
\]

Proof of Claim: Consider any pair \( U = U_1 \cup U_2, V = V_1 \cup V_2 \in T_{i_1,i_2}(c) \). We now show that either \( |U_1 \cap V_1| \geq k + 1 \) or \( |U_2 \cap V_2| \geq k + 1 \). Suppose, indirectly, otherwise. Let

\[
X := (U_1 \cap V_1) \cup (U_2 \cup V_2),
\]

\[
Y := (U_1 \cup V_1) \cup (U_2 \cap V_2).
\]

Then \( |U_1 \cap V_1| \leq k \) and \( |U_2 \cup V_2| \leq m \) imply \( X \in S^* \), and \( |U_1 \cup V_1| \leq m \) and \( |U_2 \cap V_2| \leq k \) imply \( Y \in S^* \). We then obtain the following contradiction,

\[
0 < cU - cX = c(U_1 \setminus V_1) - c(V_2 \setminus U_2) = cY - cV < 0.
\]
So indeed, for every pair \( U = U_1 \sqcup U_2, V = V_1 \sqcup V_2 \in T_{i_1,i_2}(c) \subseteq S_{k+i_1,k+i_2} \), either \(|U_1 \cap V_1| \geq k + 1\) or \(|U_2 \cap V_2| \geq k + 1\). Therefore, we can now apply the extended Erdős-Ko-Rado theorem for direct products of Frankl [3, Theorem 2], which implies

\[
\left| \frac{|T_{i_1,i_2}(c)|}{|S_{k+i_1,k+i_2}|} \right| \leq \max \left\{ \left( \frac{m - (k + i_1)}{(k + i_1) - (k + 1)} \right) \left( \frac{m - (k + i_2)}{(k + i_2) - (k + 1)} \right) \right\}
\]

from which it is easy to conclude that, as claimed,

\[
|T_{i_1,i_2}(c)| \leq \binom{m}{l} \binom{m}{k + l}.
\]

We continue with the proof of our theorem. As \( k = 7l, m = 8l^2 \) and \( l \geq 2 \), we get

\[
\binom{m}{k + l} / \binom{m}{l} = \binom{8l^2}{8l} / \binom{8l^2}{6l} \geq (\frac{4l^2}{8l})^{8l} / (8l^2)^{3l} \geq (2^{-9}l)^{2l}.
\]

Therefore

\[
|S_{k+l,k+l}| = \binom{m}{k+l} / \binom{m}{l} \geq (2^{-9}l)^{2l} \binom{m}{l} \binom{m}{k+l}.
\]

Consider any algorithm attempting to obtain a \( \rho n \)-best solution to the nonlinear optimization problem over any system \( S \), and let \( c^1, \ldots, c^q \in \mathbb{Z}^n \) be the sequence of queries to the oracle of \( S \) made by the algorithm. For each pair \( 1 \leq i_1, i_2 \leq l \) and each \( Z \in T_{i_1,i_2}(c^p) \), the number of \( Y \in S_{k+l,k+l} \) containing \( Z \), and hence satisfying \( Z \in S_Y \), is

\[
\binom{m - (k + i_1)}{l - i_1} \binom{m - (k + i_2)}{l - i_2} \leq \binom{m}{l}^2.
\]

So the number of \( Y \in S_{k+l,k+l} \) containing some \( Z \) that lies in some \( T_{i_1,i_2}(c^p) \) is at most

\[
\sum_{p=1}^{q} \sum_{i_1=1}^{l} \sum_{i_2=1}^{l} \binom{m}{l}^2 |T_{i_1,i_2}(c^p)| \leq ql^2 \binom{m}{l}^3 \binom{m}{k+l}.
\]

Therefore, if the number of oracle queries satisfies \( q < l^{-2}(2^{-9}l)^{2l} \), then there exists some \( Y \in S_{k+l,k+l} \) that does not contain any \( Z \) in any \( T_{i_1,i_2}(c^p) \). This means that any \( Z \in S_Y \) satisfies \( c^p \) attaining \( Z \) is \( \max\{c^p X : Y \in S^*_Y \} = \max\{c^p X : X \in S^*_Y \} \). Hence, whether the linear-optimization oracle presents \( S^* \) or \( S_Y \), on each query \( c^p \) it can reply with some \( X^p \in \mathcal{S}^* \) attaining

\[
c^p X^p = \max\{c^p X : X \in S^* \} = \max\{c^p X : X \in S_Y \}.
\]

So the algorithm cannot tell whether the oracle presents \( S^* \) or \( S_Y \), whether the image is \( WS^* \) or \( WS_Y \), and whether the objective-function value of every \( \rho n \)-best solution is zero or negative, let alone compute any \( \rho n \)-best solution. Therefore, with \( l \geq 2^{10} \), every algorithm that can produce a \( \rho n \)-best solution for the 2-dimensional nonlinear optimization problem (1) over every system \( S \) must make at least an exponential number

\[
q \geq l^{-2}(2^{-9}l)^{2l} \geq l^{-2}2^{2l} > 2^{l} = 2^{4\sqrt{l}}
\]

of queries to the oracle presenting \( S \) and therefore cannot run in polynomial time. \( \square \)
References


Jon Lee  
*IBM T.J. Watson Research Center, Yorktown Heights, USA*

Shmuel Onn  
*Technion - Israel Institute of Technology, Haifa, Israel*

Robert Weismantel  
*ETH, Zürich, Switzerland*