

Matrices densités aléatoires

Ion NECHITA

juin 2006

SUJET PROPOSÉ PAR STEPHANE ATTAL

Table des matières

1	Matrices densités et mécanique quantique	2
1.1	Axiomes de la mécanique quantique - systèmes fermés	2
1.2	Systèmes ouverts	3
1.3	États intriqués et entropie de von Neumann	5
2	Mesures canoniques sur l'ensemble des matrices densités	5
2.1	Action de groupe et mesures invariantes	5
2.2	Mesure canonique sur les états purs	5
3	Matrices aléatoires de Wishart	8
3.1	Introduction aux matrices aléatoires	8
3.2	Matrices de Wishart.....	8
3.3	Loi jointe des valeurs propres	9
4	Etude de la loi des matrices densités aléatoires	9
4.1	Loi des valeurs propres	9
4.2	Moments	10
4.3	Entropie	12
5	Lois limites	12
6	Applications complètement positives aléatoires	15
7	Simulations numériques	15
8	Appendices	15
8.1	L'espace projectif complexe	15
8.2	La trace partielle	16
8.3	Notation bra-ket de Dirac	16
8.4	Polynômes orthogonaux de Laguerre	17

1 Matrices densités et mécanique quantique

Le but de cette première partie est d'introduire les matrices densités, l'objet principal de notre étude. On va présenter le langage mathématique de la mécanique quantique, sans faire référence (dans la mesure du possible) à une interprétation ou une autre.

1.1 Axiomes de la mécanique quantique - systèmes fermés

La mécanique quantique, après plus d'un siècle de sa découverte, est aujourd'hui peut être la meilleure théorie physique dont nous disposons. Ses prévisions se sont montrées exactes et ont contribué de façon cruciale au développement de notre société (il suffit ici de mentionner le transistor). Même si les scientifiques qui utilisent la mécanique quantique sont tous d'accord sur le formalisme mathématique à utiliser, l'*interprétation* de la théorie est le sujet des disputes.

Dans ce qui suit, comme c'est justement le cadre mathématique qui nous intéresse, on va essayer de se débarrasser de toute interprétation en présentant le formalisme comme une théorie mathématique abstraite. Pour assurer la lisibilité du texte, on fera tout de même référence aux objets réalistes, comme "vecteur d'état", "énergie du système", etc.

Dans ce premier formalisme on s'intéressera aux systèmes *fermés*, c'est à dire aux systèmes qui sont isolés de leur environnement.

Premier axiome - états

En mécanique quantique, l'ensemble de tous les états possible d'un système physique (ou l'espace des phases) est un espace de Hilbert \mathcal{H} sur le corps des nombres complexes \mathbb{C} . Plus précisément, deux vecteurs égaux à un scalaire complexe près représentent le même état quantique. Certains ouvrages présentent la situation en disant que l'état d'un système quantique est un vecteur de norme 1 dont la phase est indéterminée. De façon équivalente, nous on va adopter le formalisme (plus rigoureux) suivant : deux vecteurs x et y d'un espace de Hilbert \mathcal{H} de dimension sont dits *équivalents* s'il existe un scalaire $\lambda \in \mathbb{C}$ tel que $y = \lambda x$. On écrit $x \sim y$.

Définition 1.1 (espace d'états pur, état pur). L'espace quotient \mathcal{H}/\sim est appelé *espace d'états purs* et il sera noté par \mathcal{E} . Un *état pur* sera une classe d'équivalence de \mathcal{E} et il sera noté par $|\varphi\rangle$ où φ est un représentant de la classe. Traditionnellement, on prend φ de norme 1 et on dit que sa phase globale est *indéterminée*. On munit \mathcal{E} de la topologie quotient.

Par la suite on va s'intéresser aux systèmes quantiques avec un nombre fini de degrés de liberté, donc tous les espaces de Hilbert qui apparaîtront seront de dimension finie. On notera avec \mathcal{H}_n un espace de Hilbert complexe de dimension n et avec \mathcal{E}_n l'espace d'états d'un système quantique à n degrés de liberté.

Bien évidemment, on a les isomorphismes $\mathcal{H}_n \approx \mathbb{C}^n$ et $\mathcal{E}_n \approx \mathbb{CP}^{n-1}$ (pour des rappels sur l'espace projectif complexe, voir l'appendice 8.1). On va utiliser implicitement ces isomorphismes dans les parties suivantes et regarder, par exemple, un état pur comme la classe d'un vecteur (non nul) de \mathbb{C}^n .

Le produit scalaire de \mathcal{H}_n ne se projette pas sur \mathcal{E}_n (à cause de l'indétermination de la phase), mais on peut définir l'application suivante, appelée *fidélité* :

$$F : \mathcal{E}_n \times \mathcal{E}_n \longrightarrow [0, 1]$$

$$(|\psi\rangle, |\varphi\rangle) \longmapsto |\langle \psi_0 | \varphi_0 \rangle|^2,$$

où φ_0 et ψ_0 sont des représentants *de norme 1* des classes $|\varphi\rangle$ et $|\psi\rangle$.

Deuxième axiome - dynamique

Un système quantique *fermé* évolue dans le temps selon :

$$|\psi_t\rangle = U_t|\psi_0\rangle, \quad (1.1)$$

où $|\psi_0\rangle$ est l'état du système au temps 0, $|\psi_t\rangle$ est l'état du système au temps t , et U_t est une matrice unitaire qui dépend du temps.

L'équation 1.8 est appelée *l'équation de Schrödinger*.

Troisième axiome - mesures

Les mesures en mécanique quantique occupent une place spéciale dans la théorie. C'est à cet endroit que les probabilités entrent en scène, faisant de la mécanique quantique une théorie non-déterministe.

On appelle *observable* toute quantité physique susceptible à être mesurée ; par exemple la position, la vitesse et le spin sont des observables. A toute observable d'un système quantique on associe un opérateur auto-adjoint A (appelé aussi observable) sur l'espace de Hilbert qui décrit le système. A admet une décomposition spectrale

$$A = \sum_i \lambda_i P_i, \quad (1.2)$$

où $\{\lambda_i\}_i$ est l'ensemble des valeurs propres (réelles) de A et P_i sont des projecteurs orthogonaux dont les rangs sont égaux aux multiplicités des λ_i .

Si on fait une mesure de l'observable A , et le système se trouve dans l'état $|\psi\rangle$, on va obtenir une des valeurs $\{\lambda_i\}_i$ avec les probabilités respectives :

$$\mathbb{P}(\lambda_i) = \|P_i\psi\|^2, \quad (1.3)$$

et, au cas où on obtient le résultat λ_i , l'état du système deviendra

$$|\psi_i\rangle = \frac{P_i\psi}{\|P_i\psi\|}. \quad (1.4)$$

Remarque 1.2. Dans les équations précédentes, comme dans tout ce qui suit, on considère que tous les vecteurs ψ , φ sont des vecteurs de \mathcal{H} de norme 1 et dont la phase est indéterminée.

1.2 Systèmes ouverts

Dans l'étude des systèmes quantiques, il se trouve qu'on a souvent à faire à des systèmes qui ne sont pas isolés de leur environnement. Ou, il se peut encore qu'on a à traiter un système dont on n'a pas accès à son intégralité, mais juste à un de ses sous-systèmes. Tous ces cas sont décrits par la théorie des systèmes quantiques ouverts, dont on va décrire le formalisme dans cette section.

Pour commencer, regardons un exemple très simple. Il vient de la théorie quantique de l'information, où on s'intéresse le plus souvent aux systèmes simples, formés des *qubits*. Un qubit est un système quantique à deux niveaux, pour lequel on met en évidence une base particulière de son espace de Hilbert, notée $\{|0\rangle, |1\rangle\}$. C'est l'équivalent quantique d'un bit d'information classique, qui permet d'avoir des superpositions de ses deux états de base, $|0\rangle$ et $|1\rangle$. Considérons maintenant un système plus grand, composé de n qubits. Une base de l'espace de Hilbert qui décrit ce système est

$$\{|x_1 x_2 \dots x_n\rangle : x_i = 0 \text{ ou } 1\}.$$

Pour mettre en évidence un système ouvert, considérons l'état suivant :

$$|\Phi^+\rangle = \frac{1}{\sqrt{2}}(|00\rangle + |11\rangle).$$

C'est un état d'un système à 4 niveaux qui est composé de deux qubits. Supposons maintenant qu'on a accès qu'au premier qubit de ce système, pour une raison ou une autre (peut être que le deuxième qubit se trouve dans une autre galaxie, etc). Que peut-on dire du premier qubit ? Ait-on le droit de parler de son état ?

Il est facile de vérifier qu'on ne peut pas trouver des états $|\varphi\rangle, |\psi\rangle \in \mathcal{H}_2^{(1)}$ telles que $|\Phi^+\rangle = |\varphi\rangle \otimes |\psi\rangle$. Ce phénomène, appelé *intrication* ou *enchevêtrement*, est une des sources de l'étrangeté quantique. On aimerait quand même être capables de décrire des systèmes ouverts. C'est à cet endroit que les matrices densités entre en jeu : elles sont des généralisations des états introduits dans la section précédente.

Désormais, on appellera les vecteurs kets des états *purs*. A un tel état pur $|\varphi\rangle$ on associe l'opérateur $|\varphi\rangle\langle\varphi|$ appelé *matrice densité*. Il est facile de voir que $|\varphi\rangle\langle\varphi|$ est auto-adjoint, positif et de trace 1. En utilisant ces propriétés, on généralise la notion d'état quantique pur en introduisant :

Premier axiome - systèmes ouverts - états

Définition 1.3 (matrice densité). L'état d'un système quantique est un opérateur ρ hermitien, positif et de trace unité sur \mathcal{H} , appelé matrice densité.

Une matrice densité se décompose toujours comme

$$\rho = \lambda_1|\varphi_1\rangle\langle\varphi_1| + \lambda_2|\varphi_2\rangle\langle\varphi_2| + \dots + \lambda_n|\varphi_n\rangle\langle\varphi_n| \quad (1.5)$$

Si tous les λ_i sont nuls sauf un seul qui vaut 1, alors ρ est de la forme $|\varphi\rangle\langle\varphi|$ et il est appelé *état pur*. Sinon, ρ est dit *mélangé*.

L'état d'un sous-système quantique est introduit à l'aide de la *trace partielle* :

Définition 1.4 (trace partielle, état d'un sous-système). Soit AB un système composé qui se trouve dans un état ρ_{AB} . Alors les sous-systèmes A et B se trouvent dans les états

$$\rho_A = \text{Tr}_B(\rho_{AB}), \quad (1.6)$$

et

$$\rho_B = \text{Tr}_A(\rho_{AB}) \quad (1.7)$$

On note \mathcal{D}_n l'ensemble des matrices densité $n \times n$, c'est à dire l'espace des matrices positives de trace 1 ; c'est l'ensemble de tous les états possibles d'un système quantique ouvert à n degrés de liberté.

Les deux autres axiomes sont obtenus comme généralisations des axiomes vus dans la section précédente.

Deuxième axiome - systèmes ouverts - dynamique

Un système quantique ouvert évolue dans le temps selon :

$$\rho_t = U_t \rho_0 U_t^*, \quad (1.8)$$

où ρ_0 est l'état du système au temps 0, ρ_t est l'état du système au temps t , et U_t est une matrice unitaire qui dépend du temps.

Troisième axiome - systèmes ouverts - mesures

Considérons un système (ouvert) qui se trouve dans un état ρ et pour lequel on veut mesurer une observable A , donnée par l'équation 1.2. On obtient une des valeurs $\{\lambda_i\}_i$ avec les probabilités respectives :

$$\mathbb{P}(\lambda_i) = \text{Tr}(P_i\rho), \quad (1.9)$$

et, au cas où on obtient le résultat λ_i , l'état du système deviendra

$$\rho_i = \frac{P_i\rho P_i}{\text{Tr}(P_i\rho)}. \quad (1.10)$$

1.3 États intriqués et entropie de von Neumann

L'intrication est un des phénomènes phares de l'information quantique.

2 Mesures canoniques sur l'ensemble des matrices densités

2.1 Action de groupe et mesures invariantes

Ici, on fait une étude générale des mesures invariantes sous l'action d'un groupe. Pour cela, fixons d'abord le cadre général. On va considérer un groupe topologique G , localement compact et séparable qui agit sur un ensemble S qui est lui aussi localement compact et séparable. Cette action est dite *transitive* si pour tous $x, y \in S$, il existe un élément $g \in G$ tel que $y = gx$. L'action est dite *propre* si pour tout élément $x \in S$, l'application $m_x : G \rightarrow S$ définie par $m_x(g) = gx$ est propre, i.e. l'image réciproque d'un compact est compacte. On munit G et S de leurs tribus boreliennes \mathcal{G} et \mathcal{S} .

Définition 2.1 (mesure invariante). Une mesure ν sur (S, \mathcal{S}) est dite *G-invariante* (à gauche) si pour tous $g \in G$ et $A \in \mathcal{S}$ on a

$$\nu(gA) = \nu(A),$$

ou encore, de façon équivalente, pour toute fonction $f \in C_c(G, S)$,

$$\int f(xs)\nu(ds) = \int f(s)\nu(ds)$$

Par exemple, la mesure de Haar sur un groupe compact G est G -invariante pour l'action à gauche de G sur lui-même. On se pose la question de l'existence et de l'unicité d'une mesure G -invariante.

Théorème 2.2 (existence et unicité de la mesure invariante). *Soit G un groupe topologique localement compact et séparable qui agit transitivement et proprement sur un espace topologique S localement compact et séparable. Alors il existe une unique (à une constante près) mesure G -invariante sur S .*

Preuve. à faire ???

2.2 Mesure canonique sur les états purs

La première étape dans notre étude sera de trouver et d'expliciter une mesure sur l'espace des états purs \mathcal{E}_n . La contrainte d'invariance unitaire va déterminer la mesure; on parlera donc de *la mesure canonique sur l'espace des états purs*.

Il vient des axiomes de la mécanique quantique que les transformations temporelles d'un système quantique fermé sont données par des matrices unitaires. On va introduire ici le groupe $\mathcal{U}(n)$ des matrices unitaires de dimension n et on va présenter ses propriétés utilisées par la suite.

Définition 2.3 (groupe des matrices unitaires). L'ensemble des matrices unitaires à n dimensions $\mathcal{U}(n)$ forme un groupe pour la multiplication des matrices. Ce groupe est compact.

Etant un groupe topologique compact, $\mathcal{U}(n)$ possède une mesure de Haar :

Proposition 2.4 (mesure de Haar sur $\mathcal{U}(n)$). Il existe une mesure de probabilité sur le groupe $\mathcal{U}(n)$, notée h et appelée mesure de Haar qui est invariante par l'action du groupe, ie. pour toute matrice unitaire U et pour tout sous ensemble mesurable B de $\mathcal{U}(n)$, on a :

$$h(U \cdot B) = h(B),$$

où $U \cdot B = \{UA | A \in B\}$.

On a déjà vu que la mesure uniforme sur \mathcal{E}_n doit être invariante par l'action du groupe $\mathcal{U}(n)$ des matrices unitaires. On va montrer l'existence et l'unicité d'une telle mesure. Pour cela, on va vérifier les hypothèses du théorème 2.2.

Pour construire l'action de $\mathcal{U}(n)$ sur \mathcal{E}_n on va factoriser l'action de $\mathcal{U}(n)$ sur \mathcal{H}_n . Pour $U \in \mathcal{U}(n)$ fixé, on a :

$$\begin{array}{ccc} \mathcal{H}_n & \xrightarrow{\Pi_n} & \mathcal{E}_n \\ m_U \downarrow & & \downarrow \tilde{m}_U \\ \mathcal{H}_n & \xrightarrow{\Pi_n} & \mathcal{E}_n \end{array}$$

avec

$$\begin{aligned} m_U : \mathcal{H}_n &\longrightarrow \mathcal{H}_n \\ \varphi &\longmapsto U\varphi \end{aligned}$$

et

$$\begin{aligned} \tilde{m}_U : \mathcal{E}_n &\longrightarrow \mathcal{E}_n \\ |\varphi\rangle &\longmapsto |U\varphi\rangle \end{aligned}$$

L'application \tilde{m}_U est bien définie, car si $\varphi, \psi \in \mathcal{H}_n$ sont tels que $\psi = \lambda\varphi$, alors $|U\psi\rangle = |\lambda U\varphi\rangle = |U\varphi\rangle$, et donc le diagramme ci-dessus commute.

On vient de construire l'action de $\mathcal{U}(n)$ sur \mathcal{E}_n :

$$\mathcal{U}(n) \times \mathcal{E}_n \longrightarrow \mathcal{E}_n \quad (2.11)$$

$$(U, |\varphi\rangle) \longmapsto |U\varphi\rangle \quad (2.12)$$

En plus, et cela n'est pas le cas de l'action de $\mathcal{U}(n)$ sur \mathcal{H}_n , cette action est *transitive* : si $|\varphi\rangle$ et $|\psi\rangle$ sont deux états purs quelconques de \mathcal{E}_n , alors il existe un unitaire U de $\mathcal{U}(n)$ tel que $|\psi\rangle = |U\varphi\rangle$. Par exemple, on peut prendre pour U la matrice de changement de base qui transforme une base commençant par φ en une base commençant par ψ . Il reste à montrer que l'action 2.11 est aussi *propre*. Fixons donc $U \in \mathcal{U}(n)$; alors on a

$$\begin{array}{ccc} \mathcal{U}(n) & \xrightarrow{\tilde{m}_U} & \mathcal{E}_n \\ & \searrow m_U & \uparrow \Pi_n \\ & & \mathcal{H}_n \end{array}$$

Comme m_U et Π_n sont continues, \tilde{m}_U l'est aussi, et pour tout compact K de \mathcal{E}_n , $\tilde{m}_U^{-1}(K)$ est fermé, et donc compact dans $\mathcal{U}(n)$. L'application \tilde{m}_U est donc propre.

L'action des unitaires sur l'espace des états purs vérifie donc les deux hypothèses du théorème 2.2. En choisissant une normalisation convenable, on a le résultat fondamental suivant :

Théorème 2.5 (la mesure invariante dans l'espace des états purs). *Sur l'espace des états purs \mathcal{E}_n il existe une unique mesure de probabilités μ_n invariante par les unitaires.*

Il est possible de voir la mesure invariante μ_n comme la mesure image de la loi d'un vecteur gaussien complexe par la projection canonique Π_n . Pour cela, étant donné que μ_n est l'unique mesure sur \mathcal{E}_n invariante par les unitaires, il suffit de montrer le résultat suivant :

Proposition 2.6. *La mesure image $(\Pi_n)_\#(\mathcal{N}_{\mathbb{C}}(0, 1)^{\otimes n})$ est invariante par l'action du groupe unitaire $\mathcal{U}(n)$ sur \mathcal{E}_n .*

Preuve. Soit $f : \mathcal{E}_n \rightarrow \mathbb{R}_+$ une fonction mesurable. Alors, pour $U \in \mathcal{U}(n)$, on a

$$\begin{aligned} \int (f \circ \tilde{m}_U)(|\varphi\rangle) \Pi_{n\#} \mathcal{N}_{\mathbb{C}}(0, 1)^{\otimes n}(d|\varphi\rangle) &= \int (f \circ \tilde{m}_U \circ \Pi_n)(z_1, \dots, z_n) \mathcal{N}_{\mathbb{C}}(0, 1)^{\otimes n}(dz_1, \dots, dz_n) \\ &= \int (f \circ \Pi_n \circ m_U)(z_1, \dots, z_n) \mathcal{N}_{\mathbb{C}}(0, 1)^{\otimes n}(dz_1, \dots, dz_n) \\ &= \int (f \circ \Pi_n)(z_1, \dots, z_n) m_{U\#} \mathcal{N}_{\mathbb{C}}(0, 1)^{\otimes n}(dz_1, \dots, dz_n) \\ &\stackrel{(*)}{=} \int (f \circ \Pi_n)(z_1, \dots, z_n) \mathcal{N}_{\mathbb{C}}(0, 1)^{\otimes n}(dz_1, \dots, dz_n) \\ &= \int f(|\varphi\rangle) \Pi_{n\#} \mathcal{N}_{\mathbb{C}}(0, 1)^{\otimes n}(d|\varphi\rangle), \end{aligned}$$

où j'utilise, à l'étape (*), le fait que la loi de n variables gaussiennes complexes indépendantes est invariante par les transformations unitaires. La plus simple façon de se convaincre de cela est d'utiliser les fonctions caractéristiques ; si $Z = (z_1, \dots, z_n) \sim \mathcal{N}_{\mathbb{C}}(0, 1)^{\otimes n}$, alors

$$\Phi_Z(\xi) = \exp\left(-\frac{1}{4}|\xi|^2\right),$$

et, pour U unitaire

$$\Phi_{UZ}(\xi) = \exp\left(-\frac{1}{4}|U^* \xi|^2\right) = \exp\left(-\frac{1}{4}|\xi|^2\right) = \Phi_Z(\xi).$$

□

On vient donc de montrer

Théorème 2.7. *La loi invariante sur \mathcal{E}_n est donnée par $\mu_n = \Pi_{n\#} \mathcal{N}_{\mathbb{C}}(0, 1)^{\otimes n}$.*

Utilisons ce résultat pour calculer la loi et l'espérance de la fidélité de l'approximation d'un état aléatoire par un état fixe $|\psi\rangle$:

$$\begin{aligned} F_{|\psi\rangle} : \mathcal{E}_n &\longrightarrow [0, 1] \\ |\varphi\rangle &\longmapsto F(|\psi\rangle, |\varphi\rangle) = \frac{|\langle \psi | \varphi \rangle|^2}{\|\psi\|^2 \|\varphi\|^2}. \end{aligned}$$

Pour cela, considérons une fonction mesurable positive $f : \mathbb{R} \rightarrow \mathbb{R}_+$ et un unitaire $U \in \mathcal{U}(n)$. On a

$$\begin{aligned} \int (f \circ F_{|\psi\rangle})(|\varphi\rangle)\mu_n(d|\varphi) &= \int (f \circ F_{|\psi\rangle} \circ \tilde{m}_U)(|\varphi\rangle)\mu_n(d|\varphi) \\ &= \int (f \circ F(|U^*\psi\rangle, \cdot))(|\varphi\rangle)\mu_n(d|\varphi) \\ &= \int (f \circ F_{|U^*\psi\rangle})(|\varphi\rangle)\mu_n(d|\varphi), \end{aligned}$$

donc la loi de $F_{|\psi\rangle}$ ne dépend pas de $|\psi\rangle$. On peut donc choisir pour ψ le premier vecteur dans la base orthonormale de $\mathbb{C}^n \approx \mathcal{H}_n$. On trouve alors

$$\begin{aligned} \int (f \circ F_{|\psi\rangle})(|\varphi\rangle)\mu_n(d|\varphi) &= \int (f \circ F_{|\psi\rangle})(|\varphi\rangle)\Pi_{n\# \setminus \mathbb{C}}(0, 1)^{\otimes n}(d|\varphi) \\ &= \int f \left(\frac{|\langle \psi | Z \rangle|^2}{\|Z\|^2} \right) \mathcal{N}_{\mathbb{C}}(0, 1)^{\otimes n}(dZ) \\ &= \int f \left(\frac{|z_1|^2}{|z_1|^2 + \dots + |z_n|^2} \right) \mathcal{N}_{\mathbb{C}}(0, 1)^{\otimes n}(dz_1, \dots, dz_n). \end{aligned}$$

On conclut que la loi de $F_{|\psi\rangle}$ est égale à la loi de $\frac{|z_1|^2}{|z_1|^2 + \dots + |z_n|^2}$, où $Z = (z_1, \dots, z_n)$ est un vecteur complexe de loi $\mathcal{N}_{\mathbb{C}}(0, 1)^{\otimes n}$. Calculons l'espérance de $F_{|\psi\rangle}$:

$$\begin{aligned} \mathbb{E}_{\mu_n} [F_{|\psi\rangle}] &= \int \left(\frac{|z_1|^2}{|z_1|^2 + \dots + |z_n|^2} \right) \mathcal{N}_{\mathbb{C}}(0, 1)^{\otimes n}(dZ) \\ &= \int \left(\frac{|z_k|^2}{|z_1|^2 + \dots + |z_n|^2} \right) \mathcal{N}_{\mathbb{C}}(0, 1)^{\otimes n}(dZ) \\ &= \frac{1}{n} \int \sum_{k=1}^n \left(\frac{|z_k|^2}{|z_1|^2 + \dots + |z_n|^2} \right) \mathcal{N}_{\mathbb{C}}(0, 1)^{\otimes n}(dZ) \\ &= \frac{1}{n} \int 1 \cdot \mathcal{N}_{\mathbb{C}}(0, 1)^{\otimes n}(dZ) = \frac{1}{n} \end{aligned}$$

3 Matrices aléatoires de Wishart

3.1 Introduction aux matrices aléatoires

3.2 Matrices de Wishart....

Les matrices de la forme X^*X , appelées *matrices de Wishart* ont été étudiées en statistique et probabilités depuis les années 1930. On se propose ici de faire un résumé des propriétés les plus importantes des matrices de Wishart.

Définition 3.1. On appelle *matrice de Wishart* une matrice $W = X^*X$ où $X \in \mathcal{M}_{k \times n}(\mathbb{C})$ a des entrées i.i.d. $\setminus_{\mathbb{C}}(0, 1)$.

3.3 Loi jointe des valeurs propres

Ici on se propose de trouver la loi jointe des valeurs propres d'une matrice de Wishart. Comme ces matrices sont auto-adjointes, on va tout d'abord calculer la densité de leur loi par rapport à la mesure de Lebesgue

$$dW = \prod_{i=1}^n dW_{ii} \prod_{1 \leq i < j \leq n} d\operatorname{Re}(W_{ij}) d\operatorname{Im}(W_{ij}). \quad (3.13)$$

On va procéder par la méthode de la transformée de Laplace.

Proposition 3.2. *Si A est une matrice $n \times n$ auto-adjointe telle que $A \leq I$ et W est distribuée sous la mesure de Wishart, alors*

$$\mathbb{E}[\exp(\operatorname{Tr}(AW))] = \det(I - A)^{-p}.$$

Calculons maintenant la transformée de Laplace de la loi de densité

$$p(W) = C_{n,k}^{(W)} \exp(-\operatorname{Tr} W) (\det W)^{k-n}.$$

On trouve le même résultat que celui de la proposition 3.2. Comme la transformée de Laplace caractérise la loi, on vient de montrer

Proposition 3.3. *La densité de la loi de Wishart de paramètres n et k par rapport à la mesure de Lebesgue 3.13 est*

$$p^{(Leb)}(W) = C_{n,k}^{(Leb)} \exp(-\operatorname{Tr} W) (\det W)^{k-n}. \quad (3.14)$$

Un changement de variables classique nous donne la loi des valeurs propres d'une matrice de Wishart (n, k) :

Proposition 3.4. *Les valeurs propres (non-ordonnées) d'une matrice de Wishart W suivent une loi de densité*

$$\Phi_{n,k}^{(W)}(\lambda_1, \dots, \lambda_n) = C_{n,k}^{(W)} \exp\left(-\sum_{i=1}^n \lambda_i\right) \prod_{i=1}^n \lambda_i^{k-n} \prod_{1 \leq i < j \leq n} (\lambda_i - \lambda_j)^2. \quad (3.15)$$

4 Etude de la loi des matrices densités aléatoires

4.1 Loi des valeurs propres

On a vu (???) que la loi $\mu_{n,k}$ est la loi d'une matrice ρ obtenue comme

$$\rho = \frac{W}{\operatorname{Tr}(W)},$$

où W est une matrice de Wishart de paramètres n et k . Si on note $(\lambda_1, \dots, \lambda_n)$ les valeurs propres de W et $(\tilde{\lambda}_1, \dots, \tilde{\lambda}_n)$ celles de ρ , on a, pour tout i ,

$$\tilde{\lambda}_i = \frac{\lambda_i}{\sum_{j=1}^n \lambda_j}.$$

Le vecteur $(\tilde{\lambda}_1, \dots, \tilde{\lambda}_n)$ est situé sur le simplexe $(n-1)$ -dimensionnel $\Sigma_{n-1} = \{(x_1, \dots, x_n) \in \mathbb{R}_+^n : \sum_{i=1}^n x_i = 1\}$. On va montrer qu'il admet une densité par rapport à la mesure de Lebesgue $dx_1 \cdots dx_{n-1}$.

On va faire ce changement de variables étape par étape pour fixer les idées. Par la suite, pour alléger l'écriture, on notera $\Delta(\lambda) = \prod_{1 \leq i < j \leq n} (\lambda_i - \lambda_j)$.

Partons donc de la densité des valeurs propres d'une matrice de Wishart :

$$\Phi_{n,k}^{(W)}(\lambda_1, \dots, \lambda_n) = C_{n,k}^{(W)} \exp\left(-\sum_{i=1}^n \lambda_i\right) \prod_{i=1}^n \lambda_i^{k-n} \Delta(\lambda)^2.$$

On fait le changement de variables $(\lambda_1, \dots, \lambda_n) \mapsto (\lambda_1, \dots, \lambda_{n-1}, S)$, où $S = \sum_{i=1}^n \lambda_i$. La jacobienne de cette transformation vaut 1, donc on a

$$\Phi_{n,k}^{(\lambda,S)}(\lambda_1, \dots, \lambda_{n-1}, S) = C_{n,k}^{(W)} \exp(-S) \prod_{i=1}^n \lambda_i^{k-n} \Delta(\lambda)^2.$$

Ensuite on normalise les valeurs propres : $(\lambda_1, \dots, \lambda_{n-1}, S) \mapsto (\lambda_1/S, \dots, \lambda_{n-1}/S, S) = (\tilde{\lambda}_1, \dots, \tilde{\lambda}_{n-1}, S)$. Le calcul de la jacobienne donne $|\det J| = 1/S^{n-1}$. On a

$$\Phi_{n,k}^{(\tilde{\lambda},S)}(\tilde{\lambda}_1, \dots, \tilde{\lambda}_{n-1}, S) = C_{n,k}^{(W)} \exp(-S) \prod_{i=1}^n (S\tilde{\lambda}_i)^{k-n} \Delta(S\tilde{\lambda})^2 S^{n-1}.$$

Cette densité se factorise (et c'est le point crucial de la méthode) en :

$$\Phi_{n,k}^{(\tilde{\lambda},S)}(\tilde{\lambda}_1, \dots, \tilde{\lambda}_{n-1}, S) = C_{n,k}^{(W)} \times \left[\prod_{i=1}^n (\tilde{\lambda}_i)^{k-n} \Delta(\tilde{\lambda})^2 \right] \times [S^{nk-1} \exp(-S)].$$

On peut en tirer deux conclusions :

- La loi des valeurs propres normalisées (par la somme) et la somme des valeurs propres sont indépendantes, donc la loi des valeurs propres normalisées est la même que la loi des valeurs propres d'une matrice de Wishart conditionnée à $\text{Tr } W = 1$.
- Pour avoir la loi des valeurs propres, il suffit de considérer la marginale par rapport à S . Comme $\int_0^\infty S^{nk-1} e^{-S} dS = \Gamma(nk)$ on trouve

$$\Phi_{n,k}(\tilde{\lambda}_1, \dots, \tilde{\lambda}_{n-1}) = C_{n,k} \prod_{i=1}^n (\tilde{\lambda}_i)^{k-n} \Delta(\tilde{\lambda})^2.$$

Attention, ici $\tilde{\lambda}_n = 1 - \sum_{i=1}^{n-1} \tilde{\lambda}_i$ n'est pas une variable. La constante $C_{n,k}$ est donné par

$$C_{n,k} = \Gamma(nk) C_{n,k}^{(W)} = \frac{\Gamma(nk)}{\prod_{j=0}^{n-1} \Gamma(n+1-j) \Gamma(k-j)}.$$

4.2 Moments

On va calculer les moments de la loi $\mu_{n,k}$ à partir des moments pour la loi de Wishart :

$$\begin{aligned}
\mathbb{E}_{n,k}^{(W)}[\mathrm{Tr}(W^q)] &= \int \left(\sum_{i=1}^n \lambda_i^q \right) \Phi_{n,k}^{(W)}(\lambda_1, \dots, \lambda_n) d\lambda_1 \cdots d\lambda_n \\
&= n \int \lambda_1^q \Phi_{n,k}^{(W)}(\lambda_1, \dots, \lambda_n) d\lambda_1 \cdots d\lambda_n \\
&= n \int S^q \tilde{\lambda}_1^q \Phi_{n,k}^{(\tilde{\lambda}, S)}(\tilde{\lambda}_1, \dots, \tilde{\lambda}_{n-1}, S) d\lambda_1 \cdots d\tilde{\lambda}_{n-1} dS \\
&= n \int \tilde{\lambda}_1^q \Phi_{n,k}(\tilde{\lambda}_1, \dots, \tilde{\lambda}_{n-1}) d\lambda_1 \cdots d\tilde{\lambda}_{n-1} \times \frac{\int_0^\infty S^{NK+q-1} e^{-S} dS}{\int_0^\infty S^{NK-1} e^{-S} dS} \\
&= \mathbb{E}_{n,k}[\mathrm{Tr}(\rho^q)] \frac{\Gamma(NK+q)}{\Gamma(NK)}.
\end{aligned}$$

Finalement, on a

$$\mathbb{E}_{n,k}[\mathrm{Tr}(\rho^q)] = \frac{\mathbb{E}_{n,k}^{(W)}[\mathrm{Tr}(W^q)]}{NK(NK+1) \cdots (nk+q-1)}.$$

Explicitement, on a (voir [4])

$$\mathbb{E}_{n,K}[\mathrm{Tr}(\rho^q)] = \frac{\Gamma(nk)}{\Gamma(nk+q)} \sum_{j=1}^q (-1)^{j-1} \frac{[K+q-j]_q [n+q-j]_q}{(q-j)!(j-1)!},$$

où $[a]_q = a(a-1) \cdots (a-q+1)$.

Dans [4] on peut aussi trouver une formule de récurrence pour les moments $M_{n,k}^{(W)}(q) = \mathbb{E}_{n,k}^{(W)}[\mathrm{Tr}(W^q)]$:

$$M_{n,k}^{(W)}(q+1) = \frac{(2q+1)(n+k)}{q+2} M_{n,k}^{(W)}(q) + \frac{(q-1)(q^2 - (k-n)^2)}{q+2} M_{n,k}^{(W)}(q-1).$$

On en déduit une relation du même type pour les moments $M_{n,k}$ de la loi $\mu_{n,k}$:

$$M_{n,k}(q+1) = \frac{(2q+1)(n+k)}{(nk+q)(q+2)} M_{n,k}(q) + \frac{(q-1)(q^2 - (k-n)^2)}{(nk+q)(nk+q-1)(q+2)} M_{n,k}(q-1).$$

Cette relation de récurrence s'avère très utile pour les calculs. L'initialisation se fait selon $M_{n,k}(0) = M_{n,k}(1) = 1$. Voici un tableau avec les premiers moments :

q	$\mathbb{E}_{n,k}^{(W)}[\mathrm{Tr}(W^q)]$	$\mathbb{E}_{n,k}[\mathrm{Tr}(\rho^q)]$
1	nk	1
2	$n^2k + nk^2$	$\frac{n+k}{nk+1}$
3	$n^3k + 3n^2k^2 + nk^3 + nk$	$\frac{n^2+3nk+k^2+1}{(nk+1)(nk+2)}$
4	$n^4k + 6n^3k^2 + 6n^2k^3 + nk^4 + 5n^2k + 5nk^2$	$\frac{n^3+6n^2k+6nk^2+k^3+5n+5k}{(nk+1)(nk+2)(nk+3)}$

4.3 Entropie

En mécanique quantique on s'intéresse souvent à la quantité $S(\lambda_1, \dots, \lambda_n) = -\sum_{i=1}^n \lambda_i \log \lambda_i$, appelée *entropie de von Neumann*. On va calculer ici l'espérance de S pour la loi $\mu_{n,K}$. Page a conjecturé en [10] que

$$\mathbb{E}_{n,K}[S(\rho)] = \left(\sum_{j=K+1}^{NK} \frac{1}{j} \right) - \frac{n-1}{2K}.$$

Depuis, ce résultat a été démontré par [13], [14], [7]. Dans les deux derniers travaux, des techniques de matrices aléatoires et les polynômes de Laguerre sont utilisés.

5 Lois limites

Dans cette partie on montre la convergence de la mesure spectrale empirique des valeurs propres vers la loi de Marchenko-Pastur, quand n et k sont grands. On va s'inspirer du résultat analogue pour les matrices de Wishart.

Commençons par introduire la loi de Marchenko-Pastur de paramètre c :

Définition 5.1. Pour $c \in]0, \infty[$, on appelle *loi de Marchenko-Pastur* de paramètre c et on note μ_c la mesure de probabilité donnée par

$$\mu_c = \max\{1-c, 0\} \delta_0 + \frac{\sqrt{(x-a)(b-x)}}{2\pi x} \mathbf{1}_{[a,b]}(x) dx, \quad (5.16)$$

où $a = (\sqrt{c}-1)^2$ et $b = (\sqrt{c}+1)^2$.

Le théorème principal est le suivant

Théorème 5.2. Soit $c \in]0, \infty[$, et $(k(n))_n$ une suite d'entiers telle que $\lim_{n \rightarrow \infty} \frac{k(n)}{n} = c$. Soit $(\rho_n)_n$ une suite des matrices densités aléatoires telle que pour tout n , ρ_n suis la loi $\mu_{n,k(n)}$. On définit la mesure spectrale empirique (renormalisée) de ρ_n par

$$L_n = \frac{1}{n} \sum_{i=1}^n \delta_{cn \lambda_i(\rho_n)}, \quad (5.17)$$

où $\lambda_1(\rho_n) \geq \dots \geq \lambda_n(\rho_n)$ sont les valeurs propres de ρ_n . Alors, presque sûrement, $(L_n)_n$ converge faiblement vers la mesure de Marchenko-Pastur μ_c .

On va prouver ce théorème en deux étapes. Considérons d'abord la mesure de probabilité (déterministe) $\bar{L}_n = \mathbb{E}_{\mu_{n,k}}[L_n]$. Montrons que \bar{L}_n converge faiblement vers μ_c .

Proposition 5.3. Soient \bar{L}_n la mesure de probabilités définie ci-dessus et μ_c la loi Marchenko-Pastur de paramètre c . Alors, on a

1. \bar{L}_n converge en moments vers μ_c quand n tends vers l'infini.
2. μ_c est caractérisée par ses moments.

On en conclut que $\bar{L}_n \Rightarrow \mu_c$.

Preuve. ???The first part of the proposition is a rather direct application of the formulas in the Wishart case.

Le deuxième point de la proposition vient du fait que la mesure de Marchenko-Pastur est à support compact.

Pour la deuxième étape de la preuve du théorème 5.2, on va utiliser des techniques de concentration de la mesure. Le résultat le plus connu dans le domaine est le suivant (Pisier)

Proposition 5.4. Soit γ_n la mesure gaussienne standard sur \mathbb{R}^n et soit $f : \mathbb{R}^n \rightarrow \mathbb{R}$ une fonction lipschitzienne avec constante de Lipschitz égale à L . Alors, pour tout $t > 0$, on a

$$\gamma_n(f - \mathbb{E}_{\gamma_n}[f] > t) \leq \exp\left(-\frac{2t^2}{\pi^2 L^2}\right), \quad (5.18)$$

et aussi

$$\gamma_n(|f - \mathbb{E}_{\gamma_n}[f]| > t) \leq 2 \exp\left(-\frac{2t^2}{\pi^2 L^2}\right). \quad (5.19)$$

Preuve. ??? à faire

Ensuite, on démontre un résultat de ce type adapté à notre situation (mais intéressant en lui même aussi) :

Proposition 5.5. Soit γ_n la mesure gaussienne standard sur \mathbb{R}^n et soit $r > 0$ un réel positif. Considérons une fonction lipschitzienne $f : \mathbb{R}^n \rightarrow \mathbb{R}$ avec une constante de Lipschitz égale à $L > 0$ et telle que f est bornée par $M > 0$ sur $\mathbb{R}^n - B_r$, où B_r est la boule de rayon r dans \mathbb{R}^n . Alors, pour tout $t > 2M\gamma_n(B_r)$, on a

$$\gamma_n(|f(x) - \mathbb{E}_{\gamma_n}[f]| > t \text{ and } \|x\| \geq r) \leq 2 \exp\left(-\frac{2(t - 2M\gamma_n(B_r))^2}{\pi^2 L^2}\right). \quad (5.20)$$

Preuve. On va montrer en fait que

$$\gamma_n(f(x) - \mathbb{E}_{\gamma_n}[f] > t \text{ and } \|x\| \geq r) \leq \exp\left(-\frac{2(t - 2M\gamma_n(B_r))^2}{\pi^2 L^2}\right), \quad (5.21)$$

pour conclure ensuite en remplaçant f par $-f$.

Pour faire bon usage du résultat de Pisier, on doit avoir f lipschitzienne dans tout \mathbb{R}^n . Cela est possible par un résultat de McShane et Whitney (cite???) sur les extensions Lipschitz minimisantes. Ils montrent en toute généralité que toute fonction f définie sur un sous ensemble $A \subset \mathbb{R}^n$ qui est L -Lipschitz sur A peut être prolongée en une fonction L -Lipschitz sur \mathbb{R}^n par

$$f(x) = \inf_{y \in A} (f(y) + L\|x - y\|), \quad \forall x \in \mathbb{R}^n - A. \quad (5.22)$$

???preuve de ce truc?

Avec ce résultat, on construit une fonction \tilde{f} , égale à f sur $\mathbb{R}^n - B_r$ et qui est L -Lipschitz sur \mathbb{R}^n . De plus, pour tout $x \in \mathbb{R}^n$, on a que $\tilde{f}(x) \leq M$.

Considérons maintenant $x \in \mathbb{R}^n$ tel que $f(x) > \mathbb{E}_{\gamma_n}[f] + t$. Comme $\tilde{f} \equiv f$ sur $\mathbb{R}^n - B_r$, on a que $\tilde{f}(x) > \mathbb{E}_{\gamma_n}[f] + t = \mathbb{E}_{\gamma_n}[\tilde{f}] + (t - (\mathbb{E}_{\gamma_n}[\tilde{f}] - \mathbb{E}_{\gamma_n}[f]))$. Mais $\mathbb{E}_{\gamma_n}[\tilde{f}] - \mathbb{E}_{\gamma_n}[f] = \mathbb{E}_{\gamma_n}[(\tilde{f} - f)\mathbf{1}_{B_r}] \leq \gamma_n(B_r)2M$. Cela implique que

$$\tilde{f}(x) > \mathbb{E}_{\gamma_n}[\tilde{f}] + (t - 2M\gamma_n(B_r)). \quad (5.23)$$

On en conclut

$$\gamma_n(f(x) - \mathbb{E}_{\gamma_n}[f] > t \text{ and } \|x\| \geq r) \leq \gamma_n(\tilde{f}(x) - \mathbb{E}_{\gamma_n}[\tilde{f}] > (t - 2M\gamma_n(B_r))), \quad (5.24)$$

et, avec le résultat de Pisier 5.4, on obtient (5.21).

On se trouve maintenant en possession de tous les éléments nécessaires pour prouver le théorème de convergence presque-sûre.

Preuve (Preuve du théorème 5.2). Soit $f \in C_b(\mathbb{R})$ une fonction continue bornée. On doit montrer que, presque sûrement,

$$\lim_{N \rightarrow \infty} |\mathbb{E}_{L_n}[f] - E_{\mu_c}[f]| = 0. \quad (5.25)$$

Rappelons-nous que $\bar{L}_n \Rightarrow \mu_c$ (proposition 5.3), donc il suffit de montrer que

$$\lim_{N \rightarrow \infty} |\mathbb{E}_{L_n}[f] - E_{\bar{L}_n}[f]| = 0 \quad (5.26)$$

est vraie presque sûrement pour toute fonction continue bornée f . En fait, on peut se restreindre à un sous espace dense de $C_b(\mathbb{R})$, disons $f \in C_c^1(\mathbb{R})$.

Notons que $\mathbb{E}_{L_n}[f] = \frac{1}{n} \sum_{i=1}^n f(cn\lambda_i(\rho_n))$ et que $E_{\bar{L}_n}[f] = \mathbb{E}_{\mu_{n,k}}[\frac{1}{n} \sum_{i=1}^n f(cn\lambda_i(\rho_n))]$. En utilisant l'écriture de la loi $\mu_{n,k}$ en termes de matrices gaussiennes aléatoires, on a que

$$\mathbb{E}_{L_n}[f] = \frac{1}{n} \sum_{i=1}^n f \left(cn\lambda_i \left(\frac{X_n^* X_n}{\text{Tr}(X_n^* X_n)} \right) \right) = \frac{1}{n} \text{Tr} \left[f \left(cn \frac{X_n^* X_n}{\text{Tr}(X_n^* X_n)} \right) \right], \quad (5.27)$$

et que $E_{\bar{L}_n}[f] = \mathbb{E}_{\gamma_{2nk}}[\mathbb{E}_{L_n}[f]]$.

On introduit maintenant la fonction $f_2 : \mathbb{R} \rightarrow \mathbb{R}$, $f_2(x) = f(x^2)$. Bien évidemment $f_2 \in C_c^1(\mathbb{R})$ et f_2 est Lipschitz avec constante $L = \|f_2'\|_\infty$. Considérons A et B deux matrices de $\mathcal{M}_{k,n}(\mathbb{C})$ et définissons \tilde{A} et $\tilde{B} \in \mathcal{M}_{k+n}(\mathbb{C})$ par

$$\tilde{A} = \begin{pmatrix} 0 & A^* \\ A & 0 \end{pmatrix}, \quad \tilde{B} = \begin{pmatrix} 0 & B^* \\ B & 0 \end{pmatrix}. \quad (5.28)$$

Il s'en suit que

$$\tilde{A}^2 = \begin{pmatrix} A^*A & 0 \\ 0 & AA^* \end{pmatrix}, \quad \tilde{B}^2 = \begin{pmatrix} B^*B & 0 \\ 0 & BB^* \end{pmatrix}, \quad (5.29)$$

et donc

$$f_2(\tilde{A}) = \begin{pmatrix} f(A^*A) & 0 \\ 0 & f(AA^*) \end{pmatrix}, \quad f_2(\tilde{B}) = \begin{pmatrix} f(B^*B) & 0 \\ 0 & f(BB^*) \end{pmatrix}. \quad (5.30)$$

Par le lemme de Klein (see, for example ???), on en deduit que

$$\|f_2(\tilde{A}) - f_2(\tilde{B})\|_{HS} \leq L \|\tilde{A} - \tilde{B}\|_{HS}. \quad (5.31)$$

Après des calculs, on obtient

$$\|f(A^*A) - f(B^*B)\|_{HS} \leq L\sqrt{2} \|A - B\|_{HS}, \quad (5.32)$$

et donc, en remplaçant A par $\sqrt{cn}A/\|A\|_{HS}$ et B par $\sqrt{cn}B/\|B\|_{HS}$, on a

$$\left\| f \left(cn \frac{A^*A}{\text{Tr}(A^*A)} \right) - f \left(cn \frac{B^*B}{\text{Tr}(B^*B)} \right) \right\|_{HS} \leq L\sqrt{2cn} \left\| \frac{A}{\|A\|_{HS}} - \frac{B}{\|B\|_{HS}} \right\|_{HS}. \quad (5.33)$$

En utilisant l'inégalité de Cauchy-Schwarz pour le produit scalaire de Hilbert-Schmidt dans ???, on obtient

$$\left| \frac{1}{n} \text{Tr} \left[f \left(cn \frac{A^*A}{\text{Tr}(A^*A)} \right) \right] - \frac{1}{n} \text{Tr} \left[f \left(cn \frac{B^*B}{\text{Tr}(B^*B)} \right) \right] \right| \leq L\sqrt{2c} \left\| \frac{A}{\|A\|_{HS}} - \frac{B}{\|B\|_{HS}} \right\|_{HS}. \quad (5.34)$$

Notons que, si $\|A\|_{HS}, \|B\|_{HS} \geq r > 0$, alors

$$\left\| \frac{A}{\|A\|_{HS}} - \frac{B}{\|B\|_{HS}} \right\|_{HS} \leq \frac{2}{r} \|A - B\|_{HS}, \quad (5.35)$$

et donc $F(A) = \frac{1}{n} \text{Tr} \left[f \left(cn \frac{A^* A}{\text{Tr}(A^* A)} \right) \right]$ est Lipschitz sur $\mathbb{R}^{2nk} - B_r$ avec constante $\frac{L\sqrt{8c}}{r}$.

Pour $t > 0$, on a

$$\mathbb{P}_{\mu_{n,k}}(|\mathbb{E}_{L_n}[f] - E_{\bar{L}_n}[f]| > t) = \gamma_{2nk}(|F(X_n) - \mathbb{E}_{\gamma_{2nk}}[F(X_n)]| > t) \quad (5.36)$$

$$\leq \gamma_{2nk}(|F(X_n) - \mathbb{E}_{\gamma_{2nk}}[F(X_n)]| > t \text{ and } \|X_n\|_{HS} \geq r) + \gamma_{2nk}(B_r). \quad (5.37)$$

Evaluons maintenant $\gamma_{2nk}(B_r)$:

$$\gamma_{2nk}(B_r) \leq \frac{1}{(2\pi)^{nk}} \text{Leb}_{2nk}(B_r) = \frac{1}{(2\pi)^{nk}} \frac{1}{2nk} \frac{2 \cdot \pi^{nk}}{\Gamma(nk)} r^{2nk} = \frac{1}{(nk)!} \left(\frac{r^2}{2} \right)^{nk}. \quad (5.38)$$

Avec la formule de Stirling, on a

$$\frac{1}{(nk)!} \left(\frac{r^2}{2} \right)^{nk} \sim \frac{1}{\sqrt{2\pi nk}} \left(\frac{er^2}{2nk} \right)^{nk}. \quad (5.39)$$

en choisissant $r = r(n) = \sqrt{n}$, on obtient

$$\sum_n \gamma_{2nk}(B_{r(n)}) < \infty \quad (5.40)$$

et, en particulier, $\lim_{n \rightarrow \infty} \gamma_{2nk}(B_{r(n)}) = 0$. Donc, pour n assez grand, les hypothèses de la proposition ?? sont satisfaites. En l'utilisant pour n tel que $t - 2\|f\|_\infty \gamma_{2nk}(B_r) \geq t/2$, on trouve

$$\gamma_{2nk}(|F(X_n) - \mathbb{E}_{\gamma_{2nk}}[F(X_n)]| > t \text{ and } \|X_n\|_{HS} \geq r) \leq 2 \exp \left(-\frac{2(t/2)^2 \cdot r}{\pi^2 8cL^2} \right). \quad (5.41)$$

Avec le même choix $r = r(n) = \sqrt{n}$ on obtient une série convergente. Enfin, on conclut en utilisant le lemme de Borel-Cantelli pour

$$\sum_n \mathbb{P}_{\mu_{n,k}}(|\mathbb{E}_{L_n}[f] - E_{\bar{L}_n}[f]| > t) < \infty. \quad (5.42)$$

6 Applications complètement positives aléatoires

7 Simulations numériques

On présente ici quelques simulations numériques réalisés dans le but de confirmer

8 Appendices

8.1 L'espace projectif complexe

L'espace projectif complexe de dimension n , noté $\mathbb{C}\mathbb{P}^n$ est l'espace projectif des droites (complexes) de \mathbb{C}^{n+1} . Plus précisément, on peut le définir comme suit :

Définition 8.1 (espace projectif complexe). Soit \sim la relation d'équivalence définie sur $(\mathbb{C}^{n+1})^*$ par $x \sim y \Leftrightarrow \exists \lambda \in \mathbb{C}^* \text{ t.q. } x = \lambda y$. On appelle *espace projectif complexe*, et on note $\mathbb{C}\mathbb{P}^n$, l'espace quotient $(\mathbb{C}^{n+1})^* / \sim$.

$\mathbb{C}\mathbb{P}^n$ est une variété complexe de dimension n (et donc de dimension réelle $2n$) munie des coordonnées

$$(z_1, z_2, \dots, z_{n+1}) \mapsto (z_1/z_k, z_2/z_k, \dots, \widehat{z_k/z_k}, \dots, z_{n+1}/z_k), \text{ si } z_k \neq 0, \quad (8.43)$$

où le chapeau marque l'absence du terme correspondant.

En tant que variétés réelles, on a l'isomorphisme suivant

$$\mathbb{C}\mathbb{P}^n \approx S^{2n+1}/U(1). \quad (8.44)$$

On en déduit que $\mathbb{C}\mathbb{P}^n$ est une variété compacte, comme quotient des variétés compactes.

8.2 La trace partielle

Ici, on se propose d'introduire la notion de *trace partielle* et de présenter quelques unes de ses propriétés.

Définition-proposition 8.2 (trace partielle). Soient \mathcal{H} et \mathcal{K} deux espaces de Hilbert, et X un opérateur de $\mathcal{B}(\mathcal{H} \otimes \mathcal{K})$. On appelle *trace partielle de X sur \mathcal{K}* l'unique opérateur Y de $\mathcal{B}(\mathcal{H})$ qui vérifie l'identité suivante :

$$\text{Tr}(X(Z \otimes I_{\mathcal{K}})) = \text{Tr}(YZ), \quad \forall Z \in \mathcal{B}(\mathcal{H}), \quad (8.45)$$

où $I_{\mathcal{K}}$ est l'opérateur identité sur \mathcal{K} .

Preuve. On doit montrer l'existence et l'unicité de l'opérateur Y défini par l'équation 8.45.

Dans le cadre de l'information quantique (espaces de Hilbert de dimensions finies), on peut donner une définition plus directe de la trace partielle :

$$\text{Tr}_2(|x_1\rangle\langle y_1| \otimes |x_2\rangle\langle y_2|) = \langle y_2|x_2\rangle \cdot |x_1\rangle\langle y_1|. \quad (8.46)$$

Par exemple, si $\{e_i\}_{i=1}^k$ est une base de \mathcal{K} , alors tout opérateur X sur $\mathcal{H} \otimes \mathcal{K}$ peut s'écrire comme

$$X = \sum_{i,j=1}^k A_{ij} \otimes |e_i\rangle\langle e_j|. \quad (8.47)$$

On peut alors expliciter la trace partielle de X sur \mathcal{K} :

$$Y = \text{Tr}_{\mathcal{K}}(X) = \sum_{i,j=1}^k \langle e_j|e_i\rangle A_{ij} = \sum_{i=1}^k A_{ii}. \quad (8.48)$$

8.3 Notation bra-ket de Dirac

La notation bra-ket a été introduite par P. Dirac pour faciliter l'écriture des équations de la mécanique quantique. On se propose ici de présenter les éléments de cet formalisme, ... ????

Commençons avec la notion de base du formalisme, le *ket* (voir aussi la section 1.1). On se place dans le cadre d'un espace de Hilbert \mathcal{H} défini sur le corps des nombres complexes. Deux vecteurs φ et ψ de \mathcal{H} sont dits *équivalents* s'il existe un scalaire $\lambda \in \mathbb{C}$ tel que $y = \lambda x$. On écrit $x \sim y$.

Définition 8.3 (vecteur ket). Soit \mathcal{H} un espace de Hilbert complexe. On appelle vecteur ket un rayon de \mathcal{H} , i.e. une classe d'équivalence de la relation \sim . On note $|\psi\rangle$ le ket associé au vecteur ψ .

Remarque 8.4. Il est coutume en mécanique quantique d'identifier un ket $|\psi\rangle$ (qui est un ensemble de vecteurs) avec un élément ψ de la classe, de norme 1 et dont la phase est *indéterminée*.

On note \mathcal{E}_n l'ensemble de kets d'un espace de Hilbert \mathcal{H} . On peut se demander comment les opérateurs de \mathcal{H} agissent sur \mathcal{E}_n . Considérons pour cela un opérateur linéaire $X : \mathcal{H} \rightarrow \mathcal{H}$ et un ket $|\psi\rangle$. On définit l'action de X sur $|\psi\rangle$ par la relation

$$X|\psi\rangle = |X\psi\rangle. \quad (8.49)$$

Il est facile de se convaincre que le membre droit de l'égalité précédente ne dépend pas du représentant ψ choisi.

Par le théorème de représentation de Riesz, un espace de Hilbert \mathcal{H} est isomorphe à son dual \mathcal{H}^* , i.e. à tout vecteur ψ de \mathcal{H} on peut associer une unique forme linéaire ψ^* comme suit

$$\psi^*(x) = \langle \psi | x \rangle, \quad \forall x \in \mathcal{H},$$

où $\langle \cdot | \cdot \rangle$ est le produit scalaire de \mathcal{H} . On est donc menés à introduire la notion de *bra* :

Définition 8.5 (vecteur bra). On appelle *vecteur bra* l'ensemble des formes linéaires associées à un vecteur ket $|\psi\rangle$. Plus précisément, on note

$$\langle \psi | = \{x^* : x \in |\psi\rangle\}. \quad (8.50)$$

Remarque 8.6. Un vecteur bra $\langle \psi |$ peut être aussi vu comme le rayon de la forme ψ^* dans l'espace dual \mathcal{H}^* .

Bien qu'on ne peut pas évaluer un vecteur bra dans un vecteur ket, on peut définir la quantité suivante, appelée *fidélité*

$$|\langle \varphi | \psi \rangle|^2 = |\varphi_0^*(\psi_0)|^2 = |\langle \varphi_0 | \psi_0 \rangle|^2, \quad (8.51)$$

où φ_0 et ψ_0 sont des représentants *de norme 1* des classes $|\varphi\rangle$ et $|\psi\rangle$.

Une des premières applications de ce formalisme est l'écriture des projecteurs orthogonaux de rang 1. Soit φ un vecteur de \mathcal{H} . L'espace vectoriel engendré par φ est le même que l'espace engendré par tout autre représentant φ' de la classe $|\varphi\rangle$. On peut donc parler du projecteur orthogonal sur l'espace engendré par un *ket* $|\varphi\rangle$, qu'on va noter $|\varphi\rangle\langle\varphi|$. Cette notation s'explique par le calcul suivant :

$$|\psi\rangle\langle\psi|(x) = |\psi_0\rangle\langle\psi_0|x\rangle, \quad (8.52)$$

ce qui est, bien évidemment, la définition du projecteur orthogonal sur l'espace engendré par le vecteur, ψ . Avec cette notation, on écrit

$$|\psi\rangle\langle\psi||\varphi\rangle = |\psi\rangle\langle\psi|\varphi\rangle. \quad (8.53)$$

??? voir cours Attal pour d'autres applications

8.4 Polynômes orthogonaux de Laguerre

Il est connu que les ensembles des polynômes orthogonaux jouent un rôle très important en théorie des matrices aléatoires. Les polynômes qui correspondent aux matrices de Wishart sont les polynômes de Laguerre, définis par ($\alpha = K - n$) :

$$L_k^{(\alpha)} = \frac{x^{-\alpha} e^x}{k!} \frac{d^k}{dx^k} (x^{k+\alpha} e^{-x}).$$

On introduit aussi l'ensemble des fonctions

$$\varphi_k^{(\alpha)} = \sqrt{\frac{k!}{\Gamma(k + \alpha + 1)}} x^\alpha e^{-x} L_k^{(\alpha)}.$$

Les polynômes de Laguerre sont orthogonaux au sens suivant :

$$\int_0^\infty L_j^{(\alpha)}(x)L_k^{(\alpha)}(x)x^\alpha e^{-x}dx = \delta_{j,k} \frac{\Gamma(k+\alpha+1)}{k!},$$

et alors

$$\int_0^\infty \varphi_j^{(\alpha)}(x)\varphi_k^{(\alpha)}(x)dx = \delta_{j,k}.$$

Si on définit la *fonction de corrélation à un point*

$$\Lambda(x) = n \int \Phi_{n,K}^{(W)}(x, \lambda_2, \dots, \lambda_n) d\lambda_2 \cdots d\lambda_n,$$

alors il est connu en théorie des matrices aléatoires (voir [8]) que

$$\Lambda(x) = K_2(x, x),$$

où

$$K_2(x, y) = \sum_{j=0}^{n-1} \varphi_j^{(\alpha)}(x)\varphi_j^{(\alpha)}(y).$$

Ici, on considère un système AB composé d'un système A à n niveaux et d'un système B à k niveaux. On s'intéresse à la loi de probabilité induite par la mesure uniforme sur l'espace \mathcal{E}_n sur \mathcal{D}_n . Plus précisément, on considère l'application

$$\begin{aligned} \tilde{T}_{n,k} : \mathcal{E}_{nk} &\longrightarrow \mathcal{D}_n \\ |\psi\rangle &\longmapsto \text{Tr}_k(|\psi\rangle\langle\psi|).. \end{aligned}$$

On note $\mu_{n,k} = \tilde{T}_{n,k\#}\nu_{nk}$, la loi image de la mesure uniforme sur l'espace des états purs \mathcal{E}_{nk} par l'application $\tilde{T}_{n,k}$.

Soit maintenant l'application

$$\begin{aligned} T_{n,k} : \mathcal{H}_{nk} \setminus \{0\} &\approx \mathcal{M}_{n \times k}(\mathbb{C}) \longrightarrow \mathcal{D}_n \\ W &\longmapsto \frac{W^*W}{\text{Tr } W} \end{aligned}$$

On a le diagramme suivant :

$$\begin{array}{ccc} \mathcal{H}_{nk} \setminus \{0\} & \xrightarrow{\Pi_{n,k}} & \mathcal{E}_{nk} \\ T_{n,k} \downarrow & \swarrow \tilde{T}_{n,k} & \\ \mathcal{D}_n & & \end{array}$$

Il est facile de vérifier que $T_{n,k} = \tilde{T}_{n,k} \circ \Pi_{n,k}$, et donc on peut construire les matrices densités à partir de \mathcal{H}_{nk} sans devoir passer par \mathcal{E}_{nk} .

Bibliographie

- [1] T. M. Cover and J. A. Thomas. *Elements of Information Theory*. John Wiley and Sons, 1991.
- [2] Robert M. Gray. *Entropy and Information Theory*. Springer, 1990.
- [3] Josef Gruska. *Quantum Computing*. McGraw-Hill, 1999.
- [4] U. Haagerup and S. Thorbjornsen. Random matrices with complex gaussian entries.
- [5] Mika Hirvensalo. *Quantum Computing*. Springer, 2004.
- [6] E. H. Lieb and M. B. Ruskai. Proof of the strong subadditivity of quantum mechanical entropy. *J. Math. Phys*, (14) :1938–1941, 1973.
- [7] L. C. Malacarne, R. S. Mendes, and E. K. Lenzi. Average entropy of a subsystem from its average tsallis entropy. *Phys. Rev. E (3)*, 65(4) :046131, 5, 2002.
- [8] M.L. Mehta. *Mandom Matrices*. Elsevier Academic Press, 2004.
- [9] Michael A. Nielsen and Isaac L. Chuang. *Quantum Computation and Quantum Information*. Cambridge University Press, 2000.
- [10] D.N. Page. Average entropy of a subsystem. *Phys. Rev. Lett*, (71) :1291, 1993.
- [11] K.R. Parthasarathy. *An introduction to quantum stochastic calculus*. BirkhÅuser, 1992.
- [12] John Preskill. Quantum information and computation. Technical report, California Institute of Technology, September, 1998.
- [13] J. SanchÃ©z-Ruiz. Simple proof of page’s conjecture on the average entropy of a subsystem. *Phys. Rev. E*, 52(5) :5653, 1995.
- [14] S. Sen. Average entropy of a quantum subsystem. *Phys. Rev. Lett.*, 77(1) :1, 1996.