

UNIVERSITÉ MARIE & LOUIS PASTEUR

Laboratoire de Mathématiques de Besançon
Master Mathématiques approfondies

Notes d'une lecture détaillée sur la complétion de matrices.

Dans le cadre d'un cours de:

Camelia Goga
Mehdi Dagdoug

Soumis par:

Hadrien Maffioli

Année académique: 2024/2025

Table des matières

1	Introduction	2
1.1	Contexte et motivation	2
1.2	But général	3
2	Rappels et résultats préliminaires	5
2.1	Quelques rappels d'algèbre linéaire	5
2.2	Inégalités de déviation	10
3	Le théorème principal	13
3.1	Inégalités à forte probabilité	13
3.2	Démonstration du théorème principal	20

1 Introduction

1.1 Contexte et motivation

Ce document vise à reprendre en détail un article de B. Recht intitulé: *A Simpler Approach to Matrix Completion* paru en 2011 dans *Journal of Machine Learning Research* [Rec09]. Le but principal sera de détailler les preuves fournies dans cet article tant du théorème principal que des résultats intermédiaires menant à ce théorème.

Commençons par décrire le problème auquel nous nous intéressons ainsi que quelques unes de ses nombreuses applications. Ce problème est connu sous le nom de **complétion de matrices**. Imaginons un tableau de taille $n_1 \times n_2$ dont on cherche à connaître l'ensemble des $n_1 n_2$ entrées. Le problème est que seulement une petite partie des entrées est observée. On se demande alors s'il est possible de retrouver toute la matrice à partir de nos seules observations.

Il est clair que sans hypothèses supplémentaires, il est impossible de retrouver en toute généralité une matrice même si l'on connaît toutes les entrées sauf une. Une hypothèse courante est que la matrice recherchée a un "faible" rang. On cherchera donc à trouver la matrice de plus faible rang dont les entrées coïncident avec celles observées.

Cette question, en lien avec la théorie associée aux données manquantes, apparaît dans de nombreux problèmes du "quotidien". Nous en donnerons trois ici, dans le but de motiver notre intérêt.

- *Netflix Prize*: Le 2 octobre 2006, Netflix lança un challenge. L'entreprise américaine offrait 1 million de dollars à l'équipe qui parviendrait à créer le meilleur algorithme pouvant prédire la note qu'un certain utilisateur attribuerait à un certain film en se basant **uniquement** sur les notes que cet utilisateur avait pu attribuer à d'autres films auparavant. Cette compétition ouverte à tous se termina le 21 Septembre 2009 lorsque le prix fut attribué à l'équipe *BellKor's Pragmatic Chaos* qui améliora l'algorithme utilisé par Netflix, le rendant capable de prédire les notes avec une marge d'erreur d'environ 10%. Ce n'était en revanche que le début des questions sur la complétion de matrices.
- *Filtrage collaboratif*: L'exemple précédent n'est en réalité qu'un cas particulier d'un domaine plus général: celui du filtrage collaboratif. Le filtrage collaboratif regroupe l'ensemble des méthodes qui visent à construire des systèmes de recommandation utilisant les opinions et évaluations du groupe pour aider l'individu. Il existe trois principaux axes de recherche dans ce domaine, dépendant chacun des données recueillies sur les utilisateurs du système :
 - ◊ Le filtrage collaboratif actif, celui utilisé par Netflix notamment, où l'on demande explicitement à l'utilisateur de renseigner ses opinions.
 - ◊ Le filtrage collaboratif passif où l'on utilise les données récoltées sur les différents utilisateurs pour déduire leurs préférences.
 - ◊ Le filtrage basé sur le contenu dont la décision de sélection ou non d'un document se base uniquement sur le contenu de celui-ci.

Dans chacun de ces cas, l'idée est de prédire la note qu'un utilisateur pourrait attribuer à un contenu en fonction des notes qu'il a déjà attribué à d'autres contenus précédents. En regroupant dans un tableau les notes attribuées par tous les utilisateurs à tous les films, il s'agit donc de compléter ce tableau dans le but de recommander le mieux possible les futurs contenus. Ce problème est par exemple abordé dans [KMO09].

- *Sondage*: Imaginons à présent la situation suivante. Lors d'un sondage auprès d'une partie de la population, les personnes interrogées ne répondent pas à l'ensemble des questions posées. On regroupe alors leurs réponses et on cherche à prédire quelles auraient pu être les réponses

manquantes aux différentes questions en fonction des réponses fournies. Les mécanismes de réponse peuvent bien sûr influencer les estimations réalisées mais nous ne rentrerons pas dans de telles considérations dans ce manuscrit. Ce problème est par exemple abordé dans [CR08] ou [CT09].

Comme nous pouvons le voir dans les exemples ci-dessus, une des principales préoccupations est la recherche d'algorithmes toujours plus performant dans la complétion de matrices. Néanmoins le besoin de résultats théoriques est également au centre d'articles paru relativement récemment. *A Simpler Approach to Matrix Completion* [Rec09] s'inscrit dans cette lignée de recherche de résultats d'existence et d'unicité de solution(s) au problème de complétion.

1.2 But général

Le principal but de l'article [Rec09] repris dans ce document est d'améliorer de précédents résultats. Ce que l'on entend par *améliorer* est un affaiblissement des hypothèses et/ou un renforcement des conclusions. Une plus grande simplicité des techniques utilisées pour démontrer de tels résultats est aussi considéré comme une amélioration.

Par exemple il a été montré dans [CR08] que sous certaines conditions, la minimisation de la norme nucléaire permettait de retrouver la plupart des matrices de "faible" rang. Comme dans de nombreux problèmes d'optimisation, la convexité de la fonction que l'on cherche à minimiser est cruciale. Malheureusement, le rang n'est pas une fonction convexe, c'est pour cette raison que nous le remplaçons par la norme nucléaire dans notre problème de minimisation. (Notée $\|\cdot\|_*$ dans la suite, nous la définirons formellement dans la section suivante.) Ceci est analogue au fait "d'approximer" le nombre d'entrées strictement positives dans un vecteur par la norme 1 de celui-ci.

Notons que l'hypothèse de faible rang mentionnée en section 1.1 est en pratique expliquée par le fait de formation de "groupes" dans la population. En effet, il paraît plutôt pertinent de supposer que tous les utilisateurs ne sont pas complètement différents, mais qu'au contraire il existe des "profils". Pour reprendre l'exemple *Netflix prize*, on peut supposer que l'on peut regrouper les films par catégorie et qu'une personne aimant un film d'action aimera tous les films d'actions. On peut ainsi créer différents profils d'utilisateurs (par exemple ceux aimant les comédies, ceux aimant les films romantiques, etc...)

Néanmoins cette seule hypothèse ne saurait être suffisante. En effet il a été montré (voir par exemple [CR08]) qu'il était impossible de retrouver une matrice dont presque toutes les entrées étaient nulles (sauf si toutes les entrées sont observées bien évidemment). Formellement, cela mène à une condition de majoration de la cohérence de la matrice recherchée, où la cohérence est une quantité définie de la façon suivante.

Définition 1.1 (Cohérence). Soit U un sous-espace vectoriel de $\mathbb{R}^n$ de dimension r et $\mathbf{P}_U$ la projection orthogonale sur U . On appelle *cohérence* de U (relativement à la base canonique $(e_i)_{1 \leq i \leq n}$) la quantité:

$$\mu(U) := \frac{n}{r} \max_{1 \leq i \leq n} \|\mathbf{P}_U e_i\|_2^2,$$

où $\|\cdot\|_2$ désigne la norme spectrale, dont nous rappellerons la définition dans la section suivante. $\diamond$

Remarque 1.2. Remarquons que l'on dispose d'un encadrement pour $\mu(U)$.

- $\mu(U)$ est en effet minorée par 1. Pour le montrer, remarquons que ($\mathbf{P}_U$ étant une projection):

$$\forall i \in \llbracket 1; n \rrbracket, \quad \|\mathbf{P}_U e_i\|_2^2 = e_i^* \mathbf{P}_U e_i$$

et puisque la trace de $\mathbf{P}_U$ vaut r (les valeurs propres sont 1 avec multiplicité r et 0 avec multiplicité $n - r$) on a :

$$\sum_{i=1}^n e_i^* \mathbf{P}_U e_i = r$$

donc le maximum de ces termes est supérieur à $\frac{r}{n}$ donc la cohérence est supérieure à 1.

- De plus $\mu(U)$ est majorée par $\frac{n}{r}$. En effet $\mathbf{P}_U$ est de norme 1 et les éléments de la base canonique sont de norme 1. Ainsi $\max_{1 \leq i \leq n} \|\mathbf{P}_U e_i\|_2^2 \leq 1$. Notons que ce maximum est atteint lorsque U contient un élément de la base canonique.

◇

Nous sommes à présent prêt à présenter le théorème principal.

Théorème 1.3 (Théorème de complétion). Soit M une matrice de taille $n_1 \times n_2$ et de rang r . Notons $U\Sigma V^*$ sa décomposition en valeurs singulières. Quitte à considérer M^* on peut supposer que $n_1 \leq n_2$. Notons de plus que dans notre décomposition on suppose que Σ est de taille $r \times r$, U de taille $n_1 \times r$ et V de taille $n_2 \times r$. Supposons que :

A0 Les sous espaces vectoriels engendrés par les lignes et les colonnes de M ont une cohérence majorée par une certaine constante $\mu_0 \geq 0$.

A1 La matrice UV^* vérifie $\|UV^*\|_\infty \leq \mu_1 \sqrt{\frac{r}{n_1 n_2}}$ pour une certaine constante $\mu_1 \geq 0$,

où $\|\cdot\|_\infty$ désigne la norme infinie dont nous rappellerons la définition en section suivante. Supposons qu'un nombre m d'entrées sont observées uniformément au hasard et notons Ω l'ensemble des indices des entrées observées. Alors si

$$m \geq 32 \max\{\mu_1^2, \mu_0\} r (n_1 + n_2) \beta \log^2(2n_2) \quad (1)$$

pour un réel $\beta > 1$, alors la matrice minimisant le problème suivant

$$\begin{aligned} & \text{minimiser} && \|X\|_* \\ & \text{sous la contrainte} && X_{i,j} = M_{i,j} \quad \forall (i,j) \in \Omega. \end{aligned} \quad (2)$$

est unique et est égal à M avec une probabilité supérieure à $1 - 6 \log(n_2)(n_1 + n_2)^{2-2\beta} - n_2^{2-2\beta^{1/2}}$. ◇

Commençons par commenter ce théorème. De façon informelle, ce résultat nous dit que sous certaines conditions il est possible de retrouver **toutes** les entrées d'une matrice avec une forte probabilité en observant seulement un "faible" nombre d'entrées. Ce que l'on entend par faible est un nombre négligeable devant le nombre total d'entrées (ici m est d'ordre une constante multipliée par $n_2 \log^2(2n_2)$ ce qui est considérablement plus faible que n_2^2 lorsque n_2 est grand). Par exemple pour $n_2 = 10^8$ ce nombre est d'ordre 10^{12} contre 10^{16} entrées au total.

Le théorème dit que pour retrouver la matrice recherchée, il suffit de minimiser la norme nucléaire de l'ensemble des matrices compatibles avec les entrées observées. Dû au caractère convexe de la norme nucléaire, il existe de nombreux algorithmes permettant d'approcher ce minimum avec une grande précision.

Bien entendu, si certaines colonnes/lignes ne sont jamais observées alors il sera impossible de retrouver la matrice en toute généralité. Nous observons les entrées uniformément au hasard pour éviter ce genre de problème et de plus, il est possible de montrer que le nombre d'entrées observées différentes

est "grand" dans la majorité des cas. La première condition **A0** permet en particulier d'éviter les cas pathologique comme par exemple lorsque la matrice recherchée ne contient presque que des zéros.

L'hypothèse **A1** est ajoutée pour des raisons techniques lors de la preuve dans le but de pouvoir minorer la norme nucléaire de matrices vérifiant la condition (2).

Notons de plus que les constantes μ_0 et μ_1 peuvent dépendre de r , n_1 et/ou n_2 . Les hypothèses **A0** et **A1** avaient été introduites auparavant par Candès et Recht [CR08] et ne sont donc pas surprenantes. En revanche, dans tous les précédents articles, d'autres hypothèses plus fortes étaient ajoutées (pour obtenir de plus faibles résultats). Par exemple dans [CT09] l'hypothèse **A0** était alourdie car une condition de "forte incohérence" était également requise ou bien dans [KMO09] on supposait que la matrice recherchée avait un rang inférieur à $\log(n_2)$.

De plus, comme nous l'avons signalé, ce résultat améliore les précédents dans le sens que le nombre d'entrées requis pour retrouver la matrice M est presque optimal. En effet, il a été prouvé dans [CT09] que le nombre minimal d'entrées nécessaires pour retrouver M était de l'ordre de $n_2 \log(n_2)$ ce qui correspond à notre résultat à un facteur $\log(n_2)$ et une constante près.

Ce résultat peut être démontré utilisant des outils d'algèbre linéaire ainsi que quelques inégalités dites de *Bernstein*. Dans la suite de ce document, nous tacherons de présenter ces outils et de fournir des démonstrations complètes de chacun des résultats préliminaires nécessaires à l'obtention de ce théorème ainsi que la preuve du théorème à proprement parler.

2 Rappels et résultats préliminaires

2.1 Quelques rappels d'algèbre linéaire

Dans cette partie, nous définirons tous les objets nécessaires à la preuve du Théorème 1.3 et nous démontrerons quelques résultats classiques d'algèbre linéaire qui nous seront utiles.

Notation: Dans toute la suite du document, les matrices seront dénotées par des lettres majuscules et les vecteurs par des lettres minuscules. Étant donné une matrice A nous noterons sa k -ième colonne par a_k . La transposée d'une matrice A (respectivement d'un vecteur x) sera notée A^* (respectivement x^*). ◇

Définition 2.1 (Norme spectrale). Soit A une matrice. On notera $\|A\|_2$ sa norme spectrale c'est à dire la norme subordonnée de la norme euclidienne sur $\mathbb{R}^n$ (on notera de même $\|x\|_2$ pour la norme euclidienne d'un vecteur x). On peut vérifier que cette norme coïncide avec la racine carrée de la plus grande valeur propre de la matrice A^*A (à noter que cette valeur propre est bien positive puisque la matrice A^*A est semi-définie positive). ◇

Proposition 2.2 (Caractérisation de la norme spectrale). On peut de plus montrer que la norme spectrale d'une matrice A coïncide avec sa plus grande valeur singulière. ◇

PREUVE: Notons $A = U\Sigma V^*$ la décomposition en valeurs singulières de la matrice A où Σ est la matrice diagonale formée des $(\sigma_1, \dots, \sigma_n)$, les valeurs singulières de A ordonnées dans l'ordre décroissant. On a alors :

$$\sup_{\|x\|_2=1} \|Ax\|_2 = \sup_{\|x\|_2=1} \|U\Sigma V^*x\|_2 = \sup_{\|x\|_2=1} \|\Sigma V^*x\|_2$$

puisque U est une matrice unitaire (donc de norme spectrale 1). En posant $y = V^*x$ on a de même que précédemment:

$$\sup_{\|x\|_2=1} \|\Sigma V^*x\|_2 = \sup_{\|y\|_2=1} \|\Sigma y\|_2.$$

Puisque Σ est diagonale on peut résoudre ce maximum (à l'aide des multiplicateurs de Lagrange par exemple) et on trouve qu'il est égal à σ_1 (atteint en $y = (1, 0, \dots, 0)^*$), ce qu'il fallait démontrer. ■

Définition 2.3 (Norme infinie). Pour une matrice A on note $\|A\|_\infty$ la norme infinie de A définie par la plus grande entrée (en module) de A . C'est à dire $\|A\|_\infty = \max_{i,j} |A_{i,j}|$. ◇

Définition 2.4 (Produit scalaire et norme de Frobenius). On rappelle que le produit scalaire naturel entre deux matrices A et B est donné par $\langle A, B \rangle = \text{Tr}(A^*B)$ et que la norme qui en découle est appelée norme de Frobenius, notée $\|\cdot\|_F$ (i.e $\|A\|_F = \langle A, A \rangle^{1/2}$). ◇

Définition 2.5 (Norme nucléaire). Pour une matrice A on note $\|A\|_*$ la norme nucléaire c'est à dire la somme des valeurs singulières. ◇

Proposition 2.6 (Caractérisation de la norme nucléaire). On peut montrer que la norme nucléaire est la norme duale de la norme spectrale. C'est à dire que pour une matrice A donnée on a :

$$\|A\|_* = \max_{\|B\|_2 \leq 1} \langle A, B \rangle.$$

◇

PREUVE: Posons $A = U\Sigma V^*$ la décomposition en valeurs singulières de la matrice A et posons $C = UV^*$ de sorte que $\|C\|_2 \leq 1$. En utilisant le fait que U et V soient des matrices unitaires (i.e $UU^* = V^*V = I$) et que $\text{Tr}(AB) = \text{Tr}(BA)$, on a donc :

$$\begin{aligned} \max_{\|B\|_2 \leq 1} \langle A, B \rangle &\geq \langle A, C \rangle \\ &= \text{Tr}(C^*A) \\ &= \text{Tr}(VU^*U\Sigma V^*) \\ &= \text{Tr}(V\Sigma V^*) \\ &= \text{Tr}(V^*V\Sigma) \\ &= \text{Tr}(\Sigma) \\ &= \sum_i \sigma_i(A) = \|A\|_* \end{aligned}$$

On a donc :

$$\max_{\|B\|_2 \leq 1} \langle A, B \rangle \geq \|A\|_* \quad (3)$$

Montrons à présent l'inégalité réciproque. Soit B une matrice quelconque telle que $\|B\|_2 \leq 1$.

$$\begin{aligned} \langle A, B \rangle &= \text{Tr}(B^*A) \\ &= \text{Tr}(B^*U\Sigma V^*) \\ &= \text{Tr}(V^*B^*U\Sigma) \\ &= \text{Tr}((U^*BV)^*\Sigma) \\ &= \langle U^*BV, \Sigma \rangle \\ &= \sum_{i,j} [U^*BV]_{i,j} \Sigma_{i,j} \\ &= \sum_i [U^*BV]_{i,i} \sigma_i \\ &= \sum_i u_i^* B v_i \sigma_i \end{aligned}$$

Or les $(u_i)_{1 \leq i \leq n}$ et $(v_i)_{1 \leq i \leq n}$ forment des bases orthonormées donc:

$$u_i^* B v_i \leq \max_{\|u\|_2 = \|v\|_2 = 1} u^* B v = \sigma_1(B).$$

(u et v sont de normes 1 donc le maximum est atteint pour les vecteurs singuliers associés à la plus grande valeur singulière: $\sigma_1(B)$.)

En repartant de la dernière égalité (puisque $\|B\|_2 = \sigma_1(B) \leq 1$), on a donc:

$$\begin{aligned} \max_{\|B\|_2 \leq 1} \langle A, B \rangle &= \max_{\|B\|_2 \leq 1} \sum_i u_i^* B v_i \sigma_i \\ &\leq \max_{\|B\|_2 \leq 1} \sum_i \sigma_1(B) \sigma_i \\ &\leq \sum_i \sigma_i \\ &= \|A\|_* \end{aligned}$$

Ainsi on a montré que $\max_{\|B\|_2 \leq 1} \langle A, B \rangle \leq \|A\|_*$ donc avec l'équation (3) on a bien que :

$$\max_{\|B\|_2 \leq 1} \langle A, B \rangle = \|A\|_*$$

■

Notation:

- On notera s.d.p pour désigner une matrice semi-définie positive, c'est à dire une matrice symétrique dont toutes les valeurs propres sont positives ou nulles.
- Les opérateurs linéaires sur les matrices seront notés avec des lettres caligraphiques. En particulier l'identité sera notée $\mathcal{I}$. La norme spectrale d'un tel opérateur sera notée $\|\mathcal{A}\|$ et est définie par $\|\mathcal{A}\| := \sup_{\|A\|_F \leq 1} \|\mathcal{A}(A)\|_F$.
- On rappelle de plus que $(e_i)_{1 \leq i \leq n}$ désigne la base canonique de $\mathbb{R}^n$.

◇

On rappelle sans démonstration le théorème classique suivant.

Théorème 2.7 (Projection orthogonale). Soit E un espace euclidien et $F \subset E$ un sous espace vectoriel. Pour tout $x \in E$, il existe un unique $y \in F$ tel que $\|x - y\| = \text{dist}(x, F) = \inf_{z \in F} \|x - z\|$. y est l'unique vecteur de F tel que $x - y \in F^\perp$. De plus, pour $(e_1, \dots, e_n)$ une base orthonormée de F , $y = \sum_{k=1}^n \langle x, e_k \rangle e_k$ et $\|x - y\|^2 = \|x\|^2 - \|y\|^2$. On a ainsi que l'application $\mathbf{P}_F$ qui à tout $x \in E$ associe cet unique $y \in F$ est bien définie. De plus cette application appelée *projection orthogonale sur F* est linéaire, de norme d'opérateur 1 et vérifie $\mathbf{P}_F \circ \mathbf{P}_F = \mathbf{P}_F$.

◇

Proposition 2.8. On dispose de la caractérisation suivante:

$$\forall x \in E, y = \mathbf{P}_F(x) \iff y \in F \text{ et } x - y \in F^\perp \iff y \in F \text{ et } \|x - y\| = \text{dist}(x, F).$$

De plus, $\mathbf{P}_F$ est auto-adjoint.

◇

PREUVE: La caractérisation est classique. Montrons que $\mathbf{P}_F$ est auto-adjoint. Soient x_1 et x_2 deux éléments de E et notons y_1 et y_2 leur projections (orthogonales) sur F respectives. On a alors par bilinéarité du produit scalaire et en utilisant la caractérisation:

$$\begin{aligned}\langle y_1, x_2 \rangle &= \langle y_1, x_2 - y_2 + y_2 \rangle \\ &= \langle y_1, x_2 - y_2 \rangle + \langle y_1, y_2 \rangle \\ &= \langle y_1, y_2 \rangle \\ &= \langle y_1 - x_1 + x_1, y_2 \rangle \\ &= \langle y_1 - x_1, y_2 \rangle + \langle x_1, y_2 \rangle \\ &= \langle x_1, y_2 \rangle.\end{aligned}$$

■

Remarque 2.9. Remarquons qu'on a même montré que $\langle y_1, x_2 \rangle = \langle y_1, y_2 \rangle = \langle x_1, y_2 \rangle$.

◇

Fixons (dans toute la suite du document) M une matrice vérifiant toutes les hypothèses du théorème 1.3. On notera $M = U\Sigma V^*$ sa décomposition en valeurs singulières. On pose F le sous espace vectoriel engendré par $(u_1, \dots, u_r)$ les colonnes de U et G le sous espace vectoriel engendré par $(v_1, \dots, v_r)$ les colonnes de V . On note alors T le sous espace vectoriel engendré par les éléments de la forme $u_k y^*$ et xv_k^* , où x et y sont quelconques. On a alors le résultat suivant.

Proposition 2.10. La projection orthogonale sur T est donnée par:

$$\forall Z, \mathcal{P}_T(Z) = \mathbf{P}_F Z + Z \mathbf{P}_G - \mathbf{P}_F Z \mathbf{P}_G.$$

◇

PREUVE: On utilise la caractérisation d'une projection orthogonale donnée plus haut (voir proposition 2.8). Les colonnes de $\mathbf{P}_F Z$ sont dans F par définition de $\mathbf{P}_F$ donc $\mathbf{P}_F Z \in \text{Vect}(u_k y^*)$. De même $Z \mathbf{P}_G \in \text{Vect}(xv_k^*)$ donc $\mathbf{P}_F Z \mathbf{P}_G \in \text{Vect}(u_k y^*, xv_k^*)$ donc $\mathcal{P}_T(Z) \in T$. De plus,

$$\begin{aligned}Z - \mathcal{P}_T(Z) &= Z - \mathbf{P}_F Z - Z \mathbf{P}_G + \mathbf{P}_F Z \mathbf{P}_G \\ &= (I_{n_1} - \mathbf{P}_F)Z(I_{n_2} - \mathbf{P}_G) \\ &= \mathbf{P}_{F^\perp} Z \mathbf{P}_{G^\perp} \in T^\perp\end{aligned}$$

ce que nous voulions démontrer.

■

Montrons à présent l'inégalité suivante.

Proposition 2.11. On a :

$$\|\mathcal{P}_T(e_a e_b^*)\|_F^2 \leq \mu_0 r \frac{n_1 + n_2}{n_1 n_2} \quad (4)$$

◇

PREUVE: Notons tout d'abord que les matrices des projections orthogonales correspondent à des matrices symétriques (car on a vu que ce sont des endomorphismes auto-adjoint) donc:

$$\mathcal{P}_T(e_a e_b^*) = \mathbf{P}_F e_a e_b^* + e_a e_b^* \mathbf{P}_G - \mathbf{P}_F e_a e_b^* \mathbf{P}_G = (\mathbf{P}_F e_a) e_b^* + e_a (\mathbf{P}_G e_b)^* - (\mathbf{P}_F e_a)(\mathbf{P}_G e_b)^*$$

Ainsi (voir Remarque 2.9) on a:

$$\begin{aligned}\|\mathcal{P}_T(e_a e_b^*)\|_F^2 &= \langle \mathcal{P}_T(e_a e_b^*), \mathcal{P}_T(e_a e_b^*) \rangle \\ &= \langle \mathcal{P}_T(e_a e_b^*), e_a e_b^* \rangle \\ &= \langle \mathbf{P}_F e_a e_b^* + e_a (\mathbf{P}_G e_b)^* - (\mathbf{P}_F e_a)(\mathbf{P}_G e_b)^*, e_a e_b^* \rangle \\ &= \langle \mathbf{P}_F e_a e_b^*, e_a e_b^* \rangle + \langle e_a (\mathbf{P}_G e_b)^*, e_a e_b^* \rangle - \langle (\mathbf{P}_F e_a)(\mathbf{P}_G e_b)^*, e_a e_b^* \rangle\end{aligned}$$

A présent remarquons que la proposition 2.8 donne de plus que:

$$\begin{aligned}
\langle (\mathbf{P}_F e_a) e_b^*, e_a e_b^* \rangle &= \|\mathbf{P}_F e_a e_b^*\|_F^2 \\
&= \text{Tr}(e_b (\mathbf{P}_F e_a)^* \mathbf{P}_F e_a e_b^*) \\
&= \text{Tr}((\mathbf{P}_F e_a)^* \mathbf{P}_F e_a) \\
&= \|\mathbf{P}_F e_a\|_2^2
\end{aligned}$$

En effectuant le même calcul pour les deux autres termes on obtient:

$$\|\mathcal{P}_T(e_a e_b^*)\|_F^2 = \|\mathbf{P}_F e_a\|_2^2 + \|\mathbf{P}_G e_b\|_2^2 - \|\mathbf{P}_F e_a\|_2^2 \|\mathbf{P}_G e_b\|_2^2.$$

De plus par définition de la cohérence (on prend le max sur l'ensemble des vecteurs de la base canonique) on a :

$$\|\mathbf{P}_F e_a\|^2 \leq \mu(F) \frac{r}{n_1}$$

Ainsi on a bien que :

$$\|\mathcal{P}_T(e_a e_b^*)\|_F^2 \leq \max\{\mu(F), \mu(G)\} r \frac{n_1 + n_2}{n_1 n_2} \leq \mu_0 r \frac{n_1 + n_2}{n_1 n_2}$$

ce qu'il fallait démontrer. ■

Pour clôre cette partie, nous montrerons quelques résultats utiles sur les matrices semi-définies positives.

Définition 2.12 (Relation d'ordre). Pour deux matrices carrées A et B , on note $A \preceq B$ pour signifier que $B - A$ est une matrice semi-définie positive. Ceci définit une relation d'ordre partiel sur l'ensemble des matrices carrées. On rappelle que l'ensemble des matrices semi-définies positives forme un cône convexe. ◇

Proposition 2.13. Soient A et B deux matrices symétriques de taille $d \times d$ et soit C une matrice quelconque (de taille $d \times d$). Si $A \succeq B$ alors $CAC^* \succeq CBC^*$. De plus si C est inversible c'est même une équivalence. ◇

PREUVE: $A \succeq B$ donc par le théorème de factorisation de Cholesky il existe une matrice V telle que $A - B = VV^*$. Ainsi,

$$C(A - B)C^* = CVV^*C^* = CV(CV)^*$$

donc par le théorème de Cholesky $C(A - B)C^*$ est semi-définie positive. Si de plus C est inversible il suffit d'appliquer ce qui est ci-dessus avec C^{-1} pour obtenir l'implication réciproque. ■

Proposition 2.14. Soient A et B deux matrices semi-définies positives. On a :

$$\|A - B\|_2 \leq \max\{\|A\|_2, \|B\|_2\}$$

◇

PREUVE: Nous allons montrer les relations suivantes:

$$-\|B\|_2 I \preceq -B \preceq A - B \preceq A \preceq \|A\|_2 I.$$

B et I commutent donc sont co-diagonalisables donc $\text{Sp}(\|B\|_2 I - B) = \rho(B) - \text{Sp}(B) \subset \mathbb{R}^+$ donc $\|B\|_2 I - B$ est s.d.p par définition.

$B + A - B = A$ est s.d.p par hypothèse. De même que précédemment $\|A\|_2 I - A$ est s.d.p. Ainsi, par transitivité de la relation d'ordre $A - B + \|B\|_2 I$ et $\|A\|_2 I - A + B$ sont s.d.p. Par commutativité, cela implique que les valeurs propres de $A - B$ sont incluses dans le compact $[-\|B\|_2, \|A\|_2]$ donc $\|A - B\|_2 \leq \max\{\|A\|_2, \|B\|_2\}$ (on a utilisé le fait que pour une matrice symétrique la norme spectrale est égale au rayon spectral). ■

Proposition 2.15. Soient A , B et C des matrices symétriques de taille $d \times d$ vérifiant $A \succeq 0$ et $B \preceq C$. On a :

$$\text{Tr}(AB) \leq \text{Tr}(AC)$$

◇

PREUVE: Traitons d'abord le cas où A est de rang 1 c'est à dire que l'on peut supposer qu'il existe un vecteur v tel que $A = vv^*$. En utilisant la caractérisation: [M est définie positive si et seulement si pour tout vecteur x , $x^*Mx \geq 0$], on a :

$$\begin{aligned} \text{Tr}(AB) &= \text{Tr}(vv^*B) = \text{Tr}(v^*Bv) = v^*Bv \\ &\leq v^*Cv = \text{Tr}(v^*Cv) = \text{Tr}(vv^*C) = \text{Tr}(AC) \end{aligned}$$

Dans le cas général on peut décomposer A sous la forme $A = \sum_{i=1}^d \lambda_i u_i u_i^*$. Puisque $A \succeq 0$, les λ_i sont positifs donc on peut poser $v_i = \sqrt{\lambda_i} u_i$ de telle sorte que $A = \sum_{i=1}^d v_i v_i^*$. Par linéarité de la trace et en utilisant le cas précédent on a bien que :

$$\text{Tr}(AB) = \sum_i \text{Tr}(v_i v_i^* B) \leq \sum_i \text{Tr}(v_i v_i^* C) = \text{Tr}(AC)$$

■

Corollaire 2.16. En appliquant la proposition précédente à $C = \|B\|_2 I$ on obtient le corollaire suivant:

$$\text{Si } A \succeq 0 \text{ alors } \text{Tr}(AB) \leq \|B\|_2 \text{Tr}(A).$$

◇

2.2 Inégalités de déviation

Dans cette partie nous introduirons des résultats de majoration sur les probabilités de déviation par rapport à la moyenne qui seront utilisés au cours de ce document.

On rappelle sans démonstration le résultat classique suivant:

Proposition 2.17 (Inégalité de Bernstein). Soient $X_1, \dots, X_n$ des variables aléatoires centrées et indépendantes. On suppose qu'il existe une constante M telle que $|X_i| \leq M$ presque sûrement pour tout i . Pour tout t positif, on a :

$$\mathbb{P}\left(\sum_{i=1}^n X_i \geq t\right) \leq \exp\left(-\frac{\frac{1}{2}t^2}{\sum_{i=1}^n \mathbb{E}[X_i^2] + \frac{1}{3}Mt}\right)$$

◇

La proposition suivante, dont la démonstration est due à Torben Hagerup et Christine Rüb (voir [HR90]), permet de raffiner ce résultat lorsque les variables X_i suivent une loi de Bernoulli.

Proposition 2.18 (Borne de Chernoff). Soient $X_1, \dots, X_n$ des variables aléatoires indépendantes et identiquement distribuées selon une loi de Bernoulli de paramètre $p \in]0; 1[$. On note $S = \sum_{i=1}^n X_i$.

Pour tout $0 < k < n$ on a :

$$\mathbb{P}(S \geq k) \leq \left(\frac{np}{k}\right)^k e^{k-np}.$$

◇

PREUVE: Soit $\epsilon \geq 0$, $t \geq 0$. Posons $m = np = \mathbb{E}[S]$. On a la suite d'inégalités suivante.

$$\begin{aligned}
\mathbb{P}(S \geq (1 + \epsilon)m) &\leq \mathbb{P}(e^{tS} \geq e^{t(1+\epsilon)m}) && (\text{par croissance de la fonction } x \mapsto e^{tx}) \\
&\leq \mathbb{E}[e^{tS}] e^{-t(1+\epsilon)m} && (\text{inégalité de Markov}) \\
&= e^{-t(1+\epsilon)m} \prod_{i=1}^n \mathbb{E}[e^{tX_i}] && (\text{par indépendance}) \\
&= e^{-t(1+\epsilon)m} \prod_{i=1}^n (pe^t + (1-p)) \\
&= e^{-t(1+\epsilon)m} (pe^t + (1-p))^n.
\end{aligned}$$

Soit $a \in [p; 1[$. En choisissant ϵ tel que $(1 + \epsilon)m = an$ la dernière inégalité devient:

$$\mathbb{P}(S \geq an) \leq e^{-tan} (pe^t + (1-p))^n.$$

Pour $t = \frac{\ln(a(1-p))}{p(1-a)}$ on a par inégalité de convexité:

$$\begin{aligned}
\mathbb{P}(S \geq an) &\leq \left(\frac{p(1-a)}{a(1-p)} \right)^{an} \left(\frac{1-p}{1-a} \right)^n \\
&= \left(\left(\frac{p}{a} \right)^a \left(1 + \frac{a-p}{1-a} \right)^{1-a} \right)^n \\
&\leq \left(\left(\frac{p}{a} \right)^a e^{a-p} \right)^n.
\end{aligned}$$

Ainsi pour $k = na$ on a bien finalement,

$$\mathbb{P}(S \geq k) \leq \left(\frac{np}{k} \right)^k \left(\frac{n-np}{n-k} \right)^{n-k} \leq \left(\frac{np}{k} \right)^k e^{k-np}.$$

■

Pour montrer une inégalité qui servira dans la prochaine section, nous aurons besoin des trois résultats suivants.

Proposition 2.19 (Inégalité de Markov pour des s.d.p). Soit X une matrice s.d.p aléatoire et A une matrice définie positive fixée. On a:

$$\mathbb{P}(X \not\leq A) \leq \text{Tr}(\mathbb{E}[X]A^{-1}).$$

◇

PREUVE: On rappelle que $X \not\leq A$ est équivalent à $A - X$ n'est pas s.d.p. Puisque A est définie positive, il existe une matrice B telle que $B^2 = A$. De plus B est inversible et symétrique. Ainsi pour tout x il existe un unique y tel que $x = By$. $X \not\leq A$ est donc équivalent à $B^{-1}(A - X)B^{-1} = I - B^{-1}XB^{-1}$ n'est pas s.d.p. Cela signifie qu'au moins une des valeurs propres de $B^{-1}XB^{-1}$ est strictement positive (car I et $B^{-1}XB^{-1}$ commutent). Ainsi $\|B^{-1}XB^{-1}\|_2 > 1$. Puisque la trace est la somme des valeurs propres, $\text{Tr}(B^{-1}XB^{-1}) \geq 0$ et si $X \not\leq A$ on vient de voir que $\text{Tr}(B^{-1}XB^{-1}) \geq 1$. Ainsi $\mathbb{1}_{\{X \not\leq A\}} \leq \text{Tr}(B^{-1}XB^{-1})$. On a donc:

$$\mathbb{P}(X \not\leq A) = \mathbb{E}[\mathbb{1}_{\{X \not\leq A\}}] \leq \mathbb{E}[\text{Tr}(B^{-1}XB^{-1})] = \text{Tr}(\mathbb{E}[X]B^{-1}B^{-1}) = \text{Tr}(\mathbb{E}[X]A^{-1}).$$

■

Nous admettons le lemme suivant pour la démonstration du prochain résultat.

Lemme 2.20 (Inégalité de Golden-Thompson). Pour toutes matrices symétriques A et B on a:

$$\text{Tr}(\exp(A + B)) \leq \text{Tr}((\exp A)(\exp B)).$$

◇

Proposition 2.21 (Borne de Chernoff non-commutative). Soient $X_1, \dots, X_n$ des matrices aléatoires symétriques indépendantes de taille $d \times d$. Soit A une matrice symétrique fixée. Pour toute matrice T inversible on a:

$$\mathbb{P} \left[\sum_{k=1}^n X_k \not\leq nA \right] \leq d \prod_{k=1}^n \|\mathbb{E}[\exp(TX_k T^* - TAT^*)]\|_2.$$

◇

PREUVE: Remarquons que $\text{Tr}(A) \leq d \|A\|_2$ puisque la trace est la somme des valeurs propres et qu'elles sont majorées par la norme spectrale pour une matrice symétrique. En utilisant les deux résultats précédents (le lemme 2.20, la proposition 2.19) et la proposition 2.13 on obtient:

$$\begin{aligned} \mathbb{P} \left(\sum_{k=1}^n X_k \not\leq nA \right) &= \mathbb{P} \left(\sum_{k=1}^n (X_k - A) \not\leq 0 \right) \\ &= \mathbb{P} \left(\sum_{k=1}^n T(X_k - A)T^* \not\leq 0 \right) \\ &= \mathbb{P} \left(\exp \left(\sum_{k=1}^n T(X_k - A)T^* \right) \not\leq I_d \right) \\ &\leq \text{Tr} \left(\mathbb{E}_{\{X_1, \dots, X_n\}} \left[\exp \left(\sum_{k=1}^n T(X_k - A)T^* \right) \right] \right) \\ &= \mathbb{E}_{\{X_1, \dots, X_n\}} \left[\text{Tr} \left(\exp \left(\sum_{k=1}^n T(X_k - A)T^* \right) \right) \right] \\ &\leq \mathbb{E}_{\{X_1, \dots, X_n\}} \left[\text{Tr} \left(\exp \left(\sum_{k=1}^{n-1} T(X_k - A)T^* \right) \exp(T(X_n - A)T^*) \right) \right] \\ &= \mathbb{E}_{\{X_1, \dots, X_{n-1}\}} \left[\text{Tr} \left(\exp \left(\sum_{k=1}^{n-1} T(X_k - A)T^* \right) \mathbb{E}_{\{X_n\}}[\exp(T(X_n - A)T^*)] \right) \right] \\ &\leq \|\mathbb{E}_{\{X_n\}}[\exp(T(X_n - A)T^*)]\|_2 \mathbb{E}_{\{X_1, \dots, X_{n-1}\}} \left[\text{Tr} \left(\exp \left(\sum_{k=1}^{n-1} T(X_k - A)T^* \right) \right) \right]. \end{aligned}$$

En itérant ce procédé puis en utilisant le corollaire 2.16, on obtient finalement:

$$\begin{aligned} \mathbb{P} \left(\sum_{k=1}^n X_k \not\leq nA \right) &\leq \prod_{k=2}^n \|\mathbb{E}[\exp(T(X_k - A)T^*)]\|_2 \times \mathbb{E}[\text{Tr}(\exp(T(X_1 - A)T^*))] \\ &\leq d \prod_{k=1}^n \|\mathbb{E}[\exp(T(X_k - A)T^*)]\|_2. \end{aligned}$$

■

Pour terminer cette partie, nous énoncerons sans démonstration le résultat suivant.

Théorème 2.22 (Inégalité de Bernstein non commutative). Soient $X_1, \dots, X_L$ des matrices aléatoires centrées et indépendantes (de taille $d_1 \times d_2$). Posons $\rho_k^2 = \max\{\|\mathbb{E}[X_k X_k^*]\|_2, \|\mathbb{E}[X_k^* X_k]\|_2\}$ et supposons qu'il existe une constante M telle que $\|X_k\|_2 \leq M$ presque sûrement pour tout k . Pour tout $\tau > 0$ on a alors:

$$\mathbb{P}\left(\left\|\sum_{k=1}^L X_k\right\|_2 > \tau\right) \leq (d_1 + d_2) \exp\left(\frac{-\tau^2/2}{\sum_{k=1}^L \rho_k^2 + M\tau/3}\right).$$

◇

La démonstration de ce résultat est basée sur les résultats préliminaires que nous avons présenté ainsi que certaines manipulations algébriques et approximations le rendant très technique.

3 Le théorème principal

3.1 Inégalités à forte probabilité

Définition 3.1. Soit $\Omega = \{(a_k, b_k)\}_{k=1}^m$ une classe d'indices échantillonnés uniformément au hasard avec remise. On définit l'opérateur $\mathcal{R}_\Omega$ par:

$$\mathcal{R}_\Omega(Z) = \sum_{k=1}^{|\Omega|} \langle e_{a_k} e_{b_k}^*, Z \rangle e_{a_k} e_{b_k}^*.$$

◇

Remarque 3.2. Notons que s'il y a des répétitions dans Ω alors $\mathcal{R}_\Omega$ n'est pas un opérateur de projection. En revanche on garde la propriété d'intérêt suivante:

X est dans le noyau de $\mathcal{R}_\Omega$ si et seulement si $X_{a,b} = 0$ pour tout indice $(a, b) \in \Omega$.

◇

Montrons à présent une série de résultats préliminaires qui seront utilisés de façon cruciale lors de la preuve du théorème 1.3.

Proposition 3.3. Le nombre maximum de répétitions de chaque entrée dans Ω est majoré par $\frac{8}{3}\beta \log(n_2)$ avec probabilité au moins $1 - n_2^{2-2\beta}$ pour $n_2 \geq 9$ et $\beta > 1$.

◇

PREUVE: Remarquons que pour un indice fixé, la probabilité qu'il apparaisse plus de t fois dans Ω suit une loi binomiale de paramètre m et $1/(n_1 n_2)$. Ainsi par la proposition 2.18 et par l'inégalité de Boole on a:

$$\begin{aligned} & \mathbb{P}(\text{une entrée est sélectionnée plus de } \frac{8}{3}\beta \log(n_2) \text{ fois}) \\ &= \mathbb{P}\left(\bigcup_{(i,j)} \{ \text{l'entrée } (i,j) \text{ est sélectionnée plus de } \frac{8}{3}\beta \log(n_2) \text{ fois} \}\right) \\ &\leq \sum_{(i,j)} \mathbb{P}(\{ \text{l'entrée } (i,j) \text{ est sélectionnée plus de } \frac{8}{3}\beta \log(n_2) \text{ fois} \}) \\ &= n_1 n_2 \mathbb{P}(\{ \text{l'entrée } (i,j) \text{ est sélectionnée plus de } \frac{8}{3}\beta \log(n_2) \text{ fois} \}) \\ &\leq n_1 n_2 \left(\frac{n_1 n_2}{m} \frac{8}{3}\beta \log(n_2)\right)^{-\frac{8}{3}\beta \log(n_2)} \exp\left(\frac{8}{3}\beta \log(n_2)\right) \\ &\leq n_2^2 \left(\frac{8}{3}\beta \log(n_2)\right)^{1-\frac{8}{3}\beta \log(n_2)} \quad \left(\text{puisque } \frac{n_1 n_2}{m} > 1 \text{ et } n_1 \leq n_2\right). \end{aligned}$$

Ainsi pour démontrer le résultat il suffit de montrer que: $(\frac{8}{3}\beta \log(n_2))^{1-\frac{8}{3}\beta \log(n_2)} \leq n_2^{-2\beta}$ ou de façon équivalente:

$$\begin{aligned} & \exp\left(\frac{8}{3}\beta \log(n_2) \left[1 - \log\left(\frac{8}{3}\beta \log(n_2)\right)\right]\right) \leq \exp(-2\beta \log(n_2)) \\ \iff & \frac{8}{3} \left(1 - \log\left(\frac{8}{3}\beta \log(n_2)\right)\right) \leq -\frac{3}{4} \\ \iff & \log\left(\frac{8}{3}\beta \log(n_2)\right) \geq \frac{7}{4} \\ \iff & \frac{8}{3}\beta \log(n_2) \geq \exp\left(\frac{7}{4}\right) \\ \iff & n_2 \geq \exp\left(\frac{3 \exp(7/4)}{8\beta}\right). \end{aligned}$$

Si $\beta \geq 1$ cela est équivalent à la condition:

$$n_2 \geq \exp\left(\frac{3 \exp(7/4)}{8}\right) \simeq 8,65.$$

Ainsi pour $n_2 \geq 9$ et $\beta \geq 1$ on a bien que:

$$\mathbb{P}(\text{une entrée est sélectionnée plus de } \frac{8}{3}\beta \log(n_2) \text{ fois}) \leq n_2^{-2\beta},$$

ce qu'il fallait démontrer. ■

Proposition 3.4. On rappelle que Ω désigne une classe de m indices tirés avec répétition uniformément au hasard. Pour tout $\beta > 1$, on a:

$$\frac{n_1 n_2}{m} \left\| \mathcal{P}_T \mathcal{R}_\Omega \mathcal{P}_T - \frac{m}{n_1 n_2} \mathcal{P}_T \right\| \leq \sqrt{\frac{16 \mu_0 r(n_1 n_2) \beta \log(n_2)}{3m}}$$

avec une probabilité supérieure à $1 - 2n_2^{-2\beta}$ dès lors que $m > \frac{16}{3} \mu_0 r(n_1 + n_2) \beta \log(n_2)$. ◇

PREUVE: Soit Z une matrice quelconque. On décompose Z dans la base canonique: $Z = \sum_{a,b} \langle Z, e_a e_b^* \rangle e_a e_b^*$.

D'après la proposition 2.8 on a donc:

$$\mathcal{P}_T(Z) = \sum_{a,b} \langle \mathcal{P}_T(Z), e_a e_b^* \rangle e_a e_b^* = \sum_{a,b} \langle Z, \mathcal{P}_T(e_a e_b^*) \rangle e_a e_b^*.$$

Ainsi,

$$\mathcal{R}_\Omega \mathcal{P}_T(Z) = \sum_{k=1}^m \langle e_{a_k} e_{b_k}^*, \mathcal{P}_T(Z) \rangle e_{a_k} e_{b_k}^* = \sum_{k=1}^m \langle \mathcal{P}_T(e_{a_k} e_{b_k}^*), Z \rangle e_{a_k} e_{b_k}^*.$$

Donc par linéarité de la projection,

$$\mathcal{P}_T \mathcal{R}_\Omega \mathcal{P}_T(Z) = \sum_{k=1}^m \langle \mathcal{P}_T(e_{a_k} e_{b_k}^*), Z \rangle \mathcal{P}_T(e_{a_k} e_{b_k}^*).$$

Posons à présent $\mathcal{T}_{a,b}$ l'opérateur défini par $\mathcal{T}_{a,b}(Z) = \langle \mathcal{P}_T(e_{a_k} e_{b_k}^*), Z \rangle \mathcal{P}_T(e_{a_k} e_{b_k}^*)$ de telle sorte que $\mathcal{P}_T = \mathcal{P}_T \circ \mathcal{P}_T = \sum_{a,b} \mathcal{T}_{a,b}$. Ainsi, par définition d'une loi uniforme on remarque que:

$$\frac{1}{n_1 n_2} \mathcal{P}_T = \frac{1}{n_1 n_2} \sum_{a,b} \mathcal{T}_{a,b} = \mathbb{E}[\mathcal{T}_{a_k, b_k}].$$

Notons que $\mathcal{T}_{a,b}$ est de rang 1 et que par l'inégalité de Cauchy-Schwarz,

$$\begin{aligned}\|\mathcal{T}_{a,b}(Z)\|_F &= |\langle Z, \mathcal{P}_T(e_a e_b^*) \rangle| \times \|\mathcal{P}_T(e_a e_b^*)\|_F \\ &\leq \|\mathcal{P}_T(e_a e_b^*)\|_F^2 \|Z\|_F.\end{aligned}$$

De plus par le cas d'égalité dans l'inégalité de Cauchy-Schwarz on tire que $\|\mathcal{T}_{a,b}\| = \|\mathcal{P}_T(e_a e_b^*)\|_F^2$. Ainsi, en utilisant la proposition 2.14 et le fait qu'une projection est de norme 1 on a:

$$\left\| \mathcal{T}_{a_k, b_k} - \frac{1}{n_1 n_2} \mathcal{P}_T \right\| \leq \max \left\{ \|\mathcal{P}_T(e_a e_b^*)\|_F^2, \frac{1}{n_1 n_2} \right\}.$$

D'après la proposition 2.11 on a donc:

$$\left\| \mathcal{T}_{a_k, b_k} - \frac{1}{n_1 n_2} \mathcal{P}_T \right\| \leq \mu_0 r \frac{n_1 + n_2}{n_1 n_2}.$$

Par ce qui précède on a de plus:

$$\mathbb{E} \left[\left(\mathcal{T}_{a_k, b_k} - \frac{1}{n_1 n_2} \mathcal{P}_T \right)^2 \right] = \mathbb{E}[(\mathcal{T}_{a_k, b_k} - \mathbb{E}[\mathcal{T}_{a_k, b_k}])^2] = \mathbb{E}[\mathcal{T}_{a_k, b_k}^2] - \mathbb{E}[\mathcal{T}_{a_k, b_k}]^2.$$

Or,

$$\begin{aligned}\mathcal{T}_{a_k, b_k}^2(Z) &= \mathcal{T}_{a_k, b_k}(\langle \mathcal{P}_T(e_{a_k} e_{b_k}^*), Z \rangle \mathcal{P}_T(e_{a_k} e_{b_k}^*)) \\ &= \langle \mathcal{P}_T(e_{a_k} e_{b_k}^*), \langle \mathcal{P}_T(e_{a_k} e_{b_k}^*), Z \rangle \mathcal{P}_T(e_{a_k} e_{b_k}^*) \rangle \mathcal{P}_T(e_{a_k} e_{b_k}^*) \\ &= \|\mathcal{P}_T(e_{a_k} e_{b_k}^*)\|_F^2 \mathcal{T}_{a_k, b_k}(Z).\end{aligned}$$

Ainsi, d'après les propositions 2.11 et 2.14 on a:

$$\begin{aligned}\left\| \mathbb{E} \left[\left(\mathcal{T}_{a_k, b_k} - \frac{1}{n_1 n_2} \mathcal{P}_T \right)^2 \right] \right\|_2 &= \left\| \mathbb{E}[\|\mathcal{P}_T(e_{a_k} e_{b_k}^*)\|_F^2 \mathcal{T}_{a_k, b_k}] - \frac{1}{n_1^2 n_2^2} \mathcal{P}_T \right\|_2 \\ &\leq \max \left\{ \left\| \mathbb{E}[\|\mathcal{P}_T(e_{a_k} e_{b_k}^*)\|_F^2 \mathcal{T}_{a_k, b_k}] \right\|_2, \frac{1}{n_1^2 n_2^2} \right\} \\ &\leq \max \left\{ \|\mathbb{E}[\mathcal{T}_{a_k, b_k}]\|_2 \mu_0 r \frac{n_1 + n_2}{n_1 n_2}, \frac{1}{n_1^2 n_2^2} \right\} \\ &\leq \max \left\{ \mu_0 r \frac{n_1 + n_2}{n_1^2 n_2^2}, \frac{1}{n_1^2 n_2^2} \right\} \\ &\leq \mu_0 r \frac{n_1 + n_2}{n_1^2 n_2^2}.\end{aligned}$$

Remarquons en outre que:

$$\begin{aligned}\mathbb{E}[\mathcal{R}_\Omega(Z)] &= \mathbb{E} \left[\sum_{k=1}^m \langle e_{a_k} e_{b_k}^*, Z \rangle e_{a_k} e_{b_k}^* \right] \\ &= \frac{1}{n_1 n_2} \sum_{a,b} \sum_{k=1}^m \langle e_a e_b^*, Z \rangle e_a e_b^* \\ &= \frac{m}{n_1 n_2} Z.\end{aligned}$$

Ainsi $\mathbb{E}[\mathcal{R}_\Omega] = \frac{m}{n_1 n_2} \mathcal{I}$. Donc:

$$\mathbb{E}[\mathcal{P}_T \mathcal{R}_\Omega \mathcal{P}_T] = \mathcal{P}_T \mathbb{E}[\mathcal{R}_\Omega] \mathcal{P}_T = \mathcal{P}_T \left(\frac{m}{n_1 n_2} \mathcal{I} \right) \mathcal{P}_T = \frac{m}{n_1 n_2} \mathcal{P}_T.$$

Toutes les hypothèses du théorème 2.22 sont donc vérifiées pour:

- $L = m, d_1 = n_1, d_2 = n_2$
- $X_k = \mathcal{T}_{a_k, b_k} - \mathbb{E}[\mathcal{T}_{a_k, b_k}]$
- $\tau = \frac{m}{n_1 n_2} \sqrt{\frac{16\mu_0 r(n_1 + n_2)\beta \log(n_2)}{3m}}$
- $M = \mu_0 r \frac{n_1 + n_2}{n_1 n_2}$
- $\rho_k^2 \leq \mu_0 r \frac{n_1 + n_2}{n_1^2 n_2^2}$

Le théorème 2.22 nous dit alors que l'inégalité $\frac{n_1 n_2}{m} \left\| \mathcal{P}_T \mathcal{R}_\Omega \mathcal{P}_T - \frac{m}{n_1 n_2} \mathcal{P}_T \right\| > \sqrt{\frac{16\mu_0 r(n_1 n_2)\beta \log(n_2)}{3m}}$ tient avec une probabilité majorée par:

$$(n_1 + n_2) \exp \left(\frac{\frac{-m^2}{2n_1^2 n_2^2} \cdot \frac{16\mu_0 r(n_1 + n_2)\beta \log(n_2)}{3m}}{m\mu_0 r \frac{n_1 + n_2}{n_1^2 n_2^2} + \frac{m\mu_0 r(n_1 + n_2)}{n_1^2 n_2^2} \sqrt{\frac{16\mu_0 r(n_1 + n_2)\beta \log(n_2)}{3m}}} \right).$$

Après simplification et le fait que $n_1 \leq n_2$ on peut majorer par :

$$2n_2 \cdot \exp \left(\frac{\frac{-16\beta \log(n_2)}{6}}{1 + \sqrt{\frac{16\mu_0 r(n_1 + n_2)\beta \log(n_2)}{3m}}} \right).$$

Notons $q = \frac{16}{3} \mu_0 r(n_1 + n_2)\beta \log(n_2)$ et $C = \frac{16\beta \log(n_2)}{6}$. Alors par hypothèse:

$$\begin{aligned} m > q &\iff \frac{q}{m} < 1 \\ &\iff 1 + \sqrt{\frac{q}{m}} < 2 \\ &\iff \frac{1}{1 + \sqrt{\frac{q}{m}}} > \frac{1}{2} \\ &\iff \frac{-C}{1 + \sqrt{\frac{q}{m}}} < \frac{-C}{2} \\ &\iff \exp \left(\frac{-C}{1 + \sqrt{\frac{q}{m}}} \right) < \exp \left(\frac{-C}{2} \right). \end{aligned}$$

On peut donc majorer la probabilité précédente par $2n_2 \cdot \exp \left(-\frac{4}{3}\beta \log(n_2) \right) = 2n_2^{1 - \frac{4}{3}\beta}$. Donc si $\beta > 1$ on obtient bien l'inégalité souhaitée avec une probabilité d'au moins $1 - 2n_2^{2-2\beta}$. ■

On montre à présent un résultat similaire en utilisant la même méthode (à savoir l'inégalité de Bernstein non commutative).

Proposition 3.5. En reprenant les mêmes notations que dans la proposition précédente, pour tout $\beta > 1$, on a:

$$\left\| \left(\frac{n_1 n_2}{m} \mathcal{R}_\Omega - \mathcal{I} \right) (Z) \right\|_2 \leq \sqrt{\frac{8\beta n_1 n_2^2 \log(n_1 + n_2)}{3m}} \|Z\|_\infty$$

avec une probabilité supérieure à $1 - (n_1 + n_2)^{1-\beta}$ dès lors que $m > 6\beta n_1 \log(n_1 + n_2)$. ◇

PREUVE: Soit Z une matrice quelconque. L'inégalité de Cauchy-Schwarz donne que $\|Z\|_2 \leq \sqrt{n_1 n_2} \|Z\|_\infty$. Remarquons de plus que:

$$\frac{n_1 n_2}{m} \mathcal{R}_\Omega(Z) - Z = \frac{1}{m} \sum_{k=1}^m (n_1 n_2 Z_{a_k, b_k} e_{a_k} e_{b_k}^* - Z)$$

est une somme de variables aléatoires centrées puisque:

$$\mathbb{E}[n_1 n_2 \langle e_{a_k} e_{b_k}^*, Z \rangle e_{a_k} e_{b_k}^*] = \frac{n_1 n_2}{n_1 n_2} \sum_{a, b} \langle e_a e_b^*, Z \rangle e_a e_b^* = Z.$$

On a alors la majoration suivante (dès lors que $n_1 \geq 2$).

$$\begin{aligned} \|n_1 n_2 Z_{a_k, b_k} e_{a_k} e_{b_k}^* - Z\|_2 &\leq \|n_1 n_2 Z_{a_k, b_k} e_{a_k} e_{b_k}^*\|_2 + \|Z\|_2 \\ &= n_1 n_2 |\langle Z, e_{a_k} e_{b_k}^* \rangle| \cdot \|e_{a_k} e_{b_k}^*\|_2 + \|Z\|_2 \\ &\leq n_1 n_2 \|Z\|_\infty + \sqrt{n_1 n_2} \|Z\|_\infty \\ &\leq \frac{3}{2} n_1 n_2 \|Z\|_\infty. \end{aligned}$$

De même on peut majorer la norme suivante en utilisant l'espérance calculée précédemment et la proposition 2.14.

$$\begin{aligned} &\|\mathbb{E}[(n_1 n_2 Z_{a_k, b_k} e_{a_k} e_{b_k}^* - Z)^* (n_1 n_2 Z_{a_k, b_k} e_{a_k} e_{b_k}^* - Z)]\|_2 \\ &= \|\mathbb{E}[n_1^2 n_2^2 Z_{a_k, b_k}^2 e_{b_k} e_{a_k}^* e_{a_k} e_{b_k}^* - n_1 n_2 Z_{a_k, b_k} e_{b_k} e_{a_k}^* Z - Z^* n_1 n_2 Z_{a_k, b_k} e_{a_k} e_{b_k}^* + Z^* Z]\|_2 \\ &= \|n_1^2 n_2^2 \mathbb{E}[Z_{a_k, b_k}^2 e_{b_k} e_{a_k}^* e_{a_k} e_{b_k}^* - 2 Z^* Z + Z^* Z]\|_2 \\ &= \left\| n_1 n_2 \sum_{c, d} Z_{c, d}^2 e_d e_c^* e_c e_d^* - Z^* Z \right\|_2 \\ &= \left\| n_1 n_2 \sum_{c, d} Z_{c, d}^2 e_d e_d^* - Z^* Z \right\|_2 \\ &\leq \max \left\{ \left\| n_1 n_2 \sum_{c, d} Z_{c, d}^2 e_d e_d^* \right\|_2, \|Z^* Z\|_2 \right\}. \end{aligned}$$

Le membre de droite peut être majoré de la façon suivante:

$$\|Z^* Z\|_2 = \|Z^*\|_2 \cdot \|Z\|_2 \leq n_1 n_2 \|Z^*\|_\infty \cdot \|Z\|_\infty = n_1 n_2 \|Z\|_\infty^2.$$

Le membre de gauche quant à lui peut être majoré de la façon suivante:

$$\left\| n_1 n_2 \sum_{c, d} Z_{c, d}^2 e_d e_d^* \right\|_2 \leq n_1 n_2 \|Z\|_\infty^2 \cdot \left\| \sum_{c, d} e_d e_d^* \right\|_2 \leq n_1^2 n_2 \|Z\|_\infty^2 \leq n_1 n_2^2 \|Z\|_\infty^2.$$

Ainsi,

$$\|\mathbb{E}[(n_1 n_2 Z_{a_k, b_k} e_{a_k} e_{b_k}^* - Z)^* (n_1 n_2 Z_{a_k, b_k} e_{a_k} e_{b_k}^* - Z)]\|_2 \leq n_1 n_2^2 \|Z\|_\infty^2$$

En majorant de la même manière $\|\mathbb{E}[(n_1 n_2 Z_{a_k, b_k} e_{a_k} e_{b_k}^* - Z)(n_1 n_2 Z_{a_k, b_k} e_{a_k} e_{b_k}^* - Z)^*]\|_2$ toutes les hypothèses du théorème 2.22 sont vérifiées pour:

- $L = m$, $d_1 = n_1$, $d_2 = n_2$

- $X_k = \frac{n_1 n_2}{m} Z_{a_k, b_k} e_{a_k} e_{b_k}^* - \frac{1}{m} Z$
- $M = \frac{3n_1 n_2}{2m} \|Z\|_\infty$
- $\rho_k^2 \leq n_1 n_2^2 \|Z\|_\infty^2$
- $\tau = \|Z\|_\infty \sqrt{\frac{8\beta n_1 n_2^2 \log(n_1 + n_2)}{3m}}$

Après majoration et simplification, l'inégalité de Bernstein non commutative donne de la même façon que dans la proposition précédente le résultat attendu. $\blacksquare$

Terminons cette partie en énonçant un dernier résultat préliminaire.

Proposition 3.6. En reprenant une fois de plus les mêmes notations, pour tout $\beta > 2$ on a :

$$\left\| \frac{n_1 n_2}{m} \mathcal{P}_T \mathcal{R}_\Omega(Z) - Z \right\|_\infty \leq \sqrt{\frac{8\beta \mu_0 r (n_1 + n_2) \log(n_2)}{3m}} \|Z\|_\infty$$

avec une probabilité supérieure à $1 - 2n_2^{2-\beta}$ dès lors que $m > \frac{8}{3}\beta \mu_0 r (n_1 + n_2) \log(n_2)$. $\diamond$

PREUVE: Pour tout indice (c, d) on tire un indice (a, b) uniformément au hasard et on définit alors la variable aléatoire $\xi_{c,d}$ par $\xi_{c,d} = \langle e_c e_d^*, n_1 n_2 \langle e_a e_b^*, Z \rangle \mathcal{P}_T(e_a e_b^*) - Z \rangle$. On peut alors montrer que ces variables aléatoires sont centrées:

$$\begin{aligned} \mathbb{E}[\xi_{c,d}] &= \mathbb{E}[\langle e_c e_d^*, n_1 n_2 \langle e_a e_b^*, Z \rangle \mathcal{P}_T(e_a e_b^*) - Z \rangle] \\ &= \frac{1}{n_1 n_2} \sum_{a,b} \langle e_c e_d^*, n_1 n_2 \langle e_a e_b^*, Z \rangle \mathcal{P}_T(e_a e_b^*) \rangle - Z_{c,d} \\ &= \sum_{a,b} \langle e_c e_d^*, \mathcal{P}_T(e_a e_b^*) \rangle \langle e_a e_b^*, Z \rangle - Z_{c,d} \\ &= \sum_{a,b} \delta_{a,c} \delta_{c,d} \langle e_a e_b^*, Z \rangle - Z_{c,d} \\ &= Z_{c,d} - Z_{c,d} = 0 \end{aligned}$$

(où $\delta_{x,y}$ vaut 1 si $x = y$ et 0 sinon.)

De plus, on peut majorer la valeur absolue de chacune de ces variables aléatoires en utilisant l'inégalité de Cauchy-Schwarz et la proposition 2.11:

$$\begin{aligned} |\xi_{c,d}| &= |\langle e_c e_d^*, n_1 n_2 \langle e_a e_b^*, Z \rangle \mathcal{P}_T(e_a e_b^*) - Z \rangle| \\ &\leq \|e_c e_d^*\|_F \cdot \|n_1 n_2 \langle e_a e_b^*, Z \rangle \mathcal{P}_T(e_a e_b^*) - Z\|_F \\ &\leq n_1 n_2 |\langle e_a e_b^*, Z \rangle| \cdot \|\mathcal{P}_T(e_a e_b^*)\|_F \\ &\leq n_1 n_2 \|Z\|_\infty \mu_0 r \frac{n_1 + n_2}{n_1 n_2} = \mu_0 r (n_1 + n_2) \|Z\|_\infty. \end{aligned}$$

Il est de plus possible de majorer le moment d'ordre 2 de ces variables aléatoires. Notons

$\zeta_{c,d} = \langle e_c e_d^*, n_1 n_2 \langle e_a e_b^*, Z \rangle \mathcal{P}_T(e_a e_b^*) \rangle$. Puisque $\xi_{c,d}$ est centrée on a que:

$$\begin{aligned}
\mathbb{E}[\xi_{c,d}^2] &= \text{Var}(\zeta_{c,d}) \\
&= \mathbb{E}[\zeta_{c,d}^2] - \mathbb{E}[\zeta_{c,d}]^2 \\
&= \frac{1}{n_1 n_2} \sum_{a,b} \langle e_c e_d^*, n_1 n_2 \langle e_a e_b^*, Z \rangle \mathcal{P}_T(e_a e_b^*) \rangle^2 - Z_{c,d}^2 \\
&= n_1 n_2 \sum_{a,b} \langle e_c e_d^*, \mathcal{P}_T(e_a e_b^*) \rangle^2 \langle e_a e_b^*, Z \rangle^2 - Z_{c,d}^2 \\
&\leq n_1 n_2 \|Z\|_\infty^2 \sum_{a,b} \langle \mathcal{P}_T(e_c e_d^*), e_a e_b^* \rangle^2 \\
&= n_1 n_2 \|Z\|_\infty^2 \|\mathcal{P}_T(e_c e_d^*)\|_F^2 \\
&\leq \mu_0 r(n_1 + n_2) \|Z\|_\infty^2
\end{aligned}$$

(On a utilisé l'égalité de Parseval puis la proposition 2.11).

Or, $\frac{n_1 n_2}{m} \mathcal{P}_T \mathcal{R}_\Omega(Z) - Z$ suit la même loi que $\frac{1}{m} \sum_{k=1}^m \xi_{c,d}^{(k)}$ (où $\xi_{c,d}^{(k)}$ sont des copies i.i.d de $\xi_{c,d}$) donc en utilisant l'inégalité de Boole on a que:

$$\begin{aligned}
&\mathbb{P} \left(\left\| \frac{n_1 n_2}{m} \mathcal{P}_T \mathcal{R}_\Omega(Z) - Z \right\|_\infty > \sqrt{\frac{8\beta \mu_0 r(n_1 + n_2) \log(n_2)}{3m}} \|Z\|_\infty \right) \\
&= \mathbb{P} \left(\bigcup_{c,d} \left| \left(\frac{n_1 n_2}{m} \mathcal{P}_T \mathcal{R}_\Omega(Z) - Z \right)_{c,d} \right| > \sqrt{\frac{8\beta \mu_0 r(n_1 + n_2) \log(n_2)}{3m}} \|Z\|_\infty \right) \\
&\leq \sum_{c,d} \mathbb{P} \left(\left| \frac{1}{m} \sum_{k=1}^m \xi_{c,d}^{(k)} \right| > \sqrt{\frac{8\beta \mu_0 r(n_1 + n_2) \log(n_2)}{3m}} \|Z\|_\infty \right)
\end{aligned}$$

Toutes les hypothèses du théorème 2.17 sont vérifiées pour:

- $n = m$
- $X_i = \xi_{c,d}^{(i)}$
- $M = \mu_0 r(n_1 + n_2) \|Z\|_\infty$

En appliquant l'inégalité de Bernstein pour $t = \sqrt{\frac{8\beta \mu_0 r(n_1 + n_2) \log(n_2)}{3m}} \|Z\|_\infty$ et en utilisant la majoration de $\mathbb{E}[\xi_{c,d}^2]$, on trouve bien après simplification (et majorations de même nature que celles effectuées dans la démonstration de la proposition 3.4) que:

$$\mathbb{P} \left(\left\| \frac{n_1 n_2}{m} \mathcal{P}_T \mathcal{R}_\Omega(Z) - Z \right\|_\infty > \sqrt{\frac{8\beta \mu_0 r(n_1 + n_2) \log(n_2)}{3m}} \|Z\|_\infty \right) \leq 2n_2^{2-\beta}$$

ce qu'il fallait démontrer. ■

Nous sommes à présent prêt à passer à la dernière partie concernant la démonstration du théorème principal.

3.2 Démonstration du théorème principal

On rappelle le théorème:

Soit M une matrice de taille $n_1 \times n_2$ et de rang r . Notons $U\Sigma V^*$ sa décomposition en valeurs singulières. Quitte à considérer M^* on peut supposer que $n_1 \leq n_2$. Notons de plus que dans notre décomposition on suppose que Σ est de taille $r \times r$, U de taille $n_1 \times r$ et V de taille $n_2 \times r$. Supposons que:

A0 Les sous espaces vectoriels engendrés par les lignes et les colonnes de M ont une cohérence majorée par une certaine constante $\mu_0 \geq 0$.

A1 La matrice UV^* vérifie $\|UV^*\|_\infty \leq \mu_1 \sqrt{\frac{r}{n_1 n_2}}$ pour une certaine constante $\mu_1 \geq 0$.

Supposons qu'un nombre m d'entrées sont observées uniformément au hasard et notons Ω l'ensemble des indices des entrées observées. Alors si

$$m \geq 32 \max\{\mu_1^2, \mu_0\} r (n_1 + n_2) \beta \log^2(2n_2)$$

pour un réel $\beta > 1$, alors la matrice minimisant le problème suivant

$$\begin{aligned} & \text{minimiser} && \|X\|_* \\ & \text{sous la contrainte} && X_{i,j} = M_{i,j} \quad \forall (i,j) \in \Omega. \end{aligned}$$

est unique et est égal à M avec une probabilité supérieure à $1 - 6 \log(n_2)(n_1 + n_2)^{2-2\beta} - n_2^{2-2\beta^{1/2}}$.

Notation: Reprenons les notations du théorème 1.3. On dit que Ω *échoue* lorsqu'il existe une matrice minimisant la condition 2 qui n'est pas la matrice M recherchée. $\diamond$

Avant de passer à la démonstration à proprement parler, donnons un dernier résultat portant sur le mode de tirage des entrées observées de la matrice.

Proposition 3.7. La probabilité que Ω échoue lorsque les m entrées sont tirées uniformément au hasard parmi les $n_1 n_2$ possibles est inférieure à celle obtenue lorsque les entrées sont tirées indépendamment avec remise. $\diamond$

Cette proposition semble naturelle puisqu'on observera en moyenne plus d'entrées en tirant uniformément au hasard les m entrées que lorsque l'on autorise les duplicats. Rigoureusement on peut le montrer de la façon suivante:

PREUVE: Notons Ω' une classe de m entrées tirées au hasard indépendamment avec remise parmi les $n_1 n_2$ possibles et notons Ω_k un ensemble de k entrées tirées uniformément au hasard parmi toutes les collections de taille k possibles. Il est clair que $\mathbb{P}(\Omega_\alpha \text{ échoue}) \geq \mathbb{P}(\Omega_\beta \text{ échoue})$ lorsque $\alpha \leq \beta$. Ainsi on a:

$$\begin{aligned} \mathbb{P}(\Omega' \text{ échoue}) &= \sum_{k=0}^m \mathbb{P}(\Omega' \text{ échoue} \mid |\Omega'| = k) \mathbb{P}(|\Omega'| = k) \\ &= \sum_{k=0}^m \mathbb{P}(\Omega_k \text{ échoue}) \mathbb{P}(|\Omega'| = k) \\ &\geq \mathbb{P}(\Omega_m \text{ échoue}) \sum_{k=0}^m \mathbb{P}(|\Omega'| = k) \\ &= \mathbb{P}(\Omega_m \text{ échoue}). \end{aligned}$$

■

Achevons ce document en démontrant le théorème 1.3.

La proposition précédente nous permet donc de considérer le théorème 1.3 dans le cas où Ω est une classe d'indices tirés indépendamment avec remise.

On suppose dans un premier temps que de nombreuses hypothèses sont vérifiées puis nous justifierons qu'elles le sont avec grande probabilité dans un second temps.

On suppose donc que:

$$\frac{n_1 n_2}{m} \left\| \mathcal{P}_T \mathcal{R}_\Omega \mathcal{P}_T - \frac{m}{n_1 n_2} \mathcal{P}_T \right\| \leq \frac{1}{2}, \quad \|\mathcal{R}_\Omega\| \leq \frac{8}{3} \beta^{1/2} \log(n_2). \quad (5)$$

Et on suppose de plus qu'il existe une matrice Y dans l'image de $\mathcal{R}_\Omega$ telle que:

$$\|\mathcal{P}_T(Y) - UV^*\|_F \leq \sqrt{\frac{r}{2n_2}}, \quad \|\mathcal{P}_{T^\perp}(Y)\|_2 < \frac{1}{2}. \quad (6)$$

Remarquons que pour toute matrice A on a par l'inégalité triangulaire:

$$\begin{aligned} \|\mathcal{R}_\Omega(Z)\|_F &= \|\mathcal{R}_\Omega(\mathcal{P}_T(Z) - \mathcal{P}_{T^\perp}(Z))\|_F \\ &= \|\mathcal{R}_\Omega(\mathcal{P}_T(Z)) - \mathcal{R}_\Omega(\mathcal{P}_{T^\perp}(Z))\|_F \\ &\geq \|\mathcal{R}_\Omega(\mathcal{P}_T(Z))\|_F - \|\mathcal{R}_\Omega(\mathcal{P}_{T^\perp}(Z))\|_F. \end{aligned}$$

Ainsi, (5) implique que pour toute matrice Z dans le noyau de $\mathcal{R}_\Omega$ on a:

$$\|\mathcal{R}_\Omega(\mathcal{P}_T(Z))\|_F \leq \|\mathcal{R}_\Omega(\mathcal{P}_{T^\perp}(Z))\|_F.$$

Notons de plus de $\mathcal{R}_\Omega^2 = \mathcal{R}_\Omega$ donc en utilisant de plus les propriétés des projections (voir proposition 2.8) on a:

$$\begin{aligned} \|\mathcal{R}_\Omega \mathcal{P}_T(Z)\|_F^2 &= \langle \mathcal{R}_\Omega \mathcal{P}_T(Z), \mathcal{R}_\Omega \mathcal{P}_T(Z) \rangle \\ &= \langle \mathcal{P}_T(Z), \mathcal{R}_\Omega^2 \mathcal{P}_T(Z) \rangle \\ &= \langle Z, \mathcal{P}_T \mathcal{R}_\Omega^2 \mathcal{P}_T(Z) \rangle \\ &= \langle Z, \mathcal{P}_T \mathcal{R}_\Omega \mathcal{P}_T(Z) \rangle \\ &= \langle Z, \mathcal{P}_T \mathcal{R}_\Omega \mathcal{P}_T(Z) - \frac{m}{n_1 n_2} \mathcal{P}_T(Z) \rangle + \frac{m}{n_1 n_2} \langle Z, \mathcal{P}_T(Z) \rangle \\ &= \frac{m}{n_1 n_2} \|\mathcal{P}_T(Z)\|_F^2 + \langle Z, \left(\mathcal{P}_T \mathcal{R}_\Omega \mathcal{P}_T - \frac{m}{n_1 n_2} \mathcal{P}_T \right) (Z) \rangle \\ &= \frac{m}{n_1 n_2} \|\mathcal{P}_T(Z)\|_F^2 + \langle \mathcal{P}_T(Z), \left(\mathcal{P}_T \mathcal{R}_\Omega \mathcal{P}_T - \frac{m}{n_1 n_2} \mathcal{P}_T \right) (\mathcal{P}_T(Z)) \rangle. \end{aligned}$$

D'après l'inégalité de Cauchy-Schwarz et l'hypothèse (5) on a que:

$$\begin{aligned} \left| \langle \mathcal{P}_T(Z), \left(\mathcal{P}_T \mathcal{R}_\Omega \mathcal{P}_T - \frac{m}{n_1 n_2} \mathcal{P}_T \right) (\mathcal{P}_T(Z)) \rangle \right| &\leq \|\mathcal{P}_T(Z)\|_F \cdot \left\| \left(\mathcal{P}_T \mathcal{R}_\Omega \mathcal{P}_T - \frac{m}{n_1 n_2} \mathcal{P}_T \right) (\mathcal{P}_T(Z)) \right\|_F \\ &\leq \frac{m}{2n_1 n_2} \|\mathcal{P}_T(Z)\|_F^2. \end{aligned}$$

Ainsi on a que:

$$\|\mathcal{R}_\Omega \mathcal{P}_T(Z)\|_F^2 \geq \frac{m}{n_1 n_2} \|\mathcal{P}_T(Z)\|_F^2 - \frac{m}{2n_1 n_2} \|\mathcal{P}_T(Z)\|_F^2 = \frac{m}{2n_1 n_2} \|\mathcal{P}_T(Z)\|_F^2.$$

Combinant ce fait avec l'hypothèse (6) et la condition 1 (puisque les constantes donnent $\frac{32 \times 9}{128} = 2,25 > 2$) on tire que:

$$\begin{aligned}
\|\mathcal{P}_{T^\perp}(Z)\|_F &\geq \frac{3}{8\beta^{1/2}\log(n_2)} \|\mathcal{R}_\Omega \mathcal{P}_{T^\perp}(Z)\|_F \\
&\geq \frac{3}{8\beta^{1/2}\log(n_2)} \|\mathcal{R}_\Omega \mathcal{P}_T(Z)\|_F \\
&\geq \frac{3}{8\beta^{1/2}\log(n_2)} \sqrt{\frac{m}{2n_1n_2}} \|\mathcal{P}_T(Z)\|_F \\
&= \sqrt{\frac{9m}{128\beta n_1n_2 \log^2(n_2)}} \|\mathcal{P}_T(Z)\|_F \\
&> \sqrt{\frac{2r}{n_2}} \|\mathcal{P}_T(Z)\|_F.
\end{aligned}$$

A présent pour $Z \in \text{Ker } \mathcal{R}_\Omega$ on choisit $U_\perp$ et $V_\perp$ tels que $[U, U_\perp]$ et $[V, V_\perp]$ soient des matrices unitaires (donc de norme spectrale égale à 1) et tels que $\langle U_\perp V_\perp^*, \mathcal{P}_{T^\perp}(Z) \rangle = \|\mathcal{P}_{T^\perp}(Z)\|_\star$. En utilisant la proposition 2.6 on a que:

$$\begin{aligned}
\|M + Z\|_\star &\geq \langle UV^* + U_\perp V_\perp^*, M + Z \rangle \\
&= \langle UV^* + U_\perp V_\perp^*, M \rangle + \langle UV^* + U_\perp V_\perp^*, Z \rangle \\
&= \langle UV^* + U_\perp V_\perp^*, U\Sigma V^* \rangle + \langle UV^* + U_\perp V_\perp^*, Z \rangle.
\end{aligned}$$

Or,

$$\langle UV^*, U\Sigma V^* \rangle = \text{Tr}(V\Sigma U^* UV^*) = \text{Tr}(V\Sigma V^*) = \text{Tr}(V^* V \Sigma) = \text{Tr}(\Sigma) = \|M\|_\star$$

et,

$$\langle U_\perp V_\perp^*, U\Sigma V^* \rangle = \text{Tr}(V\Sigma U^* U_\perp V_\perp^*) = \text{Tr}(V_\perp^* V \Sigma U_\perp^* U) = 0$$

donc,

$$\begin{aligned}
\|M + Z\|_\star &\geq \|M\|_\star + \langle UV^* + U_\perp V_\perp^*, Z \rangle \\
&= \|M\|_\star + \langle UV^* + U_\perp V_\perp^*, \mathcal{P}_T(Z) + \mathcal{P}_{T^\perp}(Z) \rangle \\
&= \|M\|_\star + \langle UV^*, \mathcal{P}_T(Z) \rangle + \langle U_\perp V_\perp^*, \mathcal{P}_{T^\perp}(Z) \rangle.
\end{aligned}$$

Mais, par hypothèse il existe une matrice A telle que $Y = \mathcal{R}_\Omega(A)$ donc:

$$\begin{aligned}
\langle \mathcal{P}_T(Y), \mathcal{P}_T(Z) \rangle + \langle \mathcal{P}_{T^\perp}(Y), \mathcal{P}_{T^\perp}(Z) \rangle &= \langle \mathcal{P}_T(Y), Z \rangle + \langle \mathcal{P}_{T^\perp}(Y), Z \rangle \\
&= \langle Y, Z \rangle \\
&= \langle \mathcal{R}_\Omega(A), Z \rangle \\
&= \langle A, \mathcal{R}_\Omega(Z) \rangle \\
&= \langle A, 0 \rangle = 0.
\end{aligned}$$

Ainsi, par l'inégalité de Cauchy-Schwarz et ce qui précède:

$$\begin{aligned}
\|M + Z\|_\star &\geq \|M\|_\star + \langle UV^* - \mathcal{P}_T(Y), \mathcal{P}_T(Z) \rangle + \langle U_\perp V_\perp^* - \mathcal{P}_{T^\perp}(Y), \mathcal{P}_{T^\perp}(Z) \rangle \\
&> \|M\|_\star - \|\mathcal{P}_T(Y) - UV^*\|_F \cdot \|\mathcal{P}_T(Z)\|_F + \|\mathcal{P}_{T^\perp}(Y)\|_F \cdot \|\mathcal{P}_{T^\perp}(Z)\|_F + \|\mathcal{P}_{T^\perp}\|_\star \\
&> \|M\|_\star - \sqrt{\frac{r}{2n_2}} \|\mathcal{P}_T(Z)\|_F + \frac{1}{2} \|\mathcal{P}_{T^\perp}(Z)\|_F \\
&\geq \|M\|_\star.
\end{aligned}$$

Cette inégalité termine la preuve dans le cas où toutes les hypothèses faites (voir (5) et (6)) sont justifiées car pour toute matrice X telle que $X - M = Z \in \text{Ker } \mathcal{R}_\Omega$ (c'est à dire telle que $X_{a,b} = M_{a,b}$

pour tout $(a, b) \in \Omega$ on a $\|X\|_* > \|M\|_*$ donc M est bien l'unique matrice minimisant $\|X\|_*$ sous la condition (2).

Montrons à présent qu'un tel Y existe avec forte probabilité. Pour cela, on partitionne $1, \dots, m$ en p parties de taille q . Par la condition (1) on peut supposer que:

$$q \geq \frac{128}{3} \max\{\mu_0, \mu_1^2\} r(n_1 + n_2) \beta \log(n_1 n_2), \quad \text{et} \quad p \geq \frac{3}{4} \log(2n_2).$$

On pose alors Ω_j l'ensemble des indices dans la j -ème partition. Par hypothèse sur le mode de tirage, ces partitions sont indépendantes les unes des autres. Supposons que:

$$\frac{n_1 n_2}{q} \left\| \mathcal{P}_T \mathcal{R}_{\Omega_k} \mathcal{P}_T - \frac{q}{n_1 n_2} \mathcal{P}_T \right\|_2 \leq \frac{1}{2} \quad (7)$$

pour tout k . On pose $W_0 = UV^*$ et on définit simultanément $Y_k = \frac{n_1 n_2}{q} \sum_{j=1}^k \mathcal{R}_{\Omega_j}(W_{j-1})$ et $W_k = UV^* - \mathcal{P}_T(Y_k)$ pour k allant de 1 à p . Puisque pour tout k , $W_k \in T$ on a par l'hypothèse (7) que:

$$\begin{aligned} \|W_k\|_F &= \|UV^* - \mathcal{P}_T(Y_k)\|_F \\ &= \left\| UV^* - \mathcal{P}_T \left(\frac{n_1 n_2}{q} \sum_{j=1}^{k-1} \mathcal{R}_{\Omega_j}(W_{j-1}) + \mathcal{R}_{\Omega_k}(W_{k-1}) \right) \right\|_F \\ &= \left\| UV^* - \mathcal{P}_T \left(\frac{n_1 n_2}{q} \sum_{j=1}^{k-1} \mathcal{R}_{\Omega_j}(W_{j-1}) \right) - \frac{n_1 n_2}{q} \mathcal{P}_T(\mathcal{R}_{\Omega_k}(W_{k-1})) \right\|_F \\ &= \left\| W_{k-1} - \frac{n_1 n_2}{q} \mathcal{P}_T(\mathcal{R}_{\Omega_k}(W_{k-1})) \right\|_F \\ &= \left\| \left(\mathcal{P}_T - \frac{n_1 n_2}{q} \mathcal{P}_T \mathcal{R}_{\Omega_k} \mathcal{P}_T \right) (W_{k-1}) \right\|_F \\ &\leq \frac{1}{2} \|W_{k-1}\|_F. \end{aligned}$$

Ainsi, $\|W_k\|_F \leq 2^{-k} \|W_0\|_F = 2^{-k} \sqrt{r}$. Puisque $p \geq \frac{3}{4} \log(2n_2) \geq \frac{1}{2} \log_2(2n_2)$ on a:

$$\|UV^* - \mathcal{P}_T(Y_p)\|_F \leq 2^{-\log_2(\sqrt{2n_2})} \sqrt{r} = \sqrt{\frac{r}{2n_2}}$$

donc la première inégalité de (6) est vérifiée pour un tel Y_p .

A présent supposons que:

$$\left\| W_{k-1} - \frac{n_1 n_2}{q} \mathcal{P}_T \mathcal{R}_{\Omega_k}(W_{k-1}) \right\|_\infty \leq \frac{1}{2} \|W_{k-1}\|_\infty \quad (8)$$

$$\left\| \left(\frac{n_1 n_2}{q} \mathcal{R}_{\Omega_j} - \mathcal{I} \right) (W_{j-1}) \right\|_2 \leq \sqrt{\frac{8n_1 n_2^2 \beta \log(n_1 + n_2)}{3q}} \|W_{j-1}\|_\infty \quad (9)$$

pour tout k allant de 1 à p .

Montrons alors que l'inégalité $\|\mathcal{P}_{T^\perp}(Y_p)\|_2 < \frac{1}{2}$ est une conséquence des hypothèses (8) et (9). De même que pour la norme de Frobenius, on a pour la norme infinie que $\|W_k\|_\infty \leq 2^{-k} \|UV^*\|_\infty$.

En utilisant successivement l'inégalité triangulaire, le fait que $W_k \in T$, $\forall k \leq p-1$, les propriétés des projections (voir Proposition 2.8 et le fait qu'une projection est de norme 1), les hypothèses (9) et (8), les condition (2) et (1) et finalement la minoration de q on a que:

$$\begin{aligned}
\|\mathcal{P}_{T^\perp}(Y_p)\|_2 &\leq \sum_{j=1}^p \left\| \frac{n_1 n_2}{q} \mathcal{P}_{T^\perp} \mathcal{R}_{\Omega_j}(W_{j-1}) \right\|_2 \\
&= \sum_{j=1}^p \left\| \mathcal{P}_{T^\perp} \left(\frac{n_1 n_2}{q} \mathcal{R}_{\Omega_j}(W_{j-1}) - W_{j-1} \right) \right\|_2 \\
&\leq \sum_{j=1}^p \left\| \left(\frac{n_1 n_2}{q} \mathcal{R}_{\Omega_j} - \mathcal{I} \right) (W_{j-1}) \right\|_2 \\
&\leq \sum_{j=1}^p \sqrt{\frac{8n_1 n_2^2 \beta \log(n_1 n_2)}{3q}} \|W_{j-1}\|_\infty \\
&\leq 2 \sum_{j=1}^p \sqrt{\frac{8n_1 n_2^2 \beta \log(n_1 n_2)}{3q}} 2^{-j} \|UV^*\|_\infty \\
&< \sqrt{\frac{32n_1 n_2^2 \beta \log(n_1 n_2)}{3q}} \sum_{j=1}^p 2^{-j} \mu_1 \sqrt{\frac{r}{n_1 n_2}} \\
&< \sqrt{\frac{32\mu_1^2 r n_2 \beta \log(n_1 + n_2)}{3q}} \sum_{j=1}^p 2^{-j} \\
&< \sqrt{\frac{32\mu_1^2 r n_2 \beta \log(n_1 + n_2)}{3q}} \\
&< \frac{1}{2}.
\end{aligned}$$

Il ne reste donc plus qu'à justifier que nos hypothèses tiennent avec une forte probabilité, or c'est précisément ce que disent les résultats de la section 3.1. En effet la première inégalité de l'hypothèse (5) est vérifiée avec une probabilité supérieure à $1 - 2n_2^{2-2\beta}$ par la proposition 3.4 tandis que la deuxième est vérifiée avec une probabilité supérieure à $1 - n_2^{2-2\beta^{1/2}}$ par la proposition 3.3. Pour tout k , (7) est vérifiée avec une probabilité supérieure à $1 - 2n_2^{2-2\beta}$ par la proposition 3.4, (8) est vérifiée avec une probabilité supérieure à $1 - 2n_2^{2-2\beta}$ par la proposition 3.6, et enfin (9) est vérifiée avec une probabilité supérieure à $1 - (n_1 + n_2)^{1-2\beta}$ par la proposition 3.5.

Par l'inégalité de Boole il nous suffit de sommer ces probabilités et le résultat tient donc avec une probabilité supérieure à $1 - \left(6n_2^{2-2\beta} + n_2^{2-2\beta^{1/2}} + (n_1 + n_2)^{1-2\beta} \right)$.

Pour obtenir la probabilité annoncée dans le théorème 1.3 il suffit donc de montrer que

$$6n_2^{2-2\beta} + (n_1 + n_2)^{1-2\beta} \leq 6\log(n_2)(n_1 + n_2)^{2-2\beta}.$$

Or, en utilisant le fait que $n_1 \leq n_2$ on a la suite d'équivalences suivante:

$$\begin{aligned}
6n_2^{2-2\beta} + (n_1 + n_2)^{1-2\beta} &\leq 6\log(n_2)(n_1 + n_2)^{2-2\beta} \\
\iff 6\left(\frac{n_2}{n_1 + n_2}\right)^{2-2\beta} + 1 &\leq 6\log(n_2)(n_1 + n_2) \\
\iff 6\left(\frac{n_1 + n_2}{n_2}\right)^{2\beta-2} + 1 &\leq 6\log(n_2)(n_1 + n_2) \\
\iff 3 \cdot 2^{2\beta-1} + 1 &\leq 6\log(n_2)(n_1 + n_2).
\end{aligned}$$

Or cette dernière inégalité est vérifiée pour n_1 et n_2 assez grands, hypothèse que l'on fait donc ici. Notons pour conclure que cette hypothèse est raisonnable car on s'intéresse en pratique à des matrices de très grande taille. On a donc bien montré que la matrice minimisant le problème suivant

$$\begin{array}{ll}
\text{minimiser} & \|X\|_{\star} \\
\text{sous la contrainte} & X_{i,j} = M_{i,j} \quad \forall (i,j) \in \Omega.
\end{array}$$

est unique et est égal à M avec une probabilité supérieure à $1 - 6\log(n_2)(n_1 + n_2)^{2-2\beta} - n_2^{2-2\beta^{1/2}}$.

Ouverture

Pour achever ce document nous pouvons rappeler quelques pistes d'amélioration indiquées dans [\[Rec09\]](#). Une première idée serait de tenter d'améliorer la constante 32 dans le nombre d'entrées requises en essayant de trouver des majorations les plus fines possibles. Deuxièmement il est possible d'adapter ce résultat au cas où les entrées observées seraient perturbées par un bruit (si celui-ci n'est pas trop important). Enfin, il serait intéressant de savoir si l'hypothèse **A1** est nécessaire ou au contraire contingente.

REFERENCES

- [CR08] Emmanuel J. Candes and Benjamin Recht.
Exact Matrix Completion via Convex Optimization. 2008. arXiv: 0805.4471 [cs.IT].
URL: <https://arxiv.org/abs/0805.4471>.
- [CT09] Emmanuel J. Candes and Terence Tao.
The Power of Convex Relaxation: Near-Optimal Matrix Completion. 2009.
arXiv: 0903.1476 [cs.IT]. URL: <https://arxiv.org/abs/0903.1476>.
- [HR90] Torben Hagerup and Christine Rüb. “A guided tour of chernoff bounds”.
In: *Information Processing Letters* 33.6 (1990), pp. 305–308. ISSN: 0020-0190.
DOI: [https://doi.org/10.1016/0020-0190\(90\)90214-I](https://doi.org/10.1016/0020-0190(90)90214-I).
URL: <https://www.sciencedirect.com/science/article/pii/002001909090214I>.
- [KMO09] Raghunandan H. Keshavan, Andrea Montanari, and Sewoong Oh.
Matrix Completion from a Few Entries. 2009. arXiv: 0901.3150 [cs.LG].
URL: <https://arxiv.org/abs/0901.3150>.
- [Rec09] Benjamin Recht. *A Simpler Approach to Matrix Completion*. 2009.
arXiv: 0910.0651 [cs.IT]. URL: <https://arxiv.org/abs/0910.0651>.