Lecture 5: Polynomial algorithms for weakly bipartite and perfect graphs

Ahmad Abdi

February 6, 2023

1 Optimization vs. separation

In the previous lecture we saw that the set covering and set packing integer programs over ideal and perfect matrices can be reduced to solving the set covering and set packing linear programs, respectively. There is a key advantage here: while integer programming is NP-hard, a linear program is polynomially solvable. That is, the running time for solving a linear program is bounded above by a polynomial in the input size. While this is a powerful guarantee for a general linear program, it may not be sufficient for combinatorial optimisation problems over a graph where the number of constraints of the associated linear program is exponential in the size of the graph. Unfortunately this was the case for the linear programs that we saw in lectures 3 and 4, as we shall explain in the coming sections.

What comes to the rescue is that while our linear program may have exponentially many constraints, we do not need to store all of them, or any of them for that matter, in memory to solve the problem. All we need to retain is a so-called *separation oracle* which given a vector x declares in polynomial time whether x is feasible or outputs an inequality of the program that is violated at x. Finding a separation oracle reduces to solving another optimisation problem, one that is typically easier to solve. Let us elaborate.

Let K be a nonempty compact convex subset of $\mathbb{R}^n$. We are interested in the following algorithmic problems:

1. *Exact Optimisation Problem*: Given $c \in \mathbb{R}^n$, find a vector x in K which maximises $c^\top x$ over K.

2. *Exact Separation Problem*: Given y in $\mathbb{R}^n$, decide if $y \in K$, and if not, find a hyperplane separating K from y, i.e. a vector $c \in \mathbb{R}^n$ such that $c^\top y > \max\{c^\top x : x \in K\}$.

A powerful and brilliant result of Khachiyan [4], based on a method of Shor [6], is that these two algorithmic problems are in fact polynomially equivalent. That is, there is a polynomial time algorithm for solving the first problem if, and only if, there is one for solving the second one. The key insight for this equivalence comes from the *Ellipsoid Method* for linear programming, which we will not focus on in this lecture.

A *convex body* is a quintuple (K, n, a_0, r, R) where $n \geq 2$, K is a convex subset of $\mathbb{R}^n$, ansd K is sandwiched between two Euclidean balls centred at a_0 of radii r and R, where $0 < r \leq R$. We have the following important result from Grötschel, Lovász and Schrijver [1] (Theorem 3.1).

Theorem 1.1. Let $\mathcal{K}$ be a class of convex bodies. There is a polynomial algorithm to solve the separation problem for the members of $\mathcal{K}$ if, and only if, there is a polynomial algorithm for solving the optimisation problem for the members of $\mathcal{K}$.

Consequently, to have a polynomial algorithm for optimising a linear objective over a nonempty compact convex set K, all we need is a *polynomial separation oracle*, which is a polynomial algorithm for solving the separation problem over K.

There is a fine print to be mindful of. Above we make no assumption of the set K. In fact, as we shall see later, K may be defined by quadratic constraints, so the maximiser of the first problem, for instance, could have irrational coordinates, so how could one even represent the answer? One could instead with a more appropriate, and correct, formulation of the first problem known as the *Weak Optimisation Problem*, in which one looks for a *rational* solution that is approximately optimal up to a prescribed additive, small factor of precision. Similarly, there is the *Weak Separation Problem*. A subtle point is that the Ellipsoid Method, as well as the theorem above, are proved for the weak variants with exponential convergence rate guarantees, which in turn allows one to solve the exact variants, where possible, also in polynomial time by essentially rounding. We refer the interested reader to the paper of Grötschel, Lovász and Schrijver [1]. The same authors also wrote a book [2] on this topic, and we refer to Chapter 4 for more details.

In the remainder of this section, we shall focus on two profound implications of Theorem 1.1 in Combinatorial Optimisation.

2 The maximum cut problem for weakly bipartite graphs

Let $G = (V, E)$ be a graph. What is the maximum size of a cut? This problem is known as the *max cut problem*, and its decision version is one of Karp's 21 NP-complete problems. Interestingly, we can model the problem via a set covering integer program. The key idea is that every circuit intersects every cut in an even number of edges, thus every odd circuit has at least one edge off of any given cut. This relation essentially characterises cuts:

Lemma 2.1. Let $B \subseteq E$ be a subset such that $(E - B) \cap C \neq \emptyset$ for every odd circuit C. Then B is contained in a cut of G.

Proof. Exercise. $\square$

By the lemma above, B is a maximum cut if, and only if, $E - B$ is a minimum edge subset that intersects every odd circuit. Denote by $\mathcal{C}(G, E)$ the family of odd circuits of G. We can formulate the maximum cut problem as the following set covering integer program for $w = \mathbf{1}$.

$$\begin{array}{lrcll} \min & w^\top x & & & \\ \text{s.t.} & x(C) & \geq & 1 & C \in \mathcal{C}(G,E) \\ & x_e & \in & \{0,1\} & e \in E \end{array} \tag{1}$$

Let us focus on instances G where this integer program is equivalent to its linear relaxation:

$$\begin{array}{lrcll} \min & w^\top x & & & \\ \text{s.t.} & x(C) & \geq & 1 & C \in \mathcal{C}(G,E) \\ & x_e & \geq & 0 & e \in E \\ & x_e & \leq & 1 & e \in E. \end{array} \tag{2}$$

$G = (V,E)$ is a *weakly bipartite* graph if for all $w \in \mathbb{R}^E$ for which (2) has an optimal solution, there exists an integral optimal solution. That is, for weakly bipartite graphs, the integer program (1) is equivalent to its linear relaxation (2) for any linear objective function. However, as the reader will have noticed, the linear program may have exponentially many constraints, so in order to solve it efficiently, we will need a polynomial separation oracle. Somewhat miraculously there is one!

Theorem 2.2. There is a polynomial algorithm for finding a minimum cost odd circuit of a graph $G = (V,E)$ with cost vector $c \in \mathbb{R}^E_{\geq 0}$.

Proof. Let C be a minimum cost odd circuit of G, and let st be any edge of the circuit. Observe that the even st-path in C is a minimum cost even st-path.Now to solve the problem, for any edge st, we find a minimum cost even st-path P_{st} – we explain how this can be done efficiently shortly. Then we add $\{st\} \cup P_{st}$ to the list of candidate minimum cost odd circuits. We do this for every edge st, and then among the candidate odd circuits, we pick the one of minimum cost.

What remains to justify is a polynomial algorithm for finding P_{st}. We saw in the lecture 1 exercises that finding a minimum cost even st-path can be reduced to a minimum cost perfect matching problem in an instance with at most twice as many vertices, where every edge cost is either 0 or c_e for some $e \in E$. Since the latter has a polynomial algorithm as explained in lecture 1, we obtain a polynomial algorithm for finding P_{st}, as required. □

We may now use the equivalence of separation and optimisation to prove the following theorem.

Theorem 2.3. There is a polynomial algorithm for finding a maximum weight cut of any weakly bipartite graph $G = (V,E)$ with edge weights $w \in \mathbb{R}^E$.

Proof. As we argued above, finding a maximum weight cut of G is equivalent to solving the integer program (1). Since G is weakly bipartite, the integer program is equivalent to its linear relaxation (2). Thus, all we need to do is solve the LP in polynomial time.

Even though the LP may have exponentially many inequalities, it has a polynomial separation oracle. To see this, let $x^\star \in \mathbb{R}^E$. We need to decide in polynomial time whether $x^\star$ is feasible for the LP (2), or output an inequality that is violated. If $x^\star \notin [0,1]^E$, then we have an obvious violated inequality. Otherwise, $x^\star \in [0,1]^E$.

We then treat $x^\star$ as edge costs of G (these are different from the edge weights w), and look for a minimum cost odd circuit C of G, which can be done in polynomial time by Theorem 2.2. If $x^\star(C) \geq 1$, we conclude that $x^\star$ is a feasible solution for the LP. Otherwise, $x^\star(C) < 1$, so we output $x(C) \geq 1$ as an inequality of the LP that is violated at the given point $x^\star$.

We presented a polynomial separation oracle for (2). By Theorem 1.1 we then have a polynomial algorithm for solving the (2), as required. □

As we proved in lecture 3, any planar graph is weakly bipartite, for example. There is a powerful characterisation of such graphs due to Guenin [3]. To describe it, let $G = (V, E)$ be a graph. We say that G has an *odd-K_5 minor* if there exist subsets $I, J \subseteq E$ such that $I \cap J = \emptyset$, $J = \delta(U)$ for some $U \subseteq V$, and the minor $G \setminus I/J$ obtained after deleting I and contracting J is isomorphic to the complete graph K_5 on five vertices.

Theorem 2.4 (Characterisation of Weakly Bipartite Graphs)**.** A graph is weakly bipartite if, and only if, it has no odd-K_5 minor.

This theorem is equivalent to the characterisation of weakly bipartite *signed* graphs we mentioned in lecture 3.

3 Computing the stability number of perfect graphs: a failed attempt

Let $G = (V, E)$ be a simple graph. Denote by $\alpha(G)$ the maximum size of a stable set of G. The problem of computing $\alpha(G)$ is equivalent to the *maximum clique problem*, which is one of Karp's 21 NP-complete problems. Consider the following linear relaxation of the parameter:

$$\begin{array}{rrlcll} \alpha^\star(G) & := \max & \mathbf{1}^\top x & & & \\ & \text{s.t.} & x(K) & \leq & 1 & \text{for every clique } K \subseteq V \text{ of } G \\ & & x_v & \geq & 0 & v \in V. \end{array} \tag{3}$$

As we saw in lecture 4, if G is a perfect graph, then $\alpha(G) = \alpha^\star(G)$. Thus, is there a polynomial algorithm for solving (3) for an arbitrary graph? As optimisation and separation are equivalent, we may equivalently ask for a polynomial separation oracle for the LP. Unfortunately, this problem is NP-hard.

Theorem 3.1. Separation over the feasible region of (3) for an arbitrary graph $G = (V, E)$ is an NP-hard problem.

Proof. Exercise. □

One can now apply the contrapositive of Theorem 1.1 to conclude the following.

Corollary 3.2. Computing $\alpha^\star(G)$ for an arbitrary graph $G = (V, E)$ is NP-hard.

There is a subtle point here. Technically, Theorem 1.1 only tells us that solving (3) for an arbitrary linear objective function is NP-hard. However, the fact that the right-hand-side values of the clique inequalities are 1, allows us to conclude the above.

Things look hopeless now. It seems that the equivalence of separation and optimisation is only giving us a negative result. However, that is not the case. This equivalence does indeed lead to a polynomial algorithm for finding $\alpha^\star(G) = \alpha(G)$ for perfect graphs. We just need to apply the result to a different (non-linear) optimisation problem. Before explaining this, however, we need to talk about a seemingly different topic.

4 The theta function and a min-max theorem

Let $G = ([n], E)$ be a graph. Denote by $\vartheta(G)$ the minimum of the largest eigenvalue of A, ranging over all real symmetric $n \times n$-matrices A where $A_{ij} = 1$ if $i = j$ or i, j are non-adjacent vertices. We have the following inequality which provides a link between this parameter and the stability number.

Lemma 4.1. $\vartheta(G) \geq \alpha(G)$.

Proof. Let $\vartheta := \vartheta(G)$ and $\alpha := \alpha(G)$. Let A be a real symmetric $n \times n$-matrix where $A_{ij} = 1$ if $i = j$ or i, j are non-adjacent vertices, and whose largest eigenvalue of A is ϑ. By the Courant-Fischer Theorem, $\vartheta = \max\{x^\top Ax : x^\top x = 1\}$. We shall exhibit a solution of to this program whose objective value is α. To this end, let S be a maximum stable set. Then the $S \times S$ principal submatrix of A is an all ones submatrix. Let $\bar{x}_i := \frac{1}{\sqrt{\alpha}}$ for all $i \in S$ and $\bar{x}_i := 0$ for all $i \in [n] - S$. Then $\bar{x}^\top \bar{x} = 1$ and

$$\bar{x}^\top A\bar{x} = \sum_{i=1}^{n}\sum_{j=1}^{n} A_{ij}\bar{x}_i\bar{x}_j = (\alpha)^2 \cdot \frac{1}{\sqrt{\alpha}} \cdot \frac{1}{\sqrt{\alpha}} = \alpha.$$

Thus, $\vartheta \geq \alpha$, as required. □

Let M be an $n \times n$ real symmetric matrix. M is *positive semidefinite (PSD)* if each eigenvalue of M is nonnegative. Equivalently, M is PSD if $x^\top Mx \geq 0$ for all $x \in \mathbb{R}^n$. We are now ready to prove a min-max theorem for $\vartheta(G)$.

Theorem 4.2. $\vartheta(G)$ is equal to the maximum trace of BJ, where B is a PSD $n \times n$-matrix of trace 1 where $B_{ij} = 0$ whenever i, j are distinct and adjacent.

Proof. Let A be a symmetric $n \times n$-matrix where $A_{ij} = 1$ if $i = j$ or i, j are non-adjacent vertices, and whose largest eigenvalue is $\vartheta(G) =: \vartheta$.

($\geq$) We shall prove something a little stronger. Let B be an arbitrary PSD $n \times n$-matrix of trace at most 1 where $B_{ij} = 0$ whenever i, j are distinct and adjacent. We claim that $\vartheta \geq \mathrm{tr}(BJ)$. To this end, note that

$$\mathrm{tr}(BJ) = \sum_{i=1}^{n}\sum_{j=1}^{n} B_{ij} = \sum_{i=1}^{n}\sum_{j=1}^{n} A_{ij}B_{ij} = \mathrm{tr}(AB)$$

where the middle equality follows from our assumptions on the entries of A, B. Thus,

$$\vartheta - \mathrm{tr}(BJ) = \vartheta - \mathrm{tr}(AB) \geq \mathrm{tr}((\vartheta I - A)B)$$

where the last equality follows from the fact that $\mathrm{tr}(B) \le 1$. Since B is real symmetric matrix, it has an orthonormal basis $v_1, \ldots, v_n \in \mathbb{R}^n$ of eigenvectors with eigenvalues $\lambda_1, \ldots, \lambda_n$, respectively. Since B is PSD, $\lambda_1, \ldots, \lambda_n \ge 0$. By the Spectral Decomposition Theorem, $B = \sum_{i=1}^n \lambda_i v_i v_i^\top$. Thus,

$$\mathrm{tr}((\vartheta I - A)B) = \sum_{i=1}^n \lambda_i \mathrm{tr}((\vartheta I - A)v_i v_i^\top) = \sum_{i=1}^n \lambda_i v_i^\top (\vartheta I - A)v_i.$$

(Note $\mathrm{tr}(Mxx^\top) = x^\top Mx$.) Since ϑ is the largest eigenvalue of A, $\vartheta I - A$ is PSD, so $v_i^\top(\vartheta I - A)v_i \ge 0$ for each i. Thus,

$$\vartheta - \mathrm{tr}(BJ) = \mathrm{tr}((\vartheta I - A)B) \ge 0,$$

as required.

($\le$) We shall construct an eligible matrix B such that $\vartheta = \mathrm{tr}(BJ)$. It shall be of the form $B = \sum_{p=1}^N \alpha_p h_p h_p^\top$ where

1. $\alpha_p \ge 0$ for each $p \in [N]$, and $\sum_{p=1}^N \alpha_p = 1$: nonnegativity of the coefficients guarantee that B is PSD,
2. $h_p \in \mathbb{R}^n$ satisfies $h_p^\top h_p = 1$ for each $p \in [N]$: this guarantees that $\mathrm{tr}(B) = 1$,
3. $\sum_{p=1}^N \alpha_p h_{p,i} h_{p,j} = 0$ for each $ij \in E$: this guarantees that $B_{ij} = 0$ whenever $ij \in E$,
4. $\sum_{p=1}^N \alpha_p (\mathbf{1}^\top h_p)^2 = \vartheta$: this guarantees that $\mathrm{tr}(BJ) = \mathbf{1}^\top B \mathbf{1} = \vartheta$.

These requirements have an intuitive geometric interpretation. Equivalently, we would like to express the vector $(0, \ldots, 0) \times (\vartheta) \in \mathbb{R}^E \times \mathbb{R}$ as a finite convex combination of vectors from the set

$$S := \left\{ (h_i h_j : ij \in E, i > j) \times (\mathbf{1}^\top h)^2 : h \in \mathbb{R}^n, h^\top h = 1 \right\}.$$

Suppose for a contradiction that this is not possible. Since the set S is compact, there exists a hyperplane separating $(0, \ldots, 0, \vartheta)$ from S. That is, there exist a coefficient vector $(d_{ij} : ij \in E, i > j) \times (d_0) \in \mathbb{R}^E \times \mathbb{R}$ and a right-hand-side $\delta \in \mathbb{R}$ such that

$$\sum_{ij \in E, i>j} d_{ij} h_i h_j + d_0 (\mathbf{1}^\top h)^2 \le \delta \qquad \forall\, h \in \mathbb{R}^n, h^\top h = 1$$

$$d_0 \vartheta > \delta$$

In particular, for $h = (1, 0, \ldots, 0) \in \mathbb{R}^n$ the first inequality gives $d_0 \le \delta$, while $\delta < d_0 \vartheta$ by the second inequality. Since $\vartheta \ge 1$, it follows that $d_0 > 0$. By rescaling the coefficient vector and the right-hand-side, we can enforce $d_0 = 1$ without changing the directions of the inequalities. Thus, for each $h \in \mathbb{R}^n, h^\top h = 1$, we have

$$\vartheta > \delta \ge \sum_{ij \in E, i>j} d_{ij} h_i h_j + (\mathbf{1}^\top h)^2 = \sum_{i=1}^n \sum_{j=1}^n c_{ij} h_i h_j = h^\top C h$$

where C is a symmetric $V \times V$-matrix where $c_{ij} = 1 + \frac{d_{ij}}{2}$ if $ij \in E$, and $c_{ij} = 1$ if $i = j$ or i, j are non-adjacent. Let $\bar{h}$ be an eigenvector of C corresponding to the largest eigenvalue, say ϑ', such that $\bar{h}^\top \bar{h} = 1$. By definition of ϑ, we must have that $\vartheta' \ge \vartheta$. However, the above inequalities imply that $\vartheta > \bar{h}^\top C \bar{h} = \vartheta'$, a contradiction. $\square$

Thus,

$$\begin{array}{rlrcll} \vartheta(G) & = \max & \operatorname{tr}(BJ) & & & \\ & \text{s.t.} & B_{ij} & = & 0 & \text{for all } ij \in E \\ & & \operatorname{tr}(B) & = & 1 & \\ & & B & \succeq & \mathbf{0} & \end{array} \tag{4}$$

We can in fact relax $\operatorname{tr}(B) = 1$ to $\operatorname{tr}(B) \leq 1$ as we already explained in the $(\geq)$ direction of the proof of Theorem 4.2. Observe that the feasible region of this optimisation problem is a nonempty compact convex set, and that the objective function is linear in B. By dropping the coordinates of B below the main diagonal, and also those corresponding to the edges of E, we obtain a convex body. Since testing the positive semidefinite property can be done in polynomial time, one readily obtains a polynomial separation oracle for (4) and the associated convex body. For more details, we refer the reader to [1]. We can now use Theorem 1.1 to conclude the following.

Theorem 4.3. There is a polynomial algorithm for computing $\vartheta(G)$ for any graph G.

5 Orthonormal representations and a successful attempt

An *orthonormal representation* of $G = ([n], E)$ is a set of vectors $v_1, \dots, v_n$ in $\mathbb{R}^d$ (for any d) such that $v_i^\top v_i = 1$ for each i, and $v_i^\top v_j = 0$ for all non-adjacent vertices i, j. We shall need the following lemma.

Lemma 5.1. Let $v_1, \dots, v_n$ be an orthonormal representation of G, and let $u_1, \dots, u_n$ be an orthonormal representation of $\overline{G}$. Then for all $c, d \in \mathbb{R}^n$ we have

$$\sum_{i=1}^{n} (c^\top v_i)^2 (d^\top u_i)^2 \leq (c^\top c)(d^\top d).$$

Proof. Exercise. □

Let M be an $n \times n$ real symmetric matrix. It is well-known that M is PSD if, and only if, $M = X^\top X$ for some real matrix X with n columns. We are now ready to provide another min-max theorem for $\vartheta(G)$.

Theorem 5.2. $\vartheta(G)$ is the maximum of $\sum_{i=1}^{n} (d^\top u_i)^2$ over all orthonormal representations $u_1, \dots, u_n$ of $\overline{G}$ and all vectors d such that $d^\top d = 1$.

Proof. Let $\vartheta := \vartheta(G)$.

$(\geq)$ First we exhibit a vector c and an orthonormal representation $v_1, \dots, v_n$ of G such that $c^\top c = 1$, and $c^\top v_i = \frac{1}{\sqrt{\vartheta}}$ for all $i \in [n]$. To this end, let A be a symmetric $n \times n$-matrix where $A_{ij} = 1$ if $i = j$ or i, j are non-adjacent vertices, and whose largest eigenvalue is ϑ. Since $\vartheta I - A$ is PSD, it can be expressed as $X^\top X$ for some matrix X with n columns $x_1, \dots, x_n$. As $\vartheta I - A$ is singular, X does not have full column rank, so there exists a vector c that is perpendicular to its columns, and $c^\top c = 1$. Let $v_i := \frac{1}{\sqrt{\vartheta}}(c + x_i)$ for each $i \in [n]$. It can be readily checked that $v_1, \dots, v_n$ is an orthonormal representation of G, and $c^\top v_i = \frac{1}{\sqrt{\vartheta}}$ for each $i \in [n]$. Now, to prove $(\geq)$, take an orthonormal representation $u_1, \dots, u_n$ of $\overline{G}$ and a vector d such that $d^\top d = 1$. Then

$$1 = (c^\top c)(d^\top d) \geq \sum_{i=1}^{n} (c^\top v_i)^2 (d^\top u_i)^2 = \frac{1}{\vartheta} \sum_{i=1}^{n} (d^\top u_i)^2$$

where the middle inequality follows from Lemma 5.1. Multiplying both sides by ϑ gives the desired inequality.

$(\leq)$ We shall exhibit an orthonormal representation $u_1, \dots, u_n$ of $\overline{G}$ and a vector $d \in \mathbb{R}^n$ such that $d^\top d = 1$, and $\vartheta \leq \sum_{i=1}^{n} (d^\top u_i)^2$. We know from Theorem 4.2 that ϑ is equal to $\mathrm{tr}(BJ)$ for some PSD $n \times n$-matrix B of trace 1 such that $B_{ij} = 0$ for all $ij \in E$. We may express B as $Y^\top Y$ for some matrix Y with n columns $y_1, \dots, y_n$. As $B_{ij} = 0$ for all $ij \in E$, we have $y_i^\top y_j = 0$ for all $ij \in E$. Subsequently, $u_i := \frac{1}{\sqrt{y_i^\top y_i}} y_i, i \in [n]$ is an orthonormal representation of $\overline{G}$. As $\mathrm{tr}(B) = 1$ and $\vartheta = \mathrm{tr}(BJ) = \mathbf{1}^\top B \mathbf{1}$, we have

$$\sum_{i=1}^{n} y_i^\top y_i = 1$$

$$\left(\sum_{i=1}^{n} y_i \right)^\top \left(\sum_{i=1}^{n} y_i \right) = \vartheta.$$

Let $d := \frac{1}{\sqrt{\vartheta}} \left(\sum_{i=1}^{n} y_i \right)$. Then $d^\top d = 1$. Moreover,

$$\begin{aligned}
\sum_{i=1}^{n} (d^\top u_i)^2 &= \left(\sum_{i=1}^{n} y_i^\top y_i \right) \left(\sum_{i=1}^{n} (d^\top u_i)^2 \right) \\
&\geq \left(\sum_{i=1}^{n} \sqrt{y_i^\top y_i} \cdot (d^\top u_i) \right)^2 \qquad \text{by the Cauchy-Schwarz inequality} \\
&= \left(d^\top \sum_{i=1}^{n} y_i \right)^2 \\
&= \vartheta,
\end{aligned}$$

as required. □

Let us now use this formula to get an upper-bound on $\vartheta(G)$ contrasting nicely with Lemma 4.1.

Lemma 5.3. $\alpha^\star(G) \geq \vartheta(G)$.

Proof. By Theorem 5.2, $\vartheta(G) = \sum_{i=1}^{n} (d^\top u_i)^2$ for some orthonormal representation $u_1, \dots, u_n$ of $\overline{G}$ and some vector d such that $d^\top d = 1$. Then for any clique K of G, the vectors $u_i, i \in K$ form an orthonormal set, so $1 = d^\top d \geq \sum_{i \in K} (d^\top u_i)^2$. Thus the vector $x_i := (d^\top u_i)^2, i \in [n]$ satisfies $x \geq \mathbf{0}$ and $x(K) \leq 1$ for every clique K, so $\vartheta(G) = \mathbf{1}^\top x \leq \alpha^\star(G)$, as required. □

In summary, we proved in this section and the previous one that $\alpha^\star(G) \geq \vartheta(G) \geq \alpha(G)$. In particular, if equality holds throughout, which is the case for any perfect graph G, then by Theorem 4.3, we can compute $\alpha(G)$ in polynomial time.

Theorem 5.4. There is a polynomial algorithm for computing the stability number of any perfect graph G.

Exercises

1. Let K be a compact convex subset of $\mathbb{R}^n$ containing $\mathbf{0}$. Define the *polar of* K, denoted $K^\star$, as the set of all vectors $y \in \mathbb{R}^n$ such that $y^\top x \le 1$ for all $x \in K$. Prove the following statements:

   (a) $K^\star$ is a compact convex subset of $\mathbb{R}^n$ containing $\mathbf{0}$,

   (b) $(K^\star)^\star = K$.

2. Let K be a compact convex subset of $\mathbb{R}^n$ containing $\mathbf{0}$. Given an algorithm for the optimisation problem over the polar $K^\star$, derive an algorithm for the separation problem over K.

3. Prove Lemma 2.1.

4. Prove there is a polynomial algorithm that finds a minimum weight signature of a weakly bipartite signed graph $(G = (V, E), \Sigma)$ with edge weights $w \in \mathbb{R}^E$.

5. Let $\mathcal{K}$ be a family of ideal clutters. Suppose the optimisation problem

$$\min\{c^\top x : x(C) \ge 1 \ \forall C \in \mathcal{C}, \mathbf{1} \ge x \ge \mathbf{0}\}$$

   is polynomially solvable for any clutter $\mathcal{C}$ in the family $\mathcal{K}$, and any rational cost vector c. Prove that the optimisation problem

$$\min\{b^\top x : x(B) \ge 1 \ \forall B \in b(\mathcal{C}), \mathbf{1} \ge x \ge \mathbf{0}\}$$

   is polynomially solvable for the blocker $b(\mathcal{C})$ of any clutter $\mathcal{C} \in \mathcal{K}$, and any rational cost vector b. (**Hint.** You may use the following fact without proof: If $\mathcal{C}$ is an ideal clutter, then the linear program $\min\{c^\top x : x(C) \ge 1 \ \forall C \in \mathcal{C}, \mathbf{1} \ge x \ge \mathbf{0}\}$ has an integral optimal solution for all c.)

6. Prove Theorem 3.1. (**Hint.** Given a graph G and an integer k, deciding whether G has a clique of size at least k is NP-complete.)

7. Prove Lemma 5.1.

8. Prove that for any n-vertex simple graph G, we have $\vartheta(G)\vartheta(\overline{G}) \ge n$.

9. Let $G = ([n], E)$ be a simple graph. Given an orthonormal representation $u_1, \ldots, u_n$ of G, define its *value* as

$$\min_{c} \max_{1 \le i \le n} \frac{1}{(c^\top u_i)^2}$$

   where the minimum ranges over all vectors $c \in \mathbb{R}^n$ such that $c^\top c = 1$. Prove that $\vartheta(G)$ is equal to the minimum value that an orthonormal representation of G can achieve. (**Hint.** Read the proof of Theorem 5.2 for getting started.)

10. Given a simple graph G with vertex set $[n]$ and an integer $k \geq 1$, denote by G^k the graph with vertex set $[n]^k$ where distinct vertices $(u_1, \ldots, u_k), (v_1, \ldots, v_k)$ are adjacent if for each $i \in [k]$, the two vertices u_i, v_i of G are either equal or adjacent. Prove that $\vartheta(G^k) = \vartheta(G)^k$.

11. Given a simple graph G, its *Shannon capacity* is $\Theta(G) := \sup_k \sqrt[k]{\alpha(G^k)}$. Prove the following statements:

    (a) $\Theta(G) \geq \alpha(G)$,

    (b) $\Theta(G) = \lim_{k\to\infty} \sqrt[k]{\alpha(G^k)}$,

    (c) $\Theta(G) \leq \alpha^\star(G)$,

    (d) $\Theta(G) \leq \vartheta(G)$.

12. Prove that the Shannon capacity of the 5-cycle is $\sqrt{5}$.

Acknowledgements

The notion of $\vartheta(G)$ is due to Lovász [5], and so are Theorem 4.2 and Theorem 5.2. The proofs I presented closely follow Lovász's original proof. His paper is an absolute delight to read. The algorithmic consequences in this lecture are all eventually due to [1].

References

[1] M. Grötschel, L. Lovász, and A. Schrijver. The ellipsoid method and its consequences in combinatorial optimization. *Combinatorica*, 1(2):169–197, 1981.

[2] M. Grötschel, L. Lovász, and A. Schrijver. *Geometric Algorithms and Combinatorial Optimization*, volume 2 of *Algorithms and Combinatorics*. Springer Berlin, Heidelberg, 1988.

[3] B. Guenin. A characterization of weakly bipartite graphs. *Journal of Combinatorial Theory, Series B*, 83(1):112 – 168, 2001.

[4] L. Khachiyan. A polynomial algorithm in linear programming. *Doklady Akademii Nauk SSSR*, 244:1093–1096, 1979.

[5] L. Lovász. On the shannon capacity of a graph. *IEEE Transactions on Information Theory*, 25(1):1–7, 1979.

[6] N. Shor. Convergence rate of the gradient descent method with dilatation of the space. *Kibernetika*, pages 80–85, 1970.