Spectral Graph Theory

Gabriel Coutinho

January 25, 2022

These are the course notes of a (under)grad course being offered at UFMG in 2021.2.

Contents

1 Symmetric matrices and adjacency of a graph

In this section, we shall introduce the basic theory of symmetric matrices, including a result generally overlooked in a first or second linear algebra course. We shall define the adjacency matrix of a graph, and then make connections between the algebraic properties of this matrix and the combinatorial properties of the graph.

1.1 Symmetric matrices

We shall work over the vector space $\mathbb{R}^n$. If $\mathbf{u}, \mathbf{v} \in \mathbb{R}^n$, then $\langle \mathbf{v}, \mathbf{u} \rangle = \mathbf{v}^T \mathbf{u}$ is an inner product (meaning, it is a positive-definite commutative bilinear form). A linear operator $\mathbf{M} : \mathbb{R}^n \to \mathbb{R}^n$ is self-adjoint if $\langle \mathbf{M}\mathbf{v}, \mathbf{u} \rangle = \langle \mathbf{v}, \mathbf{M}\mathbf{u} \rangle$ for all $\mathbf{u}$ and $\mathbf{v}$, and, because $\mathbf{M}$ can (and will) be seen as a square matrix, it follows that $\mathbf{M}$ is a self-adjoint operator if and only if $\mathbf{M} = \mathbf{M}^T$, that is, $\mathbf{M}$ is a symmetric matrix. Symmetric matrices enjoy two key important properties: they are diagonalizable by orthogonal eigenvectors, and all of their eigenvalues are real. We start proving both properties.

Lemma 1.1. *The eigenvalues of a real symmetric matrix are real numbers.*

Proof. Let $\mathbf{M}\mathbf{u} = \lambda \mathbf{u}$, with $\mathbf{u} \neq \mathbf{0}$. Some of these things could be complex numbers, so we can take the conjugate on both sides, recovering

$$\mathbf{M}\overline{\mathbf{u}} = \overline{\lambda}\overline{\mathbf{u}}.$$

Thus $\overline{\mathbf{u}}$ is an eigenvector with eigenvalue $\overline{\lambda}$. Thus

$$\lambda \mathbf{u}^T \overline{\mathbf{u}} = (\mathbf{M}\mathbf{u})^T \overline{\mathbf{u}} = \mathbf{u}^T (\mathbf{M}\overline{\mathbf{u}}) = \overline{\lambda} \mathbf{u}^T \overline{\mathbf{u}}.$$

Because $\mathbf{u}^T \overline{\mathbf{u}} \neq 0$ if $\mathbf{u} \neq 0$, then $\lambda = \overline{\lambda}$. □

Now simply assume whenever we are dealing with a symmetric matrix, its eigenvalues are real, and any eigenvector can be assumed to be real.

Lemma 1.2. *Let* $\mathbf{M}$ *be a real symmetric matrix, and assume* $\mathbf{u}$ *and* $\mathbf{v}$ *are eigenvectors associated to different eigenvalues. Then* $\mathbf{v}^T \mathbf{u} = 0$*, that is, they are orthogonal.*

Proof. Say $\mathbf{M}\mathbf{u} = \lambda \mathbf{u}$ and $\mathbf{M}\mathbf{v} = \mu \mathbf{v}$, with $\lambda \neq \mu$. It follows that

$$\lambda(\mathbf{v}^T \mathbf{u}) = \mathbf{v}^T \mathbf{M}\mathbf{u} = (\mathbf{v}^T \mathbf{M}\mathbf{u})^T = \mathbf{u}^T \mathbf{M}^T \mathbf{v} = \mathbf{u}^T \mathbf{M}\mathbf{v} = \mu(\mathbf{u}^T \mathbf{v}) = \mu(\mathbf{v}^T \mathbf{u}).$$

As $\lambda \neq \mu$, it must be that $\mathbf{v}^T \mathbf{u} = 0$. □

The lemma above already implies that if $\mathbf{M}$ is diagonalizable, then it is diagonalizable with orthogonal eigenvectors — as, in fact, we eigenvectors corresponding to distinct eigenvalues are orthogonal, and inside each eigenspace we can always find an orthogonal basis. We move forward.

A subspace U of $\mathbb{R}^n$ is said to be $\mathbf{M}$-invariant if, for all $\mathbf{u} \in U$, $\mathbf{M}\mathbf{u} \in U$. This is a key fundamental concept in linear algebra, and several results are proven by noting that certain subspaces are invariant for certain operator.

Lemma 1.3. *Let* $\mathbf{M}$ *be a real symmetric matrix. If* U *is* $\mathbf{M}$*-invariant, then* $U^\perp$ *is also* $\mathbf{M}$*-invariant.*

Proof. Note that $\mathbf{v} \in U^\perp$, by definition, if $\mathbf{v}^T\mathbf{u} = 0$ for all $\mathbf{u} \in U$. For all $\mathbf{u} \in U$ and $\mathbf{v} \in U^\perp$, note that

$$(\mathbf{M}\mathbf{v})^T\mathbf{u} = \mathbf{v}^T\mathbf{M}\mathbf{u} = \mathbf{v}^T(\mathbf{M}\mathbf{u}) = 0,$$

because $\mathbf{u} \in U$, U is $\mathbf{M}$-invariant, and so $\mathbf{M}\mathbf{u} \in U$, and $\mathbf{v} \in U^\perp$. Thus $\mathbf{M}\mathbf{v} \in U^\perp$, as we wanted. □

Let λ be such that $\det(\lambda\mathbf{I} - \mathbf{M}) = 0$. Then $\lambda\mathbf{I} - \mathbf{M}$ is singular, and therefore it contains at least one non-zero vector in its kernel. This is saying that all square matrices $\mathbf{M}$ contain at least one eigenvector for each root of $\phi_{\mathbf{M}}(x) = \det(x\mathbf{I} - \mathbf{M})$. As $\mathbf{M}$ is symmetric, we now know that all possible roots of $\phi_{\mathbf{M}}$ are real.

Lemma 1.4. *Let* U *be an* $\mathbf{M}$*-invariant subspace with dimension* ≥ 1*. Then there is one eigenvector of* $\mathbf{M}$ *in* U*.*

Proof. Let $\mathbf{P}$ be a matrix whose columns form an orthonormal basis for U. As U is $\mathbf{M}$-invariant, it follows that there is a matrix $\mathbf{N}$ so that

$$\mathbf{M}\mathbf{P} = \mathbf{P}\mathbf{N}.$$

(Stop now and think carefully why this equality is true.) In particular, $\mathbf{N} = \mathbf{P}^T\mathbf{M}\mathbf{P}$, so $\mathbf{N}$ is symmetric. Let $\mathbf{u}$ be one eigenvector of $\mathbf{N}$ with eigenvalue λ. Then

$$\mathbf{M}\mathbf{P}\mathbf{u} = \mathbf{P}\mathbf{N}\mathbf{u} = \lambda\mathbf{P}\mathbf{u},$$

and, moreover $\mathbf{P}\mathbf{u} \neq \mathbf{0}$, as the columns of $\mathbf{P}$ are linearly independent. Thus $\mathbf{P}\mathbf{u}$ is an eigenvector for $\mathbf{M}$ in U. □

These four lemmas above are all you need to prove the following result by induction as an exercise.

Theorem 1.5. *Let* $\mathbf{M}$ *be a real symmetric matrix. Then* $\mathbf{M}$ *is diagonalizable by set of orthogonal eigenvectors, all of them corresponding to real eigenvalues.*

Exercise 1.6. Write the proof of this theorem as an exercise.

Corollary 1.7. *Let* $\mathbf{v}_1, ..., \mathbf{v}_n$ *be an orthonormal basis of eigenvectors for* $\mathbf{M}$*, each corresponding to an eigenvalue* $\lambda_1, ..., \lambda_n$ *(these are not necessarily distinct). Let* $\mathbf{P}$ *be the matrix whose* i*th column is* $\mathbf{v}_i$*, and* Λ *the diagonal matrix whose* i*th diagonal element is* λ_i*. Then*

$$\mathbf{P}^T\mathbf{M}\mathbf{P} = \boldsymbol{\Lambda},$$

and

$$\mathbf{M} = \lambda_1(\mathbf{v}_1\mathbf{v}_1^T) + ... + \lambda_n(\mathbf{v}_n\mathbf{v}_n^T).$$

Proof. A linear operator is defined and determined by its action on a basis. The first equality follows from the fact that both sides act equally on the canonical basis of $\mathbb{R}^n$. The second follows from

$$\mathbf{M} = \mathbf{P}\mathbf{\Lambda}\mathbf{P}^T,$$

and, by definition of matrix product, $\mathbf{M} = \mathbf{v}_1(\lambda_1\mathbf{v}_1^T) + ... + \mathbf{v}_n(\lambda_n\mathbf{v}_n^T)$. □

You should recall right now that, because $\mathbf{v}_i$ is normalized, then $\mathbf{P}_i = \mathbf{v}_i\mathbf{v}_i^T$ is the matrix that represents the orthogonal projection onto the line spanned by $\mathbf{v}_i$, that is, $\mathbf{P}_i$ is a projection as $\mathbf{P}_i^2 = \mathbf{P}_i$, and it is an orthogonal projection as $\mathbf{P}_i$ is symmetric. Note that $\mathbf{P}_i\mathbf{P}_j = \mathbf{0}$ whenever $i \neq j$, and so any sum of the $\mathbf{P}_i$s for distinct indices will correspond to the orthogonal projection onto the space spanned by the $\mathbf{v}_i$s of the same indices. In particular $\sum_{i=1}^n \mathbf{P}_i = \mathbf{I}$.

Exercise 1.8. Assume $\mathbf{P}_i$s are orthogonal projections. Show that $\mathbf{P}_1 + \mathbf{P}_2$ is an orthogonal projection if and only if $\mathbf{P}_1\mathbf{P}_2 = \mathbf{0}$.

Show now that $\mathbf{P}_1 + ... + \mathbf{P}_k$ is an orthogonal projection if and only if $\mathbf{P}_i\mathbf{P}_j = \mathbf{0}$ for $i \neq j$.

Say $\mathbf{M}$ is an $n \times n$ symmetric matrix with distinct eigenvalues $\theta_0, ..., \theta_d$. When we write the second equation from the statement of Corollary 1.7, we can collect the terms corresponding to equal eigenvalues, and have

$$\mathbf{M} = \sum_{r=0}^{d} \theta_r \mathbf{E}_r, \tag{1}$$

where, according to the discussion above, each $\mathbf{E}_r$ corresponds to the orthogonal projection onto the θ_r eigenspace. Equation (1) is usually referred to the as the *spectral decomposition* of the matrix $\mathbf{M}$.

Exercise 1.9. Find the spectral decomposition of

$$\mathbf{M} = \begin{pmatrix} 1+\sqrt{2} & 0 & 1-\sqrt{2} & 0 \\ 0 & 1+\sqrt{2} & 0 & 1-\sqrt{2} \\ 1-\sqrt{2} & 0 & 1+\sqrt{2} & 0 \\ 0 & 1-\sqrt{2} & 0 & 1+\sqrt{2} \end{pmatrix}$$

Hint: do not try to compute the characteristic polynomial. It is easier to simply try to look and guess which are the eigenvectors and eigenvalues.

Note that the $\mathbf{E}_r$ are symmetric matrices satisfying $\mathbf{E}_r\mathbf{E}_s = \delta_{rs}\mathbf{E}_r$, and $\sum_{r=0}^d \mathbf{E}_r = \mathbf{I}$.

Exercise 1.10. Prove (or at least convince yourself) that for any polynomial $p(x)$, it follows that

$$p(\mathbf{M}) = \sum_{r=0}^{d} p(\theta_r)\mathbf{E}_r.$$

Exercise 1.11. Let $\mathbf{M}$ be a symmetric matrix, with spectral decomposition as in (1).
(A) What is the minimal polynomial of $\mathbf{M}$? (B) Prove that for each $\mathbf{E}_r$, there is a polynomial p_r of degree d so that $p_r(\mathbf{M}) = \mathbf{E}_r$. Describe this polynomial as explicitly as you can.

Exercise 1.12. Prove that two symmetric matrices $\mathbf{M}$ and $\mathbf{N}$ commute if and only if they can be simultaneously diagonalized by the same set of orthonormal eigenvectors. Is it true that if $\mathbf{M}$ and $\mathbf{N}$ commute, then there is always a polynomial p so that $p(\mathbf{M}) = \mathbf{N}$? Characterize what else you need to observe to guarantee that such polynomial exists.

Exercise 1.13. Let $\mathbf{A}$ and $\mathbf{B}$ be matrices (not necessarily squared shaped), so that both products $\mathbf{AB}$ and $\mathbf{BA}$ are defined. Prove that

$$\operatorname{tr} \mathbf{AB} = \operatorname{tr} \mathbf{BA},$$

and conclude that if $\mathbf{M}$ is a symmetric matrix with eigenvalues $\lambda_1, ..., \lambda_n$, then $\operatorname{tr} \mathbf{M}$ is equal to $\lambda_1 + ... + \lambda_n$. How about $\operatorname{tr} \mathbf{M}^2$?

1.2 The adjacency matrix of a graph

Given a graph G on a vertex set V, one can always define an arbitrary ordering to the vertices, that is, let $V = \{a_1, ..., a_n\}$, and encode the graph as a symmetric 01-matrix as follows. *The adjacency matrix* $\mathbf{A}$ of G is defined as $\mathbf{A}_{ij} = 1$ of $a_i \sim a_j$, and $\mathbf{A}_{ij} = 0$ otherwise (including the diagonal elements).

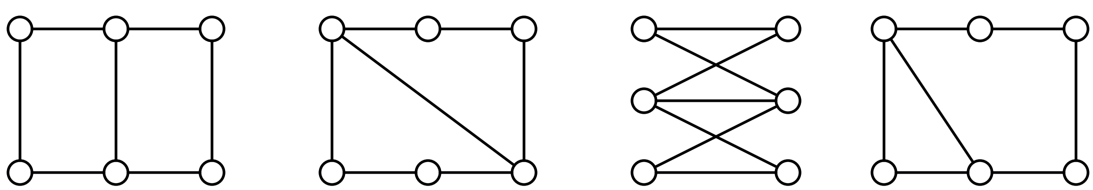
The field of spectral graph theory concerns itself with the main problem of relating spectral properties of matrices that encode adjacency in a graph (such as $\mathbf{A}$) with the combinatorial properties of the graph. We shall see several examples of such relations.

Exercise 1.14. Let G be a graph, suppose the vertices V are ordered, and let $\mathbf{A}$ be the corresponding adjacency matrix of G. Suppose you reorder the vertices by means of a permutation. Let $\mathbf{P}$ be the 01 matrix representing this permutation. Show that the new adjacency matrix obtained from this re-ordering is $\mathbf{PAP}^T$. Conclude that the eigenvalues are the same, and the only change in the eigenvectors is a permutation of its entries.

Because of this exercise, we shall simply ignore the underlying ordering, and speak of "the" adjacency matrix of G.

Given two graphs $G_1 = (V_1, E_1)$ and $G_2 = (V_2, E_2)$ on the same number of vertices, a very natural question is whether or not they encode the same combinatorial structure, which can be translated as: is there a function $f : V_1 \to V_2$ that maps edges to edges and non-edges to non-edges? Such a function, if it exists, is called a *graph isomorphism.* You can think of an isomorphism like this: draw both graphs in the plane, and try to move the vertices of one of them (without creating or destroying edges) so that the two drawings look exactly the same.

Example 1.15. Graphs G_1, G_2 and G_3 are all isomorphic, but G_4 is "different".



G_1 G_2 G_3 G_4

Two isomorphic graphs can always be seen as graphs on the same vertex set, and the isomorphism is a re-ordering that preserves adjacency and non-adjacency. Thus:

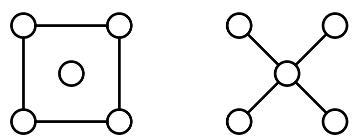
Theorem 1.16. *Let G and H be isomorphic graphs. Order their vertex sets from 1 to n, and let $\mathbf{P}$ be the permutation matrix that corresponds to the isomorphism from G to H. Then*

$$\mathbf{P}\mathbf{A}(G)\mathbf{P}^T = \mathbf{A}(H).$$

As a consequence, $\mathbf{A}(G)$ and $\mathbf{A}(H)$ have the same eigenvalues. □

Exercise 1.17. Order the vertices of G_1 and G_2 equally in terms of their geometric position. Then find the matrix $\mathbf{P}$ so that $\mathbf{P}\mathbf{A}(G_1)\mathbf{P} = \mathbf{A}(G_2)$. Compute the eigenvalues of G_1 and G_4 (using a software?) and conclude that they cannot be isomorphic.

One of the motivations of the development of spectral graph theory was the hope that two graphs would be isomorphic if and only if they had the same eigenvalues. Such a claim would immediately provide an efficient polynomial time algorithm to decide whether two graphs are isomorphic (and yet no such algorithm is known to this day). Two graphs with the same eigenvalues are called *cospectral graphs*. The following pair of graphs are the smallest known cases of cospectral but (clearly) non-isomorphic graphs. They have spectrum $2, 0^{(3)}, -2$.



G_1 G_2

This example also shows that the spectrum of a graph does not determine whether the graph is connected or not. This immediately raises the general question: what graph properties can be determined from the spectrum?

A *walk* of length r in a graph G is a sequence of $r+1$ (possibly repeated) vertices $a_0, ..., a_r$ with the property that $a_i \sim a_j$. A walk is *closed* if $a_0 = a_r$.

Lemma 1.18. *The number of distinct walks of length r from a to b in G is precisely equal to $(\mathbf{A}^r)_{ab}$.*

Exercise 1.19. Verify this result on at least 3 different graphs checking powers $r = 1, 2, 3$ for each. Then, sketch a proof by induction of this result.

Corollary 1.20. *If G has diameter D, then it must have at least $D+1$ distinct eigenvalues.*

Proof. Let

$$\mathbf{A}(G) = \sum_{r=0}^{d} \theta_r \mathbf{E}_r$$

be the spectral decomposition of $\mathbf{A}(G)$. Let W be the subspace of $\mathrm{Sym}_n(\mathbb{R})$ generated by $\{\mathbf{A}^0, \mathbf{A}, \mathbf{A}^2, ...\}$. As we saw in the past section, all powers of $\mathbf{A}$ are a linear combinations of the $\mathbf{E}_r$s, and each $\mathbf{E}_r$ is a polynomial in $\mathbf{A}$. Moreover, the matrices $\mathbf{E}_r$ are pairwise orthogonal, thus they are all linearly independent. As a consequence, $\dim W = d+1$, and $\{\mathbf{E}_0, ..., \mathbf{E}_d\}$ form a basis for W. Now observe that if $r \leq D$, then at least one entry of $\mathbf{A}^r$ is non-zero "for the first time", meaning that it was equal to 0 for all smaller powers of $\mathbf{A}$. Thus $\{\mathbf{A}^0, \mathbf{A}, \mathbf{A}^2, ..., \mathbf{A}^D\}$ form a linearly independent set in W, and $D \leq d$. □

Let us now return to the problem of deciding what can be determined by the spectrum of a graph alone. Clearly the number of vertices in a graph is determined by the spectrum. An immediate consequence of the Lemma 1.18 is that the number of edges is also determined by the spectrum.

Corollary 1.21. *Let G be a graph on n vertices, with m edges, and let $\lambda_1, ..., \lambda_n$ the eigenvalues of $\mathbf{A}(G)$. Then*

$$\lambda_1^2 + ... + \lambda_n^2 = 2m.$$

Proof. Both sides are equal to $\operatorname{tr} \mathbf{A}^2$. □

Exercise 1.22. Find a formula for the number of triangles (cycles of length 3) found as subgraphs of G that depends only on the eigenvalues of G. Explain why the number of cycles of length 4 is not determined by the spectrum alone (as you witnessed in the example above).

Exercise 1.23. Does the spectrum alone determines the length of the shortest odd cycle of a graph? Explain.

Exercise 1.24. If G has n vertices, prove that all eigenvalues lie in the interval $(-n, n)$.

Exercise 1.25. Let G be a k-regular graph (that is, all vertices have k neighbours). Prove that k is an eigenvalue for G by describing a corresponding eigenvector.

Let $\mathbf{J}$ stand for the matrix whose all entries are equal to 1. If G is a graph, let $\overline{G}$ stand for the complement graph of G, that is, the graph whose edges are precisely the non-edges of G. Then, clearly,

$$\mathbf{A}(\overline{G}) = \mathbf{J} - \mathbf{A}(G) - \mathbf{I}.$$

As immediate consequence of the past exercise, we have:

Lemma 1.26. *Let G be a k-regular graph, with eigenvalues $k = \lambda_1, ..., \lambda_n$. Then the eigenvalues of $\overline{G}$ are*

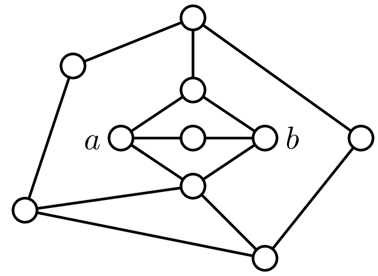
$$n-k-1,\ -\lambda_2 - 1,\ ...,\ -\lambda_n - 1.$$

Proof. The all 1s vector $\mathbf{1}$ is an eigenvector of G. Let $\mathbf{v}_2, ..., \mathbf{v}_n$ complete a basis of orthogonal eigenvectors. Then

$$(\mathbf{J} - \mathbf{A}(G) - \mathbf{I})\mathbf{1} = (n-k-1)\mathbf{1} \quad \text{and} \quad (\mathbf{J} - \mathbf{A}(G) - \mathbf{I})\mathbf{v}_i = -\lambda_i - 1,$$

as $\mathbf{J}\mathbf{v}_i = \mathbf{0}$ because $\mathbf{1}$ and $\mathbf{v}_i$ are orthogonal. □

Exercise 1.27. Assume G contains a pair of vertices a and b so that the neighbourhood of a is equal to neighbourhood of b (the rest of the graph can be anything). For example:



(a) Prove that 0 is an eigenvalue of this graph (Hint: look at a and b and try to produce one eigenvector for 0). If the example looks too complicate, forget about the 5-cycle and focus only on a, b and their neighbours.

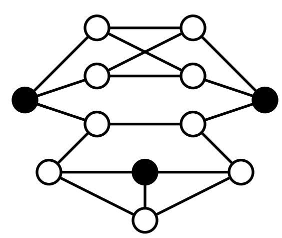
(b) What could you say if a and b shared the same neighbourhood, but were also neighbours themselves?

Exercise 1.28. Assume $G = (V, E)$ is a k-regular graph which contains a subset of vertices $U \subseteq V$ satisfying the following properties:

(a) No two vertices in U are neighbours.

(b) Any vertex in $V \backslash U$ contains exactly one neighbour in U.

Prove that if such U exists, then -1 is an eigenvalue of the graph. (Hint: recall G is assumed to be k-regular, and, again, try to produce one eigenvector. Try first in the example below, where the dark vertices are the vertices in U.)



In this next section, we shall see that two important properties about a graph can be determined from its spectrum alone: whether the graph is regular, and whether the graph is bipartite.

1.3 Perron-Frobenius (a special case)

Let $\mathbf{M}$ be a real $n \times n$ matrix with nonnegative entries. For example, the adjacency matrix of a graph. This matrix is called primitive if, for some integer k, $\mathbf{M}^k > 0$, and it is called irreducible if for all indices i and j, there is an integer k so that $(\mathbf{M}^k)_{ij} > 0$. All primitive matrices are irreducible, but the converse is not necessarily true.

Example 1.29. Consider

$$\begin{pmatrix}0&1&1\\1&0&1\\1&1&0\end{pmatrix} \quad , \quad \begin{pmatrix}0&1\\1&0\end{pmatrix} \quad , \quad \begin{pmatrix}0&1&0\\0&0&1\\1&0&0\end{pmatrix} \quad \text{and} \quad \begin{pmatrix}1&1&0\\1&1&0\\0&0&1\end{pmatrix}$$

Verify that the first is primitive, the second and third are both irreducible, but not primitive, and the fourth is neither.

Exercise 1.30. Prove that if $\mathbf{M}$ is irreducible, then $\mathbf{I} + \mathbf{M}$ is primitive.

Exercise 1.31. Let G be a graph. Show that

(a) $\mathbf{A}(G)$ is irreducible if and only if G is connected.

(b) $\mathbf{A}(G)$ is not primitive if G is bipartite.

Over the next few results, we shall actually see, amongst other things, that $\mathbf{A}(G)$ is irreducible but not primitive if and only if G is connected and bipartite. Results below are known as the Perron-Frobenius theory. This theory applies generally to matrices which are assumed to be irreducible and nothing else. We shall however add the hypothesis that the matrices are also symmetric, for the proofs become simpler and more meaningful, and our matrices will almost always be symmetric anyway.

Our first observation.

Lemma 1.32. *Let $\mathbf{M}$ be a nonnegative symmetric matrix, $\mathbf{M} \neq \mathbf{0}$. If λ is the largest eigenvalue of $\mathbf{M}$, then $\lambda > 0$.*

Proof. Follows immediately from $\operatorname{tr} \mathbf{M} \geq 0$. □

For any nonzero vector $\mathbf{u} \in \mathbb{R}^n$, and symmetric matrix $\mathbf{M}$, define

$$R_{\mathbf{M}}(\mathbf{u}) = \frac{\mathbf{u}^T \mathbf{M} \mathbf{u}}{\mathbf{u}^T \mathbf{u}}.$$

This is known as the Rayleigh quotient of $\mathbf{u}$ with respect to $\mathbf{M}$. Note that $R_{\mathbf{M}}(\alpha \mathbf{u}) = R_{\mathbf{M}}(\mathbf{u})$ for all $\alpha \neq 0$, so we shall typically assume $\mathbf{u}$ has been normalized. In a sense, this is a measurement of how much $\mathbf{M}$ displaces $\mathbf{u}$, also proportional to how much $\mathbf{M}$ stretches or shrinks $\mathbf{u}$. Therefore one should expect that this is maximum when $\mathbf{u}$ is an eigenvector of $\mathbf{M}$, corresponding to a large eigenvalue.

Lemma 1.33. *If $\mathbf{u}$ is eigenvector of $\mathbf{M}$ with eigenvalue θ, then $R_{\mathbf{M}}(\mathbf{u}) = \theta$. If λ is the largest eigenvalue of $\mathbf{M}$, then, for all $\mathbf{v} \in \mathbb{R}^n$, $R_{\mathbf{M}}(\mathbf{v}) \leq \lambda$. Equality holds for some $\mathbf{v}$ only if $\mathbf{v}$ is eigenvector for λ.*

Proof. Only the second and third assertions deserve a proof. Let $\mathbf{M} = \sum_{r=0}^{d} \theta_r \mathbf{E}_r$ be the spectral decomposition of $\mathbf{M}$. Assume θ_0 is the largest eigenvalue, and that $\mathbf{v}$ is a normalized vector. Then

$$\begin{aligned} R_{\mathbf{M}}(\mathbf{v}) = \mathbf{v}^T \mathbf{M} \mathbf{v} &= \theta_0(\mathbf{v}^T E_0 \mathbf{v}) + \theta_1(\mathbf{v}^T E_1 \mathbf{v}) + ... + \theta_d(\mathbf{v}^T E_d \mathbf{v}) \\ &\leq \theta_0((\mathbf{v}^T E_0 \mathbf{v}) + (\mathbf{v}^T E_1 \mathbf{v}) + ... + (\mathbf{v}^T E_d \mathbf{v}) = \theta_0. \end{aligned}$$

Equality holds if and only if $(\mathbf{v}^T E_r \mathbf{v}) = 0$ for all $r > 0$, which is the same as saying that $\mathbf{v}$ belongs to the θ_0 eigenspace. □

Lemma 1.34. *Let* $\mathbf{M}$ *be symmetric, non-negative and irreducible, with largest eigenvalue* λ. *There is a corresponding eigenvector* $\mathbf{u}$ *to* λ *so that* $\mathbf{u} > \mathbf{0}$.

Proof. Let $\mathbf{v}$ be a normal eigenvector for λ, and define $\mathbf{u}$ to be made from $\mathbf{v}$ by taking the absolute value at each entry (also denoted by $\mathbf{u} = |\mathbf{v}|$). Note that $\mathbf{u}$ is still normal, and, moreover

$$\lambda = R_{\mathbf{M}}(\mathbf{v}) = |R_{\mathbf{M}}(\mathbf{v})| \leq R_{\mathbf{M}}(\mathbf{u}) \leq \lambda.$$

(Second equality follows from $\lambda > 0$. First inequality from is simply the triangle inequality. Second follows from Lemma 1.33.)

Hence $R_{\mathbf{M}}(\mathbf{u}) = \lambda$, and $\mathbf{u}$ is an eigenvector for λ, with $\mathbf{u} \geq \mathbf{0}$. To see that $\mathbf{u} > \mathbf{0}$, note that as $\mathbf{M}$ is irreducible, it follows from Exercise 1.30 that $\mathbf{I} + \mathbf{M}$ is primitive, and so there is a k so that $(\mathbf{I}+\mathbf{M})^k > \mathbf{0}$. The vector $\mathbf{u}$ is also eigenvector for this matrix (with eigenvalue $(1+\lambda)^k$, but

$$\mathbf{0} < (\mathbf{I}+\mathbf{M})^k \mathbf{u} = (1+\lambda)^k \mathbf{u},$$

implying $\mathbf{u} > \mathbf{0}$. □

Lemma 1.35. *The largest eigenvalue* λ *of a symmetric, non-negative and irreducible matrix is simple (meaning, its eigenspace has dimension 1).*

Proof. From the proof of the past lemma, we know that no eigenvector for λ contains an entry equal to 0. No subspace of dimension larger than 1 can be such that all of its non-zero vectors have no non-zero entries. □

And finally:

Lemma 1.36. *Let* $\mathbf{M}$ *be symmetric, non-negative and irreducible. Let* λ *be its largest eigenvalue. Let* μ *be any other eigenvalue. Then* $\lambda \geq |\mu|$, *and, moreover, if* $-\lambda$ *is an eigenvalue, then* $\mathbf{M}^2$ *is not irreducible.*

Proof. Let $\mathbf{v}$ be an eigenvector for μ. As $\mathbf{v}$ is orthogonal to the positive eigenvector corresponding to λ, at least one entry of $\mathbf{v}$ is negative. Thus

$$|\mu| = |R_{\mathbf{M}}(\mathbf{v})| < R_{\mathbf{M}}(|\mathbf{v}|) \leq \lambda.$$

Now note that λ^2 is the largest eigenvalue of $\mathbf{M}^2$ (which is, still, symmetric and non-negative). If $-\lambda$ is eigenvalue of $\mathbf{M}$, then the eigenspace of λ^2 in $\mathbf{M}^2$ is at least 2-dimensional, thus $\mathbf{M}^2$ cannot be irreducible. □

It is quite surprising at first sight that the hypothesis on $\mathbf{M}$ being symmetric can be dropped entirely from the results above. The geometric intuition remains the same: a nonnegative irreducible matrix acts in the nonnegative orthant and there it encounters a unique direction which is an eigenvector. The proofs of these results are not hard per se, but I didn't feel they would add much to this notes. You are however invited to check any reference on spectral graph theory or non-negative matrix theory to find your favourite version of these results.

Now, to the applications.

Theorem 1.37. *Let* $\mathbf{A}$ *be the adjacency matrix of a connected graph* G*, and* $\lambda_1 \geq ... \geq \lambda_n$ *its spectrum.*

(a) G *is* k*-regular if and only if* $(1/n)(\lambda_1^2 + ... + \lambda_n^2) = \lambda_1$*, and, in this case,* $k = \lambda_1$*.*

(b) G *is bipartite if and only if* $\lambda_1 = -\lambda_n$*. If this is the case, then for all* λ_i*,* $-\lambda_i$ *is also an eigenvalue.*

Proof.

(a) Let $\mathbf{1}$ be the all 1s vector. The equality is equivalent to

$$R_{\mathbf{A}}(\mathbf{1}) = \lambda_1,$$

which, as we saw, is equivalent to $\mathbf{1}$ being an eigenvector of λ_1. This vector is eigenvector if and only if all row sums of $\mathbf{A}$ are equal, or, equivalently, all vertices have the same degree, which is going to be precisely equal to the eigenvalue λ_1.

(b) If G is bipartite, its adjacency matrix can always be written as

$$\mathbf{A} = \begin{pmatrix} \mathbf{0} & \mathbf{B} \\ \mathbf{B}^T & \mathbf{0} \end{pmatrix}.$$

If $\begin{pmatrix} \mathbf{v}_1 \\ \mathbf{v}_2 \end{pmatrix}$ is eigenvector for λ_i, then it is easy to see that $\begin{pmatrix} \mathbf{v}_1 \\ -\mathbf{v}_2 \end{pmatrix}$ is eigenvector for $-\lambda_i$. On the other hand, if $-\lambda_1$ is eigenvalue, then, from Lemma 1.36, it follows that $\mathbf{A}^2$ is not irreducible. Thus there are at least two vertices you can never walk from one to another with an even number of steps. Therefore there can be no odd cycles in this graph.

□

Corollary 1.38. *Let* λ *be the largest eigenvalue of* $\mathbf{A}(G)$*. Let* Δ *be the largest degree of* G*, and let* ∂ *be its average degree. Then*

$$\partial \leq \lambda \leq \Delta.$$

Proof. The first inequality follows from the fact that

$$\partial = R_{\mathbf{A}}(\mathbf{1}) \leq \lambda.$$

(Note in particular that this implies $\lambda \geq \delta$, where δ is the smallest degree of G). For the second, we have $\mathbf{A1} \leq \Delta\mathbf{1}$, and with $\mathbf{v}$ eigenvector for λ, we can multiply by $\mathbf{v}^T$ on the left. As $\mathbf{v} > \mathbf{0}$, the sign is preserved, and

$$\lambda \mathbf{v}^T\mathbf{1} = \mathbf{v}^T\mathbf{A1} \leq \Delta \mathbf{v}^T\mathbf{1},$$

so $\theta \leq \Delta$. □

Exercise 1.39. Prove that $\lambda \geq \sqrt{\Delta}$. (Hint: look at $\mathbf{A}^2$ and the proof above).

1.4 Eigenvalues of some classes of graphs

Consider the following classes of graphs:

(a) K_n - complete graphs on n vertices.

(b) $K_{a,b}$ - complete bipartite graphs with a vertices on one side, and b vertices on the other (in particular if $a = 1$, these are the stars).

(c) C_n - cycle graphs on n vertices.

(d) P_n - path graphs on n vertices.

Our goal here is to determine the eigenvalues (and eigenvectors) of these classes.

(a) This is easy. $\mathbf{A}(K_n) = \mathbf{J} - \mathbf{I}$. The eigenvalues of $\mathbf{J}$ are n (simple, with eigenvector $\mathbf{1}$) and 0 (all others). Thus the spectrum of K_n is $n-1$ and -1.

(b) Write
$$\mathbf{A} = \begin{pmatrix} \mathbf{0} & \mathbf{J}_{a,b} \\ \mathbf{J}_{b,a} & \mathbf{0} \end{pmatrix}.$$
There are $b-1$ vectors in the kernel of $\mathbf{J}_{a,b}$ and $a-1$ vectors in the kernel of $\mathbf{J}_{b,a}$. Each corresponding to an eigenvector for the eigenvalue 0 of $\mathbf{A}$. The two eigenvectors remaining are
$$\begin{pmatrix} \sqrt{b}\mathbf{1} \\ \sqrt{a}\mathbf{1} \end{pmatrix} \text{ and } \begin{pmatrix} \sqrt{b}\mathbf{1} \\ -\sqrt{a}\mathbf{1} \end{pmatrix},$$
corresponding to the eigenvalues $\sqrt{ab}$ and $-\sqrt{ab}$ respectively.

(c) This one is trickier. $\mathbf{A}(C_n)$ is the sum of two permutation matrices corresponding to the cycle $(123...n)$ and its inverse, say $\mathbf{P}$ and $\mathbf{P}^{-1}$. An eigenvector for a cyclic matrix can be easily built from an n-root of unity ω:
$$\mathbf{P}\begin{pmatrix} 1 \\ \omega \\ \vdots \\ \omega^{n-1} \end{pmatrix} = \begin{pmatrix} \omega^{n-1} \\ 1 \\ \vdots \\ \omega^{n-2} \end{pmatrix} = \omega^{n-1}\begin{pmatrix} 1 \\ \omega \\ \vdots \\ \omega^{n-1} \end{pmatrix} \text{ and}$$
$$\mathbf{P}^{-1}\begin{pmatrix} 1 \\ \omega \\ \vdots \\ \omega^{n-1} \end{pmatrix} = \begin{pmatrix} \omega \\ \omega^2 \\ \vdots \\ 1 \end{pmatrix} = \omega\begin{pmatrix} 1 \\ \omega \\ \vdots \\ \omega^{n-1} \end{pmatrix},$$
thus the eigenvalues are $\omega^{n-1} = \omega^{-1}$ and ω, hence the eigenvalues of $\mathbf{A}(C_n) = \mathbf{P} + \mathbf{P}^{-1}$ are $\omega^{-1} + \omega$ for all nth roots of unity, that is, $\omega = \mathrm{e}^{2\pi \mathrm{i}(k/n)}$, $k = 0, ..., n-1$. Thus the eigenvalues of C_n are
$$2\cos\left(2\pi\frac{k}{n}\right) \quad \text{for } k = 0, ..., n-1.$$
Note that 2 is always the largest (and simple) eigenvalue, and that -2 is an eigenvalue if and only if n is even. All other eigenvalues have multiplicity 2.

(d) We provide one way of finding this now. The other will come later as an exercise. Look at the cycle C_{2n+2}. Let ω be a $(2n+2)$th root of unity. Then

$$\begin{pmatrix} 1 \\ \omega \\ \vdots \\ \omega^{2n+1} \end{pmatrix} \text{ and } \begin{pmatrix} 1 \\ \omega^{-1} \\ \vdots \\ \omega^{-(2n+1)} \end{pmatrix}$$

are both eigenvalues of $\mathbf{A}(C_{2n+2})$ for $\omega + \omega^{-1}$, and so is any linear combination of them. In particular

$$\begin{pmatrix} 1 \\ \omega \\ \vdots \\ \omega^{2n+1} \end{pmatrix} - \begin{pmatrix} 1 \\ \omega^{-1} \\ \vdots \\ \omega^{-(2n+1)} \end{pmatrix} = \begin{pmatrix} 0 \\ \omega - \omega^{-1} \\ \vdots \\ \omega^{2n+1} + \omega^{-2n-1} \end{pmatrix}.$$

Note that there will be another 0 at position $n+2$, corresponding to $\omega^{n+1} - \omega^{-n-1} = -1 - (-1) = 0$. The n non-zero entries (only when $\omega \neq 1$) from positions 2 to $n+1$ are part of an eigenvector of C_{2n+2} which do not get interfered by the rest of the graph (those 0s at positions 1 and $n+2$ "disconnect" the eigenvector). Hence this part of the eigenvector is also an eigenvector for P_n (subgraph of C_{2n+2} from positions 2 to $n+1$). Therefore the spectrum of $\mathbf{A}(P_n)$ is

$$\omega + \omega^{-1} = 2\cos\left(\pi \frac{k}{n+1}\right) \quad \text{for } k = 1, ..., n.$$

1.5 Bounds to the largest and smallest eigenvalues

Later in this course we will study a powerful technique called interlacing, which allows for several interesting results relating eigenvalues and combinatorics. For now, however, we will see some examples of results one could obtain by means of ad-hoc ideas, applied to the largest and smallest eigenvalues.

Exercise 1.40. Let G be a graph with m edges, and let $\theta_1, ..., \theta_n$ be the eigenvalues of $\mathbf{A}$. Assume that they are numbered so that $\theta_i^2 \geq \theta_{i+1}^2$. Show that

$$\theta_i \leq \sqrt{\frac{2m}{i}}.$$

Recall now our usual notation for the eigenvalues of $\mathbf{A}$ with $\lambda_1 \geq ... \geq \lambda_n$. A consequence of the exercise above is that $\lambda_1 \leq \sqrt{2m}$. We can do better.

Theorem 1.41 (Hong)**.** *For a graph with n vertices and m edges,*

$$\lambda_1 \leq \sqrt{2m - (n-1)}.$$

Proof. Let $\mathbf{x}$ be an unit length eigenvector corresponding to λ_1, and let $\mathbf{x}(i)$ be the vector obtained from $\mathbf{x}$ by erasing all entries which do not correspond to neighbours of i. In particular, the ith entry of $\mathbf{x}(i)$ is also zeroed.

Note though that we still have

$$\mathbf{e}_i^{\mathsf{T}} \mathbf{A}\mathbf{x}(i) = \lambda_1 \mathbf{x}_i.$$

By Cauchy-Schwarz,

$$\lambda_1^2 \mathbf{x}_i^2 = |\mathbf{e}_i^{\mathsf{T}} \mathbf{A}\mathbf{x}(i)|^2 \leq |\mathbf{e}_i^{\mathsf{T}} \mathbf{A}|^2 |\mathbf{x}(i)|^2 = d_i \left(1 - \sum_{j \not\sim i} {\mathbf{x}_j}^2 \right)$$

and summing over all i, we get

$${\lambda_1}^2 \leq 2m - \sum_{i=1}^{n} d_i \left(\sum_{j \not\sim i} {\mathbf{x}_j}^2 \right).$$

It is not difficult to show that the subtracted term upper bounds $n-1$, with equality if and only if $d_i = 1$ or $d_i = n-1$ for all i. It follows that

$${\lambda_1}^2 \leq \sqrt{2m - (n-1)}.$$

Equality holds if and only if $X = K_n$ or $X = K_{1,n-1}$. □

We can also provide an interesting lower bound to λ_n.

Theorem 1.42 (Constantine, Hong, Powers)**.** *For a given graph with smallest eigenvalue λ_n,*

$$\lambda_n \geq -n/2.$$

Proof. Let $\mathbf{x}$ be a unit length eigenvector for λ_n, and suppose there are precisely a vertices whose coordinate is positive and b with negative. Then

$$\begin{aligned} \lambda_n = \mathbf{x}^{\mathsf{T}} \mathbf{A}\mathbf{x} &\geq \sum_{\mathbf{x}_i \mathbf{x}_j < 0} \mathbf{A}_{ij} \mathbf{x}_i \mathbf{x}_j \\ &\geq \sum_{\mathbf{x}_i \mathbf{x}_j < 0} \mathbf{x}_i \mathbf{x}_j \\ &\geq \lambda_n(K_{a,b}) \\ &= -\sqrt{ab} \\ &\geq -n/2. \end{aligned}$$

□

1.6 A result for regular graphs

The theorem in this subsection will reappear when we talk about interlacing. But it contains the nice, elementary proof below.

If the graph is regular, the size of the largest coclique can be associated with the smallest eigenvalue.

Theorem 1.43 (Delsarte)**.** *Let G be a k-regular graph, eigenvalues $k = \lambda_1 \geq .... \geq \lambda_n$. Say the size of the largest coclique of G is α. Then*

$$\alpha \leq \frac{n(-\lambda_n)}{k - \lambda_n}.$$

Proof. Let $\mathbf{x}$ be the characteristic vector of a largest coclique, meaning, $\mathbf{x}_i = 1$ if and only if i belongs to the coclique, and 0 otherwise. Note that

$$\mathbf{x}^\mathsf{T}\mathbf{A}\mathbf{x} = 0, \text{ thus } \mathbf{x}^\mathsf{T}(\mathbf{A} - \lambda_n\mathbf{I})\mathbf{x} = -\lambda_n \cdot \alpha.$$

On the other hand,

$$\mathbf{x}^\mathsf{T}(\mathbf{A} - \lambda_n\mathbf{I})\mathbf{x} = \sum_{i=1}^{n}(\lambda_i - \lambda_n)\mathbf{x}^\mathsf{T}\mathbf{E}_i\mathbf{x}.$$

All terms in the sum are non-negative, as $(\lambda_i - \lambda_n) \geq 0$ and $\mathbf{x}^\mathsf{T}\mathbf{E}_i\mathbf{x} \geq 0$. We can discard all of them but the first, and recalling that $\mathbf{1}$ is the eigenvector for λ_1, we have that $\mathbf{E}_1 = (1/n)\mathbf{J}$. Thus

$$-\lambda_n \cdot \alpha \geq \frac{\lambda_1 - \lambda_n}{n}\alpha^2,$$

and the result follows. □

1.7 Graphs with largest eigenvalue at most 2

It is still a common topic of research in spectral graph theory to classify all graphs whose spectrum lie within a constant given interval. One of the earliest and simplest results classifies all graphs whose largest eigenvalue does not surpass 2.

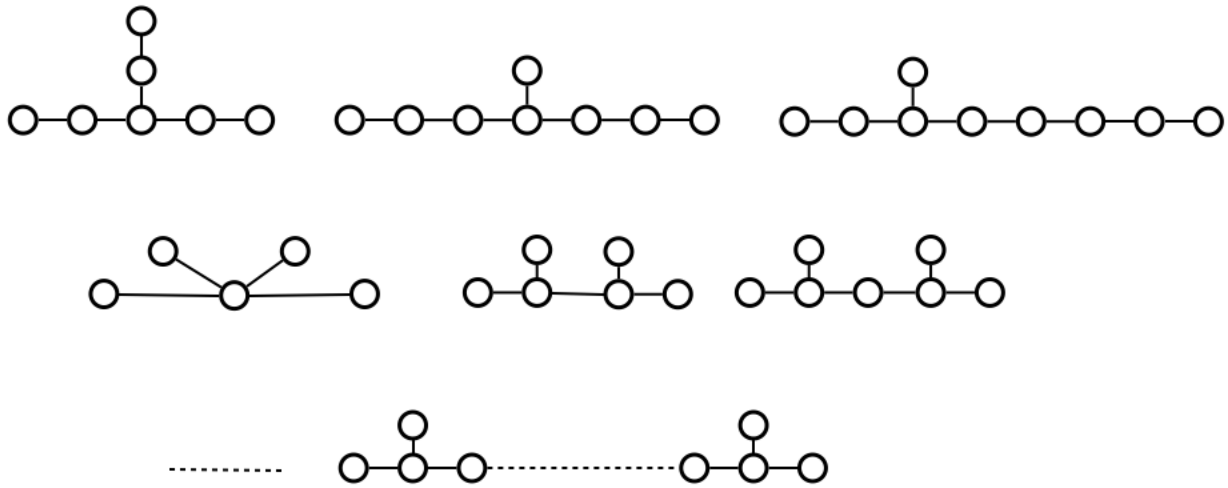
Lemma 1.44. *Let G be a connected graph, and assume H is a proper subgraph of G. Let $\lambda(G)$ be the largest eigenvalue of G, and $\lambda(H)$ the largest eigenvalue of H. Then*

$$\lambda(H) < \lambda(G).$$

Proof. Exercise! □

Exercise 1.45. Argue why any graph with largest eigenvalue at most 2 is either a tree or a cycle.

Exercise 1.46. Show that the following graphs all have largest eigenvalue equal to 2, by exhibiting a suitable eigenvector.



From the previous exercises, it follows that any graph with largest eigenvalue at most two is either a cycle, one of the trees above, or possibly a tree that contains none of the above as a subgraph. It is not difficult to see that any such tree must be a proper subgraph of the ones above.

1.8 References

Here is the set of references used to write the past few pages.

I used Chapter 8 of Godsil and Royle to write about the spectral decomposition of a symmetric matrix. This was also my reference for the basics and some exercises on the adjacency matrix.

(a) Chris Godsil and Gordon Royle. *Algebraic Graph Theory.* Springer-Verlag, New York, 2001.

Exercise 1.28 comes from Chan and Godsil "Symmetry and Eigenvectors".

I looked extensively for a nice intuitive proof of Perron-Frobenius in its full form, but the best I could do relied on using fixed point theorems. I then came up with the simplified version assuming matrices in question are symmetric. A good reference is Brouwer and Haemers, Chapter 2.

(b) Andries E Brouwer and Willem H Haemers. *Spectra of Graphs.* Universitext. Springer, New York, 2012

I also used the reference above for the spectrum of paths and cycles.

2 Graph isomorphism

Perhaps one of the nicest and most relevant applications of basic spectral graph theory is a polynomial-time algorithm to decide whether two graphs, both with simple eigenvalues, are isomorphic or not. At first one wonders if graphs have usually simple eigenvalues or not, and the answer is yes! This is no trivial result though, and was only settled in 2014 by Terence Tao and Van Vu. The consequence is that Graph Isomorphism is in P for almost all graphs. In this section, we will see how to construct such an algorithm. We will also develop some machinery which is interesting on its own.

2.1 The problem statement and a naive approach

Assume you are fed two graphs by means of their adjacency matrices. To decide whether or not these two matrices correspond to the same graph, you are basically asking the question of whether there exists a reordering of the rows and columns of one of the matrices that will leave it equal to the other.

Id est, you are given adjacency matrices $\mathbf{A}$ and $\mathbf{B}$, and you need to decide whether or not there is a permutation matrix $\mathbf{P}$ so that

$$\mathbf{P}\mathbf{A}\mathbf{P}^{\mathsf{T}} = \mathbf{B}.$$

Some remarks about this problem:

- There are only finitely many permutation matrices of size $n \times n$, thus the problem can be solved in finite time. Also, given a matrix $\mathbf{P}$, it is possible to check in polynomial time if the equality holds, so Graph Isomorphism is in NP.
- The problem can be formulated as an Integer Program (in case you know what that is). In fact, the n^2 equations you obtain by examining each entry of $\mathbf{PA} = \mathbf{BP}$ are linear on the entries of $\mathbf{P}$, and upon enforcing that $\mathbf{P}_{ij} \in \{0, 1\}$ and that the rows and columns of $\mathbf{P}$ sum to 1, you enforce that $\mathbf{P}$ is a permutation matrix.
- If $\mathbf{P}$ exists, then $\mathbf{A}$ and $\mathbf{B}$ have the same spectrum. Thus, two isomorphic adjacency matrices must have the same eigenvalues. It is not trivial to find examples of matrices $\mathbf{A}$ and $\mathbf{B}$ which (1) have the same spectrum (2) admit a fractional solution to the IP above (3) are not isomorphic.

Assume now $\mathbf{A}$ and $\mathbf{B}$ have the same spectrum. What else do we need? Or better yet: how could this not be enough? Well, here is one case where the problem becomes easy: there is one eigenvector $\mathbf{v}$ for $\mathbf{A}$ and one eigenvector $\mathbf{u}$ for $\mathbf{B}$, both corresponding to the same eigenvalue. Thus

$$\mathbf{B}\mathbf{u} = \lambda\mathbf{u} \implies \mathbf{A}(\mathbf{P}^{\mathsf{T}}\mathbf{u}) = \lambda(\mathbf{P}^{\mathsf{T}}\mathbf{u}).$$

As λ has multiplicity 1, it follows that $\mathbf{P}^{\mathsf{T}}\mathbf{u}$ is a multiple of $\mathbf{v}$. Note that multiplying by $\mathbf{P}^{\mathsf{T}}$ only rearranges the entries of $\mathbf{u}$, and does not change the length of the vector. Hence we can assume that

$$\mathbf{P}^{\mathsf{T}}\mathbf{u} = \pm\mathbf{v}.$$

If, for instance, all entries of $\mathbf{v}$ have distinct absolute value, then the problem becomes trivial: there is only one possible rearrangement that works, and thus only one candidate for the matrix $\mathbf{P}$. This is all for one eigenspace.

Therefore graph isomorphism is hard for when all eigenvectors typically have many entries with the same absolute value, and it is not possible to combine them all and refine candidate partitions for each until only one matrix $\mathbf{P}$ survives, or none.

Let us formalize a bit these observations above. Our typical notation will be that a symmetric matrix $\mathbf{A}$ is diagonalized as $\mathbf{A} = \mathbf{PDP}^T$.

Lemma 2.1. *Let $\mathbf{A}$ and $\mathbf{B}$ be symmetric matrices with the same simple eigenvalues, with corresponding diagonalizations*

$$\mathbf{A} = \mathbf{UDU}^T \quad \textit{and} \quad \mathbf{B} = \mathbf{VDV}^T.$$

There is a permutation matrix $\mathbf{P}$ so that $\mathbf{PAP}^T = \mathbf{B}$ if and only if there is a diagonal matrix $\mathbf{E}$, whose entries are ± 1, so that $\mathbf{PU} = \mathbf{VE}$.

Before continuing, recall that $\mathbf{U}^T = \mathbf{U}^{-1}$, $\mathbf{V}^T = \mathbf{V}^{-1}$ and $\mathbf{P}^T = \mathbf{P}^{-1}$, because all these matrices are orthogonal matrices.

Proof. We have $\mathbf{PAP}^T = \mathbf{B}$ if and only if

$$\mathbf{P}(\mathbf{UDU}^T)\mathbf{P}^T = \mathbf{VDV}^T, \text{ or equivalently } \mathbf{V}^T\mathbf{PUD} = \mathbf{DV}^T\mathbf{PU}.$$

Let $\mathbf{E} = \mathbf{V}^T\mathbf{PU}$. Because all entries of $\mathbf{D}$ are distinct, it is enlightening to verify that $\mathbf{E}$ must be diagonal. Not only that, $\mathbf{E}^2 = 1$, so $\mathbf{E}$ contains only ± 1s. The other direction is immediate. □

This is already enough to tell us something quite strong. Recall that an automorphism of G is a permutation of $V(G)$ that preserves adjacency and non-adjacency.

Theorem 2.2. *If G is a graph and $\mathbf{A}(G)$ has simple eigenvalues, then any non-trivial automorphism of G has order 2.*

Proof. Let $\mathbf{P}$ be the permutation matrix representing the automorphism. Thus $\mathbf{PAP}^T = \mathbf{A}$, and by the corollary above, it follows that there is a ± 1 diagonal matrix $\mathbf{E}$ so that

$$\mathbf{PU} = \mathbf{UE}.$$

Hence $\mathbf{P}^2 = (\mathbf{UEU}^T)^2 = \mathbf{I}$. □

Combinatorially, this is saying that every automorphism of a graph with simple eigenvalues is splitting the vertices into some being fixed and some being swapped. Whenever you find a graph with a different type of automorphism, you already know now that at least one of its eigenvalues is not simple.

Exercise 2.3. Prove that if $\mathbf{P}$ and $\mathbf{Q}$ represent automorphisms of a graph with simple eigenvalues, then $\mathbf{PQ} = \mathbf{QP}$.

2.2 A canonical ordering of the vertices

In this subsection, we will show how to start with a given matrix of eigenvectors of $\mathbf{A}$, call it $\mathbf{P}$, and reorder the rows and sign the columns in a canonical, unique way. We will use the path on 4 vertices as a running example. It has simple eigenvalues $(\pm 1 \pm \sqrt{5})/2$.

So picture the matrix $\mathbf{P} = ((P_{ij}))$, columns corresponding to orthonormal eigenvectors, rows to vertices, and everything indexed by the set $\{1, ..., n\}$. If the eigenvalues of P_4 are in decreasing order, and if $a \approx 0.6$ and $b \approx 0.37$, we have that

$$\mathbf{P} = \begin{pmatrix} b & a & a & b \\ a & b & -b & -a \\ a & -b & -b & a \\ b & -a & a & -b \end{pmatrix}.$$

We say that π, permutation on this set, is an *orthodox star permutation*[1] if

(a) $P_{\pi(j)j} \neq 0$ for all $j \in \{1, ..., n\}$.

(b) Amongst all π satisfying the above, the sequence $(P^2_{\pi(1)1}\ ,\ P^2_{\pi(2)2}\ ,\ \cdots\ ,\ P^2_{\pi(n)n})$ is lexicographically maximal.

For P_4, there are precisely four orthodox star permutations, each corresponding to one way of picking one entry equal to $\pm a$ in each row and column:

$\pi_1 : (1, 2, 3, 4) \mapsto (2, 1, 4, 3).$

$\pi_2 : (1, 2, 3, 4) \mapsto (2, 4, 1, 3).$

$\pi_3 : (1, 2, 3, 4) \mapsto (3, 1, 4, 2).$

$\pi_4 : (1, 2, 3, 4) \mapsto (3, 4, 1, 2).$

Theorem 2.4. *For any graph, there is at least one orthodox star permutation.*

Proof. The matrix $\mathbf{P}$ is nonsingular, thus its determinant is non-zero. The determinant of $\mathbf{P}$ is[2]

$$\sum_{\pi \text{ permutation}} \operatorname{sign}(\pi) \prod_{i=1}^{n} P_{\pi(i)i}, \tag{2}$$

thus there is at least one π so that all terms within the product are nonzero. □

Given a star permutation π, its corresponding *orthodox star basis* is defined as vectors

$$\mathbf{x}_j = P_{\pi(j)j} \cdot \mathbf{P}_j.$$

(Here $\mathbf{P}_j$ is the jth column of $\mathbf{P}$.)

[1] I am presenting the content slightly backwards from its typical formulation, so a reason for this name will come later.

[2] If you have never seen this formula before, then research it.

Let $\mathbf{X}$ be the matrix that contains the vectors $\mathbf{x}_j$ as its columns.

We depict below all four orthodox star basis matrices corresponding to P_4.

$$B_1 = \begin{pmatrix} ab & a^2 & a^2 & ab \\ a^2 & ab & -ab & -a^2 \\ a^2 & -ab & -ab & a^2 \\ ab & -a^2 & a^2 & -ab \end{pmatrix} \quad , \quad B_2 = \begin{pmatrix} ab & -a^2 & a^2 & ab \\ a^2 & -ab & -ab & -a^2 \\ a^2 & ab & -ab & a^2 \\ ab & a^2 & a^2 & -ab \end{pmatrix} \quad ,$$

$$B_3 = \begin{pmatrix} ab & a^2 & a^2 & -ab \\ a^2 & ab & -ab & a^2 \\ a^2 & -ab & -ab & -a^2 \\ ab & -a^2 & a^2 & -ab \end{pmatrix} \quad \text{and} \quad B_4 = \begin{pmatrix} ab & -a^2 & a^2 & -ab \\ a^2 & -ab & -ab & a^2 \\ a^2 & ab & -ab & -a^2 \\ ab & a^2 & a^2 & ab \end{pmatrix} \quad .$$

Given a star permutation π, and its orthodox star basis matrix $\mathbf{X}$, its corresponding *quasi-canonical basis* is obtained from $\mathbf{X}$ upon reordering its rows in a lexicographical way — that is, the new first row will be that which is the largest row in the lexicographic order of vectors, and so on. Note that there can be no ties, as $\mathbf{X}$ has full rank and thus no pair of equal rows.

The two quasi-canonical bases of P_4 are depicted below.

$$C_1 = \begin{pmatrix} a^2 & ab & -ab & -a^2 \\ a^2 & -ab & -ab & a^2 \\ ab & a^2 & a^2 & ab \\ ab & -a^2 & a^2 & -ab \end{pmatrix} \quad \text{and} \quad C_2 = \begin{pmatrix} a^2 & ab & -ab & a^2 \\ a^2 & -ab & -ab & -a^2 \\ ab & a^2 & a^2 & -ab \\ ab & -a^2 & a^2 & -ab \end{pmatrix} \quad .$$

Finally, the *canonical star basis* associated to the graph is taken from all quasi-canonical bases picking the one which is the largest with respect to the lexicographical ordering of the columns — meaning, their quasi-canonical basis matrices are compared column by column, according to the lexicographical ordering of vectors, and the largest in the lexicographic ordering of these comparisons wins.

In the case of P_4, the matrix C_2 is the canonical star basis of eigenvectors.

Theorem 2.5. *Let* $\mathbf{A}$ *and* $\mathbf{B}$ *be two adjacency matrices, both with simple eigenvalues. Then there exists* $\mathbf{P}$ *so that* $\mathbf{P}\mathbf{A}\mathbf{P}^{\mathsf{T}} = \mathbf{B}$ *if and only if they have the same spectrum, and they admit the same canonical star basis.*

Proof. Let $\mathbf{U}$ and $\mathbf{V}$ be orthogonal matrices, and $\mathbf{D}$ diagonal, so that

$$\mathbf{A} = \mathbf{U}\mathbf{D}\mathbf{U}^T \quad \text{and} \quad \mathbf{B} = \mathbf{V}\mathbf{D}\mathbf{V}^T.$$

Let $\mathbf{U}'$ and $\mathbf{V}'$ be the canonical star basis matrices of $\mathbf{U}$ and $\mathbf{V}$ respectively. In the process of starting from $\mathbf{U}$ and arriving at $\mathbf{U}'$, two permutations are relevant: the orthodox star permutation ρ and the permutation of the rows that puts it in its final form: call its permutation matrix $\mathbf{Q}_1$. Let $\mathbf{E}_1$ be the diagonal matrix whose ith diagonal entry is equal to $\mathbf{U}_{\rho(i)i}$. Define matrices $\mathbf{E}_2$ and $\mathbf{Q}_2$ analogously for $\mathbf{V}$. Thus

$$\mathbf{U}' = \mathbf{Q}_1\mathbf{U}\mathbf{E}_1 \quad \text{and} \quad \mathbf{V}' = \mathbf{Q}_2\mathbf{V}\mathbf{E}_2.$$

If $\mathbf{U}' = \mathbf{V}'$, then $\mathbf{Q}_2^\top \mathbf{Q}_1 \mathbf{U} = \mathbf{V}\mathbf{E}_2\mathbf{E}_1^{-1}$, and because the columns of $\mathbf{U}$ and $\mathbf{V}$ are normalized, it follows that $\mathbf{E}_2\mathbf{E}_1^{-1}$ is a diagonal ± 1 matrix, thus the graphs are isomorphic according to Lemma 2.1.

If the graphs are isomorphic, then there is $\mathbf{E}$, diagonal matrix with ± 1 entries, so that $\mathbf{V} = \mathbf{PUE}$. Then $\mathbf{V}' = \mathbf{Q}_2\mathbf{PUEE}_2$. Note that the matrix $\mathbf{UEE}_2$ is also an orthodox star basis for $\mathbf{U}$, and because the rows of $\mathbf{Q}_2\mathbf{PUEE}_2$ are lexicographically ordered, it follows $\mathbf{V}'$ is a quasi-canonical basis for $\mathbf{U}$. Thus $\mathbf{U}'$ is larger or equal than $\mathbf{V}'$ in the lexicographic ordering of columns. Similarly, $\mathbf{U}'$ is a quasi-canonical basis for $\mathbf{V}$. It follows that $\mathbf{U}' = \mathbf{V}'$. □

Exercise 2.6. Assume $\pi_1, ..., \pi_m$ are all orthodox star permutations which eventually result in the canonic star basis — meaning, their orthodox star basis matrices have the same set of rows. Show that $\pi_i\pi_j^{-1}$ is an automorphism of the graph for all i and j. Show that every automorphism of the graph is of this form.

2.3 How to find the canonical star basis?

We just saw that deciding whether two graphs are isomorphic boils down to computing their canonical star basis. Let us now see that this procedure can be performed in polynomial time.

When first facing the matrix $\mathbf{P}$, construct a complete bipartite graph K, with one class corresponding to the rows of $\mathbf{P}$ (vertices of the original graph), one corresponding to the columns of $\mathbf{P}$ (eigenvalues). The weight of the edge connecting row i to column j is defined to be P_{ij}^2, that is, $w(r_i c_j) = P_{ij}^2$.

In several stages of this algorithm, we will need to find a perfect matching which is lexicographically maximal. This can be done by rescaling the weights of the edges by some parameters σ_j, one for each column j, so that every lexicographically maximal perfect matching of the original graph corresponds to a maximum weight perfect matching. Computing the maximum weight perfect matching of a bipartite graph can be done efficiently by means of the Hungarian method. For the next steps, we will assume one can always find the lexicographically maximal perfect matching of a given weighted bipartite graph (which, for short, we will refer to as a lex-max-pm).

(a) For simplicity and efficiency, remove from K all edges which do not belong to a lex-max-pm.

(b) Let c_1 be the vertex in K corresponding to the first column. We want to decide to which of its neighbours c_1 will be eventually paired.

(c) To each neighbour s of c_1, we construct the graph K_s, which is obtained from K upon removing all edges incident to c_1 but for the edge sc_1, and, moreover, reassigns weights to edges:
$$w(sc_1) = P_{s1}^2 \quad \text{e} \quad w(r_i c_j) = P_{sj}P_{ij}.$$

(d) For each one of these graphs, compute their lex-max-pm.

(e) Choose s so that the corresponding lex-max-pm is lexicographically maximal amongst all winners. Make this s permanently paired to c_1.

(f) Repeat for c_j, $j \geq 2$, iteratively.

It is clear that this procedure terminates and can be carried out in polynomial time in the size of the graph. It remains to show that it actually works — that is, that is produces a canonical star basis.

The lex-max-pm computed in step (d) corresponds precisely to the s-row of K in the orthodox star permutation containing sc_1. This row is, of course, permuted to be the top row in the quasi-canonical basis. Naturally, the canonical star basis must be so that its first row is largest amongst all quasi-canonical bases, and this is exactly being done in step (e).

After this row has been set, the remaining of the algorithm works recursively, as the the comparison to find the canonical star basis is done in a lexicographically fashion.

2.4 Star bases and larger multiplicities

Any attempt to generalize the algorithm above to larger multiplicities must deal with the fact that for higher dimensional eigenspaces, there are not just only two possible choices for a basis, but rather infinitely many. So there is no natural immediate generalization to Lemma 2.1, which was key to our treatment.

With this in mind, we now start discussing which are some possible canonical bases for the eigenspaces. Again, recall our notation for the spectral decomposition of $\mathbf{A}(G)$:

$$\mathbf{A} = \sum_{k=1}^{m} \theta_r \mathbf{E}_r. \tag{3}$$

A partition $V_1 \sqcup V_2 \sqcup ... \sqcup V_m$ of $V(G)$ is a *star partition* if, for all r, the vectors $\mathbf{E}_r \mathbf{e}_a$, for $a \in V_r$, form a basis for the θ_r-eigenspace. Note that a star partition of a graph with simple eigenvalues is precisely what we called a star permutation.

Exercise 2.7. Let $\mathcal{V}$ be the partition $V_1 \sqcup V_2 \sqcup ... \sqcup V_m$ of $V(G)$. Prove that the following are equivalent.

(a) $\mathcal{V}$ is a star partition.

(b) For all r, the vectors $\mathbf{E}_r \mathbf{e}_a$, for $a \in V_r$, are linearly independent.

(c) For all r, the vectors $\mathbf{E}_r \mathbf{e}_a$, for $a \in V_r$, span the θ_r-eigenspace.

Upon assuming $\theta_1 > \theta_2 > ...\theta_m$, we say that a (ordered) partition $(X_1, ..., X_m)$ of $V(G)$ is a *feasible partition* if $|X_r| = \operatorname{tr} \mathbf{E}_r$, which is also, of course, the dimension of the θ_r-eigenspace.

Recall that we had to show a star permutation existed. Here is no different.

Theorem 2.8. *For any graph G, there exists at least one star partition of* $\mathbf{A}(G)$.

Proof. Let $\mathbf{P}$ again be a matrix of orthonormal eigenvectors, ordered according to $\theta_1 > ... > \theta_m$. Equation (2) is quite powerful. One can group the permutations according to ordered partitions of any given type. That is, all permutations are so that there is a feasible partition $\mathcal{C} = (C_1, ..., C_m)$ so that the first $|C_1|$ numbers in $\{1, ..., n\}$ are mapped to C_1 exactly, the next $|C_2|$ numbers are mapped to C_2, and so on. Therefore the determinantal expansion can be rewritten as

$$\det \mathbf{P} = \sum_{\substack{\text{feasible} \\ \text{partitions } \mathcal{C}}} \pm \prod_{r=1}^{m} \det \mathbf{P}(r)_{\mathcal{C}}$$

where $\mathbf{P}(r)_{\mathcal{C}}$ stands for the square submatrix of $\mathbf{P}$ indexed by the rows corresponding to the elements of C_r, and the columns which lie between positions $1 + \sum_{j=1}^{r-1} |C_j|$ and $\sum_{j=1}^{r} |C_j|$. It is therefore a consequence of $\det \mathbf{P} \neq 0$ that there is a feasible partition $\mathcal{C}$ so that all determinants $\det \mathbf{P}(r)_{\mathcal{C}}$ are nonzero. Each block $\mathbf{P}(r)_{\mathcal{C}}$ appears with its rows and columns scaled in the matrix $\mathbf{E}_r$ (exactly on the rows and indices indexed by C_r). Therefore the partition $\mathcal{C}$ is a star partition. □

Given a star partition $\mathcal{V} = (V_1, ..., V_m)$ for a graph G with spectral decomposition as in (3), then $\{\mathbf{E}_r \mathbf{e}_a : a \in V_r\}$ is called a star basis for the θ_r-eigenspace, and their union for $r = 1, ..., m$ is a star basis for $\mathbb{R}^n$ relative to G.

2.5 How to order star bases

After finding a star permutation in Subsection 2.2, we proceeded to find a maximal amongst all such, according to some ordering. Following the analogy, we must determine how to find a maximal star partition.

Let $\mathbf{S}$ be an eigenvector matrix whose columns form a star basis. Assume of course the columns of $\mathbf{S}$ are ordered with respect to a star partition, say $\mathbf{S} = (\mathbf{S}_1 | \cdots | \mathbf{S}_m)$. Then matrices

$$\mathbf{W}_i = \mathbf{S}_i^{\mathsf{T}} \mathbf{S}_i$$

are defined. Here is the moment I will start skipping things. It is possible to order these matrices following a certain rule so that one can then search for a star partition whose corresponding sequence of matrices $(\mathbf{W}_1, ..., \mathbf{W}_m)$ is maximal in the lexicographical ordering corresponding to the rule. Such star basis will be known as an *orthodox star partition*. Further, it is possible to order the vertices within each class, also with respect to the entries of the matrices $\mathbf{W}_i$. Once this has been done, the corresponding star basis is called an *orthodox star basis*.

What is left is very similar to what we saw before: rows of orthodox star bases matrices are ordered in the lexicographical ordering, and then matrices obtained in this way are compared in the lexicographical ordering of columns, resulting in a canonical star basis, which is an isomorphism invariant.

This entire procedure relies on the capability of finding star bases, and ordering them. Both these things are to be performed simultaneously, much like it was the case for single eigenvalues. The main complexity issue of the algorithm is that the most efficient way of ordering the weighted matrices as described above relies on computing maximal cliques, and this problem is NP-complete. Apart from this, everything else can be carried out in polynomial time, including the task of finding star bases. Of course, if the size of the cliques are bounded (if the multiplicities of the eigenvalues are bounded) then the entire problem becomes solvable in polynomial time.

These algorithms are a bit too complicate for this stage in the course, so I will leave them for you to research on your own in case you are interested.

Instead, we will see some more interesting properties about star partitions.

2.6 Star partitions

We start with a simple yet very useful observation about eigenvalues and vertex-deleted subgraphs.

Lemma 2.9. *Let G be a graph, $a \in V(G)$, and θ eigenvalue of $\mathbf{A} = \mathbf{A}(G)$ of multiplicity m. Then θ is eigenvalue of $\mathbf{A}(G-a)$ of multiplicity at least $m-1$.*

Proof. Simply generate a basis for the θ-eigenspace in which all but at most one eigenvector has a entry non-zero. This is always possible.

Now observe that the restriction to $G-a$ of the eigenvectors of $\mathbf{A}$ with 0 entry at a are also eigenvectors of $\mathbf{A}(G-a)$, for the same eigenvalue. (Prove this fact explicitly, using the function representation of eigenvectors!). □

Here is a slightly more high level way of proving this last claim in the proof above. Let $\mathcal{U}$ be the θ-eigenspace, of dimension m, and $\mathcal{V}$ the subspace of all vectors in $\mathbb{R}^V$ so that the ath entry is 0. Note that $\dim \mathcal{V} = n-1$. Then

$$n \geq \dim(\mathcal{V}+\mathcal{U}) = \dim \mathcal{V} + \dim \mathcal{U} - \dim(\mathcal{V} \cap \mathcal{U}).$$

As a consequence of the lemma above, it follows that if $U \subseteq V(G)$ and θ is not eigenvalue of $G-U$, then $|U| \geq m$. Recall now the spectral decomposition

$$\mathbf{A}(G) = \sum_{r=1}^{m} \theta_r \mathbf{E}_r.$$

We say that a partition $V_1 \sqcup ... \sqcup V_m$ of $V(G)$ is a *polynomial partition* if, for each r, θ_r is not an eigenvalue of $G - V_r$.

Our goal now is to show that polynomial partitions correspond precisely to star partitions.

For what follows, let Θ_r denote the θ_r-eigenspace.

Lemma 2.10. *Let $(V_1, ..., V_m)$ be a partition, with $|V_r| = \dim \Theta_r$. This is a star partition if and only if, for each $r = 1, ..., m$, we have*

$$\Theta_r \ \cap \ \operatorname{span}\{\mathbf{e}_a : a \notin V_r\} = \{\boldsymbol{0}\}.$$

Proof. First assume the partition is a star partition, and fix r. Let $\mathbf{x}$ lie in the intersection $\Theta_r \cap \operatorname{span}\{\mathbf{e}_a : a \notin V_r\}$. Then, at the same time, $\mathbf{E}_r\mathbf{x} = \mathbf{x}$, and also $\mathbf{x}^{\mathsf{T}}\mathbf{e}_b = 0$ for all $b \in V_r$. But this implies that

$$0 = \mathbf{x}^{\mathsf{T}}\mathbf{e}_b = (\mathbf{E}_r\mathbf{x})^{\mathsf{T}}\mathbf{e}_b = \mathbf{x}^{\mathsf{T}}\mathbf{E}_r\mathbf{e}_b,$$

hence $\mathbf{x}$ is orthogonal to a basis for the Θ_r, a contradiction.

For the other direction, it follows from the hypothesis that there is a basis $\mathbb{R}^V$ obtained from joining a basis of Θ_r with each of the $\mathbf{e}_a$, $a \notin V_r$. This means that there is also another basis for the space which is formed by taking the vector $\mathbf{e}_a$ with $a \in V_r$ and vectors orthogonal to Θ_r. Hence the Θ_r eigenspace is generated by vectors obtained from applying $\mathbf{E}_r$ to this basis, that is, from the vectors $\mathbf{E}_r\mathbf{e}_a$, with $a \in V_r$. □

Theorem 2.11. *The partition $(V_1, ..., V_m)$ is a star partition if and only if, for each $r = 1, ..., m$, θ_r is not an eigenvalue of $G - V_r$.*

Proof. Start assuming $(V_1, ..., V_m)$ is a star partition. Our goal is to show that θ_r is not an eigenvalue of $G - V_r$. Let $\mathbf{A}' = \mathbf{A}(G - V_r)$, and let $\mathbf{u}'$ be a vector with $\mathbf{A}'\mathbf{u}' = \theta_r \mathbf{u}'$. Let $\mathbf{u} \in \mathbb{R}^V$ be obtained from $\mathbf{u}'$ by adding 0s in the positions corresponding to the vertices in V_r. Our goal now is to show that $\mathbf{u}$ (and thus $\mathbf{u}'$) is the 0-vector.

In order to achieve this, we can show that $\mathbf{u}$ is an eigenvector for θ_r in G, and at the same time that it is spanned by the $\mathbf{e}_a$ with $a \notin V_r$. Then the claim ($\mathbf{u} = \mathbf{0}$) follows from the previous lemma. This second assertion follows immediately by construction. To see the first, note that $\mathbf{u}$ being an eigenvector for θ_r in G is equivalent to showing that $(\mathbf{A} - \theta_r \mathbf{I})\mathbf{u} = 0$. First, note that for all $a \notin V_r$, it follows that

$$\mathbf{e}_a^\mathsf{T} \mathbf{A} \mathbf{u} = \theta_r \mathbf{e}_a^\mathsf{T} \mathbf{u}. \quad \text{(why?)}$$

Thus $\mathbf{e}_a^\mathsf{T}(\mathbf{A} - \theta_r \mathbf{I})\mathbf{u} = 0$ and $(\mathbf{A} - \theta_r \mathbf{I})\mathbf{u}$ is spanned by the $\mathbf{e}_b$ with $b \in V_r$.

On the other hand, if $\mathbf{A}\mathbf{x} = \theta_r \mathbf{x}$, then $\mathbf{x}^\mathsf{T}\mathbf{A}\mathbf{u} = \theta_r \mathbf{x}^\mathsf{T}\mathbf{u}$, and therefore $(\mathbf{A} - \theta_r \mathbf{I})\mathbf{u}$ is orthogonal to the θ_r-eigenspace. So $(\mathbf{A} - \theta_r \mathbf{I})\mathbf{u} = \mathbf{0}$ and we are done.

Now for the converse, assume that for each r, θ_r is not eigenvalue of $G - V_r$. From Exercise 2.7, it suffices to show that $\mathbf{E}_r \mathbf{e}_a$ span the θ_r-eigenspace, for $a \in V_r$. Assume $\mathbf{A}\mathbf{x} = \theta_r \mathbf{x}$, and also $\mathbf{x}^\mathsf{T} \mathbf{E}_r \mathbf{e}_a$, and therefore $\mathbf{E}_r \mathbf{x}$ is spanned by $\mathbf{e}_b$, with $b \notin V_r$. But $\mathbf{E}_r \mathbf{x} = \mathbf{x}$, and so $\mathbf{x}$ is a vector which is 0 at the entries corresponding to V_r and also satisfies $\mathbf{A}\mathbf{x} = \theta_r \mathbf{x}$. It is immediate to verify that the restriction of $\mathbf{x}$ to $G - V_r$, say $\mathbf{x}'$, satisfies $\mathbf{A}'\mathbf{x}' = \theta_r \mathbf{x}'$, hence, by hypothesis, $\mathbf{x}' = \mathbf{0}$, and thus $\mathbf{x} = \mathbf{0}$. □

An immediate corollary if that if S is a subset of a star cell, and θ is eigenvalue of multiplicity m, then θ is eigenvalue of multiplicity $m - |S|$ in $G - S$.

Exercise 2.12. Let $(V_1, ..., V_m)$ be a star partition. Prove that the block of $\mathbf{E}_r$ corresponding to rows and columns indexed by V_r is diagonal.

Another interesting consequence of Lemma 2.10 (well, of the first part of its proof) is that, as long as one finds $U \subseteq V(G)$ so that $\{\mathbf{E}_r \mathbf{e}_a : a \in U\}$ are linearly independent, then it is possible to find a star partition so that $U \subseteq V_r$.

Exercise 2.13.

(a) Show that if such U exists, then there is U' with $U \subseteq U' \subseteq V(G)$ so that $\{\mathbf{E}_r \mathbf{e}_a : a \in U'\}$ form a basis for Θ_r.

(b) Now let $\mathbf{P}$ be an eigenvector matrix for the graph. Show that the block of $\mathbf{P}$ corresponding to the rows of U' and columns eigenvectors of θ_r is nonsingular.

(c) Using now the first part of the proof of Lemma 2.10, show that if you remove these rows and columns from (b), the remaining matrix is nonsingular.

(d) Show that there is a star partition so that $U' = V_r$.

2.7 Combinatorial connections to star partitions

This section contains several applications displaying the connection between star partitions and combinatorial properties. Most of the results have simple proofs.

Theorem 2.14. *Let $(V_1, ..., V_m)$ be a star partition, and a and b be distinct vertices in the same cell, say V_r. If the neighbours of both vertices outside V_r are the same, then either*

- *$a \sim b$, and $\theta_r = -1$, and all their neighbours are the same; or*
- *$a \nsim b$, and $\theta_r = 0$, and all their neighbours are the same.*

Proof. Consider the eigenvector equation:

$$\theta_r \mathbf{E}_r \mathbf{e}_a = \mathbf{A}\mathbf{E}_r \mathbf{e}_a = \mathbf{E}_r \mathbf{A} \mathbf{e}_a = \sum_{c \sim a} \mathbf{E}_r \mathbf{e}_c.$$

Consider then $\theta_r(\mathbf{E}_r\mathbf{e}_a - \mathbf{E}_r\mathbf{e}_b)$. Because the sums of the $\mathbf{E}_r\mathbf{e}_c$ amongst the c which are not in V_r are the same for both a and b, the result will follow from using the linear independence of the $\mathbf{E}_r\mathbf{e}_c$ for the c which are in V_r. □

A similar analysis leads to a solution to the following exercise.

Exercise 2.15.

- Let $(V_1, ..., V_m)$ be a star partition of G, and assume G has no isolated vertex. Let $a \in V_r$. Show that a has at least one neighbour outside of V_r.
- Use this to characterize the star partitions of $K_{m,n}$.
- If a has radius at least 2, what can you say about vertices at distance 2 from a?

Another way of interpreting the theorem is that when you pick two neighbours a and b in the same V_r, then $N(a) \triangle N(b)$ cannot be a subset of V_r unless $\theta_r = -1$, and analogously in case they are not neighbours with respect to the eigenvalue 0. As a consequence, we have the following.

Corollary 2.16. *Fix an integer m, and assume $\theta \notin \{0, -1\}$. There are only finitely many graphs for which the θ-eigenspace has codimension m.*

Proof. If $|\overline{V_r}| = m$, then there are at most 2^m subsets in $\overline{V_r}$, and each of these determines at most one vertex in V_r. The result follows. □

A set of vertices $U \subseteq V(G)$ is called a *dominating set* if all vertices outside of U are neighbours to at least one vertex in U. The set U is called *location dominating* if any vertex outside of U is uniquely determined by its (non-empty) neighbourhood in U. To typical problem is to find the smallest possible dominating sets in a graph.

It is not difficult to notice our results from above can be phrased in terms of dominating sets:

Corollary 2.17. *Let $(V_1, ..., V_m)$ be a star partition. Then $\overline{V_r}$ is a dominating set for all r, and if $\theta_r \notin \{0, -1\}$, then $\overline{V_r}$ is also location-dominating.*

There are several interesting other combinatorial applications of star partitions, but we stop for now. Let us move to study the connection between polynomials and graphs.

2.8 References

Here is the set of references used to write the past few pages.

Graph Isomorphism for graphs with simple eigenvalues was shown to be solvable in polynomial time in a manuscript (literally) by Leighton and Miller, in 1979.

Very soon after, the published paper by Babai, Grigoryev and Mount proved a stronger result, namely, that Graph Isomorphism is also easy for graphs whose all eigenspaces have bounded multiplicity. Their paper uses quite the non-trivial group theory.

Most of this section is based on the work of Cvetkovic, Rowlinson and Simic (*A Study of Eigenspaces of Graphs*), later surveyed on their book *Eigenspaces of Graphs*.

A good source for the original algorithm proposed by Leighton and Miller is the book *Spectral and Algebraic Graph Theory* by Dan Spielman (currently published online on a 2019 version).

3 Graph polynomials

Significant part of the algebraic graph theory of graphs revolves around studying polynomials whose definition is based on the graph. Coefficients or evaluations of such polynomials typically count things associated to the graph, but algebraic properties of them and of their roots also tend to bring interesting considerations about the graph.

One motivation to define polynomials for graphs is the hope that a given polynomial would be efficiently computable and at the same time completely identify the graph up to isomorphism. No such polynomial is known in general (otherwise graph isomorphism would be an easier problem). Another motivation possibly come (historically as well) from the famous Reconstruction Conjecture. In the next section, we will see a third and very recent relevant application of graph polynomials.

We start our section with a brief introduction to the Reconstruction Conjecture.

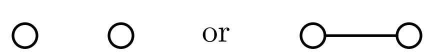
3.1 Reconstruction — an interlude

Given a graph G on n vertices, the set of n subgraphs obtained from G upon deleting each one of its vertices is called the *deck of G*. If G and its deck are presented with labelled vertices, then there is not much to ask or wonder. A totally more interesting question rises when one simply erases (or arbitrarily mixes up) the labels — we shall hence assume all graphs in this section are of such form.

Conjecture 1 (Kelly-Ulam)**.** *For any graph G on $n > 2$ vertices, G is completely determined by its deck.*

The hypothesis on $n > 2$ is necessary because the two subgraphs

could have been obtained from either of the following graphs,



but these are the only known case of such phenomenon. Several graph theorists have worked on this conjecture for the past decades, and yet a complete answer seems to be far from being found.

Partial results usually have two flavours: either one determines that graphs belonging to a certain class are reconstructible (from their deck), or one determines which properties or invariants of a graph are reconstructible. For the remainder of this section, we will mostly focus on the second type of question. But in this brief interlude, we prove the following results.

Let $\nu(H, G)$ denote the number of subgraphs of G isomorphic to H. It is not too surprising that this parameter is reconstructible.

Lemma 3.1 (Kelly)**.** *For any graphs G and H,*

$$(|V(G)| - |V(H)|)\ \nu(H, G) = \sum_{a \in V(G)} \nu(H, G \backslash a)$$

Proof. The result is trivial if $|V(H)| \geq |V(G)|$. Assume otherwise. We shall count the number of pairs (H', a) where H' is a copy of H in G, $a \in V(G)$ but $a \notin V(H')$. By choosing H' first, there are $(|V(G)| - |V(H)|)\ \nu(H, G)$ such pairs. By choosing a first, the number of copies of H not using a is precisely $\nu(H, G\backslash a)$. The result thus follows. □

Corollary 3.2. *If G has more than two vertices, the parameter $|E(G)|$ is reconstructible from the deck of G.*

Corollary 3.3. *The degree sequence of G (that is, the sequence of numbers listing the degrees of the vertices of G) is reconstructible.*

Exercise 3.4. Using Kelly's lemma, prove both corollaries above.

Theorem 3.5. *If G is a regular graph on more than 2 vertices, then G is reconstructible.*

Proof. From the degree sequence, decide whether G is regular. If it is, examine any of the graphs in its deck, and add a missing vertex so that it becomes regular. This graph will be equal to G. □

3.2 Walks

For any graph G, define $\phi_G(x)$ to be

$$\phi_G(x) = \det(x\mathbf{I} - \mathbf{A}).$$

The characteristic polynomial of a graph and of its subgraphs interplay nicely with walk counts and eigenvectors of the graph. Over the next few results, we shall make this relationship clearer.

Lemma 3.6. *If G is disconnected, and G_1 and G_2 are disjoint subgraphs of G with $G_1 \cup G_2 = G$, then*

$$\phi_G = \phi_{G_1} \cdot \phi_{G_2}.$$

This above is immediate from the block expansion of a determinant.

We now write a generating function whose coefficients are matrices:

$$W_G(x) = \sum_{k \geq 0} (\mathbf{A})^k x^k.$$

This is known as the walk generating function of G — the ij entry of the coefficient multiplying x^k counts the number of walks of length k from i to j. Rules for formal power series apply (existence of multiplicative inverses, substitutions, Laurent power series, etc.), and so we have

$$W_G(x) = \frac{1}{(\mathbf{I} - x\mathbf{A})}.$$

Notice that we are working with matrices whose coefficients are over $\mathbb{R}((x))$, but that shall mean no harm. In fact, properties about the determinant that you can prove exploring its Laplace expansion still hold true, in particular, for any $\mathbf{M}$ matrix with coefficients which are power series in x,

$$\mathbf{M} \cdot \operatorname{adj}(\mathbf{M}) = \det(\mathbf{M})\mathbf{I}. \tag{4}$$

Recall now that $\operatorname{adj}(\mathbf{M})$ is the matrix defined as

$$(\operatorname{adj} \mathbf{M})_{ij} = (-1)^{i+j} \det \mathbf{M}[j, i],$$

where $\mathbf{M}[j, i]$ stands for the matrix $\mathbf{M}$ removed of row j and column i.

Specifically, we are interested in what happens when $\mathbf{M} = (\mathbf{I} - x\mathbf{A})$. Equation (4) becomes

$$W_G(x) = \frac{\operatorname{adj}(\mathbf{I} - x\mathbf{A})}{\det(\mathbf{I} - x\mathbf{A})} = \frac{\operatorname{adj}(\mathbf{I} - x\mathbf{A})}{\det(\mathbf{I} - x\mathbf{A})}. \tag{5}$$

Corollary 3.7. *The generating function for the number of closed walks around a vertex a in the variable x is*

$$W_G(x)_{aa} = \frac{\phi_{G \backslash a}(x^{-1})}{x \cdot \phi_G(x^{-1})}.$$

Proof. Follows immediately from

$$W_G(x) = \frac{\operatorname{adj}(\mathbf{I} - x\mathbf{A})}{\det(\mathbf{I} - x\mathbf{A})} = \frac{x^{n-1} \operatorname{adj}(x^{-1}\mathbf{I} - \mathbf{A})}{x^n \det(x^{-1}\mathbf{I} - \mathbf{A})},$$

and the definition of the adjugate. □

We would also appreciate to have an expression for $W_G(x)_{ab}$. For that, we make use of an old trick due to Jacobi to arrive at an expression. For any matrix $\mathbf{M}$ with rows and columns indexed by a set V, let $\mathbf{M}_D$ stand for the submatrix with rows and columns indexed by $D \subseteq V$. The following theorem is the correct generalization of Corollary 3.7.

Theorem 3.8. *Let D be a subset of $V(G)$ (assume without loss of generality that the rows and columns indexed by D are the first). Then*

$$\det[W_G(x)]_D = \frac{1}{x^{|D|}} \frac{\phi_{G \backslash D}(x^{-1})}{\phi_G(x^{-1})}.$$

Proof. Let $\mathbf{C}$ be the matrix obtained from $\mathbf{I}$ upon replacing its first $|D|$ columns by the first $|D|$ columns of $\operatorname{adj}(\mathbf{I} - x\mathbf{A})$. Hence

$$(\mathbf{I} - x\mathbf{A}) \cdot \mathbf{C} = \begin{pmatrix} \det(\mathbf{I} - x\mathbf{A})\mathbf{I}_{|D|} & ? \\ \mathbf{0} & (\mathbf{I} - x\mathbf{A})_{\overline{D}} \end{pmatrix}.$$

Note that

$$\det \mathbf{C} = \det \operatorname{adj}(\mathbf{I} - x\mathbf{A})_D = \det[W_G(x)]_D \cdot (\det(\mathbf{I} - x\mathbf{A})^{|D|}).$$

Thus

$$\det[W_G(x)]_D = \frac{\det[(\mathbf{I} - x\mathbf{A})_{\overline{D}}]}{\det(\mathbf{I} - x\mathbf{A})} = \frac{x^{n-|D|} \det(x^{-1}\mathbf{I} - \mathbf{A})_{\overline{D}}}{x^n \det(x^{-1}\mathbf{I} - \mathbf{A})},$$

which yields the result. □

If $D = \{a, b\}$, then

$$W_G(x)_{aa}W_G(x)_{bb} - W_G(x)^2_{ab} = \frac{1}{x^2}\frac{\phi_{G\backslash ab}(x^{-1})}{\phi_G(x^{-1})},$$

therefore

$$W_G(x)_{ab} = \frac{1}{x}\frac{\sqrt{\phi_{G\backslash a}(x^{-1})\phi_{G\backslash b}(x^{-1}) - \phi_G(x^{-1})\phi_{G\backslash ab}(x^{-1})}}{\phi_G(x^{-1})}.$$

Notice in particular, from Equation (5), and replacing $y = x^{-1}$, that

$$\sqrt{\phi_{G\backslash a}(y)\phi_{G\backslash b}(y) - \phi_G(y)\phi_{G\backslash ab}(y)} = \operatorname{adj}(y\mathbf{I} - \mathbf{A})_{ab},$$

which is a polynomial (meaning: a power series with finite terms), and therefore the term inside the square root must be a perfect square (a fact that is not at all immediate at first sight).

Exercise 3.9. Let $\mathcal{P}_{ab}$ be the set of all paths from a to b. Prove that

$$\sqrt{\phi_{G\backslash a}(y)\phi_{G\backslash b}(y) - \phi_G(y)\phi_{G\backslash ab}(y)} = \sum_{P\in\mathcal{P}_{ab}} \phi_{G\backslash P}(y).$$

Hints:

(i) This will be a proof by induction.

(ii) Define $N_G(y)_{ab}$ to be the generating function for the walks that start at a, never return to it, and end at b. Find a relation between W_{ab}, N_{ab} and W_{aa}.

(iii) Find a relation between N_{ab} and W_{cb} (in $G\backslash a$), where c runs over the neighbours of a.

(iv) Apply induction.

3.3 Spectral decomposition

Say $\mathbf{A} = \sum_{r=0}^{d} \theta_r \mathbf{E}_r$. From the walk generating function $W_G(x) = \sum_{k\geq 0}(\mathbf{A})^k x^k = (\mathbf{I} - x\mathbf{A})^{-1}$, we have

$$W_G(x) = \sum_{r=0}^{d} \frac{1}{1 - x\theta_r} E_r. \tag{6}$$

Thus,

$$W_G(x^{-1}) = \sum_{r=0}^{d} \frac{x}{x - \theta_r} E_r.$$

If we focus on the diagonal entries, we have

$$\frac{x\phi_{G\backslash a}(x)}{\phi_G(x)} = W_G(x^{-1})_{aa} = \sum_{r=0}^{d} \frac{x}{x - \theta_r}(E_r)_{aa},$$

thus, multiplying both sides by $(x-\theta_r)$ and evaluating at $x=\theta_r$, yields

$$(E_r)_{aa} = \left.\frac{(x-\theta_r)\phi_{G\backslash a}(x)}{\phi_G(x)}\right|_{x=\theta_r}$$

For off-diagonal entries, we obtain

$$(E_r)_{ab} = \left.\frac{(x-\theta_r)\sqrt{\phi_{G\backslash a}(x)\phi_{G\backslash b}(x) - \phi_G(x)\phi_{G\backslash ab}(x)}}{\phi_G(x)}\right|_{x=\theta_r}.$$

Exercise 3.10. Show that if θ_r is an eigenvalue of $\mathbf{A}(G)$ with multiplicity m_r, then, for any $a \in V(G)$, its multiplicity in $\mathbf{A}(G\backslash a)$ is at least $m_r - 1$. Prove that equality holds if and only if there is at least one eigenvector for θ_r whose entry corresponding to a is non-zero.

Exercise 3.11. The goal of this exercise is to show that for any two matrices $\mathbf{M}$ and $\mathbf{N}$ so that $\mathbf{MN}$ and $\mathbf{NM}$ are defined, the following identity holds

$$\det(\mathbf{I} - \mathbf{MN}) = \det(\mathbf{I} - \mathbf{NM}).$$

To achieve this, find the two matrices that make both products below true, and finish the exercise.

$$\begin{pmatrix} \mathbf{I} & -\mathbf{M} \\ \mathbf{N} & \mathbf{I} \end{pmatrix} = \begin{pmatrix} \mathbf{I} & -\mathbf{M} \\ \mathbf{0} & \mathbf{I} \end{pmatrix} \begin{pmatrix} ? & ? \\ ? & ? \end{pmatrix} \quad \text{and} \quad \begin{pmatrix} \mathbf{I} & -\mathbf{M} \\ \mathbf{N} & \mathbf{I} \end{pmatrix} = \begin{pmatrix} \mathbf{I} & \mathbf{0} \\ \mathbf{N} & \mathbf{I} \end{pmatrix} \begin{pmatrix} ? & ? \\ ? & ? \end{pmatrix}$$

Exercise 3.12. Let $w(x)$ be the generating function whose coefficient of x^k count the total amount of all walks in the graph of length k. The goal of this exercise is to show that

$$w(x) = \frac{1}{x}\left(\frac{(-1)^n \phi_{\overline{G}}(-1-x^{-1})}{\phi_G(x^{-1})} - 1\right).$$

Recall that $\mathbf{A}(\overline{G}) = \mathbf{J} - \mathbf{I} - \mathbf{A}(G)$. You will use that $w(x) = \mathbf{1}^T W_G(x)\mathbf{1}$, that $\mathbf{J} = \mathbf{1}\mathbf{1}^T$, and finally the past exercise.

3.4 Reconstructing

In this section, we will show that the characteristic polynomial is reconstructible from the deck of the graph — that is, if the conjecture is false, then any counterexamples will have to be graphs with the same spectrum.

We would like to be able to reduce $\phi_G(x)$ somehow to an expression depending on the vertex-deleted subgraphs of G. Our best chance is then to look at Corollary 3.7, and take the trace in Equation (6). First, answer the exercise.

Exercise 3.13. Explain why $\operatorname{tr} E_r = m_r$, the multiplicity of θ_r as an eigenvalue.

Now we shall have

$$\frac{1}{\phi_G(x)}\left(\sum_{a\in V(G)} \phi_{G\backslash a}(x)\right) = \operatorname{tr}\left[x^{-1}W_G(x^{-1})\right] = \sum_{r=0}^{d} \frac{m_r}{x-\theta_r}.$$

Hence

$$\sum_{a \in V(G)} \phi_{G\backslash a}(x) = \sum_{r=0}^{d} \left(m_r(x-\theta_r)^{m_r-1}\right) \prod_{s \neq r} (x-\theta_s)^{m_s} = \phi_G(x)'.$$

This shows that we need only the characteristic polynomial of the graphs in the deck of G to recover the characteristic polynomial of G, except for its constant term. This actually will prove itself a considerably harder task, to which we devote the remaining of this subsection.

We start by actually finding a combinatorial expansion for the coefficients of $\phi(x)$, which in its own self is interesting and relevant. A *sesquivalent* subgraph H of G is a subgraph satisfying

(i) $|V(H)| = |V(G)|$.

(ii) Every connected component of H is either an isolated vertex, or an edge, or a cycle.

For each sesquivalent subgraph H of G, let $v(H)$, $e(H)$ and $c(H)$ denote the number of connected components which are, respectively, isolated vertices, edges and cycles.

Theorem 3.14 (Harary, Sachs)**.** *Let G be a simple graph, and $\mathcal{H}$ the set of all sesquivalent subgraphs of G. Then*

$$\phi_G(x) = \sum_{H \in \mathcal{H}} (-1)^{e(H)}(-2)^{c(H)} x^{v(H)}.$$

Proof. Leibniz formula for the determinant gives

$$\phi_G(x) = \det(x\mathbf{I} - \mathbf{A}) = \sum_{\sigma \in S_n} (-1)^{\epsilon(\sigma)} \prod_{i=1}^{n} (x\mathbf{I} - \mathbf{A})_{i\sigma(i)}.$$

(The sum runs over all permutations of $\{1, ..., n\}$, and $\epsilon(\sigma)$ is the number of cycles of even length in the decomposition of σ as a product of disjoint cycles.)

Consider the set of all permutations fixing precisely the points belonging to the subset $D \subseteq V(G)$. The sum of the terms corresponding to these permutations will therefore be

$$x^{|D|}(-1)^{n-|D|} \det(\mathbf{A}(G\backslash D)).$$

Each permutation of $V(G)\backslash D$ with fixed points contributes nothing to the determinant of $\mathbf{A}(G\backslash D)$. Those without will contain cycles of length two, or longer. Note that the support of the cycle structure of a permutation is a sesquivalent subgraph of $G\backslash D$. The cycles of length 2 are edges. The longer ones are the cycles of the graph. Each of the longer cycles of σ could have their orders reversed, yielding a permutation corresponding to the same sesquivalent subgraph H. Thus the total number of permutations corresponding to the sesquivalent subgraph H is $2^{c(H)}$.

Say the permutation σ corresponds to sesquivalent subgraph H. The quantity of cycles of odd length in σ has the same parity as $n - |D|$. If this is even, then total number of cycles, which is $e(H) + c(H)$, has the same parity as the number of even cycles, which is $\epsilon(\sigma)$. Otherwise, total number of cycles has opposite parity. Thus, if σ corresponds to the sesquivalent subgraph H with no isolated vertices, then

$$(-1)^{n-|D|}(-1)^{\epsilon(\sigma)} = (-1)^{e(H)+c(H)}.$$

Therefore the sum of the terms corresponding to the permutations fixing the set D will be

$$x^{|D|}(-1)^{n-|D|}\det(\mathbf{A}(G\backslash D)) = x^{|D|}\sum_H (-1)^{e(H)+c(H)}2^{c(H)}.$$

where the sum runs over the sesquivalent subgraphs of $G\backslash D$ with no isolated vertices. Varying the set D over all subsets of G will yield the desired expressions of the theorem. □

The constant term in $\phi_G(x)$, which is $(-1)^n\det(\mathbf{A}(G))$, is, according to the theorem above, equal to

$$\sum_H (-1)^{e(H)}(-2)^{c(H)}$$

where the sum runs over the sesquivalent subgraphs H of G with no isolated vertices.

Recall Kelly's lemma, which is useful to count copies of a subgraph H with $|V(H)| < |V(G)|$.

Lemma 3.15. *For any graphs G and H,*

$$(|V(G)| - |V(H)|)\ \nu(H,G) = \sum_{a\in V(G)} \nu(H, G\backslash a).$$

With a little more work, we have the following. Recall that a graph homomorphism from G_1 to G_2 is a function from $V(G_1)$ to $V(G_2)$ that preserves adjacency (but not necessarily non-adjacency).

Lemma 3.16. *G on n vertices, and H a disconnected graph on n vertices. Then $\nu(H,G)$ is reconstructible.*

Proof. Let H_1 and H_2 be disjoint subgraphs whose union is H. There are $\nu(H_1,G)\nu(H_2,G)$ homomorphisms from H to G which are injective on H_1 and H_2. Several of those however overlay images of vertices from H_1 and H_2. But we can count those. For each F on fewer than n vertices, there are $\nu(F,G)$ copies of F in G, and we can count the number of surjective homomorphisms from H to F which are injective in both H_1 and H_2. We multiply both things, and sum this for all F. We then subtract the total from $\nu(H_1,G)\nu(H_2,G)$ to recover $\nu(H,G)$. □

The result above allows us to compute the sum

$$\sum_H (-1)^{e(H)}(-2)^{c(H)}$$

for all disconnected H. The only thing remaining now to account for are the connected H.

A graph has vertex connectivity 1 if it is connected and contains a vertex whose removal disconnects the graph (a cut-vertex). A block is a maximal subgraph that does not contain a cut-vertex. For example, a tree contains $n-1$ blocks (each corresponding to an edge). The number of blocks in a 1-connected subgraph is the number of cut-vertices added by 1.

Lemma 3.17. *Let H be a 1-connected graph, on n vertices. The number of subgraphs of G with n vertices that contain the same collection of blocks of H is reconstructible.*

Proof. Assume H contains exactly two blocks H_1 and H_2 (thus $|V(H_1)| + |V(H_2)| = n+1$). Consider all homomorphisms from $H_1 \cup H_2$ to G which are injective in both H_1 and H_2. There are $\nu(H_1, G)\nu(H_2, G)$ such homomorphisms. The number of such mappings whose image is contained in a vertex deleted subgraph of G is reconstructible (see lemma above and Kelly's lemma). Thus the number of those whose image is G, obtained from overlaying only one vertex of H_1 with one of H_2, is reconstructible. These will correspond precisely to the spanning subgraphs of G which have H_1 and H_2 as their blocks. Now we can simply apply induction on the number of blocks of H to account for when H has any number of blocks. □

Using both lemmas above, one can show that:

Corollary 3.18. *If G is disconnected, then G is reconstructible. If G is a tree, then G is reconstructible.*

Exercise 3.19. Write the details proving the corollary above.

Corollary 3.20. *The number of Hamilton cycles of G can be reconstructed from the deck.*

Proof. The number of edges of G is reconstructible, so we can count the number of subgraphs of G with precisely n edges. We can also count how many of those are in vertex-deleted subgraphs, thus we can recover how many spanning subgraphs of G have precisely n edges. Out of these, we can count those which are disconnected and those which contain a cut-vertex, because they will contain a unique cycle of length $k < n$. The remaining graphs in the count will be Hamilton cycles. □

Clearly the implicit algorithm in the proof above is extremely inefficient, but there was no hope of providing an efficient algorithm that counts the number of Hamilton cycles in a graph anyway (deciding whether one exists is already itself a hard task).

Theorem 3.21. *The characteristic polynomial of G is reconstructible from the deck.*

Proof. We proved that

$$\phi_G(x)' = \sum_{a \in V(G)} \phi_{G \backslash a}(x).$$

The constant term of $\phi_G(x)$ is

$$\sum_H (-1)^{e(H)} (-2)^{c(H)}$$

where the sum runs over the sesquivalent subgraphs H of G with no isolated vertices. Those which are disconnected can be dealt with Lemma 3.16. Those which are connected correspond precisely to the Hamilton cycles of G, and this number can be reconstructed from Corollary 3.20. □

Recall the we proved that

$$\phi_{G \backslash a}(y)\phi_{G \backslash b}(y) - \phi_G(y)\phi_{G \backslash ab}(y)$$

is a perfect square of a polynomial, say $q_{ab}(y)$. If $\phi_G(y)$ is irreducible over the rationals, that it is easy to show that $\phi_{G\backslash ab}(y)$ is completely determined by $\phi_G(y)$, $\phi_{G\backslash a}(y)$, and $\phi_{G\backslash b}(y)$.

Having the eigenvalues of $G\backslash ab$, we can recover its number of edges. So we know the number of edges in G, $G\backslash a$, $G\backslash b$ and $G\backslash ab$. Hence we can find whether there is an edge between a and b in G. As a consequence:

Theorem 3.22 (Tutte). *If characteristic polynomial of G is irreducible over the rationals, then G itself is reconstructible.* □

We list two open questions related to our work in this chapter.

Problem 3.1. *Can you reconstruct of the characteristic polynomial of the Laplacian matrix from the deck?*

Problem 3.2. *Instead of the deck of G, assume you have access only to the characteristic polynomials of the graphs in the deck. Can you reconstruct $\phi_G(x)$? (It is known that this is possible if you have the characteristic polynomials of the graphs in the deck and their complement.)*

Define now the three-variable polynomial

$$\Phi_G(y, z, x) = \sum_{H \in \mathcal{H}} y^{e(H)} z^{c(H)} x^{v(H)}.$$

Note that $\phi_G(x) = \Phi_G(-1, -2, x)$.

Exercise 3.23. Prove that

$$\frac{\partial}{\partial x} \Phi_G(y, z, x) = \sum_{a \in V(G)} \Phi_{G\backslash a}(y, z, x).$$

Find expressions for

$$\frac{\partial}{\partial y} \Phi_G(y, z, x) \text{ and } \frac{\partial}{\partial z} \Phi_G(y, z, x).$$

Exercise 3.24. Verify that $\Phi_G(y, z, x)$ is reconstructible from the deck of G.

Exercise 3.25. Find a recurrence for Φ assuming G contains a cut-edge (meaning: write Φ_G in terms of Φ for some subgraphs of G.) Try the same exercise assuming G contains a cut-vertex.

Exercise 3.26. Let $\mathcal{C}_a$ be the set of cycles containing a vertex a. Explain why

$$\Phi_G = x\Phi_{G\backslash a} + y \sum_{b \sim a} \Phi_{G\backslash ab} + z \sum_{C \in \mathcal{C}_a} \Phi_{G\backslash C}.$$

Exercise 3.27. Assume all cycles of G have the same length, say c. Find a partial differential equation satisfied by Φ.

3.5 The matching polynomial of a graph

Let $\mathcal{M}(G)$ be the set of all spanning subgraphs of G whose connected components are either isolated vertices or isolated edges. The matching polynomial of a graph is defined as

$$\mu_G(x) = \sum_{M \in \mathcal{M}(G)} (-1)^{e(M)} x^{v(M)}.$$

Note that it is precisely equal to the evaluation $\Phi(-1, 0, x)$ of the polynomial $\Phi(y, z, x)$ defined in the past subsection. In fact,

Theorem 3.28. *Given a graph connected G,*

$$\mu_G(x) = \phi_G(x)$$

if and only if G is a tree.

Proof. One direction is obvious from the formula of Φ. The other I leave as a challenging exercise. □

Exercise 3.29. Verify that

$$\mu_G(x)' = \sum_{a \in V(G)} \mu_{G \backslash a}(x),$$

and, prove that, if $e = \{u, v\}$ is an edge of G, then

$$\mu_G(x) = \mu_{G \backslash e}(x) - \mu_{G \backslash uv}(x).$$

Exercise 3.30. Find recurrences for $\mu_{P_n}(x)$, $\mu_{K_n}(x)$ and $\mu_{C_n}(x)$ based on the matching polynomials of smaller graphs in each of the families. (Hint: use Exercise 3.26).

The recurrences you found in the past exercise show that matching polynomials in each of those families of graphs form what is a known as a sequence of orthogonal polynomials. We will not get into details of the theory of orthogonal polynomials, but over the next few results we will see some glimpse of it. Given polynomials $p(x)$ and $q(x)$, we define an inner product by

$$\langle p, q \rangle = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{+\infty} \mathrm{e}^{-x^2/2} p(x) q(x) \mathrm{d}x.$$

Do not get scared. Just bear with me. But maybe now it would be a good time to remember that

$$1 = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{+\infty} \mathrm{e}^{-x^2/2} \mathrm{d}x \quad \text{and} \quad 0 = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{+\infty} x \mathrm{e}^{-x^2/2} \mathrm{d}x.$$

Exercise 3.31. Prove these equalities. Hint: one of them is easy. For the other, write its square, and change variables to polar coordinates.

Lemma 3.32. *Let*

$$M(n) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{+\infty} \mathrm{e}^{-x^2/2} x^n \, dx.$$

The number of perfect matchings in K_n is equal to $M(n)$.

Proof. Integration by parts implies

$$M(n) = \frac{1}{\sqrt{2\pi}} \left[\frac{x^{n+1}}{n+1} \mathrm{e}^{-x^2/2} \right]_{-\infty}^{+\infty} + \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{+\infty} \frac{x^{n+2}}{n+1} \mathrm{e}^{-x^2/2} \mathrm{d}x.$$

The first term is 0. So it follows that $M(n) = M(n+2)/(n+1)$. As seen above, $M(1) = 0$ and $M(0) = 1$. Hence $M(\text{odd}) = 0$ and

$$M(2m) = (2m-1)!!$$

as we wanted to show. □

Recall that $(-1)^{n/2}\mu_G(0)$ is the number of perfect matchings in G. Denote this number by $\mathrm{pm}(G)$.

Theorem 3.33. *For any G, we have*

$$\mathrm{pm}(\overline{G}) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{+\infty} \mathrm{e}^{-x^2/2} \mu_G(x)\, dx$$

sketch. The proof is by induction on the number of edges in G. If G has no edges, this falls precisely in the statement of the lemma. If G has one edge, then both sides satisfy the same recursion given by the second part of Exercise 3.29. □

Exercise 3.34. Prove that

$$\frac{1}{\sqrt{2\pi}} \int_{-\infty}^{+\infty} \mathrm{e}^{-x^2/2} \mu_{K_n}(x) \mu_{K_m}(x) \mathrm{d}x = \begin{cases} m!, & \text{if } m = n; \\ 0, & \text{otherwise.} \end{cases}$$

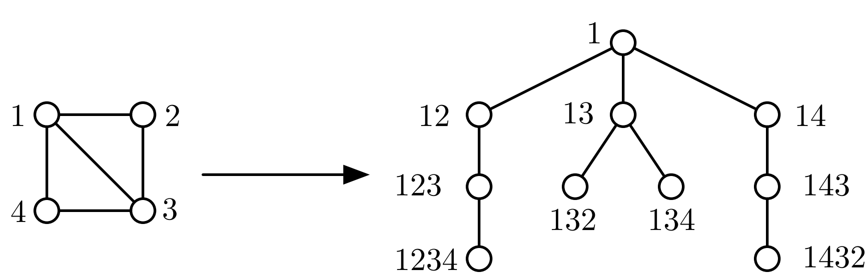
Hint: look at $K_n \cup K_m$, its complement, and the past exercise.

The conclusion from the result above is that the family $\{\mu_{K_n}(x)\}_{n \geq 0}$ is a family of orthogonal polynomials according to the inner product defined in this subsection.

3.6 Real roots

Our goal here is to show that the matching polynomial of any graph has only real roots.

Given a graph G and a vertex u in G, the *path tree* of G with respect to u is a rooted tree whose vertices correspond to the paths of G that start at u, and a the children of a vertex corresponding to path P are those vertices corresponding to paths obtained from one further edge at the end of P. For example



Theorem 3.35. *Let G be a graph, $u \in V(G)$. Let $T = T(G,u)$ be the path tree of G with respect to u. Then*

$$\mu_T(x)\mu_{G\backslash u}(x) = \mu_G(x)\mu_{T\backslash u}(x),$$

and $\mu_G(x)$ divides $\mu_T(x)$.

Proof. If G itself is already a tree, then there is nothing to prove, as $G = T$. We may assume the results holds true for vertex-deleted subgraphs of G. Thus

$$\mu_G(x) = x\mu_{G\backslash u} - \sum_{v\sim u} \mu_{G\backslash uv}(x).$$

Thus, applying induction, we have

$$\frac{\mu_G(x)}{\mu_{G\backslash u}(x)} = x - \sum_{v\sim u} \frac{\mu_{T(G\backslash u,v)\backslash v}(x)}{\mu_{T(G\backslash u,v)}(x)}.$$

Now, $T(G\backslash u, v)$ is isomorphic to the branch of $T(G,u)$ attached to u that starts at the vertex corresponding to the path uv. Thus

$$\frac{\mu_{T(G\backslash u,v)\backslash v}(x)}{\mu_{T(G\backslash u,v)}(x)} = \frac{\mu_{T(G,u)\backslash \{u,uv\}}(x)}{\mu_{T(G,u)\backslash u}(x)}.$$

Therefore

$$\frac{\mu_G(x)}{\mu_{G\backslash u}(x)} = \frac{x\mu_{T(G,u)\backslash u}(x) - \sum_{v\sim u} \mu_{T(G,u)\backslash\{u,uv\}}(x)}{\mu_{T(G,u)\backslash u}(x)} = \frac{\mu_{T(G,u)}(x)}{\mu_{T(G,u)\backslash u}(x)},$$

as wanted. For the second assertion, by induction, it follows that $\mu_{G\backslash u}(x)$ divides $\mu_{T(G\backslash u,v)}(x)$. As $T(G\backslash u, v)$ is a branch of $T(G,u)\backslash u$, it follows that $\mu_{T(G\backslash u,v)}(x)$ divides $\mu_{T(G,u)\backslash u}(x)$, so $\mu_{G\backslash u}(x)$ itself divides $\mu_{T(G,u)\backslash u}(x)$. Hence $\mu_G(x)$ divides $\mu_T(x)$. □

Corollary 3.36. *The roots of $\mu_G(x)$ are real, for any G. Moreover, they are symmetrically distributed around the origin.*

Proof. The polynomial $\mu_G(x)$ divides $\mu_T(x)$, which is equal to $\phi_T(x)$. This is the characteristic polynomial of a symmetric matrix, hence its roots are real. Therefore the roots of $\mu_G(x)$ are real.

The second part follows immediately from the fact that all exponents of x in $\mu_G(x)$ are either all odd or all even. □

Exercise 3.37. Prove that the zeros of $\mu_{G\backslash u}$ interlace those of μ_G. If G is connected, prove that the largest zero of μ_G is simple, and strictly larger than that of $\mu_{G\backslash u}$. Hint: use Theorem 3.35.

We can also bound the largest eigenvalue of μ relatively well.

Exercise 3.38. Show (again) that the largest eigenvalue of a non-negative matrix is upper bounded by its largest row sum.

Exercise 3.39. Extend the result above to argue that the largest eigenvalue of a non-negative matrix $\mathbf{M}$ is upper bounded by the largest row sum of $\mathbf{DMD}^{-1}$ for any positive diagonal matrix $\mathbf{D}$.

Exercise 3.40. Let T_Δ be a tree so that all vertices have degree $\Delta > 2$ or 1. Prove that its largest eigenvalue is upper bounded by $2\sqrt{\Delta - 1}$. Hint: Fix a vertex of degree Δ to call the root, and conjugate $\mathbf{A}(T_\Delta)$ by the diagonal matrix defined as $\mathbf{D}_{aa} = \sqrt{\Delta - 1}^{d(a)}$, where $d(a)$ is the distance from a to the root. Use the exercises above.

Exercise 3.41. Argue that any tree of maximum degree $\Delta > 1$ has its largest eigenvalue small or equal than $2\sqrt{\Delta - 1}$.

Exercise 3.42. Let G be a graph with $\Delta(G) > 1$. Show that the largest root λ of $\mu_G(x)$ satisfies

$$\sqrt{\Delta(G)} \leq \lambda \leq 2\sqrt{\Delta(G) - 1}.$$

(The upper bound should follow easily from the exercises above. The lower bound is your job to find.)

3.7 Number of matchings

The fact that the roots of $\mu_G(x)$ are real brings a combinatorial consequence. A sequence of numbers $(a_i)_{i\geq 0}$ is log-concave if $a_i^2 \geq a_{i-1}a_{i+1}$ for all $i \geq 1$. If the numbers are positive, then this is equivalent to having $(a_{i+1}/a_i)_{i\geq 0}$ non-increasing. Thus, a log-concave sequence of positive numbers is unimodal, meaning, it first increases, then stays constant, then decreases.

The binomial coefficients $\binom{n}{k}$, $k = 0, ..., n$, form a (finite) log-concave sequence. Clearly, if (a_i) and (b_i) are log-concave, so is $(a_i b_i)$.

Lemma 3.43. *If $p(x) = \sum_i a_i x^i$ is a polynomial of degree n with real roots only, then $(a_i/\binom{n}{i})$ form a log-concave sequence.*

Proof. This follows from writing

$$\frac{\mathrm{d}^{n-i-2}}{\mathrm{d}^{n-i-2}x} x^{n-i} \frac{\mathrm{d}^i}{\mathrm{d}^i x} p(x^{-1}) = \frac{1}{2} n! \left(\frac{a_i}{\binom{n}{i}} x^2 + \frac{a_{i+1}}{\binom{n}{i+1}} 2x + \frac{a_{i+2}}{\binom{n}{i+2}} \right).$$

(Fill in the details). □

Let m_k be the number of matchings in G with k edges. Note that

$$\mu_G(x) = \sum_{k\geq 0} (-1)^k m_k x^{n-2k}.$$

Corollary 3.44. *The sequence $(m_k)_{k\geq 0}$ is log-concave (and therefore unimodal).*

Proof. Assume n is even. Then $\mu_G(x) = q(x^2)$. Note that

$$p(x) = \sum_{k\geq 0} m_k x^k = x^{n/2} q(-x^{-1}),$$

which also has real roots. Similar argument for n odd. It follows then from Lemma that $(m_k)_{k\geq 0}$ is a log-concave sequence. □

3.8 Average

In this final section about the matching polynomial, we prove a remarkable result connecting $\mu(x)$ and $\phi(x)$.

Theorem 3.45. *Let G be a graph with m edges. Then*

$$\mu_G(x) = \frac{1}{2^m} \sum_F \phi_F(x),$$

where the sum runs over all 2^m signed graphs F whose underlying edges are exactly those of G.

To be a clear, $\mathbf{A}(F)$ is precisely $\mathbf{A}(G)$, except that certain symmetric off-diagonal entries have been changed to -1.

Proof. We have

$$\frac{1}{2^m} \sum_F \phi_F(x) = \frac{1}{2^m} \sum_F \sum_{\sigma \in S_n} (-1)^{\epsilon(\sigma)} \prod_{i=1}^n (x\mathbf{I} - \mathbf{A}(F))_{i\sigma(i)},$$

then

$$\frac{1}{2^m} \sum_F \phi_F(x) = \frac{1}{2^m} \sum_{\sigma \in S_n} (-1)^{\epsilon(\sigma)} \sum_F \prod_{i=1}^n (x\mathbf{I} - \mathbf{A}(F))_{i\sigma(i)},$$

Note that if σ contains a cycle with more than two vertices, then

$$\sum_F \prod_{i=1}^n (x\mathbf{I} - \mathbf{A}(F))_{i\sigma(i)} = 0,$$

as we can sum over all possible signings of this cycle having the rest constant, and later vary the rest, but the sum over all possible signings of a cycle of length larger than 2 is 0.

Thus the only permutations that contribute are those with transpositions and fixed points only, and for those the signing is irrelevant. The sum over all such permutations coincides with the matching polynomial of the graph. Therefore

$$\frac{1}{2^m} \sum_F \phi_F(x) = \frac{1}{2^m} \sum_F \sum_{M \in \mathcal{M}(G)} (-1)^{e(M)} x^{v(M)} = \frac{1}{2^m} 2^m \mu_G(x),$$

as we wished. □

Exercise 3.46. Let G be a graph, and F be obtained from G upon signing some of the edges. What exactly can be said about

$$\phi_G(x) + \phi_F(x) \quad ?$$

Exercise 3.47. Assume G is a graph with the property that every cycle of G contains at least one edge that belongs to no other cycle. Show how to compute μ_G efficiently.

3.9 References

Here is the set of references used to write the past few pages.

The main reference for this section is the book of Godsil, wherein references for all the results about the characteristic and matching polynomials can be found (Chapters 1, 2, 4 and 6).

(a) Chris D Godsil. *Algebraic Combinatorics*. Chapman & Hall, New York, 1993

(b) Chris Godsil and Gordon Royle. *Algebraic Graph Theory*. Springer-Verlag, New York, 2001

Elias Hagos proved that the characteristic polynomial is reconstructible from the characteristic polynomials of the graphs in the deck and their complements (if they are correctly paired up).

(c) Elias M Hagos. The characteristic polynomial of a graph is reconstructible from the characteristic polynomials of its vertex-deleted subgraphs and their complements. *The Electronic Journal of Combinatorics*, 7(1):12, 2000

4 Covers and Interlacing families

Two recent and relevant developments on spectral graph theory are going to be discussed in this section. Common to both is the use of signs in the adjacency matrix of a graph.

The first is a warm up, and also a review of our introductory section. It is a creative yet easy application, that solved a decades old problem originating in computer science.

The second represents an important development, that lead to the solution of several open problems.

4.1 Sensitivity

Let $Q_n = \{0,1\}^n$, and assume f is a boolean function, that is, $f : Q_n \to \{0,1\}$. For several applications in computer science, it is important to know or measure how f behaves if a coordinate of a point is changed. We say that $x, y \in Q_n$ are neighbours if they differ in one coordinate. In fact, we shall identify Q_n with a graph, henceforth called the n-dimensional hypercube.

The sensitivity of f at $x \in Q_n$, denoted by $s(f,x)$, is the number of neighbours y of x for which $f(x) \neq f(y)$. The sensitivity of the function f is the maximum sensitivity of f at all points of Q_n. In a sense, the sensitivity of a function is related to its complexity — for instance, a function which is determined by the first coordinate of an entry has sensitivity equal to 1.

There are several ways of measuring the complexity of a function, and an active field of research consists in deciding how two different ways compare. We introduce here below one other.

Every boolean function f can be uniquely represented a real multilinear polynomial. To see that, first perform the substitution $0 \to 1$ and $1 \to -1$. Let $S_n = \{1,-1\}^n$. Thus we see boolean functions as going from S_n to S_1.

Let χ_I be the function

$$\chi_I(x) = \prod_{i \in I} x_i,$$

Then, define the inner product that takes two functions and maps to a real number, given by

$$\langle f, g \rangle = 2^{-n} \sum_{x \in Q_n} f(x) g(x).$$

It is immediate to verify that $\langle \chi_I, \chi_I \rangle = 1$, and if $I \neq J$, then

$$\langle \chi_I, \chi_J \rangle = 2^{-n} \sum_{x \in Q_n} \chi_{I \Delta J}(x) = 0.$$

Because the space of all functions has dimension 2^n and we have just created 2^n orthogonal functions, they form a basis, and therefore there are coefficients α_I for each $I \subseteq [n]$ so that

$$f(x) = \sum_{I \subseteq [n]} \alpha_I \chi_I(x).$$

Moreover, each α_I is given by

$$\alpha_I = \langle f, \chi_I \rangle.$$

These are also called the Fourier transform of f at I. Let $d(f)$ denote the degree of the polynomial given above, that is,

$$d(f) = \max_{I \subseteq [n]} \{|I| : \alpha_I \neq 0\}.$$

In a sense this is also a measurement of complexity of the function f, and one wonders if there is any connection between $d(f)$ and $s(f)$. By standard methods, Nisam and Szegedy showed that $d(f) \geq \sqrt{s(f)/2}$. Our goal is to show the reverse inequality, which is an open problem that remained open for almost 30 years.

Gotsman and Linial showed the connection below.

Theorem 4.1. *For any $h : \mathbb{N} \to \mathbb{R}$, the following are equivalent.*

(a) Let $H \subseteq V(Q_n) = V$. Assume $|H| \neq 2^{n-1}$. Then we have that either $\Delta(Q_n[H]) \geq h(n)$ or $\Delta(Q_n[V - H]) \geq h(n)$.

(b) For any Boolean function $f : Q_n \to \{0, 1\}$, we have $s(f) \geq h(d(f))$.

Thus, in order to show that $s(f) \geq \sqrt{d(f)}$, it is enough to show that for any induced subgraph of the hypercube Q_n with more than 2^{n-1} vertices, the maximum degree is at least $\sqrt{n}$.

This problem is remarkably tricky. Note that there is an induced subgraph on 2^{n-1} vertices with maximum degree 0, because the hypercube is bipartite (the parity of the strings?). By adding one vertex to this graph you suddenly get something of degree n, which is as bad as it gets.

A better attempt is to split $[n]$ into (roughly) $\sqrt{n}$ sets of size $\sqrt{n}$. Say n is perfect square. Call these F_i, with $i = 1, ..., \sqrt{n}$. Consider the subset of $2^[n]$ formed by the sets of even cardinality that do not contain any F_i and those of odd cardinality that contain some F_i. It is possible to check that this family has size $2^{n-1} + 1$, and that the subgraph it induces has maximum degree $\sqrt{n}$. In the next section we show that this is best possible.

4.2 Cauchy Interlacing — a first contact

In the beginning of this course, we saw in Lemma 1.33 that the largest eigenvalue of a matrix is the maximum over a Rayleigh quotient, that is

$$\lambda_1(\mathbf{M}) = \max_{\mathbf{u} \neq 0} R_{\mathbf{M}}(\mathbf{u}).$$

Now assume $\mathbf{u}$ has been found, and we search for the maximum over all vectors orthogonal to $\mathbf{u}$. Then

$$\begin{aligned} R_{\mathbf{M}}(\mathbf{v}) &= \theta_0(\mathbf{v}^{\mathcal{T}}\mathbf{E}_0\mathbf{v}) + \theta_1(\mathbf{v}^{\mathcal{T}}\mathbf{E}_1\mathbf{v}) + \cdots + \theta_d(\mathbf{v}^{\mathcal{T}}\mathbf{E}_d\mathbf{v}) \\ &\leq \theta_1\big((\mathbf{v}^{\mathcal{T}}\mathbf{E}_0\mathbf{v}) + (\mathbf{v}^{\mathcal{T}}\mathbf{E}_1\mathbf{v}) + \cdots + (\mathbf{v}^{\mathcal{T}}\mathbf{E}_d\mathbf{v})\big) \\ &\leq \theta_1. \end{aligned}$$

Equality will be met if and only if $\mathbf{v}$ is eigenvector for θ_1. This argument generalizes for other eigenvalues and eigenvectors, and also for min instead of max if one looks for the smallest eigenvalues.

Theorem 4.2 (Cauchy's Interlacing — light version)**.** *Let* $\mathbf{A}$ *be a symmetric* $n \times n$ *matrix, and let* $\mathbf{B}$ *be a* $m \times m$ *principal square submatrix of* $\mathbf{A}$*. Let* $\theta_1 \geq ... \geq \theta_n$ *be the eigenvalues of* $\mathbf{A}$ *and* $\lambda_1 \geq ... \geq \lambda_m$ *be those of* $\mathbf{B}$*. Then, for all* k *with* $1 \leq k \leq m$*,*

$$\theta_{n-(m-k)} \leq \lambda_k \leq \theta_k.$$

Proof. Note there is $\mathbf{S}$ an $n \times m$ matrix satisfying $\mathbf{S}^T\mathbf{S} = \mathbf{I}$ and so that $\mathbf{B} = \mathbf{S}^T\mathbf{A}\mathbf{S}$. Let $\mathbf{u}_1, ..., \mathbf{u}_n$ be the eigenvectors of $\mathbf{A}$ corresponding to the θ_ks. The key thing now is to observe that, for all k, the subspace

$$\langle \mathbf{v}_1, ..., \mathbf{v}_k \rangle \cap \langle \mathbf{S}^T\mathbf{u}_1, ..., \mathbf{S}^T\mathbf{u}_{k-1} \rangle^\perp$$

contains at least one vector. Let $\mathbf{w}$ be such vector, which, in particular, implies $\mathbf{S}\mathbf{w} \in \langle \mathbf{u}_1, ..., \mathbf{u}_{k-1} \rangle^\perp$. Then, by the observation above, we have

$$\theta_k \geq \frac{(\mathbf{S}\mathbf{w})^T\mathbf{A}(\mathbf{S}\mathbf{w})}{(\mathbf{S}\mathbf{w})^T(\mathbf{S}\mathbf{w})} \geq \frac{\mathbf{w}^T\mathbf{B}\mathbf{w}}{\mathbf{w}^T\mathbf{w}} \geq \lambda_k.$$

The other inequalities follow from a symmetric argument. □

4.3 The Sensitivity Theorem

In 2019, Hao Huang shocked the graph theoretic community by providing a one-page proof of the Sensitivity Conjecture. Upon a clever yet elementary application of the interlacing principle, he showed that if H is a subgraph of the hypercube Q_n with more than 2^{n-1} vertices, then $\Delta(H) \geq \sqrt{n}$. To achieve this, he came up with a clever *signing* of the adjacency matrix of the hypercube.

Lemma 4.3. *Define the sequence* $\mathbf{A}_i$ *of matrices as follows:*

$$\mathbf{A}_1 = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} \quad \textit{and} \quad \mathbf{A}_n = \begin{pmatrix} \mathbf{A}_{n-1} & \mathbf{I} \\ \mathbf{I} & -\mathbf{A}_{n-1} \end{pmatrix}.$$

Then $\mathbf{A}_n$ *is a* $2^n \times 2^n$ *matrix whose eigenvalues are* $\pm\sqrt{n}$*, each with multiplicity* 2^{n-1}*.*

Proof. It follows immediately by induction that $\mathbf{A}_n^2 = n\mathbf{I}$. The result then follows by noticing that $\operatorname{tr} \mathbf{A}_n = 0$. □

The matrix $\mathbf{A}_n$ above is almost the adjacency matrix of the hypercube Q_n. In fact, changing all signs to $+1$, we get exactly $\mathbf{A}(Q_n)$.

Lemma 4.4. *Suppose* H *is a graph on* m *vertices, and* $\mathbf{A}$ *is an* $m \times m$ *symmetric matrix with entries* 0, 1 *and* -1*, whose nonzero entries correspond precisely to the edges of* H*. Then*

$$\Delta(H) \geq \lambda_1(\mathbf{A}).$$

Proof. Say $\mathbf{A}\mathbf{v} = \lambda_1 \mathbf{v}$, and assume $\mathbf{v}_a$ is the largest entry in absolute value of $\mathbf{v}$. Then

$$|\lambda_1 \mathbf{v}_a| = \left|(\mathbf{A}\mathbf{v})_a\right| \leq \sum_{b \sim a} |\mathbf{A}_{a,b}||\mathbf{v}_a| \leq \Delta(H)|\mathbf{v}_a|.$$

□

Theorem 4.5. *If H is a subgraph of the hypercube Q_n with more than 2^{n-1} vertices, then $\Delta(H) \geq \sqrt{n}$.*

Proof. Consider the "adjacency matrix" $\mathbf{A}_n$ of the hypercube, given in Lemma 4.3. Take a submatrix of $\mathbf{A}_n$ corresponding to the vertices in H, call it $\mathbf{B}$. Note that this is a ± 1 adjacency matrix of H, and therefore by the lemma above

$$\Delta(H) \geq \lambda_1(\mathbf{B}).$$

From interlacing, it follows that

$$\lambda_1(\mathbf{B}) \geq \lambda_{2^n - (2^{n-1}+1-1)}(\mathbf{A}_n) = \sqrt{n},$$

and the result follows. □

4.4 Covers

Start with a graph G, n vertices, with (standard) adjacency matrix $\mathbf{A}$, and assume you have also been given a signed adjacency matrix of G, say $\mathbf{B}$, as we did above for the hypercube.

Then replace every vertex of G by a pair of vertices. For each pair, decide which is the blue and which is the red, and form a new graph putting a matching between the two pairs of vertices that correspond to originally neighbouring vertices. Make it so that this matching connects vertices of the same colour if the corresponding entry in $\mathbf{B}$ was $+1$, and vertices of opposing colour if the corresponding entry in $\mathbf{B}$ was -1. This new graph is called a 2-lift or 2-cover of G.

How surprising is it to learn that the eigenvalues of the (standard) adjacency matrix of this new graph are precisely the union of the eigenvalues of $\mathbf{A}$ and those of $\mathbf{B}$?

The reason is not too sophisticated. Let $\overline{\mathbf{A}}$ be the adjacency matrix of this 2-cover. Picture $\overline{\mathbf{A}}$ as an $n \times n$ matrix, whose entries are equal to either

$$\begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} \quad \text{or} \quad \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}.$$

Multiply $\overline{\mathbf{A}}$ from left and right by a $2n \times 2n$ matrix, block diagonal, with n blocks, all of them equal to

$$\frac{1}{\sqrt{2}} \begin{pmatrix} 1 & 1 \\ 1 & -1 \end{pmatrix}.$$

Each block in $\overline{\mathbf{A}}$ gets transformed (diagonalized!), according to the rule

$$\begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} \to \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} \quad \text{or} \quad \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} \to \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}.$$

Now, rows and columns of $\overline{\mathbf{A}}$ are organized so that red and blue vertices appear intercalated. Reorganize them so that red vertices are the first n vertices, and blue vertices are the last. It follows that $\overline{\mathbf{A}}$ becomes a block matrix: one block equal to $\mathbf{A}$ and the other equal to $\mathbf{B}$.

4.5 Expander graphs

Later in this course we will discuss random walks in graphs in more details. Our goal now is to motivate the discussion in the next sections. A random walk typically converges to a limit distribution, and when this occurs, a random sample from the vertices of the graph can be well approximated by a local procedure of walking randomly with few steps. Thus, one usually wishes that this convergence happens fast. Several fields of research rely on methods to study the speed of this convergence, and on the possibility to generate graphs for which fast mixing takes place.

It is reasonable to believe that any subset of vertices in graph for which convergency occurs fast cannot be too big compared to the number of edges in its boundary. Thus it is convenient to define the edge expansion ratio of graph as

$$h(G) = \min_{S \subseteq V(G) : |S| \leq n/2} \frac{|E(S, \overline{S})|}{|S|},$$

where $E(S, \overline{S})$ denotes the set of edges between S and its complement. One therefore wishes that $h(G)$ is large for a given graph that is a candidate to have the fast mixing process.

A sequence of d-regular graphs G_i of increasing size is a *family of expander graphs* if there $\varepsilon > 0$ so that $h(G_i) \geq \varepsilon$ for all i.

Here is an example: define graphs G_m so that $V(G_m) = \mathbb{Z}_m \times \mathbb{Z}_m$. The neighbours of (x, y) are $(x \pm y, y), (x, y \pm x), (x \pm y + 1, y), (x, y \pm x + 1)$, with all operations $\mod m$. This family is due to Margulis, and is the first explicitly constructed family of expander. The expansion ratio is not easily determined.

The following result associated expansion ratio and the second eigenvalue:

Theorem 4.6 (Dodziuk, Alon-Milman). *Let G be d-regular, with eigenvalues $d = \lambda_1 \geq \cdots \geq \lambda_n$. Then*

$$\frac{d - \lambda_2}{2} \leq h(G) \leq \sqrt{2d(d - \lambda_2)}.$$

Thus, it is desirable that $d - \lambda_2$ is large, that is, λ_2 is small, in order for a graph to be an expander. How small can it be?

For a d-regular G with eigenvalues $lambda_1 \geq \cdots \geq \lambda_n$, let

$$\lambda = \max(|\lambda_2|, |\lambda_n|).$$

Theorem 4.7 (Alon-Boppana). *Let G be a d-regular graph on n vertices. Then*

$$\lambda_2 \geq 2\sqrt{d-1} - o_n(1),$$

with the understanding that $o_n(1)$ tends to 0 as $n \to \infty$, for any fixed d.

For a d-regular G with eigenvalues $lambda_1 \geq \cdots \geq \lambda_n$, let

$$\lambda = \max(|\lambda_2|, |\lambda_n|).$$

We will show the weaker version of the Theorem above for λ instead of λ_2.

Before showing the proof of this result, attempt the following easy exercise.

Exercise 4.8. Show that

$$\lambda \geq \sqrt{d}(1 - o_n(1)).$$

Hint: consider $\operatorname{tr} \mathbf{A}^2$.

Sketch of the proof.

- We have $\mathbf{A} = \mathbf{A}(G)$. Note that $\lambda(\mathbf{A}^{2\ell}) = \lambda(\mathbf{A})^{2\ell}$.
- Take two vertices a and b at distance D, and consider the vector $\mathbf{v} = \mathbf{e}_a - \mathbf{e}_b$. This vector is orthogonal to the all 1s vector, thus its Rayleigh quotient in $\mathbf{A}^{2\ell}$ lower bounds $\lambda^{2\ell}$. Assume $\ell < D/2$, thus $(\mathbf{A}^{2\ell})_{a,b} = 0$, and we have

$$\lambda^{2\ell} \geq \frac{\mathbf{v}^\mathsf{T} \mathbf{A}^{2\ell} \mathbf{v}}{2} = \frac{(\mathbf{A}^{2\ell})_{a,a} + (\mathbf{A}^{2\ell})_{a,b}}{2}.$$

- Let us lower bound $(\mathbf{A}^{2\ell})_{a,a}$. Picture a rooted tree at a, with d neighbours, each corresponding to the neighbours of a in G. To each, give them $d-1$ disjoint neighbours, each corresponding to their neighbours in G. Repeat. Note vertices of G will appear multiples times in this tree, and that is fine. Make this tree deep enough, say of depth larger than ℓ. Recall from Exercise 3.40 that this tree has largest eigenvalue smaller than $2\sqrt{d-1}$. Call this tree T.
- Any closed walk at a of length 2ℓ in T represents a closed walk at a of length 2ℓ in G. Thus

$$(\mathbf{A}(G)^{2\ell})_{a,a} \geq (\mathbf{A}(T)^{2\ell})_{a,a}.$$

  Note that the tree is the same if you start at b instead.
- Any closed walk at a of length $2k$ in T can be mapped to a sequence of decisions of going either forward or backwards down the tree. It is easy to see that the number of such sequences are counted by the Catalan numbers (check wikipedia), which are equal to $C_\ell = \binom{2\ell}{\ell}/(\ell+1)$. Each step forward had at least $d-1$ options, therefore

$$(\mathbf{A}(T)^{2\ell})_{a,a} \geq (d-1)^\ell \binom{2\ell}{\ell} \frac{1}{\ell+1}.$$

- Gathering up:

$$\lambda^{2\ell} \geq (d-1)^\ell \binom{2\ell}{\ell} \frac{1}{\ell+1}.$$

  Taking the 2ℓ-th root on both sides, and assuming that D and ℓ are largest possible, the result (eventually) follows.

□

4.6 Ramanujan graphs

A d-regular graph is a *Ramanujan graph* if $\lambda_2(G) \leq 2\sqrt{d-1}$. These are the best possible graphs you could find which guarantee a good expansion ratio. Infinite families of Ramanujan graphs therefore well sought after.

Lubotzky-Phillips-Sarnak (and independently Margulis) showed that arbitrarily large d-regular Ramanujan graphs exist when $d-1$ is prime (result later extender to prime powers).

For bipartite graphs, recall that the eigenvalues are symmetric about the origin. Thus an adjustment takes place in the definition of Ramanujan bipartite graphs: we say that a d-regular graph is a *bipartite Ramanujan graph* if all eigenvalue different than d and $-d$ lie within the interval $[-2\sqrt{d-1}, 2\sqrt{d-1}]$.

The goal of the rest of this section is to show the impressive result by Marcus, Spielman and Srivastava that infinite families of bipartite regular Ramanujan graphs exist for every fixed d.

So let us briefly review the strategy:

- Consider the matching polynomial of a graph G. It has real roots, and, moreover, if G has maximum degree d, then all roots have absolute value at most $2\sqrt{d-1}$.
- The average of the characteristic polynomials of the signed adjacency matrices of G is equal to $\mu_G(x)$.
- Start with G the complete bipartite graph, d-regular, and consider all possible signings. To each, a 2-cover is associated to the graph. The eigenvalues of the 2-covers are $d^{(1)}$,$0^{(2d-2)}$,$-d^{(1)}$, and the eigenvalues of the signed matrix.
- The average of these characteristic polynomials is μ, whose largest root is $2\sqrt{d-1}$.
- If we can show that at least one of them has largest root at most $2\sqrt{d-1}$, we choose this graph and repeat the procedure.
- The infinite family obtained this way is going to be an infinite family of bipartite Ramanujan graphs of degree d.

4.7 Interlacing families

Let $g(x) = \prod_{i=1}^{n-1}(x-\alpha_i)$, and $f(x) = \prod_{i=1}^{n}(x-\beta_i)$, with $\alpha_i, \beta_i \in \mathbb{R}$. We say that *$g$ interlaces f* if

$$\beta_1 \leq \alpha_1 \leq \beta_2 \leq \cdots \leq \beta_{n-1} \leq \alpha_{n-1} \leq \beta_n.$$

Real-rooted polynomials $f_1, \cdots, f_k$ of the same degree have a *common interlacer* if there is g so that g interlaces each f_i.

If $\beta_{i;j}$ denotes the jth smallest root of f_i, then it is easy to see that polynomials $f_1, \cdots, f_k$ have a common interlacer if and only if there are numbers $\alpha_0, \cdots, \alpha_n$ with $\beta_{i;j} \in [\alpha_{j-1}, \alpha_j]$ for all j. These numbers (except for α_0 and α_n will be the roots of a feasible interlacer g).

Lemma 4.9. *Let $f_1, \cdots, f_k$ be real-rooted polynomials of the same degree, all with positive leading coefficients. Define their sum*

$$f_\emptyset = \sum_{i=1}^{k} f_i.$$

If the polynomials f_i have a common interlacer, then there is an index j so that the largest root of f_j is upper bounded by the largest root of $f_\emptyset$.

Proof. Assume the common interlacer exists, let α be its largest root. It follows that for all i, $f_i(\alpha) \leq 0$, because the largest root of them all is at least α, and they are positive for larger values. Thus $f_\emptyset(\alpha) \leq 0$, and thus its largest root, say β, is at least α. As

$$0 = f_\emptyset(\beta) = \sum_{i=1}^{k} f_i(\beta),$$

it must be that for some j, $f_j(\beta) \geq 0$. Thus the largest root of f_j is between α and β. $\square$

Exercise 4.10. With the same assumptions (well, you may assume the polynomials have simple roots only), prove that $f_\emptyset$ is real-rooted. Also, show that there is an index j so that the mth largest root of f_j is upper bounded by the mth largest root of $f_\emptyset$.

We now want to extend the idea of a set of polynomials containing a common interlacer to the idea of a family of polynomials so that various subfamilies also have a common interlacer.

Let $S_1, \cdots, S_m$ be finite sets, and assume $\mathbf{s}_k$ denotes an element in $S = S_1 \times \cdots \times S_k$ (that is, $\mathbf{s}_k$ denotes a k-tuple of elements, with the jth element belonging to S_j, for $j = 1, ..., k$). For each such $\mathbf{s}_m$, assume polynomial $f_{\mathbf{s}_m}$ is defined, and it is real-rooted of degree n, with positive leading coefficient. For $k < m$, we define $f_{\mathbf{s}_k}$ as follows:

$$f_{\mathbf{s}_k} = \sum_{\mathbf{t} \, \in \, S_{k+1} \times \cdots \times S_m} f_{\mathbf{s}_k, \mathbf{t}}.$$

(Here we understand $\mathbf{s}_k, \mathbf{t}$ as the m-tuple obtained from concatenating the k elements from $\mathbf{s}_k$ to the $m - k$ elements from $\mathbf{t}$.)

We also interpret $f_{\mathbf{s}_0}$ as the sum of all polynomials in the family, or simply $f_\emptyset$. We say that the family $\{f_{\mathbf{s}_m}\}_{\mathbf{s}_m \in S}$ form an *interlacing family* if, for all $k \leq m - 1$, and all $\mathbf{s}_k$, we have that

$$\{f_{\mathbf{s}_k, t}\}_{t \in S_{k+1}} \quad \text{have a common interlacer.}$$

Theorem 4.11. *Let $S_1, ..., S_m$ be finite sets, and $\{f_{\mathbf{s}_m}\}$ an interlacing family. Then there is one $\mathbf{s}_m \in S_1 \times \cdots \times S_m$ so that the largest root of $f_{\mathbf{s}_m}$ is at most the largest root of $f_\emptyset$.*

Proof. This is an immediate iterative application of Lemma 4.9. The family $\{f_t\}_{t \in S_1}$ has a common interlacer and sum to $f_\emptyset$. So one applies the lemma. Fix that $s_1 \in S_1$ so that the largest root is at most that of $f_\emptyset$. Now consider the family $\{f_{s_1,t}\}_{t \in S_2}$, which sum to f_{s_1}. Apply the lemma. Repeat. In the end, there will be a sequence $((s_j))$, with $s_j \in S_j$, so that the largest root of $f_{s_1,...,s_j}$ is at most the largest root of $f_{s_1,...,s_{j-1}}$, for all j. $\square$

Assume the graph G contains m edges, and define sets $S_j = \{+1, -1\}$, for $j = 1, \cdots, m$. For any $\mathbf{s}_m \in S_1 \times \cdots \times S_m$, we will define $f_{\mathbf{s}_m}$ as the characteristic polynomial of the signed adjacency matrix of G whose kth edge has received sign $\mathbf{s}_m(k)$. Our goal is to prove that this family is an interlacing family, and therefore that there is an $f_{\mathbf{s}_m}$ whose largest root is at most that of $f_\emptyset$, which in this case is precisely a multiple of $\mu_G(x)$, whose largest root is at most $2\sqrt{d-1}$.

In order to achieve this, we need to show that several common interlacers exist. The existence of these interlacers will be a consequence of the following lemma.

Lemma 4.12. *Let $f_1, ..., f_k$ be polynomials of the same degree with positive leading coefficient. They admit a common interlacer if and only if, for all coefficients $\lambda_i \geq 0$ that sum to 1, we have that $\sum_{i=1}^{k} \lambda_i f_i$ is real rooted.*

Proof. Assume for simplicity first that all roots of the f_i are distinct. Then one direction is easy and follows straight from Exercise 4.10. In general, consider small perturbations of the polynomials to have only polynomials of simple roots, and apply the theorem to these.

For the other direction, assume first that $k = 2$ and also that the polynomials have distinct roots. So, for all t, $0 \leq t \leq 1$,

$$g_t(x) = tf_1(x) + (1-t)f_2(x)$$

is real rooted. Consider the intervals i_k, the vary from the kth largest root of $g_0(x)$ to the kth largest of $g_1(x)$. These are disjoint (why?). The result then follows.

If they have shared roots, then factor them out, apply the argument above, and add them back in. Be careful if the shared roots fall into one of the i_k intervals.

It is not too difficult to extend the result for k polynomials, but we leave it out of these notes. □

As a consequence of this result above, real-rootedness is somehow equivalent to the existence of common interlacers.

Assume for now the following result.

Theorem 4.13. *Let $p_1, \cdots, p_m$ be numbers from 0 to 1. Then the following polynomial is real rooted:*

$$\sum_{\mathbf{s}_m \in \{1,-1\}^m} \left(\prod_{i:s_i=1} p_i \right) \left(\prod_{i:s_i=-1} (1-p_i) \right) f_{\mathbf{s}_m}(x).$$

Theorem 4.14. *The polynomials $f_{\mathbf{s}_m}$ form an interlacing family.*

Proof. We fix $\mathbf{s}_k \in \{1,-1\}^k$, and we need to show that $f_{\mathbf{s}_k,1}$ and $f_{\mathbf{s}_k,-1}$ have a common interlacer. To see that, it is enough to verify that

$$tf_{\mathbf{s}_k,1} + (1-t)f_{\mathbf{s}_k,-1}$$

is real-rooted for all $t \in [0, 1]$.

This follows immediately from Theorem 4.13 by

- $p_i = \frac{1+s_i}{2}$ for i between 1 and k,
- $p_{k+1} = t$,
- $p_i = 1/2$, for $i \geq k+2$.

Thus, in the sum

$$\sum_{\mathbf{w}_m \in \{1,-1\}^m} \left(\prod_{i:w_i=1} p_i \right) \left(\prod_{i:w_i=-1} (1-p_i) \right) f_{\mathbf{w}_m}(x),$$

all terms whose prefix is different than $\mathbf{s}_k$ will be zero, and the remaining terms for $i \geq k+2$ only contribute with a constant. So this sum is a multiple of

$$t f_{\mathbf{s}_k,1} + (1-t) f_{\mathbf{s}_k,-1},$$

as we wanted. □

The result that we wanted to show now follows immediately.

- Start with $K_{d,d}$, the complete bipartite d-regular graph. Its eigenvalues are $d^{(1)}, 0^{(2d-2)}, -d^{(1)}$.
- Amongst all its signed adjacency matrices, pick that whose largest eigenvalue does not exceed that of $\mu_{K_{d,d}}$, which is $2\sqrt{d-1}$. It exists from a combination of Theorem 4.14, Theorem 4.11, Theorem 3.45, Exercise 3.42.
- Consider the covering graph determined by this signed adjacency matrix. All of its eigenvalues different from $\pm d$ lie within $\pm 2\sqrt{d-1}$, from Subsection 4.4.
- Repeat indefinitely.
- These will be a family of expanders, as $h(G) \geq (d - 2\sqrt{d-1})/2$ for all of them.
- Note that, implicitly here, we are using the fact that all these graphs are bipartite.

4.8 Real roots

Our goal in this section is to prove Theorem 4.13.

Theorem. *Let $p_1, \cdots, p_m$ be numbers from 0 to 1. Then the following polynomial is real rooted:*

$$\sum_{\mathbf{s}_m \in \{1,-1\}^m} \left(\prod_{i : s_i = 1} p_i \right) \left(\prod_{i : s_i = -1} (1 - p_i) \right) f_{\mathbf{s}_m}(x).$$

A multivariate polynomial $f \in \mathbb{R}[z_1, ..., z_n]$ is called *real stable* if either $f \equiv 0$ or it is never zero whenever the imaginary parts of all z_i are positive. Note that a univariate real stable polynomial contains only real roots. We will expose some that the polynomial in the Theorem statement is real stable.

Theory of stable polynomials is rich and vast, and I do not intend to provide here an informative survey. I will however list some useful results, and guide you through an understanding of how they can be applied to our task.

Hurwitz's theorem is an important result about what happens on non-vanishing functions converge to another function. It basically implies in our context that a polynomial obtained as the limit of a convergent sequence of stable polynomials is itself stable.

Exercise 4.15. Using this fact, argue why, if $f(x_1, ..., x_k)$ is stable, then for all $c \in \mathbb{R}$, we also have $f(x_1, ..., x_{k-1}, c)$ stable.

A very useful fact is the following lemma.

Lemma 4.16. *Let* $\mathbf{A}_1, ..., \mathbf{A}_m$ *be positive semidefinite matrices, and* $\mathbf{B}$ *Hermitian. Then*

$$f(x_1, ..., x_k) = \det(x_1\mathbf{A}_1 + ... + x_k\mathbf{A}_k + \mathbf{B})$$

is a real stable polynomial.

Sketch of proof. It is easy to see that f is a real polynomial. It follows from Hurwitz theorem that it is enough to show this result for when $\mathbf{A}_i$ is positive definite for all i, meaning, these are Hermitian matrices with positive eigenvalues.

Pick $\mathbf{z} \in \mathbb{C}^k$, write it as $\mathbf{z} = \mathbf{a} + \mathrm{i}\mathbf{b}$, and assume $\mathbf{b} > \mathbf{0}$. Let $\mathbf{Q} = \sum b_i\mathbf{A}_i$, and $\mathbf{H} = \sum a_i\mathbf{A}_i + \mathbf{B}$. It follows that $\mathbf{Q}$ is positive definite, and $\mathbf{H}$ is Hermitian, and that

$$f(\mathbf{z}) = \det(\mathbf{Q})\det(\mathrm{i}\mathbf{I} + \mathbf{Q}^{-1/2}\mathbf{H}\mathbf{Q}^{-1/2}.$$

As $\det(\mathbf{Q}) \neq 0$, it follows that $f(\mathbf{z}) = 0$ if and only if $-\mathrm{i}$ is an eigenvalue of the Hermitian matrix $\mathbf{Q}^{-1/2}\mathbf{H}\mathbf{Q}^{-1/2}$, which is not true. □

It is possible to show that differentiating with respect to one variable is is a linear map that preserves stability, and, with some more work, one obtains the following result

Lemma 4.17. *Let* $p, q \in \mathbb{R}_+$*, and consider variables* x *and* y*. Then if* $f(x, y)$ *is real stable, then so is*

$$(1 + p\partial_x + q\partial_y)f = f + p\partial_x f + q\partial_y f.$$

Exercise 4.18. Let $\mathbf{A}$ be invertible, and $\mathbf{u}$ and $\mathbf{v}$ vectors. The goal of this exercise is to show that

$$\begin{aligned}(x, y \to 0)(1 + p\partial_x + (1-p)\partial_y)\det(\mathbf{A} + x\mathbf{u}\mathbf{u}^\mathsf{T} + y\mathbf{v}\mathbf{v}^\mathsf{T}) &= \\ = p\det(\mathbf{A} + \mathbf{u}\mathbf{u}^\mathsf{T}) + (1-p)\det(\mathbf{A} + \mathbf{v}\mathbf{v}^\mathsf{T}).\end{aligned}$$

($(x, y \to 0)$ means that after applying $(1 + p\partial_x + (1-p)\partial_y)$, set variables to 0).

(a) Recall Exercise 3.11. Show that

$$\partial_t \det(\mathbf{A} + t\mathbf{u}\mathbf{u}^\mathsf{T}) = \det(\mathbf{A})(\mathbf{u}^\mathsf{T}\mathbf{A}^{-1}\mathbf{u}).$$

(b) Show that the expression on the left hand side of the equality is equal to

$$\det(\mathbf{A})(1 + p(\mathbf{u}^\mathsf{T}\mathbf{A}^{-1}\mathbf{u}) + (1-p)(\mathbf{v}^\mathsf{T}\mathbf{A}^{-1}\mathbf{v})).$$

(c) Use Exercise 3.11 again.

Theorem 4.19. *Let* $\mathbf{u}_i$*,* $\mathbf{v}_j$*,* $i, j = 1, ..., m$ *be vectors. Let* $p_1, ..., p_m$ *be real numbers between* 0 *and* 1*, and let* $\mathbf{D}$ *be positive semidefinite. Then*

$$\begin{aligned}P(x) = \sum_{\mathbf{s}_m \in \{1,-1\}^m} \left(\prod_{i:s_i=1} p_i\right)\left(\prod_{i:s_i=-1}(1-p_i)\right) \cdot \\ \cdot \det\left(x\mathbf{I} + \mathbf{D} + \sum_{i:s_i=1}\mathbf{u}_i\mathbf{u}_i^\mathsf{T} + \sum_{i:s_i=-1}\mathbf{v}_i\mathbf{v}_i^\mathsf{T}\right)\end{aligned}$$

is real rooted.

Sketch of proof. First define the multivariate polynomial

$$q(x, s_1, ..., s_m, t_1, ..., t_m) = \det\left(x\mathbf{I} + \mathbf{D} + \sum_{i:s_i=1} s_i \mathbf{u}_i \mathbf{u}_i^\mathsf{T} + \sum_{i:s_i=-1} t_i \mathbf{v}_i \mathbf{v}_i^\mathsf{T}\right).$$

It is real stable from Lemma 4.16. Then, successively apply the operator from Exercise 4.18 taking each pair of variables (s_i, t_i). The resulting polynomial, upon setting s_i and t_i to 0, will be exactly the polynomial of the statement. □

The only thing left now is to show that the polynomial in the statement of Theorem 4.13 is of the form of $P(x)$ as above.

Proof of Theorem 4.13. We will show that

$$\sum_{\mathbf{s}_m \in \{1,-1\}^m} \left(\prod_{i:s_i=1} p_i\right)\left(\prod_{i:s_i=-1} (1-p_i)\right) \det(x\mathbf{I} + \Delta(G)\mathbf{I} - \mathbf{A}_{\mathbf{s}_m})$$

is real rooted. Let $\mathbf{D}$ be a diagonal matrix of degrees. Note that

$$\mathbf{D} - \mathbf{A}_\mathbf{s} = \sum_{ab \in E(G)} (\mathbf{e}_a \pm \mathbf{e}_b)(\mathbf{e}_a \pm \mathbf{e}_b)^\mathsf{T},$$

where the sign is + if $\mathbf{s}(ab) = 1$, and − if $\mathbf{s}(ab) = -1$. Thus $\Delta\mathbf{I} - \mathbf{A}_\mathbf{s}$ is of the form $\mathbf{E} + \sum_{i:s_i=1} \mathbf{u}_i \mathbf{u}_i^\mathsf{T} + \sum_{i:s_i=-1} \mathbf{v}_i \mathbf{v}_i^\mathsf{T}$ with $\mathbf{E}$ positive semidefinite, and the result follows. □

5 Eigenvalues and the structure of graphs

5.1 Rayleigh quotients and Interlacing

Given a symmetric matrix $\mathbf{M}$, we recall the definition of the Rayleigh quotient of $\mathbf{M}$ with respect to a non-zero vector $\mathbf{v}$:

$$R_{\mathbf{M}}(\mathbf{v}) = \frac{\mathbf{v}^T\mathbf{M}\mathbf{v}}{\mathbf{v}^T\mathbf{v}}.$$

We will always assume vectors whose Rayleigh quotient is being taken are non-zero. As we have seen, if $\mathbf{v}$ is an eigenvector with corresponding eigenvalue θ, then

$$R_{\mathbf{M}}(\mathbf{v}) = \theta.$$

We also saw that if $\lambda_1 \geq ... \geq \lambda_n$ are the eigenvalues of $\mathbf{M}$ with corresponding eigenprojectors E_rs, and assuming $\mathbf{v}$ is normalized, then

$$R_{\mathbf{M}}(\mathbf{v}) = \mathbf{v}^T\left(\sum_{r=1}^{n}\lambda_r E_r\right)\mathbf{v} = \sum_{r=1}^{n}\lambda_r(\mathbf{v}^T E_r \mathbf{v}) \leq \lambda_1\left(\sum_{r=0}^{d}\mathbf{v}^T E_r \mathbf{v}\right) = \lambda_1$$

for all vectors $\mathbf{v}$, and equality holds if and only if $\mathbf{v}$ belongs to the λ_1 eigenspace.

Lemma 5.1. *Let* $\mathbf{M}$ *be a symmetric matrix, with largest eigenvalue* λ_1 *and smallest eigenvalue* λ_n*. Then*

$$\lambda_1 = \max_{\mathbf{v}\in\mathbb{R}^n} R_{\mathbf{M}}(\mathbf{v}) \quad \textit{and} \quad \lambda_n = \min_{\mathbf{v}\in\mathbb{R}^n} R_{\mathbf{M}}(\mathbf{v}).$$

□

Examining more carefully how we bounded the Rayleigh quotient, it is not hard to see that all eigenvalues can be defined as a max or min of the Rayleigh quotient over certain subspaces. Let L_r denote the orthogonal complement to the sum of the eigenlines corresponding to the largest eigenvalues all the way to λ_r, that is

$$L_r = \text{null }(E_1 + E_1 + ... + E_{r-1}).$$

Likewise, define S_r to correspond to the orthogonal complement to the sum of the eigenlines corresponding to the smallest eigenvalues all the way to λ_{r+1}, that is

$$S_r = \text{null }(E_{r+1} + E_{r+2} + ... + E_n).$$

It follows immediately that

$$\lambda_r = \max_{\mathbf{v}\in L_r} R_{\mathbf{M}}(\mathbf{v}) = \min_{\mathbf{v}\in S_r} R_{\mathbf{M}}(\mathbf{v}).$$

The expression of λ_r can be made with the subspaces L_r and S_r implicitly defined, via a min-max formula.

Lemma 5.2 (Courant–Fischer–Weyl min-max principle)**.** *Let* $\mathbf{M}$ *be a symmetric matrix, with eigenvalues* $\lambda_1 \geq ... \geq \lambda_n$*. Then*

$$\lambda_k = \min_{\substack{\textit{subspace } U \\ \dim U = n-k+1}} \max_{\mathbf{v}\in U} R_{\mathbf{M}}(\mathbf{v}) = \max_{\substack{\textit{subspace } U \\ \dim U = k}} \min_{\mathbf{v}\in U} R_{\mathbf{M}}(\mathbf{v}).$$

Proof. I will show the first equality only, as the second is analogous. Note that we have already seen that there is a subspace U of dimension $n-k+1$ so that

$$\lambda_k = \max_{\mathbf{v} \in U} R_{\mathbf{M}}(\mathbf{u}),$$

this subspace is simply the orthogonal complement of the sum of the eigenlines corresponding to the largest $k-1$ eigenvalues. The result will now follow if we verify that, for all subspaces U of dimension $n-k+1$, we have

$$\lambda_k \leq \max_{\mathbf{v} \in U} R_{\mathbf{M}}(\mathbf{u}).$$

To see this, let U be a subspace of dimension $n-k+1$, and let V be the sum of the eigenlines corresponding to the largest k eigenvalues. As $\dim U + \dim V$ exceeds n, it follows that $U \cap V \neq \emptyset$. Let $\mathbf{v}$ belong to this intersection. Then

$$R_{\mathbf{M}}(\mathbf{v}) \geq \lambda_k \sum_{r=1}^{k} \mathbf{v}^T E_r \mathbf{v} \geq \lambda_k,$$

as we wanted. □

Exercise 5.3. We've seen that $\Delta \geq \lambda_1$, the largest eigenvalue of $\mathbf{A}$. If the $(d_1, ..., d_n)$ is the degree sequence in decreasing order, then you can now show that $d_i \geq \lambda_i$.

Such min-max formula provides an alternative and meaningful definition of eigenvalues. For graph theory, it is hard to find interesting applications of this formula by itself. We can use it however to prove a strong result.

Theorem 5.4 (Cauchy's Interlacing)**.** *Let $\mathbf{A}$ be a symmetric $n \times n$ matrix and $\mathbf{S}$ be an $n \times m$ matrix satisfying $\mathbf{S}^T\mathbf{S} = \mathbf{I}$. Let $\mathbf{B} = \mathbf{S}^T\mathbf{A}\mathbf{S}$. Let $\theta_1 \geq ... \geq \theta_n$ be the eigenvalues of $\mathbf{A}$ and $\lambda_1 \geq ... \geq \lambda_m$ be those of $\mathbf{B}$. Then*

(a) For all k with $1 \leq k \leq m$,

$$\theta_{n-(m-k)} \leq \lambda_k \leq \theta_k$$

(b) If equality holds in either of the inequalities above for some λ_k eigenvalue of $\mathbf{B}$, then there is a λ_k-eigenvector $\mathbf{v}$ of $\mathbf{B}$ so that $\mathbf{S}\mathbf{v}$ is an eigenvector for λ_k in $\mathbf{A}$.

(c) Let $\mathbf{v}_1, ..., \mathbf{v}_m$ be an orthogonal basis of eigenvectors of $\mathbf{B}$, with $\mathbf{v}_i$ corresponding to λ_i. If for some $\ell \in \{1, ..., m\}$ we have that $\lambda_k = \theta_k$ for all $k = 1, ..., \ell$ (or $\lambda_k = \theta_{n-(m-k)}$ for all $k = \ell, ..., n$), then $\mathbf{S}\mathbf{v}_k$ is an θ_k eigenvector for $\mathbf{A}$ for $k = 1, ..., \ell$ (respectively for $k = \ell, ..., n$).

(d) If there is an $\ell \in \{1, ..., m\}$ so that $\lambda_k = \theta_k$ for all $k = 1, ..., \ell$, and $\lambda_k = \theta_{n-(m-k)}$ for all $k = \ell + 1, ..., n$, then $\mathbf{SB} = \mathbf{AS}$. In this case, interlacing is called tight.

Proof. Let $\mathbf{u}_1, ..., \mathbf{u}_n$ be the eigenvectors of $\mathbf{A}$ corresponding to the θ_ks. The key thing now is to observe that, for all k, the subspace

$$\langle \mathbf{v}_1, ..., \mathbf{v}_k \rangle \cap \langle \mathbf{S}^T\mathbf{u}_1, ..., \mathbf{S}^T\mathbf{u}_{k-1} \rangle^\perp$$

contains at least one vector. Let $\mathbf{w}$ be such vector, which, in particular, implies $\mathbf{Sw} \in \langle \mathbf{u}_1, ..., \mathbf{u}_{k-1} \rangle^{\perp}$. Then, by Lemma 5.2, we have

$$\theta_k \geq \frac{(\mathbf{Sw})^T \mathbf{A} (\mathbf{Sw})}{(\mathbf{Sw})^T (\mathbf{Sw})} \geq \frac{\mathbf{w}^T \mathbf{B} \mathbf{w}}{\mathbf{w}^T \mathbf{w}} \geq \lambda_k.$$

If $\theta_k = \lambda_k$, then $\mathbf{w}$ and $\mathbf{Sw}$ are eigenvectors for $\mathbf{B}$ and $\mathbf{A}$ respectively. Item (iii) follows easily by induction. Finally, with tight interlacing, we can guarantee that $\mathbf{Sv}_1, ..., \mathbf{Sv}_m$ are all eigenvectors for $\mathbf{A}$ with the same eigenvalues they have in $\mathbf{B}$. Therefore $\mathbf{SBv}_k = \mathbf{ASv}_k$ for all k, and as the set of eigenvectors form a basis, the two matrices are equal. □

The basic principle for applying interlacing is to carefully chose the matrix $\mathbf{S}$.

Exercise 5.5. Let $\mathbf{A}$ be a $n \times n$ symmetric matrix, with eigenvalues $\theta_1 \geq ... \geq \theta_n$. Let $\mathbf{B}$ be a principal submatrix, size $(n-1) \times (n-1)$, with eigenvalues $\lambda_1 \geq ... \geq \lambda_{n-1}$. Show that, for all k,

$$\theta_{k+1} \leq \lambda_k \leq \theta_k.$$

(This is actually the reason why the result above is called interlacing.)

The stability number of a graph is the size of the largest subset of vertices which contains no edge inside of it — known as an independent or stable set.

Exercise 5.6. Let $\alpha(G)$ be the stability number of a graph. Then

$$\alpha(G) \leq |\{k : \theta_k \geq 0\}| \quad \text{and} \quad \alpha(G) \leq |\{k : \theta_k \leq 0\}|.$$

This follows easily if you note that an independent set corresponds to a block of 0s in $\mathbf{A}(G)$. Write the details.

Consider now the Petersen graph. By using interlacing, we will show two interesting facts about it. The Petersen graph has eigenvalues $3, 1^{(5)}, -2^{(4)}$. If its incidence matrix is $\mathbf{N}$, then the adjacency matrix is $\mathbf{NN}^T - 3\mathbf{I}$, and the adjacency matrix of its line graph is $\mathbf{N}^T\mathbf{N} - 2\mathbf{I}$.

Exercise 5.7. Find the spectrum of its line graph.

If the Petersen graph contains a Hamilton cycle, then its line graph contains an induced cycle C_{10}. This means that we can delete 5 vertices of its line graph, and find C_{10}. The eigenvalues of C_{10} are

$$2, \pm\left(\frac{1 \pm \sqrt{5}}{2}\right), -2$$

whereas 2 and -2 are simple, and the others each have multiplicity 2.

Exercise 5.8. Use interlacing now to show that the Petersen graph does not have a Hamilton cycle.

Finally, an application of the method of finding a vector in an intersection of subspace by looking at the dimension. The graph K_{10} has 45 edges, and therefore it would be possible for the edges of K_{10} to be partitioned into copies of the Petersen graph. However, this is not possible. To see that, assume K_{10} contains already two disjoint copies of Pete, say G and H. The eigenspace corresponding to 1 in G has dimension 5, and the same for that of 1 in H. As both eigenspaces are orthogonal to the line spanned by $\mathbf{1}$, then there must be at least one vector, say $\mathbf{w}$, who is simultaneously an eigenvector for $\mathbf{A}(G)$ and $\mathbf{A}(H)$. Thus

$$\mathbf{A}(\mathbf{J} - \mathbf{I} - \mathbf{A}(G) - \mathbf{A}(H))\mathbf{w} = -3\mathbf{w}.$$

This means that the complement of G and H in K_{10} has eigenvalue -3, and therefore cannot be isomorphic to the Petersen graph.

5.2 Partitions - cliques, cocliques, colourings

Consider a partition of the vertex set of a graph G, with characteristic matrix $\mathbf{P}$ (meaning, rows of $\mathbf{P}$ are vertices, and columns are parts of the partition, with 1s and 0s indicating whether a vertex belongs or not to a part).

A partition of the vertex set of a graph with characteristic matrix $\mathbf{P}$ is called equitable with respect to $\mathbf{A}(G)$ if the column space of $\mathbf{P}$ is $\mathbf{A}(G)$-invariant, that is, if there is a matrix $\mathbf{B}$ so that

$$\mathbf{AP} = \mathbf{PB}.$$

Combinatorially, this means that the number of neighbours a vertex a has in a class C of the partition is determined uniquely by the class that contains a. In other words, any two vertices in a given class have the same number of neighbours in any other given class (including their own).

All graphs contain at least one equitable partition of its vertex set: that in which all classes are singletons. If the graph is regular, then the partition that contains only one class is also equitable.

Given any partition with characteristic matrix $\mathbf{P}$, it is always possible to scale each column of $\mathbf{P}$ so that it becomes a normal vector. If $\mathbf{S}$ is the matrix obtained in this manner, it is immediate to verify that

$$\mathbf{S}^T\mathbf{S} = \mathbf{I},$$

and therefore $\mathbf{S}$ is suitable to be used as in Theorem 5.4.

For a first example of this property, we derive another bound to the independence number of a graph.

Corollary 5.9 (Ratio bound for independent sets)**.** *Let G be k-regular on n vertices, with smallest eigenvalue θ_n. Then*

$$\alpha(G) \leq \frac{n(-\theta_n)}{k - \theta_n}.$$

If equality holds, then the partition of the vertex set into any maximum independent set and its complement is equitable, and in particular, there is a τ eigenvector which is constant in each class of this partition.

Proof. Let $\mathbf{P}$ be the characteristic matrix of a partition that contains two class: one is a maximum independent set, and the other is its complement. Let $\mathbf{S}$ be the normalized characteristic matrix. Then

$$\mathbf{S}^T\mathbf{A}\mathbf{S} = \begin{pmatrix} 0 & \frac{\alpha k}{\sqrt{\alpha}\sqrt{n-\alpha}} \\ \frac{\alpha k}{\sqrt{\alpha}\sqrt{n-\alpha}} & \frac{(n-\alpha)k - k\alpha}{n-\alpha} \end{pmatrix} = \begin{pmatrix} 0 & \frac{\sqrt{\alpha}k}{\sqrt{n-\alpha}} \\ \frac{\sqrt{\alpha}k}{\sqrt{n-\alpha}} & k - \frac{k\alpha}{n-\alpha} \end{pmatrix}.$$

Clearly, the eigenvalues are k and $(-k\alpha)/(n-\alpha)$. Due to interlacing, it follows that

$$(-k\alpha)/(n-\alpha) \geq \theta_n,$$

which rearranges to

$$\alpha \leq \frac{n(-\theta_n)}{k - \theta_n}.$$

If you can't compute the eigenvalues easily, you can simply compare the determinant of $\mathbf{S}^T\mathbf{A}\mathbf{S}$ with the product of the largest and smallest eigenvalues of $\mathbf{A}$.

If equality holds, and because the largest eigenvalue of $\mathbf{A}$ and $\mathbf{S}^T\mathbf{A}\mathbf{S}$ are also equal, we have that (iv) in Theorem 5.4 applies. Moreover, (ii) of said theorem implies the assertion about the τ-eigenvector. □

It is quite surprising that this bound is met in several interesting cases, although it is not a good approximation for α in the general case (no such hope exists).

Exercise 5.10. Let δ be the smallest degree of G. If G is any graph (not necessarily regular), with largest eigenvalue θ_1 and smallest eigenvalue θ_n. Show that

$$\alpha \leq \frac{n(-\theta_1\theta_n)}{\delta^2 - \theta_1\theta_n}.$$

Hint: let k be the average degree in the independent set, and proceed as above.

Exercise 5.11. Let G be k-regular on n vertices, with eigenvalues $\theta_1 \geq ... \geq \theta_n$. Assume G contains an induced subgraph H with n' vertices and m' edges. Show that

$$\theta_2 \geq \frac{2m'n - (n')^2k}{n'(n-n')} \geq \theta_n.$$

Characterize what happens if equality holds in either side.

Exercise 5.12. Let $\omega(G)$ be the size of a maximum clique in G, that is, the size of the largest subgraph of G which is isomorphic to a complete graph. Assume G is k-regular. Find an upper bound to ω using the eigenvalues of G.

We now devote our attention to the chromatic number of G. A colouring of $V(G)$ is an assignment of colours to the vertex set of G so that any two neighbours receive different colours. It is always possible to colour a graph with n colours. A graph is 2-colourable if and only if it is bipartite. The chromatic number of a graph $\chi(G)$ is the minimum number of colours necessary to colour the vertices of G.

Just like α and ω, χ is hard to approximate, so any simple formulas using the spectrum of G can only bound, and even so not that well in the general case. However, this is quite significant to the best one could do.

Exercise 5.13. Explain why

$$\alpha \cdot \chi \geq n \quad \text{and also} \quad \omega \leq \chi.$$

The first inequality in the exercise above immediately implies a spectral lower bound to χ in regular graphs, using the upper bound to α. As it turns out, we can ignore the requirement of the graph to be regular.

Theorem 5.14 (Hoffman)**.** *Let G be a graph with chromatic number χ, largest eigenvalue θ_1 and smallest eigenvalue θ_n. Then*

$$\chi(G) \geq 1 - \frac{\theta_1}{\theta_n}.$$

Proof. Let $\mathbf{P}$ be the characteristic matrix of a colouring. To prove this result, it won't be enough to simply scale the columns of $\mathbf{P}$ and proceed with interlacing (in fact, try to do this). Instead, we shall first scale the rows of $\mathbf{P}$. Let $\mathbf{D}$ be a diagonal matrix whose diagonal entries are taken from the Perron eigenvector $\mathbf{v}$ of G. Let $\mathbf{S}$ be the obtained from $\mathbf{DP}$ upon multiplying from the right by a diagonal matrix $\mathbf{E}$ which effects to normalizing its columns. Thus $\mathbf{S}^T\mathbf{S} = \mathbf{I}$, and we proceed with interlacing now. We have $\mathbf{B} = \mathbf{S}^T\mathbf{A}\mathbf{S}$ with 0s in the diagonal, as the support of $\mathbf{S}$ corresponds to a colouring of G. Note that $\mathbf{B}$ is $m \times m$ with $m = \chi$. We also note that θ is an eigenvalue of $\mathbf{B}$, because $\mathbf{S}^T\mathbf{A}\mathbf{S}(\mathbf{E}^{-1}\mathbf{1}) = \mathbf{S}^T\mathbf{A}\mathbf{v} = \theta_1\mathbf{E}^{-1}\mathbf{1}$. Hence, by interlacing,

$$0 = \operatorname{tr}\mathbf{B} = \lambda_1 + \lambda_2 + ... + \lambda_m \geq \theta_1 + (\chi - 1)\theta_n.$$

□

Exercise 5.15. What can you say if equality holds in this bound?

Note that in the last line of the proof, our bound was quite crude. An immediate improvement is to say

Corollary 5.16. *Let G be a graph with chromatic number χ, and eigenvalues $\theta_1 \geq ... \geq \theta_n$. Then*

$$\theta_1 + \theta_n + \theta_{n-1} + ... + \theta_{n-(\chi-2)} \leq 0.$$

□

Exercise 5.17. In this exercise, you will show that if $\theta_2 > 0$, then

$$\chi(G) \geq 1 - \frac{\theta_{n-\chi+1}}{\theta_2}.$$

I will give you a hint. Let $\mathbf{P}$ be the partition matrix of an optimal colouring. Let $\mathbf{v}_1$ be the Perron eigenvector, and $\mathbf{D}$ the diagonal matrix which contains its entries in the diagonal. Consider

$$\ker(\mathbf{P}^T\mathbf{D}) \cap \langle \mathbf{v}_n, ..., \mathbf{v}_{n-\chi+1} \rangle.$$

Prove that this intersection contains a vector, define a diagonal matrix with this vector, and also define $\mathbf{A}' = \mathbf{A} - (\theta_1 - \theta_2)\mathbf{v}_1\mathbf{v}_1^T$. Now proceed as in the proof of Hoffman's theorem.

Exercise 5.18. If m_n is the multiplicity of θ_n, verify now that

$$\chi \geq \min\{1 + m_m, 1 - \frac{\theta_n}{\theta_2}\}.$$

It is quite remarkable that in the results above, while exploring the connection between the eigenvalues of $\mathbf{A}$ and the independence number of G or its chromatic number, nowhere the fact that the entries of $\mathbf{A}$ are restricted to 1s and 0s was used. The only real constraint is that the non-zero entries are restricted to positions corresponding to edges in the graph. In fact, one can vary the entries of $\mathbf{A}$, and as long as the eigenvalue expressions increase (for chromatic number) or decrease (for independence number), better bounds will be obtained. This leads to the interesting topic of applications of semidefinite programming to algebraic graph theory.

There is a famous upper bound for χ:

Theorem 5.19 (Brooks)**.** *Let G be a graph with maximum degree Δ. Then $\chi(G) \leq \Delta$, unless G is a complete graph or an odd cycle, in which cases $\Delta + 1$ colours suffice.* □

This is one of the classical theorems in graph theory. Its proof is certainly not trivial (only purely combinatorial proofs are known, so you will have to research that on your own). I am sure you remember that $\theta_1 \leq \Delta$. It turns out, we can somehow strengthen the statement of Brooks theorem for several graphs.

Theorem 5.20 (Wilf)**.** *If G is a graph with chromatic number χ and largest eigenvalue θ_1, then*

$$\chi \leq 1 + \theta_1.$$

Equality holds if and only if G is an odd cycle or the complete graph.

Proof. Let G' be a subgraph of G which is χ-critical, meaning, the subgraph whose removal of any vertex decreases the chromatic number. In this subgraph, the degree of any vertex is at least $\chi - 1$ (why?). Thus its largest eigenvalue is at least $\chi - 1$. By interlacing, the largest eigenvalue of G is at least $\chi - 1$. □

Exercise 5.21. Finish the proof of the theorem above.

Exercise 5.22. One final exercise here: show that $\theta_n \geq -n/2$, for any connected graph.

Problem 5.1. *Find a spectral proof of Brooks theorem.*

5.3 Other eigenvalues

The second largest eigenvalue of $\mathbf{A}^2$ seems to be related to how "random" the graph looks like. This is very vague, I know, reason why it will be better explained looking at the results. Let us assume G is a regular graph of degree k, with eigenvalues $\theta_1 \geq \theta_2 \geq ... \geq \theta_n$. Let $\lambda > 0$ be such that λ^2 is the second largest eigenvalue of A^2, that is, $\lambda = \max\{|\theta_2|, |\theta_n|\}$.

Theorem 5.23. *Let G be k-regular, and λ as before. Let S and T be two subsets of $V(G)$, of respective sizes s and t. Let $e(S,T)$ be the number of edges from S to T. Then*

$$\left| e(S,T) - \frac{kst}{n} \right| \leq \lambda \sqrt{st\left(1-\frac{s}{n}\right)\left(1-\frac{t}{n}\right)} \leq \lambda\sqrt{st}.$$

Proof. Let

$$\mathbf{A} = \sum_{r=1}^{n} \theta_i E_i$$

be the spectral decomposition of $\mathbf{A}$ into 1-dimensional eigenspaces. Let χ_S and χ_T be the characteristic vectors of sets S and T. It follows that

$$e(S,T) = \chi_S^T \mathbf{A} \chi_T = \sum_{r=1}^{n} \theta_i (\chi_S^T E_i \chi_T).$$

Note that $(\chi_S^T E_1 \chi_T) = st/n$ therefore

$$\left| e(S,T) - \frac{kst}{n} \right| = \left| \sum_{r=2}^{n} \theta_i (\chi_S^T E_i \chi_T) \right| \leq \lambda \sum_{r=2}^{n} |\chi_S^T E_i \chi_T|.$$

By Cauchy Schwarz (applied twice),

$$\begin{aligned} \left| e(S,T) - \frac{kst}{n} \right| &\leq \lambda \sum_{r=2}^{n} \sqrt{\chi_S^T E_i \chi_S}\sqrt{\chi_T^T E_i \chi_T} \\ &\leq \sqrt{\sum_{r=2}^{n} \chi_S^T E_i \chi_S}\sqrt{\sum_{r=2}^{n} \chi_T^T E_i \chi_T} \\ &= \sqrt{s - \frac{s^2}{n}}\sqrt{t - \frac{t^2}{n}}. \end{aligned}$$

□

That is, if λ is small compared to k, then between any two subsets of vertices of the graph, the number of edges tends to be the “expected” number, had every edge been put randomly and independently in the graph.

Exercise 5.24. Can you use the result above (or its proof method?) to show the ratio bound for cocliques without going through interlacing?

If $u \in V(G)$, let $N(u)$ denote the neighbourhood of u.

Exercise 5.25. Let T be a subset of $V(G)$ of size t. Show that

$$\sum_{u \in V(G)} \left(|N(u) \cap T| - \frac{kt}{n} \right)^2 \leq \frac{t(n-t)}{n} \lambda^2.$$

From Theorem 5.23, you would indeed expect that a large ratio k/λ implies that the graph "looks" random. If that is indeed the case, the diameter would also be relatively small. We can turn this intuition into a result.

Theorem 5.26. *Let G be a k-regular connected graph on n vertices, $n \geq 2$. Let d be its diameter, $\theta_1 = k$ its largest eigenvalue, and $\lambda = \max\{\theta_2, |\theta_n|\}$. Assume G is not bipartite, thus $\lambda < k$. Then*

$$d \leq \left\lfloor \frac{\log(n-1)}{\log(k/\lambda)} \right\rfloor + 1.$$

Proof. Let m be an integer, with

$$m > \frac{\log(n-1)}{\log(k/\lambda)}.$$

Thus, $k^m > (n-1)\lambda^m$. We will show that this implies that all entries of A^m are positive, therefore $d \leq m$. To see that, we will again apply Cauchy-Schwarz twice. Note that

$$\begin{aligned}
(A^m)_{a,b} = \mathbf{e}_a^T A^m \mathbf{e}_b &= \frac{k^m}{n} + \sum_{r=2}^{n} \theta_r^m (\mathbf{e}_a^T \mathbf{E}_r \mathbf{e}_b) \\
&\geq \frac{k^m}{n} - \lambda^m \sum_{r=2}^{n} |\mathbf{e}_a^T \mathbf{E}_r \mathbf{e}_b| \\
&\geq \frac{k^m}{n} - \lambda^m \sum_{r=2}^{n} \sqrt{\mathbf{e}_a^T \mathbf{E}_r \mathbf{e}_a} \sqrt{\mathbf{e}_b^T \mathbf{E}_r \mathbf{e}_b} \\
&\geq \frac{k^m}{n} - \lambda^m \sqrt{\sum_{r=2}^{n} \mathbf{e}_a^T \mathbf{E}_r \mathbf{e}_a} \sqrt{\sum_{r=2}^{n} \mathbf{e}_b^T \mathbf{E}_r \mathbf{e}_b} \\
&= \frac{k^m}{n} - \lambda^m \left(1 - \frac{1}{n}\right).
\end{aligned}$$

This last term is positive if $k^m > (n-1)\lambda^m$. □

We now proceed to show how to associate the third eigenvalue of a graph to matchings. First, a result due to Tutte whose proof is purely combinatorial, and therefore we skip:

Theorem 5.27. *A graph G has no perfect matching if and only if there is a subset $S \subseteq V(G)$ so that the subgraph of G induced by $V \backslash S$ has more than $|S|$ odd components (that is, a connected component with an odd number of vertices).*

(Note however that one direction of the Theorem is very easy to show).

Again, we will be dealing with regular graphs (for the last time).

Theorem 5.28. *A connected k-regular graph G on n vertices, n even, and eigenvalues $\theta_1 \geq ... \geq \theta_n$, has a perfect matching if*

$$\theta_3 \leq \begin{cases} k - 1 + \frac{3}{k+1} & \text{if } k \text{ is even,} \\ k - 1 + \frac{3}{k+2} & \text{if } k \text{ is odd.} \end{cases}$$

Proof. Assume there is no perfect matching. By (the difficult direction of) Tutte's theorem, there is a set S of size s so that $V \backslash S$ has at least $s+2$ odd components (why not $s+1$ only?). Let $G_1, ..., G_q$ be each of one these, each of size n_i. Then

$$\sum_{i=1}^{q} e(G_i, S) \leq ks.$$

As $s \geq 1$, $e(G_i, S) \geq 1$, this implies $e(G_i, S) < k$ and $n_i > 1$ for at least three values of i. Say $i = 1, 2, 3$, ordered in such way that the largest eigenvalues of $\mathbf{A}(G_i)$, say λ_i, satisfy $\lambda_1 \geq \lambda_2 \geq \lambda_3$. Upon taking the union of these three graphs, we find $\theta_3 \geq \lambda_3$.

We now look at G_3. We have that its average degree is

$$\partial_3 = \frac{2|E(G_3)|}{n_3} = \frac{kn_3 - e(G_3, S)}{n_3} = k - \frac{e(G_3, S)}{n_3}.$$

Note that $e(G_3, S) < k$, and $n_3 > 1$, so $k < n_3$. If k is even, $e(G_3, S)$ is even. If k is odd, then $k \leq n_3 - 2$. Thus

$$\partial_3 \geq \begin{cases} k - \frac{k-2}{k+1} & \text{if } k \text{ is even,} \\ k - \frac{k-1}{k+2} & \text{if } k \text{ is odd.} \end{cases}$$

As G_3 is not regular, because $e(G_3, S)$, we have $\partial_3 < \lambda_3 \leq \theta_3$, as wished. □

As we have seen over several previous results, the hypothesis of a graph being regular comes in very handy when dealing with the eigenvalues of the adjacency matrix. Our goal is to introduce another type of adjacency matrix that shall overcome this necessity.

5.4 References

Here is the set of references used to write the past few pages.

W. Haemers's paper "Interlacing Eigenvalues of Graphs" is a standard reference for applications of interlacing to combinatorics.

More interlacing resources are Brouwer and Haemers's textbook "Spectra of Graphs", and Godsil and Royle's "Algebraic Graph Theory", Chapter 9.

For the theorem associating eigenvalues and matchings, the reference is Brouwer and Haemers's paper "Eigenvalues and Perfect Matchings".

The diameter bound is due to Fan Chung "Diameters and Eigenvalues".

6 Optimization for cliques and colourings

6.1 Geometry of inequalities

We are in $\mathbb{R}^n$, and given a few vectors, we will concern ourselves with some special linear combinations of these vectors.

If $\{\mathbf{u}_1, ..., \mathbf{u}_m\}$ are vectors and $\alpha_1, ..., \alpha_m$ are scalars, then

$$\alpha_1\mathbf{u}_1 + ... + \alpha_m\mathbf{u}_m$$

is a linear combination of the vectors, defined to be

(i) *affine combination* if $\sum \alpha_i = 1$,

(ii) *conical combination* if $\alpha_i \geq 0$ for all i,

(iii) *convex combination* if $\sum \alpha_i = 1$ and $\alpha_i \geq 0$ for all i.

The *affine hull* of the set $\{\mathbf{u}_1, ..., \mathbf{u}_m\}$ is the set of all affine combinations, and, similarly, the *convex hull* is the set of all convex combinations.

A subset of $\mathbb{R}^n$ is called *convex* if the segment between any two points of the set is entirely contained in the set. A *cone* is a set subset C of $\mathbb{R}^n$ so that if $\mathbf{v} \in C$, then $\alpha\mathbf{v} \in C$ for all $\alpha \geq 0$.

Exercise 6.1. Show that the convex hull of a set of vectors is precisely equal to the smallest convex set that contains those points.

A *hyperplane* $\mathcal{H}$ in $\mathbb{R}^n$ is the set of vectors that lie in the kernel of a linear functional, meaning, there is $\mathbf{a}$ so that

$$\mathcal{H} = \{\mathbf{x} : \mathbf{a}^\mathsf{T}\mathbf{x} = 0\}.$$

If instead of 0 we put a non-zero scalar β, then we call it an *affine hyperplane*. If we replace the equality sign by $\leq$, then the region defined is called a *half-space*.

Exercise 6.2. Verify that the intersection of convex sets is convex.

Exercise 6.3. Prove that a half-space is a convex set.

The intersection of a finite number of half-spaces is called a *polyhedron*, that is, any polyhedron $\mathbb{P}$ can be determined by a matrix $\mathbf{A}$ and a vector $\mathbf{b}$ as follows:

$$\mathbb{P} = \{\mathbf{x} : \mathbf{A}\mathbf{x} \leq \mathbf{b}\}.$$

A *polytope* is a bounded polyhedron. It is not hard to believe that a set is a polytope if and only if it is the convex hull of a finite number of points, though the proof is less straightforward.

Given $\mathbf{A}$ and $\mathbf{b}$, a major and important problem is to decide whether the polyhedron determined by the inequalities $\mathbf{A}\mathbf{x} \leq \mathbf{b}$ is empty or not. To that avail, we introduce the Theorem of Alternatives.

Theorem 6.4. *Given* $\mathbf{A}$ *and* $\mathbf{b}$*, exactly one of the following holds.*

(1) The system $\mathbf{A}\mathbf{x} \leq \mathbf{b}$ *has a solution.*

(2) There is $\mathbf{z}$ *with* $\mathbf{z} \geq \mathbf{0}$*,* $\mathbf{z}^{\mathsf{T}}\mathbf{A} = \mathbf{0}$*, and* $\mathbf{b}^{\mathsf{T}}\mathbf{z} < 0$*.*

Proof. Clearly both cannot be true at the same time. So we must show both cannot be false at the same time. The proof is by induction on the number of columns of $\mathbf{A}$. Within the induction there is an implicit algorithm known as the Fourier-Motzkin elimination procedure.

For the base case, note that if $\mathbf{A}$ has zero columns, then (1) simply states that $\mathbf{b} \geq \mathbf{0}$. If that fails, one coordinate is negative, say ith, so simply choose $\mathbf{z} = \mathbf{e}_i$.

Assume the theorem holds for any matrix with $n-1$ columns. Consider now $\mathbf{A}$ a $m \times n$ matrix. Assume wlog that the inequalities were scaled so that the magnitude of the non-zero coefficients of x_n is 1. Let I_+ be the row indices corresponding to positive coefficients of x_n, I_- to negative, and I_0 to the 0 coefficients. Therefore the system writes as

$$\begin{aligned} \sum_{k=1}^{n-1} \mathrm{A}_{ik}\mathrm{x}_k + \mathrm{x}_n &\leq \mathrm{b}_i, \quad \text{for all } i \in I_+, \\ \sum_{k=1}^{n-1} \mathrm{A}_{ik}\mathrm{x}_k - \mathrm{x}_n &\leq \mathrm{b}_i, \quad \text{for all } i \in I_-, \\ \sum_{k=1}^{n-1} \mathrm{A}_{ik}\mathrm{x}_k &\leq \mathrm{b}_i, \quad \text{for } i \in I_0. \end{aligned}$$

This implies

$$\max_{i \in I_-} \left(\sum_{k=1}^{n-1} \mathrm{A}_{ik}\mathrm{x}_k - \mathrm{b}_i \right) \leq \min_{i \in I_+} \left(\mathrm{b}_i - \sum_{k=1}^{n-1} \mathrm{A}_{ik}\mathrm{x}_k \right),$$

so the original system has a solution if and only if the following system also does

$$\begin{aligned} \sum_{k=1}^{n-1} (\mathrm{A}_{ik} + \mathrm{A}_{jk})\mathrm{x}_k &\leq \mathrm{b}_i + \mathrm{b}_j, \quad \text{for } i \in I_- \text{ and } j \in I_+, \\ \sum_{k=1}^{n-1} \mathrm{A}_{ik}\mathrm{x}_k &\leq \mathrm{b}_i, \quad \text{for } i \in I_0. \end{aligned}$$

If the original has no solution, the one above also is not, and, by induction, there are w_{ij}, $i \in I_i$ and $j \in I_+$, and , v_i, $i \in I_0$, so that, for all indices, $w_{ij} \geq 0$, $v_i \geq 0$, and

$$\sum_{i \in I_-, j \in I_+} (\mathrm{A}_{ik} + \mathrm{A}_{jk}) w_{ij} + \sum_{i \in I_0} \mathrm{A}_{ik} v_i = 0, \quad \text{for all } k,$$

and

$$\sum_{i \in I_-, j \in I_+} (\mathrm{b}_i + \mathrm{b}_j) w_{ij} + \sum_{i \in I_0} \mathrm{b}_i v_i < 0.$$

Now, simply define the vector $\mathbf{z}$ with

$$\begin{aligned} \mathrm{z}_i &= \sum_{j \in I_+} w_{ij}, \quad &\text{for } i \in I_-, \\ \mathrm{z}_j &= \sum_{i \in I_-} w_{ij}, \quad &\text{for } j \in I_+, \\ \mathrm{z}_i &= v_i, \quad &\text{for } i \in I_0. \end{aligned}$$

It satisfies $\mathbf{z} \geq \mathbf{0}$, $\mathbf{z}^\mathsf{T}\mathbf{A} = \mathbf{0}$, and $\mathbf{b}^\mathsf{T}\mathbf{z} < \mathbf{0}$. □

Exercise 6.5. Farka's lemma says that $\mathbf{A}\mathbf{x} = \mathbf{b}$ and $\mathbf{x} \geq \mathbf{0}$ has no solution if and only if there is $\mathbf{y}$ so that $\mathbf{y}^\mathsf{T}\mathbf{A} \geq \mathbf{0}$ and $\mathbf{b}^\mathsf{T}\mathbf{y} < 0$. Prove this (using the Theorem of the Alternatives).

6.2 Linear programs

A *linear program* is an optimization problem that attempts to find the maximum or minimum of a linear functional whose viability region is a polyhedron. Upon certain operations, an equivalent formulation to any LP has the form

$$\begin{aligned} \max \quad & \mathbf{c}^\mathsf{T}\mathbf{x} \\ \text{subject to} \quad & \mathbf{A}\mathbf{x} \leq \mathbf{b} \\ & \mathbf{x} \geq \mathbf{0}. \end{aligned}$$

The vector $\mathbf{x}$ corresponds to variables. The objective function is linear, and so must be the remaining constraints.

Optimization of combinatorial structures can be modelled with LPs provided integrality constraints are added to the variables $\mathbf{x}$. In this case, the optimization problem is called an *integer program*. For example, let $\mathbf{N}$ be the incidence matrix of a graph, with rows corresponding to vertices, and columns to edges. Then

$$\begin{aligned} \max \quad & \mathbf{1}^\mathsf{T}\mathbf{x} \\ \text{subject to} \quad & \mathbf{N}\mathbf{x} \leq \mathbf{1} \\ & \mathbf{x} \geq \mathbf{0} \\ & \mathbf{x} \in \mathbb{Z}^m \end{aligned}$$

yields an optimum solution that corresponds to the edges of a matching of maximum size.

A linear program can satisfy one of three possibilities. Either there is a vector $\mathbf{x}$ that satisfies the constraints and attains a maximum, or there is not, and this happens either because there are vectors satisfying the constraints of arbitrarily large objective value, or because there is no vector at all satisfying the constraints (this is because the objective function is continuous and the region of feasibility is a closed set).

To any linear program, one can define a dual program, as follows

$$\text{(P)} \quad \begin{aligned} \max \quad & \mathbf{c}^\mathsf{T}\mathbf{x} \\ \text{s.t.} \quad & \mathbf{A}\mathbf{x} \leq \mathbf{b} \\ & \mathbf{x} \geq \mathbf{0}. \end{aligned} \qquad \Bigg| \qquad \text{(D)} \quad \begin{aligned} \min \quad & \mathbf{b}^\mathsf{T}\mathbf{y} \\ \text{s.t.} \quad & \mathbf{A}^\mathsf{T}\mathbf{y} \geq \mathbf{c} \\ & \mathbf{y} \geq \mathbf{0}. \end{aligned}$$

This dual program was written carefully so that one can always guarantee that any feasible solution to the primal has objective value less or equal than any feasible solution to the dual. In fact, if $\mathbf{x}$ and $\mathbf{y}$ are a pair of primal-dual solutions, it follows that

$$\mathbf{c}^\mathsf{T}\mathbf{x} \leq (\mathbf{y}^\mathsf{T}\mathbf{A})\mathbf{x} = \mathbf{y}^\mathsf{T}(\mathbf{A}\mathbf{x}) \leq \mathbf{y}^\mathsf{T}\mathbf{b}.$$

This is known as **weak duality**, and it implies, in particular, that if any of the primal or dual is unbounded, then the other must be infeasible.

Perhaps surprisingly, if both have feasible solutions, then their respective optima have same objective value.

Theorem 6.6 (Strong duality)**.** *Assume (P) and (D) are feasible. Then their optima solutions have the same objective values.*

Proof. Consider the following big system of equations

$$\begin{pmatrix} \mathbf{A} & \\ & -\mathbf{A}^\mathsf{T} \\ -\mathbf{c}^\mathsf{T} & \mathbf{b}^\mathsf{T} \end{pmatrix} \begin{pmatrix} \mathbf{x} \\ \mathbf{y} \end{pmatrix} \leq \begin{pmatrix} \mathbf{b} \\ -\mathbf{c} \\ 0 \end{pmatrix}, \quad \text{with } \mathbf{x}, \mathbf{y} \geq \mathbf{0}.$$

Assume the theorem is false. Then, because of weak duality, the system above has no solution. By a variant of the Theorem of the Alternatives, there is a vector $\begin{pmatrix} \mathbf{u} & \mathbf{v} & q \end{pmatrix}^\mathsf{T} \geq \mathbf{0}$ so that

$$\begin{pmatrix} \mathbf{u} & \mathbf{v} & q \end{pmatrix}^\mathsf{T} \begin{pmatrix} \mathbf{A} & \\ & -\mathbf{A}^\mathsf{T} \\ -\mathbf{c}^\mathsf{T} & \mathbf{b}^\mathsf{T} \end{pmatrix} \geq \mathbf{0} \quad \text{and} \quad \begin{pmatrix} \mathbf{u} & \mathbf{v} & q \end{pmatrix}^\mathsf{T} \begin{pmatrix} \mathbf{b} \\ -\mathbf{c} \\ 0 \end{pmatrix} < 0.$$

Thus

$$\mathbf{u}^\mathsf{T}\mathbf{A} \geq q\mathbf{c}^\mathsf{T}, \quad \mathbf{A}\mathbf{v} \leq q\mathbf{b}, \quad \mathbf{u}^\mathsf{T}\mathbf{b} - \mathbf{v}^\mathsf{T}\mathbf{c} < 0.$$

This immediately leads to a contradiction. □

Exercise 6.7. Prove that it is only needed to assume that either the primal or the dual is feasible to get that the other is also feasible.

Exercise 6.8. Prove the complementary slackness conditions, ie, if $\mathbf{x}$ and $\mathbf{y}$ are a pair of respective optima solutions for a primal-dual pair of LPs as above, then it is not possible that a variable is non-zero while the corresponding inequality is not satisfied with equality. (Hint: write extra variables $\mathbf{u}$ so that $\mathbf{A}\mathbf{x} + \mathbf{u} = \mathbf{b}$ and variables $\mathbf{v}$ with $\mathbf{A}^\mathsf{T}\mathbf{y} - \mathbf{v} = \mathbf{c}$.)

There are two main reasons why we are discussing linear programs. The first is that eventually I will say something like "Remember the results we had for variables whose corresponding vector lies in the positive orthant of $\mathbb{R}^n$? They work pretty much the same for this other cone over here".

6.3 Fractional chromatic number

The other reason is that we can already associate this theory to some of the concepts we saw before

Let $\mathbf{M}$ be a matrix whose rows are indexed by the vertices of a graph G, and the columns by all cliques of the graph. The integer program

$$\begin{aligned} \max \quad & \mathbf{1}^{\mathsf{T}}\mathbf{x} \\ \text{subject to} \quad & \mathbf{M}^{\mathsf{T}}\mathbf{x} \leq \mathbf{1} \\ & \mathbf{x} \geq \mathbf{0} \\ & \mathbf{x} \in \mathbb{Z}^n \end{aligned}$$

is asking for the maximum number of vertices of the graph so that no two of them belong to the same clique, ie, a maximum coclique of the graph. If we drop the integrality constraints, we obtain an LP, to which we can write the dual

$$\text{(P)} \quad \begin{aligned} \max \quad & \mathbf{1}^{\mathsf{T}}\mathbf{x} \\ \text{s.t.} \quad & \mathbf{M}^{\mathsf{T}}\mathbf{x} \leq \mathbf{1} \\ & \mathbf{x} \geq \mathbf{0} \end{aligned} \qquad \Bigg| \qquad \text{(D)} \quad \begin{aligned} \min \quad & \mathbf{1}^{\mathsf{T}}\mathbf{y} \\ \text{s.t.} \quad & \mathbf{M}\mathbf{y} \geq \mathbf{1} \\ & \mathbf{y} \geq \mathbf{0}. \end{aligned}$$

Note that if we add integrality constraints to the dual, we will be finding the smallest number of cliques necessary to cover the vertices of G, meaning, $\chi(\overline{G})$. For this reason, the optimum of these LPs is denoted by $\chi_f(\overline{G})$, with "f" standing for fractional.

It is an immediate consequence that $\alpha(G) \leq \chi_f(\overline{G})$ (or, if you prefer, $\omega(G) \leq \chi_f(G)$.

Exercise 6.9. Show that

$$\chi_f(G) \geq \frac{|V(G)|}{\alpha(G)}.$$

Show that equality occurs if the graph is vertex-transitive (meaning: there is automorphism mapping any vertex to any vertex).

It seems plausible to believe optimization over certain subsets of vector spaces plays a big role in the theory behind α or ω, and χ or $\overline{\chi}$. We saw another example of this connection before: Hoffman's bound $\chi(G) \geq 1 - \theta_1(\mathbf{A})/\theta_n(\mathbf{A})$ held true even if we varied the non-zero entries of $\mathbf{A}$. Our goal over the next sections is to understand how these seemingly unrelated approaches are in fact very much related.

6.4 Positive semidefinite matrices

A real matrix $\mathbf{M}$ is positive semidefinite if it satisfies the following properties:

- $\mathbf{M}$ is symmetric.
- $\mathbf{v}^T\mathbf{M}\mathbf{v} \geq 0$ for all $\mathbf{v}$.

If the inequality is strict for all non-zero $\mathbf{v}$, then $\mathbf{M}$ is called positive definite. The only thing we want now is a characterization.

This is probably one of the most famous "exercises" in linear algebra.

Theorem 6.10. *Let* $\mathbf{M}$ *be a symmetric matrix. The following are equivalent.*

(a) $\mathbf{M}$ *is positive semidefinite.*

(b) The eigenvalues of $\mathbf{M}$ *are non-negative.*

(c) There exists a matrix $\mathbf{B}$ *so that* $\mathbf{M} = \mathbf{B}^T\mathbf{B}$.

(d) For all positive semidefinite matrices $\mathbf{A}$*, we have* $\langle \mathbf{M}, \mathbf{A}\rangle \geq 0$.

Proof. Assume (a). Let $\mathbf{Mv} = \theta\mathbf{v}$. Then $0 \leq \mathbf{v}^T\mathbf{Mv} = \theta\mathbf{v}^T\mathbf{v}$, thus $\theta \geq 0$. Assume (b). We diagonalize $\mathbf{M}$ as

$$\mathbf{M} = \mathbf{P}^T\mathbf{DP}.$$

As $\mathbf{D} \geq 0$, we have

$$\mathbf{M} = \mathbf{P}^T\sqrt{\mathbf{D}}\sqrt{\mathbf{D}}\mathbf{P} = (\sqrt{\mathbf{D}}\mathbf{P})^T(\sqrt{\mathbf{D}}\mathbf{P}).$$

Assume (c). Then

$$\langle \mathbf{M}, \mathbf{A}\rangle = \operatorname{tr}\mathbf{MA} = \operatorname{tr}\mathbf{B}^T\mathbf{BA} = \operatorname{tr}\mathbf{BAB}^T.$$

As $\mathbf{A}$ is psd, we have $\operatorname{tr}\mathbf{BAB}^T \geq 0$. Finally, assume (d). Take $\mathbf{A} = \mathbf{vv}^T$, which is clearly psd for any $\mathbf{v}$. We have $0 \leq \langle \mathbf{M}, \mathbf{vv}^T\rangle = \mathbf{v}^T\mathbf{Mv}$, as wished. □

Exercise 6.11. Show that $\mathbf{M}$ is positive semidefinite if and only if its principal minors are non-negative (use interlacing?). Recall, a principal minor is a determinant of a square submatrix symmetric about the main diagonal.

Exercise 6.12. Let $\mathbf{M} \succcurlyeq \mathbf{0}$. Given vectors $\mathbf{u}$ and $\mathbf{v}$, and $t \in \mathbb{R}$, write $(\mathbf{u} + t\mathbf{v})^\mathsf{T}\mathbf{M}(\mathbf{u} + t\mathbf{v})$ explicitly. You know that this is ≥ 0 for all t. Use this to conclude that

$$(\mathbf{u}^\mathsf{T}\mathbf{Mv})^2 \leq (\mathbf{u}^\mathsf{T}\mathbf{Mu})(\mathbf{v}^\mathsf{T}\mathbf{Mv}).$$

(This is probably our favourite proof of *Cauchy-Schwarz*'s).

Suppose the symmetric matrix $\mathbf{M}$ is written as

$$\mathbf{M} = \begin{pmatrix} \mathbf{A} & \mathbf{B} \\ \mathbf{C} & \mathbf{D} \end{pmatrix},$$

with $\mathbf{A}$ invertible. This leads to the factorization

$$\begin{pmatrix} \mathbf{A} & \mathbf{B} \\ \mathbf{C} & \mathbf{D} \end{pmatrix} = \begin{pmatrix} \mathbf{I} & \mathbf{0} \\ \mathbf{CA}^{-1} & \mathbf{I} \end{pmatrix} \begin{pmatrix} \mathbf{A} & \mathbf{0} \\ \mathbf{0} & \mathbf{D} - \mathbf{CA}^{-1}\mathbf{B} \end{pmatrix} \begin{pmatrix} \mathbf{I} & \mathbf{A}^{-1}\mathbf{B} \\ \mathbf{0} & \mathbf{I} \end{pmatrix}.$$

The matrix $\mathbf{S} = \mathbf{D} - \mathbf{CA}^{-1}\mathbf{B}$ is called the Schur complement of $\mathbf{A}$ in $\mathbf{M}$.

Theorem 6.13. *Given* $\mathbf{M}$ *in block form as above, with* $\mathbf{A}$ *invertible, then* $\mathbf{M} \succcurlyeq \mathbf{0}$ *if and only if* $\mathbf{A}$ *and* $\mathbf{S}$ *are also.*

Proof. Follows immediately from noting that $\mathbf{M} = \mathbf{P}^\mathsf{T} \begin{pmatrix} \mathbf{A} & \\ & \mathbf{S} \end{pmatrix} \mathbf{P}$, with a matrix $\mathbf{P}$ that is non-singular. □

This result gives the famous *Cholesky decomposition* of a matrix:

Theorem 6.14. *If* $\mathbf{A}$ *is positive semidefinite, there is a lower triangular matrix* $\mathbf{L}$ *so that* $\mathbf{A} = \mathbf{L}\mathbf{L}^\mathsf{T}$.

Proof. By induction on n. If $\mathrm{A}_{1,1} = 0$, then the first row and column of $\mathbf{A}$ are zero (why?), thus we can apply induction to the submatrix of $\mathbf{A}$ obtained upon deleting them. Otherwise, we have

$$\mathbf{A} = \begin{pmatrix} a & \mathbf{b}^\mathsf{T} \\ \mathbf{b} & \mathbf{A}_1 \end{pmatrix},$$

thus, there is a lower triangular $\mathbf{P}$ with

$$\mathbf{P}^{-\mathsf{T}}\mathbf{A}\mathbf{P}^{-1} = \begin{pmatrix} a & \mathbf{0} \\ \mathbf{0} & \mathbf{A}_1 - a^{-1}\mathbf{b}\mathbf{b}^\mathsf{T} \end{pmatrix},$$

where both diagonal blocks are positive semidefinite. By induction, both admit a Cholesky decomposition, which regrouped with $\mathbf{P}$ gives the Cholesky decomposition of $\mathbf{A}$. □

Exercise 6.15. How to use this decomposition to decide (efficiently) whether a given matrix is positive semidefinite?

6.5 Kronecker and Schur

If $\mathbf{A}$ and $\mathbf{B}$ are matrices, their *Kronecker product* $\mathbf{A}\otimes\mathbf{B}$ is the matrix obtained from replacing the ij entry of $\mathbf{A}$ by $\mathrm{A}_{i,j}\mathbf{B}$. You are invited to verify that the Kronecker product is bilinear, but, most notably, it satisfies

$$(\mathbf{A}\otimes\mathbf{B})(\mathbf{C}\otimes\mathbf{D}) = \mathbf{AC}\otimes\mathbf{BD},$$

provided all products are well defined. In particular, if $\mathbf{u}$ is eigenvector of $\mathbf{A}$ and $\mathbf{v}$ is one of $\mathbf{B}$, then $\mathbf{u}\otimes\mathbf{v}$ is eigenvector of $\mathbf{A}\otimes\mathbf{B}$ — its corresponding eigenvalue is the product of the original eigenvalues. This fact alone immediately implies the following:

Lemma 6.16. *If* $\mathbf{A}$ *and* $\mathbf{B}$ *are positive semidefinite, then so it is* $\mathbf{A}\otimes\mathbf{B}$.

The operator $\mathrm{vec}(\mathbf{A})$ takes $\mathbf{A}$, with n columns, to the column vector

$$\begin{pmatrix} \mathbf{A}\mathbf{e}_1 \\ \vdots \\ \mathbf{A}\mathbf{e}_n \end{pmatrix}.$$

Exercise 6.17. Convince yourself that

$$\mathrm{vec}(\mathbf{A}\mathbf{M}\mathbf{B}^\mathsf{T}) = (\mathbf{A}\otimes\mathbf{B})\,\mathrm{vec}(\mathbf{M}).$$

In other words, the linear map $\mathbf{M}\mapsto\mathbf{A}\mathbf{M}\mathbf{B}^\mathsf{T}$ is represented by $\mathbf{A}\otimes\mathbf{B}$.

The *Schur product* $\mathbf{A}\circ\mathbf{B}$ of two matrices with the same shape is defined as the entry-wise product (your high school dream).

Exercise 6.18. Prove that if $\mathbf{A}$ and $\mathbf{B}$ are positive semidefinite, then $\mathbf{A} \circ \mathbf{B}$ is also. (Kronecker and Schur are together in this section and it's not a coincidence).

Our goal in the next section is to look at

$$\begin{aligned} \max \quad & \mathbf{c}^\mathsf{T}\mathbf{x} \\ \text{subject to} \quad & \mathbf{A}\mathbf{x} \leq \mathbf{b} \\ & \mathbf{x} \geq \mathbf{0}, \end{aligned}$$

and replace $\mathbf{x}$, vector in $\mathbb{R}^n$ with non-negative entries, by a positive semidefinite matrix. Of course, we need to build some theory to grasp how each component of the program above will find its analagous. Hopefully, we will gain more power in optimizing, while still maintaining the possibility to solve the programs.

6.6 Cones

An Euclidean space is a vector space over the reals of finite dimension and equipped with an inner product. For example, $\mathbb{R}^n$ and $\mathbb{S}^n$ are Euclidean spaces, and the respective inner products are given by $\langle \mathbf{x}, \mathbf{y} \rangle = \mathbf{x}^\mathsf{T}\mathbf{y}$, and $\langle \mathbf{A}, \mathbf{B} \rangle = \operatorname{tr} \mathbf{AB}$. A cone $\mathbb{K}$ is a subset so that for all $\mathbf{x} \in \mathbb{K}$, we have $\alpha\mathbf{x} \in \mathbb{K}$ for all $\alpha > 0$.

We introduce four properties about cones that shall be useful to us. A cone $\mathbb{K}$ is called...

- *...closed* if no point that is the limit of a sequence of points in $\mathbb{K}$ is outside of $\mathbb{K}$ — for example, $\mathbb{R}^n_+$ is closed, but $\mathbb{R}^n_{++}$ is not;
- *...pointed* if $\mathbb{K}$ contains no line about the origin, equivalently, if $\mathbf{x} \in \mathbb{K}$ and $-\mathbf{x} \in \mathbb{K}$, then $\mathbf{x} = \mathbf{0}$;
- *...convex* if $\mathbb{K} + \mathbb{K} \subseteq \mathbb{K}$, i.e., the line segment between any two points in $\mathbb{K}$ belongs entirely to $\mathbb{K}$ — for example, $\mathbb{S}^n_+$ is convex (a fact that is trivial only if you use the correct condition from Theorem **??**...);
- *...a cone with non-empty interior* if there is at least one point in $\mathbb{K}$ so that a ball of positive radius around that point is entirely contained in $\mathbb{K}$ — for example, the vectors $\mathbf{x}$ in $\mathbb{R}^n$ with $\mathbf{x} \geq \mathbf{0}$ and $\mathrm{x}_1 = 0$ for a cone that contains no interior point.

Exercise 6.19. Consider the *Lorentz cone*, defined as

$$\mathbb{K} = \{(\mathbf{x}, t) \in \mathbb{R}^n \times \mathbb{R} : \|\mathbf{x}\| \leq t\}.$$

Verify if it satisfies each of the four properties above. Then, (try to) do the same for the cone of *copositive matrices*, defined as

$$\mathbb{P} = \{\mathbf{M} \in \mathbb{S}^n : \mathbf{x}^\mathsf{T}\mathbf{M}\mathbf{x} \geq 0 \text{ for all } \mathbf{x} \geq 0\}.$$

If $\mathbb{V}_1$ and $\mathbb{V}_2$ are Euclidean spaces, let $\mathbf{c} \in \mathbb{V}_1$, $\mathcal{A} : \mathbb{V}_1 \to \mathbb{V}_2$ be linear, and $\mathbf{b} \in \mathbb{V}_2$. Let $\mathbb{K}$ be a closed, pointed and convex cone, with non-empty interior. A *conic program* is an optimization problem that can be written in the form

$$\begin{aligned} \max \quad & \langle \mathbf{c}, \mathbf{x} \rangle \\ \text{subject to} \quad & \mathcal{A}(\mathbf{x}) = \mathbf{b} \\ & \mathbf{x} \in \mathbb{K}. \end{aligned}$$

(In some contexts, a conic program is presented with "sup" instead of "max", to highlight the fact that it can be feasible and bounded while having no optimal solution. I will present all programs with "max" and "min", understanding that those might not exist even if the program is feasible and bounded).

For a linear program, $\mathbb{K} = \mathbb{R}^n_+$. For a semidefinite program, $\mathbb{K} = \mathbb{S}^n_+$. In both cases you can view $\mathbb{V}_1$ or $\mathbb{V}_2$ as isomorphic to $\mathbb{R}^k$ for some k, though in the second case this has to be made explicit via an isomorphism such as $\text{vec} : \mathbb{R}^{n \times n} \to \mathbb{R}^{n^2}$, as we discussed.

In order to better understand conic programs (and their duals), let us introduce some more theory. Given $\mathcal{A} : \mathbb{V}_1 \to \mathbb{V}_2$, linear, its *adjoint* is the (unique) linear transformation $\mathcal{A}^* : \mathbb{V}_2 \to \mathbb{V}_1$ that satisfies

$$\langle \mathcal{A}(\mathbf{x}), \mathbf{y} \rangle_{\mathbb{V}_2} = \langle \mathbf{x}, \mathcal{A}^*(\mathbf{y}) \rangle_{\mathbb{V}_1}.$$

Note that when $\mathbb{V}_1 = \mathbb{R}^n$ and $\mathbb{V}_2 = \mathbb{R}^m$, with the conventional inner product, and $\mathcal{A}$ is represented by a $m \times n$ matrix $\mathbf{A}$, then $\mathcal{A}^*$ is given by $\mathbf{A}^\mathsf{T}$.

Exercise 6.20. Consider the linear operator $\text{diag} : \mathbb{R}^{n \times n} \to \mathbb{R}^n$, that acts as

$$\text{diag}(\mathbf{X}) = \sum_{i=1}^{n} \mathbf{X}_{ii} \mathbf{e}_i.$$

Describe its adjoint (hint: we will denote it by Diag).

Given $\mathbf{x}, \mathbf{y} \in \mathbb{V}$, we say that

$$\mathbf{x} \succcurlyeq_{\mathbb{K}} \mathbf{y} \iff \mathbf{x} - \mathbf{y} \in \mathbb{K}.$$

Exercise 6.21. You are invited to prove (or at least think about and convince yourself of) the following properties:

(1) $\mathbf{x} \succcurlyeq_{\mathbb{K}} \mathbf{y} \iff \mathbf{y} \succcurlyeq_{-\mathbb{K}} \mathbf{x}$

(2) $\succcurlyeq_{\mathbb{K}}$ is a partial order, meaning, for all $\mathbf{x}, \mathbf{y}, \mathbf{z}$ in $\mathbb{V}$, we have $\mathbf{x} \succcurlyeq_{\mathbb{K}} \mathbf{x}$; $\mathbf{x} \succcurlyeq_{\mathbb{K}} \mathbf{y}$ and $\mathbf{y} \succcurlyeq_{\mathbb{K}} \mathbf{x}$ imply $\mathbf{x} = \mathbf{y}$; and $\mathbf{x} \succcurlyeq_{\mathbb{K}} \mathbf{y}$ and $\mathbf{y} \succcurlyeq_{\mathbb{K}} \mathbf{z}$ imply $\mathbf{x} \succcurlyeq_{\mathbb{K}} \mathbf{z}$.

(3) $\mathbf{x} \succcurlyeq_{\mathbb{K}} \mathbf{y}$ imply $\alpha\mathbf{x} \succcurlyeq_{\mathbb{K}} \alpha\mathbf{y}$ for all $\alpha > 0$.

(4) $\mathbf{x} \succcurlyeq_{\mathbb{K}} \mathbf{y}$ and $\mathbf{u} \succcurlyeq_{\mathbb{K}} \mathbf{v}$ imply $\mathbf{x} + \mathbf{u} \succcurlyeq_{\mathbb{K}} \mathbf{y} + \mathbf{v}$.

Exercise 6.22. If $\mathcal{A}$ is a linear map from $\mathbb{V}_1$ to $\mathbb{V}_2$, and $\mathbb{K}$ is a convex cone in $\mathbb{V}_1$, show that $\mathcal{A}\mathbb{K}$ is a convex cone in $\mathbb{V}_2$.

The *dual of a cone* $\mathbb{K} \subseteq \mathbb{V}$ is defined as

$$\mathbb{K}^* = \{\mathbf{x} \in \mathbb{V} : \langle \mathbf{x}, \mathbf{s} \rangle \geq 0 \text{ for all } \mathbf{s} \in \mathbb{K}\}.$$

Lemma 6.23. *If $\mathbb{K}$ is a cone, then $\mathbb{K}^*$ is a closed convex cone.*

Proof. First, if $\mathbf{x} \in \mathbb{K}^*$, then $\alpha\mathbf{x} \in \mathbb{K}^*$ for all $\alpha > 0$, as $\langle \alpha\mathbf{x}, \mathbf{s} \rangle = \alpha\langle \mathbf{x}, \mathbf{s} \rangle \geq 0$ for all $\mathbf{s} \in \mathbb{K}$.

Second, $\mathbb{K}^*$ is closed because it is the intersection of closed sets (research this fact from analysis if you want).

Third, $\mathbb{K}^*$ is convex because it is the intersection of convex sets (remember that you already showed that half-spaces are convex). □

Exercise 6.24. Which are the cones $(\mathbb{R}^n_+)^*$ and $(\mathbb{S}^n_+)^*$? Could you also describe the duals of the two cones appearing in Exercise 6.19 ?

A cone that is equal to its dual is called a *self-dual cone*.

We end this section stating an important result about cones and their duals. Its proof is beyond the scope of this course.

Theorem 6.25. *If $\mathbb{K}$ is a closed, convex cone, then so it is $\mathbb{K}^*$, and $\mathbb{K}^{**} = \mathbb{K}$. Moreover,*

(i) If K is pointed, then $\mathbb{K}^*$ has non-empty interior.

(ii) If K has non-empty interior, then $\mathbb{K}^*$ is pointed.

6.7 Duality

Say $\mathbb{V}_1$ and $\mathbb{V}_2$ are Euclidean spaces, let $\mathbf{c} \in \mathbb{V}_1$, $\mathcal{A} : \mathbb{V}_1 \to \mathbb{V}_2$ be linear, and $\mathbf{b} \in \mathbb{V}_2$. Let $\mathbb{K} \subseteq \mathbb{V}_1$ and $\mathbb{L} \subseteq \mathbb{V}_2$ be closed, convex cones. Consider the conic program

$$\begin{array}{rl} \max & \langle \mathbf{c}, \mathbf{x} \rangle_{\mathbb{V}_1} \\ \text{subject to} & \mathcal{A}(\mathbf{x}) \preccurlyeq_{\mathbb{L}} \mathbf{b} \\ & \mathbf{x} \in \mathbb{K}. \end{array}$$

Recall linear program duality, where from

$$\begin{array}{crl|crl} & \max & \mathbf{c}^\mathsf{T}\mathbf{x} & & \min & \mathbf{b}^\mathsf{T}\mathbf{y} \\ \text{(P)} & \text{s.t.} & \mathbf{A}\mathbf{x} \leq \mathbf{b} & \text{(D)} & \text{s.t.} & \mathbf{A}^\mathsf{T}\mathbf{y} \geq \mathbf{c} \\ & & \mathbf{x} \geq \mathbf{0} & & & \mathbf{y} \geq \mathbf{0} \end{array}$$

we obtained immediately that

$$\mathbf{c}^\mathsf{T}\mathbf{x} \leq \mathbf{y}^\mathsf{T}\mathbf{A}\mathbf{x} \leq \mathbf{y}^\mathsf{T}\mathbf{b}.$$

Let us now define a conic program crafted in such way that the objective value of any of its feasible solutions is an upper bound on the objective value of the conic program above.

- We know that $\mathbf{b} - \mathcal{A}(\mathbf{x}) \in \mathbb{L}$, thus $\langle \mathbf{b} - \mathcal{A}(\mathbf{x}), \mathbf{y} \rangle_{\mathbb{V}_2} \geq 0$ for all $\mathbf{y} \in \mathbb{L}^*$.
- Therefore, $\langle \mathbf{b}, \mathbf{y} \rangle_{\mathbb{V}_2} \geq \langle \mathcal{A}(\mathbf{x}), \mathbf{y} \rangle_{V_2} = \langle \mathbf{x}, \mathcal{A}^*(\mathbf{y}) \rangle_{V_1}$.

- Now, if we require $\mathcal{A}^*(\mathbf{y}) - \mathbf{c} \in \mathbb{K}^*$, it follows that, for all $\mathbf{x} \in \mathbb{K}$, we have $\langle \mathcal{A}^*(\mathbf{y}) - \mathbf{c}, \mathbf{x} \rangle_{\mathbb{V}_1} \geq 0$, hence $\langle \mathcal{A}^*(\mathbf{y}), \mathbf{x} \rangle_{\mathbb{V}_1} \geq \langle \mathbf{c}, \mathbf{x} \rangle_{\mathbb{V}_1}$.

We arrive at the following primal-dual formulation for conic programs:

$$\text{(P)} \quad \begin{array}{rl} \max & \langle \mathbf{c}, \mathbf{x} \rangle_{\mathbb{V}_1} \\ \text{s.t.} & \mathcal{A}(\mathbf{x}) \preccurlyeq_{\mathbb{L}} \mathbf{b} \\ & \mathbf{x} \in \mathbb{K}. \end{array} \qquad \Bigg| \qquad \text{(D)} \quad \begin{array}{rl} \min & \langle \mathbf{b}, \mathbf{y} \rangle_{\mathbb{V}_2} \\ \text{s.t.} & \mathcal{A}^*(\mathbf{y}) \succcurlyeq_{\mathbb{K}^*} \mathbf{c} \\ & \mathbf{y} \in \mathbb{L}^*. \end{array}$$

The discussion above has shown the weak duality theorem:

Theorem 6.26. *For a pair of primal-dual conic programs such as above, if $\mathbf{x}$ is feasible for (P) and $\mathbf{y}$ is feasible for (D), then $\langle \mathbf{c}, \mathbf{x} \rangle_{\mathbb{V}_1} \leq \langle \mathbf{b}, \mathbf{y} \rangle_{\mathbb{V}_2}$.* □

Strong duality for linear programs guarantees that optimum solutions for the primal and dual, if they exist, must be equal. A similar theorem holds true also for semidefinite programs: under some natural and mild extra conditions, equality also holds for the optimum of the primal and dual formulations. I will not show this theorem in this course: you are however invited the research it on your own.

All semidefinite programs we will study in this course satisfy these mild conditions to guarantee strong duality, therefore from here on we assume that for a pair of primal-dual SDPs, if they are bounded, they achieve their optimum values, which coincide.

Exercise 6.27. Write the dual of the conic program

$$\begin{array}{rl} \max & \langle \mathbf{c}, \mathbf{x} \rangle_{\mathbb{V}_1} \\ \text{subject to} & \mathcal{A}(\mathbf{x}) = \mathbf{b} \\ & \mathbf{x} \in \mathbb{K}. \end{array}$$

Exercise 6.28. Using Theorem 6.25, verify that the dual of the dual is the primal.

Let us now come back to something more concrete. Assume $\mathbb{V}_1 = \mathbb{S}^n$, and $\mathbb{V}_2 = \mathbb{R}^m$. Have $\mathbb{K} = \mathbb{S}^n_+$. How do linear functions from $\mathbb{S}^n$ to $\mathbb{R}^m$ look like?

Well, if $f : \mathbb{S}^n \to \mathbb{R}$ is a linear functional, then there exists a unique $\mathbf{M}$ so that, for all $\mathbf{X} \in \mathbb{S}^n$, $f(\mathbf{X}) = \langle \mathbf{X}, \mathbf{M} \rangle$. This can be easily obtained — in fact, if $\{\mathbf{M}_1, ..., \mathbf{M}_\ell\}$ form an orthonormal basis, then

$$\mathbf{M} = \sum f(\mathbf{M}_i)\mathbf{M}_i.$$

Therefore, if $\mathcal{A} : \mathbb{S}^n \to \mathbb{R}^m$ is linear, there are matrices $\mathbf{A}_1$,..., $\mathbf{A}_m$ so that, for all $\mathbf{X}$,

$$\mathcal{A}(\mathbf{X}) = \begin{pmatrix} \langle \mathbf{A}_1, \mathbf{X} \rangle \\ \vdots \\ \langle \mathbf{A}_m, \mathbf{X} \rangle \end{pmatrix}.$$

Exercise 6.29. Given $\mathcal{A}$ with matrices $\mathbf{A}_i$ as above, describe who is $\mathcal{A}^*(\mathbf{y})$ for $\mathbf{y} \in \mathbb{R}^m$.

Exercise 6.30. Given the positive semidefinite program

$$\begin{aligned} \max \quad & \langle \mathbf{C}, \mathbf{X} \rangle \\ \text{subject to} \quad & \langle \mathbf{A}_i, \mathbf{X} \rangle \leq \mathrm{b}_i \text{ for } i = 1, ..., m \\ & \mathbf{X} \succcurlyeq \mathbf{0}, \end{aligned}$$

write its dual.

Exercise 6.31. Assume a positive semidefinite program (P) and its dual (D) have feasible solution with the same objective value. What would complementary slackness conditions say about them?

6.8 Three SDP formulations

Lovász Theta

We now introduce the Lovász Theta parameter of a graph. Given a graph G, with adjacency matrix $\mathbf{A}$, let $\mathbf{z}$ denote the characteristic vector of an independent set S. If $\mathbf{Z} = (1/|S|)\mathbf{z}\mathbf{z}^\mathsf{T}$, then $\operatorname{tr} \mathbf{Z} = 1$, $\mathbf{Z} \succcurlyeq \mathbf{0}$, and $\mathbf{Z}_{ij} = 0$ for all edges ij of the graph. As such, it is a feasible solution to the semidefinite program

$$\begin{aligned} \max \quad & \langle \mathbf{J}, \mathbf{X} \rangle \\ \text{subject to} \quad & \mathrm{X}_{ij} = 0 \text{ for all edges } ij \text{ of the graph,} \\ & \operatorname{tr} \mathbf{X} = 1, \\ & \mathbf{X} \succcurlyeq \mathbf{0}, \end{aligned}$$

and its objective value is equal to $|S|$.

Exercise 6.32. You are invited to check that all constraints are of the form $\langle *, \mathbf{X} \rangle = *$.

Exercise 6.33. Write the dual of the program above.

The common optimum value of the program above and its dual is defined as the Lovász theta parameter of the graph, denoted by $\vartheta(G)$.

Eigenvalues

Recall that given $\mathbf{M} \in \mathbb{S}^n$ with largest eigenvalue θ and smallest eigenvalue τ, we have

$$\theta = \max_{\mathbf{x} \in \mathbb{R}^n,\ \|\mathbf{x}\|=1} \mathbf{x}^\mathsf{T} \mathbf{M} \mathbf{x},$$

and

$$\tau = \min_{\mathbf{x} \in \mathbb{R}^n,\ \|\mathbf{x}\|=1} \mathbf{x}^\mathsf{T} \mathbf{M} \mathbf{x},$$

Consider now that SDP

$$\begin{aligned} \max \quad & \langle \mathbf{M}, \mathbf{X} \rangle \\ \text{subject to} \quad & \operatorname{tr} \mathbf{X} = 1, \\ & \mathbf{X} \succcurlyeq \mathbf{0}. \end{aligned}$$

Exercise 6.34. Write its dual. Conclude (trivially!) that the dual has an optimal solution with objective value θ. Knowing this, conclude now that the primal also has an optimal solution with objective value θ.

Exercise 6.35. Write an SDP formulation for τ and repeat the steps of the previous exercise.

Sum of eigenvalues

Largest (or smallest) eigenvalues are optima of SDPs. The same can be said about the sum of the k largest eigenvalues. In this section, assume $\lambda_j(\mathbf{X})$ stands for the jth largest eigenvalue of $\mathbf{X}$.

Theorem 6.36. *Let $\mathbf{X} \in \mathbb{S}^n$, $\mu \in \mathbb{R}$, $k \in [n]$. The following are equivalent.*

(1) $\sum_{j=1}^{k} \lambda_j(\mathbf{X}) \leq \mu$.

(2) There are $\mathbf{Y} \in \mathbb{S}^n_+$ and $\eta \in \mathbb{R}$ so that

$$\mu - k\eta \geq \operatorname{tr} \mathbf{Y} \quad \textit{and} \quad \mathbf{Y} - \mathbf{X} + \eta \mathbf{I} \succcurlyeq \mathbf{0}.$$

Proof. From (2) to (1), we have

$$\mathbf{Y} - \mathbf{X} \succcurlyeq -\eta \mathbf{I} \implies \lambda_j(\mathbf{Y}) - \lambda_j(\mathbf{X}) \geq -\eta.$$

(This implication is not trivial — due to Ky Fan.) Thus, immediately, $\sum_{j=1}^{k} \lambda_j(\mathbf{X}) \leq \mu$.

From (1) to (2), have $\eta = \lambda_k(\mathbf{X})$, and

$$\mathbf{Y} = \sum_{j=1}^{k} (\lambda_j(\mathbf{X}) - \eta) \mathbf{v}_j \mathbf{v}_j^{\mathsf{T}},$$

where $\mathbf{v}_j$ are the eigenvectors of $\mathbf{X}$ corresponding to λ_j. Thus $\mu - k\eta - \operatorname{tr}(\mathbf{Y}) \geq 0$, and $\mathbf{Y} - \mathbf{X} + \eta \mathbf{I} \succcurlyeq \mathbf{0}$. ☐

Exercise 6.37. Given $\mathbf{M}$, show that $\lambda_1(\mathbf{M}) + ... + \lambda_k(\mathbf{M})$ is equal to

$$\begin{aligned} \max \quad & \langle \mathbf{M}, \mathbf{X} \rangle \\ \text{subject to} \quad & \langle \mathbf{I}, \mathbf{X} \rangle = k \\ & \mathbf{X} \preccurlyeq \mathbf{I} \\ & \mathbf{X} \succcurlyeq \mathbf{0}. \end{aligned}$$

(Use weak duality!).

7 Theta

In this section, we focus on perhaps the most standard application of semidefinite program to graph theory.

Let $G = (V, E)$ be a graph on n vertices. Recall: a *coclique* (or stable set, or independent set) S is a subset of V that induces no edge from $E(G)$. A *clique* is the complement of a clique. A *colouring* of G is a partition of V into cocliques (having a colour given to each one of them). A *clique partition* is a partition of G into cliques. With this:

(i) $\alpha(G)$ is the size of the largest coclique in G, called the *independence number* or *coclique number*.

(ii) $\omega(G)$ is the size of the largest clique, called the *clique number*.

Note that $\alpha(G) = \omega(\overline{G})$.

(iii) $\chi(G)$ is the minimum number of colours needed to colour $V(G)$ with cocliques.

(iv) $\theta(G)$ is the minimum number of cliques needed to partition $V(G)$.

Note that $\theta(G) = \chi(\overline{G})$.

7.1 Theta and its dual

From the end of last section, we have defined the parameter

$$\begin{aligned} \vartheta(G) = \qquad & \max \quad \langle \mathbf{J}, \mathbf{X} \rangle \\ & \text{subject to} \quad \mathrm{X}_{ij} = 0 \; \forall \, ij \in E(G), \\ & \qquad \operatorname{tr} \mathbf{X} = 1, \\ & \qquad \mathbf{X} \succcurlyeq \mathbf{0}. \end{aligned}$$

It was easy to see that $\alpha(G) \leq \vartheta(G)$, because if $\mathbf{z}$ denote the characteristic vector of a coclique S, then $\mathbf{Z} = (1/|S|)\mathbf{z}\mathbf{z}^{\mathsf{T}}$ is feasible for the formulation above, with objective value equal to $|S|$.

From SDP duality, we also obtained that

$$\begin{aligned} \vartheta(G) = \qquad & \min \quad \lambda \\ & \text{subject to} \quad \mathrm{Y}_{ij} = 0 \; \forall \, ij \notin E(G) \\ & \qquad \lambda \mathbf{I} + \mathbf{Y} - \mathbf{J} \succcurlyeq \mathbf{0}. \end{aligned}$$

Our goal now is to introduce other alternative ways to look at this parameter $\vartheta(G)$. We would like to understand more precisely why $\vartheta(G) \leq \chi(\overline{G})$, for example, or better yet, what is the connection between ϑ and Hoffman lower bound for the chromatic number.

7.2 Largest eigenvalue

Theorem 7.1. *For any graph G, it follows that*

$$\vartheta(G) = \max\{\lambda_{\max}(\mathbf{B}) : \operatorname{diag}(\mathbf{B}) = \mathbf{1},\ \mathbf{B} \succcurlyeq \mathbf{0},\ \mathbf{B}_{ij} = 0\ \forall\, ij \in E(G)\}.$$

Proof. We will show that each feasible solution to the formulation above gives one for the max formulation of the ϑ parameter with the same objective value, and vice-versa.

Let $\mathbf{B}$ be feasible to the program above, and let $\mathbf{Y}$ be the projection onto an eigenline corresponding to $\lambda_{\max}(\mathbf{B})$. Then $\langle \mathbf{B}, \mathbf{Y}\rangle = \lambda_{\max}(\mathbf{B})$. On the other hand,

$$\langle \mathbf{B}, \mathbf{Y}\rangle = \langle \mathbf{B} \circ \mathbf{Y}, \mathbf{J}\rangle.$$

Now observe that $\mathbf{B} \circ \mathbf{Y}$ has trace equal to the trace of $\mathbf{Y}$, which is 1. It is also positive semidefinite (one of Schur's theorems), and from the definition of $\mathbf{B}$, we also have $(\mathbf{B}\circ\mathbf{Y})_{ij} = 0$ for all $ij \in E(G)$. Thus $\mathbf{B} \circ \mathbf{Y}$ os feasible for the max formulation of ϑ, with objective value $\lambda_{\max}(\mathbf{B})$.

Now consider an optimum solution in the max formulation of ϑ. If any of its diagonal entries is 0, make them equal to 1. Call this new matrix $\mathbf{X}$. Then, let $\mathbf{D}$ be the diagonal matrix that contains its diagonal. Look at

$$\mathbf{B} = \mathbf{D}^{-1/2}\mathbf{X}\mathbf{D}^{-1/2}.$$

It has constant diagonal equal to 1, it is positive-semidefinite, and it is so that $\mathbf{B}_{ij} = 0$ for all $ij \in E(G)$. If $\mathbf{d} = \operatorname{diag}(\mathbf{X}')$, then

$$\frac{\sqrt{\mathbf{d}}^{\mathsf{T}}\mathbf{B}\sqrt{\mathbf{d}}}{\sqrt{\mathbf{d}}^{\mathsf{T}}\sqrt{\mathbf{d}}} = \frac{\vartheta(G)}{1} = \vartheta(G).$$

□

With this result, we can show that the best possible Hoffman ratio is actually equal to $\vartheta(\overline{G})$ (and therefore that $\vartheta(\overline{G}) \leq \chi(G)$).

Corollary 7.2. *Given a matrix $\mathbf{A}$ with 0 diagonal, define*

$$\tilde{\mathbf{A}} = \begin{cases} [-\lambda_{\min}(\mathbf{A})]^{-1}\,\mathbf{A} & \text{if } \mathbf{A} \neq 0 \\ \mathbf{A} & \text{if } \mathbf{A} = 0 \end{cases}.$$

Given G, we have

$$\vartheta(\overline{G}) = \max\{1 + \lambda_{\max}(\tilde{\mathbf{A}}) : \mathrm{A}_{ij} = 0\ \forall\ ij \notin E(G).\}.$$

Proof. From the Theorem, we have

$$\vartheta(\overline{G}) = \max\{\lambda_{\max}(\mathbf{I} + \mathbf{A}) : \mathbf{A} \succcurlyeq -\mathbf{I}, \mathrm{A}_{ij} = 0\ \forall\ ij \notin E(G)\}.$$

For any $\mathbf{A}$ feasible, just note that the best scaling so that $\succcurlyeq -\mathbf{I}$ is maintained is always obtained from dividing $\mathbf{A}$ by $-\lambda_{\min}$. Hence the result follows.

□

7.3 Conic optimization for cocliques and colourings

In this section, we will see how to formulate the problems of finding α and χ as conic programs. Second, we will introduce more parameters between α and χ. They will be motivated by the well known vector colourings.

A matrix $\mathbf{M}$ is called *completely positive* if, for some k, there are nonnegative vectors $\mathbf{x}_1, ..., \mathbf{x}_k \in \mathbb{R}^n_+$ so that

$$\mathbf{M} = \mathbf{x}_1\mathbf{x}_1^\mathsf{T} + \ldots + \mathbf{x}_k\mathbf{x}_k^\mathsf{T}$$

Clearly, completely positive matrices are positive semidefinite, but the converse does not hold.

Exercise 7.3. Why not?

We will denote

$$\mathbb{S}^n_{\text{clyp}} = \{\mathbf{M} \in \mathbb{S}^n : \mathbf{M} \text{ is completely positive.}\}$$

The completely positive rank of a completely positive matrix $\mathbf{M}$ is the minimum k so that $\mathbf{M}$ writes as a sum of k nonnegative rank-1 projectors, just as above.

Exercise 7.4. If $\mathbf{M}$ is a $n \times n$ completely positive, show that its completely positive rank is upper bounded by $\binom{n+1}{2}$. (Hint: this is actually the dimension of the space of symmetric $n \times n$ matrices.)

A symmetric matrix $\mathbf{M}$ is called *copositive* if, for all $\mathbf{x} \in \mathbb{R}^n_+$, we have

$$\mathbf{x}^\mathsf{T}\mathbf{M}\mathbf{x} \geq 0.$$

Clearly, all positive semidefinite matrices are copositive, but the converse does not hold.

Exercise 7.5. Why not?

We will denote

$$\mathbb{S}^n_{\text{copo}} = \{\mathbf{M} \in \mathbb{S}^n : \mathbf{M} \text{ is copositive.}\}$$

Theorem 7.6. *The sets $\mathbb{S}^n_{copo}$ and $\mathbb{S}^n_{clyp}$ are closed, convex conves, and duals of each other.*

Proof. I leave it as an exercise to show that both sets are convex cones. If $\mathbf{M} \notin \mathbb{S}^n_{\text{copo}}$, then there is $\mathbf{x} \geq \mathbf{0}$ with $\mathbf{x}^\mathsf{T}\mathbf{M}\mathbf{x} < 0$. Any small variation around $\mathbf{M}$ will not change this fact, thus $\mathbb{S}^n_{\text{copo}}$ is closed. To see that $\mathbb{S}^n_{\text{clyp}}$ is closed, we will show that the limit of any convergent sequence in it belongs to it. Let $(\mathbf{M}^{(k)})_{k \geq 0} \in \mathbb{S}^n_{\text{clyp}}$ be a convergent sequence, converging to $\mathbf{M}$. Let $\mathbf{A}^{(k)} \geq \mathbf{0}$ be so that

$$\mathbf{M}^{(k)} = \mathbf{A}^{(k)}(\mathbf{A}^{(k)})^\mathsf{T}.$$

Thus the ith row of $\mathbf{A}^{(k)}$, for $k \to \infty$, form a bounded sequence of vectors. Thus this sequence has a convergent subsequence, say to $\mathbf{a}_i$, with $\mathbf{a}_i \geq \mathbf{0}$. If $\mathbf{A}$ has these as its columns, it follows that

$$\mathbf{M} = \mathbf{A}\mathbf{A}^\mathsf{T}.$$

Finally, if $\mathbf{M} \in \mathbb{S}^n_{\text{copo}}$, and $\mathbf{N} = \sum_{i=1}^k \mathbf{x}_k \mathbf{x}_k^\mathsf{T}$ with $\mathbf{x}_i \geq \mathbf{0}$, then

$$\langle \mathbf{M}, \mathbf{N} \rangle = \sum_{i=1}^{k} \langle \mathbf{M}, \mathbf{x}_i \mathbf{x}_i^\mathsf{T} \rangle \geq 0,$$

hence $\mathbf{M} \in (\mathbb{S}^n_{\text{clyp}})^*$. For the converse, if $\mathbf{M} \notin \mathbb{S}^n_{\text{copo}}$, there is $\mathbf{x} \geq \mathbf{0}$ with $\mathbf{x}^\mathsf{T}\mathbf{M}\mathbf{x} < 0$. Consequently, $\mathbf{M} \notin (\mathbb{S}^n_{\text{clyp}})^*$. Because they are closed convex cones, we also get

$$(\mathbb{S}^n_{\text{copo}})^* = \mathbb{S}^n_{\text{clyp}}.$$

□

Exercise 7.7. Describe as best as you can the cones $\mathbb{S}^2_{\text{copo}}$ and $\mathbb{S}^2_{\text{clyp}}$.

7.4 Conic program for the independence number

We now consider the following primal-dual pair of conic programs.

$$\text{(P)} \quad \begin{array}{rl} \max & \langle \mathbf{J}, \mathbf{X} \rangle \\ \text{s.t.} & \mathrm{X}_{ij} = 0 \ \forall\ ij \in E(G) \\ & \operatorname{tr} \mathbf{X} = 1 \\ & \mathbf{X} \in \mathbb{S}^n_{\text{clyp}} \end{array} \qquad \Bigg| \qquad \text{(D)} \quad \begin{array}{rl} \min & \lambda \\ \text{s.t.} & \lambda \mathbf{I} - \mathbf{J} - \mathbf{Y} \in \mathbb{S}^n_{\text{copo}} \\ & \mathrm{Y}_{ij} = 0 \ \forall\ ij \notin E(G). \end{array}$$

We will now show that the optimum of both programs above is $\alpha(G)$.

Lemma 7.8. *The optimum of (P) is $\geq \alpha(G)$.*

Proof. Let S be a maximum sized coclique, with characteristic vector $\mathbf{x}$. Simply note now that

$$\mathbf{X} = \frac{1}{\alpha(G)} \mathbf{x}\mathbf{x}^\mathsf{T}$$

is feasible for (P), with objective value $\alpha(G)$. □

It remains to show that the optimum of (D) is $\leq \alpha(G)$. To achieve this, we will use an influential result due to Motzkin and Straus.

Theorem 7.9. *For any graph G,*

$$\frac{1}{\alpha(G)} = \min\{\mathbf{x}^\mathsf{T}(\mathbf{A} + \mathbf{I})\mathbf{x} : \mathbf{1}^\mathsf{T}\mathbf{x} = 1,\ \mathbf{x} \geq \mathbf{0}\}.$$

Proof. Let $f(G)$ be the optimum of this program. If S is a coclique, with characteristic vector $\mathbf{x}$, then $(1/|S|)\mathbf{x}$ is feasible, and

$$\frac{1}{|S|^2} \mathbf{x}^\mathsf{T}(\mathbf{A} + \mathbf{I})\mathbf{x} = \frac{1}{|S|},$$

thus $f(G) \leq \alpha^{-1}$.

The other direction is more involved, and goes by induction on $n = |V(G)|$. If $n = 1$, the result is trivial. Assume now $f(H) \geq \alpha(G)^{-1}$ for all proper induced subgraphs H of G. Let $\mathbf{y}$ be an optimum giving $f(G)$. If one entry $\mathrm{y}_i = 0$, then we are lucky, and setting $H = G - i$, and $\overline{\mathbf{y}}$ the restriction of $\mathbf{y}$ to H, we have

$$\alpha(G)^{-1} = \alpha(H)^{-1} \leq f(H) \leq \overline{\mathbf{y}}^{\mathsf{T}}(\mathbf{A}(H) + \mathbf{I})\overline{\mathbf{y}} = \mathbf{y}^{\mathsf{T}}(\mathbf{A}(G) + \mathbf{I})\mathbf{y} = f(G).$$

Let $\mathbf{y}$ be a minimizer for $f(G)$, and $\mathbf{y} > \mathbf{0}$. Pick $ij \in E(G)$ and define $\mathbf{z}$ with $\mathrm{z}_i = \mathrm{y}_i + \varepsilon$, $\mathrm{z}_j = \mathrm{y}_j - \varepsilon$, and otherwise $\mathbf{z}_\ell = \mathbf{y}_\ell$. Clearly

$$\mathbf{z}^{\mathsf{T}}(\mathbf{A}+\mathbf{I})\mathbf{z} = f(G) + \varepsilon\big[\mathbf{y}^{\mathsf{T}}(\mathbf{A}+\mathbf{I})(\mathbf{e}_i - \mathbf{e}_j) + (\mathbf{e}_i - \mathbf{e}_j)^{\mathsf{T}}(\mathbf{A}+\mathbf{I})\mathbf{y}\big] + \varepsilon^2(\mathbf{e}_i - \mathbf{e}_j)^{\mathsf{T}}(\mathbf{A}+\mathbf{I})(\mathbf{e}_i - \mathbf{e}_j)$$

The term on ε^2 vanishes, as $ij \in E(G)$. The term on ε vanishes because $\mathbf{y}$ is a minimizer. Thus, we can choose ε so that $\mathbf{z}$ becomes 0 at a coordinate, and apply induction on the graph with the vertex corresponding to this coordinate removed. □

Exercise 7.10. Show that

$$\max\{\mathbf{x}^{\mathsf{T}}\mathbf{A}\mathbf{x} : \mathbf{1}^{\mathsf{T}}\mathbf{x} = 1,\ \mathbf{x} \geq \mathbf{0}\} = \frac{1}{2}\left(1 - \frac{1}{\omega(G)}\right).$$

Exercise 7.11. In the program

$$\begin{array}{rl} \min & \lambda \\ \text{subject to} & \lambda\mathbf{I} - \mathbf{J} - \mathbf{Y} \in \mathbb{S}^n_{\text{copo}} \\ & \mathrm{Y}_{ij} = 0\ \forall\ ij \notin E(G), \end{array}$$

argue why there is an optimum solution with $\mathbf{Y}$ having all entries corresponding to edges equal to a constant.

Theorem 7.12. *The optimum of (D) is at most $\alpha(G)$.*

Proof. From the exercise above, we need to show that

$$\alpha(G) \geq \min\{\lambda : \lambda\mathbf{I} - \mathbf{J} - z\mathbf{A} \in \mathbb{S}^n_{\text{copo}},\ \lambda, z \in \mathbb{R}\}.$$

From Motzkin Straus, it follows that, for all $\mathbf{x}$ with $\mathbf{1}^{\mathsf{T}}\mathbf{x} = 1$ and $\mathbf{x} \geq \mathbf{0}$, we have

$$\alpha(G)\mathbf{x}^{\mathsf{T}}(\mathbf{A} + \mathbf{I})\mathbf{x} \geq 1 = \mathbf{x}^{\mathsf{T}}\mathbf{J}\mathbf{x}.$$

Therefore, for all $\mathbf{x} \geq \mathbf{0}$, we have

$$\mathbf{x}^{\mathsf{T}}(\alpha\mathbf{A} + \alpha\mathbf{I} - \mathbf{J})\mathbf{x} \geq 0.$$

Thus the matrix $(\alpha\mathbf{A} + \alpha\mathbf{I} - \mathbf{J})$ is copositive, and thus a feasible solution to the program above, with value $\alpha(G)$. □

Exercise 7.13. Show that

$$\alpha(G)^{-1} = \min\{\langle \mathbf{A} + \mathbf{I}, \mathbf{X}\rangle : \langle \mathbf{J}, \mathbf{X}\rangle = 1,\ \mathbf{X} \in \mathbb{S}^n_{\text{clyp}}\}.$$

7.5 Conic programming for the fractional chromatic number

We now consider the following primal-dual pair of conic programs.

$$
\text{(P)}\quad \begin{array}{rl} \max & \langle \mathbf{J}, \mathbf{X} \rangle \\ \text{s.t.} & \mathrm{X}_{ij} + \mathrm{Z}_{ij} = 0 \ \forall\, ij \in E(G) \\ & \mathrm{tr}(\mathbf{X} + \mathbf{Z}) = 1 \\ & \mathbf{X} \in \mathbb{S}_+^n,\ \mathbf{Z} \in \mathbb{S}_{\text{copo}}^n \end{array}
\qquad\Bigg|\qquad
\text{(D)}\quad \begin{array}{rl} \min & \lambda \\ \text{s.t.} & \lambda \mathbf{I} - \mathbf{J} - \mathbf{Y} \in \mathbb{S}_+ \\ & \lambda \mathbf{I} - \mathbf{Y} \in \mathbb{S}_{\text{clyp}} \\ & \mathrm{Y}_{ij} = 0 \ \forall\, ij \notin E(G). \end{array}
$$

We will now show that the optimum of both programs above is $\chi_f(\overline{G})$.

Lemma 7.14. *The optimum of (D) is $\leq \chi_f(\overline{G})$.*

Proof. Assume $\mathbf{y}$ is an optimum for

$$\min\{\mathbf{1}^T\mathbf{y} : \mathbf{M}\mathbf{y} = \mathbf{1},\ \mathbf{y} \geq \mathbf{0}\}.$$

Recalling that $\overline{\chi}_f = \mathbf{1}^\mathsf{T}\mathbf{y}$, consider the matrix

$$\mathbf{W} = \overline{\chi}_f \sum_i \mathrm{y}_i \mathbf{x}_{K_i}\mathbf{x}_{K_i}^\mathsf{T} = \sum_{i,j} \mathrm{y}_i\mathrm{y}_j \mathbf{x}_{K_i}\mathbf{x}_{K_i}^\mathsf{T},$$

where the sums run over all cliques of the graph. This matrix is clearly completely positive, and, moreover, its diagonal is constant equal to $\overline{\chi}_f$. We also have $\mathrm{W}_{ij} = 0$ for all $ij \in E(\overline{G})$. So it remains to show that $\mathbf{W} - \mathbf{J} \succcurlyeq \mathbf{0}$. Recalling that $\mathbf{M}\mathbf{y} = \mathbf{1}$, it follows that

$$\mathbf{W} - \mathbf{J} = \sum_{i,j} \mathrm{y}_i\mathrm{y}_j \mathbf{x}_{K_i}\mathbf{x}_{K_i}^\mathsf{T} - \sum_{i,j} \mathrm{y}_i\mathrm{y}_j \mathbf{x}_{K_i}\mathbf{x}_{K_j}^\mathsf{T}.$$

Thus, given any $\mathbf{x} \in \mathbb{R}^n$, and having $a_i = \mathbf{x}_{K_i}^\mathsf{T}\mathbf{x}$, it follows that

$$\mathbf{x}^\mathsf{T}(\mathbf{W} - \mathbf{J})\mathbf{x} = \sum_{i,j} \mathrm{y}_i\mathrm{y}_j(a_i^2 - a_i a_j) = \sum_{i<j} \mathrm{y}_i\mathrm{y}_j(a_i^2 - 2a_i a_j + a_j^2) \geq 0.$$

□

Lemma 7.15. *The optimum of (D) is $\geq \chi_f(\overline{G})$.*

Proof. Consider the dual formulation for $\overline{\chi}_f$, and let $\mathbf{x}$ be an optimum solution, that is, $\mathbf{x} \geq \mathbf{0}$, $\mathbf{M}^\mathsf{T}\mathbf{x} \leq \mathbf{1}$, and

$$\overline{\chi}_f = \mathbf{1}^\mathsf{T}\mathbf{x}.$$

Let $\lambda, \mathbf{Y}$ be optimum solution for (D). Note that

$$\lambda\mathbf{I} - \mathbf{Y} = \sum_{i=1}^k \mathbf{z}_i\mathbf{z}_i^\mathsf{T},$$

for some $\mathbf{z}_i \geq \mathbf{0}$. Thus, each $\mathbf{z}_i$ is supported in a clique of the graph (because $\mathrm{Y}_{ij} = 0$ for all non edges). As $\lambda\mathbf{I} - \mathbf{Y} \succcurlyeq \mathbf{J}$, by applying $\mathbf{x}$ on both sides, we obtain

$$\chi_f^2 \leq \sum_{i=1}^k (\mathbf{x}^\mathsf{T}\mathbf{z}_i)^2.$$

Assume $\mathbf{z}_i$ is supported in the clique C. Let $\mathbf{x}_C$ be the vector obtained form $\mathbf{x}$ upon making all entries outside of C to be zero. Note that $\mathbf{1}^\mathsf{T}\mathbf{x}_C \leq 1$. We now write $\mathbf{x}^\mathsf{T}\mathbf{z}_i = \sqrt{\mathbf{x}_C}^\mathsf{T}(\sqrt{\mathbf{x}} \circ \mathbf{z}_i)$, and apply Cauchy-Schwarz, to obtain

$$(\mathbf{x}^\mathsf{T}\mathbf{z}_i)^2 \leq (\mathbf{1}^\mathsf{T}\mathbf{x}_C) \cdot [\mathbf{x}^\mathsf{T}(\mathbf{z}_i \circ \mathbf{z}_i)] \leq \mathbf{x}^\mathsf{T}(\mathbf{z}_i \circ \mathbf{z}_i).$$

Thus

$$\chi_f^2 \leq \sum_{i=1}^{k} (\mathbf{x}^\mathsf{T}\mathbf{z}_i)^2 \leq \mathbf{x}^\mathsf{T}\left(\sum_{i=1}^{k} (\mathbf{z}_i \circ \mathbf{z}_i)\right) = \mathbf{x}^\mathsf{T}(\lambda\mathbf{1}) = \lambda \cdot \chi_f.$$

□

Exercise 7.16. Find a feasible solution for (P) with objective value $\chi_f(\overline{G})$.

7.6 Vector colourings and theta variants

A graph G has an γ-vector colouring if there is an assignment of unit vectors $\{\mathbf{v}_a : a \in V(G)\} \subseteq \mathbb{R}^d$ so that, if $a \sim b$, then

$$\langle \mathbf{a}, \mathbf{b} \rangle \leq \frac{1}{1-\gamma}.$$

The smallest $\gamma > 1$ so that G has an γ-vector colouring is called the *vector chromatic number* of G, to be denoted by $\chi_{\text{vec}}(G)$ (if G has no edge, this is defined to be equal to 1). The existence of an γ-vector colouring is equivalent to the existence of a matrix $\mathbf{W}$ so that

(i) $\mathbf{W} \succcurlyeq \mathbf{0}$.

(ii) $\operatorname{diag} \mathbf{W} = \mathbf{1}$.

(iii) $W_{ij} \leq 1/(1-\gamma)$ for all $ij \in E(G)$.

The existence of such $\mathbf{W}$ is equivalent to the existence of $\mathbf{Y}$ so that

$$(\gamma - 1)\mathbf{W} = \gamma\mathbf{I} - \mathbf{Y} - \mathbf{J},$$

where $Y_{ij} \geq 0$ if $ij \in E(G)$, and $Y_{ii} = 0$. Given that $\gamma > 1$, it follows that the vector chromatic number of G is the optimum of the SDP

$$\begin{aligned} \min \quad & \gamma \\ \text{subject to} \quad & \gamma\mathbf{I} - \mathbf{Y} - \mathbf{J} \succcurlyeq \mathbf{0} \\ & Y_{ij} \geq 0 \quad \text{for all } ij \notin E(\overline{G}). \end{aligned}$$

This formulation is dual to

$$\begin{aligned} \max \quad & \langle \mathbf{J}, \mathbf{X} \rangle \\ \text{subject to} \quad & \operatorname{tr} \mathbf{X} = 1 \\ & X_{ij} = 0 \quad \text{for all } ij \in E(\overline{G}) \\ & \mathbf{X} \geq \mathbf{0},\ \mathbf{X} \succcurlyeq \mathbf{0}. \end{aligned}$$

Given both formulations above, and recalling the formulation for $\gamma(G)$ as a conic program over $\mathbb{S}^n_{\text{clyp}}$, which in particular belongs to the set of matrices which are non-negative and positive semidefinite, the corollary below follows immediately:

Corollary 7.17. *For any graph G,*

$$\omega(G) \leq \chi_{vec}(G) \leq \vartheta(\overline{G}).$$

Note that $\chi_{\text{vec}}(G)$ can be computed efficiently.

Exercise 7.18. Prove in an elementary way that $\omega(G) \leq \chi_{\text{vec}}(G)$. Hint: consider an optimal γ-vector colouring, and sum the vectors corresponding to a maximum clique. Then consider the norm of this resulting vector.

If we had enforced that the vector colouring satisfies inner product inequality with an equality, we would have arrived at what is known as the *strict vector chromatic number*. This is denoted by $\chi_{\text{svec}}(G)$, and formally defined as the lest $\gamma > 1$ so that there is an assignment of unit vectors $\{\mathbf{v}_a : a \in V(G)\} \subseteq \mathbb{R}^d$ so that, if $a \sim b$, then

$$\langle \mathbf{a}, \mathbf{b} \rangle = \frac{1}{1-\gamma}.$$

Exercise 7.19. Verify that $\chi_{\text{svec}}(G) = \vartheta(\overline{G})$. Hint: look at the SDP formulations above.

An γ-strict vector colouring which further satisfies

$$\langle \mathbf{a}, \mathbf{b} \rangle \geq \frac{1}{1-\gamma}.$$

for all $a, b \in V(G)$ is called a *rigid vector colouring*. The least γ so that G has a rigid vector colouring is denoted by $\chi_{\text{rvec}}(G)$. Note that

$$\chi_{\text{vec}}(G) \leq \chi_{\text{svec}}(G) \leq \chi_{\text{rvec}}(G),$$

as these are all minimization problems with increasing degrees of restrictions. Moreover, it is immediate to verify that $\chi_{\text{rvec}}(G)$ is defined by the following pair of primal-dual SDPs:

$$\text{(P)} \quad \begin{array}{rl} \min & \gamma \\ \text{s.t.} & \gamma\mathbf{I} - \mathbf{Y} - \mathbf{J} \succcurlyeq \mathbf{0} \\ & Y_{ij} = 0 \quad \text{for all } ij \notin E(\overline{G}) \\ & \mathbf{Y} \leq \mathbf{0}. \end{array} \qquad \Bigg| \qquad \text{(D)} \quad \begin{array}{rl} \max & \langle \mathbf{J}, \mathbf{X} \rangle \\ \text{s.t.} & \operatorname{tr} \mathbf{X} = 1 \\ & X_{ij} \leq 0 \quad \text{for all } ij \in E(\overline{G}) \\ & \mathbf{X} \succcurlyeq \mathbf{0}. \end{array}$$

Exercise 7.20. Check the definition of $\chi_f(\overline{G})$ as a conic program, and prove that $\chi_f(G) \geq \chi_{\text{rvec}}(G)$.

Historically, the vector chromatic number was introduced as a variant of ϑ by Schrijver, and it was originally denoted by ϑ' (and called, later, by Schrijver's Theta). The parameter χ_{rvec} was introduced by Szegedy also as a ϑ variant. If one decides to denote $\chi_{\text{vec}}(G) = \vartheta^-(\overline{G})$ and $\chi_{\text{rvec}}(G) = \vartheta^+(\overline{G})$, the inequalities seen above take the pleasant form

$$\gamma(G) \leq \vartheta^-(G) \leq \vartheta(G) \leq \vartheta^+(G) \leq \chi_f(\overline{G}).$$

8 The Laplacian matrix

8.1 Basics

Let G be a graph, and define $\mathbf{D}(G)$ to be the diagonal matrix whose entries correspond to the degrees of the vertices of G. Define the Laplacian matrix of G by

$$\mathbf{L} = \mathbf{L}(G) = \mathbf{D}(G) - \mathbf{A}(G).$$

Theorem 8.1. *The Laplacian matrix is positive semidefinite. Moreover, the multiplicity of* 0 *as an eigenvalue of* $\mathbf{L}$ *is equal to the number of connected components of* G*.*

Proof. To see this, assume G has been oriented, meaning, each edge has been assigned a direction, thus becoming an arc. Let $\mathbf{N}$ be the corresponding vertex by arc incidence matrix, so that an entry is 0 if the arc does not touch the vertex, $+1$ if the vertex is the head of the arc, and -1 if it is the tail. It is immediate to see that

$$\mathbf{L} = \mathbf{N}\mathbf{N}^T.$$

(Note that this does not depend on the choice for the orientation.)

Following, $\mathbf{N}^T\mathbf{v} = 0$ if and only if $\mathbf{L}\mathbf{v} = 0$. It is immediate to see that $\mathbf{N}^T\mathbf{v} = 0$ if and only if $\mathbf{v}$ is constant on each connected component of G, whence the result follows (and describes essentially the unique eigenvector for 0 in a connected graph — the constant vector). □

Exercise 8.2. Assume G is regular, and let $\theta_1 \geq ... \geq \theta_n$ be the eigenvalues of $\mathbf{A}(G)$, with corresponding eigenbasis $\{\mathbf{v}_1, ..., \mathbf{v}_n\}$. Find an expression of the eigenvalues of $\mathbf{L}(G)$, and find a corresponding eigenbasis.

Exercise 8.3. Let $0 = \lambda_1 \leq ... \leq \lambda_n$ be the eigenvalues of $\mathbf{L}(G)$. Find the eigenvalues of $\mathbf{L}(\overline{G})$. Use this exercises to find the eigenvalues of $\mathbf{L}(K_{n,m})$ (this is the complete bipartite graph with n vertices on one side and m on the other).

As we have seen, $\mathbf{L}(G)$ is positive semidefinite. It follows that $\mathbf{x}^T\mathbf{L}\mathbf{x} \geq 0$ for all $\mathbf{x} \in \mathbb{R}^n$. We can moreover find a useful and meaningful expression for this. As $\mathbf{L} = \mathbf{N}\mathbf{N}^T$ where $\mathbf{N}$ is the incidence matrix of an orientation of the graph, it follows that

$$\mathbf{x}^T\mathbf{L}\mathbf{x} = (\mathbf{N}^T\mathbf{x})^T(\mathbf{N}^T\mathbf{x}) = \sum_{uv \in E(G)} (\mathbf{x}_u - \mathbf{x}_v)^2.$$

Exercise 8.4. Assume G is connected, on n vertices, and let λ_2 be its second smallest Laplacian eigenvalue. We certainly know (from the minimax principle for eigenvalues) that

$$\lambda_2 = \min_{\mathbf{v} \perp \mathbf{1}} \frac{\sum_{ab \in E(G)} (\mathbf{v}_a - \mathbf{v}_b)^2}{\sum_{a \in V(G)} \mathbf{v}_a^2}.$$

What I want you to show is that

$$\lambda_2 = \min_{\mathbf{v} \neq \gamma \mathbf{1}} \frac{n \sum_{ab \in E(G)} (\mathbf{v}_a - \mathbf{v}_b)^2}{\sum_{a<b} (\mathbf{v}_a - \mathbf{v}_b)^2}.$$

(The minimum is simply being taken over all vectors which are just not constant.)

Also, and without much difficulty now, prove that

$$\lambda_n = \max_{\mathbf{v} \neq \alpha \mathbf{1}} \frac{n \sum_{ab \in E(G)} (\mathbf{v}_a - \mathbf{v}_b)^2}{\sum_{a<b} (\mathbf{v}_a - \mathbf{v}_b)^2}.$$

Exercise 8.5. Can you prove lower and upper bounds to the eigenvalues os $\mathbf{L}$ with regards to the number of vertices, edges, and minimum, average and maximum degrees of the graph?

8.2 Trees

A spanning tree of a connected graph G on n vertices is a subset of its edges that connects all vertices without forming any cycle. Necessarily, any spanning tree will contain $n-1$ edges.

A first result we shall see about the Laplacian matrix is actually a quite surprising one: we can efficiently count how many spanning trees any graph has.

Let $\tau(G)$ denote the number of spanning trees of the graph G. Recall the notation for edge deletion and contraction: $G \backslash e$ is the graph G with e removed, and G/e is the graph G with e removed and its incident vertices identified.

Lemma 8.6. *For any graph G and edge e, we have*

$$\tau(G) = \tau(G \backslash e) + \tau(G/e).$$

Exercise 8.7. Why?

We can now state the Matrix-Tree Theorem (due to Kirchkoff).

Theorem 8.8. *Let G be a graph, Laplacian $\mathbf{L}$. Let $a \in V(G)$, and $\mathbf{L}[a]$ denote the submatrix of $\mathbf{L}$ obtained upon deleting row and column corresponding to a. Then*

$$\tau(G) = \det \mathbf{L}[a].$$

Proof. This will be a proof by induction on the number of edges. You should check a few base cases on your own. Let us now assume G has m edges, and the result holds for any graph on fewer edges. Let $e \in E(G)$, with $e = \{a, b\}$. In G/e, vertices a and b are identified — let c be the name they receive in this case. If we show that

$$\det \mathbf{L}(G)[a] = \det \mathbf{L}(G \backslash e)[a] + \det \mathbf{L}(G/e)[c],$$

then, by induction and the lemma above, we will be done. So this equality above is now our task. In computing $\det \mathbf{L}(G)[a]$, we will perform row expansion in the row corresponding to b. Note that all terms of this expansion coming from an off-diagonal position will appear exactly the same in $\det \mathbf{L}(G \backslash e)[a]$. The only problem is the diagonal position — it is one unit larger in $\mathbf{L}(G)[a]$ then in $\mathbf{L}(G \backslash e)[a]$. Now the submatrix corresponding to excluding row and column b from $\mathbf{L}(G \backslash e)[a]$ is precisely $\mathbf{L}(G/e)[c]$, that is

$$\det \mathbf{L}(G)[a] = \det \mathbf{L}(G \backslash e)[a] + \det \mathbf{L}(G \backslash e)[a, b] = \det \mathbf{L}(G \backslash e)[a] + \det \mathbf{L}(G/e)[c],$$

as wished. □

Exercise 8.9. Very easily now you can verify that the number of spanning trees of K_n vertices is n^{n-2}.

Exercise 8.10. Prove that the number of spanning trees of G that contain a given edge $e = ab$ is equal to $\det \mathbf{L}[a, b]$.

As we have all learned before, for any square matrix $\mathbf{M}$,

$$\mathbf{M}\,\mathrm{adj}(\mathbf{M}) = \det(\mathbf{M})\mathbf{I},$$

where $\mathrm{adj}(\mathbf{M})_{ij} = (-1)^{i+j} \det \mathbf{M}(j, i)$. As we have just seen from above, all diagonal entries of $\mathrm{adj}\,\mathbf{L}(G)$ are equal to $\tau(G)$.

However for any G, $\det \mathbf{L}(G) = 0$. If we now assume G is connected, we know that there is essentially only one eigenvector to the eigenvalue 0, thus the equality

$$\mathbf{L}(G) \cdot \mathrm{adj}\,\mathbf{L}(G) = \mathbf{0}$$

implies that all columns of $\mathrm{adj}\,\mathbf{L}(G)$ are constant, and therefore all entries of $\mathrm{adj}\,\mathbf{L}(G)$ are equal to $\tau(G)$. It is immediate to verify all comments above hold if G is disconnected, in which case $\tau(G) = 0$.

Corollary 8.11. *For any graph G, we have*

$$\mathrm{adj}\,\mathbf{L}(G) = \tau(G)\mathbf{J}.$$

□

Exercise 8.12. Prove that for any graph G with Laplacian eigenvalues $\lambda_1 \leq ... \leq \lambda_n$, it holds that

$$\tau(G) = \frac{1}{n} \prod_{i=2}^{n} \lambda_i.$$

Hint: let $\psi(x)$ be the characteristic polynomial of $\mathbf{L}$. Arrive at the result realizing that

$$\prod_{i=1}^{n} (x - \lambda_i) = \psi(x) = \det(x\mathbf{I} - \mathbf{L}).$$

8.3 Representation, springs and energy

A *representation of a graph* is a map $\rho : V(G) \to \mathbb{R}^m$ (you can think of it as an m-dimensional drawing of G). You can associate ρ to an $n \times m$ matrix $\mathbf{R}$ — each row is the image of the corresponding vertex. A representation ρ is called balanced if $\sum_{a \in V} \rho(a) = \mathbf{0}$, that is, if $\mathbf{1}^T\mathbf{R} = \mathbf{0}$. Upon assuming we can freely translate a representation, we can always assume it is balanced. Moreover, we shall also assume the columns of $\mathbf{R}$ are linearly independent (otherwise we simply look at the representation to the subspace of $\mathbb{R}^m$ and rewrite ρ upon a change of basis so that $\mathbf{R}$ has fewer columns).

Now imagine a physical model, in which the vertices have been placed in $\mathbb{R}^m$. Some of them, $U \subseteq V(G)$, are "nailed", some of them, $V(G) - U$, are free. The edges are *springs*. For now, identical springs, with spring constant 1. By Hooke's law, the force the spring between a and b exerts in a is equal to $\rho(b) - \rho(a)$. Note that a configuration is in equilibrium if and only if the net force at each vertex in $V - U$ is 0. This is equivalent to requiring, for all $a \in V - U$, that

$$\sum_{b \sim a} \rho(b) - \rho(a) = 0 \iff \deg(a)\rho(a) - \sum_{b \sim a} \rho(b) = \mathbf{0}.$$

In other words, **LR** must have a rectangle of 0s in the rows corresponding to the vertices in $V - U$. Once the entries of **R** corresponding to vertices in U have been determined, finding the remaining entries of **R** so that this holds is equivalent to solving m systems of equation whose coefficient matrix is $\mathbf{L}[U]$. All these systems have unique solutions if the graph is connected and $U \neq$, because $\mathbf{L}[U]$ is positive definite.

Exercise 8.13. Let $\mathbf{L}[U]$ denote the submatrix of **L** obtained upon removing rows and columns corresponding to the vertices in subset U. Assume the graph is connected, and U non-empty. Prove that all eigenvalues of $\mathbf{L}[U]$ are positive.

Exercise 8.14. Convince yourself that nothing really changes if we assume the spring between a and b to have spring constant ω_{ab}.

Physics also teaches us that vertices will settle in the position the minimizes the potential energy. The potential energy of a spring with constant ω and stretched to a length ℓ is $(1/2)\omega\ell^2$ (we will ignore the fraction). Thus, the potential energy of a configuration is

$$\mathcal{E}(\rho) = \sum_{ab \in E(G)} \omega_{ab} ||\rho(a) - \rho(b)||^2.$$

Let W be a diagonal matrix, indexed by $E(G)$, whose diagonal entry is equal to ω_{ab}. As before, let **N** be the incidence matrix of an orientation of G, and **R** the matrix of the representation. It is immediate to verify that

$$\mathcal{E}(\rho) = \operatorname{tr} \mathbf{R}^T \mathbf{N} \mathbf{W} \mathbf{N}^T \mathbf{R}.$$

Note that $\mathbf{N}\mathbf{W}\mathbf{N}^T$ is simply a weighted Laplacian (and if **W** has positive diagonal and the graph is connected, then $\mathbf{N}\mathbf{W}\mathbf{N}^T$ is positive semidefinite, and 0 is a simple eigenvalue with eigenvector $\mathbf{1}$).

A representation is called *orthogonal* if the columns of **R** are orthonormal. In this case, $\mathbf{R}^T\mathbf{R} = \mathbf{I}$. Requiring a representation to be orthogonal is also a way of imposing a shape. We do not need to "nail" vertices in this case, as the following theorem shows.

Theorem 8.15. *Let G be a graph, with a weighted Laplacian matrix $\mathcal{L}$, with eigenvalues $0 = \lambda_1 < \lambda_2 \leq \lambda_3 \leq ... \leq \lambda_n$. The minimum energy of a balanced orthogonal representation into $\mathbb{R}^m$ is equal to*

$$\sum_{r=2}^{m+1} \lambda_r.$$

Proof. To any orthogonal representation into $\mathbf{R}^k$ whose first column is $\mathbf{1}$ corresponds a balanced orthogonal representation in $\mathbb{R}^{k-1}$ with the same energy, obtained upon ignoring this first column. Thus the minimum energy of a balanced orthogonal representation into $\mathbb{R}^m$ is equal to the minimum energy of an orthogonal representation into $\mathbb{R}^{m+1}$ whose first column is $\mathbf{1}$. Let $\mathbf{R}$ be the matrix of one such representation. Its energy is

$$\operatorname{tr} \mathbf{R}^T \mathbf{L} \mathbf{R},$$

which, by interlacing, is at least $\sum_{r=1}^{m+1} \lambda_r$. Recall that $\lambda_1 = 1$. Moreover, one representation meeting this energy exists: simply write the eigenvectors corresponding $\lambda_1, ..., \lambda_{m+1}$ as the columns of $\mathbf{R}$. □

An immediate consequence is a method to draw graphs in $\mathbb{R}^m$ that is balanced and will somehow look "rigid" and "having a volume" — the so called spring embedding. Just pick the eigenvectors of $\mathbf{L}$ corresponding to $\lambda_2, ..., \lambda_{m+1}$, line them up as columns of a matrix, and map each vertex to the corresponding row. You should try this method to draw graphs in $\mathbb{R}^2$ and $\mathbb{R}^3$ using your favourite software.

8.4 Electrical currents

We define a physical model of a graph in which each edge corresponds to a wire. Let us say each edge has weight ω_{ab}, and this will mean to us that its resistance is $1/\omega_{ab}$ (a small weight corresponds to a big resistance). Ohm's law says that the potential drop across a resistor is equal to the current flowing times the resistance. If the current from a to b is $i(a, b)$ and the potentials in a and b are $v(a)$ and $v(b)$, then

$$v(a) - v(b) = \frac{i(a, b)}{\omega_{ab}}.$$

If $\mathbf{N}$ is the incidence matrix of an orientation of G, $\mathbf{W}$ the diagonal matrix with edge weights, $\mathbf{v}$ the vector with vertex potentials, and $\mathbf{i}$ the vector of edge currents, we now have

$$\mathbf{i} = \mathbf{W} \mathbf{N}^T \mathbf{v}.$$

Let $\mathbf{j}$ be a vector indexed by vertices whose ath entry denotes the net current entering or leaving the network at a. Recall that by Kirchkoff's law, the current entering a node is equal to the current exiting. Thus

$$\mathbf{j} = \mathbf{N} \mathbf{i}.$$

All together, and again making $\mathbf{L} = \mathbf{N} \mathbf{W} \mathbf{N}^T$ the weighted Laplacian matrix, we have

$$\mathbf{j} = \mathbf{L} \mathbf{v}.$$

As a consequence of this fact, it must be that $\mathbf{1}^T \mathbf{j} = 0$.

On the other hand, assume now that we are given a vector indicating currents entering and leaving the network satisfying $\mathbf{1}^T \mathbf{j} = 0$. Is this enough to find the voltages that correspond to the system?

A solution for $\mathbf{v}$ can be found computing the pseudo-inverse $\mathbf{L}^+$ of $\mathbf{L}$. That is, if $\mathbf{L}$ has spectral decomposition

$$\mathbf{L} = \sum_{i=1}^{n} \lambda_i \mathbf{E}_i,$$

with $0 = \lambda_1 < \lambda_2 \leq ... \leq \lambda_n$, then, given $\mathbf{j}$, with $\mathbf{1}^T\mathbf{j} = 0$, a solution for $\mathbf{v}$ can be found as

$$\mathbf{v} = \left(\sum_{i=2}^{n} \frac{1}{\lambda_i} \mathbf{E}_i \right) \mathbf{j}.$$

(Note that if $\mathbf{v}$ is as above, then $\mathbf{v} + \alpha\mathbf{1}$ also satisfies $\mathbf{L}(\mathbf{v} + \alpha\mathbf{1}) = \mathbf{j}$ for any α).

Now assume a and b are neighbours, and imagine one unit of current is pushed into a, and one unit extracted from b (meaning: $\mathbf{j}_a = -\mathbf{j}_b = 1$, 0 elsewhere, or simply $\mathbf{j} = \mathbf{e}_a - \mathbf{e}_b$). We can solve which potential arrangement at all vertices allows for this, and the difference of potential between a and b is defined as their *effective resistance.* In other words

$$R_{\text{eff}}(a, b) = (\mathbf{e}_a - \mathbf{e}_b)^T \mathbf{L}^+ (\mathbf{e}_a - \mathbf{e}_b).$$

Exercise 8.16. Suppose you have two edges, ab and cd. Prove that the difference of potential between c and d when you push one unit of current at a a and remove it at b is the same as the difference of potential between a and b when you push one unit of current at a c and remove it at d.

8.5 Connectivity and interlacing

Again, $\mathbf{L}$ is the Laplacian matrix, and $\lambda_1 \leq \lambda_2 \leq ... \leq \lambda_n$ its eigenvalues.

Over the next few sections, we will learn that λ_2 carries powerful information about a graph. We start with a bound associating λ_2 to the connectivity. Given a graph G, it is k-vertex-connected if it has more than k vertices, and remains connected whenever fewer than k vertices are removed. The *vertex connectivity* of a graph, denoted $\kappa_0(G)$, is the largest k to so that G is k-vertex connected. For all graphs which are not complete, this definition is equivalent to saying that $\kappa_0(G)$ is the smallest size of a subset of vertices whose removal disconnects G.

Computing the vertex connectivity of a graph is not difficult — Menger's theorem says that the size of a minimum cut in a graph is equal to the maximum number of disjoint paths that can be found between any pair of vertices. I invite you to prove this result using linear programming duality.

Nevertheless, an eigenvalue bound can always be useful.

Theorem 8.17. *Suppose $U \subseteq V(G)$. Then*

$$\lambda_2(G) \leq \lambda_2(G \backslash U) + |U|.$$

Proof. Let $\mathbf{v}'$ be a normalized $\lambda_2(G\backslash U)$ eigenvector of $\mathbf{L}(G\backslash U)$. Let $\mathbf{v}$ be the extension of

$\mathbf{v}'$ to $\mathbf{R}^{V(G)}$, adding 0s in the remaining entries. By Courant-Fisher-Weyl, we have

$$\begin{aligned}\lambda_2(G) &\leq \sum_{ab \in E(G)} (\mathbf{v}_a - \mathbf{v}_b)^2 \\ &\leq \sum_{a \in U} \sum_{b \sim a} \mathbf{v}_b^2 + \sum_{ab \in E(G \backslash U)} (\mathbf{v}_a - \mathbf{v}_b)^2 \\ &\leq |U| + \lambda_2(G \backslash U).\end{aligned}$$

□

If U is a cut-set, then $G \backslash U$ is disconnected, thus 0 has multiplicity bigger than 1, and therefore

$$\lambda_2(G) \leq \kappa_0(G).$$

This immediately implies that $\lambda_2(G) \leq \delta(G)$.

Exercise 8.18. Prove that for all trees on more than 2 vertices, $\lambda_2 \leq 1$. Prove that equality holds if and only if the tree is a star.

Interlacing "works" for $\mathbf{L}$, but the problem is that the submatrices of $\mathbf{L}$ are not Laplacian matrices of subgraphs. If we would like to related the eigenvalues of $\mathbf{L}(G)$ with those of the Laplacians of subgraphs, we must use different methods. The following exercise can be proved elementarily, just as we did above.

Exercise 8.19. Let G be a graph, and $e \in E(G)$. Prove that

$$\lambda_2(G \backslash e) \leq \lambda_2(G) \leq \lambda_2(G \backslash e) + 2.$$

Show that equality holds in the second bound if and only if G is complete.

8.6 Partitioning and cuts

Even though finding the minimum cut in a graph amount to an easy task, other problems involving cuts or partitions into clusters are significantly harder. We introduce three problems related to edge cuts:

(a) bipartition width: finding the minimum over all $e(U, \overline{U})$ where $U \subseteq V(G)$, and $|U| = \lfloor n/2 \rfloor$.

(b) maxcut: finding the maximum cut, meaning a non-empty propoer subset U of $V(G)$ so that $e(U, \overline{U})$ is maximized.

(c) finding the conductance, meaning, the minimum over all $e(U, \overline{U})/|U|$, with $U \subseteq V(G)$, $0 < |U| \leq n/2$.

These parameters are all NP-hard to compute, but we can find some interesting bounds or approximations using the eigenvalues λ_2 or λ_n. We start with an easy observation.

Lemma 8.20. *For all $U \subseteq V(G)$, we have*

$$\lambda_2 \frac{|U|(n-|U|)}{n} \leq e(U,\overline{U}) \leq \lambda_n \frac{|U|(n-|U|)}{n}.$$

Proof. Both bounds follow immediately from Exercise 8.4. □

This immediately leads to a lower bound to the bipartition width of the graph, caled $\mathrm{bw}(G)$. We have

$$\mathrm{bw}(G) \geq \frac{1}{4} n \lambda_2(G).$$

It also implies an immediate upper bound to the maxcut, labelled $\mathrm{mc}(G)$. We have

$$\mathrm{mc}(G) \leq \frac{1}{4} n \lambda_n(G).$$

Both these bounds can be made stronger by solving semidefinite programs. I won't get into details, but I will hint where in the expression we are allowed to put "new variables".

Theorem 8.21. *Let G be a graph, of even order n. Then*

$$bw(G) \geq \frac{1}{4} n \max_{\mathbf{v} \perp \mathbf{1}} \min_{\mathbf{u} \perp \mathbf{1}} \frac{\langle (\mathbf{L} + \boldsymbol{diag}(\mathbf{c}))\mathbf{u}, \mathbf{u} \rangle}{\langle \mathbf{u}, \mathbf{u} \rangle},$$

Proof. Let S be a set of cardinality $n/2$ with $e(S,\overline{S}) = \mathrm{bw}(G)$, and define $\mathbf{w} \in \mathbb{R}^V$ to be $+1$ in S and -1 in $\overline{S}$. Note that $\mathbf{w} \perp \mathbf{1}$. Also,

$$\langle \mathbf{diag}(\mathbf{v})\mathbf{w}, \mathbf{w} \rangle = 0.$$

Therefore

$$\begin{aligned} \frac{\langle (\mathbf{L} + \mathbf{diag}(\mathbf{v}))\mathbf{w}, \mathbf{w} \rangle}{\langle \mathbf{w}, \mathbf{w} \rangle} &= \frac{\langle \mathbf{L}\mathbf{w}, \mathbf{w} \rangle}{\langle \mathbf{w}, \mathbf{w} \rangle} = \frac{\sum_{ab \in E} (\mathbf{w}_a - \mathbf{w}_b)^2}{\sum_{a \in V} \mathbf{w}_a^2} \\ &= \frac{4e(S,\overline{S})}{n} = \frac{4}{n} \mathrm{bw}(G). \end{aligned}$$

□

Exercise 8.22. Let $\mathbf{Q}$ be a $n \times (n-1)$ matrix with orthonormal columns and $\mathbf{1}$ in its left kernel. Argue why we also have

$$\mathrm{bw}(G) \geq \frac{1}{4} n \max_{\mathbf{v} \perp \mathbf{1}} \lambda_1(\mathbf{Q}^T(\mathbf{L} + \mathbf{diag}(\mathbf{v})\mathbf{Q}).$$

Exercise 8.23. Prove that

$$\mathrm{mc}(G) \leq \frac{1}{4} n \min_{\mathbf{v} \perp \mathbf{1}} \lambda_n((\mathbf{L} + \mathbf{diag}(\mathbf{v})).$$

(Hint: it is similar to the Theorem above).

For the third parameter we defined, the conductance, denoted by $\Phi(G)$ and also called the isoperimetric number, Lemma 8.20 implies that

$$\Phi(G) \geq \lambda_2/2.$$

For this parameter, we can bound it from the other side as well.

Theorem 8.24. *Given a graph G, we have*

$$\Phi(G) < \sqrt{2\Delta(G)\lambda_2(G)}.$$

Proof. We consider a normalized eigenvector $\mathbf{v}$ for λ_2, and we assume without loss of generality that the vertices are ordered, meaning $V(G) = \{1, 2, ..., n\}$, in such a way that $\mathbf{v}_i \geq \mathbf{v}_{i+1}$ for all i. Let V_+ be the vertices with $\mathbf{v}_i > 0$, and assume $\mathbf{v}$ is signed so that $|V_+| \leq n/2$. Also, define $\mathbf{u}$ vector with $\mathbf{u}_i = \mathbf{v}_i$ if $\mathbf{v}_i > 0$, and $\mathbf{u}_j = 0$ otherwise. We finally define E_+ the set of edges incident to one vertex in V_+.

To each i, $1 \leq i \leq |V(G)|$, we consider the cut

$$C_i = \{\{j, k\} \in E(G) : 1 \leq j \leq i < k \leq n\}.$$

Let

$$\alpha = \min_{1 \leq i \leq n} \frac{|C_i|}{\min\{i, n-i\}},$$

whence $\alpha \geq \Phi(G)$. Let $\mathbf{P}$ be the projection onto the subspace spanned by the characteristic vectors of the vertices in V_+, that is, $\mathbf{P}\mathbf{v} = \mathbf{u}$. We now have

$$
\begin{aligned}
\lambda_2 = \frac{\mathbf{v}^T\mathbf{P}\mathbf{L}\mathbf{v}}{\mathbf{v}^T\mathbf{P}\mathbf{v}} &= \frac{(\mathbf{N}^T\mathbf{P}\mathbf{v})^T(\mathbf{N}^T\mathbf{v})}{\sum_{i\in V_+}\mathbf{v}_i^2} \\
&= \frac{\sum_{ij\in E_+}(\mathbf{u}_i-\mathbf{u}_j)(\mathbf{v}_i-\mathbf{v}_j)}{\sum_{i\in V_+}\mathbf{v}_i^2} \\
&> \frac{\sum_{ij\in E_+}(\mathbf{u}_i-\mathbf{u}_j)^2}{\sum_{i\in V_+}\mathbf{u}_i^2} \\
&= \frac{\sum_{ij\in E_+}(\mathbf{u}_i-\mathbf{u}_j)^2\sum_{ij\in E_+}(\mathbf{u}_i+\mathbf{u}_j)^2}{\sum_{i\in V_+}\mathbf{u}_i^2\sum_{ij\in E_+}(\mathbf{u}_i+\mathbf{u}_j)^2} \\
&\geq \frac{\left(\sum_{i\sim j}|\mathbf{u}_i^2-\mathbf{u}_j^2|\right)^2}{2\Delta\left(\sum_{i\in V_+}\mathbf{u}_i^2\right)^2} \\
&\geq \frac{\left(\sum_i|\mathbf{u}_i^2-\mathbf{u}_{i+1}^2||C_i|\right)^2}{2\Delta\left(\sum_{i\in V_+}\mathbf{u}_i^2\right)^2} \\
&\geq \frac{\left(\sum_i|\mathbf{u}_i^2-\mathbf{u}_{i+1}^2|\alpha i\right)^2}{2\Delta\left(\sum_{i\in V_+}\mathbf{u}_i^2\right)^2} \\
&\geq \frac{\alpha^2}{2\Delta} \\
&\geq \frac{\Phi^2}{2\Delta}
\end{aligned}
$$

□

Exercise 8.25. Justify in details each of the steps of the inequality chain above.

Note that not only the result above gives a bound, but it also has an implicit algorithm in its proof. In fact, we were able to efficiently find a set of vertices U so that

$$\Phi \leq \frac{e(U,\overline{U})}{|U|} \leq \sqrt{2\Delta\lambda_2} \leq 2\sqrt{\Delta\Phi}.$$

8.7 Normalized Laplacian

As we now know,

$$\mathbf{L} = \mathbf{D} - \mathbf{A}.$$

We assume G has no isolated vertices. We define $\mathbf{Q}$ as

$$\mathbf{Q} = \mathbf{D}^{-1/2}\mathbf{L}\mathbf{D}^{-1/2} = \mathbf{I} - \mathbf{D}^{-1/2}\mathbf{A}\mathbf{D}^{-1/2}.$$

Note that it is positive semidefinite.

Exercise 8.26. Prove that
$$R_{\mathbf{Q}}(\mathbf{u}) = \frac{\sum_{ab \in E} (\mathbf{v}_a - \mathbf{v}_b)^2}{\sum_{a \in V} \mathbf{v}_a^2 d(a)},$$
where $\mathbf{v} = \mathbf{D}^{-1/2}\mathbf{u}$.

We also prove some basic properties.

Theorem 8.27. *Let G be a graph with no isolated vertices, and denote the eigenvalues of $\mathbf{Q}$ by $\mu_1 \leq ... \leq \mu_n$. Then*

(i) $\sum_j \mu_j = n$.

(ii) For $n \geq 2$, $\mu_2 \leq n/(n-1)$, and equality holds if and only if G is the complete graph. Also, $\mu_n \geq n/(n-1)$.

(iii) For a graph not complete, we have $\mu_2 \leq 1$.

(iv) The multiplicity of 0 as an eigenvalue is the number of connected components of G.

(v) We have $\mu_n \leq 2$, and equality holds if and only if G is bipartite. In this case, for all μ eigenvalue of $\mathbf{Q}$, $2 - \mu$ is also eigenvalue.

Exercise 8.28. Prove the properties above.

Exercise 8.29. Let G be a graph, connected, diameter d. Prove that
$$\mu_2 \geq \frac{1}{d \sum_{a \in V} d(a)}.$$

8.8 Random Walks

Let G be a weighted graph (edge weights are given by a function ω). We assume a walker is sitting at a vertex, and then decides to move with certain probability. At each step, this walker hops from one vertex to another with probability proportional to the edge weight of the corresponding edge. This model is equivalent to a Markov chain defined on a finite measurable state space.

We will be specially interested in the expected behaviour of a random walker, rather than on some fixed particular instance of this experiment.

Let $\mathbf{p}_t \in \mathbb{R}^V$ denote the probability distribution of the walker at time t. As such, $\mathbf{p}_t(a) \geq 0$ for all $a \in V$, and
$$\sum_{a \in V} \mathbf{p}_t(a) = 1.$$
Because the probability of arriving at a at $t+1$ steps is only determined by the probability distribution at time t and the edge weights, we have
$$\mathbf{p}_{t+1}(a) = \sum_{b \sim a} \frac{\omega_{ab}}{d_b} \mathbf{p}_t(b).$$

with d_b meaning the sum of the weights of edges incident to b. Equivalently,

$$\mathbf{p}_{t+1} = \mathbf{A}\mathbf{D}^{-1}\mathbf{p}_t,$$

where $\mathbf{A}$ is the weighted adjacency matrix and $\mathbf{D}$ the diagonal matrix of (weighted) degrees. Let $\mathbf{W} = \mathbf{A}\mathbf{D}^{-1}$. It is immediate to verify that $\mathbf{p}_{t+k} = \mathbf{W}^k\mathbf{p}_t$, where $\mathbf{p}_0$ typically stands for the starting distribution. We also see that

$$\mathbf{D}^{-1/2}\mathbf{W}\mathbf{D}^{1/2} = \mathbf{I} - \mathbf{Q},$$

where $\mathbf{Q}$ is the normalized Laplacian. If there is a probability that the walker does not move, say $1/2$, we now have

$$\mathbf{p}_{t+1} = (1/2)\mathbf{I}\mathbf{p}_t + (1/2)\mathbf{A}\mathbf{D}^{-1}\mathbf{p}_t.$$

Let $\mathbf{Z} = (1/2)(\mathbf{I} + \mathbf{A}\mathbf{D}^{-1})$. You can now see that

$$\mathbf{D}^{-1/2}\mathbf{Z}\mathbf{D}^{1/2} = \mathbf{I} - (1/2)\mathbf{Q},$$

Matrices $\mathbf{Z}$ or $\mathbf{W}$ are not symmetric, but they are both similar to a symmetric matrix. This gives that they are diagonalizable with real eigenvectors.

Exercise 8.30. If $\mathbf{v}$ is λ-eigenvector of $\mathbf{Q}$, to which eigenpair of $\mathbf{W}$ or $\mathbf{Z}$ they relate? Later, prove that all eigenvalues of $\mathbf{W}$ are between -1 and 1, and those of $\mathbf{Z}$ lie between 0 and 1.

Exercise 8.31. What do you obtain if you replace $1/2$ by another probability?

A random walk $\mathbf{W}$ converges to a distribution $\mathbf{p}$ if for any given ε and any distribution $\mathbf{q}$, there is an n so that

$$||\mathbf{W}^n\mathbf{q} - \mathbf{p}|| < \varepsilon.$$

Exercise 8.32. Can you show that if $\mathbf{W}$ converges to $\mathbf{p}$, then $\mathbf{p}$ is "stable", meaning, $\mathbf{W}\mathbf{p} = \mathbf{p}$? Later, prove that every graph contains a stable distribution, and if the graph is connected, this is unique (it also does not depend on the probability of staying put).

Exercise 8.33. Show that if there is any probability that a walker stays put (we call these random walks "lazy"), then the random walk will converge to the stable distribution. Describe the graphs for which a non-lazy random walk does not converge.

Example 8.34. Imagine now the following experiment. A deck of n cards c_1,...,c_n is lying on a table. We will shuffle these cards in a very stupid way: at each time step, we select i, j from 1 to n uniformly at random and exchange the positions of cards i and j (that includes choosing $i = j$ and doing nothing). How fast does this procedure produces a good shuffling?

The graph here is the one whose vertex set corresponds to the permutations on n elements. Vertices adjacent if one can be obtained from the other by applying a transposition (this is called the Cayley Graph $\mathrm{Cay}(S_n, T)$).

The weights here are simply determined:

$$\mathbf{p}_{t+1} = (1/n)\mathbf{I}\mathbf{p}_t + [(n-1)/n]\mathbf{A}\mathbf{D}^{-1}\mathbf{p}_t,$$

with $\mathbf{A}_{\sigma\tau} = 1$ if $\sigma\tau^{-1}$ is a transposition, and $= 0$ otherwise.

We thus want to know how fast $\mathbf{p}_t$ becomes the stable distribution.

Exercise 8.35. What is the stable distribution of the example above?

We can now provide a reasonably good estimate.

Theorem 8.36. *Let $\mathbf{W}$ be the transition matrix of a random walk, having laziness probability $1-\rho$, and with eigenvalues $1 = \omega_1 \geq \omega_2 \geq \ldots \geq \omega_n$. Assume $\omega = \max\{|\omega_2|, |\omega_n|\} < 1$, meaning, the random walk converges, say, to $\mathbf{p}$. Let $\mathbf{p}_0 = \mathbf{e}_a$, meaning, the walk starts at a. Then*

$$|\mathbf{p}_t(b) - \mathbf{p}(b)| < \sqrt{\frac{d(b)}{d(a)}}\omega^t.$$

Proof. Let

$$\mathbf{Q} = \sum_{i=1}^{n} \lambda_i \mathbf{F}_i$$

be the spectral decomposition of the normalized Laplacian. Then

$$\mathbf{W} = \mathbf{I} - \rho \mathbf{D}^{1/2}\mathbf{Q}\mathbf{D}^{-1/2} = \sum_{i=1}^{n}(1-\rho\lambda_i)\mathbf{D}^{1/2}\mathbf{F}_i\mathbf{D}^{-1/2} = \sum_{i=1}^{n}\omega_i \mathbf{E}_i.$$

Note that

$$\mathbf{E}_1 = \frac{1}{\sum_{a \in V} d(a)}\mathbf{D}^{1/2}\mathbf{D}^{1/2}\mathbf{1}\mathbf{1}^T\mathbf{D}^{1/2}\mathbf{D}^{-1/2} = \frac{1}{\sum_{a \in V} d(a)}\mathbf{D}\mathbf{1}\mathbf{1}^T.$$

Then

$$\begin{aligned}
|\mathbf{p}_t(b) - \mathbf{p}(b)| &= |\mathbf{e}_b^T \mathbf{W}^t \mathbf{e}_a - \mathbf{e}_b^T \mathbf{E}_1 \mathbf{e}_b| \\
&= \left| \sum_{i=2}^{n} \omega_i^t (\mathbf{e}_b^T \mathbf{E}_i \mathbf{e}_a) \right| \\
&\leq \omega^t \sum_{i=2}^{n} \left| \mathbf{e}_b^T \mathbf{E}_i \mathbf{e}_a \right| \\
&\leq \omega^t \sqrt{\frac{d(b)}{d(a)}} \sum_{i=2}^{n} \left| \mathbf{e}_b^T \mathbf{F}_i \mathbf{e}_a \right| \\
&\leq \omega^t \sqrt{\frac{d(b)}{d(a)}} \sum_{i=2}^{n} \sqrt{\mathbf{e}_b^T \mathbf{F}_i \mathbf{e}_b}\sqrt{\mathbf{e}_a^T \mathbf{F}_i \mathbf{e}_a} \\
&\leq \omega^t \sqrt{\frac{d(b)}{d(a)}} \sqrt{\sum_{i=2}^{n} \mathbf{e}_b^T \mathbf{F}_i \mathbf{e}_b}\sqrt{\sum_{i=2}^{n} \mathbf{e}_a^T \mathbf{F}_i \mathbf{e}_a} \\
&\leq \omega^t \sqrt{\frac{d(b)}{d(a)}} \sqrt{1 - \mathbf{e}_b^T \mathbf{F}_1 \mathbf{e}_b}\sqrt{1 - \mathbf{e}_a^T \mathbf{F}_1 \mathbf{e}_a} \\
&< \omega^t \sqrt{\frac{d(b)}{d(a)}}.
\end{aligned}$$

□

Exercise 8.37. Assume G is connected and non-bipartite, with an initial probability distribution $\mathbf{q}$. Let $\mathbf{W}$ be the transition matrix of a non-lazy random walk, and, as before, let $\omega = \max\{|\omega_2|, |\omega_n|\}$. Let $\mathbf{p}$ be the stable distribution. Prove that

$$||\mathbf{p}_t - \mathbf{p}|| < \omega^t.$$

We end this section with a nice exercise.

Exercise 8.38. Let $\mathbf{L}$ be the combinatorial Laplacian, with eigenvalues $\lambda_1 \leq ... \leq \lambda_n$, and $\mathbf{Q}$ the normalized version, with eigenvalues $\mu_1 \leq ... \leq \mu_n$. Let Δ and δ be the largest and smallest degrees of the graph. Verify that

$$\frac{\lambda_i}{\Delta} \leq \mu_i \leq \frac{\lambda_i}{\delta}.$$

(Hint: Use the Courant-Fisher-Weyl theorem — and apply the transformation $\mathbf{D}^{1/2}$).

8.9 References

Here is the set of references used to write the past few pages.

The initial material on Laplacian matrix was mostly based on Godsil and Royle's, Chapter 13.

Fan Chung's book "Spectral Graph Theory" is the standard reference on the Normalized Laplacian.

Bojan Mohar has several articles about Laplacian matrices: "Some Applications of Laplace Eigenvalues of Graphs", "The Laplacian Spectrum of Graphs", "Eigenvalues in combinatorial optimization" (with S. Poljak), and others.

Finally, I also acknowledge D. Spielman's course notes (2018), specially for the last section on random walks.