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

$$\mathbf{PAP}^{\mathsf{T}} = \mathbf{B}.$$

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

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

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. $\square$

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* $\mathbf{PAP}^\mathsf{T} = \mathbf{B}$ *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

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

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 $\mathbf{U}' = \mathbf{V}'$. $\square$

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!). $\square$

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 $a \in V_r$. $\square$

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).