Solving Linear Programs with Fast Online Learning Algorithms

ICML 2023

Wenzhi Gao[1], Dongdong Ge[1], Chunlin Sun[2], Yinyu Ye[2]

July 1, 2023

[1]Shanghai University of Finance and Economics
[2]Stanford University

Problem of interest

We want solve an LP with m constraints and n variables

$$\begin{aligned} \max_{\mathbf{x}} \quad & \langle \mathbf{c}, \mathbf{x} \rangle \\ \text{subject to} \quad & \mathbf{A}\mathbf{x} \leq \mathbf{b} \\ & \mathbf{0} \leq \mathbf{x} \leq \mathbf{1} \end{aligned}$$

- The formulation covers many types of LPs
  including set covering, set packing, resource allocation...
- Seeking to approximately optimize it with help of **online algorithms**

Dual problem and optimality condition

The dual problem is given by

$$\begin{aligned} \min_{\mathbf{y},\mathbf{s}} \quad & \langle \mathbf{b}, \mathbf{y} \rangle + \langle \mathbf{1}, \mathbf{s} \rangle \\ \text{subject to} \quad & \mathbf{s} \geq \mathbf{c} - \mathbf{A}^\top \mathbf{y} \\ & \mathbf{y}, \mathbf{s} \geq \mathbf{0} \end{aligned}$$

and optimality condition exhibits a strong relation between primal and dual.

$$x_j^* \in \begin{cases} \{0\}, & c_j < \langle \mathbf{a}_j, \mathbf{y}^* \rangle \\ [0,1], & c_j = \langle \mathbf{a}_j, \mathbf{y}^* \rangle \\ \{1\}, & c_j > \langle \mathbf{a}_j, \mathbf{y}^* \rangle \end{cases}$$

- Primal largely determined by the dual through the optimality condition
- We adopt a dual-based approach

Dual reformulation and online algorithm

The dual problem can be reformulated as a finite-sum problem

$$\min_{\mathbf{y}\geq\mathbf{0}} \ \frac{1}{n}\sum_{j=1}^{n}\langle\mathbf{d},\mathbf{y}\rangle + [c_j - \langle\mathbf{a}_j,\mathbf{y}\rangle]_+ =: \frac{1}{n}\sum_{j=1}^{n} f(\mathbf{y},j), \quad (\mathbf{d}=\mathbf{b}/n)$$

We propose two popular online algorithms to optimize the dual problem

Explicit online subgradient applies projected subgradient method

$$\mathbf{y}^{k+1} = \arg\min_{\mathbf{y}\geq\mathbf{0}}\left\{\left\langle f'\left(\mathbf{y}^k,k\right),\mathbf{y}-\mathbf{y}^k\right\rangle + \frac{1}{2\gamma}\left\|\mathbf{y}-\mathbf{y}^k\right\|^2\right\}$$

Implicit online proximal point applies proximal point method

$$\mathbf{y}^{k+1} = \arg\min_{\mathbf{y}\geq\mathbf{0}}\left\{f\left(\mathbf{y},k\right) + \frac{1}{2\gamma}\left\|\mathbf{y}-\mathbf{y}^k\right\|^2\right\}$$

- Iterate through a **permutation** of columns in an "online fashion".
- Primal value x^k is obtained using $\mathbf{y}^k, \mathbf{y}^{k+1}$ and the optimality condition

Theoretical guarantees of online algorithms

Theorem

Let $\hat{\mathbf{x}}$ *be the output of the online algorithm, then*

$$\mathbb{E}[\langle \mathbf{c}, \hat{\mathbf{x}}^* \rangle - \langle \mathbf{c}, \hat{\mathbf{x}} \rangle] \leq \mathcal{O}\Big(\frac{m \log n}{K} + \sqrt{\frac{n}{K}} \log n + \sqrt{\frac{mn}{K}}\Big)$$

$$\mathbb{E}[\|[\mathbf{A}\hat{\mathbf{x}} - \mathbf{b}]_+\|] \leq \mathcal{O}\Big(\frac{m}{K} + \sqrt{\frac{mn}{K}}\Big).$$

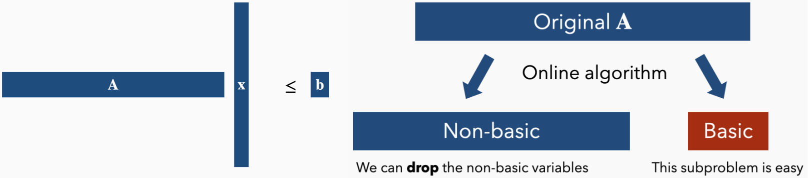
after we go though the permutation of LP columns K **times**.

- Sub-linear optimality and constraint violation bound
- We can improve the bound by $\sqrt{K}$ by copying each LP variable K times
- Implementation can be very cheap if we exploit sparsity of $\mathbf{A}$.

Application: LP Sifting/Column generation

Our method can help solving large-scale LPs with column generation.

Given an LP with $n \gg m$, column generation



- finds a subset of variables **likely** to take > 0 (basic) in $\mathbf{x}^*$
- primal output $\hat{\mathbf{x}}$ of online algorithm can be used as a likelihood estimate
- the dual solution can also be used for dual stabilization

Numerical results: approximate LP solver

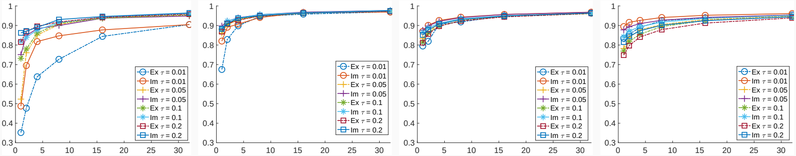


Figure 1: Left to right: LP of different sizes. The x-axis represents parameter K (how many times we copy each LP column) ranging in $\{1, 2, 4, 8, 16, 32\}$. The y-axis shows the relative optimality.

Numerical results: LP Sifting

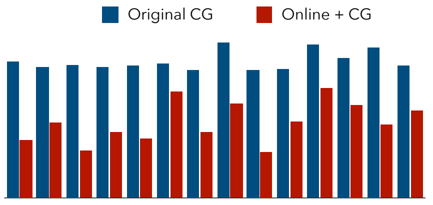
We choose LP datasets with more variables than rows

Table 1: Datasets collected from large-scale LP benchmark

Dataset	Row	Column	Dataset	Row	Column
rail507	507	6.4e+04	rail516	516	4.8e+04
rail582	582	5.6e+04	rail2586	2586	9.2e+05
rail4284	4284	1.0e+06	scpm1	5000	5.0e+05
scpn2	5000	1.0e+06	scpl4	2000	2.0e+05
scpj4scip	1000	1.0e+05	scpk4	2000	1.0e+05
s82	87878	1.7e+06	s100	14733	3.6e+05
s250r10	10962	2.7e+05	-	-	-

Numerical results: LP Column generation

Data	Acc	Data	Acc
scpm1	100%	rail507	90%
scpn2	100%	rail516	88%
scpl4	100%	rail2586	94%
scpk4	100%	rail4284	96%



- Online algorithm achieves high accuracy in basis status identification.
- LP column generation benefits from it

Summary

- We propose to apply online algorithms to solve LPs
- Online algorithms can approximately solve LPs
- Online algorithms also benefits exact LP solving

Thanks for listening !