Assignment 2 — Separation

Koen Moors and Jeroen van Riel

June 2023

1 – Relation between special sets and stable sets

Intuitively, the edges in G model which negative points may not occur together in a special set. Formally, it is easily seen as follows. Let $S \subseteq Y$ be a stable set of G, then for any pair $y_1, y_2 \in S$, we have that $\{y_1, y_2\} \notin E$, hence $\text{conv}(\{y_1, y_2\}) \cap \text{conv}(X) \neq \emptyset$, which shows that S satisfies the definition of a special set for (X, Y). Conversly, the same argument shows that any special set for (X, Y) is also a valid stable set of G.

2 – Integer programming formulation for stable sets

Let $G = (V, E)$ be an undirected graph. We can now solve the following integer program to find the maximum stable set.

$$
\begin{aligned}
\max & \sum_{v \in V} x_v & & \quad (1a)\\
\text{s.t. } & x_u + x_v \leq 1 & \text{for all } \{u, v\} \in E, & \quad (1b)\\
& x_v \in \{0, 1\} & \text{for all } v \in V. & \quad (1c)
\end{aligned}
$$

Here, x_v is a variable that indicates whether vertex v is part of the stable set. We now show that every optimal solution of (1) corresponds to a maximum stable set and vice versa.

Let $x \in \{0, 1\}^n$ be the solution of (1) and let $S = \{v \in V : x_v = 1\}$. Let $u, v \in S$ with $u \neq v$. Note that $x_u = x_v = 1$. To show that S is a stable set, we need to show that $\{u, v\} \notin E$. Suppose $\{u, v\} \in E$. Then $x_u + x_v \leq 1$ by (1b), but this contradicts the fact that $x_u = x_v = 1$ and hence $\{u, v\} \notin E$ and S is a stable set.

What is remaining is to show that S is maximum. Suppose there exists a stable set S' with $|S'| > |S|$. Then there exist distinct $u, v \in S'$ with $\{u, v\} \in E$ and $x_u + x_v > 1$. If this would not be the case, the integer program would have found the solution $|S'|$ because it satisfies all constraints and has a higher objective value due to (1a). However, this contradicts the stable set property of S'.

Therefore S is the maximum stable set if and only if the corresponding x is an optimal solution of (1).

3 – Stengthening the formulation

Consider the odd cycle inequality

$$\sum_{v \in V'} x_v \leq \frac{k-1}{2}. \qquad (2)$$

Part 1

Let $S \subseteq V$ be a stable set of G and define $S' = S \cap V'$. Suppose $|S'| \geq (k+1)/2$ for some odd k. Because S is a stable set, two distinct vertices $x, y \in S'$ cannot be incident to the same edge. Hence, each node $v \in S'$ is incident to exactly two unique edges $e_1, e_2 \in E'$ from the cycle, so the total number of unique edges in E' must be at least $2|S'| \geq k+1$. However, it is easily seen that $|E'| = k$ for any induced cycle with k vertices. This contradiction shows that $|S'| \leq (k-1)/2$, which shows that (2) must hold for any feasible $x \in [0,1]^V$ corresponding to a stable set.

Part 2

- Observe that each e_i with $i \in \{1, \ldots, k\}$ is contained in $P_I(C)$. Together with the zero vector, they form an affinely independent set, showing that $\dim(P_I(C)) = k$.
- Let F denote the face of $P_I(C)$ induced by the odd cycle inequality (2). The points S satisfying (2) with equality form a proper affine subspace of $\mathbb{R}^{V'}$, so $\dim(F) \leq \dim(S) \leq k-1$.
- Note that we may consider C_k instead of C up to relabeling of vertices, so let $V_k = \{1, \ldots, k\}$ denote the vertices in C_k corresponding to C. It is convenient to reorder them further by defining a permutation as follows. Let $r := (k+1)/2$ denote the number of odd numbers up to k. For $1 \leq i \leq r$ let $\sigma(i) = 2i - 1$ and for $r+1 \leq i \leq k$, let $\sigma(i) = 2(i-r)$. For example, $k = 7$ yields $\sigma = (1\,3\,5\,7\,2\,4\,6)$.

  Let $l := (k-1)/2$ denote the length of a largest stable set in C_k. We define k vectors $x^{(1)}, \ldots, x^{(k)}$ as follows. Let

  $$x_n^{(j)} = \mathbb{1}\{0 \leq n - j \mod k < l\},$$

  where $\mathbb{1}\{\cdot\}$ denotes the indicator function. Each $x^{(i)}$ corresponds to a sequence of consecutive vertices $(\sigma(i), \ldots, \sigma(i+l))$ that form a longest stable set in C.

  For notational convenience, we collect these points as columns in a matrix A_k. For example, $k = 7$ yields

  $$A_7 = \begin{pmatrix} 1 & 0 & 0 & 0 & 0 & 1 & 1 \\ 1 & 1 & 0 & 0 & 0 & 0 & 1 \\ 1 & 1 & 1 & 0 & 0 & 0 & 0 \\ 0 & 1 & 1 & 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 1 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 & 1 & 1 & 0 \\ 0 & 0 & 0 & 0 & 1 & 1 & 1 \end{pmatrix}.$$

- We now show that $x^{(1)}, \ldots, x^{(k)}$ are linearly independent and thus affinely independent, by showing that A_k is non-singular by showing that A_k^{-1} exists. For example, the inverse of A_7 is given by

  $$A_7^{-1} = \begin{pmatrix} 0.33 & 0.33 & 0.33 & -0.67 & 0.33 & 0.33 & -0.67 \\ -0.67 & 0.33 & 0.33 & 0.33 & -0.67 & 0.33 & 0.33 \\ 0.33 & -0.67 & 0.33 & 0.33 & 0.33 & -0.67 & 0.33 \\ 0.33 & 0.33 & -0.67 & 0.33 & 0.33 & 0.33 & -0.67 \\ -0.67 & 0.33 & 0.33 & -0.67 & 0.33 & 0.33 & 0.33 \\ 0.33 & -0.67 & 0.33 & 0.33 & -0.67 & 0.33 & 0.33 \\ 0.33 & 0.33 & -0.67 & 0.33 & 0.33 & -0.67 & 0.33 \end{pmatrix}.$$

In general, the first row of A_k^{-1} is defined by

$$(A_k^{-1})_{1n} = \begin{cases} 1/l - 1 \text{ for } n \in \{l+1, k\}, \\ 1/l \text{ otherwise.} \end{cases}$$

Recall that $(A_k)_{nj} = x_n^{(j)}$. Let $I_j := \{n : 0 \leq n - j \mod k < l\}$ denote the indices of the entries with a 1 in $x^{(j)}$. We first compute the first row of $A_k^{-1}A_k$, which is given by

$$(A_k^{-1}A_k)_{1j} = \sum_{n=1}^{k} (A_k^{-1})_{1n}(A_k)_{nj} = \sum_{n \in I_j} (A_k^{-1})_{1n}.$$

For $j = 1$, we obtain

$$(A_k^{-1}A_k)_{11} = \sum_{n=1}^{l} 1/l = 1.$$

For $2 \leq j \leq l+1$, we have

$$l+1 \in I_j, k \notin I_j,$$

so

$$(A_k^{-1}A_k)_{1j} = (1/l - 1) + \sum_{n \in I_j \setminus \{l+1\}} 1/l = 0.$$

For $l+2 \leq j \leq k$, we have

$$l+1 \notin I_j, k \in I_j,$$

so

$$(A_k^{-1}A_k)_{1j} = (1/l - 1) + \sum_{n \in I_j \setminus \{k\}} 1/l = 0.$$

Now the rth row of $A_k^{-1}A_k$ is just first row shifted r positions to the right, formally defined by

$$(A_k^{-1})_{r,n} = (A_k^{-1})_{1,(n-r \mod k)+1},$$

which will produce a 1 in the rth position of $(A_k^{-1}A_k)_{r\cdot}$, by similar arguments as above, hence $A_k^{-1}A_k = I$.

- Observe that $y^{(1)}, \ldots, y^{(k)}$, defined by $y_{\sigma(i)}^{(j)} = x_i^{(j)}$ for each $i, j \in \{1, \ldots, k\}$, are contained in $P_I(C)$ and satisfy the odd cycle inequality with equality. Because reordering of entries does not change affine independence, they form a set of affinely independent points in F, which shows that $\dim(F) = k-1$, hence we conclude that F is a facet of $P_I(C)$.

4 – Finding odd cycle inequalities

Given an undirected graph $G = (V, E)$, we can, in principle, strengthen the LP relaxation of (1) by adding all odd cycle inequalities to (1). In general, this approach is intractable though, because there might be exponentially many odd cycles in G. Therefore, one usually does not add all odd cycles to (1), but only those being 'useful' within the following procedure:

Start with the formulation $P(G)$, i.e., without odd cycle inequalities, and use it within branch-andbound to find a maximum stable set. The branch-and-bound algorithm is then enhanced as follows. Whenever a solution x^* of an LP relaxation is computed, one checks whether there exists some odd cycle inequality that is violated by x^*. If a violated inequality exists, it is added to the model to strengthen the formulation.

This causes the next challenge: how to identify violated odd cycle inequalities? To solve this, we use the following algorithm

---

Algorithm 1 Finding odd cycles

Input: an undirected graph $G = (V, E)$, a point $x^* \in P(G)$;
Output: an odd cycle $C = (V', E')$ in G for which $\sum_{v \in V'} x_v^* - (|V'| - 1)/2$ is as large as possible, or the correct statement that there does not exist any odd cycle in G.

Make (bipartite) auxiliary graph G'' with

- vertex set $V'' = V^+ \cup V^-$ where V^+ and V^- are copies of V,
- edge set $E'' = \{\{v^+, u^-\}, \{v^-, u^+\} : \{v, u\} \in E\}$,
- edge weights $w(v^+, u^-) = w(v^-, u^+) = 1 - x_v - x_u$ for every edge $\{v, u\} \in E$.

for $v \in V$ **do**
  find the shortest path from v^+ to v^- using e.g. Dijkstra's algorithm,
  record its length as ℓ_v,
  if there is no path from v^+ to v^-, set $\ell_v = \infty$
end for
Return: odd cycle $C = (V', E')$ such that ℓ_v is minimum for all $v \in V'$ or none if $\min_{v \in V} \ell_v = \infty$.

We first explain why this algorithm allows us to find a violated odd cycle inequality. Given a solution x^*, it finds an odd cycle that maximizes $\sum_{v \in V'} x_v^* - (|V'| - 1)/2$. Suppose there is an odd cycle $C = (V', E')$ of size k that violates the odd cycle inequality (2). Then $\sum_{v \in V'} x_v^* - (|V'| - 1)/2 > 0$. Therefore, if we maximize this quantity we find the odd cycle inequality that is most 'useful' to add to problem (1).

Note that this algorithm runs in polynomial time. First, the construction of the auxiliary graph G'' only takes a polynomial amount of operations. Then, for each $v \in V$ we apply Dijkstra's algorithm, which is $\mathcal{O}(|V||E|)$. Finally, finding the cycle from the values of ℓ_v can also be done relatively easy.

We now prove the correctness of the algorithm. To do so, we need to show the following

- if G does not contain an odd cycle, we find $\min_{v \in V} \ell_v = \infty$,
- if G contains an odd cycle, the algorithm always returns an odd cycle,

- if the algorithm returns odd cycle C, there is no odd cycle $\tilde{C}$ in G such that $\sum_{v\in\tilde{V}'} x_v^* - (|\tilde{V}'|-1)/2 > \sum_{v\in V'} x_v^* - (|V'|-1)/2$

Suppose G does not contain an odd cycle. Then for every vertex $v \in V$ there is no path in G from v to v of odd size. Note that for every vertex $v \in V$, every path from v^+ to v^- in G'' has an odd length, since they are on opposing sides of a bipartite graph. Moreover, every path from v^+ to v^- in G'' corresponds to a path from v to v in G of equal length by leaving out the $^+$ and $^-$ indicators in the edges, since these edges are present in G by construction of E''. Then if there exists a path from v^+ to v^- in G'', it is necessarily of an odd size and it corresponds to an odd length path from v to v in G. By contradiction, we then have that there is no path from v^+ to v^- in G'' and therefore $\ell_v = \infty$ for all $v \in V$.

Now suppose G does contain an odd cycle $C = (V', E')$. Let $v \in V'$. Then in auxiliary graph G'' there is a path from v^+ to v^-, given by the edges in E' where the $^+$ and $^-$ indicators are altered for neighbouring vertices in V'. Therefore $\ell_v \leq |V'|$ odd and hence the algorithm will always return a cycle. For reasons mentioned above, this cycle will always be odd.

Finally, suppose the algorithm returns cycle $C = (V', E')$, but there exists a cycle $\tilde{C} = (\tilde{V}', \tilde{E}')$ with

$$\begin{aligned}
&\sum_{v\in\tilde{V}'} x_v^* - \frac{|\tilde{V}'|-1}{2} > \sum_{v\in V'} x_v^* - \frac{|V'|-1}{2} \\
\Leftrightarrow &\sum_{v\in\tilde{V}'} 2x_v^* - |\tilde{V}'| > \sum_{v\in V'} 2x_v^* - |V'| \\
\Leftrightarrow &\sum_{v\in\tilde{V}'} (2x_v^* - 1) > \sum_{v\in V'} (2x_v^* - 1) \\
\Leftrightarrow &\sum_{v\in\tilde{V}'} (1 - 2x_v^*) < \sum_{v\in V'} (1 - 2x_v^*)\,.
\end{aligned}$$

Let $v \in \tilde{V}'$. Then using the path in G'' that corresponds to the cycle $\tilde{C}$ in G, we have that

$$\ell_v = \sum_{\{u,w\}\in\tilde{E}'} (1 - x_u - x_w) = |\tilde{E}'| - \sum_{\{u,w\}\in\tilde{E}'} (x_u + x_w) = |\tilde{V}'| - 2\sum_{v\in\tilde{V}'} x_v = \sum_{v\in\tilde{V}'} (1 - 2x_v^*)\,,$$

where we used that $|\tilde{E}'| = |\tilde{V}'|$ and that each vertex occurs twice in the edge set of a cycle. Now note that because our algorithm found the shortest path, we have

$$\ell_v = \sum_{v\in\tilde{V}'} (1 - 2x_v^*) < \sum_{v\in V'} (1 - 2x_v^*) = \min_{u\in V} \ell_u,$$

which leads to a contradiction and hence finishes the proof.

5 – A branch-and-cut algorithm

See the attached .py file.