The Integrality Gap of Points and Lines

Raymond Chen

November 19 2019

Introduction

Consider the following problem: Given a set $X = \{x_1, x_2, ..., x_n\}$ of points in the 2D plane, and a set $Y = \{y_1, y_2, ..., y_m\}$ of lines, find the minimal number of lines in Y such that every point in X is contained in at least one line of the solution. Note that for any pair (x_i, x_j), the line passing through the two is not necessarily in Y.

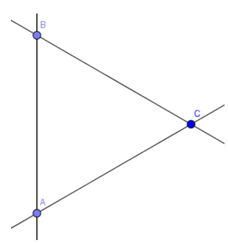
This problem is a specific instance of Set-Coverage, and is still NP-hard. For this paper I would like to understand the integrality gap of the standard set-coverage linear programming formulation for this problem. More specifically, I will show that for any constant $C > 0$, the integrality gap is greater than that constant. (The integrality gap is the supremum of $\frac{OPT}{LP}$ across all instances of Points and Lines).

Let $x_i \in y_j$ symbolize point x_i touching line y_j. Moreover, let y_j be a decision variable between 0 and 1 for each line.

LP:

$$min \Sigma_{i=1}^{m} y_i$$
$$\forall x \in X, \Sigma_{y_i \in Y \ s.t. \ x \in y_i} y_i \geq 1$$

As a very basic example to show that this LP is not always exact, consider the following:



Then the optimal solution requires two out of the three possible lines, but by assigning $\frac{1}{2}$ to each line, we can acheive an LP solution of just $\frac{3}{2}$. So we know that the integrality gap is at least $\frac{2}{\frac{3}{2}} = \frac{4}{3}$

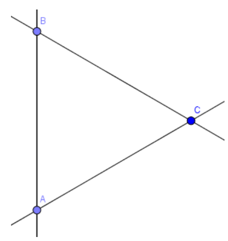
Hitting Set

We first show that the integrality gap of a related problem, covering lines by points, for the standard formulation, is not upper bounded by a constant. In covering lines by points, given a set $Y = \{y_1, y_2, ...y_m\}$ of lines, $X = \{x_1, x_2, ..., x_n\}$ of points, we want to minimize the number of points needed such that every line in Y contains at least one point of our solution.

Let x_i represent our decision variable between 0 and 1 for every point.
LP:

$$min \Sigma_{i=1}^{n} x_i$$
$$\forall y \in Y, \Sigma_{x_i \in X \ s.t. \ x_i \in y} x_i \geq 1$$

Looking at this figure again:



We see that OPT requires at least two points, but for LP we can assign $\frac{1}{2}$ to every point to get a value of $\frac{3}{2}$. So OPT is at least $\frac{4}{3}$.

In fact, Alon's[1] paper has a construction that shows that this gap is not upper bounded by any constant.

In order to see this example we introduce the following from Alon's paper:

Let $[k]^d$ be the set of all d-dimensional points with integer coordinates between 1 and k.

A combinatorial line is a set of k points, a set I $\subset \{1, 2, ..., d\}$ and corresponding $a_{i_1}, a_{i_2}, ..., a_{i_j}$ of integer values between 1 and k:
  We fix coordinates $a_{i_1}, ..., a_{i_j}$ for all points in the combinatorial line
  In all other positions we have $a_{j_1} = a_{j_2} = ...a_{j_n} = e$ for the eth point in the combinatorial line.

For example, a combinatorial line in $[4]^4$ could look like:
  (1, 2, 1, 1)
  (2, 2, 2, 2)
  (3, 2, 2, 2)
  (4, 2, 4, 4)

Density Hales-Jewett Theorem: Given k, δ, there exists $d_0 = d_0(k, \delta)$ such that $\forall d \geq d_0$, any set of at least $\delta * k^d$ points contains a combinatorial line in $[k]^d$.

Construction: Now for any positive integer $k \geq 2$, consider $X = [k]^d, Y =$ {combinatorial lines of $[k]^d$} where d is $d(k, \frac{1}{2})$ by the Hales-Jewett Theorem, and we just take combinatorial lines to mean the lines passing through the combinatorial line sets.

By Hales-Jewett, we know that any set of $\frac{k^d}{2}$ points in $[k]^d$ contains a combinatorial line. And the complement of any set which has $\frac{k^d}{2}$ points as well must also contain a combinatorial line. Thus, we know that the optimal solution must have at least $\frac{k^d}{2}$ points.

However, in LP, we can assign $\frac{1}{k}$ to every point in $[k]^d$. Since every combinatorial line contains exactly k points, we know that every combinatorial line is covered in the LP. So the LP soluition is at most k^{d-1}.

Combining these two results, we see that in the d-dimensional instance of Points and Lines, the integrality gap is at least $\frac{\frac{k^d}{2}}{k^{d-1}} = \frac{k}{2}$.

Alon uses the following lemma to translate the d-dimensional construction to the 2D plane:

"For every positive integer d, there are d vectors $v_1, ..., v_d$ in the plane so that for any every two nontrivial sequences of integers $(k_1, ..., k_d)$ and $(k'_1, k'_2, ..., k'_d)$ with $|k_i|$, $|k'_i|$ < k for all i, the two vectors $\Sigma_i k_i v_i$ and $\Sigma_i k'_i v_i$ have the same direction if and only if $(k_1, ..., k_d)$ and $(k'_1, k'_2, ..., k'_d)$ all have the same direction... Moreover, there are such vectors v_i in which all coordinates are integers of absolute value at most $(2k-1)^{2d}$."

Thus, we can project the point $(a_1, a_2, ..., a_d)$ to $a_1 v_1 + a_2 v_2 + ... + a_d v_d$ in the 2D plane.

So we know that the integrality gap is at least $\frac{k}{2}$ for any integer k in the planar instance of hitting set.

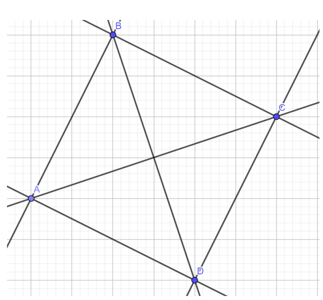
Translating to Points and Lines

Using these results, it would be nice if this construction could be applied to the Points and Lines problem. In fact, using Primal-Dual transforms from Computational Geometry[2], we can convert Hitting Set directly to covering Points by Lines.

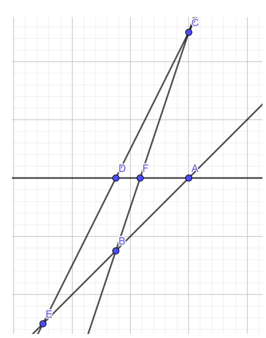
For any line y = mx + c, let's represent it as the point (m, -c).
For any point (a, b) let's represent it as the line y = ax - b.

We can apply this transformation as long as no lines are vertical. As long as there are a finite number of lines, we can rotate the whole plane so that no lines are vertical.

Note that if n lines intersect in the original graph, these are represented as n collinear points. As an example of this transformation consider:



(a) Covering lines by points

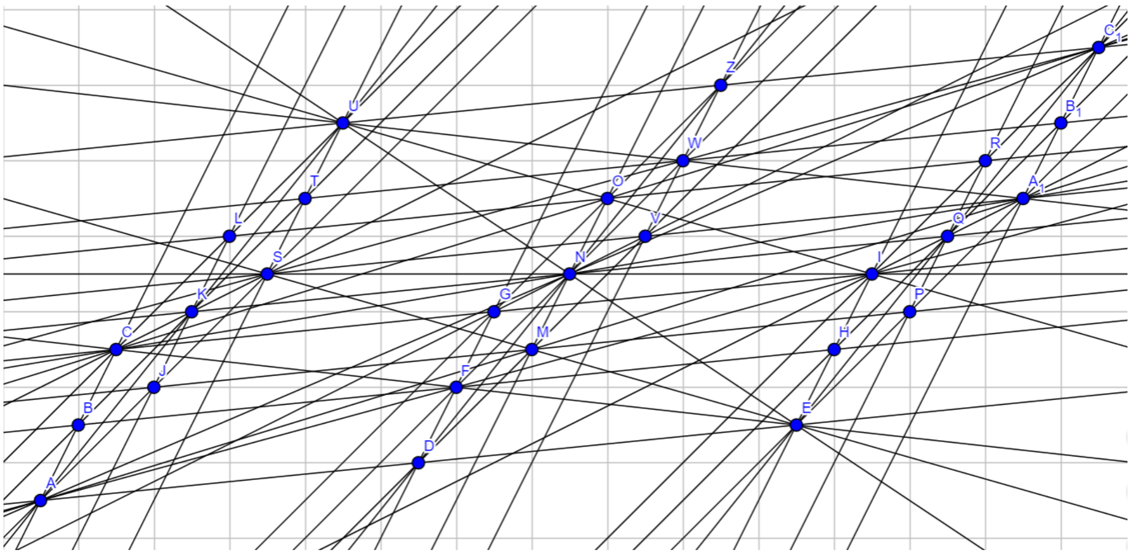


(b) Covering points by lines

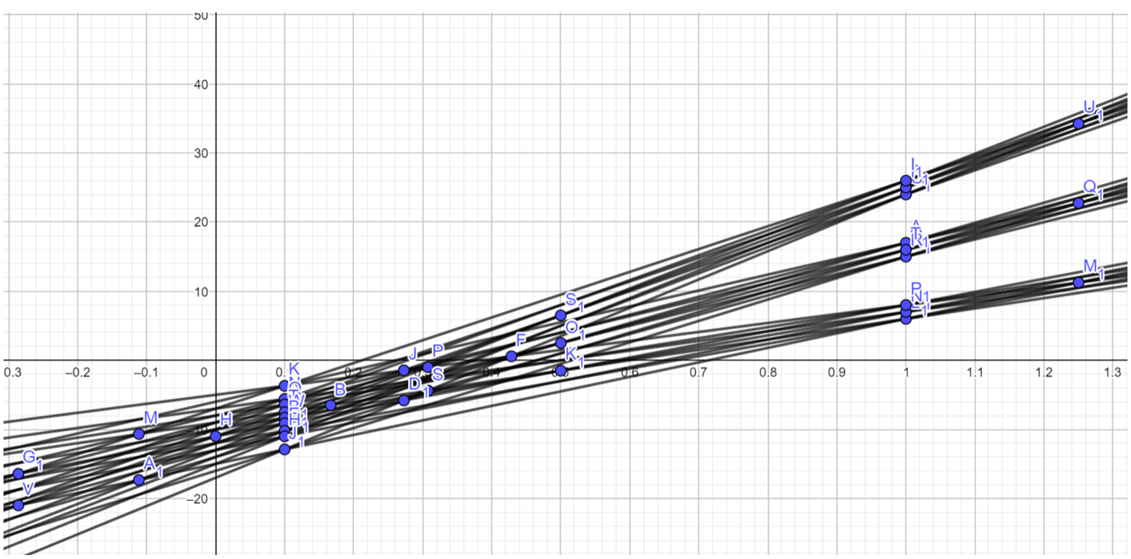
If a line contains k points in the original image, then k lines pass through it in the transform.

So the integrality gap is preserved when translating from covering lines by points to covering points by lines. Thus, we can apply this transformation to Alon's example to see that for any positive integer k, the integrality gap of covering points by lines is at least k.

Note that in practice it's extremely hard to come up with concrete examples. For example, to translate $[3]^3$ to covering lines by points to covering points by lines looks like:



Covering lines by points



Covering points by lines

And this figure only demonstrates $\sim \frac{3}{2}$ gap. To show a gap of say 50, we would need to find $d(100, .5)$, and the transform would require billons of lines.

Conclusion:

In this paper, we've shown that there cannot exist a constant upper bound for the integrality gap for Covering Points by Lines. Perhaps this LP formulation is still salvagable, and the integrality gap is upper bounded by a function of the number of points. Since in order to construct an integrality gap of k, we needed $[2k]^d$ points for some abstract d, which is an exponentially scaling number of points. It also could be that a better LP formulation gives us better integrality gaps.

I found this problem interesting because it's very similar to vertex cover, but creating systems of colinear points allows for more complex problems that vertex cover allows. Since we know that vertex cover has a 2-approximation, it would have been nice to be able to use this LP to come up with a rounding approximation that leads to a constant factor approximation. Through researching this problem, I was able to learn results about Ramsey Theory and Computational Geometry that I found very fascinating! I think that the result that I have shown would be interesting to those interested in geometric algorithms. This problem was interesting because it invoked many techniques we used in class, including Linear Progamming and Primal-Dual, although it required additional tools to complete the construction.

Works Cited

1. N. Alon, A non-linear lower bound for planar epsilon-nets; https://www.tau.ac.il/ nogaa/PDFS/epsnet1.pdf

2. D. Mount, Computational Geometry, http://graphics.stanford.edu/courses/cs268-16-fall/Notes/cmsc754-lects.pdf, 43-44.

3. Collaborated with Professor Chakrabarty