

Diviser pour mieux Régner : Séparation Spectrale des Communautés

ESTEVE Victor, GARCIA Hugo, SALA Raphaël
P. MAILLARD

Juin 2024

Table des matières

1	Introduction	2
1.1	Problématique	2
1.2	Le partitionnement spectral	3
1.3	Résultats non-asymptotique et asymptotique	4
2	Perturbation	4
2.1	Théorème de Davis-Kahan	5
2.2	Caractérisation Spectrale	5
2.3	Démonstration du Théorème 1	7
2.4	Démonstration Proposition 1	8
3	Etude non-asymptotique	9
3.1	Structures Discrètes : ε -Net et ε -Separated	9
3.2	Variable aléatoire sous-gaussienne	11
3.3	Majoration de $\ R_n\ $	14
3.4	Démonstration du Résultat 1	16
4	Étude asymptotique	17
4.1	Théorème de Wigner et loi du demi cercle	17
4.2	Equation canonique K1	18
4.3	Démonstration Propriété 5	19
4.4	Démonstration du Résultat 2	19
4.5	Proportion d'erreur maximale	20
5	Discussion	21

Depuis l'apparition du COVID-19, la recherche de clusters est devenue un domaine prolifique. En effet, comment identifier des clusters au sein d'un groupe d'individus interconnectés ? De nombreux algorithmes ont été développés pour répondre à cette question. Dans ce rapport, nous nous inspirons du modèle de Stochastic Block Model et explorons son application en utilisant des concepts de la théorie spectrale et des matrices aléatoires.

Soit un ensemble de $n > 0$ individus, avec n pair, divisé en deux communautés égales, C_1 et C_2 , telles que $|C_1| = |C_2| = \frac{n}{2}$. Considérons deux probabilités $1 > p > q > 0$. Nous définissons un modèle aléatoire où deux individus sont reliés avec une probabilité p s'ils appartiennent à la même communauté, ou avec une probabilité q s'ils appartiennent à des communautés différentes.

Pour au mieux modéliser le graphe donné, nous numérotons chaque noeud de 1 à n et représentons le graphe par sa matrice d'adjacence.

Remarque 1. *Tout graphe non orienté a sa matrice d'adjacence symétrique.*

Soit $X_n = (x_{ij})_{i,j \in \llbracket 1, n \rrbracket}$ la matrice d'adjacence stochastique et symétrique, relatif à notre problème. Elle est défini par,

La matrice d'adjacence peut prendre différentes formes selon la numérotation des sommets choisis. Néanmoins, toutes les propriétés énoncées restent valides, quelle que soit la configuration de X_n .¹ Ainsi, par souci de simplicité, nous représentons X_n avec les communautés déjà organisées par bloc dans la matrice, c'est-à-dire que les points de C_1 sont numérotés dans l'intervalle $\llbracket 1, \frac{n}{2} \rrbracket$, et les points de C_2 dans $\llbracket \frac{n}{2} + 1, n \rrbracket$. Nous avons donc,

Remarque 2. La diagonale de X_n est systématiquement nulle. En effet, un membre ne peut pas être en relation avec lui-même, ce qui signifie que le graphe associé ne contient pas de boucle.

2

1.2 Le partitionnement spectral

Définition 2. Soit X_n la matrice d'adjacence, nous définissons $\hat{X}_n := \mathbb{E}(X_n)$. Dans le cas organisé présenté plus tôt, nous avons $\hat{X}_n = \begin{pmatrix} p & | & q \\ q & | & p \end{pmatrix} - pI_n$.

\hat{X}_n est particulièrement intéressante puisque les communautés se distinguent immédiatement. Cette remarque se reflète dans son caractère intrinsèque, les valeurs propres.

Propriété 1. Les valeurs propres λ_i et vecteurs propres u_i associés à \hat{X} sont donnés par :

$$\begin{aligned} \lambda_1 &= \frac{p+q}{2}n - p, & \text{avec } u_1 &= (1, \dots, 1)^T \\ \lambda_2 &= \frac{p-q}{2}n - p, & \text{avec } u_2 &= (a_1, a_2, \dots, a_n)^T \text{ tel que } \forall i \in \llbracket 1, n \rrbracket, \begin{cases} a_i = 1 & \text{si } i \in C_1 \\ a_i = -1 & \text{si } i \in C_2 \end{cases} \\ \lambda_3 &= \lambda_4 = \dots = \lambda_n = -p \end{aligned}$$

Démonstration. Nous prouverons seulement la cas avec \hat{X}_n organisé comme vu précédemment, la preuve générale se fait de la même manière.

Soit $(\mu_i)_{1 \leq i \leq n}$ les valeurs propres de $B = \begin{pmatrix} p & | & q \\ q & | & p \end{pmatrix}$.

En multipliant par v_1 et v_2 , nous obtenons respectivement qu'ils sont associés aux valeurs propres

$$\mu_1 = \frac{p+q}{2}n \quad \text{et} \quad \mu_2 = \frac{p-q}{2}n$$

De plus, $\dim(\ker B) = n - 2$: toutes les lignes sont égales soit à la première, soit à la dernière ligne de B . D'où pour tout $i \in \llbracket 3, n \rrbracket$, $\mu_i = 0$.

Enfin en revenant à la définition de \hat{X} ,

$$\hat{X}v_i = \begin{pmatrix} p & | & q \\ q & | & p \end{pmatrix} v_i - pI_n v_i = \mu_i v_i - p v_i = \lambda_i v_i$$

□

Nous remarquons que l'information des communautés peut être facilement retrouvé à partir du vecteur propre u_2 associé à la deuxième valeur propre de \hat{X}_n .

Le problème est que dans un cas concret, nous ne disposons pas de \hat{X}_n dû à la numérotation aléatoire. Néanmoins, nous avons remarqué que même à travers son caractère aléatoire, la matrice X_n , évaluée² à un cas concret, aura la même propriété que la matrice \hat{X}_n pour repérer les communautés. En effet si l'on regarde v_2 , le vecteur propre associé à la deuxième valeur propre de X_n évalué, celui-ci est scindé en deux parties comme u_2 .

Nous définissons donc,

$$\tilde{C}_1 = \{s \in \llbracket 1, n \rrbracket \mid (v_2)_s > 0\} \quad \tilde{C}_2 = \{s \in \llbracket 1, n \rrbracket \mid (v_2)_s < 0\}$$

Avec \tilde{C}_1 et \tilde{C}_2 les communautés prédites. Cette méthode de recherche de communautés est appelée méthode par algorithme spectral.

2. X_n étant une matrice stochastique, on parlera de matrice X_n évalué pour évoquer une matrice d'adjacence classique relatif au problème suivant la même forme que X_n

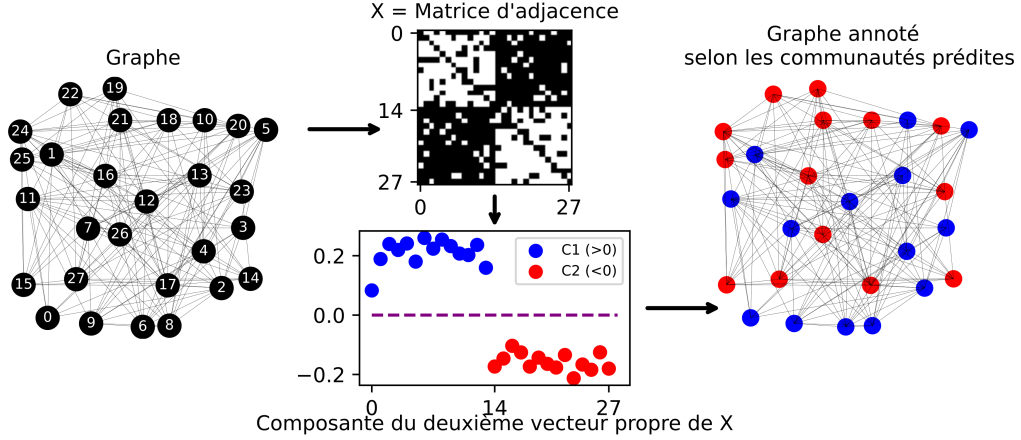


FIGURE 1 – Processus estimant les communautés

Le caractère aléatoire de la matrice X_n ne permet pas de connaître exactement les valeurs de ses valeurs propres et vecteurs propres avant d'effectuer une réalisation.

Dans cet article, nous essayons de prouver l'efficacité de la méthode par algorithme spectral en donnant une majoration efficace malgré le caractère aléatoire du nombre d'erreurs créées par l'algorithme. Nous traiterons aussi les limites de cette méthode dans deux cas critiques ; lorsque q est proche de p et lorsque p et q sont proche de 0

1.3 Résultats non-asymptotique et asymptotique

Pour trouver une majoration fiable du nombre d'erreur, deux méthodes peuvent être utilisées. Une méthode non-asymptotique permet un résultat fiable utilisable peu importe n , utile pour de l'analyse avec un "petit" n . À l'inverse, une méthode asymptotique est plus précise mais ne marche que pour un n très grand, utile pour de grands échantillons.

Notons m le nombre d'erreurs commises par la méthode.

Définition 3 (Erreur). Soit $u_2 \in \mathbb{R}^n$ avec $u_{2,i} = 1$ si $i \in C_1$ et $u_i = -1$ si $i \in C_2$. Soit v_2 un vecteur propre de X_n associé à sa 2ème valeurs propre. On note m le nombre d'erreurs défini par,

$$m = \min_{\eta \in \{-1, 1\}} (\text{Card}\{i \in \llbracket 1, n \rrbracket \text{ tel que } u_{2,i} \eta v_{2,i} < 0\})$$

Résultat 1 (Non-asymptotique). Soit $0 < q < p < 1$, et m le nombre d'erreurs de l'algorithme spectral. En posant $Q = \max_{a \in \{q, p\}} \left\{ \sqrt{\frac{1-2a}{\ln \frac{1-a}{a}}} \right\}$, nous avons pour tout n , avec une probabilité de $1 - 4e^{-3^2} > 99.95\%$,

$$m \leq \frac{46656Q^2}{\min\{q, \frac{p-q}{2}\}^2}$$

Résultat 2 (Asymptotique). Soit $0 < q < p < 1$, et m le nombre d'erreurs de l'algorithme spectral,

$$\mathbb{P}(m \leq \frac{16(q(1-q) + p(1-p))}{\min\{q, \frac{p-q}{2}\}^2}) \xrightarrow{n \rightarrow \infty} 1$$

2 Perturbation

Un Théorème clef dans notre étude des perturbations, et particulièrement central à notre raisonnement, est le Théorème de Davis-Kahan. Tout au long de cette section, nous introduisons les outils nécessaires à sa démonstration. Nous continuons également à nous concentrer sur ce Théorème dans les sections suivantes.

La norme utilisée tout au long du rapport est la norme opérateur. Ce choix fait sens lorsque l'on s'intéresse à la théorie spectrale. En effet, les matrices considérées étant symétriques, leur norme correspond à leur rayon spectral.

2.1 Théorème de Davis-Kahan

Théorème 1 (Davis-Kahan). *Soit $\hat{X} \in \mathbb{M}_n(\mathbb{R})$ symétrique avec $\lambda_1 \geq \dots \geq \lambda_n$ ses valeurs propres et $(u_i)_{1 \leq i \leq n}$ les vecteurs propres unitaires associés. Soit $R \in \mathbb{M}_n(\mathbb{R})$ symétrique et $X := \hat{X} + R$ dont $\{v_i\}_{1 \leq i \leq n}$ sont les vecteurs propres unitaires associés, respectivement, aux valeurs propres $\mu_1 \geq \dots \geq \mu_n$. Posons, θ_i le plus petit angle entre les vecteurs u_i et v_i , alors,*

$$\sin(\theta_i) \leq \frac{2\|R\|}{\min_{j \neq i} |\lambda_i - \lambda_j|}$$

Ce Théorème est particulièrement intéressant puisqu'à partir de quelques considérations trigonométriques, nous pouvons évaluer l'influence du bruit sur une matrice respectant les précédentes conditions, dans notre cas \hat{X} .

Proposition 1. *En considérant les mêmes conditions que précédemment,*

$$\|u_i - \tau v_i\| \leq \frac{2\sqrt{2}\|R\|}{\min_{i \neq j} |\lambda_i - \lambda_j|}, \quad \text{avec } \tau = \begin{cases} +1 & \text{si } \theta_i \in [-\frac{\pi}{2}, \frac{\pi}{2}] \\ -1 & \text{si } \theta_i \in]\frac{\pi}{2}, \frac{3\pi}{2}] \end{cases}$$

2.2 Caractérisation Spectrale

Afin de prouver ce puissant résultat, commençons par caractériser ce qui détermine, intrinsèquement, les matrices carrées : les valeurs propres.

Théorème 2 (Rayleigh-Ritz). *Soit $A \in \mathbb{M}_n(\mathbb{R})$ symétrique, avec $\lambda_1 \geq \dots \geq \lambda_n$ ses valeurs propres. Alors,*

$$\lambda_1 = \max_{x \in \mathbb{R}^n, \|x\|=1} x^T A x$$

$$\lambda_n = \min_{x \in \mathbb{R}^n, \|x\|=1} x^T A x$$

Démonstration. La décomposition spectrale assure l'existence d'une matrice orthogonale Q , composée des vecteurs propres de A , et d'une matrice diagonale D , contenant les valeurs propres correspondantes, tel que, $A = Q^T D Q$

Posons, $y = Qx$, donc

$$x^T A x = y^T D y = \sum_{i=1}^n \lambda_i y_i^2$$

Comme $\lambda_n \leq \lambda_i \leq \lambda_1, \forall i \in \llbracket 1, n \rrbracket$

$$\lambda_n \sum_{i=1}^n y_i^2 \leq \sum_{i=1}^n \lambda_i y_i^2 \leq \lambda_1 \sum_{i=1}^n y_i^2.$$

Puisque Q est orthogonale, $\|y\|^2 = \|Q^T x\|^2 = \|x\|^2$, et donc,

$$\lambda_n \|x\|^2 \leq x^T A x \leq \lambda_1 \|x\|^2$$

C'est-à-dire,

$$\lambda_n \leq x^T A x, \quad \forall x \in \mathbb{R}^n, \|x\| = 1$$

$$\lambda_1 \geq x^T A x, \quad \forall x \in \mathbb{R}^n, \|x\| = 1$$

En particulier, nous avons égalité lorsque $x = v_n$ et $x = v_1$ respectivement. D'où le résultat. \square

Ce théorème ne nous servira pas directement, toutefois il permet d'introduire et d'intuiter un résultat plus fort. Lorsque l'on se place dans une base orthogonale et que l'on applique une matrice symétrique réelle, les valeurs propres décrivent avec quelle intensité elles étirent l'espace. Le théorème suivant permet de se placer dans n'importe quelle sous-espace vectoriel de \mathbb{R}^n pour pouvoir saisir l'effet de ces valeurs propres.

Théorème 3 (Courant-Fischer min-max). *Soit $A \in \mathbb{M}_n(\mathbb{R})$ symétrique, avec $\lambda_1 \leq \dots \leq \lambda_n$ ³ ses valeurs propres et V_i un sous-espace vectoriel de \mathbb{R}^n de dimension i . Alors, $\forall i \in \llbracket 1, n \rrbracket$,*

$$\lambda_i = \min_{V_i \subset \mathbb{R}^n} \max_{\substack{x \in V_i \\ \|x\|=1}} x^T A x$$

Démonstration. Considérons $\{v_j\}_{1 \leq j \leq n}$, l'ensemble des vecteurs propres de A , qui forme une base orthonormée de \mathbb{R}^n . $\forall i \in \llbracket 1, n \rrbracket$, soit $F = \text{Vect}\{v_1, \dots, v_i\} \subset \mathbb{R}^n$, nous avons bien $\dim F = i$.

Pour tout $x \in F$ avec $\|x\| = 1$, il existe $\{x_j\}_{1 \leq j \leq i}$, tel que l'on peut écrire $x = \sum_{j=1}^i x_j v_j$. Donc à l'instar de la précédente preuve,

$$x^T A x = \sum_{j=1}^i \lambda_j x_j^2 \leq \lambda_i \sum_{j=1}^i x_j^2 = \lambda_i$$

Nous avons donc, avec l'égalité atteinte pour $x = v_i$ et $V_i = F$, que,

$$\min_{V_i \subset \mathbb{R}^n} \max_{\substack{x \in V_i \\ \|x\|=1}} x^T A x \leq \lambda_i$$

Réciproquement, soit $F \subset \mathbb{R}^n$ tel que $\dim F = i$ et $G = \text{Vect}\{v_i, \dots, v_n\}$. D'après le Théorème de Grassmann,

$$\dim F \cap G = \dim F + \dim G - \dim F + G \geq i + n - i + 1 - n = 1$$

Cela nous assure l'existence de $x \in F \cap G$ tel que $\|x\| = 1$. Cette fois-ci, nous avons

$$x^T A x = \sum_{j=i}^n \lambda_j x_j^2 \geq \lambda_i \sum_{j=i}^n x_j^2 = \lambda_i$$

Nous avons donc, pour F quelconque, avec égalité si $x = v_i \in F$,

$$\max_{\substack{x \in V_i \\ \|x\|=1}} x^T A x \geq \lambda_i$$

Finalement, par double inégalité,

$$\lambda_i = \min_{V_i \subset \mathbb{R}^n} \max_{\substack{x \in V_i \\ \|x\|=1}} x^T A x$$

□

Enfin, avant de commencer la preuve, nous établissons une inégalité importante dans la théorie spectrale découlant du précédent résultat.

Théorème 4 (Inégalité de Weyl). *Soit $A, B \in \mathbb{M}_n(\mathbb{R})$ symétrique avec $\lambda_1 \geq \dots \geq \lambda_n$ et $\mu_1 \geq \dots \geq \mu_n$ les valeurs propres des matrices A et B , respectivement. Soit η_i la i -ème valeur propre de $A+B$, alors $\forall i \in \llbracket 1, n \rrbracket$,*

$$\mu_n \leq \eta_i - \lambda_i \leq \mu_1$$

3. Par soucis de lisibilité, l'ordre est inversé par rapport aux autres Théorèmes.

Démonstration. Selon le Théorème de Courant-Fischer, nous avons $\forall i \in \llbracket 1, n \rrbracket$,

$$\begin{aligned}
\eta_i &= \min_{V_i \subseteq \mathbb{R}^n} \max_{\substack{x \in V_i \\ \|x\|=1}} x^T (A + B)x \\
&= \min_{V_i \subseteq \mathbb{R}^n} \max_{\substack{x \in V_i \\ \|x\|=1}} (x^T Ax + x^T Bx) \\
&\leq \min_{V_i \subseteq \mathbb{R}^n} \max_{\substack{x \in V_i \\ \|x\|=1}} (x^T Ax) + \mu_1 \quad (\text{D'après le Théorème 2}) \\
&= \lambda_i + \mu_1
\end{aligned}$$

De manière similaire, soit $x_0 = \arg \max_{\substack{x \in V_i \\ \|x\|=1}} x^T Ax$. Alors nous avons,

$$\begin{aligned}
\max_{\substack{x \in V_i \\ \|x\|=1}} (x^T Ax + x^T Bx) &\geq x_0^T Ax_0 + x_0^T Bx_0 \\
&\geq \max_{\substack{x \in V_i \\ \|x\|=1}} x^T Ax + \mu_n.
\end{aligned}$$

En prenant le minimum sur tous les sous-espaces $V_i \subset \mathbb{R}^n$, il s'ensuit que,

$$\eta_i \geq \lambda_i + \mu_n.$$

En combinant les deux résultats, nous concluons l'inégalité de Weyl

$$\begin{aligned}
\lambda_i + \mu_n &\leq \eta_i \leq \lambda_i + \mu_1 \\
\Leftrightarrow \mu_n &\leq \eta_i - \lambda_i \leq \mu_1
\end{aligned}$$

□

2.3 Démonstration du Théorème 1

Démonstration (Davis-Kahan). \hat{X} est symétrique donc $(u_i)_{1 \leq i \leq n}$ forme une base orthonormée. Ecrivons chaque v_i dans celle-ci, c'est-à-dire que $\forall i \in \llbracket 1, n \rrbracket$, $\exists \{c_j\}_{1 \leq j \leq n}$ tel que,

$$v_i = \sum_{j=1}^n c_j u_j.$$

Avec , $\forall j \in \llbracket 1, n \rrbracket$,

$$c_j = \langle u_j, v_i \rangle = \|u_j\| \cdot \|v_i\| \cdot \cos(\angle(v_i, u_j)) = \cos(\theta_i)$$

Examinons comment v_i se comporte en tant qu'approximation d'un vecteur propre de \hat{X} associé à λ_i . Par linéarité,

$$\begin{aligned}
\hat{X} v_i &= \sum_{j=1}^n \lambda_j c_j u_j \\
\hat{X} v_i - \lambda_i v_i &= \sum_{j=1}^n (\lambda_j - \lambda_i) c_j u_j \\
\|\hat{X} v_i - \lambda_i v_i\|^2 &= \left\langle \sum_{j=1}^n (\lambda_j - \lambda_i) c_j u_j, \sum_{k=1}^n (\lambda_k - \lambda_i) c_k u_k \right\rangle \\
&= \sum_{j=1}^n \sum_{k=1}^n (\lambda_j - \lambda_i) (\lambda_k - \lambda_i) c_j c_k \langle u_j, u_k \rangle \\
&= \sum_{j=1}^n (\lambda_j - \lambda_i)^2 c_j^2 \quad (\{u_k\}_{1 \leq k \leq n} \text{ est orthonormée})
\end{aligned}$$

Posons $\delta = \min_{j \neq i} |\lambda_i - \lambda_j|$.

$$\|\hat{X}v_i - \lambda_i v_i\|^2 \geq \delta^2 \sum_{j \neq i} c_j^2 = \delta^2 \sum_{j=1}^n c_j^2 - \delta^2 c_i^2$$

En remarquant que $\sum_{j=1}^n c_j^2 = \langle v_i, v_i \rangle = 1$,

$$\|\hat{X}v_i - \lambda_i v_i\|^2 \geq \delta^2 (1 - c_i^2) = \delta^2 \sin^2(\theta_i)$$

Ainsi,

$$\sin(\theta_i) \leq \frac{\|\hat{X}v_i - \lambda_i v_i\|}{\delta}$$

Enfin,

$$\|\hat{X}v_i - \lambda_i v_i\| = \|(X - R)v_i - \lambda_i I v_i\| = \|(X - I\lambda_i)v_i - Rv_i\| \leq \|(\mu_i - \lambda_i)v_i\| + \|Rv_i\|$$

D'après la Remarque [4] et le Théorème 4, nous avons, $(\mu_i - \lambda_i) \leq \|R\|$, donc,

$$\|\hat{X}v_i - \lambda_i v_i\| \leq \| \|R\| v_i \| + \|Rv_i\| \leq 2\|R\|$$

Finalement, en rassemblant le tout, nous obtenons le résultat attendu.

$$\sin(\theta_i) \leq \frac{2\|R\|}{\delta}$$

□

2.4 Démonstration Proposition 1

Démonstration. Prenons dans un premier temps $\theta_i \in [-\frac{\pi}{2}, \frac{\pi}{2}]$.

$$\langle u_i, v_i \rangle = \frac{-\|u_i - v_i\|^2 + \|u_i\|^2 + \|v_i\|^2}{2} = \frac{-\|u_i - v_i\|^2}{2} + 1$$

Ainsi,

$$-2(\langle u_i, v_i \rangle - 1) = \|u_i - v_i\|^2 \iff \|u_i - v_i\| = \sqrt{2}\sqrt{1 - \langle u_i, v_i \rangle} = \sqrt{2}\sqrt{1 - \cos(\theta_i)}$$

Or $\theta_i \in [-\frac{\pi}{2}, \frac{\pi}{2}]$ donc $\cos(\theta_i) \in [0, 1]$. Ainsi,

$$\|u_i - v_i\| = \sqrt{2}\sqrt{1 - \cos(\theta_i)} \leq \sqrt{2}\sqrt{1 - \cos^2(\theta_i)} \leq \sqrt{2}\sin(\theta_i)$$

Enfin d'après le Théorème 1, nous avons $\|u_i - v_i\| \leq \frac{2\sqrt{2}\|R\|}{\min_{i \neq j} |\lambda_i - \lambda_j|}$.

Prenons maintenant $\theta_i \in]\frac{\pi}{2}, \frac{3\pi}{2}[$

$$\langle u_i, v_i \rangle = \frac{\|u_i + v_i\|^2 - \|u_i\|^2 - \|v_i\|^2}{2} = \frac{\|u_i + v_i\|^2}{2} - 1$$

Ainsi,

$$2(\langle u_i, v_i \rangle + 1) = \|u_i + v_i\|^2 \iff \|u_i + v_i\| = \sqrt{2}\sqrt{1 + \langle u_i, v_i \rangle} = \sqrt{2}\sqrt{1 + \cos(\theta_i)}$$

Or $\theta_i \in]\frac{\pi}{2}, \frac{3\pi}{2}[$ donc $\cos(\theta_i) \in [-1, 0[$. Ainsi,

$$\|u_i + v_i\| = \sqrt{2}\sqrt{1 + \cos(\theta_i)} \leq \sqrt{2}\sqrt{\sin^2(\theta_i) + \cos^2(\theta_i) + \cos(\theta_i)} \leq \sqrt{2}\sin(\theta_i)$$

Enfin, d'après le Théorème 1, nous avons $\|u_i + v_i\| \leq \frac{2\sqrt{2}\|R\|}{\min_{i \neq j} |\lambda_i - \lambda_j|}$.

□

4. mat sym vp abs = norm

3 Etude non-asymptotique

Posons $R_n = X_n - \hat{X}_n$. Le Théorème 1 montre l'importance de diminuer la norme de R_n afin que la borne de la Proposition 1 soit minimisée au mieux. En effet, réduire le bruit permet une meilleure approximation.

Cette section est particulièrement longue car à l'instar de la précédente, elle introduit des outils et résultats, dans le but de fournir la dernière pièce manquante pour répondre à notre problématique. Pour une première lecture, il est préférable de se laisser guider.

3.1 Structures Discrètes : ε -Net et ε -Separated

Souvent, voire systématiquement, nous sommes confrontés à de l'analyse dans des espaces continus. Etudier chaque point de ces espaces peut s'avérer fastidieux, donc un manière de procéder est d'approximer notre espace par un recouvrement d'un nombre fini de boules.

La définition qui suit, établit la définition dans un cadre général. Nous définissons Condition 1 pour alléger la lecture.

Condition 1. Soit K un sous-espace d'un espace métrique muni d'une distance d , $\varepsilon > 0$ et $N \subset K$

Définition 4 (ε -net). On appelle N un ε -net de K , si Condition 1 et tel que

$$\forall x \in K, \exists x_0 \in N, d(x, x_0) \leq \varepsilon.$$

Autrement dit, N est un ε -net si et seulement si K peut être entièrement recouvert par des boules fermées de rayons ε et de centres situés dans N .

Définition 5 (ε -separated). On appelle N un ε -separated de K , si Condition 1 et tel que

$$\forall x_0, y_0 \in N, x_0 \neq y_0, d(x_0, y_0) > \varepsilon.$$

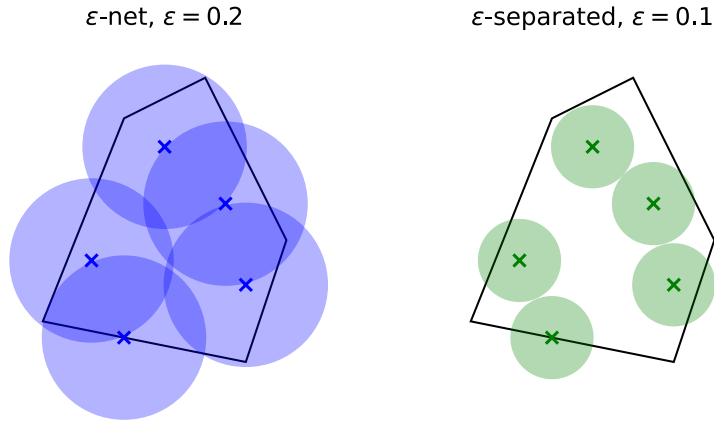


FIGURE 2 – Exemple de ε -net et ε -separated sur un polygone

Cet exemple illustre les précédentes définitions. On observe que pour un même ensemble de points, la valeur de ε peut transformer un ensemble d'un ε -net en un ε -separated. Naturellement, nous définissons la taille minimale et maximale de ces deux objets.

Définition 6 (Recouvrement & Empilement). Si Condition 1,

— Le nombre de recouvrements de K , noté $\mathcal{N}(K, \varepsilon, d)$, est défini par

$$\mathcal{N}(K, \varepsilon, d) = \min\{|N| : N \text{ est un } \varepsilon\text{-net de } K\}.$$

— Le nombre d'empilements de K , noté $\mathcal{P}(K, \varepsilon, d)$, est défini par

$$\mathcal{P}(K, \varepsilon, d) = \max\{|N| : N \text{ est un } \varepsilon\text{-separated de } K\}.$$

Dans la suite du rapport, nous travaillons avec $K \subset \mathbb{R}^n$ donc nous choisissons la distance euclidienne, d . Ainsi, pour simplifier la lecture,

$$\mathcal{N}(K, \varepsilon, d) = \mathcal{N}(K, \varepsilon)$$

$$\mathcal{P}(K, \varepsilon, d) = \mathcal{P}(K, \varepsilon)$$

Remarque 3. Si $K \subset \mathbb{R}^n$ ⁵, $\mathcal{P}(K, \varepsilon)$ est le nombre maximal de boules fermées disjointes de rayon $\frac{\varepsilon}{2}$ et de centres situés dans K .

Théorème 5. Posons $B = B_{\|\cdot\|_2}(0_{\mathbb{R}^n}, 1)$. Nous avons, $\forall \varepsilon > 0$,

$$\mathcal{N}(B, \varepsilon) \leq \left(\frac{2}{\varepsilon} + 1\right)^n$$

Démonstration. Soit $\varepsilon > 0$ et P un ε -separated de B tel que $|P| = \mathcal{P}(B, \varepsilon)$. Son existence est assurée par compacité de B ⁶.

P est également un ε -net de B . En effet, supposons par l'absurde $\exists x \in B$ tel que $d(x, x_0) > \varepsilon, \forall x_0 \in P$ alors $P \cup \{x\}$ est un ε -separated, donc $|P \cup \{x\}| > |P| = \mathcal{P}(B, \varepsilon)$, c'est absurde. On en déduit que $\mathcal{N}(B, \varepsilon) \leq |P|$

Enfin, nous avons que

$$|P| \cdot \text{vol}\left(\frac{\varepsilon}{2}B\right) \leq \text{vol}\left(\left(1 + \frac{\varepsilon}{2}\right)B\right)$$

Nous sommes en n dimensions donc, $\forall r > 0$, $\text{vol}(rB) = r^n \text{vol}(B)$, d'où

$$\mathcal{N}(B, \varepsilon) \leq |P| \leq \frac{\left(1 + \frac{\varepsilon}{2}\right)^n}{\left(\frac{\varepsilon}{2}\right)^n} = \left(\frac{2}{\varepsilon} + 1\right)^n$$

□

Un résultat préliminaire permet d'illustrer l'utilisation de la Section 3.1 afin de majorer la norme d'une matrice. Cependant, il nous est nécessaire d'exprimer la norme opérateur par un produit scalaire.

Propriété 2. Soit $A \in \mathbb{M}_n(\mathbb{R})$ et S^{n-1} désigne la sphère unité dans \mathbb{R}^n . La norme opérateur de A , s'exprime à l'aide du produit scalaire canonique comme,

$$\|A\| = \sup_{x, y \in S^{n-1}} |\langle Ax, y \rangle|$$

Démonstration. Soit $x, y \in S^{n-1}$, d'après Cauchy-Schwarz, $\langle Ax, y \rangle \leq \|Ax\| \cdot \|y\|$. En prenant le sup,

$$\sup_{x, y \in S^{n-1}} |\langle Ax, y \rangle| \leq \sup_{x, y \in S^{n-1}} \|Ax\| \cdot \|y\| = \sup_{x \in S^{n-1}} \|Ax\| = \|A\|$$

Enfin,

$$\sup_{x, y \in S^{n-1}} |\langle Ax, y \rangle| \geq \sup_{x \in S^{n-1}} \left| \left\langle Ax, \frac{Ax}{\|Ax\|} \right\rangle \right| = \sup_{\|x\|=1} \frac{1}{\|Ax\|} \|Ax\|^2 = \sup_{\|x\|=1} \|Ax\| = \|A\|$$

D'où le résultat par double inégalité. □

5. Toute espace normé satisfait cette remarque.

6. Supposons par l'absurde que $|P| = +\infty$, alors il existe une infinité de boules distantes de ε dans B , donc B n'est pas borné. C'est absurde.

Théorème 6. Soit $A \in \mathbb{M}_n(\mathbb{R})$ et $\varepsilon \in]0, \frac{1}{2}[$ tel que N soit un ε -net de S^{n-1} , alors nous avons,

$$\|A\| \leq \frac{1}{1-2\varepsilon} \max_{x,y \in N} \langle Ax, y \rangle$$

Démonstration. Soit $x, y \in S^{n-1}$, ainsi, $\exists x_0, y_0 \in N$ tel que

$$\|x - x_0\|_2 \leq \varepsilon \quad \text{et} \quad \|y - y_0\|_2 \leq \varepsilon$$

$$\begin{aligned} |\langle Ax, y \rangle| &= |\langle A(x - x_0 + x_0), y - y_0 + y_0 \rangle| \\ &= |\langle A(x - x_0), y \rangle + \langle Ax_0, y - y_0 \rangle + \langle Ax_0, y_0 \rangle| \\ &\leq |\langle A(x - x_0), y \rangle| + |\langle Ax_0, y - y_0 \rangle| + |\langle Ax_0, y_0 \rangle| \quad (\text{Inégalité triangulaire}) \\ &\leq \|A(x - x_0)\|_2 \|y\|_2 + \|Ax_0\|_2 \|y - y_0\|_2 + |\langle Ax_0, y_0 \rangle| \quad (\text{Cauchy-Schwarz}) \\ &\leq \|A\| \cdot \|x - x_0\|_2 \cdot 1 + \|A\| \cdot 1 \cdot \|y - y_0\|_2 + |\langle Ax_0, y_0 \rangle| \quad (\text{Inégalité norme opérateur}) \\ &\leq 2\varepsilon \|A\| + |\langle Ax_0, y_0 \rangle|. \end{aligned}$$

En passant au sup des deux cotés, puisque $|N| \leq \infty$, et en appliquant la Propriété 2, nous obtenons,

$$\begin{aligned} \sup_{x,y \in S^{n-1}} |\langle Ax, y \rangle| &\leq 2\varepsilon \|A\| + \max_{x_0, y_0 \in N} |\langle Ax_0, y_0 \rangle| \\ (1-2\varepsilon)\|A\| &\leq \max_{x_0, y_0 \in N} |\langle Ax_0, y_0 \rangle| \\ \|A\| &\leq \frac{1}{1-2\varepsilon} \max_{x_0, y_0 \in N} |\langle Ax_0, y_0 \rangle|. \end{aligned}$$

□

3.2 Variable aléatoire sous-gaussienne

Jusqu'à présent, nous n'avons pas abordé l'aspect aléatoire des graphes. X_n , ne correspond pas exactement à \hat{X}_n car elle est affectée par un certain bruit. Nous exprimons X_n en ajoutant un terme de perturbation : $X_n = \hat{X}_n + R_n$ avec $R_n \in \mathbb{M}_n(\mathbb{R})$. Les variables aléatoires (v.a.) qui composent la matrice de perturbation R_n peuvent être vues comme des transformations d'une variable de Bernoulli. Plus précisément, soit $0 \leq a \leq 1$ et définissons $B_a \sim Be(a) - a$. La matrice R_n est ainsi définie par,

$$R_{ij} = \begin{cases} 0 & \text{si } i = j, \\ B_p & \text{sinon, si } (i, j \leq \frac{n}{2}) \text{ ou } (i, j > \frac{n}{2}), \\ B_q & \text{sinon} \end{cases}$$

Ces variables appartiennent à une catégorie de v.a. nommées *sous-gaussiennes*. Les variables sous-gaussiennes, sont celles dont la fonction génératrice des moments quadratiques est bornée.

Définition 7 (Variable aléatoire sous-gaussienne). Soit une variable aléatoire X telle que $\mathbb{E}(X) = 0$. Elle est dite sous-gaussienne s'il existe $K > 0$ tel que,

$$\mathbb{E}\left(e^{X^2/K^2}\right) < 2$$

Nous pouvons définir une norme pour ces v.a., correspondant au plus petit K satisfaisant la Définition 7. Autrement dit,

Définition 8 (Norme sous-gaussienne). L'application $\|\cdot\|_{\psi_2} : X \mapsto \inf_{t>0} \left\{ \mathbb{E}(e^{X^2/t^2}) \leq 2 \right\}$ est une norme, sur l'espace des v.a. sous-gaussiennes.

Démonstration. La preuve étant lourde et que nous nous servons uniquement de l'homogénéité pour la suite, nous négligeons la démonstration. Une preuve détaillée peut être trouvée dans les notes [4] □

Il existe différentes manières de caractériser ces v.a. sous-gaussiennes. Nous en présentons deux, dont nous nous servons pour la suite.

Propriété 3. Soit X une v.a. telle que $\mathbb{E}(X) = 0$. Alors, les propositions sont équivalentes,

- (i) X est sous-gaussienne.
- (ii) Il existe $K_2 > 0$ tel que $\forall \lambda \in \mathbb{R}, \mathbb{E}(e^{\lambda X}) \leq e^{\lambda^2 K_2^2}$
- (iii) Il existe $K_3 > 0$ tel que $\forall \lambda > 0, \mathbb{P}(|X| \geq \lambda) \leq 2 \exp\left(-\frac{\lambda^2}{K_3^2}\right)$

Démonstration. (i) \implies (ii) Si X est sous-gaussienne, alors en posant $K = \|X\|_{\psi_2}$, ceci vérifie la Définition 7. Utilisons la formule de Taylor avec reste intégral de Laplace, $e^{\lambda X} = 1 + \lambda X + \int_0^1 (1-y) \lambda^2 X^2 e^{y\lambda X} dy$, d'où en passant à l'espérance,

$$\mathbb{E}(e^{\lambda X}) = 1 + \int_0^1 (1-y) \mathbb{E}((\lambda X)^2 e^{y\lambda X}) dy \leq 1 + \left[y - \frac{y^2}{2}\right]_0^1 \mathbb{E}(\lambda^2 X^2 e^{|\lambda X|}) = 1 + \frac{\lambda^2}{2} \mathbb{E}(X^2 e^{|\lambda X|})$$

Puis d'après l'inégalité de Young⁷, nous avons : $|\lambda X| \leq \frac{K^2 \lambda^2}{2} + \frac{X^2}{2K^2}$, donc,

$$\begin{aligned} \mathbb{E}(X^2 e^{|\lambda X|}) &\leq \exp\left(\frac{\lambda^2 K^2}{2}\right) \mathbb{E}\left(X^2 \exp\left(\frac{X^2}{2K^2}\right)\right) \\ &\leq K^2 \exp\left(\frac{\lambda^2 K^2}{2}\right) \mathbb{E} \exp\left(\frac{X^2}{K^2}\right) \quad \left(z \leq \exp \frac{z}{2}, \forall z \in \mathbb{R}\right) \\ &\leq 2K^2 \exp\left(\frac{\lambda^2 K^2}{2}\right) \end{aligned}$$

En appliquant pour la dernière étape $1 + z \leq e^z, \forall z \in \mathbb{R}$, nous majorons le tout,

$$\mathbb{E} e^{\lambda X} \leq 1 + \lambda^2 K^2 \exp\left(\frac{\lambda^2 K^2}{2}\right) \leq (1 + \lambda^2 K^2) \exp\left(\frac{\lambda^2 K^2}{2}\right) \leq \exp\left(\frac{3\lambda^2 K^2}{2}\right)$$

D'où le résultat en considérant $K_2 > 0$ tel que $\sqrt{\frac{3}{2}}K \leq K_2$.

A l'instar de la Définition 8, on note

$$\|X\|_\eta := \inf_{t>0} \left\{ \forall \lambda \in \mathbb{R}, \mathbb{E}(e^{\lambda X}) \leq e^{\lambda^2 t^2} \right\}$$

Donc d'après la précédente inégalité,

$$\|X\|_\eta \leq \sqrt{\frac{3}{2}} \|X\|_{\psi_2} \tag{1}$$

(ii) \implies (iii) Soit $K_2 = \|X\|_\eta$, nous avons bien pour tout $\lambda \in \mathbb{R}, \mathbb{E}(e^{\lambda X}) \leq \exp(\lambda^2 K_2^2)$. Soit $t > 0$ que l'on définira,

$$\mathbb{P}(X \geq \lambda) = \mathbb{P}(e^{tX} \geq e^{t\lambda}) \leq \mathbb{E}(e^{tX}) e^{-\lambda t} \leq e^{t^2 K_2^2 - \lambda t}$$

Minimisons la borne selon t , c'est-à-dire posons $t = \frac{\lambda}{2K_2^2}$, donc,

$$\mathbb{P}(X \geq \lambda) \leq \exp\left(-\frac{\lambda^2}{4K_2^2}\right)$$

Puis, $\mathbb{P}(X \leq -\lambda)$ est un cas analogue et via l'inégalité de Boole, nous obtenons le résultat souhaité avec $K_3 > 0$ tel que $2K_2 \leq K_3$

7. C'est une application de l'inégalité arithmético-géométrique, où $\forall a, b \in \mathbb{R}$ et $K > 0$, en définissant $a' = a/K$ et $b' = Kb$, nous avons $2ab = 2a'b' \leq a'^2 + b'^2 = K^{-1}a + Kb$

Cette fois-ci, définissons

$$\|X\|_\mu := \inf_{t>0} \left\{ \forall \lambda > 0, \mathbb{P}(|X| \geq \lambda) \leq 2 \exp\left(-\frac{\lambda^2}{t^2}\right) \right\}$$

Donc d'après la précédente inégalité,

$$\|X\|_\mu \leq 2\|X\|_\eta \quad (2)$$

(iii) \implies (i) Soit $K_3 = \|X\|_\mu$, posons $c = \frac{1}{K_1^2}$. Pour tout $a > 0$ tel que $0 < a < c$, nous avons,

$$\mathbb{E}\left(e^{aX^2}\right) \leq 1 + \int_0^\infty 2ate^{at^2} \cdot \mathbb{P}(|X| > t) dt$$

D'après (iii), $\mathbb{P}(|X| \geq t) \leq 2e^{-ct^2}$, donc l'intégrale est majorée par,

$$\int_0^\infty 2ate^{at^2} \cdot 2e^{-ct^2} dt = 4a \int_0^\infty te^{-(c-a)t^2} dt = 2a \left[-\frac{e^{-(c-a)t^2}}{c-a} \right]_0^\infty = \frac{2a}{c-a}$$

Soit $K = \sqrt{\frac{1}{a}} > 0$, c'est-à-dire $a = \frac{1}{K^2}$, alors,

$$\frac{2a}{c-a} = \frac{2 \cdot \frac{1}{K^2}}{\frac{1}{K_3^2} - \frac{1}{K^2}} = \frac{2}{K^2 \left(\frac{K^2 - K_3^2}{K_3^2 K^2} \right)} = \frac{2K_3^2}{K^2 - K_3^2}$$

D'où,

$$\mathbb{E}(e^{X^2/K^2}) = \mathbb{E}(e^{aX^2}) \leq 1 + \frac{2K_3^2}{K^2 - K_3^2} \leq 2 \implies 2K_3^2 \leq K^2 - K_3^2 \iff \sqrt{3}K_3 \leq K$$

A la fin, nous avons bien $K_3 \leq K$, c'est-à-dire $a < c$, ce qui conclut la preuve.

Nous obtenons une dernière relation de cette implication, $\|X\|_{\psi_2} \leq \sqrt{3}\|X\|_\mu$, qui lorsque l'on combine avec (1) et (2), nous obtenons,

$$\|X\|_{\psi_2} \leq \sqrt{3}\|X\|_\mu \leq 2\sqrt{3}\|X\|_\eta \leq 3\sqrt{2}\|X\|_{\psi_2} \quad (3)$$

□

Nous cernons mieux le comportement des v.a. sous-gaussiennes. Il en est de même lorsque nous les additionnons. En effet, le résultat est également une v.a. sous-gaussienne, dont le comportement global est contrôlé par le comportement individuel de chacun de ses termes.

Propriété 4. Soit $\{X_i\}_{1 \leq i \leq n}$ des v.a. sous-gaussiennes. Alors, $\sum_{i=1}^n X_i$ est une v.a. sous-gaussienne et

$$\left\| \sum_{i=1}^n X_i \right\|_{\psi_2}^2 \leq 18 \sum_{i=1}^n \|X_i\|_{\psi_2}^2,$$

Démonstration. Pour commencer, montrons que $\sum_{i=1}^n X_i$ est sous-gaussienne. Pour tout $\lambda \in \mathbb{R}$, nous avons, par indépendance des X_i ,

$$\mathbb{E} \exp \left(\lambda \sum_{i=1}^n X_i \right) = \prod_{i=1}^n \mathbb{E} (e^{\lambda X_i}) \leq \prod_{i=1}^n e^{\lambda^2 \|X_i\|_\eta^2}$$

D'après (3), $\|X\|_\eta \leq \frac{3\sqrt{2}}{2\sqrt{3}}\|X\|_{\psi_2} = \sqrt{\frac{3}{2}}\|X\|_{\psi_2}$. Donc,

$$\begin{aligned}
\mathbb{E} \exp \left(\lambda \sum_{i=1}^n X_i \right) &\leq \prod_{i=1}^n \exp \left(\lambda^2 \frac{3}{2} \|X_i\|_{\psi_2}^2 \right) \\
&= \exp \left(\lambda^2 \frac{3}{2} \sum_{i=1}^n \|X_i\|_{\psi_2}^2 \right) \\
&= e^{\lambda^2 K_2^2} \quad \text{avec } K_2^2 = \frac{3}{2} \sum_{i=1}^n \|X_i\|_{\psi_2}^2
\end{aligned}$$

Ainsi, d'après la Propriété 3, $\sum_{i=1}^n X_i$ est sous-gaussienne. De plus, d'après (3), mise au carré, nous obtenons le résultat souhaité,

$$\left\| \sum_{i=1}^N X_i \right\|_{\psi_2}^2 \leq (2\sqrt{3})^2 \left\| \sum_{i=1}^N X_i \right\|_{\eta}^2 \leq 12K_2^2 = 18 \sum_{i=1}^n \|X_i\|_{\psi_2}^2$$

□

Lorsque l'on s'intéresse aux v.a. de la forme B_a , c'est-à-dire les coefficients de R_n , il est possible de les majorer efficacement malgré leur nature aléatoire.

Proposition 2. *Pour $0 < p < 1$,*

$$\|B_p\|_{\psi_2} \leq \sqrt{\frac{1-2p}{\ln \frac{1-p}{p}}}$$

Démonstration. Considérons la fonction $f_B : p \in]0, 1[\mapsto \sqrt{\frac{1-2p}{\ln \frac{1-p}{p}}}$, que nous utiliserons aussi dans la suite.

Définissons également la fonction g définie par,

$$g(p) = \mathbb{E} \exp \left(\frac{B_p^2}{f_B(p)^2} \right) = p \exp \left(\frac{(1-p)^2 \ln \frac{1-p}{p}}{1-2p} \right) + (1-p) \exp \left(\frac{p^2 \ln \frac{1-p}{p}}{1-2p} \right).$$

Par une analyse de fonction, nous montrons que $g'(p) < 0$ pour $p \in]0, \frac{1}{2}[$ et $g'(p) > 0$ pour $p \in]\frac{1}{2}, 1[$. En comparant la croissance d'un polynôme à celle de la fonction exponentielle, nous obtenons les limites suivantes,

$$\lim_{p \rightarrow 0^+} g(p) = 2 \quad \text{et} \quad \lim_{p \rightarrow 1^-} g(p) = 2.$$

Par conséquent, pour tout $p \in]0, 1[$, nous avons $g(p) < 2$. Ainsi, il s'ensuit que,

$$\|B_p\|_{\psi_2} \leq f_B(p).$$

□

3.3 Majoration de $\|R_n\|$

Le théorème suivant est inspiré du Théorème 4.4.5 de [5]. Toutefois, bien que plus spécifique, nous explicitons une constante, facilitant ainsi son application directe à notre étude.

Théorème 7. *Soit $A \in \mathbb{M}_n(\mathbb{R})$ et $\{a_{ij}\}_{1 \leq i, j \leq n}$ où $0 \leq a_{ij} \leq 1$ tel que $A_{ij} \sim B_{a_{ij}}$. Alors, pour tout $t > 0$, avec une probabilité d'au moins $1 - 2e^{-t^2}$,*

$$\|A\| \leq 18\sqrt{2} \left(\max_{1 \leq i, j \leq n} \|B_{a_{ij}}\|_{\psi_2} \right) (2\sqrt{n} + t),$$

Démonstration. Soit $\varepsilon = \frac{1}{4}$ et $\mathcal{N}_1, \mathcal{N}_2$ deux ε -net de $S^{n-1} \subset B^n$. D'après le Théorème 5,

$$|\mathcal{N}_1| \leq 9^n \quad \text{et} \quad |\mathcal{N}_2| \leq 9^n \quad (4)$$

Le choix de ε est un compromis entre la densité du ε -net et de son cardinal.

D'après le Théorème 6, nous avons le coeur du résultat,

$$\|A\| \leq 2 \max_{x, y \in \mathcal{N}} \langle Ax, y \rangle$$

Donc, fixons $(x, y) \in \mathcal{N}_1 \times \mathcal{N}_2$. Nous avons d'après la proposition 4,

$$\|\langle Ax, y \rangle\|_{\psi_2}^2 = \left\| \sum_{i=1}^n \sum_{j=1}^n A_{ij} x_i y_j \right\|_{\psi_2}^2 \leq 18 \sum_{i=1}^n \sum_{j=1}^n \|A_{ij} x_i y_j\|_{\psi_2}^2$$

Posons $K = \max_{1 \leq i, j \leq n} \|B_{a_{ij}}\|_{\psi_2}$, donc,

$$\|\langle Ax, y \rangle\|_{\psi_2}^2 \leq 18K^2 \sum_{i=1}^n x_i^2 \sum_{j=1}^n y_j^2 = 18K^2$$

Sachant que $\langle Ax, y \rangle$ est sous-gaussienne, d'après la Propriété 3 et la relation (3), nous obtenons une inégalité de concentration que nous majorons selon la précédente inégalité. Soit $u \geq 0$,

$$\mathbb{P}(\langle Ax, y \rangle \geq u) \leq 2 \exp\left(-\frac{u^2}{6\|\langle Ax, y \rangle\|_{\psi_2}^2}\right) \leq 2 \exp\left(-\frac{u^2}{108K^2}\right) \quad (5)$$

Donc, en parcourant l'ensemble des points de \mathcal{N}_1 et \mathcal{N}_2 , nous obtenons,

$$\begin{aligned} \mathbb{P}\left(\max_{\substack{x \in \mathcal{N}_1 \\ y \in \mathcal{N}_2}} \langle Ax, y \rangle \geq u\right) &\leq \sum_{x \in \mathcal{N}_1} \sum_{y \in \mathcal{N}_2} \mathbb{P}(\langle Ax, y \rangle \geq u) \quad (\text{inégalité de Boole}) \\ &\leq 9^{n+n} \cdot 2 \exp\left(-\frac{u^2}{108K^2}\right) \quad (\text{D'après, (4) et (5)}) \end{aligned}$$

Choisissons $u = 9\sqrt{2}K(2\sqrt{n} + t) \geq 0$, donc,

$$\frac{u^2}{108K^2} \geq \frac{3}{2}(4n + t^2) \geq 3(n + n) + t^2$$

Nous pouvons combiner le tout et obtenir la majoration souhaitée,

$$\begin{aligned} \mathbb{P}(\|A\| \geq 2u) &\leq \mathbb{P}\left(\max_{\substack{x \in \mathcal{N}_1 \\ y \in \mathcal{N}_2}} \langle Ax, y \rangle \geq u\right) \\ &\leq 9^{n+n} \cdot 2e^{-3(n+n)-t^2} \\ &\leq 2e^{-t^2} \end{aligned}$$

□

Cette estimation nécessite l'indépendance des coefficients dont R_n ne dispose pas. En effet, R_n est symétrique. Néanmoins, cela ne nous dérange pas tant que ça.

Lemme 1. Soit $A \in \mathbb{M}_n(\mathbb{R})$ symétrique et $\{a_{ij}\}_{1 \leq i \leq j \leq n}$ où $0 \leq a_{ij} \leq 1$ tel que $A_{ij} \sim B_{a_{ij}}$. Alors, pour tout $t > 0$, avec une probabilité d'au moins $1 - 4e^{-t^2}$,

$$\|A\| \leq 6 \left(\max_{1 \leq i, j \leq n} \|B_{a_{ij}}\|_{\psi_2} \right) (2\sqrt{n} + t),$$

Démonstration. Décomposons A en 2 matrices triangulaires, A_I et A_S , inférieure et supérieure. Le choix de la diagonale importe peu, ses coefficients ne sont pas liés avec d'autres A_{ij} . Ainsi, $A = A_I + A_S$ donc

$$\|A\| \leq \|A_I\| + \|A_S\|$$

En appliquant le Théorème 7 sur chaque terme de la majoration, et d'après l'inégalité de Boole,

$$\mathbb{P} \left(\|A\| \geq 18\sqrt{2} \left(\max_{1 \leq i, j \leq n} \|B_{a_{ij}}\|_{\psi_2} \right) (2\sqrt{n} + t) \right) \leq 2 \cdot 2e^{-t^2}$$

D'où le résultat, en prenant l'évènement contraire. \square

3.4 Démonstration du Résultat 1

Démonstration (Résultat 1). Nous pouvons maintenant appliquer le Théorème 1 aux matrices \hat{X}_n et X_n .

Soit $\lambda_1 \geq \dots \geq \lambda_n$ les valeurs propres de \hat{X}_n . Considérons u_2 , un vecteur propre de norme \sqrt{n} correspondant à la deuxième valeur propre de \hat{X}_n , ainsi que v_2 , un vecteur propre de norme \sqrt{n} correspondant à la deuxième valeur propre de X_n . Définissons alors leur vecteur unitaire, $\bar{u}_2 = \frac{1}{\sqrt{n}}u_2$ et $\bar{v}_2 = \frac{1}{\sqrt{n}}v_2$. Posons $\delta = \min_{j \neq 2} |\lambda_2 - \lambda_j| = \min\{q, \frac{p-q}{2}\}$. D'après la Propriété 1, il existe un $\tau \in \{-1, 1\}$ tel que,

$$\|\bar{u}_2 - \tau \bar{v}_2\| \leq \frac{2\sqrt{2}\|R_n\|}{\delta n}$$

D'après la Proposition 2 et le Lemme 1, avec une probabilité d'au moins $1 - 4e^{-3^2} > 99.95\%$ ⁸,

$$\|R\| \leq 18\sqrt{2} \left(\max_{1 \leq i, j \leq n} \|B_{a_{ij}}\|_{\psi_2} \right) (2\sqrt{n} + 3) \leq 54\sqrt{2n} \max_{a \in \{q, p\}} \{f_B(a)\}$$

Donc, en posant $Q = \max_{a \in \{q, p\}} \{f_B(a)\}$

$$\begin{aligned} \|\bar{u}_2 - \tau \bar{v}_2\| &\leq \frac{2\sqrt{2} \cdot 54\sqrt{2n}Q}{\delta\sqrt{n}} \\ \Rightarrow \|u_2 - \tau v_2\| &\leq \frac{216Q}{\delta} \\ \Rightarrow \|u_2 - \tau v_2\|^2 &\leq \frac{46656Q^2}{\delta^2} \\ \Rightarrow \sum_{i=1}^n (u_{2,i} - \tau v_{2,i})^2 &\leq \frac{46656Q^2}{\delta^2} \end{aligned}$$

Par définition de m , le nombre d'erreur, nous savons que,

$$m \leq \text{Card} \{i \in \llbracket 1, n \rrbracket \text{ tel que } u_{2,i}\tau v_{2,i} < 0\}$$

En se rappelant de la Propriété 1, $u_{2,i} = \pm 1$, donc,

$$m \leq \text{Card} \{i \in \llbracket 1, n \rrbracket \text{ tel que } (u_{2,i} - \tau v_{2,i})^2 > 1\}$$

Nous obtenons donc le résultat souhaité,

$$m \leq \sum_{i=1}^n (u_{2,i} - \tau v_{2,i})^2 \leq \frac{46656Q^2}{\delta^2}$$

\square

8. Et $n \geq 9$ puisque par soucis de lisibilité il est préférable de majorer 3 par \sqrt{n}

4 Étude asymptotique

La recherche de la valeur exacte de la norme de R_n est rendu impossible de par son caractère aléatoire. Néanmoins, une convergence de la norme lorsque n tend vers l'infini semble possible, il est donc pertinent d'étudier son comportement asymptotique.

L'étude des propriétés spectrales asymptotiques des matrices aléatoires est un domaine de recherche très complet dont un des théorèmes clefs est le Théorème de Wigner.

4.1 Théorème de Wigner et loi du demi cercle

Définition 9. Soit W_n une matrice aléatoire de taille n et $(\lambda_i)_{i \in \llbracket 1, n \rrbracket}$ ses valeurs propres classées dans l'ordre croissant, on définit sa mesure spectrale, ou distribution de ses valeurs propres μ_{W_n} tel que,

$$\mu_{W_n} = \frac{1}{n} \sum_{i=1}^n \delta_{\lambda_i}$$

Théorème 8 (Théorème de Wigner (Admis)). Soit $W_n = (w_{i,j})_{1 \leq i,j \leq n}$ une matrice aléatoire, W_n est dite matrice de Wigner si elle est symétrique et que ses coefficients sont i.i.d. avec $\mathbb{E}(w_{i,j}) = 0$ et $\mathbb{V}(w_{i,j}) = \sigma^2$. Dans ce cas là, la mesure spectrale $\mu_{\frac{1}{\sqrt{n}}W_n}$ converge presque sûrement vers la loi du demi cercle

$$\mu_{sc}^{\sigma^2}(x) = \frac{1}{2\pi\sigma^2} \sqrt{4\sigma^2 - x^2}$$

Nous avons ainsi une convergence en probabilité de $\frac{1}{\sqrt{n}}\|W_n\|$ vers $\max(|x| \text{ tel que } \mu_{sc}^{\sigma^2}(x) > 0) = 2\sigma$ qu'on notera à partir de maintenant $\frac{1}{\sqrt{n}}\|W_n\| \xrightarrow[n \rightarrow \infty]{\mathbb{P}} 2\sigma$

Il est clair que R_n n'est pas une matrice de Wigner. Cependant lorsqu'on regarde sa mesure spectrale avec n suffisamment grand, une distribution en forme de demi-cercle apparaît. Un lien entre R_n et le théorème de Wigner peut donc être conjecturé. Une piste de recherche que nous avons explorée consiste à approximer R_n par une matrice aléatoire i.i.d. dont l'espérance et la variance de cette matrice sont définies comme la moyenne des espérances et des variances, respectivement, des coefficients de R_n . Nous obtenons ainsi une matrice de Wigner avec paramètre $\sigma^2 = \frac{p(1-p)+q(1-q)}{2}$.

La méthode est prometteuse numériquement. Nous essayons donc de prouver son efficacité.

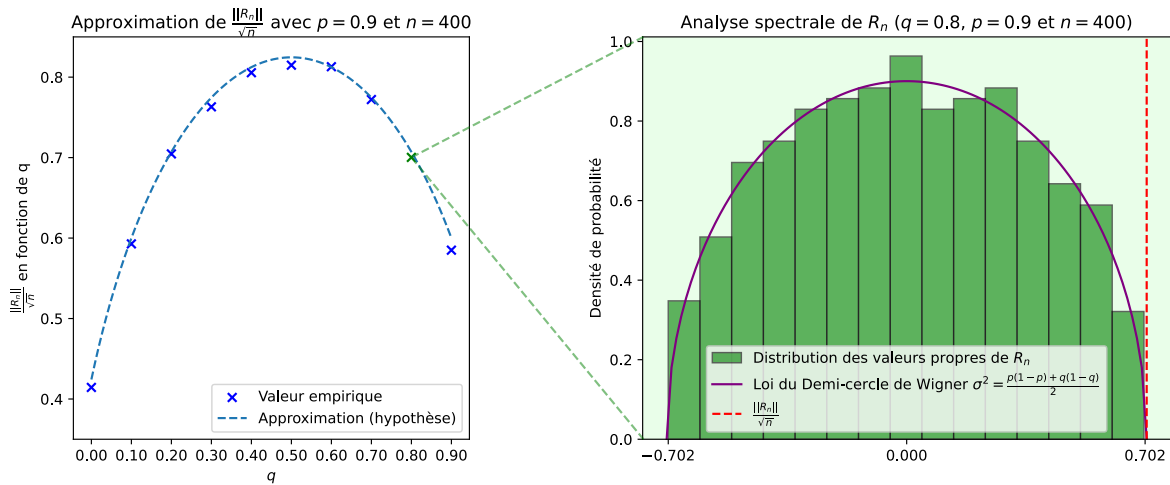


FIGURE 3 – Approximation de $\|R_n\|$ par une matrice de Wigner

Propriété 5. Soit R_n une matrice aléatoire de bruit de taille n , alors on a

$$\|R_n\| \xrightarrow[n \rightarrow \infty]{\mathbb{P}} \sqrt{2n(q(1-q) + p(1-p))}$$

Généralement, la méthode des moments est employée pour la preuve, inspirée de la démonstration du Théorème de Wigner. Nous allons ici utiliser une approche différente passant par la transformée de Stieltjes permettant de nombreuses généralisations⁹.

Définition 10 (Transformée de Stieltjes). Soit W_n une matrice aléatoire et μ_n la distribution de ses valeurs propres. On définit la transformée de Stieltjes de μ_n comme :

$$s_{\mu_n}(z) = \int \frac{d\mu_n(x)}{x - z} = \frac{1}{n} \sum_{i=1}^n \frac{1}{\lambda_i(W_n) - z}$$

Remarque 4. La transformée de la loi du demi-cercle est $s_{\mu_{sc}}(z) = \frac{-2 + \sqrt{z^2 - 4}}{2} = \frac{-1}{z + \sigma^2 s_{\mu_{sc}}(z)}$ ¹⁰

Théorème 9 (Continuité de Stieltjes). μ_n converge presque sûrement vers la loi de distribution μ si et seulement si $s_{\mu_n}(z)$ converge presque sûrement vers $s_{\mu}(z)$ pour tout z dans $\mathbb{C}^+ = \{z \in \mathbb{C} | \text{Im}(z) > 0\}$.

Nous pouvons donc utiliser la continuité de Stieltjes pour montrer que la mesure spectrale de R_n tend vers la loi demi-cercle. Pour cela il suffit de montrer que la transformée de la mesure spectrale de R_n tend vers $S_{\mu_{sc}}$.

4.2 Equation canonique K1

Pour faciliter le calcul de la transformée de la mesure spectrale μ_{R_n} , nous allons utiliser les travaux de L.V Girko [2] sur la classification des transformées de Stieltjes appliquée aux matrices aléatoires, plus précisément sur la classe K1.

Propriété 6 (Équation canonique K1). Soit $W_n = (w_{i,j})_{i,j \in \llbracket 1, n \rrbracket}$ une matrice aléatoire symétrique où chaque coefficient suit une loi discrète bornée et est indépendant des autres¹¹ telle que pour tout $i, j \in \llbracket 1, n \rrbracket$,

$$\mathbb{E}(w_{i,j}) = a_{i,j} \quad \text{et} \quad \mathbb{V}(w_{i,j}) = \sigma_{i,j}^2$$

En notant $A_n = (a_{i,j})_{i,j \in \llbracket 1, n \rrbracket}$, δ_{ij} le symbole de Kronecker, et $L = \{z \in \mathbb{C} \mid \text{Im}(z)\text{Im}(c_k) > 0, \forall k \in \llbracket 1, n \rrbracket\}$, nous avons pour tout k , $c_k(z)$ est l'unique solution dans L du système d'équations,

$$c_k(z) = ([A_n - zI_n - \Omega]^{-1})_{kk}$$

Avec $\Omega = (\Omega_{i,j})_{i,j \in \llbracket 1, n \rrbracket}$ définie par,

$$\Omega_{i,j} = \delta_{ij} \sum_{k=1}^n c_k(z) \sigma_{k,j}^2$$

Alors la mesure spectrale μ_{W_n} converge presque sûrement vers la mesure spectrale μ de transformée de $s_{\mu}(z)$ avec,

$$s_{\mu}(z) = \int \frac{d\mu(x)}{x - z} = \frac{1}{n} \sum_{i=1}^n c_i(z)$$

9. La méthode des moments donne cependant plus d'information sur la vitesse de convergence (voir Theoreme 3 [3])

10. On exprime souvent la transformée de Stieltjes sous sa forme auto-cohérente. Nous n'allons pas démontrer ici l'équivalence de ces deux notations

11. Le théorème est plus général dans l'ouvrage de Girko, mais nous nous limitons à un cas simple.

4.3 Démonstration Propriété 5

Démonstration (Propriété 5). Pour faciliter les calculs, nous allons normaliser la matrice R_n . On définit alors $\bar{R}_n := \frac{1}{\sqrt{n}} R_n$, satisfaisant toutes les conditions de la Propriété 6 avec dans notre cas A une matrice nulle. La transformé de Stieljes de la mesure spectrale de \bar{R}_n converge donc presque sûrement vers $S_{\mu_{\bar{R}_n}}(z)$ tel que,

$$S_{\mu_{\bar{R}_n}}(z) = \frac{1}{n} \sum_{i=1}^n c_i(z)$$

et pour tout $i \in \llbracket 1, n \rrbracket$,

$$c_i(z) = ([-zI_n - \Omega]^{-1})_{ii} = \frac{-1}{z + \Omega_{i,i}}$$

Avec,

$$\Omega_{i,j} = \delta_{ij} \sum_{k=1}^n c_k(z) \sigma_{k,j}^2 \quad \text{Et} \quad \sigma_{k,i}^2 = \begin{cases} \frac{p(1-p)}{n} & \text{si } (k, i) \in \llbracket 1, \frac{n}{2} \rrbracket^2 \cup \llbracket \frac{n}{2} + 1, n \rrbracket^2 \\ \frac{q(1-q)}{n} & \text{sinon} \end{cases}$$

Nous avons donc,

$$c_i(z) = \frac{-1}{z + \sum_{k=1}^n c_k(z) \sigma_{k,i}^2}$$

Nous nous rendons compte que pour tout $i, j \in \llbracket 1, \frac{n}{2} \rrbracket$, $c_i(z) = c_j(z)$, et de même, si $i, j \in \llbracket \frac{n}{2} + 1, n \rrbracket$. Donc simplifions,

$$S_{\mu_{\bar{R}_n}}(z) = \frac{1}{2} (c_1(z) + c_n(z))$$

Avec,

$$\begin{cases} c_1(z) &= \frac{-1}{z + \frac{q(1-q)}{2} c_n(z) + \frac{p(1-p)}{2} c_1(z)} \\ c_n(z) &= \frac{-1}{z + \frac{q(1-q)}{2} c_1(z) + \frac{p(1-p)}{2} c_n(z)} \end{cases}$$

Ce système d'équations est appelé équation quadratique vectorielle (QVE)[1]. Dans ce cas $c_1(z) = c_n(z)$ est l'unique solution du système dans L. Nous avons donc,

$$S_{\mu_{\bar{R}_n}}(z) = c_1(z) = \frac{-1}{z + \frac{q(1-q)+p(1-p)}{2} c_1(z)} = \frac{-1}{z + \frac{q(1-q)+p(1-p)}{2} S_{\mu_{\bar{R}_n}}(z)}$$

Nous reconnaissons ici la transformée de Stieljes de la loi du demi-cercle. D'après la continuité de la transformée, nous savons donc que $\mu_{\bar{R}_n}$ converge presque sûrement vers la distribution du demi cercle de paramètre $\sigma^2 = \frac{p(1-p)+q(1-q)}{2}$. Nous avons donc, comme supposé,

$$\begin{aligned} \|\bar{R}_n\| &\xrightarrow[n \rightarrow \infty]{\mathbb{P}} 2\sqrt{\frac{q(1-q) + p(1-p)}{2}} \\ \Leftrightarrow \|R_n\| &\xrightarrow[n \rightarrow \infty]{\mathbb{P}} \sqrt{2n(q(1-q) + p(1-p))} \end{aligned}$$

□

4.4 Démonstration du Résultat 2

Démonstration (Résultat 2). De la même manière que dans la section 3, nous pouvons maintenant appliquer le Théorème 1 aux matrices \hat{X}_n et X_n .

Soit $\lambda_1 \geq \dots \geq \lambda_n$ les valeurs propres de \hat{X}_n . Considérons u_2 , un vecteur propre de norme \sqrt{n} correspondant à la deuxième valeur propre de \hat{X}_n , ainsi que v_2 , un vecteur propre de norme \sqrt{n} correspondant à

la deuxième valeur propre de X_n . Définissons alors leur vecteur unitaire, $\bar{u}_2 = \frac{1}{\sqrt{n}}u_2$ et $\bar{v}_2 = \frac{1}{\sqrt{n}}v_2$. Posons $\delta = \min_{j \neq 2} |\lambda_2 - \lambda_j| = \min\{q, \frac{p-q}{2}\}$. D'après la Proposition 5,

$$\|R_n\| \xrightarrow[n \rightarrow \infty]{\mathbb{P}} \sqrt{2n(q(1-q) + p(1-p))}$$

En se rappelant que le nombre d'erreurs m est majoré par $m \leq \sum_{i=1}^n (u_{2,i} - \tau v_{2,i})^2$, nous développons de la même manière que dans la section 3.4 (Démonstration du Résultat 1),

$$\begin{aligned} \|\bar{u}_2 - \tau \bar{v}_2\| &\leq \frac{2\sqrt{2}\sqrt{2n(q(1-q) + p(1-p))}}{\delta n} \\ \Rightarrow \|u_2 - \tau v_2\| &\leq \frac{4\sqrt{(q(1-q) + p(1-p))}}{\delta} \\ \Rightarrow \|u_2 - \tau v_2\|^2 &\leq \frac{16(q(1-q) + p(1-p))}{\delta^2} \\ \Rightarrow \sum_{i=1}^n (u_{2,i} - \tau v_{2,i})^2 &\leq \frac{16(q(1-q) + p(1-p))}{\delta^2} \end{aligned}$$

Nous obtenons donc le résultat souhaité,

$$\mathbb{P}\left(m \leq \frac{16(q(1-q) + p(1-p))}{\delta^2}\right) \xrightarrow[n \rightarrow \infty]{} 1$$

□

4.5 Proportion d'erreur maximale

Grâce à l'étude asymptotique du nombre d'erreurs, nous étudions l'efficacité de la méthode par l'algorithme spectrale avec n grand. Pour cela nous définissons la proportion d'erreur maximale.

Définition 11 (Proportion d'erreur maximale). *Soit m le nombre d'erreurs, nous avons ainsi $\frac{m}{n}$ la proportion d'erreurs. Notons $Pr(n, p, q)$ la proportion d'erreurs maximale définie comme,*

$$Pr(n, p, q) = \frac{16(q(1-q) + p(1-p))}{n \min\{q, \frac{p-q}{2}\}^2}$$

Nous avons donc,

$$\mathbb{P}\left(\frac{m}{n} \leq Pr(n, p, q)\right) \xrightarrow[n \rightarrow \infty]{} 1$$

Nous observons que pour p et q fixés, la proportion d'erreur tend vers 0 lorsque n tend vers l'infini. La méthode par l'algorithme spectral gagne donc en efficacité lorsque n tend vers l'infini. Cependant nous remarquons une perte d'efficacité de l'algorithme spectral lorsque p et q sont proches de 0, ou lorsque p est proche de q .

Pour étudier ces cas critiques, nous étudions la comportement asymptotique de deux fonctions : $Pr\left(n, \frac{p}{\varphi(n)}, \frac{q}{\varphi(n)}\right)$ et $Pr\left(n, p, p - \frac{1}{\varphi(n)}\right)$ avec $\varphi(n)$ une fonction croissante et p, q fixés.

Propriété 7. *Soit $\varphi(n) = n^\alpha$ avec $\alpha \in \mathbb{R}^+$. Nous considérons 3 comportements asymptotiques de $Pr\left(n, p, p - \frac{1}{\varphi(n)}\right)$ en fonction de α .*

1. Si $\alpha < 0.5$: La proportion d'erreur maximale tend vers 0.
2. Si $\alpha = 0.5$: La proportion d'erreur maximale est constante pour tout n .
3. Si $\alpha > 0.5$: La proportion d'erreur maximale diverge vers l'infini.

Démonstration. Pour prouver cette proposition, nous calculons simplement la limite de $Pr\left(n, p, 1 - \frac{1}{n^\alpha}\right)$ nous considérons ici que $\min\left\{p - \frac{1}{n^\alpha}, \frac{1}{2n^\alpha}\right\} = \frac{1}{2n^\alpha}$ pour n grand.

$$\begin{aligned}\lim_{n \rightarrow \infty} Pr\left(n, \frac{p}{n^\alpha}, \frac{q}{n^\alpha}\right) &= \lim_{n \rightarrow \infty} \frac{16(p(1-p) + (p - \frac{1}{n^\alpha})(1 - p + \frac{1}{n^\alpha}))}{n^{\frac{1}{4n^{2\alpha}}}} \\ &= \lim_{n \rightarrow \infty} \frac{64(p(1-p) + p(1-p))}{n^{1-2\alpha}} \\ &= \lim_{n \rightarrow \infty} \frac{124p(1-p)}{n^{1-2\alpha}}\end{aligned}$$

La proportion d'erreur max tend donc vers 0 si $\alpha < 0.5$, est constante si $\alpha = 0.5$, et tend vers l'infini si $\alpha > 0.5$. \square

Propriété 8. Soit $\alpha \in \mathbb{R}_*^+$, On considérera 3 comportements asymptotiques de $Pr\left(n, \frac{p}{n^\alpha}, \frac{q}{n^\alpha}\right)$ en fonction de α .

1. Si $\alpha < 1$: La proportion d'erreur maximale tend vers 0.
2. Si $\alpha = 1$: La proportion d'erreur maximale est constant pour tout n .
3. Si $\alpha > 1$: La proportion d'erreur maximale diverge vers l'infini.

Démonstration. De la même manière que précédemment, nous calculons simplement la limite de $Pr\left(n, \frac{p}{n^\alpha}, \frac{q}{n^\alpha}\right)$ lorsque n tend vers l'infini. Il y a 2 possibilités en fonction de la valeur de $\min\left\{\frac{p}{n^\alpha}, \frac{p-q}{2n^\alpha}\right\}$.

$$\begin{aligned}\lim_{n \rightarrow \infty} Pr\left(n, \frac{p}{n^\alpha}, \frac{q}{n^\alpha}\right) &= \begin{cases} \lim_{n \rightarrow \infty} \frac{16(\frac{q}{n^\alpha}(1 - \frac{q}{n^\alpha}) + \frac{p}{n^\alpha}(1 - \frac{p}{n^\alpha}))}{n^{\frac{q^2}{4n^{2\alpha}}}} & \text{si } q \leq \frac{p-q}{2} \\ \lim_{n \rightarrow \infty} \frac{16(\frac{q}{n^\alpha}(1 - \frac{q}{n^\alpha}) + \frac{p}{n^\alpha}(1 - \frac{p}{n^\alpha}))}{n^{\frac{(p-q)^2}{4n^{2\alpha}}}} & \text{si } q \geq \frac{p-q}{2} \end{cases} \\ &= \begin{cases} \lim_{n \rightarrow \infty} \frac{16(\frac{q}{n^\alpha} + \frac{p}{n^\alpha})}{n^{1-2\alpha}q^2} & \text{si } q \leq \frac{p-q}{2} \\ \lim_{n \rightarrow \infty} \frac{64(\frac{q}{n^\alpha} + \frac{p}{n^\alpha})}{n^{1-2\alpha}(p-q)^2} & \text{si } q \geq \frac{p-q}{2} \end{cases} \\ &= \begin{cases} \lim_{n \rightarrow \infty} \frac{16(p+q)}{n^{1-\alpha}q^2} & \text{si } q \leq \frac{p-q}{2} \\ \lim_{n \rightarrow \infty} \frac{64(p+q)}{n^{1-\alpha}(p-q)^2} & \text{si } q \geq \frac{p-q}{2} \end{cases}\end{aligned}$$

Donc ceci, tend donc vers 0 si $\alpha < 1$, est constante si $\alpha = 1$, et tend vers l'infini si $\alpha > 1$. \square

5 Discussion

Une manière de généraliser le problème est de changer la taille des communautés. Prenons par exemple 2 communautés respectivement de taille βn et $(1 - \beta)n$ avec $\beta \in]0, 1[$, est-il possible de retrouver les communautés ? De la même manière que pour le problème initial, nous définissons $X_n = \hat{X}_n + R_n$, u_2 (resp v_2) le vecteur propre de \hat{X}_n (resp X_n) associé à sa deuxième valeur propre.

Ici aussi u_2 permet de connaître les communautés, par exemple dans le cas où \hat{X}_n est organisé, nous avons $u_2 = \left(\underbrace{1 \cdots 1}_{\beta n \text{ fois}} \mid \underbrace{-1 \cdots -1}_{(1-\beta n) \text{ fois}} \right)^T$. De la même manière que pour le cas à deux communautés de mêmes tailles, nous voulons montrer que v_2 admet les mêmes propriétés que u_2 .

Nous allons donc étendre les Résultats 1 (non-asymptotique) et 2 (asymptotique), et majorer le nombre d'erreurs à l'aide du Théorème 1 et de l'étude de la norme de R_n .

Dans le cas non-asymptotique, Section 3, la majoration de $\|R_n\|$ est toujours valide. La borne de Résultat 1, ne dépend que d'un coefficient de la matrice R_n : $\max_{1 \leq i, j \leq n} \|R_{ij}\|_{\psi_2}$

Enfin pour le cas asymptotique, nous cherchons, comme dans la précédente section, la mesure spectrale de R_n en utilisant la transformée de Stieltjes. Ainsi nous sommes de nouveau amenés à traiter l'équation K1. Après simplification de celle-ci nous obtenons que la transformée de la mesure spectrale de \bar{R}_n converge presque sûrement vers $S_{\mu_{\bar{R}_n}}(z)$ tel que,

$$S_{\mu_{\bar{R}_n}}(z) = \beta c_1(z) + (1 - \beta)c_2(z)$$

Avec

$$\begin{cases} \beta c_1(z) = \frac{-\beta}{z+p(1-p)\beta c_1(z)+(1-\beta)q(1-q)c_2(z)} \\ (1-\beta)c_2(z) = \frac{-(1-\beta)}{z+(1-\beta)p(1-p)c_2(z)+\beta q(1-q)c_1(z)} \end{cases}$$

Contrairement au cas où les communautés sont de même taille, cette QVE est plus difficile à résoudre, nous allons donc la résoudre numériquement. Pour faire cela, posons $p(1-p) = P$, $q(1-q) = Q$, $(1-\beta)c_2(z) = x$ et $\beta c_1(z) = y$. Nous obtenons donc,

$$\begin{cases} x = \frac{-(1-\beta)}{z+Px+Qy} \\ y = \frac{-\beta}{z+Py+Qx} \end{cases} \Leftrightarrow \begin{cases} y = -\frac{xz+X^2P+(1-\beta)}{xQ} \\ zy + Py^2 + Qxy + \beta = 0 \end{cases}$$

En fusionnant les équations, nous obtenons un polynôme de degré 4,

$$[P^3 - Q^2P]x^4 + [2P^2z - PQz - Q^2z]x^3 + [Pz^2 + 2P^2(1-\beta) - Qz^2 + Q^2(2\beta-1)]x^2 + [2Pz(1-\beta) - Qz + Qz(\beta-1)]x + P(1+\beta^2-2\beta) = 0$$

En trouvant les racines du polynôme numériquement, nous obtenons quatre couples de valeurs possibles pour x et y , nous prenons donc l'unique solution dans \mathbb{L} .

Nous avons donc avec z , p et q fixés, la valeur de $S_{\mu_{\bar{R}_n}} = x + y$. Le résultat étant loin de la transformée de la loi du demi-cercle, on ne peut pas utiliser la continuité de la transformée de Stieljes pour connaître la mesure spectrale. Nous devons donc trouver une formule d'inversion pour passer de la transformée de Stieljes à la mesure spectrale.

Propriété 9 (Inversion de Stieljes-Perron (admis)). *Soit $\mu(x)$ une mesure spectrale et $S_\mu(z)$ sa transformée de Stieljes, alors,*

$$\lim_{\varepsilon \rightarrow 0} \Im \left(\frac{S(x + \varepsilon i)}{\pi} \right) = \mu(x)$$

Nous pouvons donc utiliser la formule d'inversion de Stieljes-Perron pour connaître, pour x , p et q fixés, la valeur de $\mu_{\bar{R}_n}(x)$.

Formes remarquables de mesures spectrales avec $\beta = \frac{1}{3}$
n=3000

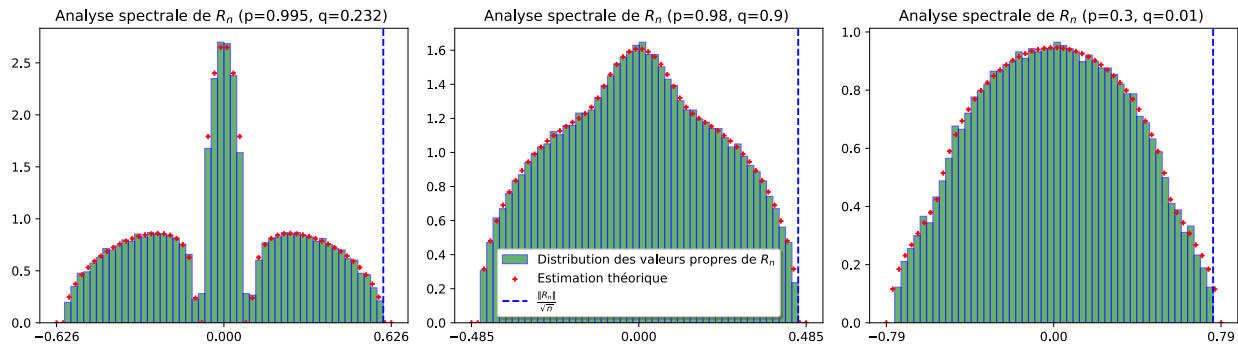


FIGURE 4 – Estimations et formes atypiques de $\mu_{\bar{R}_n}$ avec $\beta = \frac{1}{3}$

La mesure spectrale de \bar{R}_n converge presque sûrement vers $\mu_{\bar{R}_n}(x)$, donc $\|R_n\| \xrightarrow[n \rightarrow \infty]{\mathbb{P}} \sqrt{n} \max \{|x|, \mu_{\bar{R}_n}(x) > 0\}$. Cette technique est très précise mais ne permet pas d'avoir une formule analytique de la norme de R_n . Une autre méthode peut cependant être utilisée pour avoir une majoration efficace de $\|R_n\|$ en utilisant des résultats déjà connus. Pour cela nous allons utiliser le théorème d'entrelacement de Cauchy. Afin de ne pas

rallonger le rapport, nous omettons la preuve, bien qu'il soit intéressant de noter qu'elle découle directement du Théorème 3 (Courant Fisher, min-max).

Théorème 10 (Théorème d'entrelacement de Cauchy). *Soit $A \in \mathbb{M}_n(\mathbb{R})$ une matrice symétrique et B une sous-matrice de A ¹². Soit λ_i les valeurs propres de A et α_j les valeurs propres de B classé par ordre croissant¹³,*

$$\lambda_1 \leq \alpha_1 \leq \lambda_2 \leq \dots \leq \alpha_{n-1} \leq \lambda_n$$

Un corollaire immédiat du Théorème d'entrelacement de Cauchy est que pour A une matrice symétrique et B une sous sous matrice de A , $\|A\| \leq \|B\|$

Ainsi, dans le cas où les deux communautés sont de tailles différentes, nous associons la matrice de perturbation R . Considérons alors \tilde{R} , la sous-matrice de R correspondant au cas où les deux communautés ont la même taille, c'est-à-dire de taille $\max\{\beta, 1 - \beta\} \cdot n$. Nous avons donc,

$$\|R\| \leq \|\tilde{R}\| \xrightarrow[n \rightarrow \infty]{\mathbb{P}} \sqrt{2(p(1-p) + q(1-q))} \cdot \sqrt{2 \max\{\beta, (1-\beta)\} n}$$

Ainsi, par une preuve analogue à la démonstration du résultat 2, nous obtenons avec m le nombre d'erreur,

$$\mathbb{P} \left(m \leq \max\{\beta, (1-\beta)\} \cdot \frac{32(p(1-p) + q(1-q))}{\min\{q, \frac{p-q}{2}\}^2} \right) \xrightarrow[n \rightarrow \infty]{} 1$$

Cette méthode est moins précise que la recherche exacte de la norme par la mesure spectrale, mais elle est efficace car elle démontre que toutes les propriétés de la proportion d'erreur maximale sont préservées même lorsque les communautés sont de tailles différentes.

FIN.

Références

- [1] Krüger Torben/ László Erdős/ Oskari Ajanki. *Quadratic vector equations on complex upper half-plane*. AMS, 2019.
- [2] Vyacheslav.L Girko. *Mathematics and its Applications : Theory of Stochastic Canonical Equation Volume 1*. Springer Science+Business Media, 2001.
- [3] Konstantin Avrachenkov/ Laura Cottatellucci / Arun Kadavankandy. *Spectral Properties of Random Matrices for Stochastic Block Model*. INRIA, 2015.
- [4] Courtney Paquette. High dimensional probability notes, 2019. Assistant Professor, Mathematics and Statistics.
- [5] Roman Vershynin. *High-Dimensional Probability : An Introduction with Applications in Data Science*. Cambridge University Press, Cambridge, 2018.

¹². Matrice obtenue en supprimant un ensemble quelconque de lignes ainsi que les colonnes correspondantes de A .

¹³. De la même manière que pour le théorème de Courant Fisher, on inverse le sens habituel par soucis de lisibilité