| CS602 Combinatorial Optimization via Linear and Convex Programming | Jan-Apr 2022 |
|---|---|
| Mid-semester exam | Total Marks: 40 |

**Que 1 [4 marks].** We want to write a linear program to check whether a given graph is 2-colorable, that is, whether we can color each vertex either red or blue such that any two adjacent vertices get different colors.

Let the variable $b_v$ indicate whether vertex $v$ is colored blue or not. Similarly, let the variable $r_v$ indicate whether vertex $v$ is colored red or not. We write the following linear constraints.

$$\begin{aligned} b_v + r_v &= 1 \text{ for each vertex } v \\ b_v, r_v &\geq 0 \text{ for each vertex } v \\ \sum_{v \in C} b_v &\leq 1 \text{ for each clique } C \text{ in the graph} \\ \sum_{v \in C} r_v &\leq 1 \text{ for each clique } C \text{ in the graph} \end{aligned}$$

We claim that the graph is 2-colorable if and only if the above linear constraints have a feasible solution. It turns out that the claim is false. Which of the following graphs demonstrate that the claim is false?

- any bipartite graph
- the complete graph on 5 vertices
- the complete graph on 3 vertices
- cycle of length 5

**Ans.** Cycle of length 5.

It is easy to verify that a cycle of length 5 is not 2-colorable, we need at least three colors to color its vertices. On the other hand the given set of linear constraints for this particular graph have a feasible solution. We can set $b_v = r_v = 1/2$ for every vertex in the cycle. The only cliques in this graph are the edges. Hence, we have two constraints for each edge $(u, v)$: $b_u + b_v \leq 1$ and $r_u + r_v \leq 1$. Both conditions are clearly satisfied. Hence, this example demonstrates that our claim was false.

**Que 2 [4 marks].** Consider the following three dimensional polytope P with 5 constraints.

$$\begin{aligned} z &\geq 0 \\ y + z &\leq 1 \\ z - y &\leq 1 \\ x - z &\leq 1 \\ x + z &\geq -1 \end{aligned}$$

Let us eliminate variable $z$ using Fourier Motzkin elimination. That will give us the projection of P on the $(x, y)$-plane. What kind of region will this projection be?

- Pentagon
- Hexagon
- Quadrilateral
- Unbounded region

**Ans.** For Fourier Motzkin elimination let's sort the constraints int two buckets (i) $z \leq$ and (ii) $z \geq$.

$$\begin{aligned} z \geq& 0 \\ z \geq& x - 1 \\ z \geq& -x - 1 \end{aligned}$$

$$\begin{aligned} z \leq& 1 - y \\ z \leq& 1 + y \end{aligned}$$

To eliminate $z$ we will combine each constraint from the first bucket with every constraint from the second bucket. That gives us $3 \times 2 = 6$ combinations. So, the 2-dimensional region we will get will be defined by 6 constraints. Hence, our first guess can be a hexagon. But, remember that some of the constraints might be redundant. Also, we cannot rule out an unbounded region. So, let us find the constraints and see. The 6 constraints we get are

$$\begin{aligned} 0 \leq& 1 - y \\ 0 \leq& 1 + y \\ x - 1 \leq& 1 - y \\ x - 1 \leq& 1 + y \\ -x - 1 \leq& 1 - y \\ -x - 1 \leq& 1 + y \end{aligned}$$

Rewriting,

$$\begin{aligned} y \leq& 1 \\ y \geq& -1 \\ x + y \leq& 2 \\ x - y \leq& 2 \\ y - x \leq& 2 \\ -x - y \leq& 2 \end{aligned}$$

Let's plot. The shaded region in Figure 1 is a hexagon.

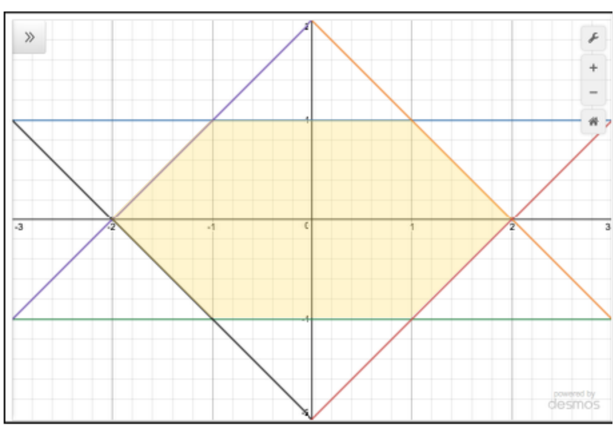


Figure 1: Yellow shaded region shows the feasible region defined by the 6 constraints

**Que 3.** Suppose you want to promote an ad on twitter. You have identified a set of $n$ influencers on twitter, each of which have a significant number of followers you are targeting. You have obtained the lists of their followers, let these lists be $L_1, L_2, \ldots, L_n$ . Note that these lists may have many common followers among them. Let's say the influencers are somewhat diverse in the sense that any follower is present in **at most 3** out of these lists.

Each influencer is asking for a price for promoting your ad, let these prices be $p_1, p_2, \ldots, p_n$. Your goal is to choose a subset of influencers with minimum total price such that every follower in the union of the $n$ lists is covered by at least one chosen influencer.

We write the following LP.

$$\min \sum_{i=1}^{n} p_i x_i \text{ subject to}$$
$$x_i \geq 0 \text{ for each } 1 \leq i \leq n$$
$$\sum_{i \text{ followed by } f} x_i \geq 1 \text{ for each follower } f$$

The last summation is over all influencers $i$ such that the follower $f$ is in his/her list $L_i$. Recall the assumption that for any follower there are at most 3 such influencers.

(a) Write the dual LP for this LP [3 marks].

(b) Write the complementary slackness conditions [2 marks].

(c) Describe a primal dual algorithm, along similar lines as the vertex cover algorithm we discussed in the class [4 marks].

(d) Prove that your algorithm is a 3-approximation algorithm [3 marks].

**Ans.** (a) Dual LP

$$\max \sum_{f} y_f \text{ subject to}$$
$$y_f \geq 0 \text{ for each follower } f$$
$$\sum_{f \text{ follows } i} y_f \leq p_i \text{ for } 1 \leq i \leq n$$

(b) Complementary slackness conditions

$$x_i > 0 \implies \sum_{f \text{ follows } i} y_f = p_i \quad \text{for } 1 \leq i \leq n$$

$$y_f > 0 \implies \sum_{i \text{ followed by } f} x_i = 1 \quad \text{for each follower } f$$

(c) It's analogous to the vertex cover algorithm seen in the class.

1. Initialize $I \leftarrow \phi$ (empty set). All dual variables $y_f \leftarrow 0$.
2. Take any follower $f$ that is not covered by the set $I$ of influencers. Increase $y_f$ till the point that one of the dual constraints become tight.
3. If the dual constraint for any influencer $i$ is tight, then include $i$ in $I$ or equivalently, set $x_i = 1$.
4. Stop if all followers are covered, otherwise go to 2.

(d) Note that $y$ is always a feasible solution for the dual LP, as we never violate any dual constraints. Suppose $I$ is the set of influencers computed at the end of the algorithm. The total price for this set will be

$\sum_{i \in I} p_i$. Recall that an influencer is included in $I$ only when the corresponding dual constraint has become tight. Hence,

$$\sum_{i \in I} p_i = \sum_{i \in I} \sum_{f \text{ follows } i} y_f = \sum_f y_f \times (\text{number of influencers in } I \text{ followed by } f) \leq \sum_f y_f \times 3$$

The last inequality comes from the assumption that each follower $f$ follows at most 3 influencers. Now, we can say

$$\begin{aligned}
\sum_{i \in I} p_i &\leq 3 \times (\text{value of a dual feasible solution}) \\
&\leq 3 \times (\text{dual LP optimal value}) \\
&= 3 \times (\text{primal LP optimal value}) \\
&\leq 3 \times (\text{ price for optimal set of influencers})
\end{aligned}$$

Here we are using the fact that the cost of any dual feasible solution is at most the optimal dual LP cost. Moreover, the primal LP optimal value is at most the price for optimal set of influencers.

**Que 4 [5 marks].** Let's slightly change the problem statement of the previous influencer problem. Now, we don't have to cover every follower, but let's say we are given a target $N$ for the number of followers we need to cover. The goal is to choose a subset of influencers with minimum total price such that we cover at least $N$ distinct followers. Write a linear program for this and argue that an integral optimal solution to this LP will give us the desired optimal subset of influencers.

**Ans.** We can write the following LP. We will introduce additional variables $z_f$ for each follower $f$, which can take values between 0 and 1 and denotes whether $f$ is covered by some influencer or not. We will replace the inequality $\sum_{i \text{ followed by } f} x_i \geq 1$ with $\sum_{i \text{ followed by } f} x_i \geq z_f$.

$$\begin{aligned}
\min \sum_{i=1}^{n} p_i x_i &\text{ subject to} \\
x_i \geq 0 &\text{ for each } 1 \leq i \leq n \\
1 \geq z_f \geq 0 &\text{ for each follower } f \\
\sum_f z_f \geq N & \\
\sum_{i \text{ followed by } f} x_i \geq z_f &\text{ for each follower } f
\end{aligned}$$

Note that enforcing $1 \geq z_f$ is necessary. If we allow $z_f$ to be, say 2, it would mean we are allowed to count a follower twice if it is covered by two influencers, and in that case $\sum_f z_f \geq N$ will not ensure that we have covered $N$ distinct followers.

Let's argue that any integral feasible solution will correspond to a set of influencers that covers at least $N$ followers. Consider any integral feasible solution $(x, z)$. We can interpret this solution as follows. Think of $\{f : z_f = 1\}$ as the set of followers chosen to be covered. The constraint $\sum_f z_f \geq N$ ensures that at least $N$ followers are chosen to be covered. The constraint $\sum_{i \text{ followed by } f} x_i \geq z_f$ ensures that if $f$ is chosen to be covered then there is at least one influencer covering it. The objective function is trying to minimize the total price as usual.

**Que 5 [5 marks]** Recall the vertex cover LP and the dual LP we wrote in the class.

$$\min \sum_{u \in V} w_u x_u \text{ subject to}$$
$$x_u \geq 0 \text{ for } u \in V$$
$$x_u + x_v \geq 1 \text{ for } (u,v) \in E$$

Dual LP

$$\max \sum_{e \in E} y_e \text{ subject to}$$
$$y_e \geq 0 \text{ for } e \in E$$
$$\sum_{e \in \delta(u)} y_e \leq w_u \text{ for } u \in V$$

Give an example of a graph with weights on vertices where the weight of the optimal vertex cover is almost twice of the optimal dual value [3 marks].

The primal dual algorithm we saw in the class gives a guarantee of approximation factor 2. It does that by bounding the gap between the obtained vertex cover and a dual feasible solution. Argue that any algorithm that goes via bounding the gap between the obtained vertex cover and a dual feasible solution cannot achieve approximation factor 1.99 (or less) [2 marks].

**Ans.** For the example asked in the question, take the complete graph on $n$ vertices, with all vertex weights as 10. Since it is a complete graph, we need at least $n-1$ vertices in any vertex cover (if we leave out any two vertices the edge between them will not be covered). Thus, the weight of the optimal vertex cover is $10(n-1)$.

Now, let's see the optimal dual value. Note that it always is equal to the primal LP optimal value. We claim that the solution $x_u = 1/2$ for every vertex $u \in V$ is a the LP optimal solution. It's easy to see that this solution satisfies the primal LP constraints. The primal object value for this solution will be $1/2 \times 10 \times n = 5n$. To be convinced that this is indeed the optimal solution, let's see a dual solution achieving the same objective value. The dual solution is $y_e = 10/(n-1)$ for every edge $e$. It is easy to verify that $\sum_{e \in \delta(u)} y_e = 10 = w_u$. The dual objective value here is $nchoose2 \times 10/(n-1) = 5n$.

Clearly, the weight of the optimal vertex cover $10(n-1)$ is almost twice of the optimal dual value $5n$.

We have shown an example where the gap between any vertex cover and any dual solution is at least a factor of $2(1-1/n)$. Whatever approximation guarantee the algorithm gives, it must work on every instance of the problem. Hence, any kind of primal dual algorithm cannot give an approximation guarantee better than $2(1-1/n)$. In particular, one cannot achieve approximation factor 1.99.

**Que 6 [10 marks]** Let $P$ and $Q$ be $n$-dimensional polytopes. Prove the following minimax equality.

$$\max_{x \in P} \min_{y \in Q} x^T y = \min_{y \in Q} \max_{x \in P} x^T y$$

You can use any results from the class, for example, strong duality in any form or the minimax equality proved in the last lecture. If you want to use strong duality, the following hints may be useful.

**Hint 1**: Let the corners of $P$ be $p_1, p_2, \ldots, p_k$ where $p_i = (p_{i,1}, p_{i,2}, \ldots, p_{i,n})$ for $1 \leq i \leq k$ . The polytope $P$ can be described as follows.

$$x_j = \lambda_1 p_{1,j} + \lambda_2 p_{2,j} + \cdots + \lambda_k p_{k,j} \text{ for } 1 \leq j \leq n$$
$$\lambda_1 + \lambda_2 + \cdots + \lambda_k = 1$$
$$\lambda_1, \lambda_2, \ldots, \lambda_k \geq 0.$$

**Hint 2**: $\max\{p_1^T y, p_2^T y, \ldots, p_k^T y\}$ is same as

$$\begin{aligned}&\min z \text{ subject to}\\ &z \geq p_{i,1}y_1 + p_{i,2}y_2 + \cdots + p_{i,n}y_n \text{ for } 1 \leq i \leq k\end{aligned}$$

**Ans.** Let's first see an approach with strong duality. Let's start with $\min_{y \in Q} \max_{x \in P} x^T y$.

**Observation 1** *For a fixed $y$, $\max_{x \in P} x^T y$ will be achieved at one of the corners of $P$.*

To see the above observation, you can view $x^T y = y^T x$ as a linear function in $x$. We know that for any polytope and any objective function, there is always a corner that achieves the optimal value. Suppose the corners of $P$ are $p_1, p_2, \ldots, p_k$. Then, we can write

$$\max_{x \in P} x^T y = \max\{p_1^T y, p_2^T y, \ldots, p_k^T y\}.$$

Now, observe that $\max\{p_1^T y, p_2^T y, \ldots, p_k^T y\}$ is same as the optimal value of the following linear program.

$$\begin{aligned}&\min z \text{ subject to}\\ &z \geq p_i^T y = p_{i,1}y_1 + p_{i,2}y_2 + \cdots + p_{i,n}y_n \text{ for } 1 \leq i \leq k\end{aligned}$$

Now, we need to minimize over $y \in Q$. We can say that $\min_{y \in Q} \max_{x \in P} x^T y = \min_{y \in Q} \max\{p_1^T y, p_2^T y, \ldots, p_k^T y\}$ is same as

$$\begin{aligned}&\min z \text{ subject to}\\ &z \geq p_i^T y = p_{i,1}y_1 + p_{i,2}y_2 + \cdots + p_{i,n}y_n \text{ for } 1 \leq i \leq k\\ &y \in Q.\end{aligned}$$

Now, let's express the constraint $y \in Q$ as linear constraints. We will enforce that $y$ is a convex combination of corners of $Q$. Let the corners of $Q$ be $q_1, q_2, \ldots, q_\ell$, where $q_i = (q_{i,1}, q_{i,2}, \ldots, q_{i,n})$ for $1 \leq i \leq \ell$. The above optimization program can be rewritten as

$$\begin{aligned}&\min z \text{ subject to}\\ &z \geq p_{i,1}y_1 + p_{i,2}y_2 + \cdots + p_{i,n}y_n \text{ for } 1 \leq i \leq k\\ &y = \lambda_1 q_1 + \lambda_2 q_2 + \cdots + \lambda_\ell q_\ell\\ &\lambda_1 + \lambda_2 + \cdots + \lambda_\ell = 1\\ &\lambda_1, \lambda_2, \ldots, \lambda_\ell \geq 0.\end{aligned}$$

Note that the equation $y = \lambda_1 q_1 + \lambda_2 q_2 + \cdots + \lambda_\ell q_\ell$ actually means $n$ equations, as $y$ and $q_i$'s are $n$-dimensional vectors. Let's write these $n$ equations explicitly. The program becomes

---

**LP1:**

$$\begin{aligned}&\min z \text{ subject to}\\ &z \geq p_{i,1}y_1 + p_{i,2}y_2 + \cdots + p_{i,n}y_n \text{ for } 1 \leq i \leq k\\ &y_j = \lambda_1 q_{1,j} + \lambda_2 q_{2,j} + \cdots + \lambda_\ell q_{\ell,j} \text{ for } 1 \leq j \leq n\\ &\lambda_1 + \lambda_2 + \cdots + \lambda_\ell = 1\\ &\lambda_1, \lambda_2, \ldots, \lambda_\ell \geq 0.\end{aligned}$$

---

To clarify, in the above LP, the variables are $z$, $y_1, y_2, \ldots, y_n$, and $\lambda_1, \lambda_2, \ldots, \lambda_\ell$. $q_{i,j}$s and $p_{i,j}$s are all constants.

By the same arguments, $\max_{x\in P}\min_{y\in Q} x^T y$ is same as the optimal value of the following linear program.

---

**LP2:**

$$\begin{aligned}
&\max u \text{ subject to}\\
&u \leq q_{i,1}x_1 + q_{i,2}x_2 + \cdots + q_{i,n}x_n \text{ for } 1 \leq i \leq \ell\\
&x_j = \alpha_1 p_{1,j} + \alpha_2 p_{2,j} + \cdots + \alpha_k p_{k,j} \text{ for } 1 \leq j \leq n\\
&\alpha_1 + \alpha_2 + \cdots + \alpha_k = 1\\
&\alpha_1, \alpha_2, \ldots, \alpha_k \geq 0.
\end{aligned}$$

---

Now, we will just argue that LP2 is the dual of LP1, which will mean that the optimal values of the two LPs are same and we will get the desired equality. Let's rewrite LP1 in the standard form.

$$\begin{aligned}
\min z &\text{ subject to}\\
z - p_{1,1}y_1 - p_{1,2}y_2 + \cdots - p_{1,n}y_n &\geq 0\\
z - p_{2,1}y_1 - p_{2,2}y_2 + \cdots - p_{2,n}y_n &\geq 0\\
&\vdots\\
z - p_{k,1}y_1 - p_{k,2}y_2 + \cdots - p_{k,n}y_n &\geq 0\\
y_1 - q_{1,1}\lambda_1 - q_{2,1}\lambda_2 - \cdots - q_{\ell,1}\lambda_\ell &= 0\\
y_2 - q_{1,2}\lambda_1 - q_{2,2}\lambda_2 - \cdots - q_{\ell,2}\lambda_\ell &= 0\\
&\vdots\\
y_n - q_{1,n}\lambda_1 - q_{2,n}\lambda_2 - \cdots - q_{\ell,n}\lambda_\ell &= 0\\
\lambda_1 + \lambda_2 + \cdots + \lambda_\ell &= 1\\
\lambda_1, \lambda_2, \ldots, \lambda_\ell &\geq 0.
\end{aligned}$$

LP1 has $1+n+\ell$ variables and $k+n+1$ constraints (except the non-negativity constraints). Let's take $k+n+1$ dual variables, $\alpha_1, \alpha_2, \ldots, \alpha_k$, $x_1, x_2, \ldots, x_n$, and $u$. We will have $\alpha_1, \alpha_2, \ldots, \alpha_k \geq 0$ in the dual LP because the first $k$ primal constraints are inequalities. While the other dual variables $x_1, x_2, \ldots, x_n$, and $u$ will be unrestricted (no non-negativity constraint) because the last $n+1$ constraints are equalities.

The dual objective comes from the RHS numbers in the primal constraints. There is exactly one primal constraint with nonzero RHS value, and the corresponding dual variable has been taken as $u$. Hence, the dual objective function is $u$. Since the primal is minimizing, the dual will be maximizing.

For each primal variable, we need to add a dual constraint. For the primal variable $z$, the dual constraint we get is

$$\alpha_1 + \alpha_2 + \cdots + \alpha_k = 1.$$

The 1 in the RHS comes from the primal objective.

Now, let's write a dual constraint for each primal variable $y_j$. The RHS will be zero as $y_j$ does not participate in the objective function. And the constraint will be equality because $y_j$ is unrestricted in the primal (no $\geq 0$ constraint). For $1 \leq j \leq n$,

$$-p_{1,j}\alpha_1 - p_{2,j}\alpha_2 - \cdots - p_{k,j}\alpha_k + x_j = 0$$

Now, let's write a dual constraint for each primal variable $\lambda_i$. The RHS will be zero as $\lambda_i$ does not participate in the objective function. And the constraint will be inequality ($\leq$) because we have $\lambda_i \leq 0$ and the dual is a maximization LP. For $1 \leq i \leq \ell$,

$$-q_{i,1}x_1 - q_{i,2}x_2 - \cdots - q_{i,n}x_n + u \leq 0.$$

Putting everything together,

$$\begin{aligned}
\max u \text{ subject to} \\
\alpha_1, \alpha_2, \ldots, \alpha_k \geq 0. \\
\alpha_1 + \alpha_2 + \cdots + \alpha_k = 1 \\
-p_{1,j}\alpha_1 - p_{2,j}\alpha_2 - \cdots - p_{k,j}\alpha_k + x_j = 0 \text{ for } 1 \leq j \leq n \\
-q_{i,1}x_1 - q_{i,2}x_2 - \cdots - q_{i,n}x_n + u \leq 0 \text{ for } 1 \leq i \leq \ell
\end{aligned}$$

This is same as LP2.

**Alternate approach:** Another way to prove the desired equation is via the minimax equality seen in the class. Let $\Delta_n$ be the $n$-dimensional probability simplex, i.e.,

$$\Delta_\ell = \{\lambda \in \mathbb{R}^\ell : \sum_{i=1}^{\ell} \lambda_i = 1 \text{ and } \lambda_i \geq 0 \text{ for each } 1 \leq i \leq \ell\}.$$

Let $R$ be a $k \times \ell$ real matrix. Let $R_{i,j}$ be the $(i,j)$-th entry of $R$. We had seen the equality

$$\max_{\alpha \in \Delta_k} \min_{\lambda \in \Delta_\ell} \alpha^T R \lambda = \max_{\alpha \in \Delta_k} \min_{\lambda \in \Delta_\ell} \alpha^T R \lambda.$$

Now, suppose the corners of the polytope $P$ are $p_1, p_2, \ldots, p_k$. We can write

$$P = \{\alpha_1 p_1 + \alpha_2 p_2 + \cdots + \alpha_k p_k : (\alpha_1, \alpha_2, \ldots, \alpha_k) \in \Delta_k\}.$$

Let $\mathsf{P}$ be the $n \times k$ matrix whose $i$-th column is $p_i$. The above can be rewritten as

$$P = \{\mathsf{P}\alpha : \alpha \in \Delta_k\}.$$

Similarly, let $\mathsf{Q}$ be the $n \times \ell$ matrix whose $i$-th column is $q_i$, the $i$-th corner of polytope $Q$. We can similarly write

$$Q = \{\mathsf{Q}\lambda : \lambda \in \Delta_k\}.$$

Now, let's go back to our desire equation. From the above discussion,

$$\begin{aligned}
\max_{x \in P} \min_{y \in Q} x^T y &= \max_{\alpha \in \Delta_k} \min_{\lambda \in \Delta_\ell} (\mathsf{P}\alpha)^T (\mathsf{Q}\lambda) \\
&= \max_{\alpha \in \Delta_k} \min_{\lambda \in \Delta_\ell} \alpha^T \mathsf{P}^T \mathsf{Q} \lambda \\
&= \max_{\alpha \in \Delta_k} \min_{\lambda \in \Delta_\ell} \alpha^T \mathsf{R} \lambda
\end{aligned}$$

Here we have defined the matrix $\mathsf{R}$ as $\mathsf{P}^T\mathsf{Q}$. Similarly, we can obtain

$$\min_{y \in Q} \max_{x \in P} x^T y = \min_{\lambda \in \Delta_\ell} \max_{\alpha \in \Delta_k} \alpha^T \mathsf{R} \lambda.$$

Now, we can use the equality seen in the class and conclude

$$\max_{x \in P} \min_{y \in Q} x^T y = \min_{y \in Q} \max_{x \in P} x^T y.$$