

## UN PROBLÈME D'APPROXIMATION MATRICIELLE : QUELLE EST LA MATRICE BISTOCHASTIQUE LA PLUS PROCHE D'UNE MATRICE DONNÉE ? \*

PAWOUMODOM L. TAKOUDA<sup>1</sup>

**Abstract.** We are interested in the following work in the doubly stochastic matrix nearness problem. Instances of this problems occurs in different fields: aggregation of preferences in operational research, calculus of variations and shape optimisation, etc. We propose here a direct study *via* the projection theorem and a numerical resolution inspired by the alternating projections algorithm of Boyle-Dykstra.

**Résumé.** Nous nous intéressons dans ce travail au problème d'approximation d'une matrice donnée par une matrice bistochastique. Des instances de ce problème peuvent apparaître dans différents domaines : en recherche opérationnelle dans un problème d'agrégation de préférence, en calcul de variations et optimisation de forme entre autres. Nous en proposons dans cet article une étude directe *via* le théorème de projection et une résolution numérique inspirée par la méthode de projections alternées de Boyle-Dykstra.

**Mots Clés.** Approximation matricielle, projections alternées.

**Classification Mathématique.** 90C25.

### 1. INTRODUCTION

Dans cet article, nous nous intéressons au problème suivant : étant donnée une matrice  $M$ , carrée d'ordre  $n$ , trouvez-en la matrice bistochastique la plus proche,

---

Reçu le 23 janvier 2003, accepté le 19 novembre 2004.

\* *Version abrégée du rapport interne MIP 02-21* [16].

<sup>1</sup> Department of Management Sciences, University of Waterloo, 200 university avenue west, Waterloo, Ontario N2L 3G1, Canada ; [ptakouda@uwaterloo.ca](mailto:ptakouda@uwaterloo.ca)

c'est-à-dire

$$(\mathcal{P}) \quad \begin{cases} \text{minimiser } \|M - B\| \\ \text{tel que } B \in \mathbb{B}_n, \end{cases} \quad (1)$$

où  $\mathbb{B}_n$  désigne l'ensemble des matrices bistochastiques et  $\|\cdot\|$  désigne la norme de Fröbenius sur les matrices. Ce problème rentre dans le cadre des problèmes d'approximation matricielle étudiés par Higham [10]. Il présente la particularité que l'ensemble réalisable (sur lequel l'on veut approximer)  $\mathbb{B}_n$  s'écrit comme **intersection d'un sous-espace et d'un cône convexe fermés**. Des instances de ce type apparaissent, notamment, en théorie des graphes, en optimisation sous contraintes de semi-définie positivité, en recherche opérationnelle, en statistiques et finance (voir [17]).

La section 2 rappelle quelques notations et notions que nous utilisons. Dans la section 3, après avoir proposé une nouvelle preuve du théorème de Birkhoff, nous faisons une première étude du problème d'approximation par matrices bistochastiques et nous présentons brièvement l'algorithme de Boyle-Dykstra. La section 4 présente la mise en œuvre de cet algorithme pour notre problème et des tests numériques suivis d'une conclusion en section 5.

## 2. RAPPELS ET NOTATIONS

Dans toute la suite, nous utiliserons, sauf indication contraire, des notions et notations classiques d'analyse matricielle et analyse convexe (voir [12, 13]). En particulier, l'identification de  $\mathcal{M}_n(\mathbb{R})$  à  $\mathbb{R}^{n^2}$  sera faite ligne par ligne, soit au sens suivant :

$$M = (a_{ij})_{i,j} \mapsto x = (a_{11}, \dots, a_{1n}, a_{21}, \dots, a_{2n}, \dots, a_{n1}, \dots, a_{nn}).$$

Nous notons  $e$  le vecteur de  $\mathbb{R}^n$  défini par  $e = (1, \dots, 1)^T$ ,  $rg(M)$ , le rang de la matrice  $M$ ,  $tr(M)$  sa trace.

Nous nous plaçons dans l'espace  $\mathcal{M}_n(\mathbb{R})$  muni d'une structure d'espace de Hilbert par le produit scalaire de Fröbenius

$$\langle A, B \rangle = tr(A^T B) \text{ dont la norme associée est : } \|A\| = \sqrt{tr(A^T A)}.$$

**Lemme 2.1** [5, 12, 15]. *Considérons une partie  $\mathcal{C}$  convexe fermée non vide de  $(\mathbb{H}, \langle \cdot, \cdot \rangle)$ , espace de Hilbert.*

*Pour tout point  $x$  de  $\mathbb{H}$ , il existe un et un seul point  $\bar{c}$  de  $\mathcal{C}$  tel que :*

$$\|x - \bar{c}\| = \inf\{\|x - c\|, c \in \mathcal{C}\},$$

caractérisé par :

$$\begin{cases} \bar{c} \in \mathcal{C}, \\ \langle x - \bar{c}, c - \bar{c} \rangle \leq 0 \quad \forall c \in \mathcal{C}. \end{cases} \quad (2)$$

Si  $\mathcal{C}$  est un sous espace fermé de  $\mathbb{H}$ , alors la caractérisation (2) devient :

$$\begin{cases} \bar{c} \in \mathcal{C}, \\ x - \bar{c} \in \vec{\mathcal{C}}^\perp, \end{cases} \quad (3)$$

où  $\vec{\mathcal{C}}^\perp$  est le sous-espace orthogonal de la direction  $\vec{\mathcal{C}}$  du sous-espace  $\mathcal{C}$ .

Le point  $\bar{c}$  est appelé *projeté* de  $x$  sur l'ensemble  $\mathcal{C}$ , noté  $\bar{c} = \mathcal{P}_{\mathcal{C}}(x)$  ou  $\mathcal{P}_{\mathcal{C}}x$  définissant ainsi un opérateur de projection  $\mathcal{P}_{\mathcal{C}} : \mathbb{H} \mapsto \mathbb{H}$ . De (2) et (3), on déduit :

- (i) si  $x \in \mathbb{H}$ , on a :  $x - \mathcal{P}_{\mathcal{C}}(x) \in \mathcal{N}_{\mathcal{C}}(\mathcal{P}_{\mathcal{C}}(x))$ , cône normal à  $\mathcal{C}$  en  $\mathcal{P}_{\mathcal{C}}(x)$ ;
- (ii) on suppose que  $\mathcal{C}$  est un sous-espace vectoriel (resp. affine), alors  $\mathcal{P}_{\mathcal{C}}$  est *linéaire* (resp. *affine*).

### 3. APPROXIMATION PAR MATRICES BISTOCHASTIQUES : ÉTUDE PRÉLIMINAIRE

#### 3.1. LE POLYTOPE $\mathbb{B}_n$ DES MATRICES BISTOCHASTIQUES

##### 3.1.1. Définitions

Soit  $M$  une matrice carrée d'ordre  $n$  ( $n \in \mathbb{N}^*$ ).

**Définition 3.1.**  $M$  est appelée *matrice bistochastique* si on a :

- (1)  $M \geq 0$  (c'est-à-dire toutes les composantes sont positives);
- (2)  $Me = e$ ;
- (3)  $M^T e = e$ .

Pour  $n \in \mathbb{N}^*$  fixé, nous noterons  $\mathbb{B}_n$  l'ensemble des matrices bistochastiques. C'est un polyèdre convexe compact. En effet, il est immédiat que  $\mathbb{B}_n \equiv \{x \in \mathbb{R}^{n^2} \mid \mathcal{A}x = b, x \geq 0\}$  où  $\mathcal{A} \in \mathcal{M}_{2n, n^2}$  est définie sous la forme blocs suivante :

$$\mathcal{A} = \begin{pmatrix} e^T & 0 & \dots & 0 \\ 0 & e^T & \dots & 0 \\ \vdots & \vdots & \ddots & 0 \\ 0 & \dots & \dots & e^T \\ I_n & \dots & \dots & I_n \end{pmatrix}, \quad (4)$$

et  $b = (1, \dots, 1)^T \in \mathbb{R}^{2n}$ .

Il y a eu énormément de travaux mathématiques concernant les matrices bistochastiques, concernant notamment leur géométrie, la conjecture de van Der Waerden et leur utilisation dans différents domaines (voir [17]).

Prenons par exemple  $n = 2$ . On peut facilement voir que les matrices de  $\mathbb{B}_2$  sont celles qui peuvent se mettre sous la forme

$$M = \begin{pmatrix} a & 1-a \\ 1-a & a \end{pmatrix}, \text{ avec } a \in [0, 1].$$

On peut alors écrire  $M = aI_2 + (1 - a)P_1$ , avec  $I_2$  matrice identité et  $P_1 = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}$ .

L'ensemble  $\mathbb{B}_2$  est donc simplement le segment d'extrémités  $I_2$  et  $P_1$ , qui en sont par conséquent les points extrémaux. Nous allons nous intéresser aux points extrémaux de  $\mathbb{B}_n$  au prochain paragraphe.

### 3.1.2. Points extrémaux

Un point d'un convexe est extrémal s'il ne peut s'écrire comme combinaison convexe d'autres points du convexe. Pour déterminer ceux de  $\mathbb{B}_n$ , nous utilisons le lemme suivant prouvé dans [17].

**Lemme 3.2** [17]. *Soit  $P$  un polyèdre convexe dans  $\mathbb{R}^n$ .*

*Si  $P$  est de la forme*

$$P = \{x \mid Dx = \delta, x \geq 0\},$$

*avec  $D$  une matrice  $m \times n$  et  $\delta$  un vecteur donnés, alors les propositions suivantes sont équivalentes :*

- (1)  $\bar{x}$  élément non nul de  $P$  est point extrémal de  $P$  ;
- (2) les colonnes de  $D$  correspondant aux composantes non nulles de  $\bar{x}$  sont linéairement indépendantes.

De (4), il vient immédiatement que :

$$\text{rg}(A) = 2n - 1.$$

D'autre part, pour  $M = (a_{ij})_{i,j}$  une matrice bistochastique,

- (1) on a :  $\forall(i, j), 0 \leq a_{ij} \leq 1$ , et  $M$  n'a aucune ligne (resp. colonne) nulle ;
- (2) si l'une des composantes de  $M$  vaut 1, alors les autres composantes de la ligne et de la colonne auxquelles elle appartient sont toutes égales à 0.

Soit donc  $S$  une matrice bistochastique, supposons qu'elle est un *point extrémal* de  $\mathbb{B}_n$ . Alors, d'après le lemme 3.2, les colonnes de  $A$  (voir (4)) correspondant aux composantes non nulles de  $S$  doivent être linéairement indépendantes. On en déduit :

- $S$  a au maximum  $2n - 1$  composantes non nulles, puisque le rang de  $A$  est  $2n - 1$ .
- $S$  a au moins une ligne composée d'un seul élément non nul, qui vaut alors 1.
- Si on supprime la ligne et la colonne de  $S$  contenant cet unique élément non nul, la sous-matrice obtenue est bistochastique et est un point extrémal de  $\mathbb{B}_{n-1}$ .

On peut alors montrer par récurrence que :

**Théorème 3.3** (voir [17]). *Pour que  $S$  soit un point extrémal de  $\mathbb{B}_n$ , il faut que toutes les lignes (et colonnes) de  $S$  aient une et une seule composante non nulle (qui vaut alors 1).*

*Cette proposition est suffisante.*

Une matrice dont toutes les lignes (et les colonnes) ont exactement une composante non nulle égale à 1 est une **matrice de permutation**. Le théorème 3.3 est exactement le Théorème de Birkhoff (1946). La preuve que nous venons d'en donner, qui s'apparente à la classe des preuves dites *géométriques* est assez originale. En effet, le théorème y apparaît comme une conséquence de lemme 3.2 qui est un résultat de programmation linéaire.

### 3.1.3. Approximation par matrices bistochastiques

Soit un espace vectoriel normé (de matrices)  $(\mathbb{E}, \|\cdot\|)$  et  $\mathcal{S}$ , une partie de  $\mathbb{E}$ . Considérons dans cet espace le problème d'optimisation (matriciel) suivant : étant donné  $x \in \mathbb{E}$ , on cherche  $\bar{x}$  solution de

$$\begin{cases} \inf \|x - s\| \\ \text{t.q. } s \in \mathcal{S}. \end{cases}$$

C'est un problème d'approximation (matriciel) d'après Higham [10]. Il consiste à étudier l'existence et l'unicité de points  $\bar{x}$ , puis à en déterminer une caractérisation "explicite", ou le cas échéant un procédé algorithmique permettant de le calculer. Notre problème d'approximation par matrices bistochastiques est donc une instance pour laquelle on se place dans l'espace de Hilbert  $\mathbb{E} = \mathcal{M}_n(\mathbb{R})$  et on prend  $\mathcal{S} = \mathbb{B}_n$ .

Puisque  $\mathbb{B}_n$  est un convexe compact de  $\mathcal{M}_n(\mathbb{R})$ , on a (voir Lem. 2.1) :

**Proposition 3.4.** *Soit  $M \in \mathcal{M}_n(\mathbb{R})$ .*

*Il existe une et une seule matrice bistochastique  $\overline{\overline{M}}$  solution de (1), caractérisée par :*

$$\begin{cases} \overline{\overline{M}} \in \mathbb{B}_n; \\ \text{tr}((M - \overline{\overline{M}})^T (S - \overline{\overline{M}})) \leq 0, \quad \forall S \in \mathbb{B}_n. \end{cases} \quad (5)$$

D'après le Théorème de Birkhoff (Lem. 3.3), la caractérisation (5) est équivalente à la suivante :

$$\begin{cases} \overline{\overline{M}} \in \mathbb{B}_n; \\ \text{tr}((M - \overline{\overline{M}})^T (P - \overline{\overline{M}})) \leq 0, \quad \text{pour toute matrice } P \text{ de permutation.} \end{cases} \quad (6)$$

Si l'on voulait expliciter  $\overline{\overline{M}}$  en utilisant la caractérisation (6), on serait amené à résoudre un système d'équations ou inéquations, comportant au moins  $n!$  inéquations. Il est facile d'en conclure que cette caractérisation a toutes les chances de ne pas nous permettre de calculer "*explicitement*"  $\overline{\overline{M}}$ . Et ceci, même pour des

petites valeurs de  $n$ . En effet, pour  $n = 2$ , le problème se ramène à : trouver  $a \in [0, 1]$  tel que

$$\overline{M} = \begin{pmatrix} a & 1-a \\ 1-a & a \end{pmatrix} \text{ et } \text{tr}((M - \overline{M})^T(P - \overline{M})) \leq 0 \text{ pour } P = I_2, P_1 \quad (7)$$

qui n'est pas forcément "facile" à résoudre. Nous reviendrons sur ce cas  $n = 2$  un peu plus loin pour en donner une solution "explicite".

Manifestement en tout cas, l'approche directe ne semble pas la plus à même de fournir une caractérisation explicite de la solution du problème (1). Nous allons donc chercher une solution algorithmique au problème, en utilisant les méthodes de projection que nous introduisons ci-après.

### 3.2. LA MÉTHODE DE PROJECTIONS ALTERNÉES DE BOYLE-DIKSTRA

Les méthodes de projections, motivées par la résolution de problèmes de faisabilité convexe (voir [1, 3, 6]), permettent en général de trouver **un** point appartenant à une intersection de convexes. Lorsque l'on cherche en fait un point particulier, comme c'est le cas ici (nous cherchons un projeté), elles peuvent s'avérer inefficaces. Néanmoins, l'algorithme de projections alternées de Boyle-Dykstra a la particularité de calculer le projeté d'un point sur une intersection de convexes quelconques. Introduit par Dykstra en 1983 dans le cas où les convexes sont des cônes, étendu avec Boyle [4] en 1986 à des convexes quelconques, cet algorithme a été étudié par Bauschke et Borwein [2] et utilisé par Glunt *et al.* [8, 9], Escalante [7], Higham [11].

Considérons dans un espace de Hilbert  $(\mathbb{H}, \langle \cdot, \cdot \rangle)$ , deux convexes fermés  $\mathcal{A}$  et  $\mathcal{B}$ . Notons  $\mathcal{P}_{\mathcal{A}}$  et  $\mathcal{P}_{\mathcal{B}}$  les opérateurs de projections respectivement sur  $\mathcal{A}$  et  $\mathcal{B}$ . L'algorithme de von Neumann consiste en le schéma : on construit quatre suites :  $(a_n)$ ,  $(b_n)$  (appelées *suites principales*) et  $(p_n)$ ,  $(q_n)$  (appelées *suites auxiliaires*) comme suit :

$$\left\{ \begin{array}{l} a_0 = 0 ; b_0 = x \in \mathbb{H} ; p_0 = 0 ; q_0 = 0 ; \\ \mathbf{a}_{n+1} = \mathcal{P}_{\mathcal{A}}(\mathbf{b}_n + \mathbf{p}_n) \\ p_{n+1} = b_n + p_n - a_{n+1} \\ \mathbf{b}_{n+1} = \mathcal{P}_{\mathcal{B}}(\mathbf{a}_{n+1} + \mathbf{q}_n) \\ q_{n+1} = a_{n+1} + q_n - b_{n+1} \\ \text{avec } b_0 = x \in \mathbb{H} \text{ et } a_0 = 0. \end{array} \right. \quad (8)$$

Notons le calcul à chaque itération des vecteurs  $p_n$  et  $q_n$ . Ceux-ci sont calculés après projection sur chaque convexe et représentent, d'un point de vue géométrique, le *déplacement* effectué pour aller du nouvel itéré au point dont cet itéré est le projeté. On rappelle que ce vecteur appartient au cône normal au convexe ( $\mathcal{A}$  ou  $\mathcal{B}$ ) sur lequel on a projeté, au point résultat de la projection.

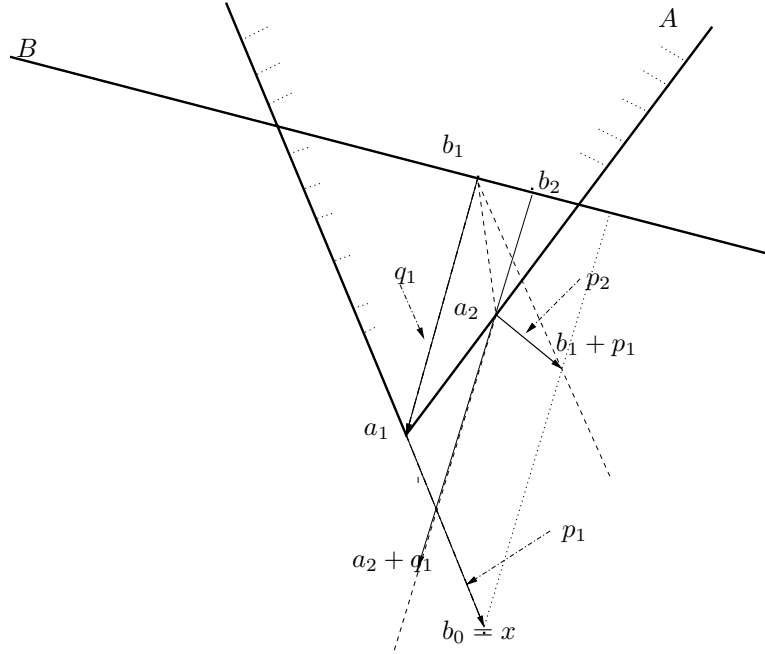


FIGURE 1. Illustration de l'algorithme de Boyle-Dykstra  
– Illustration of the Boyle-Dykstra's algorithm.

La figure 1 ci-dessus donne une illustration de l'algorithme de Boyle-Dykstra. Une itération de l'algorithme (par exemple, celle qui permet de passer de  $b_1$  à  $a_2$ ) peut être décrite de la manière suivante :

- on déplace le point courant (par exemple  $b_1$  sur la figure) dans la dernière direction normale ( $p_1$ ) au convexe sur lequel on doit projeter ( $\mathcal{A}$ ) gardée en mémoire ;
- on effectue la projection (sur  $\mathcal{A}$ ) du point obtenu ( $b_1 + p_1$ ) ;
- on garde en mémoire la nouvelle direction normale ( $p_2$ ) obtenue ainsi que le résultat de la projection ( $a_2$ ) qui est le nouvel itéré courant.

En 1994, Borwein et Bauschke [2] ont proposé une superbe analyse de cette méthode de projections alternées dans le cas de deux convexes. Leur résultat le plus important du point de vue de notre travail ici est le suivant :

**Théorème 3.5** [2]. *Soient  $\mathbb{H}$  un espace de Hilbert,  $\mathcal{A}$ ,  $\mathcal{B}$  deux convexes fermés de  $\mathbb{H}$  et  $x$  un point de  $\mathbb{H}$ . On définit les suites de Dykstra de la même manière qu'en (8).*

*Si  $\mathcal{A} \cap \mathcal{B} \neq \emptyset$ , alors*

$$\|b_n - a_n\|, \|b_n - a_{n+1}\| \rightarrow 0 \text{ et } a_n, b_n \rightarrow \mathcal{P}_{\mathcal{A} \cap \mathcal{B}}(x). \quad (9)$$

On pourra se reporter à [2] pour une preuve de ce résultat, où il est exprimé sous une forme plus générale.

Ces résultats sont intéressants pour nous puisque d'une part, le second justifie l'usage d'un algorithme de Boyle-Dykstra pour le calcul du projeté sur une intersection de convexes. D'autre part, le premier permet la mise en œuvre d'un test d'arrêt efficace lors de l'implémentation numérique de l'algorithme. Notons ensuite que l'algorithme de Boyle-Dykstra a une vitesse de convergence sous-linéaire. En fait, il a été montré (voir [2]) que l'algorithme converge au mieux linéairement, la vitesse linéaire n'étant atteinte que lorsqu'on a une intersection de sous-espaces.

Pour terminer, signalons que lorsque  $\mathcal{A}$  et/ou  $\mathcal{B}$  est un sous-espace,  $\mathcal{P}_{\mathcal{A}}$  et/ou  $\mathcal{P}_{\mathcal{B}}$  est linéaire (ou affine) et le calcul des  $p_n$  et  $q_n$  est inutile. Ainsi, l'algorithme de Boyle-Dykstra apparaît alors comme une généralisation de l'algorithme bien connu de projections alternées de Von-Neumann.

#### 4. APPROXIMATION PAR MATRICES BISTOCHASTIQUES PAR PROJECTIONS ALTERNÉES

Nous revenons à notre problème d'approximation par des matrices bistochastiques. Il est facile de voir que

$$\mathbb{B}_n = \mathcal{M}_n(\mathbb{R})^+ \cap \mathcal{LC1},$$

où  $\mathcal{M}_n(\mathbb{R})^+ = \{M \in \mathcal{M}_n(\mathbb{R}) \mid M \geq 0\}$  et  $\mathcal{LC1} = \{M \in \mathcal{M}_n(\mathbb{R}) \mid Me = e, M^T e = e\}$ . On remarque aussi que  $\mathcal{M}_n(\mathbb{R})^+$  et  $\mathcal{LC1}$  sont des ensembles convexes fermés : le premier est un cône convexe et le second un sous-espace affine. Cette écriture de  $\mathbb{B}_n$  nous permet alors d'appliquer une méthode de type Boyle-Dykstra à la résolution de notre problème d'approximation, *via* le calcul des projections  $\mathcal{P}_{\mathcal{M}_n(\mathbb{R})^+}$  et  $\mathcal{P}_{\mathcal{LC1}}$ .

La première de ces projections est triviale :

$$\forall A \in \mathcal{M}_n(\mathbb{R}), \mathcal{P}_{\mathcal{M}_n(\mathbb{R})^+}(A) = A^+$$

où  $A^+ = (\max(a_{ij}, 0))_{i,j}$  si  $A = (a_{ij})_{i,j}$ .

##### 4.1. PROJECTION SUR $\mathcal{LC1}$

**Définition 4.1.** Une matrice  $M$  est dite **bistochastique généralisée** ou **lc1** si elle vérifie :

- (1)  $Me = e$ ;
- (2)  $M^T e = e$ .

Les matrices bistochastiques généralisées forment le sous-espace affine  $\mathcal{LC1}$  que nous avons introduit précédemment. On est toujours placé dans l'espace de Hilbert  $(\mathcal{M}_n(\mathbb{R}), \langle \langle \cdot, \cdot \rangle \rangle)$ . Déterminons la projection  $\mathcal{P}_{\mathcal{LC1}}$ .



Puisque  $\mathcal{LC1}$  est un sous-espace affine et fermé (immédiat),  $\overline{M} = \mathcal{P}_{\mathcal{LC1}}(M)$  est caractérisé par (voir Lem. 2) :

- $\overline{M} \in \mathcal{LC1}$ ;
- $M - \overline{M} \in \overrightarrow{\mathcal{LC1}}^\perp$ .

Soit l'application linéaire  $l : M \in \mathcal{M}_n(\mathbb{R}) \mapsto (Me, M^T e)$ . On a :  $\mathcal{LC1} = l^{-1}(\{(e, e)\})$ , d'où  $\overrightarrow{\mathcal{LC1}} = \ker(l)$  et

$$\overrightarrow{\mathcal{LC1}}^\perp = (\ker(l))^\perp = \text{im}(l^T).$$

Or  $l^T$ , application linéaire de  $\mathbb{R}^n \times \mathbb{R}^n$  dans  $\mathcal{M}_n(\mathbb{R})$ , vérifie :

$$\forall (u, v) \in \mathbb{R}^n \times \mathbb{R}^n, \forall M \in \mathcal{M}_n(\mathbb{R}) \quad \langle l^T(u, v), M \rangle = \langle (u, v), l(M) \rangle_{n \times n}$$

où  $\langle \cdot, \cdot \rangle_{n \times n}$  désigne le produit scalaire de  $\mathbb{R}^n \times \mathbb{R}^n$  :  $\langle (u, v), (u', v') \rangle_{n \times n} = u^T u' + v^T v'$ . Par suite :

$$\begin{aligned} \langle l^T(u, v), M \rangle &= \langle (u, v), (Me, M^T e) \rangle_{n \times n}, \\ &= \langle u, Me \rangle + \langle v, M^T e \rangle, \\ &= \langle \langle ue^T, M \rangle \rangle + \langle \langle ve^T, M^T \rangle \rangle, \\ &= \langle \langle ue^T, M \rangle \rangle + \langle \langle ev^T, M \rangle \rangle, \\ &= \langle \langle ue^T + ev^T, M \rangle \rangle. \end{aligned}$$

D'où,

$$\forall (u, v) \in \mathbb{R}^n \times \mathbb{R}^n, \quad l^T(u, v) = ue^T + ev^T.$$

**Proposition 4.2.** *On a :*

$$\overrightarrow{\mathcal{LC1}}^\perp = \{ue^T + ev^T; u \in \mathbb{R}^n, v \in \mathbb{R}^n\}.$$

On en déduit que  $\overline{M} = \mathcal{P}_{\mathcal{LC1}}(M)$  vérifie :

$$\begin{cases} \overline{M}e = e, \\ \overline{M}^T e = e, \\ M - \overline{M} = ue^T + ev^T, u, v \in \mathbb{R}^n. \end{cases} \quad (10)$$

Par suite,

$$M - \overline{M} = ue^T + ev^T \Rightarrow \overline{M} = M - ue^T - ev^T. \quad (11)$$

Cette dernière relation injectée dans les deux premières équations de (10) conduit à :

$$\begin{cases} nu + ev^T e = Me - e, \\ eu^T e + nv = M^T e - e. \end{cases} \quad (12)$$

En appliquant une méthode de substitution à ce système linéaire,  $u$  et  $v$  sont solutions de

$$(nI_n - ee^T)u = Me - J_n Me, \quad (13)$$

et

$$(nI_n - ee^T)v = M^T e - J_n M^T e. \quad (14)$$

Les vecteurs  $u$  et  $v$  sont donc solutions de systèmes linéaires qui ne diffèrent que par leurs seconds membres. Notons  $E_n$  la matrice commune aux systèmes (13) et (14). Elle s'écrit :

$$E_n = \begin{pmatrix} n-1 & -1 & \dots & -1 \\ -1 & n-1 & \dots & -1 \\ \vdots & \ddots & \ddots & \vdots \\ -1 & \dots & -1 & n-1 \end{pmatrix}.$$

**Proposition 4.3.** *La matrice  $E_n$  est de rang  $n-1$ . Son noyau est l'espace de dimension 1 engendré par le vecteur  $e$ .*

Le calcul du rang est élémentaire. Puis, il suffit de remarquer que

$$E_n e = (nI_n - ee^T)e = nI_n e - e(e^T e) = ne - ne = 0.$$

En observant que :

$$E_n \left( \frac{1}{n} M e \right) = M e - J_n M e,$$

on obtient l'ensemble  $S_{13}$  des solutions de (13)

$$S_{13} = \left\{ \frac{1}{n} M e + k e, k \in \mathbb{R} \right\},$$

et de manière analogue,

$$S_{14} = \left\{ \frac{1}{n} M^T e + k' e, k' \in \mathbb{R} \right\}.$$

Ainsi si  $u$  et  $v$  résolvent (12),

$$u = \frac{1}{n} M e + k e \text{ et } v = \frac{1}{n} M^T e + k' e, k, k' \in \mathbb{R}^n.$$

En réinjectant ces informations dans (12), on aboutit à :

$$(k + k')e = -\frac{1}{n}(I_n + J_n M)e, \text{ soit } k + k' = -\frac{1}{n^2}e^T(I_n + J_n M)e,$$

où on a posé :  $J_n = \frac{1}{n}ee^T$ . Donc  $u$  et  $v$  sont déterminés par :

$$\begin{cases} u = \frac{1}{n} M e + k e, \\ v = \frac{1}{n} M^T e + k' e, \\ (k + k')e = -\frac{1}{n}(I_n + J_n M)e. \end{cases} \quad (15)$$

Si on pose  $W_n = I_n - J_n$ , à partir de (11) en utilisant (15), on obtient :

$$\overline{M} = W_n M W_n + J_n. \quad (16)$$

Réciproquement, on a bien :

**Proposition 4.4.**  $\overline{M}$  vérifie la relation de caractérisation de la proposition 2, d'où :

$$\forall M \in \mathbf{E}, \quad \mathcal{P}_{\mathcal{LC1}}(M) = W_n M W_n + J_n. \quad (17)$$

*Démonstration.*

- $\overline{M} \in \mathcal{LC1}$ .

Immédiat d'après  $I_n, J_n \in \mathcal{LC1}$ . On déduit :

- $M - \overline{M} \in \overrightarrow{\mathcal{LC1}}^\perp$ .

Soit  $B \in \mathcal{LC1}$ . On doit montrer que :  $\langle\langle M - \overline{M}, B \rangle\rangle = cte$ . Par définition,

$$\langle\langle M - \overline{M}, B \rangle\rangle = tr((M - \overline{M})^T B) = tr(B^T (M - \overline{M})).$$

On a :

$$B^T (M - \overline{M}) = B^T M - B^T W_n M W_n - J_n,$$

ce qui implique :

$$\langle\langle M - \overline{M}, B \rangle\rangle = tr(B^T M) - tr(B^T W_n M W_n) - 1.$$

Or :

$$\begin{aligned} W_n M W_n &= M - M J_n - J_n M + J_n M J_n ; \\ B^T (W_n M W_n) &= B^T M - B^T M J_n - J_n M + J_n M J_n, \quad \text{car } B^T J_n = J_n. \\ tr(B^T (W_n M W_n)) &= tr(B^T M) - tr(M J_n B^T) - tr(J_n M) + tr((J_n)^2 M), \\ &= tr(B^T M) - tr(M J_n). \end{aligned}$$

Ainsi,

$$\begin{aligned} \langle\langle M - \overline{M}, B \rangle\rangle &= tr(B^T M) - tr(B^T M) + tr(M J_n) - 1, \\ &= tr(M J_n) - 1. \end{aligned}$$

Les matrices  $J_n$  et  $M$  étant fixées,  $tr(M J_n) - 1$  est une constante (indépendante de  $B$ ). Par suite, on a :

$$\langle\langle M - \overline{M}, B \rangle\rangle = cte, \quad \forall B \in \mathcal{LC1},$$

d'où

$$\forall B, B' \in \mathcal{LC1}, \langle\langle M - \overline{M}, B - B' \rangle\rangle = 0.$$

Donc  $M - \overline{M} \in \overrightarrow{\mathcal{LC1}}^\perp$ . □

Nous obtenons un résultat qui a été trouvé de deux manières différentes par Khoury [14] et Glunt *et al.* [9]. Khoury a utilisé une approche purement géométrique (en fait algébrique) tandis que Glunt *et al.* se sont placés dans un contexte d'optimisation convexe et attachés à la résolution du système de Karush-Kuhn-Tucker correspondant au problème d'optimisation.

#### 4.2. QUELQUES REMARQUES

Nous pouvons dès à présent énoncer quelques résultats sur notre problème d'approximation par des matrices bistochastiques en utilisant la projection sur  $\mathcal{LC1}$ .

Dans toute cette partie, nous notons, pour une matrice  $M$  donnée de  $\mathcal{M}_n(\mathbb{R})$ ,

$$\overline{M} = \mathcal{P}_{\mathcal{LC1}}(M) \text{ et } \overline{\overline{M}} = \mathcal{P}_{\mathbb{B}_n}(M).$$

Nous nous intéressons aux cas où  $\overline{M} = \overline{\overline{M}} = \mathcal{P}_{\mathbb{B}_n}$ . Puisque sur l'espace  $\mathcal{LC1}$ , les contraintes  $\overline{M}e = e$  et  $\overline{M}^T e = e$  sont déjà satisfaites, il reste en fait à s'assurer que  $\overline{M}$  a toutes ses composantes positives.

Tout d'abord, reprenons le cas  $n = 2$ . Dans ce cas,  $\mathcal{LC1}$  est la droite (dimension 1) passant par  $I_2$  et  $P_1$  (qui sont définies en section 3). Le problème se ramène alors à celui de projeter sur le segment  $[I_2; P_1]$ , quand l'on sait projeter sur la droite  $(I_2 P_1)$  sous-jacente. Ainsi,

**Proposition 4.5.** *Si  $n = 2$ , on a :*

$$\overline{\overline{M}} = \mathcal{P}_{\mathcal{LC1}}(M) = \begin{cases} \overline{M} = W_2 M W_2 + J_2 & \text{si } \overline{M} \geq 0, \\ I_2 & \text{si } \overline{M} \not\geq 0 \text{ et } \|I_2 - \overline{M}\|_F < \|P - \overline{M}\|_F, \\ P_1 & \text{si } \overline{M} \not\geq 0 \text{ et } \|I_2 - \overline{M}\|_F > \|P - \overline{M}\|_F. \end{cases}$$

Pour  $n = 2$ , la projection sur  $\mathcal{LC1}$  est donc explicite. Et, pour  $n$  quelconque, on a une forme explicite pour certaines matrices.

Notons  $E_{ij}$  les matrices de la base canonique de  $\mathcal{M}_n(\mathbb{R})$ .

On a alors :

**Proposition 4.6.** *Si  $M = 0_{\mathcal{M}_n(\mathbb{R})}$  ou  $M = E_{ij}$ ,  $i = 1, \dots, n$ ,  $j = 1, \dots, n$ , alors*

$$\overline{\overline{M}} = \overline{M} = W_n M W_n + J_n.$$

*Démonstration.* Le cas  $M = 0_{\mathcal{M}_n(\mathbb{R})}$  est immédiat. Pour  $M = E_{ij}$ , il nous suffit de montrer que pour  $i, j$  fixés,

$$\overline{E_{ij}} = W E_{ij} W + J_n \geq 0.$$

Nous allons tout simplement calculer explicitement les composantes de  $\overline{E_{ij}}$  et vérifier qu'elles sont toutes positives.

On pose :  $D = W E_{ij} W = (d_{kl})_{k,l}$ .

Par définition, on a :

$$\begin{aligned}
d_{kl} &= \sum_{u=1}^n \sum_{v=1}^n w_{ku} e_{uv} w_{vl}, \\
&= \sum_v w_{kk} e_{kv} w_{vl} + \sum_{u \neq k} \sum_v w_{ku} e_{uv} w_{vl}, \\
&= w_{kk} e_{kl} w_{ll} + \sum_{v \neq l} w_{kk} e_{kv} w_{vl} + \sum_{u \neq k} w_{ku} e_{ul} w_{ll} + \sum_{u \neq k} \sum_{v \neq l} w_{ku} e_{uv} w_{vl}, \\
&= \left(1 - \frac{1}{n}\right)^2 e_{kl} + \left(1 - \frac{1}{n}\right) \left(-\frac{1}{n}\right) \sum_{v \neq l} e_{kv} + \left(1 - \frac{1}{n}\right) \left(-\frac{1}{n}\right) \sum_{u \neq k} e_{ul} \\
&\quad + \frac{1}{n^2} \sum_{u \neq k} \sum_{v \neq l} e_{uv}, \\
&= \left(1 - \frac{1}{n}\right)^2 e_{kl} + \left(\frac{1}{n} - 1\right) \left(\frac{1}{n}\right) \left(\sum_{v \neq l} e_{kv} + \sum_{u \neq k} e_{ul}\right) + \frac{1}{n^2} \sum_{u \neq k} \sum_{v \neq l} e_{uv}.
\end{aligned}$$

On en déduit :

- si  $(k, l) = (i, j)$  alors  $d_{ij} = \left(1 - \frac{1}{n}\right)^2$  ;
- si  $k = i, l \neq j, d_{il} = \frac{1}{n} \left(\frac{1}{n} - 1\right) = \frac{1}{n^2} - \frac{1}{n}$  ;
- si  $k \neq i, l = j, d_{kj} = \frac{1}{n} \left(\frac{1}{n} - 1\right) = \frac{1}{n^2} - \frac{1}{n}$  ;
- si  $k \neq i, l \neq j, d_{kl} = \frac{1}{n^2}$ .

Comme  $J_n$  a toutes ses composantes égales à  $\frac{1}{n}$ , on a :  $\overline{E}_{ij} = (\overline{e}_{kl})_{k,l}$  tel que :

$$\begin{cases} \overline{e}_{ij} = \left(1 - \frac{1}{n}\right)^2 + \frac{1}{n}, \\ \overline{e}_{il} = \frac{1}{n^2} & l \neq j, \\ \overline{e}_{kj} = \frac{1}{n^2} & k \neq i, \\ \overline{e}_{kl} = \frac{1}{n^2} + \frac{1}{n} & k \neq i, l \neq j. \end{cases}$$

Il va de soi qu'on a :

$$\overline{E}_{ij} \geq 0.$$

D'où le résultat. □

Signalons qu'au passage, nous avons montré que pour  $M = (m_{ij})$ , on a  $\overline{M} = (\overline{m}_{ij})_{i,j}$  avec :

$$\overline{m}_{ij} = \left(1 - \frac{1}{n}\right)^2 m_{ij} + \frac{1}{n} \left(\frac{1}{n} - 1\right) \left(\sum_{k \neq i} m_{kj} + \sum_{l \neq j} m_{kl}\right) + \frac{1}{n^2} \sum_{k \neq i} \sum_{l \neq j} m_{kl} + \frac{1}{n}.$$

Plus généralement, on peut montrer

**Proposition 4.7.** *Soit  $M = (a_{ij}) \in E$  tel que  $a_{ij} \geq 0$ ,  $\forall i, j$  et  $\sum_{i,j} a_{ij} \leq 1$ .*

*Alors,*

$$\mathcal{P}_{\mathbb{B}_n}(M) = \mathcal{P}_{\mathcal{LC}_1}(M) = W_n M W_n + J_n.$$

*Démonstration.* La justification découle directement du lemme suivant dû à Zarantonello [19] :

**Lemme 4.8.** *Si  $\mathcal{P}$  est un opérateur de projection dans un Hilbert, alors :*

$$\left\| \mathcal{P}\left(\sum_1^k \alpha_i x_i\right) - \sum_1^k \alpha_i \mathcal{P}(x_i) \right\|^2 \leq \frac{1}{2} \sum_{i,j=1}^k \alpha_i \alpha_j \langle \mathcal{P}(x_i) - \mathcal{P}(x_j), (I - \mathcal{P})(x_i) - (I - \mathcal{P})(x_j) \rangle, \quad (18)$$

pour toutes familles finies  $\{x_i\}_i$  de vecteurs et  $\{\alpha_i\}_i$  de réels positifs tels que  $\sum_1^k \alpha_i = 1$ .

Il suffit d'appliquer (18) à la décomposition :

$$M = \sum_{ij} a_{ij} E_{ij} + (1 - \sum_{ij} a_{ij}) 0_{\mathcal{M}_n(\mathbb{R})}. \quad \square$$

### 4.3. TESTS NUMÉRIQUES

Nous avons donc proposé l'algorithme structuré comme suit pour résoudre le problème d'approximation par matrices bistochastiques :

**Initialisation**  $B^0 = M$

$Q^0 = 0$

Précision  $\varepsilon$

**Itération**

$A^{k+1} = W_n B^k W_n + J_n$  [=  $\mathcal{P}_{\mathcal{LC}_1}(B^k)$ ]

$B^{k+1} = (A^{k+1} + Q^k)^+$  [=  $\mathcal{P}_{\mathcal{M}_n(\mathbb{R})^+}(A^{k+1})$ ]

$Q^{k+1} = (A^{k+1} + Q^k) - (A^{k+1} + Q^k)^+$

**Test d'arrêt** si  $\|A^{k+1} - B^{k+1}\| < \varepsilon$  Stop

sinon retour à Itération

où  $M$  est la matrice que l'on cherche à approcher par une matrice bistochastique.

Cet algorithme est tout simplement une adaptation de l'algorithme (8) à notre cas. Nous l'avons écrit en tenant compte du fait que l'un de nos convexes est un sous-espace, et qu'il est donc inutile d'en calculer les composantes normales à chaque itération. Le test d'arrêt est basé sur le fait qu'on doit avoir  $\lim_{k \rightarrow +\infty} \|A^{k+1} - B^{k+1}\| = 0$  (voir Th. 3.5).

Nous avons testé l'algorithme pour différentes matrices. Nous avons obtenu les résultats exprimés par les figures suivantes. Du fait des applications que nous

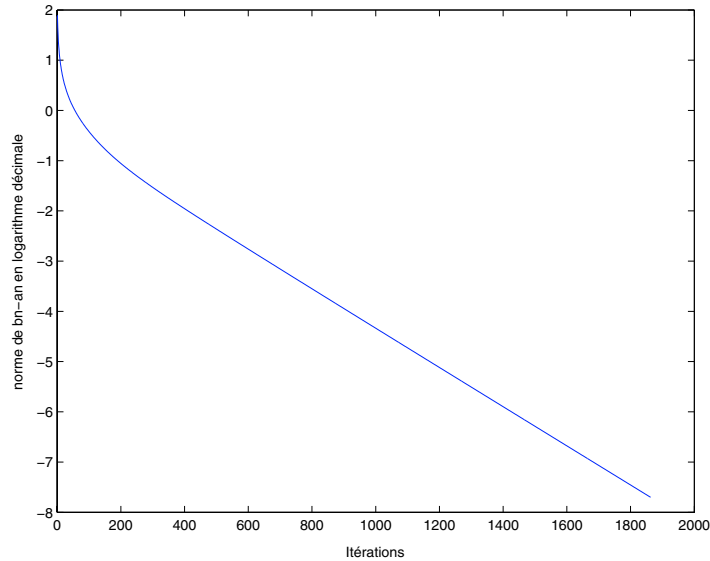


FIGURE 2.  $\log_{10} \|B^k - A^k\|$  en fonction des itérations pour une matrice aléatoire avec  $n = 200$  et  $\varepsilon = 10^{-10}$  –  $\log_{10} \|B^k - A^k\|$  vs. – Iterations for a random matrix with  $n = 200$  and  $\varepsilon = 10^{-10}$ .

avons en vue, nous avons considéré des matrices dont les composantes sont comprises entre 0 et 1 générées aléatoirement. En particulier, les quatre dernières figures sont réalisées avec des valeurs moyennes obtenues après 10 tests pour chaque cas.

Les figures 2 et 3 illustrent la convergence de  $\|A^{k+1} - B^{k+1}\|$  vers 0 pour une matrice  $M$  de dimension  $n = 200$ , pour lesquelles la précision demandée était  $10^{-10}$  et  $10^{-15}$ . On observe bien sur ces figures la convergence sous-linéaire de l'algorithme de Boyle-Dykstra. Notons que les temps de calculs sont de l'ordre de 387 s et 655 s respectivement.

Nous avons étudié ensuite le comportement de l'algorithme par rapport à la taille de la matrice que l'on veut approcher. La figure 4 représente l'évolution du nombre d'itérations en fonction de la dimension de la matrice, tandis que la figure 5 celle du temps de calcul en fonction également de la dimension.

Il apparaît, au vu des exemples que nous avons traités, que l'algorithme se comporte assez bien, et que le nombre d'itérations n'explose pas lorsqu'on augmente la taille de la matrice traitée. Notons qu'une itération de l'algorithme consiste en un remplacement des composantes négatives d'une matrice par zéro, et en quatre produits d'une matrice par la matrice  $J_n$ .

Nous avons ensuite étudié le comportement de l'algorithme pour les matrices creuses. Malheureusement, il semble que l'approximation par matrices bistochastiques ne conserve pas dans l'absolu le caractère creux de la matrice de départ. Ceci est probablement dû aux produits matriciels effectué à chaque projection

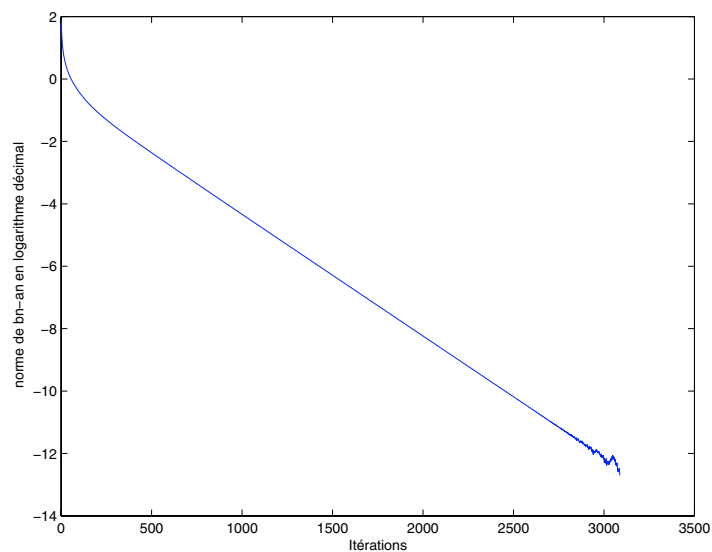


FIGURE 3.  $\log_{10} \|B^k - A^k\|$  en fonction des itérations pour une matrice aléatoire avec  $n = 200$  et  $\varepsilon = 10^{-20} - \log_{10} \|B^k - A^k\|$  vs. – Iterations for a random matrix with  $n = 200$  and  $\varepsilon = 10^{-15}$ .

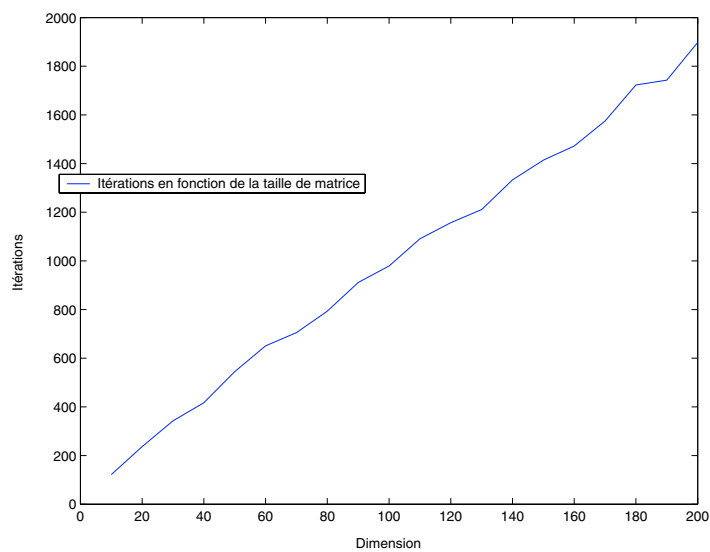


FIGURE 4. Itérations en fonction de la taille de la matrice – Iterations *versus* dimension of the matrix.



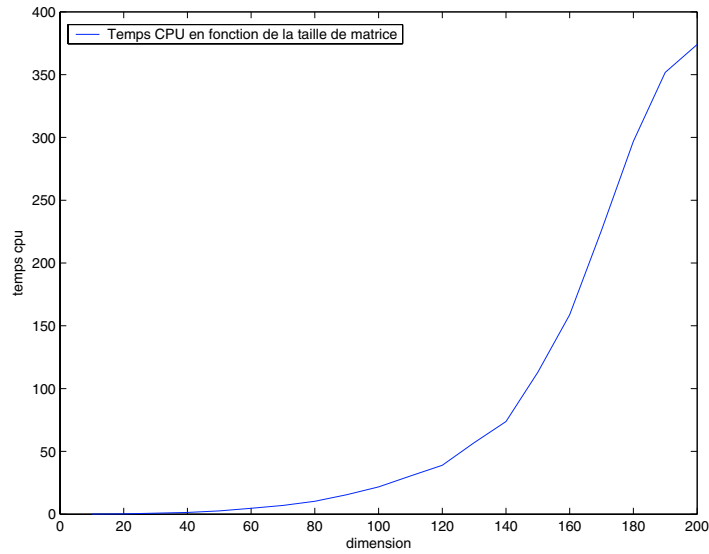


FIGURE 5. Temps CPU en fonction de la taille de la matrice – CPU time *versus* dimension of the matrix.

sur  $\mathcal{LC1}$ . Compte tenu de la proposition 4.6 sur la projection des matrices de la base canonique. Par exemple, on obtient à partir de la matrice  $E_{11}$  de dimension 4, la matrice solution

$$\frac{1}{16} \begin{pmatrix} 13 & 1 & 1 & 1 \\ 1 & 5 & 5 & 5 \\ 1 & 5 & 5 & 5 \\ 1 & 5 & 5 & 5 \end{pmatrix},$$

qui, contrairement à  $E_{11}$ , est dense. Les figures 6 et 7 représentent une illustration du comportement pour des matrices creuses d'ordre 100.

Pour ces tests, l'exploitation des matrices creuses intrinsèques à Matlab a été utilisée. Toutefois, l'exploitation du caractère creux de la matrice à projeter devrait être étudiée de plus près.

## 5. CONCLUSION ET PERSPECTIVES

Nous venons d'étudier le problème d'approximation par des matrices bistochastiques. Il ressort de cette étude que pour une matrice donnée  $M$ , il existe une et une seule matrice la plus proche de  $M$ . Cette matrice possède une caractérisation qui, malheureusement, ne peut permettre d'obtenir une formule "explicite" de cette matrice bistochastique, sauf dans certains cas particuliers. Cela étant, nous avons proposé une mise en œuvre d'un algorithme adapté de Boyle-Dijkstra qui permet de calculer cette matrice bistochastique itérativement.

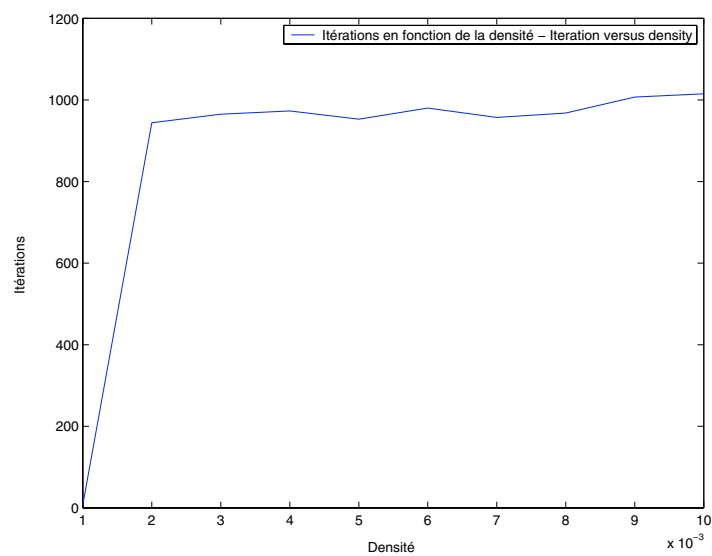


FIGURE 6. Itérations en fonction de la densité de la matrice – Iterations *versus* density of the matrix.

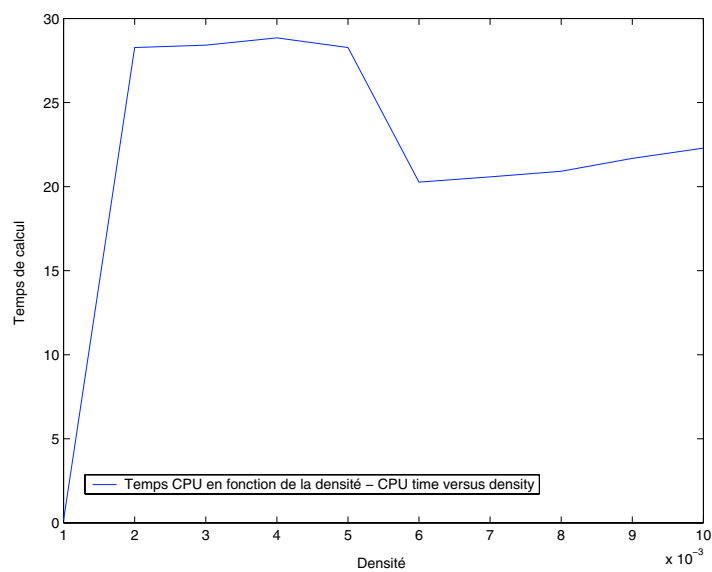


FIGURE 7. Temps CPU en fonction de la densité de la matrice – CPU time *versus* density of the matrix.

Cet algorithme a été utilisé dans le cadre d'une modélisation matricielle du problème d'agrégation de préférences, travail publié dans *Mathématiques et Sciences Humaines* [18].

Il existe de nombreuses perspectives à ce travail. Outre l'exploitation du caractère creux de la matrice à projeter, il sera intéressant d'étudier l'influence de la manière dont  $\mathbb{B}_n$  est écrite sous la forme d'intersection de convexes. Nous avons choisi ici une écriture intersection de deux convexes, il est possible d'avoir un écriture avec jusqu'à  $2n + n^2$  convexes, chaque écriture menant à un algorithme différent. De plus, il existe d'autres techniques (points intérieurs, relaxation lagrangienne, etc.) qui peuvent s'appliquer à ce problème d'approximation par matrices bistochastiques. Une comparaison de ces techniques serait d'un très grand intérêt.

*Remerciements.* Je voudrais remercier Jean-Baptiste Hiriart-Urruty et Marcel Mongeau (mes directeurs de thèse) pour les nombreuses discussions fructueuses que nous avons eues à propos de cette étude. Je voudrais associer à ces remerciements Combettes et Bauschke qui m'ont ouvert au monde des méthodes de projections.

## RÉFÉRENCES

- [1] H. Bauschke, *Projections Algorithms and Monotone Operators*. Ph.D. thesis, Simon Fraser University (1996).
- [2] H. Bauschke and J. Borwein, Dykstra's alternating projection algorithm for two sets. *J. Approx. Theory* **79** (1994) 418–443.
- [3] H. Bauschke and J. Borwein, On projection algorithms for solving convex feasibility problems. *SIAM Rev.* **38** (1996) 367–426.
- [4] J.P. Boyle and R.L. Dykstra, A method for finding projections onto the intersection of convex sets in Hilbert spaces, in *Advances in Order Restricted Statistical Inference*, edited by R.L. Dykstra, T. Robertson and F.T. Wright. Springer-Verlag. *Lect. Notes Statist.* (1985) 28–47.
- [5] H. Brezis, *Analyse fonctionnelle. Théories et Applications*. Masson (1983).
- [6] P. Combettes, Hilbertian convex feasibility problem: Convergence of projection methods. *Appl. Math. Optim.* **35** (1997) 311–330.
- [7] R. Escalante, Dykstra's algorithm for a constrained least-squares matrix problem. *Numerical Linear Algebra Appl.* **3** (1996) 459–471.
- [8] W. Glunt, T. Hayden, S. Hong and J. Wells, An alternating projection algorithm for computing the nearest Euclidian distance matrix. *SIAM J. Matrix Anal. Appl.* **11** (1990) 589–600.
- [9] W. Glunt, T. Hayden and R. Reams, The nearest “doubly stochastic” matrix to a real matrix with the same first moment. *Numer. Linear Algebra Appl.* **5** (1998) 475–482.
- [10] N.J. Higham, Matrix nearness problems and applications. In *Applications of Matrix Theory*, edited by M.J.C. Gover and S. Barnett. Oxford University Press (1989) 1–27.
- [11] N.J. Higham, Computing the nearest correlation matrix – a problem from finance. *IMA J. Numer. Anal.* **22** (2002) 329–343.
- [12] J.-B. Hiriart-Urruty and C. Lemaréchal, *Convex analysis and minimization algorithms*. Grundlehren der mathematischen Wissenschaften 305 & 306. Springer-Verlag, Berlin, Heidelberg (1993). (New printing in 1996).
- [13] R.B. Horn and C.R. Johnson, *Matrix Analysis*. Cambridge University Press (1985). (Reprinted in 1991, 1992).
- [14] R.N. Khoury, Closest matrices in the space of generalized doubly stochastic matrices. *J. Math. Anal. Appl.* **222** (1998) 562–568.

- [15] R. Rockafeller and R.J.-B. Wets, *Variational Analysis*. Grundlehren der mathematischen Wissenschaften 317. Springer-Verlag, Berlin, Heidelberg (1998).
- [16] P. Takouda, *Un problème d'approximation matricielle : quelle est la matrice bistochastique la plus proche d'une matrice donnée ?* Technical report, Laboratoire MIP, Université Paul Sabatier, Toulouse 3, 2002. Research Report MIP 02–21. Accessible on the web at the url : [http ://mip.ups-tlse.fr/publi/2002.html](http://mip.ups-tlse.fr/publi/2002.html). Submitted.
- [17] P. Takouda, *Problèmes d'approximations matricielle linéaire conique : approches par projections et via optimisation sous contraintes de semi-définie positivité*. Ph.D. thesis, Université Paul Sabatier - Toulouse III (Septembre 2003).
- [18] P. Takouda, Résolution d'un problème d'agrégation de préférence en approximant par des matrices bistochastiques. *Mathématiques et Sciences Humaines, "Recherche opérationnelle et aide à la décision"*, 41<sup>e</sup> année **161** (2003) 77–97. Aussi rapport interne numéro 03–08, du laboratoire MIP de l'Université Paul Sabatier, Toulouse.
- [19] E. Zarantonello, Projections on convex sets in Hilbert spaces and spectral theory, in *Contributions to Nonlinear Functionnal Analysis*, edited by E.E. Zarantonello, number 27 in University of Wisconsin. Mathematics Research Center Publications, Academic Press, New york (1971) 1–38. *Proceeding on the special session on Optimization and Nonlinear Analysis*, Jerusalem (May 1995).