Continuous cutting plane algorithms

Didier Chételat

CERC Day
16/08/2022

ILPs

Imagine you would like to derive a lower bound for an ILP

$$\begin{aligned} \min \quad & c^t x \\ \text{s.t.} \quad & Ax \geq b \\ & x \geq 0, \quad x \in \mathbb{Z}^n \end{aligned}$$

One possibility would be to drop the integer constraints and compute the LP optimal value

$$\begin{aligned} \min \quad & c^t x \\ \text{s.t.} \quad & Ax \geq b \\ & x \geq 0 \end{aligned}$$

Chvátal-Gomory cutting planes

There is a way to improve this lower bound by the way of cutting planes. The simplest generic procedure available is Chvátal-Gomory cuts, which can be applied if A, b are all integer.

Write the LP in equality form,

$$\begin{aligned} \min \quad & c^t x \\ \text{s.t.} \quad & [A, -I] \begin{bmatrix} x \\ z \end{bmatrix} = b \\ & x, z \geq 0 \end{aligned}$$

and let denote by B the optimal basis matrix of the LP.

Chvátal-Gomory cutting planes

Then the *m* Chvátal-Gomory cuts are

$$(\lceil B^{-1} \rceil A - \lfloor B^{-1} A \rfloor) x \geq \lceil B^{-1} \rceil b - \lfloor B^{-1} b \rfloor$$

If we add these cuts to the LP, this yields a new LP with the same integer hull (same ILP),

$$\begin{aligned} \min \quad & c^t x \\ \text{s.t.} \quad & Ax \geq b \\ & (\lceil B^{-1} \rceil A - \lfloor B^{-1} A \rfloor) x \geq \lceil B^{-1} \rceil b - \lfloor B^{-1} b \rfloor \\ & x \geq 0 \end{aligned}$$

but better lower bound.

Chvátal-Gomory cutting planes

Okay, now I would like to point out two things. First, we can write this more compactly as follows. Define the function $f_{B^{-1}}(y) : \mathbb{R}^m \to \mathbb{R}^{2m}$,

$$f_{B^{-1}}(y) = [y, \lceil B^{-1} \rceil y - \lfloor B^{-1} y \rfloor]$$

Then the LP can be written more compactly as

$$\begin{aligned} \min \quad & c^t x \\ \text{s.t.} \quad & f_{B^{-1}}(A) x \geq f_{B^{-1}}(b) \\ & x \geq 0 \end{aligned}$$

(Notation throughout:
$f(A) = [f(A_1), \ldots, f(A_n)]$)

Chvátal-Gomory valid inequalities

The second point is that we can generalize this. The inequalities

$$(\lceil W \rceil A - \lfloor W A \rfloor) x \geq \lceil W \rceil b - \lfloor W b \rfloor$$

are in fact "valid inequalities" for any W - that is, are guaranteed to not cut the integer hull. In other words, solving

$$\begin{array}{ll} \min & c^t x \\ \text{s.t.} & f_W(A) x \geq f_W(b) \\ & x \geq 0 \end{array}$$

yields a valid lower bound for the ILP for *any* W.
You recover the "classical" Chvátal-Gomory cuts by taking $W=B^{-1}$.

Higher order valid inequalities

Because the inequalities $f_W(A)x \geq f_W(b)$ are valid, the ILP

$$\begin{aligned} \min \quad & c^t x \\ \text{s.t.} \quad & f_W(A)x \geq f_W(b) \\ & x \geq 0, \quad x \in \mathbb{Z}^n \end{aligned}$$

is equivalent to the original ILP. But we can apply Chvátal-Gomory cuts to it too: yielding the even tighter ILP

$$\begin{aligned} \min \quad & c^t x \\ \text{s.t.} \quad & f_{W_2} \circ f_{W_1}(A)x \geq f_{W_2} \circ f_{W_1}(b) \\ & x \geq 0, \quad x \in \mathbb{Z}^n \end{aligned}$$

Higher order valid inequalities

In fact, we can repeat this for as many rounds as we like, say, K rounds

$$\begin{aligned} \min \quad & c^t x \\ \text{s.t.} \quad & f_{W_K} \circ \cdots \circ f_{W_1}(A)x \geq f_{W_K} \circ \cdots \circ f_{W_1}(b) \\ & x \geq 0, \quad x \in \mathbb{Z}^n \end{aligned}$$

The ``classical" Chvátal-Gomory cuts then correspond to taking $W_1=B^{-1}$; adding back the cuts, solving the LP, obtaining a basis matrix B_1 and taking $W_2=B_1^{-1}$; etc. That is, the classical cuts correspond to a specific sequence of weights, $W_1=B^{-1}$, $W_2=B_1^{-1}$, ..., $W_K=B_{K-1}^{-1}$.

Continuous cuts optimization

Continuous cuts optimization

The classical weights (B^{-1}, B_1^{-1}, ..., B_{K-1}^{-1}) are good because they yield a high lower bound

$$\mathrm{LP}(A, b, c;\ W_1, \ldots, W_K)$$

$$= \begin{array}{ll} \min & c^t x \\ \text{s.t.} & f_{W_K} \circ \cdots \circ f_{W_1}(A)x \geq f_{W_K} \circ \cdots \circ f_{W_1}(b) \\ & x \geq 0, \quad x \in \mathbb{Z}^n \end{array}$$

But are they optimal? No!

Continuous cuts optimization

How about then trying to optimize the weights?

$$\max_{W_1,\ldots,W_K} \; \mathrm{LP}(A, b, c;\; W_1, \ldots, W_K)$$

⇨ Continuous optimization problem.

Algorithm

Repeat

1. Solve $\mathrm{LP}(A, b, c;\ W_1, \ldots, W_K)$ and obtain solution x*;

2. Take gradients steps to minimize the violations of each cut at x*

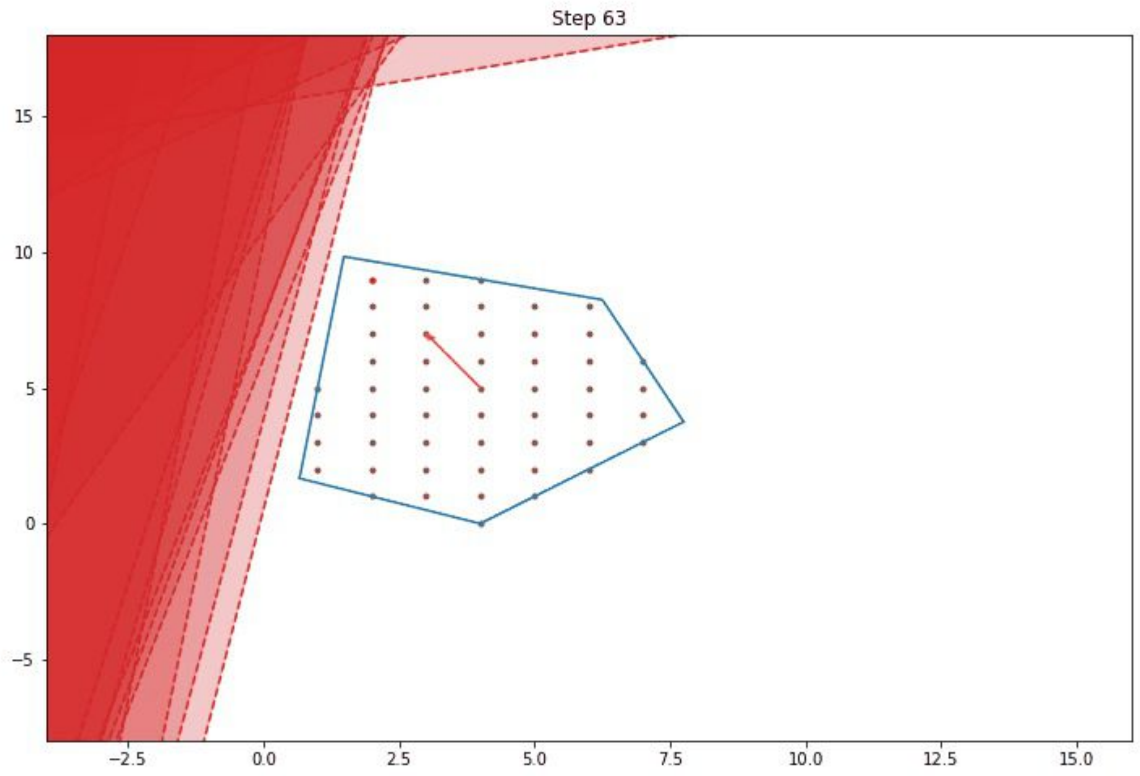
$$\mathcal{L} = \sum_i \left[f_{W_K} \circ \cdots \circ f_{W_1}(A)_i x^* - f_{W_K} \circ \cdots \circ f_{W_1}(b)_i \right]$$

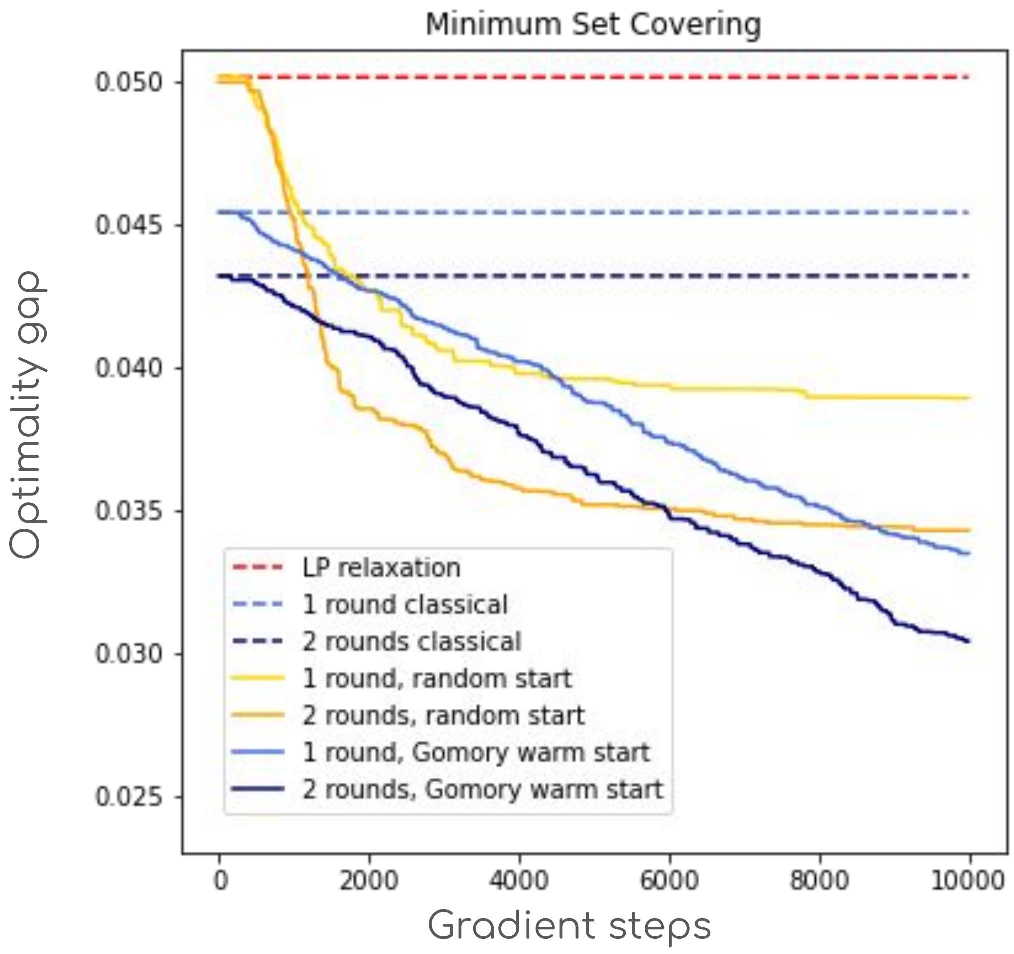
until

$$f_{W_K} \circ \cdots \circ f_{W_1}(A)_i x^* < f_{W_K} \circ \cdots \circ f_{W_1}(b)_i$$

for some *i* (i.e. until some inequality cuts off x*).

32 rank-1 GMI inequalities, randomly initialized





1000 variables
500 constraints

Subadditive neural networks

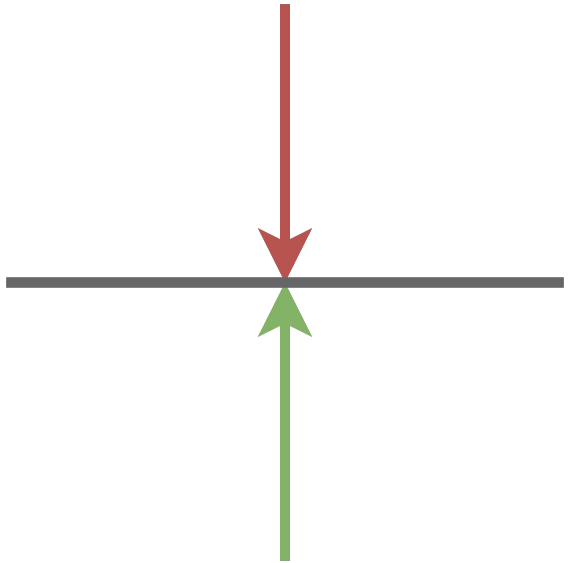
LP duality

As is well known, every LP

$$\begin{aligned} \min \quad & c^t x \\ \text{s.t.} \quad & Ax = b \\ & x \geq 0 \end{aligned}$$

has an associated equivalent dual LP

$$\begin{aligned} \max \quad & w^t b \\ \text{s.t.} \quad & w^t A \leq c \end{aligned}$$



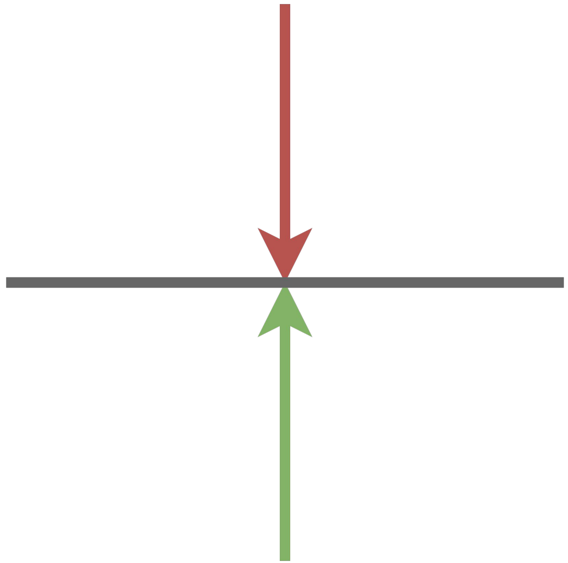
ILP duality

Every ILP

$$\begin{aligned} \min \quad & c^t x \\ \text{s.t.} \quad & Ax = b \\ & x \geq 0, \quad x \in \mathbb{Z}^n \end{aligned}$$

has an an associated equivalent infinite-dimensional, “continuous” problem

$$\begin{aligned} \max \quad & f(b) \\ \text{s.t.} \quad & f(A) \leq c \\ & f \text{ is subadditive, non-decreasing} \end{aligned}$$



Deep dual optimization

Let's go back to our cuts optimization. By LP duality,

$$\begin{array}{ll} \min & c^t x \\ \text{s.t.} & f_{W_K} \circ \cdots \circ f_{W_1}(A)x \\ & \quad \geq f_{W_K} \circ \cdots \circ f_{W_1}(b) \\ & x \geq 0 \end{array} = \begin{array}{ll} \max\limits_{w} & w^t f_{W_K} \circ \cdots \circ f_{W_1}(b) \\ \text{s.t.} & w^t f_{W_K} \circ \cdots \circ f_{W_1}(A) \leq c \\ & w \geq 0 \end{array}$$

So define

$$f_{w,W_1,\ldots,W_K}(y) \equiv w^t f_{W_K} \circ \cdots \circ f_{W_1}(y)$$

Deep dual optimization

Hence we really have

$$\max_{W_1,\ldots,W_K} \mathrm{LP}(A, b, c;\ W_1, \ldots, W_K)$$

$$= \begin{array}{cl} \displaystyle\max_{w \geq 0, W_1, \ldots, W_K} & f_{w, W_1, \ldots, W_K}(b) \\ \text{s.t.} & f_{w, W_1, \ldots, W_K}(A) \leq c \end{array}$$

Looks suspiciously like the subadditive dual!

$$\begin{array}{cl} \max & f(b) \\ \text{s.t.} & f(A) \leq c \\ & f \text{ is subadditive, non-decreasing} \end{array}$$

Deep dual optimization

Theorem: For any $(w \geq 0, W_1, \ldots, W_K)$, $f_{w,W_1,\ldots,W_K}(y)$ is subadditive and non-decreasing.

Proof: $f_{w,W_1,\ldots,W_K}(y) \equiv w^t f_{W_K} \circ \cdots \circ f_{W_1}(y)$, and each function is subadditive and non-decreasing.

Deep dual optimization

So continuous optimization of Chvátal-Gomory valid inequalities

= Solving the subadditive dual with a ``subadditive neural net"

$$f_{w,W_1,...,W_K}(y) \equiv w^t f_{W_K} \circ \cdots \circ f_{W_1}(y)$$

composed of ``Chvátal-Gomory" layers

$$f_W(y) = [y, \lceil W \rceil y - \lfloor W y \rfloor]$$

(except the last one, which is a linear layer.)

Deep dual optimization

Moreover:

- Classical Chvátal-Gomory cuts = ``layer-by-layer`` greedy training of the neural net

- Our approach = end-to-end training (all layers trained simultaneously)

Other layers

- Chvátal-Gomory cuts:

$$f_W(y) = [y, \lceil W \rceil y - \lfloor W y \rfloor]$$

- Gomory-mixed-integer cuts:

$$f_{W,v}(y) = \left[y, \min\left(\frac{\{Wy\}}{\{v\}}, \frac{1-\{Wy\}}{1-\{v\}}\right) + \max\left(\frac{-W}{\{v\}}, \frac{W}{1-\{v\}}\right) y\right]$$

- In general, any <u>subadditive</u> and <u>non-decreasing</u> layer would work (other cuts?)

Thank you!