Méthodes spectrales pour la détection de communautés dans les graphes

Guillaume Lecué*

1 Introduction

Exemple d'un problème de détection de communauté : On dispose d'un graphe où chaque nœud représente une personne et un lien du graphe entre deux nœuds représente une connexion entres ces deux personnes établie, par exemple, à partir d'un réseau social d'échange de messages. On cherche à identifier dans ce graphe des groupes de noeuds particulièrement connectés. Ces groupes peuvent représenter des groupes d'amis ou des personnes partageant des intérêts communs qui ont donc tendance à souvent s'échanger des messages. On dispose donc d'un graphe et le problème de détection de communauté est de trouver des groupes de nœuds plus densément connectés entre eux qu'avec le reste du graphe. Une fois les communautés détecter, on peut chercher les influenceurs en leur sein et leur envoyer des pubs ou les démarcher directement.

Définition 1.1. *Un* ***graphe*** *est un couple* (V, E) *où* V *est un ensemble dont les éléments sont appelés les* ***noeuds du graphe*** *et* $E \subset V \times V$ *est un ensemble dont les éléments sont appelés* ***arêtes du graphe****. Un* ***graphe pondéré*** *est un triplet* (V, E, W) *tel que* (V, E) *est un graphe et* $W \in \mathbb{R}^{|V|\times|V|}$*. Les entrées de* W *sont appelées les* ***poids du graphe****. On dit qu'un graphe* (V, E) *est* ***non orienté*** *quand* $(i, j) \in E$ *implique* $(j, i) \in E$ *et qu'un graphe pondéré* (V, E, W) *est* ***non orienté*** *quand* (V, E) *est un graphe non orienté et* $W = W^\top$*.*

Il existe plusieurs types de graphes selon qu'ils sont orientés (ou non) et pondérés (ou non) comme représenté dans la Figure 3. On donne aussi quelques exemples classiques de graphes dans les Figures 1 et 2.

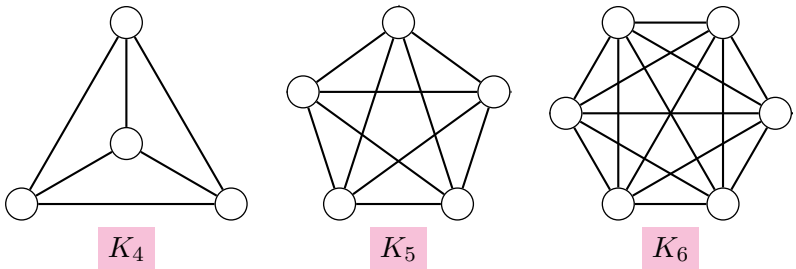

FIGURE 1 – 3 exemples de graphes complets

*CREST, ENSAE. Bureau 3029, 5 avenue Henry Le Chatelier. 91 120 Palaiseau. Email: guillaume.lecue@ensae.fr.

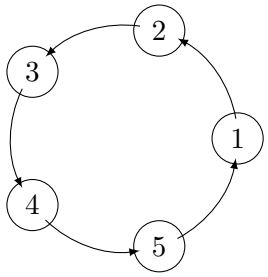

FIGURE 2 – Exemple d'un ring graph

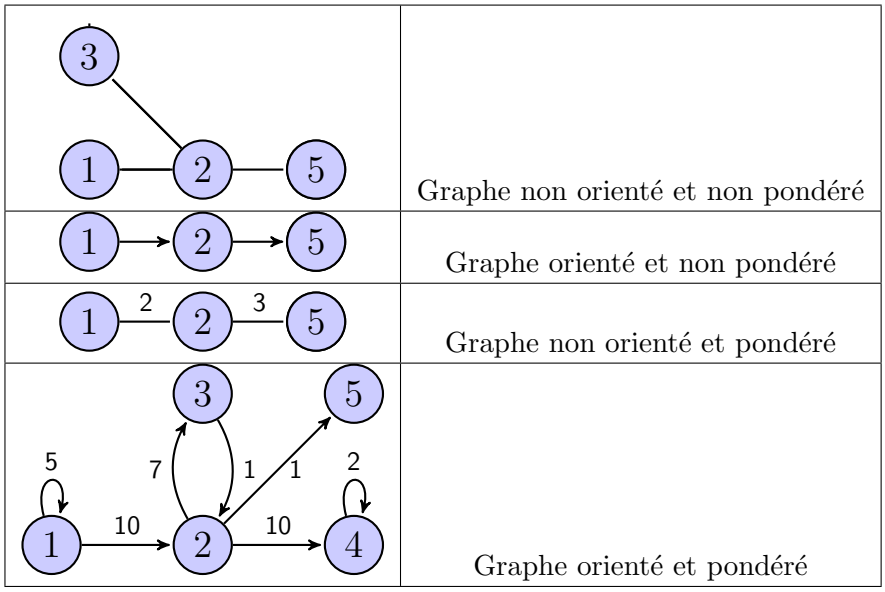

FIGURE 3 – Exemples de graphes orienté ou non et pondéré ou non

On représente souvent les propriétés de connexions du graphe à l'aide d'une matrice.

Définition 1.2. *Soit $G = (V, E)$ un graphe. La **matrice d'adjacence** de G est donnée par $A \in \{0,1\}^{|V|\times|V|}$ où pour tout $(i,j) \in \{1,\dots,|V|\}^2$, on a $A_{ij} = 1$ quand $(i,j) \in E$ et $A_{ij} = 0$ quand $(i,j) \notin E$. La matrice d'adjacence d'un graphe pondéré (V, E, W) est $A \in \mathbb{R}^{|V|\times|V|}$ où on a $A_{ij} = W_{ij}$ quand $(i,j) \in E$ et $A_{ij} = 0$ quand $(i,j) \notin E$.*

Si un graphe est non orienté alors sa matrice d'adjacence est symétrique. Pour le problème de détection de communautés, on s'intéresse aux propriétés de connectivité dans un graphe. La première définition est que tous les nœuds du graphe sont connectés.

Définition 1.3. *Un graphe $G = (V, E)$ est dit **connecté** ou **connexe** quand on peut trouver un chemin d'arêtes liant tous les nœuds du graphe. Une **composante connexe** de G est un sous-graphe de G connecté et sans arête le liant avec son complémentaire. Une **partition connexe** de G est un ensemble fini de sous-graphes de G formant une partition des nœuds de G telle que les sous-graphes associés à cette partition sont des composantes connexes.*

On peut aussi étendre cette définition aux graphes pondérés. On retrouve dans ce cas la définition précédente pour un graphe non pondéré quand on lui ajoute la matrice de poids constants $W = (1)_{(i,j)\in V\times V}$.

Définition 1.4. *Un graphe pondéré $G = (V, E, W)$ est dit* ***connecté*** *ou* ***connexe*** *quand on peut trouver un chemin d'arêtes liant tous les nœuds du graphe et $W_{ij} \neq 0$ pour toutes les arêtes (i, j) de ce chemin. Une* ***composante connexe*** *de G est un sous-graphe de G connecté et sans arête (de poids non nul) le liant avec son complémentaire. Une* ***partition connexe*** *d'un graphe G est un ensemble fini de sous-graphes de G formant une partition des nœuds de G telle que les composantes de cette partition sont connexes.*

2 Le Laplacien d'un graphe et quelques propriétés de son spectre

Dans cette section, on s'intéresse au lien entre les propriétés de connectivité d'un graphe et le spectre de son Laplacien. On définit d'abord le Laplacien d'un graphe.

Définition 2.1. *Soit G un graphe (pondéré ou non). On note par V l'ensemble des nœuds de G et par A sa matrice d'adjacence. La* ***matrice de Laplace*** *ou* ***Laplacien*** *de G est donnée par $L = D - A$ où $D = \mathrm{diag}(d)$, $d = (d_i)_{i \in V}$ et $d_i = \sum_{j \in V} A_{ij}$ pour tout $i \in V$ est le degré du i-ième noeud.*

Si G est un graphe non orienté alors son Laplacien est symétrique. Dans ce cas, le spectre de L va jouer un rôle clef pour la méthode spectral de détection des communautés.

Proposition 2.2. *Soit $G = (V, E, W)$ un graphe pondéré non orienté. On suppose que $W_{ij} \geq 0$ pour tout $(i, j) \in V \times V$. On a :*

1) pour tout $f \in \mathbb{R}^{|V|}$, $f^\top L f = \frac{1}{2} \sum_{(i,j) \in E} W_{ij}(f_i - f_j)^2$

2) L est symétrique positive

3) 0 est valeur propre de L et $L\mathbf{1} = 0$ où $\mathbf{1} = (1)_{i \in V}$.

Démonstration. **1)** On note par $A = (W_{ij}\mathbf{1}_{(i,j) \in E})$ la matrice d'adjacence de G. Comme G est non orienté, on a $W_{ij} = W_{ji}$ pour tout $(i, j) \in V \times V$ et donc pour tout $f \in \mathbb{R}^{|V|}$,

$$
\begin{aligned}
f^\top L f = f^\top (D - A) f &= \sum_{i \in V} d_i f_i^2 - \sum_{(i,j) \in E} f_i W_{ij} f_j \\
&= \sum_{i \in V} \sum_{j \in V} W_{ij} \mathbf{1}_{(i,j) \in E} f_i^2 - \sum_{(i,j) \in E} f_i W_{ij} f_j = \frac{1}{2} \sum_{(i,j) \in E} W_{ij}(f_i^2 - 2 f_i f_j + f_j^2) \\
&= \frac{1}{2} \sum_{(i,j) \in E} W_{ij}(f_i - f_j)^2.
\end{aligned}
$$

2) On a pour tout $f \in \mathbb{R}^{|V|}$,

$$
\big\langle f, Lf \big\rangle = f^\top L f = \frac{1}{2} \sum_{(i,j) \in E} W_{ij}(f_i - f_j)^2 \geq 0.
$$

On a donc bien $L \succeq 0$.

3) Par définition du degré d'un noeud, on a pour tout $i \in V$,

$$
(L\mathbf{1})_i = \sum_{j \in V} L_{ij} = d_i - \sum_{j \in V} A_{ij} = 0.
$$

On a donc bien 0 comme valeur propre de L et $\mathbf{1}$ est un vecteur propre associé à la valeur propre 0. ∎

La multiplicité de la valeur propre 0 de L donne le nombre de composantes connexes de G et les indicatrices des nœuds des composantes connexes sont des vecteurs propres engendrant cet espace propre de L.

Proposition 2.3. *Soit $G = (V, E, W)$ un graphe pondéré non orienté dont les poids sont positifs. La multiplicité de la valeur propre 0 du Laplacien de G est égal au nombre de composante connexes de G. De plus l'espace propre associé à la valeur propre 0 de L, càd son noyau $\mathrm{Ker}(L)$, est engendré par $\mathbf{1}_{V_1}, \ldots, \mathbf{1}_{V_k}$ où $V_1 \sqcup \cdots \sqcup V_k$ est la partition de nœuds de V en composantes connexes de G.*

Démonstration. On commence par le cas d'une seule composante connexe, càd quand G est connexe. Le noyau de L est formé de tous les éléments $f \in \mathbb{R}^{|V|}$ tels que $Lf = 0$. Si $Lf = 0$ alors $f^\top Lf = 0$ et donc d'après la Proposition 2.2, on aura $\sum_{(i,j)\in E} W_{ij}(f_j - f_i)^2 = 0$. Par ailleurs, les poids W_{ij} sont positifs ou nuls, alors, si $\sum_{(i,j)\in E} W_{ij}(f_j - f_i)^2 = 0$, on a pour tout $(i, j) \in E$, $W_{ij}(f_j - f_i)^2 = 0$ donc soit $W_{ij} = 0$ soit $f_j = f_i$. Par ailleurs, G est connexe donc, si on prend n'importe quel couple de nœuds $i, j \in V$ il existe un chemin représenté par des indices de noeuds $i_1, \ldots, i_p \in V$ tels que $(i, i_1) \in E, (i_1, i_2) \in E, \ldots, (i_p, j) \in E$ et $W_{ii_1} > 0, W_{i_1 i_2} > 0, \cdots, W_{i_k j} > 0$. Le long de ce chemin, on a soit $W_{pq} = 0$ soit $f_p = f_q$ pour $(p, q) \in \{(i, i_1), (i_1, i_2), \cdots, (i_p, j)\}$; mais comme $W_{pq} \neq 0$ on a $f_p = f_q$. Ceci étant vrai pour toutes les arêtes du chemin, on en déduit que les extrémités f_i et f_j sont égales. Donc, f est un vecteur constant. On a donc bien que $\mathrm{Ker}(L)$ est engendré par $\mathbf{1} = (1)_{i\in V}$ l'indicatrice de l'unique composante connexe de G.

On considère le cas général où G a k composantes connexes. Quitte à réordonner les nœuds de V, on peut écrire le Laplacien de G sous forme de matrice par blocs :

$$L = \left[\begin{array}{c|c|c|c} L_1 & 0 & \cdots & 0 \\ \hline 0 & L_2 & \cdots & 0 \\ \hline \cdots & \cdots & \ddots & 0 \\ \hline 0 & 0 & \cdots & L_k \end{array}\right]$$

où pour tout $i = 1, \ldots, k$, L_i est le Laplacien de la i-ième composante connexe $G_i = (V_i, E_i, W^{(i)})$ de G où $V_i \subset V$ est l'ensemble des nœuds de la i-ième composante connexe de G, $E_i = \{(p, g) : p, q \in V_i\}$ est l'ensemble de ses arêtes et $W^{(i)} = (W_{pq} : p, q \in V_i)$ est sa matrice de poids.

On note $\mathbf{1}_{V_i} \in \mathbb{R}^{|V|}$ l'indicatrice des nœuds de G_i. On veut montrer que $\mathrm{Ker}(L)$ est engendré par $\mathbf{1}_{V_i}, i = 1, \cdots, k$. Pour tout $i = 1, \ldots, k$, L_i est le Laplacien de G_i donc 0 est valeur propre de L_i et $(1)_{p\in V_i}$ engendre le noyau de L_i. Donc $\mathbf{1}_{V_i}$ est dans le noyau de L. On a donc $\mathrm{vect}(\mathbf{1}_{V_i}, i = 1, \cdots, k) \subset \mathrm{Ker}(L)$. Par ailleurs, étant donnée la structure par bloc de L, on voit que $Lf = 0$ si et seulement si $L_i f = 0$ pour tout $i = 1, \ldots, k$ et comme $\ker(L_i) = \mathrm{vect}((\mathbf{1})_{\mathbf{p}\in\mathbf{V_i}})$, on a $f_{|V_i} = \alpha_i(\mathbf{1})_{\mathbf{p}\in\mathbf{V_i}}$ pour un certain $\alpha_i \in \mathbb{R}$. On en déduit que $f = \sum_i \alpha_i \mathbf{1}_{V_i} \in \mathrm{vect}(\mathbf{1}_{V_i}, i = 1, \cdots, k)$. ∎

Remarque 2.4. *La Proposition 2.3 s'applique aussi aux graphes non pondérés non orientés. Il suffit de l'appliquer à la matrice de poids $W = (1)_{(i,j)\in V\times V}$.*

Exemple : Calculons la multiplicité de la valeur propre 0 du Laplacien d'un graphe complet et vérifions qu'elle est bien égale à 1 étant donné que ce graphe est connexe. On note K_n le graphe complet de n sommets. Le Laplacien de K_n est donné par

$$L = \mathrm{diag}(n, \cdots, n) - (1)_{n\times n} = \begin{bmatrix} n-1 & -1 & \cdots & -1 \\ -1 & n-1 & \cdots & -1 \\ \cdots & \cdots & \ddots & -1 \\ -1 & -1 & \cdots & n-1 \end{bmatrix}.$$

Si $f \in \mathbb{R}^n$ est tel que $Lf = 0$ alors $nf = (\langle f, (1)_1^n \rangle)_{i=1,\dots,n} = \langle f, (1)_1^n \rangle (1)_1^n$. Donc $f \in \text{vect}((1)_1^n)$ et 0 est de multiplicité 1 pour L.

Le Laplacien d'un graphe et la loi de refroidissement de Newton. On a vu que le Laplacien $L = D - A$ d'un graphe joue un rôle essentiel sur les propriétés de connectivité du graphe. Il apparaît aussi lorsqu'on étudie d'autres propriétés d'un graphe. On peut se poser la question sur l'origine de son nom et en particulier s'il a un lien avec le Laplacien qu'on rencontre en physique quand on étudie l'équation de la chaleur. Il se trouve qu'il y a bien un lien entre les deux notion et qu'on peut voir le Laplacien d'un graphe comme une version discrète du Laplacien en physique. Le lien unissant les deux approches est la loi de refroidissement de Newton ou de transfert de chaleur (*Newton's law of cooling* en anglais) disant que *la chaleur se transfert d'un point à un autre proportionnellement à la différence de température entre les deux points.*

On imagine que les nœuds de notre graphe $G = (V, E)$ ont des températures données par la famille $(T_i)_{i \in V}$ (T_i est la température du nœud i à la date t). On laisse évoluer les transferts de chaleur sur ce graphe en fonction de la loi de Newton : 'le gradient de température entre deux points connectés est proportionnel à la différence de température entre ces deux points' : pour tout nœud $i \in V$, la température en ce nœud va évoluer de la manière suivante

$$\frac{dT_i}{dt} = -\kappa \sum_{j \in V} A_{ij}(T_i - T_j). \tag{1}$$

On peut ensuite regarder l'évolution du vecteur des températures des noeuds du graphe $T = (T_i)_{i \in V}$ et développer cette égalité pour faire apparaître le Laplacien :

$$\frac{dT}{dt} = -\kappa \left(\sum_{j \in V} A_{ij}(T_i - T_j) \right)_{i \in V} = -\kappa \left(T_i \sum_{j \in V} A_{ij} - \sum_{j \in J} A_{ij} T_j \right)_{i \in V} = -\kappa (D - A) T.$$

On obtient donc une équation différentielle sur l'évolution des températures des nœuds du graphe de la forme $dT/dt + \kappa LT = 0$ où $L = D - A$ est le Laplacien du graphe. Si on rappelle l'équation de la chaleur $\partial u / \partial t = \alpha \Delta u$ où Δ est le Laplacien $\partial_1^2 + \cdots \partial_d^2$, on peut identifier L et $-\Delta$. C'est de cette analogie que L tire son nom.

3 Principe des méthodes spectrales en détection de communautés

On donne ici deux idées introduisant plus ou moins formellement la méthode spectrale pour trouver des communautés dans des graphes. On suppose qu'on dispose d'un graphe $G = (V, E, W)$ pondéré non-orienté.

3.1 Cadre idéal de composantes connexes

Cadre idéal : Quand il s'agit de détecter des communautés dans un graphe, en quelque sorte, le cadre idéal a lieu quand ces communautés ne sont pas connectées entre elles, càd quand le graphe admet une partition en composantes connexe et que chaque composante connexe constitue une communauté. C'est un cadre idéal car il n'apparaît presque jamais en pratique, vue qu'on a toujours quelques liens inter-communautés.

Même si le cadre idéal est un cadre qui n'a presque jamais lieu en pratique, il est un bon guide pour comprendre comment solutionner le problème de détection de communautés dans des situations plus générales.

On se place alors dans le cadre idéal (en première approximation). Le Laplacien de $G = (V, E, W)$ et son noyau ont donc des formes particulières. En effet, d'après la Proposition 2.3, le noyau de L est engendré par $\mathbf{1}_{V_1}, \dots, \mathbf{1}_{V_k}$ où $V_1 \sqcup \cdots \sqcup V_k$ est la partition de nœuds de V en les k composantes connexes de G.

Il n'est cependant pas facile de trouver les indicatrices $\mathbf{1}_{V_i}, i = 1, \dots, k$ à partir de L. Ce qui est plus facile est de trouver une base orthonormale de $\mathrm{Ker}(L)$: $u_1, \dots, u_k \in \mathbb{R}^{|V|}$ (on peut d'ailleurs prendre $u_1 = (1)_{i=1}^n$). La **méthode spectrale** procède ensuite en deux étapes :

1) on clusterise les $|V|$ vecteurs lignes $(y_i)_{i \in V}$ de la matrice $[u_1|u_2|\cdots|u_k]$ de taille $|V| \times k$ en k clusters $C_1, \cdots, C_k$
2) on retourne la partition $V_1 \sqcup \cdots \sqcup V_k$ des neouds de V où $V_p = \{i \in V : y_i \in C_p\}$ pour $p = 1, \dots, k$. Ceux sont les communautés de noeuds qu'on a détecté.

On s'attend à ce que la partition $V_1 \sqcup \cdots \sqcup V_k$ des noeuds de V donne les k composantes connexes de G . En effet, $u_1, \dots, u_k$ et $\mathbf{1}_{V_1}, \dots, \mathbf{1}_{V_k}$ sont deux bases orthonormales de $\mathrm{Ker}(L)$. Il existe alors une matrice de rotation (matrice orthogonale) $R \in \mathcal{O}(|V|)$ telle que $u_i = R\mathbf{1}_{V_i}, i = 1, \dots, k$. Hors si on effectue un clustering sur la matrice $[\mathbf{1}_{V_1}|\cdots|\mathbf{1}_{V_k}]$ à k clusters, on se dit que les clusters seront formés des lignes qui sont toutes égales entre elles : celles ayant un 1 au même endroit et 0 ailleurs – il y en a exactement k comme ça. C'est donc la partition de G en ses composantes connexes. Par ailleurs, on se dit aussi que la rotation R devrait conserver ce clustering des lignes et que donc clusteriser les vecteurs lignes de la matrice $[u_1|u_2|\cdots|u_k]$ devrait aussi redonner les composante connexes de G comme ce clustering le fait sur $[\mathbf{1}_{V_1}|\cdots|\mathbf{1}_{V_k}]$.

Voilà pour ce qui est de l'intuition derrière une méthode spectrale courament utilisée en détection de communautés. En pratique, G ne sera pas partitionable en composantes connexes ; ainsi, on ne regardera pas le noyau de L mais plustôt son espace propre associé à ses k plus petites valeurs propres. On cherchera ensuite une base de k vecteurs propres $u_1, \dots, u_k$ de cet espace propre. C'est alors les $|V|$ lignes de la matrice $[u_1|\cdots|u_k]$ qu'on clusterisera en k groupes. Des indices de ligne de ces k groupes on extraira un clustering des noeuds du graphe, càd on aura un estimateur des communautés.

3.2 Point de vue "graph cut"

On donne dans cette section, un autre point de vue sous-jacent aux méthodes spectrales. L'idée ici est que détecter des communautés est en fait équivalent à trouver une partition des nœuds minimisant la somme des poids portés par les arêtes inter-communautés tout en maximisant la taille de ces communautés.

Pour formaliser cette approche, on introduit quelque fonctions, appelées **fonctions de modularité**, qui quantifie les notions de masse de poids intra et inter communautés qu'on cherchera à optimiser sur tous les "**cuts**", càd partitions, du graphe.

Mincut problem : Étant donné deux ensembles $A, B \subset V$ de nœuds, on définit la masse totale entre A et B par

$$W(A, B) = \sum_{\substack{(i,j) \in E \\ (i,j),(j,i) \in A \times B}} W_{ij} = \sum_{\substack{(i,j) \in E \\ (i,j) \in A \times B}} W_{ij} + W_{ji} = 2 \sum_{\substack{(i,j) \in E \\ (i,j) \in A \times B}} W_{ij}$$

où on a utilisé la symétrie de W dans la dernière égalité. Ainsi on peut définir une fonction associant à chaque partition $V_1 \sqcup \cdots \sqcup V_k$ des nœuds du graphe la masse totale des poids intra-

communauté :

$$(V_1, \dots, V_k) \in \mathcal{P}_k(V) \longrightarrow \mathrm{cut}(V_1, \dots, V_k) = \frac{1}{2} \sum_{i=1}^{k} W(V_i, V_i^c).$$

où $\mathcal{P}_k(V)$ est l'ensemble des partitions de V ayant au plus k éléments et V_i^c est le complémentaire de V_i dans V.

Pour le problème à deux communautés, on peut récrire le problème sous la forme suivante. Chaque partition de V en deux communautés $V_1 \sqcup V_2$ (ici $V_2 = V_1^c$) est décrite par un vecteur $x \in \{0,1\}^{|V|}$ d'appartenance tel que $x_i = 1$ si $i \in V_1$ et $x_i = 0$ si $i \in V_2$. On a alors

$$\mathrm{cut}(V_1, V_2) = W(V_1, V_2) = \sum_{\substack{i:x_i=1 \\ j:x_j=0 \\ (i,j)\in E}} W_{ij} + W_{ji} = 2 \sum_{(i,j)\in E} x_i W_{ij}(1 - x_j).$$

Le problème du mincut à deux classes peut donc s'écrire : trouver $x^* \in \{0,1\}^{|V|}$ solution du problème

$$x^* \in \operatorname*{argmin}_{x\in\{0,1\}^{|V|}} \sum_{(i,j)\in E} x_i W_{ij}(1 - x_j).$$

Ce problème de min-cut à deux classes peut se résoudre de manière efficace grâce à l'algorithme de Stoer-Wagner. Mais en pratique l'algorithme renvoie souvent une partition dont une classe se réduit à un seul noeud (le point de plus petit degré, càd, celui le moins connecté au graphe). C'est pas vraiment l'idée qu'on se fait d'une communauté. On va donc forcer les communautés à être de taille suffisante en introduisant une nouvelle fonction de modularité.

Ratiocut problem : Pour éviter les solutions triviales du mincut problem, on introduit une nouvelle fonction de modularité :

$$(V_1, \dots, V_k) \in \mathcal{P}_k(V) \longrightarrow \mathrm{RatioCut}(V_1, \dots, V_k) = \frac{1}{2} \sum_{i=1}^{k} \frac{W(V_i, V_i^c)}{|V_i|} \tag{2}$$

qui force la taille des communautés à ne pas être trop petite. Le ratio $W(V_i, V_i^c)/|V_i|$ est sensé réaliser une balance entre la connectivité de V_i avec son complémentaire et la taille de V_i.

En général minimiser le ratiocut (2) sur toutes les partitions de $\mathcal{P}_k(V)$ est NP-hard. On va alors récrire ce problème sous forme vectoriel (pour $k = 2$) ou matriciel (k général) et montrer qu'une relaxation convexe de ce problème donne la méthode spectrale de détection de communauté vue au chapitre précédent.

Ratiocut pour $k = 2$: On regarde d'abord le problème du ratiocut dans le cas de deux communautés. On cherche alors une solution au problème

$$\min_{A\subset V} \mathrm{RatioCut}(A, A^c). \tag{3}$$

On récrit ce problème sous une forme vectorielle. Pour tout $A \subset V$, on construit un vecteur $f^A = (f_i^A)_{i\in V} \in \mathbb{R}^{|V|}$ définie pour tout $i \in V$ par

$$f_i^A = \begin{cases} \sqrt{\frac{|A^c|}{|A|}} & \text{si } i \in A \\ -\sqrt{\frac{|A|}{|A^c|}} & \text{si } i \notin A. \end{cases} \tag{4}$$

Proposition 3.1. *Soit L le Laplacien d'un graphe pondéré non orienté. On a :*

1) $\langle f^A, \mathbf{1}\rangle = 0$ *et* $\|f^A\|_2^2 = n$
2) $(f^A)^\top L f^A = n\text{RatioCut}(A, A^c)$.

Démonstration. On a

$$\begin{aligned}
(f^A)^\top L f^A &= \frac{1}{2}\sum_{(i,j)\in E} W_{ij}(f_i^A - f_j^A)^2 \\
&= \frac{1}{2}\sum_{\substack{(i,j)\in E\\ i\in A, j\in A^c}} W_{ij}\left(\sqrt{\frac{|A^c|}{|A|}} + \sqrt{\frac{|A|}{|A^c|}}\right)^2 + \frac{1}{2}\sum_{\substack{(i,j)\in E\\ j\in A, i\in A^c}} W_{ij}\left(-\sqrt{\frac{|A|}{|A^c|}} - \sqrt{\frac{|A^c|}{|A|}}\right)^2 \\
&= \frac{1}{2}\sum_{\substack{(i,j)\in E\\ i\in A, j\in A^c}} W_{ij}\left(\frac{n}{|A|} + \frac{n}{|A^c|}\right) + \frac{1}{2}\sum_{\substack{(i,j)\in E\\ j\in A, i\in A^c}} W_{ij}\left(\frac{n}{|A^c|} + \frac{n}{|A|}\right) \\
&= \frac{n}{2|A|}\sum_{\substack{(i,j)\in E\\ (i,j),(j,i)\in A\times A^c}} W_{ij} + \frac{n}{2|A^c|}\sum_{\substack{(i,j)\in E\\ (i,j),(j,i)\in A\times A^c}} W_{ij} = n\frac{W(A,A^c)}{|A|} + n\frac{W(A^c,A)}{|A^c|} \\
&= n\text{RatioCut}(A, A^c)
\end{aligned}$$

où on a utilisé le calcul

$$\begin{aligned}
\left(\sqrt{\frac{|A^c|}{|A|}} + \sqrt{\frac{|A|}{|A^c|}}\right)^2 &= \frac{|A^c|}{|A|} + \frac{|A|}{|A^c|} + 2\sqrt{\frac{|A^c||A|}{|A||A^c|}} \\
= \frac{|A^c|}{|A|} + \frac{|A|}{|A^c|} + 2 &= \frac{|A^c| + |A|}{|A|} + \frac{|A| + |A^c|}{|A^c|} = \frac{n}{|A^c|} + \frac{n}{|A|}.
\end{aligned}$$

De plus, on a

$$\|f^A\|_2^2 = \sum_{i\in A}\frac{|A^c|}{|A|} + \sum_{i\in A^c}\frac{|A|}{|A^c|} = |A| + |A^c| = n$$

et aussi $\langle f^A, \mathbf{1}\rangle = 0$ car

$$\sum_{i\in V} f_i^A = \sum_{i\in A}\sqrt{\frac{|A^c|}{|A|}} - \sum_{i\in A^c}\sqrt{\frac{|A|}{|A^c|}} = \sqrt{|A||A^c|} - \sqrt{|A||A^c|} = 0.$$

■

Il y a donc équivalence entre les trois problèmes :

$$\min_{A\subset V}(f^A)^\top L f^A, \tag{5}$$

$$\min_{A\subset V}\left((f^A)^\top L f^A : \|f^A\|_2^2 = n, \langle f^A, \mathbf{1}\rangle = 0\right) \tag{6}$$

et le ratiocut problem

$$\min_{A\subset V}\text{RatioCut}(A, A^c). \tag{7}$$

Les trois problèmes sont des problèmes d'optimisation discrète. Comme (7) est NP-hard en général c'est aussi le cas pour les deux autres (5) et (6). On va alors chercher une relaxation convexe pour

(6). Le problème ici est que l'espace de recherche "$A \subset V$" est discret. On va alors le "convexifier" simplement en le remplaçant par $\mathbb{R}^{|V|}$. On considère alors le problème "relaché de (6)" :

$$\min_{f \in \mathbb{R}^{|V|}} \left(f^\top L f : \|f\|_2^2 = 1, \langle f, \mathbf{1} \rangle = 0 \right). \tag{8}$$

Montrons que (8) est bien la méthode spectrale introduite au chapitre précédent. Comme L est symétrique, on peut trouver une base orthonormale de vecteurs propres. Or on sait que $\mathbf{1} = (1)_{i \in V}$ est vecteur propre associé à la valeur propre 0 de L et que, comme L est positive (voir Proposition 2.2), 0 est la plus petite valeur propre de L, (8) revient donc à chercher un vecteur propre associé à la deuxième plus petite valeur propre de L (qui peut être aussi 0 si 0 est de multiplicité plus grande que 2 dans le spectre de L).

Définition 3.2. *Soit L le Laplacien d'un graphe pondéré non orienté. Le **Fiedler vector** de L est un vecteur propre associé à la deuxième plus petite valeur propre de L. On appelle aussi la deuxième plus petite valeur propre de L la **connectivité algébrique**.*

Proposition 3.3. *Les vecteurs de Fiedler de L sont les solutions du problème (8).*

Démonstration. Comme L est symétrique, on peut écrire $L = UDU^\top$ où U est une matrice orthogonale ayant pour vecteurs colonnes les vecteurs propres de L noté $(u_i)_{i \in V}$ et $D = \mathrm{diag}(\lambda)$ où $\lambda = (\lambda_i)_{i \in V}$ est le spectre de L. On note $V = \{1, \dots, n\}$ et $\lambda = (\lambda_i)_{i=1}^n$ tel que $0 = \lambda_1 \leq \lambda_2 \leq \cdots \leq \lambda_n$. On a $u_1 = \mathbf{1}$ et $\mathrm{vect}(u_2, \cdots, u_n) = \mathrm{vect}(u_1)^\perp$.

Pour tout $f \in \mathbb{R}^{|V|}$ tel que $\|f\|_2^2 = 1$ et $\langle f, \mathbf{1} \rangle = 0$, on a

$$\begin{aligned} f^\top L f &= f^\top \left(\sum_{i \in V} \lambda_i u_i u_i^\top \right) f = \sum_{i \in V} \lambda_i \langle u_i, f \rangle^2 = \sum_{i=2}^n \lambda_i \langle u_i, f \rangle^2 \\ &\geq \lambda_2 \sum_{i=2}^n \langle u_i, f \rangle^2 = \lambda_2 \|f\|_2^2 = \lambda_2 \end{aligned} \tag{9}$$

car $\lambda_1 = 0$ et, comme $\langle f, \mathbf{1} \rangle = 0$ on a $\langle f, u_1 \rangle = 0$, donc $1 = \|f\|_2^2 = \sum_{i=2}^n \langle u_i, f \rangle^2$. Or pour $f = u_2$, on a $f^\top L f = \lambda_2$ donc cette borne inférieure est atteinte par un deuxième vecteur propre de L. Seuls ces vecteurs propres normalisés peuvent atteindre cette borne sinon il est facile de voir que $f^\top L f > \lambda_2$ en regardant le cas d'égalité dans (9). ∎

On approche donc une solution du RatioCut problem par la recherche d'un deuxième vecteur propre, Fiedler vector, de L. Une fois calculé un Fiedler vector u_2, on doit en déduire des communautés pour G. Comme u_2 est sensé être une valeur approchante du vecteur d'appartenance aux communautés f^{A^*} où $A^* \subset V$ est une solution du RatioCut problem, on espère que les coordonnées de u_2 vont se concentrées autour de deux valeurs (car f^{A^*} prend exactement 2 valeurs). On fait alors un clustering à deux classes C_1, C_2 sur $\mathbb{R}$ des coordonnées de u_2 pour en déduire les communautés $\{i : u_{2i} \in C_1\}$, $\{i : u_{2i} \in C_2\}$ de G. C'est exactement la méthode spectrale dans le cas $k = 2$ quand on prend pour premier vecteur propre $u_1 = \mathbf{1}$.

RatioCut pour k communautés : Dans le cas $k = 2$, à chaque partition $A \sqcup A^c = V$, on a associé une fonction f^A définie dans (4). Ici pour le cas général, à chaque partition, on associe une matrice de la manière suivante. Soit $V_1 \sqcup \cdots \sqcup V_k$ une partition de V, on construit $H := H(V_1, \dots, V_k) \in \mathbb{R}^{|V| \times k}$ tel que ces k vecteurs colonnes sont donnés par $h_1, \dots, h_k$ définis pour tout $i \in V, j = 1, \dots, k$ par

$$h_{ij} = \begin{cases} \frac{1}{\sqrt{|V_j|}} & \text{si } i \in V_j \\ 0 & \text{sinon.} \end{cases}$$

Proposition 3.4. *La matrice H définie ci-dessus à partir d'une partition $V_1 \sqcup \cdots \sqcup V_k$ de V vérifie :*

1) $H^\top H = I_k$

2) $\mathrm{RatioCut}(V_1, \ldots, V_k) = \mathrm{Tr}(H^\top L H)$.

Démonstration. **1)** On a pour tout $p, q = 1, \ldots, k$,

$$(H^\top H)_{pq} = \sum_{i \in V} H_{ip} H_{iq} = \sum_{i \in V} \frac{1}{\sqrt{|V_p||V_q|}} I(i \in V_p) I(i \in V_q).$$

Alors si $p \neq q$, comme $V_p \cap V_q = \emptyset$ on a $I(i \in V_p)I(i \in V_q) = 0$ pour tout $i \in V$ et donc $(H^\top H)_{pq} = 0$. Quand $p = q$, on a $I(i \in V_p)I(i \in V_q) = 1$ pour tout $i \in V_p$ et sinon $I(i \in V_p)I(i \in V_q) = 0$ donc

$$(H^\top H)_{pp} = \sum_{i \in V_p} \frac{1}{|V_p|} I(i \in V_p) = 1.$$

On a donc bien $H^\top H = I_k$.

2) On rappelle que $h_1, \ldots, h_k$ sont les vecteurs colonnes de H. On a

$$\mathrm{Tr}(H^\top L H) = \sum_{p=1}^{k} (H^\top L H)_{pp} = \sum_{p=1}^{k} h_p^\top L h_p.$$

Pour tout $p = 1, \ldots, k$, on a

$$\begin{aligned} h_p^\top L h_p &= \frac{1}{2} \sum_{(i,j) \in E} W_{ij} (h_{ip} - h_{jp})^2 = \frac{1}{2} \sum_{(i,j) \in E} W_{ij} \left(\frac{I(i \in V_p)}{\sqrt{|V_p|}} - \frac{I(j \in V_p)}{\sqrt{|V_p|}} \right)^2 \\ &= \frac{1}{2} \sum_{\substack{(i,j) \in E \\ i,j \in V_p}} W_{ij} \left(\frac{1}{\sqrt{|V_p|}} - \frac{1}{\sqrt{|V_p|}} \right)^2 + \frac{1}{2} \sum_{\substack{(i,j) \in E \\ i \in V_p, j \notin V_p}} \frac{W_{ij}}{|V_p|} + \frac{1}{2} \sum_{\substack{(i,j) \in E \\ j \in V_p, i \notin V_p}} \frac{W_{ij}}{|V_p|} \\ &= \sum_{\substack{(i,j) \in E \\ i \in V_p, j \notin V_p}} \frac{W_{ij}}{|V_p|} = \frac{W(V_p, V_p^c)}{2|V_p|}. \end{aligned}$$

On a donc

$$\mathrm{Tr}(H^\top L H) = \sum_{p=1}^{k} \frac{W(V_p, V_p^c)}{2|V_p|} = \mathrm{RatioCut}(V_1, \cdots, V_k).$$

■

Ainsi d'après la Proposition 3.4, il y a équivalence entre les trois problèmes :

$$\min_{\substack{V_1 \sqcup \cdots \sqcup V_k = V \\ H = H(V_1, \ldots, V_k)}} \mathrm{Tr}(H^\top L H), \tag{10}$$

$$\min_{\substack{V_1 \sqcup \cdots \sqcup V_k = V \\ H^\top H = I_k}} \mathrm{Tr}(H^\top L H) \tag{11}$$

et

$$\min_{V_1 \sqcup \cdots \sqcup V_k = V} \mathrm{RatioCut}(V_1, \ldots, V_k). \tag{12}$$

Ces deux problèmes sont en général NP-hard. On va alors effectuer une relaxation convexe de la contrainte de (11) par

$$\min_{\substack{H\in\mathbb{R}^{|V|\times k}\\ H^\top H=I_k}} \mathrm{Tr}(H^\top LH). \tag{13}$$

En écrivant la SVD de $L = UDU^\top$ où $D = \mathrm{diag}(\lambda)$, $\lambda = (\lambda_i)_{i\in V}$ est le spectre de L tel que $0 = \lambda_1 \leq \lambda_2 \leq \cdots \leq \lambda_n$ et $U = [u_1|\cdots|u_n]$ quand $V = \{1,\ldots,n\}$, on a

$$\mathrm{Tr}(H^\top LH) = \sum_{i\in V} \lambda_i \|a_i\|_2^2$$

où, pour tout $i \in V$, a_i est le i-ième vecteur ligne de $U^\top H$ càd $a_i = (\langle u_i, h_j\rangle)_{j=1,\ldots,k}$. Par ailleurs, comme $(h_1,\ldots,h_k)$ est une famille de vecteurs orthonormaux (car $H^\top H = I_k$), on a

$$\sum_{i\in V} \|a_i\|_2^2 = \sum_{j=1}^{k}\sum_{i\in V}\langle u_i, h_j\rangle^2 = \sum_{j=1}^{k}\|h_j\|_2^2 = k,$$

$\|a_i\|_2^2 \leq \|u_i\|_2^2 = 1$ pour tout $i \in V$ et que

$$\min_{0\leq c_i\leq 1:\sum_i c_i = k} \sum_{i\in V}\lambda_i c_i = \sum_{i=1}^{k}\lambda_i$$

qui est atteint lorsque $\mathrm{vect}(h_1,\ldots,h_k) = \mathrm{vect}(u_1,\ldots,u_k)$. On voit donc que les solutions de (13) sont les matrices H orthogonales dont les colonnes engendrent le sev de dimension k associé à k plus petites valeurs propres de L.

On retombe bien sur le problème de trouver k vecteurs propres de L associés aux k plus petites valeurs propres de L. Ensuite, on clusterise les n vecteurs lignes de H en k clusters dont on déduit les communautés de L.

3.3 Pseudo-algorithme de la méthode spectrale à k communautés

Soit G un graphe non orienté et pondéré ou pas. On note par A sa matrice d'adjacence et par $L = D - A$ son Laplacien où D est la matrice diagonale des degrés des nœuds de G. La méthode spectrale pour la détection de k communautés dans G a pour pseudo-code :

1 **Input :** L Laplacien du graphe G à n nœuds et k le nombre de communautés
2 **Output :** Partition des nœuds de G en k communautés
3 Trouver une base orthogonale de k vecteurs propres de L associés aux plus petites valeurs propres de L.
4 On note H la matrice de taille $n \times k$ ayant pour vecteurs colonnes les k vecteurs propres de L précédemment calculés.
5 On clusterise en k clusters les n vecteurs lignes $y_1,\ldots,y_n \in \mathbb{R}^k$ de H : $C_1,\ldots,C_k$
6 On retourne les k communautés $\{i \in \{1,\ldots,n\} : y_i \in C_p\}, p = 1,\ldots,k$.

Algorithm 1: Méthode spectrale pour le détection de communautés basée sur l'étude du spectre du Laplacien du graphe.

4 Modèles probabilistes de graphes et méthodes spectrales pour la détection de communautés

Idée : Dans les chapitres précédents, on a supposé que tout le graphe est observé. Il y a cependant des cas où certains liens n'ont pas pu être observés ou établis avec certitude (liens entre gènes où compagnies, etc.). Dans ce cas, une possibilité, est de supposer un modèle probabiliste sous-jacent à nos observations. En quelques sorte on observe que partiellement la matrice d'adjacence d'un graphe mais on souhaite toujours identifier une structure de communautés au sein de ce graphe. On va ici utiliser la méthode spectrale des sections précédentes mais appliquées seulement au Laplacien de la matrice d'adjacence partiellement observée.

4.1 Modèles probabilistes de graphes

On présente dans cette section deux modèles probabilistes de graphes aléatoires.

Le modèle de Erdös-Rényi : On dit que le graphe aléatoire G suit le modèle d'Erdös-Rényi à n nœuds et de paramètre p, et on note $G \sim G(n,p)$, quand G est un graphe non orienté non pondéré sur n nœuds dont la matrice d'adjacence est donnée par $A = (\delta_{ij})_{1\leq i,j\leq n}$ où $\delta_{ij} = \delta_{ji}$, $\delta_{ii} = 1$ et $(\delta_{ij} : j > i)$ sont des variables aléatoires de Bernoulli de paramètre p indépendantes.

Stochastic Block Model (SBM) : On dit qu'un graphe aléatoire suit le SBM sur n noeuds et de paramètres p, q où $0 \leq q < p \leq 1$, et on note $G \sim G(n,p,q)$ quand $G = (V,E)$ est un graphe non orienté et non pondéré de matrice d'adjacence $A = (\delta_{ij})_{1\leq i,j\leq n}$ telle qu'il existe une partition $V_1 \sqcup V_1^c = V$ des sommets de G pour laquelle on a pour tout $i < j$

$$\delta_{ij} \sim \begin{cases} \text{Bern}(p) & \text{si } i,j \in V_1 \text{ ou } i,j \in V_1^c \text{ et } i \neq j \\ \text{Bern}(q) & \text{si } (i,j) \in V_1 \times V_1^c \text{ ou } (i,j) \in V_1^c \times V_1 \end{cases} \tag{14}$$

et $\delta_{ij} = \delta_{ji}$ et $\delta_{ii} = 1$.

Autrement dit, dans un SBM, les nœuds appartenant à la même communauté sont connectés avec proba p et s'ils n'appartiennent pas à la même communauté alors ils sont connectés avec probabilité q. Comme $p > q$ on s'attend à ce qu'il y ait une plus forte densité de nœuds intra-communautés qu'en inter-communautés.

Le SBM modélise donc les graphes organisés en deux communautés au sens où on l'a définit au début : une communauté a une plus forte densité de liens en inter qu'en externe avec son complémentaires. Le problème qu'on va chercher à résoudre est le suivant : étant donné un graphe G tiré selon le SBM de paramètre (n,p,q), comment retrouver les deux communautés V_1 et V_1^c de G ?

4.2 Méthodes spectrales pour le SBM à deux classes de même taille basé sur la recherche d'un vecteur de Fiedler

On considère un graphe $G = (V,E)$ tiré selon le SBM de paramètre (n,p,q) avec $p > q$ et $V = \{1,\dots,n\}$. On note par A la matrice d'adjacence (observée) de G comme définie dans (14). Comme vu au chapitre précédent la méthode spectrale consiste ici à trouver un vecteur de Fiedler du Laplacien de G càd de $L = D - A$ où $D = \text{diag}(d_1,\dots,d_n)$ où $d_i = \sum_{j=1}^n A_{ij}$ est le degré du i-ième nœud de G. Une fois un vecteur de Fiedler obtenu on clusterise ces coordonnées en deux groupes qui vont nous donner les deux communauté de G idéalement. C'est ce qu'on aimerait démontrer ici avec grande probabilité.

On réécrit la méthode spectrale plus simplement quand les deux communautés V_1 et V_1^c ont même cardinal grâce au lemme suivant.

Lemme 4.1. *Soit $A \in \mathbb{R}^{n\times n}$ la matrice d'adjacence d'un graphe $G = (V, E)$ tiré selon le SBM de paramètre (n, p, q) où n est pair. On suppose que les deux communautés V_1, V_1^c sous-jacentes au modèle SBM sont de même taille $n/2$. Il y a équivalence entre :*

1) u_2 est un vecteur de Fiedler de $\mathbb{E}L = \mathbb{E}(D - A)$

2) u_2 est un vecteur propre associé à la plus grande valeur propre de

$$\mathbb{E}A - \left[n\left(\frac{p+q}{2}\right) + (1-p)\right]\left(\frac{\mathbf{1}}{\sqrt{n}}\right)\left(\frac{\mathbf{1}}{\sqrt{n}}\right)^\top \tag{15}$$

où $\mathbf{1} = (1)_{i=1}^n$.

Démonstration. Quitte à réordonner les noeuds de V, on peut écrire $\mathbb{E}A$ sous la forme

$$\mathbb{E}A = \left[\begin{array}{cccc|cccc} 1 & p & \cdots & p & q & q & \cdots & q \\ p & 1 & \cdots & p & q & q & \cdots & q \\ \cdots & \cdots & \ddots & \cdots & \cdots & \cdots & \cdots & \cdots \\ p & p & \cdots & 1 & q & \cdots & \cdots & q \\ \hline q & q & \cdots & q & 1 & p & \cdots & p \\ q & q & \cdots & q & p & 1 & \cdots & p \\ \cdots & \cdots & \cdots & \cdots & \cdots & \cdots & \ddots & \cdots \\ q & \cdots & \cdots & q & p & p & \cdots & 1 \end{array}\right].$$

On voit que la matrice moyenne des degré est alors proportionnelle à l'identité car pour tout noeud $i \in \{1, \ldots, n\}$, le degré moyen du nœud i est $\mathbb{E}d_i = 1 + (n/2 - 1)p + (n/2)q$. On a donc $\mathbb{E}D = (1 + (n/2 - 1)p + (n/2)q)I_n$.

Comme $\mathbb{E}D$ est proportionnel à l'identité, on voit que f est un vecteur propre de $\mathbb{E}L$ de valeur propre λ si et seulement si c'est un vecteur propre de $\mathbb{E}A$ de valeur propre $[1 + (n/2 - 1)p + (n/2)q] - \lambda$. Ainsi chercher un vecteur de Fiedler pour $\mathbb{E}L$ est équivalent à chercher un vecteur propre pour $\mathbb{E}A$ pour sa deuxième valeur propre.

Par ailleurs, comme 0 est la plus petite valeur propre de $\mathbb{E}L$ (car c'est le Laplacien du graphe dont la matrice d'adjacence est donnée par $\mathbb{E}A$) associé au vecteur propre $\mathbf{1} = (1)_1^n$, $1 + (n/2 - 1)p + (n/2)q$ est la plus grande valeur propre de $\mathbb{E}A$ associé au même vecteur propre $\mathbf{1} = (1)_1^n$. En écrivant la décomposition en valeur singulières de $\mathbb{E}A = \sum_i([1 + (n/2-1)p + (n/2)q] - \lambda_i)u_i \otimes u_i$, où $0 = \lambda_1 \leq \lambda_2 \leq \cdots \leq \lambda_n$ sont les valeurs propres de $\mathbb{E}L$ associées aux vecteurs propres respectifs $u_1, \ldots, u_n$ (formant une base orthonormale de $\mathbb{R}^n$), on voit que $\mathbb{E}A - [1+(n/2-1)p+(n/2)q]u_1 \otimes u_1$ a pour plus grande valeur propre $[1 + (n/2 - 1)p + (n/2)q] - \lambda_2$ qui est la deuxième valeur propre la plus grande de $\mathbb{E}A$. Il est donc équivalent de chercher un vecteur propre de $\mathbb{E}A$ associé à sa deuxième plus grande valeur propre que de chercher un vecteur propre associé à la plus grande valeur propre de $\mathbb{E}A - [1 + (n/2 - 1)p + (n/2)q]u_1 \otimes u_1$. Hors, on peut prendre $u_1 = \mathbf{1}/\|\mathbf{1}\|_2$. Ce qui conclut la preuve. ∎

Dans le cas de deux communautés de même taille, la méthode spectrale consiste donc, d'après le Lemme 4.1, à chercher un vecteur propre associé à la plus grande valeur propre de la matrice (15) (et ensuite à clusteriser ses coordonnées).

Cependant, on ne connaît ni $\mathbb{E}A$, ni p ni q. On va donc estimer ces quantités : on utilise A pour estimer $\mathbb{E}A$ et pour $(p + q)/2$, on remarque que pour

$$\lambda = \frac{2}{n(n-1)}\sum_{i<j} A_{ij} \tag{16}$$

on a

$$\mathbb{E}(\lambda) = \frac{2}{n(n-1)}\left[\left(\binom{n}{2}^2 - \binom{n}{2}\right)p + \binom{n}{2}^2 q\right] = \frac{p+q}{2} - \frac{p-q}{n-1}.$$

Ainsi λ estime $(p+q)/2$ à un terme en $1/n$ près. On va alors estimer un plus grand vecteur propre de (15) par un plus grand vecteur propre de $A - n\lambda(\mathbf{1}/\sqrt{n})(\mathbf{1}/\sqrt{n})^\top$ càd une solution au problème

$$\max_{x:\|x\|_2\leq 1} \langle x, (A - \lambda J_n)x\rangle \tag{17}$$

où $J_n = (1)_{n\times n}$.

4.3 Méthodes spectrales pour le SBM à deux classes de même taille basé sur une relaxation SDP

Le problème de détection de communautés s'écrit, initialement, comme un problème de recherche d'une solution au problème discret

$$\min_{A\subset V}\left((f^A)^\top \mathbb{E}L f^A : \|f^A\|_2^2 = 1, \langle f^A, \mathbf{1}\rangle = 0\right) \tag{18}$$

initialement introduit dans (6) où $\mathbb{E}L$ était entièrement connu (et pas seulement au travers d'une seule observation L). Ce problème a été relaxé en supprimant la contrainte que f devait être de la forme f^A pour un certain $A \subset V$ (où f^A est défini dans (4)). Cette relaxation a motivé l'introduction de la recherche d'un vecteur de Fiedler du Laplacien. On peut cependant proposer d'autres relaxations convexes.

Par exemple, dans le cas de deux classes de même taille, on a vu que chercher un vecteur de Fiedler de $\mathbb{E}L$ est équivalent à chercher un premier vecteur propre de (15) qui peut être estimé par $A - \lambda J_n$ où λ est défini par (16) et $J_n = (1)_{n\times n}$. On peut alors espérer qu'une solution au problème

$$\max_{A\subset V}\left((f^A)^\top (A - \lambda J_n) f^A\right) \tag{19}$$

peut être une bonne solution approchante d'une solution du problème initial (18). Par ailleurs, sous l'hypothèse de communautés de taille égale, on voit que $f^A \in \{-1,1\}^n$ est tel que $(f^A)_i = 1$ si $i \in V_1$ et $(f^A)_i = -1$ si $i \notin V_1$. Le problème (19) peut donc se réécrire comme

$$\max_{x\in\{-1,1\}^n}\left(x^\top (A - \lambda J_n)x\right) \tag{20}$$

qui est lui-même équivalent à

$$\max_{x\in\{-1,1\}^n}\langle A - \lambda J_n, xx^\top\rangle \tag{21}$$

où on utilise ici le produit scalaire entre deux matrices donné par $\langle A, B\rangle = \sum_{i,j} A_{ij}B_{ij} = \mathrm{Tr}(AB^\top)$. Ce problème est un problème combinatoire et donc potentiellement difficile à résoudre directement. On va utiliser une relaxation convexe de type SDP pour l'approcher. Pour cela, on voit $xx^\top$ dans le problème (21) comme une variable matricielle ayant les propriété suivantes :

i) $xx^\top = (x_i x_j)_{1\leq i,j\leq n}$ est symétrique
ii) $xx^\top$ est positive car $\langle xx^\top y, y\rangle = \langle x, y\rangle^2 \geq 0$ pour tout $y \in \mathbb{R}^n$
iii) $\mathrm{diag}(xx^\top) = \mathrm{diag}(x_1^2, \ldots, x_n^2) \preceq I_n$ car $x_i^2 = 1$ pour tout $n = 1, \ldots, n$.

(remarque : on a aussi $\text{rang}(xx^\top) = 1$ mais c'est une propriété qu'on ne peut pas facilement implémenter et qui est souvent source de complexité algorithmique ; on ne va donc pas la garder pour construire un problème SDP).

On va alors faire une relaxation convexe de (21) en cherchant une solution au problème

$$\max_{\substack{Z \succeq 0 \\ \text{diag}(Z) \preceq I_n}} \langle A - \lambda J_n, Z \rangle \tag{22}$$

où on rappelle que $Z \succeq 0$ signifie que Z est symétrique et que $\langle Zy, y \rangle \geq 0$ pour tout $y \in \mathbb{R}^n$. On utilise aussi que $A \succeq B$ quand $A - B \succeq 0$.

Si $\hat{Z}$ est solution de (22), on prend ensuite un plus grand vecteur propre (un vecteur propre associé à sa plus grande valeur propre) de $\hat{Z}$ dont on prend le signe. Les coordonnées de signe 1 forment une communauté et les autres de signe -1 forment l'autre communauté. On espère que cette procédure puisse bien estimer une solution $x^* \in \{-1, 1\}^n$ du problème (20) mais aussi et surtout du problème initial (18) avec grange probabilité. C'est l'objet des sections suivantes de le démontrer.

On s'assure d'abord que le vecteur d'appartenance aux communautés $\bar{x} \in \{-1, 1\}^n$ défini par $\bar{x}_i = 1$ si $i \in V_1$ et $\bar{x}_i = -1$ si $i \notin V_1$ est bien tel que $\bar{x}(\bar{x})^\top$ est l'unique solution du problème relaxé

$$\max_{\substack{Z \succeq 0 \\ \text{diag}(Z) \preceq I_n}} \langle \mathbb{E}A - \alpha u_1 \otimes u_1, Z \rangle. \tag{23}$$

où $\alpha = [(p+q)/2]n + (1-p)$ et $u_1 = \mathbf{1}/\sqrt{n}$. On rappelle aussi que $u \otimes v = uv^\top = (u_i v_j)_{1 \leq i,j \leq n} \in \mathbb{R}^{n \times n}$.

Proposition 4.2. *Le problème (23) admet une unique solution donnée par $\bar{x} \otimes \bar{x}$.*

Démonstration. On note $B = \mathbb{E}A - \alpha u_1 \otimes u_1$. On commence par étudier le spectre de B. On a

$$B = \left[\begin{array}{c|c} p & q \\ \hline q & p \end{array}\right] - \left[n\left(\frac{p+q}{2}\right) + (1-p)\right] u_1 \otimes u_1 + (1-p) I_n.$$

On note $u_2 = \bar{x}/\sqrt{n}$. On a

$$\left[\begin{array}{c|c} p & q \\ \hline q & p \end{array}\right] = n\left(\frac{p+q}{2}\right) u_1 \otimes u_1 + n\left(\frac{p-q}{2}\right) u_2 \otimes u_2.$$

Comme $\langle u_1, u_2 \rangle = 0$ et que $\|u_1\|_2 = \|u_2\| = 1$, (u_1, u_2) forme le début d'une base orthonormale qu'on peut compléter : soit $(u_i)_{i=3}^n$ tel que $(u_i)_{i=1}^n$ forme une base orthonormale de $\mathbb{R}^n$. On écrit

$$(1-p) I_n = (1-p) u_1 \otimes u_1 + (1-p) u_2 \otimes u_2 + (1-p) \sum_{i=3}^n u_i \otimes u_i.$$

On en déduit que

$$B = \left[n\left(\frac{p-q}{2}\right) + (1-p)\right] u_2 \otimes u_2 + (1-p) \sum_{i=3}^n u_i \otimes u_i.$$

On voit ensuite grâce au théorème spectral qu'il y a équivalence entre les deux problèmes :
1)

$$\bar{Z} \in \operatorname*{argmax}_{Z \succeq 0, \text{diag}(Z) \preceq I_n} \langle B, Z \rangle \tag{24}$$

2) $\bar{Z} = \bar{X}(\bar{X})^\top$ et

$$\bar{X} \in \underset{\substack{X\in\mathbb{R}^{n\times n}\\ X_{i\bullet}\in B_2^n}}{\operatorname{argmax}} \left\langle B, XX^\top \right\rangle \tag{25}$$

où, pour tout $i = 1, \dots, n$, $X_{i\bullet}$ est le $i-$ième vecteur ligne de X et $X_{i\bullet} \in B_2^n$ signifie que $\|X_{i\bullet}\|_2 \leq 1$.

Soit $X \in \mathbb{R}^{n\times n}$ tel que pour tout $i = 1, \dots, n$ on a $X_{i\bullet} \in B_2^n$. On note par $X_{\bullet j}$ le j-ième vecteur colonne de X. On a

$$XX^\top = \left(\sum_{k=1}^n X_{ik}X_{jk}\right)_{1\leq i,j\leq n} = \sum_{k=1}^n X_{\bullet k} \otimes X_{\bullet k}$$

car $X_{\bullet k} \otimes X_{\bullet k} = (X_{ik}X_{jk})_{1\leq i,j\leq n}$. On remarque que $\left\langle u\otimes u, v\otimes v\right\rangle = \operatorname{Tr}(uu^\top vv^\top) = \left\langle u, v\right\rangle^2$ pour tout $u, v \in \mathbb{R}^n$. Comme $p > q$ et que $(u_i)_{i=1}^n$ est une base orthonormale de $\mathbb{R}^n$, on a

$$\begin{aligned}
&\left\langle B, XX^\top \right\rangle \\
&= \left\langle \left[n\left(\frac{p-q}{2}\right) + (1-p)\right] u_2 \otimes u_2 + (1-p)\sum_{i=3}^n u_i \otimes u_i, \sum_{k=1}^n X_{\bullet k} \otimes X_{\bullet k} \right\rangle \\
&= \left[n\left(\frac{p-q}{2}\right) + (1-p)\right] \sum_{k=1}^n \left\langle u_2, X_{\bullet k}\right\rangle^2 + (1-p)\sum_{i=3}^n\sum_{k=1}^n \left\langle u_i, X_{\bullet k}\right\rangle^2 \\
&\leq \left[n\left(\frac{p-q}{2}\right) + (1-p)\right] \sum_{i=2}^n\sum_{k=1}^n \left\langle u_i, X_{\bullet k}\right\rangle^2 \\
&\leq \left[n\left(\frac{p-q}{2}\right) + (1-p)\right] \sum_{i=1}^n\sum_{k=1}^n \left\langle u_i, X_{\bullet k}\right\rangle^2 \\
&\leq \left[n\left(\frac{p-q}{2}\right) + (1-p)\right] \sum_{k=1}^n \|X_{\cdot k}\|_2^2 = \left[n\left(\frac{p-q}{2}\right) + (1-p)\right] \sum_{i=1}^n \|X_{i\bullet}\|_2^2 \\
&= n\left[n\left(\frac{p-q}{2}\right) + (1-p)\right]
\end{aligned} \tag{26}$$

Par ailleurs, en explorant le cas d'égalité dans la majoration de $\left\langle B, XX^\top \right\rangle$ dans (26), on voit que nécessairement $\left\langle u_i, X_{\bullet k}\right\rangle = 0$ pour tout $i = 1$ et $i = 3, \dots, n$ et $k = 1, \dots, n$. On doit aussi avoir $\|X_{i\bullet}\|_2 = 1$ pour tout $i = 1, \dots, n$. Donc les colonnes $X_{\bullet k}$ de X sont nécessairement portées par u_2. En écrivant X comme étant de la forme $[a_1u_2|a_2u_2|\cdots|a_nu_2]$ pour des réels $a_1, \dots, a_n$. On a $\|X_{i\bullet}\|_2 = 1$ pour tout $i = 1, \dots, n$ si et seulement si $\left(\sum_{i=1}^n a_i^2\right)^{1/2} |u_{2i}| = 1$ pour tout $i = 1, \dots, n$ et comme $|u_{2i}| = 1/\sqrt{n}$, on a donc nécessairement $\left((1/n)\sum_{i=1}^n a_i^2\right)^{1/2} = 1$.

D'un autre côté, on voit que pour $\bar{X}$ ayant pour vecteurs colonnes $a_iu_2, i = 1, \dots, n$ pour des réels $a_1, \dots, a_n$ tels que $\left((1/n)\sum_{i=1}^n a_i^2\right)^{1/2} = 1$, on a $\bar{X}_{i\bullet} = (a_1u_{2i}, a_2u_{2i}, \cdots, a_nu_{2i})^\top \in \mathbb{R}^n$ est tel que $\left\|\bar{X}_{i\bullet}\right\|_2 = \left(\sum_{i=1}^n a_i^2\right)^{1/2} |u_{2i}| = 1$ pour tout $i = 1, \dots, n$ car $|u_{2i}| = 1/\sqrt{n}$. Donc $\bar{X}$ est bien dans l'ensemble de contrainte de (25). Par ailleurs, $\bar{X}(\bar{X})^\top = (\sum_i a_i^2)u_2 \otimes u_2 = nu_2 \otimes u_2$ alors

$$\left\langle B, \bar{X}(\bar{X})^\top \right\rangle = n\left[n\left(\frac{p-q}{2}\right) + (1-p)\right].$$

car $\left\langle u_2\otimes u_2, u_2\otimes u_2\right\rangle = \left\langle u_2, u_2\right\rangle^2 = 1$. Donc la borne $n\left[n\left((p-q)/2\right) + (1-p)\right]$ est atteinte par $\bar{X}$ dans (26).

Donc les solutions de (25) sont toutes de la forme $\bar{X} = [a_1 u_2 | \cdots | a_n u_2]$ où $\sum_i a_i^2 = n$. Dans ce cas, les solution de (24) sont de la forme $\bar{Z} = \bar{X}(\bar{X})^\top = (\sum_i a_i^2) u_2 \otimes u_2 = n u_2 \otimes u_2 = \bar{x} \otimes \bar{x}$ qui est donc unique. ■

Comme $\bar{x} \otimes \bar{x}$ est de rang 1 et que son unique espace propre associé à la valeur propre n est engendré par $\bar{x}$, on voit bien qu'en prenant le signe de $\lambda \bar{x}$ (un plus grand vecteur propre de $\bar{x} \otimes \bar{x}$) pour tout $\lambda \in \mathbb{R}$ on obtient $\bar{x}$ ou $-\bar{x}$. On peux donc retrouver les communautés du graphes en prenant le signe d'un plus grand vecteur propre d'une solution au problème (23). Ceci motive la méthode spectrale qui consiste à prendre le signe d'un plus grand vecteur propre d'une solution du problème (23) où $\mathbb{E}A - \alpha u_1 \otimes u_1$ a été remplacé par son estimation $A - \lambda J_n$. On écrit cette méthode spectrale en pseudo-code :

1 **Input :** A : matrice d'adjacence d'un graphe distribué selon le SBM à deux communautés de même taille.
2 **Output :** Partition des nœuds de G en 2 communautés
3 Calcul de λ donné dans (16)
4 Résolution du problème SDP de (22)
5 Recherche d'un plus grand vecteur propre $\hat{x}$ de $\hat{Z}$ solution de (22)
6 On prend le signe de $\hat{x}$
7 On retourne les 2 communautés $\hat{V}_1 = \{i \in \{1, \dots, n\} : \text{sign}(\hat{x}_i) = 1\}$ et son complémentaire $\hat{V}_1^c$.

Algorithm 2: Méthode spectrale pour le détection de communautés basée sur une relaxation SDP.

5 Analyse de la convergence de la méthode spectrale obtenue par relaxation SDP

Problème : On observe la matrice d'adjacence d'un graphe distribué selon un SBM à deux communautés de même taille. On souhaite retrouver ces deux communautés à partir de A. On a introduit dans la section précédente une méthode spectrale basée sur une relaxation SDP du problème d'origine (voir Algorithme 2). L'objectif de cette section est de prouver que cet algorithme fournit un bon estimateur du vecteur d'appartenance $\bar{x}$.

On rappelle quelques notations : λ est défini dans (16),

$$\hat{Z} \in \underset{\substack{Z \succeq 0 \\ \text{diag}(Z) \preceq I_n}}{\text{argmax}} \langle A - \lambda J_n, Z \rangle \tag{27}$$

où $J_n = (1)_{n \times n}$ et $\hat{x}$ un plus grand vecteur propre de $\hat{Z}$. On veut montrer que $\hat{x}$ est proche de $\bar{x}$ le vecteur d'appartenance aux deux communautés V_1 et V_1^c : $\bar{x}_i = 1$ quand $i \in V_1$ et $\bar{x}_i = -1$ quand $i \notin V_1$.

Théorème 5.1. *Soit $\epsilon \in (0,1)$ tel que $n \geq 10^4 \epsilon$. Soit A la matrice d'adjacence d'un graphe distribué selon le SBM de paramètre (n,p,q) où $p > q$. On suppose que $\max(p(1-p), q(1-q)) \geq 20/n$. On suppose que $p \geq a/n$, $q \geq b/n$ et $(a-b)^2 \geq 10^4 \epsilon^{-2}(a+b)$. Soit $\hat{Z}$ une solution de (27). Avec probabilité au moins $1 - e^3 5^{-n}$, on a*

$$\left\| \hat{Z} - \bar{x}(\bar{x})^\top \right\|_2^2 \leq \epsilon n^2 = \epsilon \left\| \bar{x}(\bar{x})^\top \right\|_2^2.$$

Ensuite, on passe de l'estimation de la matrice $\bar{x}(\bar{x})^\top$ en norme 2 à l'estimation d'un plus grand vecteur propre de $\bar{x}(\bar{x})^\top$ de la manière suivante.

Corollaire 5.2. *Sous les hypothèse du Théorème 5.1. Si $\hat{x}$ est un plus grand vecteur propre de $\hat{Z}$ tel que $\|\hat{x}\|_2 = \sqrt{n}$ alors*

$$\min_{\alpha=\pm 1} \|\alpha\hat{x} - \bar{x}\|_2^2 \leq 8\epsilon n = 8\epsilon \|\bar{x}\|_2^2 .$$

La preuve du Corollaire 5.2 s'appuie sur le théorème de Davis-Kahane aussi connu sous le nom de "$\sin\theta$-theorem" qu'on rappelle maintenant.

Théorème 5.3 (Davis-Kahan)**.** *Soit A et B deux matrices symétriques de même dimensions. Soit i. On suppose que la i-ième plus grande valeur propre $\lambda_i(A)$ de A est bien séparée du reste du spectre de A :*

$$\min_{j:j\neq i} |\lambda_i(A) - \lambda_j(A)| = \delta > 0.$$

Si $u_i(A)$ est un vecteur propre unitaire de A associé à la valeur propre $\lambda_i(A)$ et que $u_i(B)$ est un vecteur propre unitaire associé à la $i-$ième plus grande valeur propre de B, on a

$$\min_{\alpha\in\pm 1} \|u_i(A) - \alpha u_i(B)\|_2 \leq 2^{3/2} \frac{\|A-B\|_2}{\delta}.$$

Démonstration du Corollaire 5.2. On applique Davis-Kahan à $A = \bar{x}(\bar{x})^\top$ et $B = \hat{Z}$. Comme A est de rang 1 et que sa plus grande valeur propre est $\lambda_1(A) = n$, le spectral gap de A est $\delta = n > 0$. On a alors d'après le Théorème de Davis-Kahan que

$$\min_{\alpha\in\pm 1} \|u_1(A) - \alpha u_1(B)\|_2 \leq 2^{3/2} \frac{\|A-B\|_2}{\delta}.$$

où $u_1(A) = \bar{x}/\sqrt{n}$ et $u_1(B) = \hat{x}/\sqrt{n}$. Autrement dit,

$$\min_{\alpha=\pm 1} \|\alpha\hat{x} - \bar{x}\|_2 \leq 2^{3/2}\sqrt{n} \left\|\hat{Z} - \bar{x}(\bar{x})^\top\right\|_2 \leq 2^{3/2}\sqrt{n\epsilon}$$

où on a utilisé le Théorème 5.1 dans la dernière inégalité. ∎

5.1 Schéma de la preuve du Théorème 5.1

On note

$$\mathcal{M}_{opt} = \left\{Z \in \mathbb{R}^{n\times n} : Z \succeq 0, \mathrm{diag}(Z) \preceq I_n\right\}$$

l'ensemble de contrainte de la procédure SDP (27). On veut montrer que

$$\hat{Z} \in \operatorname*{argmax}_{Z\in\mathcal{M}_{opt}} \langle A - \lambda J_n, Z\rangle$$

est proche de $\bar{Z} = \bar{x}(\bar{x})^\top$.

On va procéder en 3 étapes :

1) on montre que $\bar{Z} \in \mathrm{argmax}_{Z\in\mathcal{M}_{opt}} \langle \mathbb{E}(A - \lambda J_n), Z\rangle$ en modifiant très légèrement la preuve de la Proposition 4.2 vu que $\mathbb{E}\lambda J_n = [(p+q)/2]J_n$ est presque égale à $\alpha u_1 \otimes u_1 = [(p+q)/2 + (1-p)/n]J_n$.

2) On montre que $(A - \lambda J_n)$ est proche de $\mathbb{E}(A - \lambda J_n)$ uniformément sur $\mathcal{M}_{opt}$, càd avec probabilité $1 - e^3 5^{-n}$,

$$\sup_{Z\in\mathcal{M}_{opt}} \left|\langle (A - \lambda J_n) - \mathbb{E}(A - \lambda J_n), Z\rangle\right| \leq c_0\epsilon \tag{28}$$

où c_0 est une constante absolue.

3) On montre une inégalité de courbure de la fonction objectif à l'optimum : pour tout $Z \in \mathcal{M}_{opt}$,

$$\left\langle \mathbb{E}(A - \lambda J_n), \bar{Z} \right\rangle - \left\langle A - \lambda J_n, Z \right\rangle \geq c_1 \left\| Z - \bar{Z} \right\|_1 \tag{29}$$

Les deux points importants sont **2)** et **3)**. Tous l'aspect probabiliste du problème se trouve dans le point **2)**. On commence par détailler ce point-là.

5.2 Aspect probabiliste de la preuve

On montre dans cette section le point **2)** du schéma de la preuve du Théorème 5.1. Cet argument s'appuie sur l'inégalité de Grothendieck qu'on rappelle maintenant d'abord sous sa forme générale puis sous sa forme matricielle.

Théorème 5.4 (Inégalité de Grothendieck)**.** *Soit $B = (b_{ij})_{i,j}$ une matrice de $\mathbb{R}^{p \times q}$. On suppose que*

$$\left| \sum_{i,j} b_{ij} s_i t_j \right| \leq 1$$

pour tout $s_i, t_j \in \{-1, 1\}$ alors, pour tout espace de Hilbert H et tous vecteurs $u_i, v_j \in H$ tels que $\|u_i\|_2 \leq 1, \|v_j\|_2 \leq 1$ on a

$$\left| \sum_{i,j} b_{ij} \langle u_i, v_j \rangle \right| \leq K_G$$

où $K_G \leq 1.783$.

Théorème 5.4 est la forme la plus connue de l'inégalité de Grothendieck. Il en existe de multiple formulations et des généralisation, comme celle du théorème de Nesterov très utilisé pour trouver des solutions approchantes à des problèmes combinatoires grâce à des relaxation convexes menant à des SDP. La forme de l'inégalité de Grothendieck que nous allons utiliser est une reformulation sous forme matricielle du Théorème 5.4 dans le cas carré $p = q = n$ et pour $H = \mathbb{R}^n$. On note

$$\mathcal{M}_1 = \{st^\top : s, t \in \{-1, 1\}^n\} \text{ et } \mathcal{M}_G = \{XY^\top \in \mathbb{R}^{n \times n} : \text{ rows } X_{i\bullet}, Y_{i\bullet} \in B_2^n\}$$

où B_2^n est la boule unité Euclidienne de $\mathbb{R}^n$. Pour tout $B \in \mathbb{R}^{n \times n}$, on a

$$\sum_{i,j} b_{ij} s_i t_j = \langle B, st^\top \rangle \text{ et } \langle B, XY^\top \rangle = \sum_{i,j} b_{ij} \langle X_{i\bullet}, Y_{j\bullet} \rangle$$

On a clairement, $\mathcal{M}_1 \subset \mathcal{M}_G$ alors

$$\sup_{Z \in \mathcal{M}_1} \langle B, Z \rangle \leq \sup_{Z \in \mathcal{M}_G} \langle B, Z \rangle.$$

L'inégalité de Grothendieck montre que l'inégalité inverse est aussi vraie à constante près.

Corollaire 5.5. *Il existe une constante absolue $K_G \leq 1.783$ telle que pour toute matrice $B \in \mathbb{R}^{n \times n}$, on a*

$$\sup_{Z \in \mathcal{M}_G} \langle B, Z \rangle \leq K_G \sup_{Z \in \mathcal{M}_1} \langle B, Z \rangle.$$

On a donc d'après l'inégalité de Grothendieck que pour tout $B \in \mathbb{R}^{n\times n}$,

$$\sup_{Z\in\mathcal{M}_1} \langle B, Z\rangle \le \sup_{Z\in\mathcal{M}_G} \langle B, Z\rangle \le K_G \sup_{Z\in\mathcal{M}_1} \langle B, Z\rangle.$$

On peux écrire le sup sur $\mathcal{M}_1$ comme une norme d'opérateur car pour tout $B \in \mathbb{R}^{n\times n}$

$$\sup_{Z\in\mathcal{M}_1} \langle B, Z\rangle = \sup_{s,t\in B_\infty^n} \langle B, st^\top\rangle = \sup_{s,t\in B_\infty^n} \langle Bt, s\rangle = \sup_{t\in B_\infty^n} \|Bt\|_1 = \|B\|_{\infty\to 1}\,.$$

Pour le problème qu'on cherche à résoudre – contrôler le supremum d'un processus empirique indexé par $\mathcal{M}_{opt}$ – on ne regarde que les matrices symétriques positive de $\mathcal{M}_G$ et on peut démontrer le résultat suivant.

Proposition 5.6. *On a $\mathcal{M}_{opt} \subset \mathcal{M}_G$.*

Démonstration. Si $Z \in \mathcal{M}_{opt}$, comme $Z \succeq 0$, il existe $X \in \mathbb{R}^{n\times n}$ tel que $Z = XX^\top$. Par ailleurs, $\mathrm{diag}(Z) \preceq I_n$. Or $\mathrm{diag}(Z) = \mathrm{diag}(\|X_{1\bullet}\|_2^2, \cdots, \|X_{n\bullet}\|_2^2)$ donc $\|X_{i\bullet}\|_2^2 \le 1$ pour tout $i = 1, \ldots, n$. Alors $Z \in \mathcal{M}_G$. ∎

On déduit de l'inégalité de Grothendieck et de la Proposition 5.6 que

$$\begin{aligned}\sup_{Z\in\mathcal{M}_{opt}} \left|\langle (A-\lambda J_n) - \mathbb{E}(A-\lambda J_n), Z\rangle\right| &\le K_G \sup_{Z\in\mathcal{M}_1} \left|\langle (A-\lambda J_n) - \mathbb{E}(A-\lambda J_n), Z\rangle\right| \\ &\le K_G \|A-\mathbb{E}A\|_{\infty\to 1} + K_G|\lambda - \mathbb{E}\lambda|\,\|J_n\|_{\infty\to 1}\,. \qquad (30)\end{aligned}$$

Il reste donc à majorer $\|A-\mathbb{E}A\|_{\infty\to 1}$ et $|\lambda-\mathbb{E}\lambda|$ avec grande probabilité. Pour cela, on va utiliser l'inégalité de concentration de Bernstein suivie d'une "union bound". On rappelle ici l'inégalité de Bernstein.

Théorème 5.7. *Soit $Z_1, \ldots, Z_m$ des variables aléatoires indépendantes centrées telles que $|Z_i| \le b$ pour tout $i = 1, \ldots, m$ presque sûrement. On note $\sigma^2 = (1/m)\sum_{i=1}^m \mathbb{E}Z_i^2$. Pour tout $t > 0$,*

$$\mathbb{P}\left[\frac{1}{m}\sum_{i=1}^m Z_i \ge t\right] \le \exp\left(\frac{-mt^2}{2\sigma^2 + 2bt/3}\right).$$

L'inégalité de Bernstein et l'union bound donnent le résultat suivant.

Proposition 5.8. *On pose $\bar{p} = [2/(n(n-1))]\sum_{i<j}\mathrm{var}(A_{ij})$. Si $\bar{p} > 9/n$ alors avec probabilité au moins $1 - e^3 5^{-n}$, $\|A-\mathbb{E}A\|_{\infty\to 1} \le 6n(n-1)\sqrt{\bar{p}/n}$.*

Démonstration. On a

$$\|A-\mathbb{E}A\|_{\infty\to 1} = \max_{s_i,t_j\in\{-1,1\}} \sum_{i,j}(A_{ij}-\mathbb{E}A_{ij})s_it_j.$$

Soit $s, t \in \{-1,1\}^n$. Comme $A_{ii} = 1$ et $A_{ij} = A_{ji}$, on a

$$\sum_{i,j}(A_{ij}-\mathbb{E}A_{ij})s_it_j = \sum_{i<j}(A_{ij}-\mathbb{E}A_{ij})(s_it_j + s_jt_i)$$

∎