Approximation algorithms for the covering-type k-violation linear program

Yotaro Takazawa Shinji Mizuno Tomonari Kitahara

January 6, 2018

Abstract

We study the covering-type k-violation linear program where at most k of the constraints can be violated. This problem is formulated as a mixed integer program and known to be strongly NP-hard. In this paper, we present a simple $(k+1)$-approximation algorithm using a natural LP relaxation. We also show that the integrality gap of the LP relaxation is $k+1$. This implies we can not get better approximation algorithms using the LP-relaxation as a lower bound of the optimal value.

1 Introduction

In a usual optimization problem, we need satisfy all the constraints in the problem. On the other hand, requirement about constraints can be sometimes soft when we model problems. From this context, a model, where we do not need satisfy a specified number of constraints, is considered in many optimization problems such as linear programming [1, 8] , facility location [2], clustering [3] and set covering problems [5, 7]. In this paper, as a special case of the LP where we can violate at most k of the constraints, we study the covering-type k-violation linear program (CKVLP) which is formulated as a mixed-integer program (MIP) as follows.

$$\begin{array}{lll} \min & \sum_{j \in N} c_j x_j & \\ \text{s.t.} & \sum_{j \in N} a_{ij} x_j + b_i z_i \geq b_i & \forall i \in M, \\ & \sum_{i \in M} z_i \leq k, & \\ & x_j \geq 0 & \forall j \in N, \\ & z_i \in \{0,1\} & \forall i \in M. \end{array} \tag{1}$$

where $c_j \geq 0$ $(j \in N)$, $a_{ij} \geq 0$ $(i \in M, j \in N)$, $b_i \geq 0$ $(i \in M)$ and $k \in \{0, 1, \ldots, m-1\}$.

CKVLP has applications such as a probabilistically constrained portfolio optimization problem, see Qiu et al. [6] and references therein. When k or $m-k$ are constant, we can get an optimal solution in polynomial time. In general case, Qiu et al. show that CKVLP is strongly NP-hard and it is hard to solve some realistic instances by CPLEX, and they develop an efficient exact algorithm based on MIP. Dinitz and Gupta [4] present two approximation algorithms which are based on a reduction to a problem similar to the set cover problem and use a Lagrangian relaxation approach. Their approximation ratios are $O(f \log f)$ and $O(\log \Delta \log f)$ where $f = \max_{i \in M} |\{j \in N \mid a_{ij} > 0\}|$ and $\Delta = \max_{j \in N} |\{i \in M \mid a_{ij} > 0\}|$. Takazawa et al. [9] present a primal-dual $\max\{f,\ k+1\}$-approximation algorithm for the partial covering 0–1 integer program, where constraints $x_j \in \{0,1\}$ $(\forall j \in N)$ are added to CKVLP.

In this paper, we present an LP rounding $(k+1)$-approximation algorithms for CKVLP. We also show that the integerality gap of the LP is $k+1$. This implies we can not get a better approximation algorithm as long as we use an LP-relaxation as a lower bound of the optimal value [10]. We also present a simple $(m-k)$-approximation algorithm.

2 A $(k+1)$-approximation algorithm

In this section, we present a $(k+1)$-approximation algorithm for CVKLP based on a natural LP relaxation and show this approximation ratio matches the integrality gap of the LP. A straightforward LP relaxation of (1) is expressed as follows.

$$\begin{array}{lll} \min & \sum_{j \in N} c_j x_j & \\ \text{s.t.} & \sum_{j \in N} a_{ij} x_j + b_i z_i \geq b_i & \forall i \in M, \\ & \sum_{i \in M} z_i \leq k & \\ & x_j \geq 0 & \forall j \in N, \\ & 1 \geq z_i \geq 0 & \forall i \in M. \end{array} \tag{2}$$

In our algorithm, we first solve the LP relaxation problem and obtain an optimal LP solution $(\boldsymbol{x}^L, \boldsymbol{z}^L)$. Then, by rounding and scaling the LP solution $(\boldsymbol{x}^L, \boldsymbol{z}^L)$, we obtain a MIP solution $(\boldsymbol{x}^M, \boldsymbol{z}^M)$. The algorithm is presented as follows.

Algorithm 1

Step 1: Solve the LP relaxation problem (2) and let an optimal solution for (2) be $(\boldsymbol{x}^L, \boldsymbol{z}^L)$.

Step 2: Let $(i_1, \ldots, i_m)$ be a permutation of M such that

$$z^L_{i_1} \geq z^L_{i_2} \geq \ldots \geq z^L_{i_m}$$

and set

$$\delta = \frac{1}{1 - z^L_{i_{k+1}}}.$$

Step 3: Set the solution $(\boldsymbol{x}^M, \boldsymbol{z}^M)$ as follows.

$$\boldsymbol{x}^M = \delta \boldsymbol{x}^L$$

and

$$z^M_i = \begin{cases} 1 & \text{if } i \in \{i_1 \ldots i_k\}, \\ 0 & \text{if } i \in \{i_{k+1} \ldots i_m\}. \end{cases}$$

Output $(\boldsymbol{x}^M, \boldsymbol{z}^M)$.

Lemma 1. *$(\boldsymbol{x}^M, \boldsymbol{z}^M)$ is a feasible solution of* (1).

Proof. It suffices to show $\sum_{j\in N} a_{ij}x_j^M \geq b_i$ for all $i \in \{i_{k+1} \dots i_m\}$. Fix any $i \in \{i_{k+1} \dots i_m\}$. If $z_i^L = 0$, then we have that

$$\sum_{j\in N} a_{ij}x_j^M = \delta \sum_{j\in N} a_{ij}x_j^L \geq \delta b_i \geq b_i.$$

Let us consider the case when $z_i^L > 0$. From Step 2 we obtain that

$$\sum_{j\in N} a_{ij}x_j^M = \delta \sum_{j\in N} a_{ij}x_j^L = \frac{1}{1 - z_{i_{k+1}}^L} \sum_{j\in N} a_{ij}x_j^L.$$

Since $(\boldsymbol{x}^L, \boldsymbol{z}^L)$ is a feasible solution of (2), $\sum_{j\in N} a_{ij}x_j^L \geq (1 - z_i^L)b_i$ holds. Therefore we get

$$\sum_{j\in N} a_{ij}x_j^M = \frac{1}{1 - z_{i_{k+1}}^L} \sum_{j\in N} a_{ij}x_j^L \leq \frac{1 - z_i^L}{1 - z_{i_{k+1}}^L} b_i \geq b_i.$$

□

Theorem 1. *$(\boldsymbol{x}^M, \boldsymbol{z}^M)$ is a feasible solution of* (1) *such that*

$$\sum_{j\in N} c_j x_j^M \leq \delta OPT \leq (k+1)OPT,$$

where OPT is the optimal value of (1).

Proof. Let OPT_{LP} be the optimal value of (2). Then we have that

$$\sum_{j\in N} c_j x_j^M = \delta OPT_{LP} \leq \delta OPT.$$

Hence, it suffices to show that $\delta \leq k+1$. Suppose that $z_{i_{k+1}}^L > \frac{k}{k+1}$. Then, we obtain that

$$\sum_{i=1}^{m} z_i^L \geq \sum_{\ell=1}^{k+1} z_{i_\ell}^L > (k+1)\frac{k}{k+1} = k.$$

since $z_{i_1}^L \geq \dots \geq z_{i_{k+1}}^L$ holds from Step 2. Then $(\boldsymbol{x}^L, \boldsymbol{z}^L)$ is infeasible to (2) and this is a contradiction. Therefore, we have that $z_{i_{k+1}}^L \leq \frac{k}{k+1}$ and

$$\delta = \frac{1}{1 - z_{i_{k+1}}^L} \leq \frac{1}{1 - \frac{k}{k+1}} = k+1.$$

□

Next, we show that the integrality gap of the LP relxation (2) is $k+1$. The integrality gap of the relaxation problem (2) is defined as the supremum of the ratio of the optimal values of (1) and (2) for all instances [10] . From the proof of Theorem 1, we have that

$$OPT \leq \sum_{j\in N} c_j x_j^M \leq (k+1)OPT_{LP}.$$

Therefore, we can say that the integrality gap is bounded above by $k+1$. Then, we show that this bound is tight by introducing the following instance.

$$\begin{array}{lll} \min & x_1 & \\ \text{s.t.} & x_1 + z_i \geq 1 & \forall i \in \{1, \ldots, k+1\}, \\ & x_1 + z_i \geq \frac{1}{k+1} & \forall i \in \{k+2, \ldots, m\}, \\ & \sum_{i=1}^{m} z_i \leq k & \\ & x_1 \geq 0, & \\ & z_i \in \{0, 1\} & \forall i \in \{1, \ldots, m\}, \end{array} \tag{3}$$

where $m \geq 2$ and $k \in \{1, \ldots, m-1\}$. When we set $x_1 = \frac{1}{k+1}$, we can not satisfy $m-k$ constraints for any $\boldsymbol{z}$. Hence, an optimal solution of this problem is $(x_1^*, \boldsymbol{z}^*) = (1, \mathbf{0})$. The objective value of $(x_1^*, \boldsymbol{z}^*)$ is 1. Now consider an LP relaxation of (3) and a feasible solution $(x_1^L, \boldsymbol{z}^L)$ of the LP such that $x_1^L = \frac{1}{k+1}$ and

$$z_i^L = \begin{cases} \frac{k}{k+1} & \text{if } i \in \{1, \ldots, k+1\}, \\ 0 & \text{if } i \in \{k+2, \ldots, m\}. \end{cases}$$

The objective value of $(x_1^L, \boldsymbol{z}^L)$ is $\frac{1}{k+1}$. Thus, the optimal value of (3) is at least $k+1$ times that of its LP relaxation. From this observation and Theorem 1, the integrality gap of the problem (2) is $k+1$.

3 Discussion

When k is big, the performance of the algorithm gets worse. However, since our approximation ratio matches the integrality gap, in order to get a better approximation algorithm, we need a strong LP relaxation of CVKLP or other approaches. On the other hand, when k is close to m, we can easily get a better approximation algorithm, that is, an $(m-k)$-approximation algorithm as follows.

Algorithm 2

Step 1: For all $i \in M$ solve the problem, $\min \sum_{j \in N} c_j x_j$ such that $\sum_{j \in N} a_{ij} x_j \geq b_i$ and $x_j \geq 0$ $(j \in N)$. Let $\boldsymbol{x}^i$ be an optimal solution and $d^i = \sum_{j \in N} c_j x_j^i$ be the otpimal value.

Step 2: Let $(i_1, \ldots, i_m)$ be a permutation of M such that

$$d_{i_1} \leq d_{i_2} \leq \ldots \leq d_{i_m}.$$

Step 3: Set the solution $(\boldsymbol{x}^M, \boldsymbol{z}^M)$ as follows.

$$\boldsymbol{x}^M = \sum_{\ell=1}^{m-k} \boldsymbol{x}^{i_\ell}$$

and

$$z_i^M = \begin{cases} 0 & \text{if } i \in \{i_1 \ldots i_{m-k}\}, \\ 1 & \text{if } i \in \{i_{m-k+1} \ldots i_m\}. \end{cases}$$

Output $(\boldsymbol{x}^M, \boldsymbol{z}^M)$.

Theorem 2. *Algorithm 2 is an $(m-k)$-approximation algorithm CKVLP.*

Proof. Let $(\boldsymbol{x}^M, \boldsymbol{z}^M)$ be the output of Algorithm 2. Then, $(\boldsymbol{x}^M, \boldsymbol{z}^M)$ is an feasible solution of (1) since it crealy satisfies $m-k$ constraints. Suppose $d_{i_{m-k}}$ is greater than the optimal value ($= OPT$). Then, an optimal solution satisfies at least $m-k$ constraints but the objective value is less than $d_{i_{m-k}}$. This is the contradiction with the definition of $d_{i_{m-k}}$. We have that

$$\sum_{j \in N} c_j x_j^M = \sum_{\ell=1}^{m-k} d_{i_\ell} \leq (m-k) d_{i_{m-k}} \leq (m-k) OPT.$$

□

Acknowledgment

This research is supported in part by Grant-in-Aid for Science Research (A) 26242027 of Japan Society for the Promotion of Science and Grant-in-Aid for Young Scientist (B) 15K15941.

References

[1] Timothy M Chan. Low-dimensional linear programming with violations. *SIAM Journal on Computing*, 34(4):879–893, 2005.

[2] Moses Charikar, Samir Khuller, David M Mount, and Giri Narasimhan. Algorithms for facility location problems with outliers. In *Proceedings of the twelfth annual ACM-SIAM symposium on Discrete algorithms*, pages 642–651. Society for Industrial and Applied Mathematics, 2001.

[3] Ke Chen. A constant factor approximation algorithm for k-median clustering with outliers. In *Proceedings of the twelfth annual ACM-SIAM symposium on Discrete algorithms*, volume 8, pages 826–835, 2008.

[4] Michael Dinitz and Anupam Gupta. *Packing interdiction and partial covering problems*, pages 157–168. Springer Berlin Heidelberg, Berlin, Heidelberg, 2013.

[5] Rajiv Gandhi, Samir Khuller, and Aravind Srinivasan. Approximation algorithms for partial covering problems. *Journal of Algorithms*, 53(1):55–84, 2004.

[6] Feng Qiu, Shabbir Ahmed, Santanu S Dey, and Laurence A Wolsey. Covering linear programming with violations. *INFORMS Journal on Computing*, 26(3):531–546, 2014.

[7] Yingli Ran, Yishuo Shi, and Zhao Zhang. Local ratio method on partial set multi-cover. *Journal of Combinatorial Optimization*, 34(1):302–313, 2017.

[8] Thomas Roos and Peter Widmayer. k-violation linear programming. *Information Processing Letters*, 52(2):109–114, 1994.

[9] Yotaro Takazawa, Shinji Mizuno, and Tomonari Kitahara. An approximation algorithm for the partial covering 0–1 integer program. *Discrete applied mathematics.*

[10] Vijay V Vazirani. Introduction to lp-duality. In *Approximation Algorithms*, pages 93–107. Springer, 2003.