Optimizing Low Dimensional Functions over the Integers

Daniel Dadush, Arthur Léonard,
Lars Rohwedder, José Verschae

IPCO 2023

Inspiration

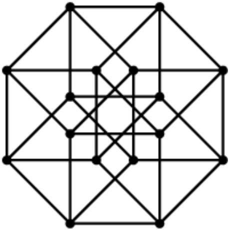
[Hunkenschröder, Pokutta, Weismantel, 2022]

$$\min_{x \in \{0,1\}^n} g(Wx)$$

$g : \mathbb{R}^m \to \mathbb{R}$ is a nice convex function

$W \in \mathbb{Z}^{n \times m}$ matrix (known or unknown),

$\|W\|_\infty \leq \Delta$, $m \ll n$.



Inspiration

[Hunkenschröder, Pokutta, Weismantel, 2022]

$$\min_{x \in \{0,1\}^n} g(Wx)$$

$g : \mathbb{R}^m \to \mathbb{R}$ is a nice convex function
$W \in \mathbb{Z}^{n \times m}$ matrix (known or unknown),
$\|W\|_\infty \leq \Delta$, $m \ll n$.

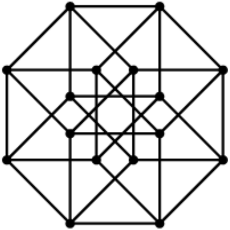
$\to$ The algorithm runs in time $O(n(m\Delta)^{O(m^2)})$ when W is known and g is separable convex.

The Problem

Our contribution : an algorithm to compute an optimal solution to :

$$\min \quad c^\top x + g(Wx)$$
$$l_i \leq x_i \leq u_i \text{ for all } i \in \{1, \ldots, n\}$$
$$x \in \mathbb{Z}^n$$

$W \in \mathbb{Z}^{m \times n}$, $\|W\|_\infty \leq \Delta$, $m \ll n$.



The Problem

Our contribution : an algorithm to compute an optimal solution to :

$$\begin{aligned} \min \quad & c^\top x + g(Wx) \\ & l_i \leq x_i \leq u_i \text{ for all } i \in \{1, \dots, n\} \\ & x \in \mathbb{Z}^n \end{aligned}$$

$W \in \mathbb{Z}^{m \times n}$, $\|W\|_\infty \leq \Delta$, $m \ll n$.

The complexity of this algorithm is :

$$O(n^{O(m)} \cdot (m\Delta)^{O(m^2)} \cdot Q)$$

Requirement on g

We suppose we can do **oracle queries**. Given :

- a partition of the variables $I \dot\cup J = [n]$ with $|I| \leq m$
- a fixing $z \in \mathbb{Z}^J$ of the J-variables

it solves the following problem in time Q :

$$
\begin{aligned}
\min \quad & c_I x + g(W_I x + W_J z) & \\
\text{s.t.} \quad & l_i \leq x_i \leq u_i & \text{for all } i \in I, \\
& x \in \mathbb{Z}^I &
\end{aligned}
$$

Requirement on g

We suppose we can do **oracle queries**. Given :

- a partition of the variables $I \dot\cup J = [n]$ with $|I| \leq m$
- a fixing $z \in \mathbb{Z}^J$ of the J-variables

it solves the following problem in time Q :

$$\begin{aligned} \min \quad & c_I x + g(W_I x + W_J z) & \\ \text{s.t.} \quad & l_i \leq x_i \leq u_i & \text{for all } i \in I, \\ & x \in \mathbb{Z}^I & \end{aligned}$$

Remark

If g is convex and accessible by function and gradient evaluation, oracle queries can be implemented using the algorithm in [Kannan, 1983] in $O(m^{O(m)} \langle \text{input} \rangle^{O(1)})$

Application : Mixed-Integer linear programming

$$\begin{aligned} \min \quad & c^\top x + d^\top y \\ & Wx + By = b \\ & l_i \leq x_i \leq u_i \text{ for all } i \in \{1, \ldots, n\} \\ & x \in \mathbb{Z}^n, y \in P \subset \mathbb{R}^h \end{aligned}$$

The polytope P imposes constraints on the continuous variables.

Application : Mixed-Integer linear programming

$$\begin{aligned}\min \quad & c^\top x + d^\top y \\ & Wx + By = b \\ & l_i \leq x_i \leq u_i \text{ for all } i \in \{1, \ldots, n\} \\ & x \in \mathbb{Z}^n, y \in P \subset \mathbb{R}^h\end{aligned}$$

The polytope P imposes constraints on the continuous variables.

Encoding

$$g(Wx) := \begin{cases} \min\{d^\top y : By = b - Wx, y \in P\} & \text{if it exists,} \\ \infty & \text{otherwise.}\end{cases}$$

Application : Variable-sized knapsack

$$\max \quad \left\{ \sum_{i=1}^{n} p_i x_i - g\left(\sum_{i=1}^{n} w_i x_i \right) \; : \; x_i \in \{0, \ldots, u_i\} \text{ for all } i \right\}$$

- p_i : the value of item i
- w_i : the space needed for item i
- x_i : the number of item i that we take
- u_i : the number of available item i

Application : Variable-sized knapsack

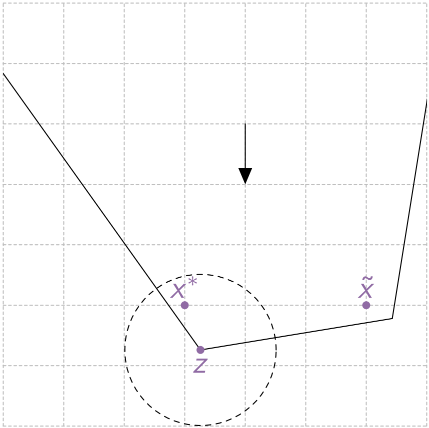
$$\max \quad \left\{ \sum_{i=1}^{n} p_i x_i - g\left(\sum_{i=1}^{n} w_i x_i\right) \ : \ x_i \in \{0, \ldots, u_i\} \text{ for all } i \right\}$$

- p_i : the value of item i
- w_i : the space needed for item i
- x_i : the number of item i that we take
- u_i : the number of available item i

If g is convex, our algorithm works in time $(n + w_{\max})^{O(1)}$.

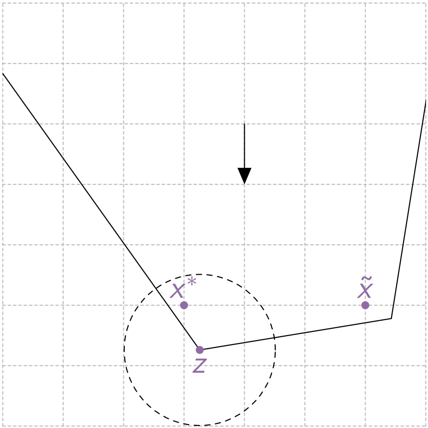
A proximity bound

$$\max\{c^\top x \ : \ Ax = b \text{ and } l_i \leq x_i \leq u_i \text{ for all } i\}$$



A proximity bound

$$\max\{c^\top x \ : \ Ax = b \text{ and } l_i \leq x_i \leq u_i \text{ for all } i\}$$



$$\|x^* - z\|_1 \leq m(2m\Delta + 1)^m$$

A proximity bound

Proximity lemma [Eisenbrand, Weismantel, 2020]

Let z be an optimal **vertex** solution to

$$\max\{c^\top x \ : \ Ax = b \text{ and } l_i \leq x_i \leq u_i \text{ for all } i\}$$

where $A \in \mathbb{Z}^{m \times n}$ has entries of size at most Δ.

If there exists an optimal integer solution $\tilde{x}$, then there exists an optimal integer solution x^* with :

$$\|x^* - z\|_1 \leq m(2m\Delta + 1)^m$$

A proximity bound

Proximity lemma [Eisenbrand, Weismantel, 2020]

Let z be an optimal **vertex** solution to

$$\max\{c^\top x \;:\; Ax = b \text{ and } l_i \leq x_i \leq u_i \text{ for all } i\}$$

where $A \in \mathbb{Z}^{m \times n}$ has entries of size at most Δ.

If there exists an optimal integer solution $\tilde{x}$, then there exists an optimal integer solution x^* with :

$$\|x^* - z\|_1 \leq m(2m\Delta + 1)^m$$

They give an algorithm solving the IP in time $O(n(m\Delta)^{O(m^2)})$

A proximity bound

Proximity lemma [Eisenbrand, Weismantel, 2020]

Let z be an optimal **vertex** solution to

$$\max\{c^\top x \ : \ Ax = b \text{ and } l_i \leq x_i \leq u_i \text{ for all } i\}$$

where $A \in \mathbb{Z}^{m\times n}$ has entries of size at most Δ.

If there exists an optimal integer solution $\tilde{x}$, then there exists an optimal integer solution x^* with :

$$\|x^* - z\|_1 \leq m(2m\Delta + 1)^m$$

If $u = \infty$, the running time is $O(n(m\Delta)^{O(m)})$.

Easier case

$$\begin{aligned} \min \quad & c^\top x + g(Wx) \\ & 0 \leq x_i \text{ for all } i \in \{1, \ldots, n\} \\ & x \in \mathbb{Z}^n \end{aligned}$$

Easier case

$$\begin{aligned} \min \quad & c^\top x + g(Wx) \\ & 0 \le x_i \text{ for all } i \in \{1, \ldots, n\} \\ & x \in \mathbb{Z}^n \end{aligned}$$

Let $\tilde{x}$ be an optimal solution, and $b^* = W\tilde{x}$.

Easier case

$$\begin{aligned} \min \quad & c^\top x + g(Wx) \\ & 0 \le x_i \text{ for all } i \in \{1, \dots, n\} \\ & x \in \mathbb{Z}^n \end{aligned}$$

Let $\tilde{x}$ be an optimal solution, and $b^* = W\tilde{x}$.

Let z be an optimal **vertex** solution of

$$\min\{c^\top x \ : \ Wx = b^*, x \in \mathbb{R}^n_{\ge 0}\}$$

Easier case

$$\begin{aligned} \min \quad & c^\top x + g(Wx) \\ & 0 \leq x_i \text{ for all } i \in \{1, \ldots, n\} \\ & x \in \mathbb{Z}^n \end{aligned}$$

Let $\tilde{x}$ be an optimal solution, and $b^* = W\tilde{x}$.

Let z be an optimal **vertex** solution of

$$\min\{c^\top x \ : \ Wx = b^*, x \in \mathbb{R}^n_{\geq 0}\}$$

Proximity lemma

$\exists$ an optimal integer solution x^* with $\|x^* - z\|_1 = O(m\Delta)^m$

Easier case

The vector z has at least $n - m$ zero components.

If T is the set of the zero components :

$$\|x_T^*\|_1 \leq \|x_T^* - z_T\|_1 \leq \|x^* - z\|_1 \leq O(m\Delta)^m$$

Since the entries of W are bounded by Δ :

$$\|W_T x_T^*\|_1 \leq m\Delta \|x_T\|_1 \leq O(m\Delta)^{m+1}$$

We guess T and $b^{(T)} := W_T x_T^*$ among the $n^m \times O(m\Delta)^{m(m+1)}$ choices.

Easier case

The problem is now divided into :

$$\min \quad \left\{ c_T^\top x_T \ : \ W_T x_T = b^{(T)} \text{ and } x_T \in \mathbb{Z}_{\geq 0}^{|T|} \right\}$$

and

$$\min \quad \left\{ c_L^\top x_L + g(W_L x^L + b^{(T)}) \ : \ x_L \in \mathbb{Z}_{\geq 0}^{|L|} \right\}$$

The first can be solved in $O(n(m\Delta)^m)$ using the algorithm of Eisenbrand and Weismantel.

The second corresponds to an **oracle query**.

Final running time : $O(n^{O(m)} \cdot (m\Delta)^{O(m^2)} \cdot Q)$

Dealing with upper bounds

$$
\begin{aligned}
\min \quad & c^\top x + g(Wx) \\
& l_i \leq x_i \leq u_i \text{ for all } i \in \{1, \dots, n\} \\
& x \in \mathbb{Z}^n
\end{aligned}
$$

Dealing with upper bounds

$$
\begin{aligned}
\min \quad & c^\top x + g(Wx) \\
& l_i \leq x_i \leq u_i \text{ for all } i \in \{1, \ldots, n\} \\
& x \in \mathbb{Z}^n
\end{aligned}
$$

Let $\tilde{x}$ be an optimal solution, and $b^* = W\tilde{x}$.

Dealing with upper bounds

$$\begin{aligned} \min \quad & c^\top x + g(Wx) \\ & l_i \leq x_i \leq u_i \text{ for all } i \in \{1, \ldots, n\} \\ & x \in \mathbb{Z}^n \end{aligned}$$

Let $\tilde{x}$ be an optimal solution, and $b^* = W\tilde{x}$.
Let z be an optimal **vertex** solution of

$$\min\{c^\top x \ : \ Wx = b^*, l \leq x \leq u\}$$

Dealing with upper bounds

$$\begin{aligned} \min \quad & c^\top x + g(Wx) \\ & l_i \le x_i \le u_i \text{ for all } i \in \{1, \ldots, n\} \\ & x \in \mathbb{Z}^n \end{aligned}$$

Let $\tilde{x}$ be an optimal solution, and $b^* = W\tilde{x}$.
Let z be an optimal **vertex** solution of

$$\min\{c^\top x \ : \ Wx = b^*, l \le x \le u\}$$

Proximity lemma

$\exists$ an optimal integer solution x^* with $\|x^* - z\|_1 = O(m\Delta)^m$

Dealing with upper bounds

Let z be an optimal vertex solution of

$$\min\{c^\top x \ : \ Wx = b^*, l \leq x \leq u\}$$

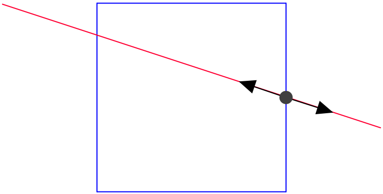
Dealing with upper bounds

Let z be an optimal vertex solution of

$$\min\{c^\top x \ : \ Wx = b^*, l \leq x \leq u\}$$

The vector z is tight on $n - m$ components :

How to guess if $z_i = l_i$ or u_i for these ?



Dealing with upper bounds

We deduced the value of z on $n - m$ components.

We use the same algorithm as in the un-upper-bounded case.

Final running time :

$$O(n^{O(m)} \cdot (m\Delta)^{O(m^2)} \cdot Q)$$

Extensions

When g is nice and separable convex, and W is unknown, [Hunkenschröder, Pokutta, Weismantel, 2022] give a $O(n(m\Delta)^{O(m^3)})$-time algorithm if we are given a value and gradient evaluation oracle for $x \mapsto g(Wx)$.

Extensions

When g is nice and separable convex, and W is unknown, [Hunkenschröder, Pokutta, Weismantel, 2022] give a $O(n(m\Delta)^{O(m^3)})$-time algorithm if we are given a value and gradient evaluation oracle for $x \mapsto g(Wx)$.

We improved it for any separable convex function in

$$O(n(m\Delta)^{O(m^2)}).$$

$\rightarrow$ Details in the paper !

Open Problems

Our contribution : an $O(n^{O(m)} \cdot (m\Delta)^{O(m^2)} \cdot Q)$-time algorithm to compute an optimal solution to :

$$\begin{aligned} \min \quad & c^\top x + g(Wx) \\ & l_i \leq x_i \leq u_i \text{ for all } i \in \{1, \ldots, n\} \\ & x \in \mathbb{Z}^n \end{aligned}$$

where $W \in \mathbb{Z}^{m \times n}$ has entries bounded by Δ in absolute value.

Open Problems

Our contribution : an $O(n^{O(m)} \cdot (m\Delta)^{O(m^2)} \cdot Q)$-time algorithm to compute an optimal solution to :

$$\begin{aligned} \min \quad & c^\top x + g(Wx) \\ & l_i \leq x_i \leq u_i \text{ for all } i \in \{1, \dots, n\} \\ & x \in \mathbb{Z}^n \end{aligned}$$

where $W \in \mathbb{Z}^{m \times n}$ has entries bounded by Δ in absolute value.

Open problems : Can we get rid of the $O(m)$ exponent ?

$\rightarrow$ Can be removed when $c = 0$ or $u = \infty$.

Open Problems

Our contribution : an $O(n^{O(m)} \cdot (m\Delta)^{O(m^2)} \cdot Q)$-time algorithm to compute an optimal solution to :

$$\begin{aligned} \min \quad & c^\top x + g(Wx) \\ & l_i \leq x_i \leq u_i \text{ for all } i \in \{1, \dots, n\} \\ & x \in \mathbb{Z}^n \end{aligned}$$

where $W \in \mathbb{Z}^{m \times n}$ has entries bounded by Δ in absolute value.

Open problems : Can we get rid of the $O(m)$ exponent ?

$\rightarrow$ Can be removed when $c = 0$ or $u = \infty$.

If $u = \infty$, Can we reduce the $O(m^2)$ exponent to $O(m)$?

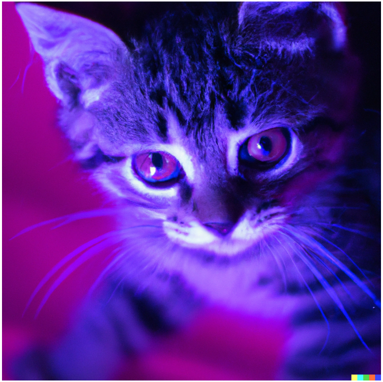
Questions ?

Thank you for your attention !

Questions ?

Thank you for your attention !

Make the kitten happy : ask a question !



(credit : Dall-E)

Empty slide

Application : Integer compressed sensing

$$\min \quad \{\|b - Wx\|_2 \ : \ x \in \{0, 1, \ldots, u_i\}^n, \|x\|_1 \leq \sigma\}$$

Dealing with upper bounds

Let z be an optimal vertex solution of

$$\min\{c^\top x \ : \ Wx = b^*, l \leq x \leq u\}$$

whose dual problem is :

$$\begin{aligned} \max \quad & {b^*}^\top y + l^\top s^l - u^\top s^u \\ & c - W^\top y = s^l - s^u \\ & s^l, s^u \in \mathbb{R}^n_{\geq 0} \\ & y \in \mathbb{R}^m \end{aligned}$$

Dealing with upper bounds

Let z be an optimal vertex solution of

$$\min\{c^\top x \ : \ Wx = b^*, l \leq x \leq u\}$$

whose dual problem is :

$$\begin{aligned} \max \quad & {b^*}^\top y + l^\top s^l - u^\top s^u \\ & c - W^\top y = s^l - s^u \\ & s^l, s^u \in \mathbb{R}^n_{\geq 0} \\ & y \in \mathbb{R}^m \end{aligned}$$

For an optimal vertex solution of the dual, $s^l_i = s^u_i = 0$ for at least m components such that $(W^\top)_i$ are linearly independent rows.

Dealing with upper bounds

$$\min\{c^\top x \;:\; Wx = b^*, l \le x \le u\}$$

$$\begin{aligned} \max \quad & {b^*}^\top y + l^\top s^l - u^\top s^u \\ & c - W^\top y = s^l - s^u \\ & s^l, s^u \in \mathbb{R}^n_{\ge 0} \\ & y \in \mathbb{R}^m \end{aligned}$$

By guessing these rows amongst $O(n^m)$ choices, we recover y exactly, thus $c - W^\top y$.

- If $(c - W^\top y)_i < 0$ then $s_i^u > 0$ and $z_i = u_i$
- If $(c - W^\top y)_i > 0$ then $s_i^l > 0$ and $z_i = l_i$