

Estimation de l'intensité conditionnelle d'un processus ponctuel

Auteurs:

Alexandre Lecestre

Responsable:

Édith Gabriel

Avignon, mars - juillet 2019

INTRODUCTION

Ce rapport présente le travail que j’ai réalisé en stage de mars à juillet 2019 à Avignon. Lors de ce stage, j’étais encadré par Édith Gabriel, Jérôme Coville et Joël Chadœuf au sein du Laboratoire de Mathématiques d’Avignon (LMA) et de l’unité de recherche “Biostatistique et processus spatiaux” de l’INRA (BioSP). Le sujet du stage était l’estimation de l’intensité conditionnelle d’un processus ponctuel, dans la continuité des résultats de Gabriel *et al.* [2].

La statistique spatiale et notamment la théorie des processus ponctuels est largement utilisée pour modéliser des phénomènes réels. On a des exemples d’applications dans différents domaines comme la cosmologie [6], l’écologie [4], l’épidémiologie [1], la criminologie [9], etc. Concrètement, la récolte de données peut-être fastidieuse ou coûteuse en pratique ce qui ne permet pas toujours d’observer la zone d’intérêt dans sa totalité mais uniquement sur une zone restreinte. L’objectif est alors de prédire l’intensité du processus ponctuel sur le reste du domaine d’intérêt, conditionnellement à sa réalisation dans la zone observée.

La méthode développée dans [2] utilise le BLUP (*Best Linear Unbiased Predictor*), qui est un estimateur classique en géostatistique appelé estimateur de krigeage, dans le cas d’un processus stationnaire. Pour calculer le BLUP, on assigne un “poids” à chaque point observé dans la fenêtre d’observation et on prend la somme de ces poids comme estimation. Construire le BLUP revient à trouver la fonction de poids qui donne donc le meilleur estimateur linéaire non-biaisé. Pour cela on va d’abord déterminer la condition nécessaire pour avoir un estimateur sans biais, puis utiliser la méthode des multiplicateurs de Lagrange pour déterminer la fonction de poids qui minimise la variance de notre estimateur (parmi les estimateurs sans biais). La fonction de poids associée au BLUP est alors solution d’une équation intégrale de Fredholm. Il n’existe pas de formule analytique pour la résolution de ces équations dans le cas général. La difficulté pratique est donc d’automatiser la résolution de cette équation pour avoir une méthode applicable de manière simple, notamment par des non-spécialistes sur des données réelles.

Après quelques notions de théorie des processus ponctuels, je présente l’essentiel du travail réalisé durant mon stage. J’ai d’abord appliqué la démarche évoquée précédemment au cas de processus marqués (marque continue et multitype). Puis, j’ai adapté la construction du BLUP pour utiliser des notions d’angles lors de l’estimation. En pratique, on estime en mettant des poids sur l’ensemble des couples de points parmi les points observés. Pour chaque couple de points, on a un angle formé avec le point où on fait la prédiction ce qui permet de prendre compte des alignements qui apparaissent naturellement sur des données réelles.

I—THÉORIE DES PROCESSUS PONCTUELS

Toute cette partie s’inspire directement du bouquin de Chiu, Stoyan, Kendall et Mecke [7]. On se restreint ici à des notions de base ou des notions qui seront utilisés dans la suite du rapport. On peut se référer au livre pour un aperçu plus complet et détaillé de la théorie des processus ponctuels.

1 Définitions et notations

On donne dans cette section la définition formelle d’un processus ponctuel et le vocabulaire de base utilisé dans la littérature.

1.1 Cadre général

Définition 1. *Processus ponctuel sur \mathbb{R}^d*

Un processus ponctuel sur \mathbb{R}^d est une variable aléatoire à valeurs dans un espace mesurable $[\mathbb{N}, \mathcal{N}]$, où \mathbb{N} est la classe de toutes les suites φ de points de \mathbb{R}^d , et qui vérifie les propriétés suivantes :

- *la suite φ est localement finie,*
- *la suite est simple, i.e. $i \neq j \Rightarrow x_i \neq x_j$.*

La σ -algèbre \mathcal{N} est la plus petite σ -algèbre sur \mathbb{N} telle que la fonction $\varphi \mapsto \varphi(B)$ est mesurable pour tout borélien borné B de \mathbb{R}^d , où $\varphi(B)$ désigne le nombre de points de φ appartenant à B .

Définition 2. *Ensemble de configuration*

On appelle ensemble de configuration tout élément Y de la σ -algèbre \mathcal{N} .

On notera P la mesure de probabilités sur $[\mathbb{N}, \mathcal{N}]$ définie par le processus ponctuel ϕ .

Définition 3. *Lois de dimension finie*

Les lois de dimension finie sont les probabilités de la forme

$$\mathbb{P}(\phi(B_1) = n_1, \dots, \phi(B_k) = n_k)$$

où B_1, \dots, B_k sont des boréliens et n_1, \dots, n_k des entiers positifs.

La loi de ϕ est entièrement déterminée par la famille des lois de dimension finie pour tout $k = 1, 2, \dots$. On peut même se contenter de la famille des probabilités de vide (sur l'ensemble des compacts) pour caractériser un processus ponctuel.

Définition 4. *Probabilités de vide*

La probabilité de vide du borélien B est définie par

$$v_B = P(\{\varphi \in \mathbb{N} : \varphi(B) = 0\}) = \mathbb{P}(\phi(B) = 0).$$

Par définition, les notations suivantes sont équivalentes pour toute fonction mesurable positive f ,

$$\mathbb{E} \left[\sum_{x \in \phi} f(x) \right] = \int_{\mathbb{N}} \sum_{x \in \varphi} f(x) P(d\varphi) = \int_{\mathbb{N}} \int_{\mathbb{R}^d} f(x) \varphi(dx) P(d\varphi). \quad (1)$$

Définition 5. *Stationnarité*

On dit que ϕ est stationnaire si pour tout $Y \in \mathcal{N}$ et $x \in \mathbb{R}^d$,

$$\mathbb{P}(\phi \in Y) = \mathbb{P}(\phi_x \in Y),$$

où $\phi_x = \{x_n + x\}$. Autrement dit, ϕ et ϕ_x ont même loi.

Définition 6. *Isotropie*

On dit que ϕ est isotrope si pour toute rotation \mathbf{r} ,

$$\mathbb{P}(\phi \in Y) = \mathbb{P}(\mathbf{r}\phi \in Y).$$

Si un processus est stationnaire et isotrope, on dit qu'il est invariant par déplacement (isométrie directe).

Définition 7. *Mesure d'intensité*

La mesure d'intensité Λ du processus ponctuel ϕ est la mesure positive définie par $\Lambda(B) = \mathbb{E}[\phi(B)]$, pour tout borélien B . De plus, si Λ est absolument continue par rapport à la mesure de Lebesgue v_d , alors $\lambda = \frac{d\Lambda}{dv_d}$ est appelé densité d'intensité de ϕ et on a donc

$$\Lambda(B) = \int_B \lambda(x) dx.$$

La mesure d'intensité permet de simplifier les calculs par rapport à (1) grâce au théorème de Campbell.

Théorème 1. *Campbell (1909)*

Soit un processus ponctuel ϕ et une fonction mesurable positive f . Alors, on a

$$\mathbb{E} \left[\sum_{x \in \phi} f(x) \right] = \int f(x) \Lambda(dx).$$

Définition 8. Mesures de moments

On définit la mesure de n -ième moment $\mu^{(n)}$ de ϕ par

$$\int_{\mathbb{R}^{nd}} f(x_1, \dots, x_n) \mu^{(n)}(d(x_1, \dots, x_n)) = \mathbb{E} \left[\sum_{x_1, \dots, x_n \in \phi} f(x_1, \dots, x_n) \right].$$

On a $\mu^{(n)}(B_1 \times \dots \times B_n) = \mathbb{E}[\phi(B_1 \times \dots \times B_n)]$ et en particulier

$$\mu^{(n)}(B^n) = \mathbb{E}[\phi(B)^n],$$

d'où le terme “mesure de n -ième moment”. De même on définit $\alpha^{(n)}$, la mesure de moment factoriel d'ordre n , par

$$\int f(x_1, \dots, x_n) \alpha^{(n)}(d(x_1, \dots, x_n)) = \mathbb{E} \left[\sum_{x_1, \dots, x_n \in \phi}^{\neq} f(x_1, \dots, x_n) \right]$$

où \sum^{\neq} indique la somme prise sur tous les n -uplets de points distincts, i.e. on ne prend pas en compte ceux avec deux membres égaux ou plus. En particulier, on a

$$\alpha^{(n)}(B^n) = \mathbb{E}[\phi(B)(\phi(B) - 1) \dots (\phi(B) - n + 1)],$$

d'où le terme “moment factoriel d'ordre n ”.

Définition 9. Densités produits

Si $\alpha^{(n)}$ est localement finie et absolument continue par rapport à la mesure de Lebesgue v_{nd} , on note $\varrho^{(n)} = \frac{d\alpha^{(n)}}{dv_{nd}}$ la densité produit d'ordre n telle que

$$\mathbb{E} \left[\sum_{x_1, \dots, x_n}^{\neq} f(x_1, \dots, x_n) \right] = \int \dots \int f(x_1, \dots, x_n) \varrho^{(n)}(x_1, \dots, x_n) dx_1 \dots dx_n.$$

On peut remarquer que si elle existe, la densité produit $\varrho^{(n)}$ est invariante par permutation, i.e.

$$\varrho^{(n)}(x_1, \dots, x_n) = \varrho^{(n)}(x_{\sigma(1)}, \dots, x_{\sigma(n)})$$

pour toute permutation σ de $\{1, \dots, n\}$ et $x_1, \dots, x_n \in \mathbb{R}^d$.

Définition 10. Fonctions de corrélation

La mesure $\alpha^{(n)}$ est toujours dominée par $\Lambda^{\otimes n}$. Si de plus Λ est dominée par ν_d , alors on définit g_n , la fonction de corrélation d'ordre n , par

$$g_n = \frac{d\alpha^{(n)}}{d\Lambda^{(n)}}$$

et telle que

$$\begin{aligned} \alpha^{(n)}(d(x_1, \dots, x_n)) &= \varrho^{(n)}(x_1, \dots, x_n) dx_1 \dots dx_n \\ &= g_n(x_1, \dots, x_n) \lambda(x_1) \dots \lambda(x_n) dx_1 \dots dx_n. \end{aligned}$$

1.2 Cas stationnaire

Si on considère un processus ponctuel ϕ stationnaire, alors on peut adopter des notations plus simples que dans le cas général. En particulier, l'invariance par translation donne

$$g_n(x_1, \dots, x_n) = g_n(0, x_2 - x_1, \dots, x_n - x_1).$$

On peut donc toujours considérer que le point x_1 est l'origine. On note en pratique $g_2(x_2 - x_1)$. De plus, si ϕ est isotrope, seul la distance entre x_1 et x_2 importe donc on rencontre très souvent la notation $g_2(r)$ pour les modèles invariants par déplacement.

1.3 Processus marqué

Un processus ponctuel marqué sur \mathbb{R}^d est une suite aléatoire $\Psi = \{[x_n; m_n]\}$ où l'ensemble des points x_n constitue un processus ponctuel dans \mathbb{R}^d et m_n est la marque correspondant au point x_n . On appellera “processus de terrain” (pour *ground process*) le processus ponctuel $\phi_g = \{x_n\}$. Les marques m_n appartiennent à un ensemble donné (espace des marques) \mathbb{M} qu'on supposera être un espace Polonais. La tribu des boréliens de \mathbb{M} est noté \mathcal{M} . On redéfinit la translation et la rotation d'un processus par rapport au cas non-marqué. Le processus translaté de $\Psi = \{[x_n, m_n]\}$ par $x \in \mathbb{R}^d$ est donné par

$$\Psi_x = \{[x_n + x, m_n]\}.$$

La translation agit uniquement sur le processus de terrain. De même, la rotation de Ψ par \mathbf{r} est définie par

$$\mathbf{r}\Psi = \{[\mathbf{r}x_n, m_n]\}.$$

Définition 11. *Intensité d'un processus ponctuel marqué*
 Pour $B \in \mathcal{B}^d$ et $L \in \mathcal{M}$, on définit l'intensité Λ par

$$\Lambda(B \times L) = \mathbb{E}[\Psi(B \times L)],$$

$$\text{où } \Psi(B \times L) = \sum_{[x, m] \in \Psi} \mathbb{1}_B(x) \mathbb{1}_L(m).$$

On a aussi une version du théorème de Campbell pour les processus marqués. On a,

$$\mathbb{E} \left[\sum_{[x, m] \in \phi} f(x, m) \right] = \int f(x, m) \Lambda(dx, dm),$$

pour toute fonction mesurable positive f sur $\mathbb{R}^d \times \mathbb{M}$.

Pour $L \in \mathcal{M}$ fixé, la mesure $\Lambda(\cdot \times L)$ est absolument continue par rapport à la

mesure d'intensité Λ_g du processus de terrain. Par conséquent, on peut montrer que

$$\Lambda(d(x,m)) = M_x(dm)\Lambda_g(dx),$$

où M_x est une loi de probabilité $[\mathbb{M}, \mathcal{M}]$, qu'on peut voir comme la distribution des marques en x .

Parmi les cas les plus fréquents de processus marqués on retrouve les processus multitypes, où l'espace des marques est fini, et les processus à marque continue, où l'espace des marques est un intervalle de \mathbb{R} . Les processus à marque continue peuvent notamment être utilisés pour modéliser des phénomènes spatio-temporels. Il existe quelques notions spécifiques aux processus multitypes.

Considérons $\mathbb{M} = \{1, \dots, M\}$. On peut alors interpréter le processus ϕ comme M processus en interaction. On définit les fonctions de corrélation de paire croisées de la manière suivante. Soit

$$\mu_{k_1, k_2}^{(2)}(B_1 \times B_2) = \mathbb{E}[\phi_{k_1}(B_1)\phi_{k_2}(B_2)],$$

pour $k_1 \neq k_2$. On sait que $\mu_{k_1, k_2}^{(2)}$ est absolument continue par rapport à $\Lambda_{k_1} \otimes \Lambda_{k_2}$ donc il existe une fonction g_{k_1, k_2} , la dérivée de Radon-Nikodym de $\mu_{k_1, k_2}^{(2)}$ par rapport à $\Lambda_{k_1} \otimes \Lambda_{k_2}$, telle que

$$\mu_{k_1, k_2}^{(2)}(B_1 \times B_2) = \int_{B_1 \times B_2} g_{k_1, k_2}(x_1, x_2) \Lambda_{k_1}(dx_1) \Lambda_{k_2}(dx_2).$$

Si Λ_{k_1} et Λ_{k_2} sont absolument continues par rapport à la mesure de Lebesgue, on a

$$\mathbb{E}[\phi_{k_1}(B_1)\phi_{k_2}(B_2)] = \int_{B_1 \times B_2} g_{k_1, k_2}(x_1, x_2) \lambda_{k_1}(x_1) \lambda_{k_2}(x_2) dx_1 dx_2.$$

Par convention, on notera également $g_{k, k}$ la fonction de corrélation de paire du processus ϕ_k .

2 Modèles

Un avantage de la méthode utilisée par Gabriel *et al.* et qu'elle ne nécessite pas de spécifier de modèles. On va tout de même en exposer quelques uns ici, qu'on utilisera pour simuler des processus ponctuels et pour avoir plusieurs exemples en tête.

2.1 Processus de Poisson

Un processus ponctuel de Poisson homogène ϕ , d'intensité λ , est caractérisé par deux propriétés fondamentales.

- Le nombre de points contenu dans un borélien B suit une loi de Poisson (de paramètre $\lambda v_d(B)$), i.e.

$$\mathbb{P}(\phi(B) = m) = \frac{\mu^m}{m!} \exp(-\mu) \text{ où } \mu = \lambda v_d(B).$$

- La répartition des points est indépendante, i.e. si B_1, \dots, B_k sont des boréliens disjoints, alors $\phi(B_1), \dots, \phi(B_k)$ sont des variables aléatoires indépendantes.

Un processus de Poisson homogène est stationnaire et isotrope. Cependant, le modèle est généralisable au cas inhomogène. Pour une mesure de Radon Λ , sans atome sur \mathbb{R}^d , un processus de Poisson ϕ de mesure d'intensité Λ est un processus ponctuel qui vérifie les propriétés suivantes.

- Le nombre de points contenu dans un borélien B suit une loi de Poisson, i.e.

$$\mathbb{P}(\phi(B) = n) = \frac{\Lambda(B)^n}{n!} \exp(-\Lambda(B)).$$

- La répartition des points est indépendante de la même manière que dans le cas homogène.

On peut remarquer que la répartition indépendante des points donne $\alpha^{(n)}(B_1 \times \dots \times B_k) = \Lambda(B_1) \dots \Lambda(B_k)$ pour tout processus de Poisson, homogène et inhomogène. Par conséquent, on a $g_2(x_1, x_2) = 1$ pour tout processus de Poisson.

2.2 Processus de Neyman-Scott

Un processus de Neyman-Scott est un processus “construit” en deux étapes. Un premier processus ponctuel donne les “points parents”, puis les points parents induisent la réalisation de processus ponctuels indépendants et de même loi qui donnent les “points enfants”. Chaque point parent donne naissance à un cluster. Le processus de Neyman-Scott est l'ensemble des points enfants sans les points parents.

- Les centres des clusters (appelés points parents) sont issus d'un processus de Poisson homogène d'intensité λ_p .
- Les points enfants d'un cluster représentatif N_0 sont aléatoires en nombre et répartis de manière indépendante selon la même loi de densité $f(y)$ autour de l'origine.
- Les points parents n'apparaissent pas dans le nuage de points observé, seulement les points enfants.

Un tel processus est stationnaire, car le processus parent l'est. Il a pour intensité

$$\lambda = \lambda_p \bar{c},$$

où

$$\bar{c} = \sum_{n=0}^{\infty} n p_n,$$

et

$$p_n = \mathbb{P} \left(N_0(\mathbb{R}^d) = n \right).$$

On note P la loi de ϕ et \mathbf{c}_0 la loi ('de Palm') du cluster N_0 . Alors

$$P_0 = P * \mathbf{c}_0,$$

où la loi de Palm \mathbf{c}_0 est donnée par

$$\mathbf{c}_0(Y) = \frac{1}{\bar{c}} \mathbb{E} \left[\sum_{x \in N_0} \mathbb{1}_Y(N_0 - x) \right], \forall Y \in \mathcal{N}.$$

Il existe énormément de modèles utilisés pour modéliser une grande variété de phénomènes concrets. Il serait donc inutile de vouloir en faire une liste exhaustive. Les exemples précédents permettent juste de mieux comprendre les processus simulés pour réaliser des prédictions dans la suite.

II – ESTIMATION DE L'INTENSITÉ CONDITIONNELLE POUR DES PROCESSUS MARQUÉS

La problématique qui nous intéresse est l'estimation de l'intensité d'un processus en un point où il n'a pas été observé conditionnellement à sa réalisation dans une certaine fenêtre observation. Plus formellement, on note

- $W \in \mathcal{B}(\mathbb{R}^d)$ la fenêtre d'observation,
- ϕ un processus ponctuel dans \mathbb{R}^d
- et ν la mesure de Lebesgue sur \mathbb{R}^d .

Dans toute cette partie, on suppose que ϕ admet une fonction d'intensité λ et une fonction de corrélation de paire g_2 , toutes les deux continues.

Dans [2], Gabriel *et al.* considèrent un processus ϕ stationnaire et estiment son intensité en un point $x_0 \notin W$ conditionnellement à son observation sur W . Cette quantité qu'on appelle "intensité conditionnelle" est définie par

$$\lambda(x_0|\phi_W) = \lim_{\nu(B) \rightarrow 0} \frac{\mathbb{E}[\phi(B \oplus x_0)|\phi_W]}{\nu(B)}$$

où B est une surface élémentaire autour de 0. Pour cela, ils utilisent le BLUP qui est de la forme suivante

$$\hat{\lambda}(x_0|\phi_W) = \sum_{x \in \phi_W} w(x; x_0)$$

où $x_0 \notin W$ et $\phi_W = \phi \cap W$. On veut alors que la fonction de poids w vérifie une condition nécessaire et suffisante pour que l'estimateur $\hat{\lambda}$ soit sans biais, et on veut également que w minimise l'erreur quadratique de $\hat{\lambda}$ parmi les estimateurs linéaires sans biais.

Dans le cas stationnaire (Proposition 2.1 [2]), la fonction de poids w du BLUP satisfait la condition $\int_W w(x; x_0) dx = 1$ et vérifie l'équation intégrale

$$w(x; x_0) + \int_W w(y; x_0) k(x, y) dy = f(x; x_0)$$

avec

$$k(x, y) = \lambda \left(g(x - y) - \frac{1}{\nu(W)} \int_W g(x - y) dx \right)$$

et

$$f(x; x_0) = \frac{1}{\nu(W)} + \lambda \left(g(x - x_0) - \frac{1}{\nu(W)} \int_W g(x - x_0) dx \right).$$

La fonction de poids w est donc uniquement définie de manière implicite par cette équation qui fait intervenir l'intensité λ et la fonction de corrélation de paire g . En pratique, il faut bien souvent estimer ces quantités. Il existe déjà des fonctions implémentées dans R qui permettent de calculer des estimateurs empiriques et également d'inférer des modèles paramétriques à partir d'un estimateur empirique. On ne s'intéressera donc pas directement à ce point dans la suite et on considérera que les différentes quantités relatives au processus observé sont connues.

On peut également remarquer qu'il faut résoudre autant d'équations que de points x_0 auxquels on veut prédire l'intensité conditionnelle et on ne sait pas calculer en temps raisonnable une solution exacte de cette équation. Il faut donc recourir à des méthodes numériques de résolution approchée comme proposé dans [2]. Dans mon cas, je n'ai utilisé uniquement la méthode des éléments finis pour obtenir la fonction de poids de mon estimateur.

Dans la suite, les résultats importants sont mis en évidence tandis que le détail des preuves et les différents calculs se trouvent en annexe. La démarche est toujours la même ; obtenir la condition pour avoir un estimateur sans biais, puis minimiser l'erreur quadratique sous la contrainte de non-biais avec la méthode des multiplicateurs de Lagrange et finalement résoudre de manière approchée l'équation intégrale obtenue avec la méthode des éléments finis.

Dans cette section on va appliquer cela dans le cas de processus ponctuels marqués $\phi = \{[x_1, m_1], [x_2, m_2], \dots\}$ dans $\mathbb{R}^d \times \mathbb{M}$. On considère d'abord des processus à marque continue ($\mathbb{M} = \mathbb{R}$) et dans un second temps des processus multitypes ($\mathbb{M} = \{1, 2, \dots, M\}, M \in \mathbb{N}$).

3 Processus à marque continue

On considère un processus marqué ϕ dont l'espace des marques est continue $\mathbb{M} = \mathbb{R}$. Soit W un borélien borné de \mathbb{R}^d . On note $\phi_W = \{[x, m] \in \phi : x \in W\}$ et on note Λ_g l'intensité du "ground process" : le processus ponctuel $\phi_g = \{x_i : \exists m, [x, m] \in \phi\}$. En pratique on appliquera la méthode sur des points observés donc on fait l'hypothèse naturelle $\Lambda_g(W) > 0$. De plus, on supposera que le processus marqué admet une fonction d'intensité $\lambda(x, m)$ et une fonction de corrélation de paire $g(x, m, y, n)$ toutes deux continues.

3.1 Équation intégrale

Proposition 1. *Pour $x_0 \notin W$, le prédicteur*

$$\hat{\lambda}(x_0, m_0 | \phi_W) = \sum_{[x, m] \in \phi_W} w(x, m; x_0, m_0)$$

est le BLUP de $\lambda(x_0, m_0 | \phi_W)$ et la fonction de poids w est solution de l'équation

$$\lambda(x, m) \left\{ w(x, m; x_0, m_0) + \int_{\mathbb{R} \times W} w(y, n; x_0, m_0) K(x, m, y, n) dndy - f(x, m; x_0, m_0) \right\} = 0 \quad (2)$$

où

$$K(x, m, y, n) = g(x, m, y, n) - \frac{1}{\Lambda_g(W)} \int \lambda(x, m) g(x, m, y, n) dmdx$$

et

$$f(x, m; x_0, m_0) = \lambda(x_0, m_0) \left[g(x, m, x_0, m_0) + \frac{1}{\Lambda_g(W)} - \frac{1}{\Lambda_g(W)} \int_{\mathbb{R} \times W} \lambda(x, m) g(x, m, x_0, m_0) dmdx \right].$$

La preuve se trouve en annexe en section A.1.

3.2 Résolution par la méthode des éléments finis

On ne peut appliquer la méthodes des éléments finis uniquement sur un domaine borné or on considère jusqu'ici que l'espace des marques $\mathbb{M} = \mathbb{R}$ ne l'est pas. En pratique, on observe nécessairement des marques bornées si on se place sur un domaine borné. Il faut donc se restreindre à un intervalle fermé pour les marques afin d'appliquer la méthode des éléments finis. Autrement, c'est comme dans le cas non-marqué.

On note V_h l'espace d'approximation qui est un sous-espace de dimension finie de $L^2(W \times \mathbb{M})$. On note $N = \dim(V_h)$ et $\{\phi_i\}_{i=1, \dots, N}$ une base de V_h . On projette les différentes fonctions sur l'espace d'approximation :

$$w(x, m) \approx \sum_{i=1}^N \omega_i \phi_i(x, m)$$

et

$$K(x, m, x', m') \approx \sum_{1 \leq i, j \leq N} K_{i,j} \phi_i(x, m) \phi_j(x', m').$$

On définit

- la matrice de masses par

$$M_{i,j} = \Lambda(\mathbb{1}_{W \times \mathbb{M}} \phi_i \phi_j),$$

- le vecteur des poids par

$$\Omega_l = \omega_l$$

- et le vecteur F par

$$F_l = \int_{W \times \mathbb{M}} \lambda(x, m) \phi_l(x, m) f(x, m; x_0, m_0) dx dm.$$

On a alors l'équation matricielle

$$(M + MKM)\Omega = F$$

qui donne

$$\Omega = (M + MKM)^{-1} F.$$

Le détail du calcul se trouve en section B.1. On peut remarquer que la matrice $M + MKM$ ne dépend pas du point (x_0, m_0) où on fait la prédiction. On a donc une seule matrice de cette forme à stocker en mémoire et à inverser (ce qui est souvent limitant en temps et en mémoire). Une fois qu'on a prédit pour un point, ce n'est pas embêtant de le faire pour toute une zone de prédiction.

3.3 Discussion

L'aspect théorique du cas d'un processus ponctuel à marque continue est assez simple, assez semblable au cas stationnaire vu dans [2]. Cependant, c'est beaucoup plus compliqué de visualiser les données pour un processus à marque continue. Par conséquent, la représentation de la prédiction est très compliquée de par la dimension des données. J'ai donc décidé de ne pas inclure de figures ici, également faute de temps. En revanche, ce cas permet de bien voir comment la méthode fonctionne pour construire l'estimateur BLUP puisqu'on applique la même procédure dans le cas multitype et les différentes équations ont la même forme que dans le cas d'une marque continue.

Au delà de la représentation des résultats, le cas d'une marque continue fait apparaître quelques difficultés spécifiques. Il faut notamment considérer un espace des marques borné en pratique pour appliquer la méthode des éléments finis, donc il faudrait se pencher sur la façon de choisir l'espace des marques pour appliquer la méthode avec des données réelles. De plus, on peut également s'intéresser à la discrétisation de la fenêtre de prédiction au niveau des marques.

4 Processus multitype

Un processus multitype est un processus ponctuel marqué où l'ensemble des marques \mathbb{M} est un ensemble fini. On note dans cette partie $\mathbb{M} = \{1, \dots, M\}$. Cela revient au même que de considérer M processus ponctuels non marqués.

Soit M processus ponctuels ϕ_1, \dots, ϕ_M dans \mathbb{R}^d . Pour $k, k' \in \{1, \dots, M\}$, on suppose que ϕ_k admet

- une fonction d'intensité λ_k continue,
- une fonction de corrélation de paire g_k continue,

On suppose également que ϕ_k et $\phi_{k'}$ admettent une fonction de corrélation de paire croisée $g_{k,k'}$ pour tous $k, k' \in \{1, \dots, M\}$. Pour simplifier les notations, on notera $g_{k,k} = g_k$.

4.1 Équation intégrale

Pour $k \in \{1, \dots, M\}$, soit W_k un borélien borné de \mathbb{R}^d . Pour chaque processus ϕ_k , on observe une réalisation de ce processus dans la fenêtre W_k . Comme dans le cas précédent, on fait l'hypothèse naturelle $\Lambda_k(W_k) > 0, \forall k$. On estime l'intensité conditionnelle de ϕ_{k_0} au point $x_0 \notin W_{k_0}$. Le BLUP de $\lambda_{k_0}(x_0 | \phi_{1W_1}, \dots, \phi_{MW_M})$ est défini par la proposition suivante.

Proposition 2. *Pour $x_0 \notin W_{k_0}$, le prédicteur*

$$\hat{\lambda}_{k_0}(x_0 | \phi_{1W_1}, \dots, \phi_{MW_M}) = \sum_{k=1}^M \sum_{x \in \phi_{kW_k}} w_k(x; x_0)$$

est le BLUP de $\lambda_{k_0}(x_0 | \phi_{1W_1}, \dots, \phi_{MW_M})$. La fonction de poids $w(x, x_0)$ est solution du système d'équations

$$\lambda_k(x) \cdot \left\{ w_k(x) + \sum_{k'=1}^M \int_{W_{k'}} w_{k'}(y) \lambda_{k'}(y) K_{k,k'}(x, y) dy - f_k(x; x_0) \right\} = 0, \forall k \quad (3)$$

où

$$K_{k,k'}(x, y) = \left(g_{k,k'}(x, y) - \frac{1}{\Lambda_k(W_k)} \int_{W_k} \lambda_k(x') g_{k,k'}(x', y) dx' - \frac{\mathbb{1}_{k=k'}}{\Lambda_k(W_k)} \right),$$

et

$$f_k(x; x_0) = \lambda_{k_0}(x_0) \left(g_{k,k_0}(x, x_0) - \frac{1}{\Lambda_k(W_k)} \int_{W_k} \lambda_k(x') g_{k,k_0}(x', x_0) dx' \right),$$

et vérifie la condition

$$\begin{aligned}
& \frac{1}{\Lambda_{k_1}(W_{k_1})} \left[\lambda_{k_0}(x_0) \int_{W_{k_1}} \lambda_{k_1}(x) (g_{k_1, k_0}(x, x_0) - 1) dx \right. \\
& - \sum_{k'=1}^M \int_{W_{k_1} \times W_{k'}} w_{k'}(x') \lambda_{k_1}(x) \lambda_{k'}(x') (g_{k_1, k'}(x, x') - 1) dx dx' \\
& \left. - \int_W w_{k_1}(x) \lambda_{k_1}(x) dx \right] = \frac{1}{\Lambda_{k_2}(W_{k_2})} \left[\lambda_{k_0}(x_0) \int_{W_{k_2}} \lambda_{k_2}(x) (g_{k_2, k_0}(x, x_0) - 1) dx \right. \\
& - \sum_{k'=1}^M \int_{W_{k_2} \times W_{k'}} w_{k'}(x') \lambda_{k_2}(x) \lambda_{k'}(x') (g_{k_2, k'}(x, x') - 1) dx dx' \\
& \left. - \int_{W_{k_2}} w_{k_2}(x) \lambda_{k_2}(x) dx \right],
\end{aligned} \tag{4}$$

pour tous $k_1, k_2 \in \{1, \dots, M\}$.

La preuve se trouve en section A.2. On a ici un système de M équations intégrales. Cela n'empêche pas d'utiliser la méthode des éléments finis pour en obtenir une solution approchée.

4.2 Résolution par la méthode des éléments finis

Soit V_1, \dots, V_L les espaces d'approximation utilisés. Pour $k = 1, \dots, M$, V_k est un sous-espace de dimension-finie de $L^2(W_k)$. On note $N_k = \dim(V_k)$ et $\{\phi_{ik}\}_{i=1, \dots, N_k}$ est une base de V_k . On projette les différentes fonctions qui apparaissent dans l'équation sur les espaces V_k correspondant. Pour $k \in \{1, \dots, M\}$, on définit :

- le vecteur $\Omega_k = (\omega_{ik})_i$ par

$$w_k(x) \approx \sum_{j=1}^{N_k} \omega_{jk} \phi_{jk}(x),$$

- la matrice des masses (symétrique) M_k par

$$(M_k)_{ij} \approx \Lambda_k(\mathbb{1}_{W_k} \phi_{ik} \phi_{jk}),$$

- la matrice $\kappa^{k, k'}$ par

$$K_{k, k'}(x, y) \approx \sum_{i_1=1}^{N_k} \sum_{i_2=1}^{N_{k'}} \kappa_{i_1, i_2}^{k, k'} \phi_{i_1 k}(x) \phi_{i_2 k'}(y),$$

- le vecteur F_k par

$$(F_k)_l = \int_{W_k} \lambda_k(x) \phi_{lk}(x) f_k(x; x_0) dx.$$

On obtient à partir de (3) les équations matricielles

$$M_k \Omega_k + \sum_{k'=1}^M M_k \kappa^{k,k'} M_{k'} \Omega_{k'} = F_k, \forall k \in \{1, \dots, M\}.$$

On peut alors réécrire les M équations matricielles en une seule équation matricielle en considérant des matrices par blocs, i.e.

$$AL \times \Omega = F$$

avec

$$AL = \begin{pmatrix} M_1 + M_1 \kappa^{1,1} M_1 & M_1 \kappa^{1,2} M_2 & \dots & M_1 \kappa M_M \\ M_2 \kappa^{2,1} M_1 & M_2 + M_2 \kappa^{2,2} M_2 & \dots & M_2 \kappa^{2,M} M_M \\ \vdots & \vdots & \ddots & \vdots \\ M_M \kappa^{M,1} M_1 & M_M \kappa^{M,2} M_2 & \dots & M_M + M_M \kappa^{M,M} M_M \end{pmatrix},$$

$$F = \begin{pmatrix} F_1 \\ \vdots \\ F_M \end{pmatrix} \text{ et } \Omega = \begin{pmatrix} \Omega_1 \\ \vdots \\ \Omega_M \end{pmatrix}.$$

Le calcul détaillé se trouve en section B.2.

4.3 Résultats sur données simulées

On a simulé un processus de Cox multitype dirigé par un processus booléen de disques. En pratique,

- on simule 3 processus de Poisson inhomogènes $\phi_{0,i}, i = 1, 2, 2$ indépendants d'intensités $\lambda_{0,i}(x), i = 1, 2, 2$,
- on simule un processus booléen de disques ϕ_D tel que les centres des disques est un processus de Poisson homogène d'intensité λ_b et les disques ont un rayon R_b ,
- on réalise un thinning, les points de $\phi_{0,i}$ qui se retrouvent à l'intérieur d'un disque de ϕ_W sont retirés avec probabilité $1 - p_i$.

On peut alors calculer les intensités et les fonctions de corrélations de paire (croisées ou non) du processus de Cox multitype ainsi obtenu. Les intensités respectives des processus finaux sont

$$\lambda_i(x) = \lambda_{0,i}(x) \left(e^{-\lambda_b \pi R_b^2} + (1 - e^{-\lambda_b \pi R_b^2}) p_i \right), i = 1, 2, 3$$

et les fonctions de corrélation de paire sont données par

$$g_{i,j}(r) = \frac{A + B(p_1 + p_2) + (1 - A - 2B)p_1 p_2}{(e^{-\lambda_b \pi R_b^2} + (1 - e^{-\lambda_b \pi R_b^2}) p_i)(e^{-\lambda_b \pi R_b^2} + (1 - e^{-\lambda_b \pi R_b^2}) p_j)}, i, j = 1, 2, 3$$

où $A = e^{-\lambda_b S_r}$, $B = (1 - e^{-\lambda_b(\pi R_b^2 - s_r)})e^{-\lambda_b \pi R_b^2}$, et S_r est la surface de l'union de deux cercles de rayon R_b de centres séparés d'une distance r , et s_r est la surface de l'intersection de deux cercles de rayon R_b de centres séparés de r .

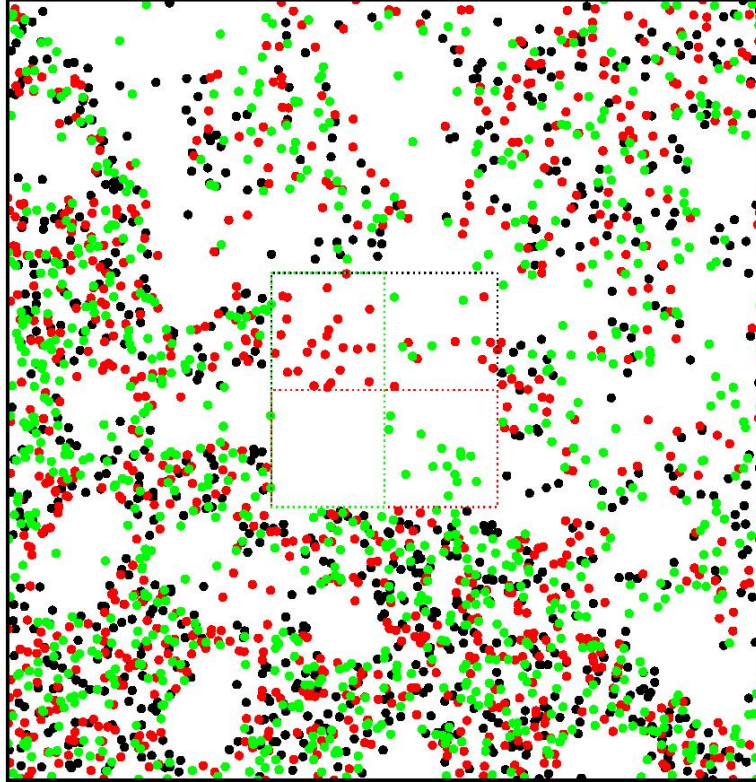
Les quantités S_r et s_r sont reliés par la formule $S_r = 2\pi R_b^2 - s_r$ et elles sont données par

$$s_r = \begin{cases} 2R_b^2 \left[\cos^{-1} \left(\frac{r}{2R_b} \right) - \left(1 - \frac{r}{2R_b} \right) \frac{r}{2R_b} \right] & r \leq 2R_b, \\ 0 & r > 2R_b. \end{cases}$$

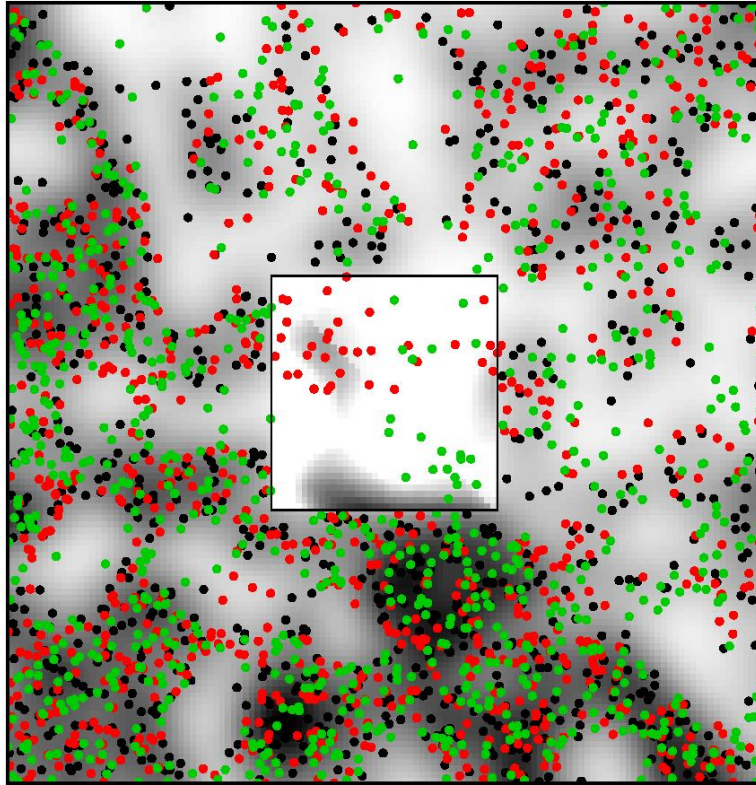
et

$$S_r = \begin{cases} 2R_b^2 \left[\cos^{-1} \left(-\frac{r}{2R_b} \right) + \left(1 - \frac{r}{2R_b} \right) \frac{r}{2R_b} \right] & r \leq 2R_b, \\ 2\pi R_b^2 & r > 2R_b. \end{cases}$$

Pour le graphique obtenu ci-dessous, on a pris la même fonction d'intensité pour les trois processus initialement qui est de la forme $\lambda_{0,i}(x) = \lambda_- \mathbb{1}_{x+y \leq A} + \lambda_+ \mathbb{1}_{x+y > A}$ où A est le côté du carré sur le graphique. Cette fonction d'intensité ne prend que deux valeurs possibles, la plus grande sur la moitié inférieure gauche de la zone d'observation ($\lambda_- > \lambda_+$).



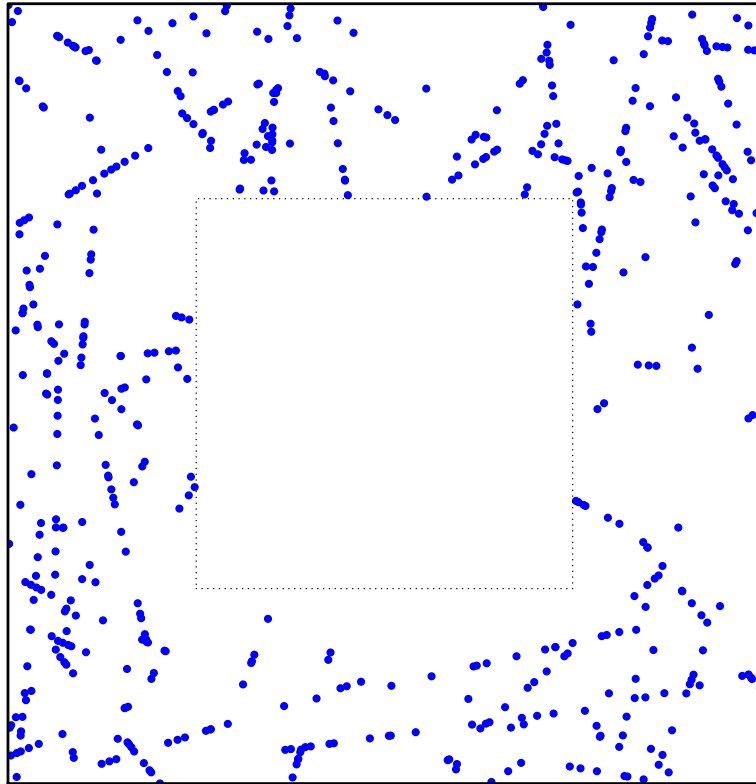
Pour la prédiction, on observe les trois processus dans la zone délimitée par le carré extérieur et le carré noir intérieur en pointillé. De plus, on observe le processus vert et le processus rouge respectivement sur les moitiés droite et haute du carré intérieur. On applique alors la méthode pour prédire l'intensité conditionnelle du processus noir sur l'ensemble du carré intérieur.



Même si le résultat n'est pas extrêmement convaincant, on peut tout de même observer que un des objectifs de la méthode est atteint. On arrive à tirer de l'information à partir de l'observation des processus rouge et vert pour prédire l'intensité conditionnelle du processus noir. C'est assez flagrant dans la partie supérieure gauche du carré intérieure avec le groupe de points du processus rouge.

III—CAS DIRECTIONNEL

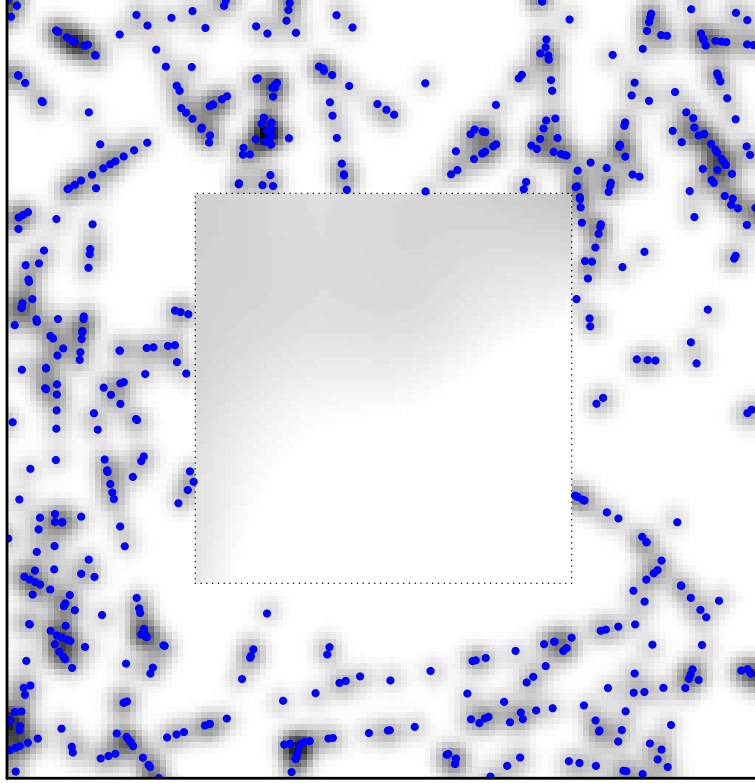
L'estimateur BLUP est déterminé entièrement par la fenêtre d'observation W , la fonction d'intensité λ et la fonction de corrélation de paires g_2 . L'estimateur n'est pas performant pour des processus dirigés pas des fibres par exemple comme c'est le cas sur le graphique ci-dessous.



Le nuage de points ci-dessus a été obtenu à partir d'une simulation. On a d'abord considéré un processus booléen de segments : les milieux des segments correspondent à un processus de Poisson homogène et les directions des segments sont indépendantes et identiquement distribuées selon une loi uniforme. Les segments sont tous de même longueur. Conditionnellement à ce processus de segments, on considère des processus de Poisson homogène indépendants et identiquement distribués sur chaque segment. La figure ci-dessus est une réalisation du processus

ponctuel ainsi défini.

On a alors un processus ponctuel stationnaire et on peut donc appliquer l'estimateur de Gabriel *et al.* [2]. On obtient alors le résultat suivant.



On est capable en regardant le nuage de points de voir que les points sont répartis le long de segments mais on voit aussi que l'estimateur ne le détecte pas et donne donc une prédiction plutôt médiocre qui tient uniquement compte de la densité de points dans les différentes zones mais pas de leur alignement. En effet, on a des poids qui ne dépendent uniquement du point x_0 où on prédit et du point x auquel on associe le poids $w(x; x_0)$ or on ne peut pas avoir de notion d'angle, donc d'alignement, avec uniquement un couple de points.

5 Définition du nouvel estimateur

Pour tenir compte d'interactions d'ordre supérieur à 2 pour “capter” les alignements qu'on peut voir ci-dessus par exemple, notre nouvel estimateur est de la

forme

$$\hat{\lambda}(x_0|\phi_W) = \sum_{x_1, x_2 \in \phi_W}^{\neq} w(x_1, x_2; x_0).$$

Comme avant on veut obtenir le meilleur estimateur sans biais. Donc on détermine la condition pour avoir un estimateur non biaisé et à nouveau on fait de l'optimisation sous contrainte pour minimiser l'erreur quadratique de notre estimateur. On notera l'estimateur BBUP dans la suite (pour *Best Bilinear Unbiased Predictor*).

5.1 Équation intégrale

On observe à nouveau un processus ϕ sur une fenêtre W , borélien borné de \mathbb{R}^d . Maintenant, on suppose de plus que ϕ admet des fonctions de corrélations d'ordre 3 et 4 continues. Autrement dit, on suppose que $\alpha^{(3)}$ et $\alpha^{(4)}$ sont absolument continues respectivement par rapport à $\Lambda^{\otimes 3}$ et $\Lambda^{\otimes 4}$ où Λ est la mesure d'intensité du processus, elle-même absolument continue par rapport à la mesure de Lebesgue sur \mathbb{R}^d . On note g_3 et g_4 les densités

$$\alpha^{(3)}(d(x_1, x_2, x_3)) = g_3(x_1, x_2, x_3) \lambda(x_1) \lambda(x_2) \lambda(x_3) dx_1 dx_2 dx_3$$

et

$$\alpha^{(4)}(d(x_1, x_2, x_3, x_4)) = g_4(x_1, x_2, x_3, x_4) \lambda(x_1) \lambda(x_2) \lambda(x_3) \lambda(x_4) dx_1 dx_2 dx_3 dx_4.$$

Ces quantités interviennent dans le calcul de l'erreur quadratique comme on le voit dans la proposition suivante.

Proposition 3. *Pour $x_0 \notin W$, le prédicteur*

$$\hat{\lambda}(x_0|\phi_W) = \sum_{x_1, x_2 \in \phi_W}^{\neq} w(x_1, x_2; x_0)$$

est le BBUP de $\lambda(x_0|\phi_W)$. On définit

$$p(x_1, x_2) = w(x_1, x_2; x_0) + w(x_2, x_1; x_0) \quad (5)$$

et cette nouvelle fonction de poids p est solution de l'équation

$$\begin{aligned} \lambda(x_1) \lambda(x_2) & \left[p(x_1, x_2) g_2(x_1, x_2) \right. \\ & + \int_W \lambda(x_3) (p(x_1, x_3) + p(x_2, x_3)) K_3(x_1, x_2, x_3) dx_3 \\ & \left. + \int_{W^2} \lambda(x_3) \lambda(x_4) p(x_3, x_4) K_4(x_1, x_2, x_3, x_4) dx_3 dx_4 - f(x_1, x_2; x_0) \right] = 0 \end{aligned} \quad (6)$$

où

$$K_3(x_1, x_2, x_3) = g_3(x_1, x_2, x_3) - \frac{g_2(x_1, x_2)}{\alpha^{(2)}(W)} \int_{W^2} \lambda(y_1) \lambda(y_2) g_3(y_1, y_2, x_3) dy_1 dy_2,$$

$$K_4(x_1, x_2, x_3, x_4) = \frac{1}{2} \left(g_4(x_1, x_2, x_3, x_4) - \frac{g_2(x_1, x_2)}{\alpha^{(2)}(W)} \int_{W^2} \lambda(y_1) \lambda(y_2) g_4(y_1, y_2, x_3, x_4) dy_1 dy_2 \right),$$

et

$$f(x_1, x_2; x_0) = \lambda(x_0) \left(g_3(x_1, x_2, x_0) + 2 \frac{g_2(x_1, x_2)}{\alpha^{(2)}(W^2)} - \frac{g_2(x_1, x_2)}{\alpha^{(2)}(W^2)} \int_{W^2} \lambda(x_1) \lambda(x_2) g_3(y_1, y_2, x_0) dy_1 dy_2 \right).$$

En pratique on va résoudre l'équation pour obtenir p et non w . Finalement, cela n'a pas d'importance puisqu'on a $\hat{\lambda}(x_0 | \phi_W) = \sum_{x_1, x_2 \in \phi \cap W}^{\neq} w(x_1, x_2; x_0) = \frac{1}{2} \sum_{x_1, x_2 \in \phi \cap W}^{\neq} p(x_1, x_2)$ et donc toute fonction w qui vérifie (5) donne le BBUP. On prendra alors tout simplement $w(x_1, x_2; x_0) = \frac{1}{2} p(x_1, x_2)$.

5.2 Méthode des éléments finis

Soit V_h un espace d'approximation, sous-espace de dimension finie de $L^2(W)$. On note $N = \dim(V_h)$ and $\{\phi_i\}_{i=1, \dots, N}$ une base de V_h . On projette sur cette espace différentes fonctions intervenant dans l'équation. On définit

- la matrice $P \in (\mathbb{R}^N)^{\otimes 2}$ telle que

$$p(x_1, x_2) \approx \sum_{1 \leq i_1, i_2 \leq N} P_{i_1, i_2} \phi_{i_1}(x_1) \phi_{i_2}(x_2),$$

- la matrice $G \in (\mathbb{R}^N)^{\otimes 2}$ telle que

$$g_2(x_1, x_2) \approx \sum_{1 \leq i_1, i_2 \leq N} G_{i_1, i_2} \phi_{i_1}(x_1) \phi_{i_2}(x_2) = G((\phi_j(x_1))_{1 \leq j \leq N}, (\phi_j(x_2))_{1 \leq j \leq N}),$$

- le tenseur $T \in (\mathbb{R}^N)^{\otimes 3}$ tel que

$$\begin{aligned} K_3(x_1, x_2, x_3) &\approx \sum_{1 \leq i_1, i_2, i_3 \leq N} T_{i_1, i_2, i_3} \phi_{i_1}(x_1) \phi_{i_2}(x_2) \phi_{i_3}(x_3) \\ &= T((\phi_j(x_1))_{1 \leq j \leq N}, (\phi_j(x_2))_{1 \leq j \leq N}, (\phi_j(x_3))_{1 \leq j \leq N}), \end{aligned}$$

- le tenseur $Q \in (\mathbb{R}^N)^{\otimes 4}$ tel que

$$\begin{aligned} K_4(x_1, x_2, x_3, x_4) &\approx \sum_{1 \leq i_1, i_2, i_3, i_4 \leq N} Q_{i_1, i_2, i_3, i_4} \phi_{i_1}(x_1) \phi_{i_2}(x_2) \phi_{i_3}(x_3) \phi_{i_4}(x_4) \\ &= Q((\phi_j(x_1))_{1 \leq j \leq N}, (\phi_j(x_2))_{1 \leq j \leq N}, (\phi_j(x_3))_{1 \leq j \leq N}, (\phi_j(x_4))_{1 \leq j \leq N}), \end{aligned}$$

- la “matrice de masses d’ordre deux” $M^{(2)} \in (\mathbb{R}^N)^{\otimes 2}$ par

$$M_{i,j}^{(2)} = \Lambda(\mathbb{1}_W \phi_i \phi_j),$$

- le “tenseur de masses d’ordre trois” $M^{(3)} \in (\mathbb{R}^N)^{\otimes 3}$ par

$$M_{i,j,k}^{(3)} = \Lambda(\mathbb{1}_W \phi_i \phi_j \phi_k),$$

- et la “matrice source” $F \in (\mathbb{R}^N)^{\otimes 2}$ par

$$F_{l_1, l_2} = \int_{W^2} \lambda(x_1) \lambda(x_2) \phi_{l_1}(x_1) \phi_{l_2}(x_2) f(x_1, x_2; x_0) dx_1 dx_2.$$

Si on note alors $AL \in (\mathbb{R}^N)^{\otimes 4}$ le tenseur suivant

$$AL = \left(G(M_{j_1, l_1}^{(3)}, M_{j_2, l_2}^{(3)}) + T(M_{l_1}^{(2)}, M_{j_1, l_2}^{(3)}, M_{j_2}^{(2)}) \right. \\ \left. + T(M_{l_1, j_1}^{(3)}, M_{l_2}^{(2)}, M_{j_2}^{(2)}) + Q(M_{j_1}^{(2)}, M_{l_1}^{(2)}, M_{j_2}^{(2)}, M_{l_2}^{(2)}) \right)_{i_1, i_2, j_1, j_2},$$

alors P est solution de l’équation tensorielle

$$AL \times P = F,$$

$$\text{où } (AL \times P)_{i_1, i_2} = \sum_{j_1, j_2} AL_{i_1, i_2, j_1, j_2} P_{j_1, j_2}.$$

5.3 Remarques, problèmes, ouvertures ...

Cette nouvelle méthode ne présente pas spécialement plus de difficultés que l’ancienne d’un point de vue théorique. En revanche, c’est beaucoup plus compliqué en pratique. En effet, on peut d’abord remarquer que pour un même espace d’approximation de dimension N , on doit stocker un tenseur de dimension N^4 pour le BBUP alors qu’on reste en N^2 pour le BLUP. Cela entraîne donc une limite plus contraignante en terme d’espace de stockage mais également en temps pour appliquer la méthode dans un temps raisonnable. De plus, on a également besoin de renseigner les fonctions de corrélations d’ordre 3 et 4. Il faut donc l’estimer lorsqu’on travaille sur des données réelles, ce qui est bien plus difficile que pour l’intensité et la fonction de corrélation de paire pour lesquels il existe des estimateurs bien définis et des méthodes automatiques sur différents packages R pour calculer ces estimateurs. En revanche pour l’estimation de g_3 et g_4 , il existe très peu de références dans la littérature et aucune méthode déjà implémentée sous R par exemple. Pour l’ordre 3, on peut essayer de se référer à [3] pour estimer g_3 en pratique mais cela reste encore plus compliqué pour l’ordre 4.

CONCLUSION

L'objectif initial était d'étendre l'approche développée par Gabriel *et al.*[2] dans le cas d'un processus ponctuel stationnaire à d'autres cas. Je me suis d'abord concentré sur les processus marqués avec le cas multitype et le cas d'une marque continue. Puis je me suis intéressé à l'adaptation de l'estimateur pour tenir compte d'interaction d'ordre 3 ce qui permettrait de travailler sur des aspects plus directionnels, ce qu'on ne peut pas faire en se restreignant aux couples de points. En section 3, on s'est intéressé au cas d'un processus à marque continue. Ce cas là ne présente pas de difficultés particulières d'un point de vue théorique par rapport à ce qui est fait dans [2]. Cela devient plus compliqué en pratique, notamment pour gérer l'approximation sur l'espace des marques mais je ne me suis pas attardé sur cette problématique.

Un autre genre de processus ponctuels marqués est celui des processus dont l'espace des marques n'est pas continue mais discret, voir fini dans notre cas. On parle alors de processus multitypes ou multivariés. Le cas multitype est un peu confus puisque tout est répété en autant fois que le nombre de types. On a différentes fenêtres d'observation, différentes fonctions d'intensité, de corrélation de paire (croisée ou non) et différentes fonctions de poids également. Néanmoins, on peut toujours raisonner comme avant en appliquant la méthode des multiplicateurs de Lagrange pour obtenir ici un système d'équations intégrales vérifiée par les fonctions de poids de l'estimateur, et la méthode des éléments finis pour obtenir une solution approchée du système d'équations obtenu. Les résultats obtenus sont plutôt satisfaisants et on peut voir clairement qu'on arrive à tirer de l'information à partir des types qui ne sont pas celui qu'on cherche à prédire.

Finalement, dans la partie 5 on a adapté l'estimateur BLUP qui ne permet pas de faire de prédiction correcte dans le cas de processus ponctuels portés par un processus de fibres. C'est probablement la partie la plus intéressante car la plus novatrice mais c'est aussi celle où on rencontre le plus de difficultés. Faute de temps et suite à ces difficultés, je n'ai pas pu vraiment tester cette méthode en pratique pendant de mon stage, pas même sur des données simulées. En revanche, cela a permis de se poser différentes questions et d'avoir des pistes de réflexion également d'intérêt pour les méthodes précédentes.

La première limite qu'on rencontre lorsqu'on souhaite appliquer le nouvel estimateur est qu'on a besoin de connaître les fonctions de corrélations d'ordre 3 et 4 (g_3 et g_4). En pratique, pour appliquer la méthode sur des données réelles il faudrait estimer ces interactions. L'estimation de l'intensité et de la fonction de corrélation de paire sont largement étudiées déjà et des méthodes automatiques

implémentées sous R existent. En revanche, je n’ai rien trouvé dans la littérature pour une méthode simple ou déjà implémentées sous un langage de programmation pour l’estimation de g_3 et g_4 . Ce "vide" dans la littérature soulève également la question de savoir quelle qualité on peut espérer pour l’estimation des fonctions de corrélation g_3 et g_4 puisque, même dans le cas où on peut estimer les fonctions λ, g_2, g_3 et g_4 intervenant dans la définition de l’estimateur, on obtient pas en pratique l’estimateur qu’on voulait. Premièrement, on n’aura pas la bonne équation intégrale puisqu’on aura estimé ces fonctions et dans un second temps, la fonction de poids finalement obtenue est une solution approchée de cette même équation intégrale “approchée”. Il s’agirait donc d’étudier la sensibilité de l’estimateur par rapport à la qualité de l’estimation des différentes fonctions et par rapport à la qualité de l’approximation avec la méthode des éléments finis. On peut voir en particulier pour cette dernière problématique que la question reste pertinente dans le cas le plus simple d’un processus ponctuel stationnaire. Finalement, on a aussi des problématiques “numériques” puisqu’on peut se demander ce qu’est un temps raisonnable de calcul pour la méthode en pratique, sachant que cette question est directement liée à la qualité de l’approximation voulue pour la méthode des éléments finis qui est elle même limitée par l’espace mémoire disponible en machine.

References

- [1] Viktor Beneš, Karel Bodlák, Jesper Møller, and Rasmus Waagepetersen. A case study on point process modelling in disease mapping. *Image Analysis & Stereology*, 24(3):159–168, 2011. 2
- [2] E. Gabriel, J. Coville, and J. Chadœuf. Estimating the intensity function of spatial point processes outside the observation window. *Spatial Statistics*, 22:225 – 239, 2017. Spatio-temporal Statistical Methods in Environmental and Biometrical Problems. 2, 10, 11, 13, 21, 25
- [3] K H. Hanisch. Reduction of n-th moment measures and the special case of the third moment measure of stationary and isotropic planar point processes. *Statistics: A Journal of Theoretical and Applied Statistics*, 14:421–435, 01 1983. 24
- [4] Richard Law, Janine Illian, David F. R. P. Burslem, Georg Gratzer, C. V. S. Gunatilleke, and I. A. U. N. Gunatilleke. Ecological information from spatial patterns of plants: insights from point process theory. *Journal of Ecology*, 97(4):616–628, 2009. 2
- [5] Tomáš Mrkvička and Ilya Molchanov. Optimisation of linear unbiased intensity

estimators for point processes. *Annals of the Institute of Statistical Mathematics*, 57(1):71–81, Mar 2005. 44

- [6] Jerzy Neyman and Elizabeth L. Scott. Statistical approach to problems of cosmology. *Journal of the Royal Statistical Society: Series B (Methodological)*, 20(1):1–29, 1958. 2
- [7] Dietrich Stoyan, Wilfrid S Kendall, and Joseph Mecke. Stochastic geometry and its applications, 1986. 3
- [8] Eberhard Zeidler. Nonlinear functional analysis and its applications i. 1986. 46
- [9] Jiancang Zhuang and Jorge Mateu. A semiparametric spatiotemporal hawkes-type point process model with periodic background for crime data. *Journal of the Royal Statistical Society: Series A (Statistics in Society)*, 182(3):919–942, 2019. 2

A Équations de Fredholm

La partie A regroupe les calculs qui permettent d'aboutir aux différentes équations de Fredholm par la méthode des multiplicateurs de Lagrange. Dans chaque cas, on admet l'existence d'une fonction de poids minimisant l'erreur quadratique parmi les fonctions donnant un estimateur sans-biais.

A.1 Processus à marque continue

On considère un processus ponctuel avec des marques continues sur \mathbb{R} . On suppose qu'on a

$$\Lambda(d(x,m)) = \lambda(x,m)dm dx$$

et une fonction de corrélation de paire $g(x,m,y,n)$ continue. On note $\Lambda_g(dx) = \Lambda(dx \times \mathbb{R})$. On fixe x_0 dans la fenêtre de prédiction et une marque $m_0 \in \mathbb{R}$. On considère l'estimateur

$$\hat{\lambda}(x_0, m_0 | \phi_W) = \sum_{[x,m] \in \phi \cap W} w_{x_0, m_0}(x, m).$$

On notera juste $w(x,m)$ dans la suite. On a

$$\mathbb{E} [\hat{\lambda}] = \int_{\mathbb{R} \times W} w(x,m) \lambda(x,m) dm dx$$

donc pour obtenir un estimateur sans biais il faut que la fonction de poids w vérifie la condition suivante

$$\int_{\mathbb{R} \times W} w(x,m) \lambda(x,m) dm dx = \lambda(x_0, m_0). \quad (7)$$

Par le calcul, on a

$$\begin{aligned} \mathbb{E} [\hat{\lambda}^2] &= \int_{\mathbb{R} \times W} w^2(x,m) \lambda(x,m) dm dx \\ &\quad + \int_{\mathbb{R}^2 \times W^2} w(x,m) w(x',m') \lambda(x,m) \lambda(x',m') g(x,m,x',m') dm dm' dx dx' \end{aligned}$$

et

$$\mathbb{E}[\lambda \hat{\lambda}] = \lambda(x_0, m_0) \int w(x,m) \lambda(x,m) g(x,m,x_0, m_0) dm dx.$$

On veut minimiser l'erreur quadratique de l'estimateur $\hat{\lambda}$ ce qui revient à minimiser la fonctionnelle

$$\begin{aligned} F(w) &= \mathbb{E} [\hat{\lambda}^2] - 2\mathbb{E}[\hat{\lambda}\lambda] \\ &= \int w^2(x,m)\lambda(x,m)dm dx \\ &\quad + \int w(x,m)w(x',m')\lambda(x,m)\lambda(x',m')g(x,m,x',m')dm dm' dx dx' \\ &\quad - 2\lambda(x_0,m_0) \int w(x,m)\lambda(x,m)g(x,m,x_0,m_0)dm dx. \end{aligned}$$

On veut minimiser l'erreur quadratique sous la contrainte (7) donc on va appliquer la méthode des multiplicateurs de Lagrange. On cherche un extremum de

$$\begin{aligned} T(w) &= \int w^2(x,m)\lambda(x,m)dm dx \\ &\quad + \int w(x,m)w(x',m')\lambda(x,m)\lambda(x',m')g(x,m,x',m')dm dm' dx dx' \\ &\quad - 2\lambda(x_0,m_0) \int w(x,m)\lambda(x,m)g(x,m,x_0,m_0)dm dx \\ &\quad + \mu \left(\int w(x,m)\lambda(x,m)dm dx - \lambda(x_0,m_0) \right) \end{aligned}$$

où μ est une constante telle que notre estimateur est sans biais. En notant $\alpha = w + \epsilon$, on a

$$\begin{aligned} T(\alpha) &\approx T(w) + 2 \int w(x,m)\epsilon(x,m)\lambda(x,m)dm dx \\ &\quad + 2 \int w(x,m)\epsilon(x',m')\lambda(x,m)\lambda(x',m')g(x,m,x',m')dm dm' dx dx' \\ &\quad - 2\lambda(x_0,m_0) \int \epsilon(x,m)\lambda(x,m)g(x,m,x_0,m_0)dm dx \\ &\quad + \mu \int \epsilon(x,m)\lambda(x,m)dm dx \\ &= 2 \int_{\mathbb{R} \times W} \lambda(x,m)\epsilon(x,m)h(x,m)dm dx \end{aligned}$$

où

$$\begin{aligned} h(x,m) &= w(x,m) + \int_{\mathbb{R} \times W} w(x',m')\lambda(x',m')g(x,m,x',m')dm' dx' \\ &\quad - \lambda(x_0,m_0)g(x,m,x_0,m_0) + \frac{\mu}{2}. \end{aligned}$$

Si w minimise T , alors on a $T(\alpha) = T(w) + o(\epsilon^2)$. Donc on cherche w tel que

$$\lambda(x, m) \left\{ w(x, m) + \int w(x', m') \lambda(x', m') g(x, m, x', m') dm' dx' - \lambda(x_0, m_0) g(x, m, x_0, m_0) + \frac{\mu}{2} \right\} = 0.$$

On intègre pour obtenir la valeur de μ

$$\begin{aligned} \lambda(x_0, m_0) + \int_{\mathbb{R}^2 \times W^2} w(x, m) \lambda(x, m) w(x', m') \lambda(x', m') g(x, m, x', m') dm dm' dx dx' \\ - \lambda(x_0, m_0) \int_{\mathbb{R} \times W} \lambda(x, m) g(x, m, x_0, m_0) dm dx \\ + \frac{\mu}{2} \Lambda(W \times \mathbb{R}) = 0. \end{aligned}$$

Comme $\Lambda_g(W) = \Lambda(W \times \mathbb{R})$, on a alors l'équation

$$\begin{aligned} \lambda(x, m) \left\{ w(x, m) + \int w(x', m') \lambda(x', m') g(x, m, x', m') dm' dx' \right. \\ - \lambda(x_0, m_0) g(x, m, x_0, m_0) \\ + \frac{1}{\Lambda_g(W)} \lambda(x_0, m_0) \int \lambda(x, m) g(x, m, x_0, m_0) dm dx \\ - \frac{1}{\Lambda_g(W)} \lambda(x_0, m_0) \\ \left. - \frac{1}{\Lambda_g(W)} \int_{\mathbb{R}^2 \times W^2} \lambda(x, m) w(x', m') \lambda(x', m') g(x, m, x', m') dm dm' dx dx' \right\} = 0. \end{aligned}$$

Finalement, la fonction de poids de notre estimateur est solution de l'équation intégrale

$$\lambda(x, m) \left\{ w(x, m) + \int_{\mathbb{R} \times W} w(y, n) K(x, m, y, n) dn dy - f(x, m; x_0, m_0) \right\} = 0$$

où

$$K(x, m, y, n) = g(x, m, y, n) - \frac{1}{\Lambda_g(W)} \int \lambda(x, m) g(x, m, y, n) dm dx$$

et

$$f(x, m; x_0) = \lambda(x_0, m_0) \left[g(x, m, x_0, m_0) + \frac{1}{\Lambda_g(W)} - \frac{1}{\Lambda_g(W)} \int_{\mathbb{R} \times W} \lambda(x, m) g(x, m, x_0, m_0) dm dx \right].$$

A.2 Processus multitype

On considère M processus ponctuels ϕ_1, \dots, ϕ_M observés respectivement sur les fenêtres d'observation W_1, \dots, W_M . On fixe $k_0 \in \{1, \dots, M\}$ et $x_0 \notin W_{k_0}$. Le BLUP est sous la forme

$$\hat{\lambda}_{k_0}(x_0 | \phi_{1W_1}, \dots, \phi_{MW_M}) = \sum_{k=1}^M \sum_{x \in \phi_{kW_k}} w(x, k; x_0, k_0).$$

Pour simplifier les notations, on notera juste

- $\hat{\lambda}$ pour $\hat{\lambda}_{k_0}(x_0 | \phi_{1W_1}, \dots, \phi_{MW_M})$,
- $\lambda(x_0 | \phi_W)$ pour $\lambda_{k_0}(x_0 | \phi_{1W_1}, \dots, \phi_{MW_M})$,
- et $w_k(x)$ pour $w(x, k; x_0, k_0)$.

On a

$$\mathbb{E}[\hat{\lambda}] = \sum_{k=1}^M \int_{W_k} w_k(x) \lambda_k(x) dx$$

et $\mathbb{E}[\lambda_{k_0}(x_0 | \phi_{W_1}, \dots, \phi_{W_M})] = \lambda_{k_0}(x_0)$ donc l'estimateur $\hat{\lambda}$ est non-biaisé si et seulement si

$$\sum_{k=1}^M \int_{W_k} w_k(x) \lambda_k(x) dx = \lambda_{k_0}(x_0). \quad (8)$$

Par le calcul, on a

$$\begin{aligned} \mathbb{E}[\hat{\lambda}^2] &= \sum_{k=1}^M \int_{W_k} w_k^2(x) \lambda_k(x) dx \\ &\quad + \sum_{1 \leq k, k' \leq M} \int_{W_k \times W_{k'}} w_k(x) w_{k'}(x') \lambda_k(x) \lambda_{k'}(x') g_{kk'}(x, x') dx dx' \end{aligned}$$

et

$$\mathbb{E}[\hat{\lambda} \lambda(x_0 | \phi_W)] = \lambda_{k_0}(x_0) \sum_{k=1}^M \int_{W_k} w_k(x) \lambda_k(x) g_{k_0, k}(x_0, x) dx.$$

On veut minimiser l'erreur quadratique sous la contrainte d'avoir un estimateur

non biaisé. Cela revient à minimiser la quantité

$$\begin{aligned}
F(w) &= \mathbb{E} [\hat{\lambda}^2] - 2\mathbb{E}[\lambda(x_0|\phi_W)\hat{\lambda}] \\
&= \sum_{k=1}^M \int_{W_k} w_k^2(x) \lambda_k(x) dx \\
&\quad + \sum_{1 \leq k, k' \leq M} \int_{W_k \times W_{k'}} w_k(x) w_{k'}(x') \lambda_k(x) \lambda_{k'}(x') g_{k,k'}(x, x') dx dx' \\
&\quad - 2\lambda_{k_0}(x_0) \sum_{k=1}^M \int_{W_k} w_k(x) \lambda_k(x) g_{k_0,k}(x_0, x) dx
\end{aligned}$$

conditionnellement à (8). On applique la méthode des multiplicateurs de Lagrange, on cherche alors à minimiser

$$\begin{aligned}
T(w) &= \sum_{k=1}^M \int_{W_k} w_k^2(x) \lambda_k(x) dx \\
&\quad + \sum_{1 \leq k, k' \leq M} \int_{W_k \times W_{k'}} w_k(x) w_{k'}(x') \lambda_k(x) \lambda_{k'}(x') g_{k,k'}(x, x') dx dx' \\
&\quad - 2\lambda_{k_0}(x_0) \sum_{k=1}^M \int_{W_k} w_k(x) \lambda_k(x) g_{k_0,k}(x_0, x) dx \\
&\quad + \mu \left(\sum_{k=1}^M \int_{W_k} w_k(x) \lambda_k(x) dx - \lambda_{k_0}(x_0) \right).
\end{aligned}$$

où μ est une constante telle que $\mathbb{E}[\hat{\lambda}] = \lambda_{k_0}(x_0)$. On pose $\alpha = w + \varepsilon$ (à voir comme un vecteur, on peut faire composante par composante sinon), on a

$$\begin{aligned}
T(\alpha) \approx T(w) + 2 \sum_{k=1}^M \int_{W_k} \varepsilon_k(x) \lambda_k(x) \left[w_k(x) - \lambda_{k_0}(x_0) g_{k,k_0}(x, x_0) + \frac{\mu}{2} \right. \\
\left. + \sum_{k'=1}^M \int_{W_{k'}} w_{k'}(x') \lambda_{k'}(x') g_{k,k'}(x, x') dx' \right] dx.
\end{aligned}$$

Donc w réalise un minimum de T si pour tout $k \in \{1, \dots, M\}$ et $x \in W_k$, on a

$$\begin{aligned}
\lambda_k(x) \left[w_k(x) - \lambda_{k_0}(x_0) g_{k,k_0}(x, x_0) + \frac{\mu}{2} \right. \\
\left. + \sum_{k'=1}^M \int_{W_{k'}} w_{k'}(x') \lambda_{k'}(x') g_{k,k'}(x, x') dx' \right] = 0.
\end{aligned} \tag{9}$$

On obtient ensuite la valeur de μ en intégrant, i.e.

$$\begin{aligned} \int_{W_k} w_k(x) \lambda_k(x) dx + \sum_{k'=1}^M \int_{W_k \times W_{k'}} w_{k'}(x') \lambda_k(x) \lambda_{k'}(x') g_{k,k'}(x, x') dx dx' \\ - \lambda_{k_0}(x_0) \int_{W_k} \lambda_k(x) g_{k,k_0}(x, x_0) dx + \mu \frac{\Lambda_k(W_k)}{2} = 0. \end{aligned}$$

On a donc

$$\begin{aligned} \frac{\mu}{2} = \frac{1}{\Lambda_k(W_k)} \left\{ \lambda_{k_0}(x_0) \int_{W_k} \lambda_k(x) g_{k,k_0}(x, x_0) dx \right. \\ \left. - \int_{W_k} w_k(x) \lambda_k(x) dx \right. \\ \left. - \sum_{k'=1}^M \int_{W_k \times W_{k'}} w_{k'}(x') \lambda_k(x) \lambda_{k'}(x') g_{k,k'}(x, x') dx dx' \right\} \end{aligned}$$

pour tout $k \in \{1, \dots, M\}$. On a donc une nouvelle condition sur les poids w_k puisque μ ne doit pas dépendre de k . En remplaçant μ par la valeur obtenue dans (9) on retrouve

$$\begin{aligned} \lambda_k(x) \cdot \left\{ w_k(x) - \lambda_{k_0}(x_0) g_{k,k_0}(x, x_0) \right. \\ \left. + \sum_{k'=1}^M \int_{W_{k'}} w_{k'}(x') \lambda_{k'}(x') g_{k,k'}(x, x') dx' \right. \\ \left. - \frac{1}{\Lambda_k(W_k)} \left[\int_{W_k} w_k(x) \lambda_k(x) dx - \lambda_{k_0}(x_0) \int_{W_k} \lambda_k(x) g_{k,k_0}(x, x_0) dx \right. \right. \\ \left. \left. + \sum_{k'=1}^M \int_{W_k \times W_{k'}} w_{k'}(x') \lambda_k(x) \lambda_{k'}(x') g_{k,k'}(x, x') dx dx' \right] \right\} = 0. \end{aligned}$$

Finalement, le meilleur estimateur linéaire non-biaisé a des poids $(w_k(x))_k$ qui vérifient pour tout $k \in \{1, \dots, M\}$

$$\begin{aligned} \lambda_k(x) \cdot \left\{ w_k(x) \right. \\ \left. + \sum_{k'=1}^M \int_{W_{k'}} w_{k'}(y) \lambda_{k'}(y) \left(g_{k,k'}(x, y) - \frac{1}{\Lambda_k(W_k)} \int_{W_k} \lambda_k(x') g_{k,k'}(x', y) dx' - \frac{\mathbb{1}_{k=k'}}{\Lambda_k(W_k)} \right) dy \right\} \\ = \lambda_k(x) \lambda_{k_0}(x_0) \left(g_{k,k_0}(x, x_0) - \frac{1}{\Lambda_k(W_k)} \int_{W_k} \lambda_k(x') g_{k,k_0}(x', x_0) dx' \right) \end{aligned}$$

et avec la condition

$$\begin{aligned} \frac{1}{\Lambda_{k_1}(W_{k_1})} & \left[\lambda_{k_0}(x_0) \int_{W_{k_1}} \lambda_{k_1}(x) g_{k_1,k_0}(x, x_0) dx \right. \\ & - \sum_{k'=1}^M \int_{W_{k_1} \times W_{k'}} w_{k'}(x') \lambda_{k_1}(x) \lambda_{k'}(x') g_{k_1,k'}(x, x') dx dx' \\ & \left. - \int_W w_{k_1}(x) \lambda_{k_1}(x) dx \right] = \frac{1}{\Lambda_{k_2}(W_{k_2})} \left[\lambda_{k_0}(x_0) \int_{W_{k_2}} \lambda_{k_2}(x) g_{k_2,k_0}(x, x_0) dx \right. \\ & - \sum_{k'=1}^M \int_{W_{k_2} \times W_{k'}} w_{k'}(x') \lambda_{k_2}(x) \lambda_{k'}(x') g_{k_2,k'}(x, x') dx dx' \\ & \left. - \int_{W_{k_2}} w_{k_2}(x) \lambda_{k_2}(x) dx \right] \end{aligned}$$

pour tous $k_1, k_2 \in \{1, \dots, M\}$. On peut résumer l'équation par l'écriture suivante

$$\lambda_k(x) \cdot \left\{ w_k(x) + \sum_{k'=1}^M \int_{W_{k'}} w_{k'}(y) \lambda_{k'}(y) K_{k,k'}(x, y) dy - f(x; x_0) \right\} = 0, \forall k \in \{1, \dots, M\}, \forall x \in W_k,$$

avec

$$K_{k,k'}(x, y) = \left(g_{k,k'}(x, y) - \frac{1}{\Lambda_k(W_k)} \int_{W_k} \lambda_k(x') g_{k,k'}(x', y) dx' - \frac{\mathbb{1}_{k=k'}}{\Lambda_k(W_k)} \right),$$

et

$$f_k(x; x_0) = \lambda_{k_0}(x_0) \left(g_{k,k_0}(x, x_0) - \frac{1}{\Lambda_k(W_k)} \int_{W_k} \lambda_k(x') g_{k,k_0}(x', x_0) dx' \right).$$

A.3 “Cas directionnel”

On considère un processus ponctuel ϕ dans \mathbb{R}^d et W un borélien de \mathbb{R}^d qui correspond à la fenêtre d'observation de ϕ et on suppose que $\Lambda(W) < \infty$. On considère les estimateurs de la forme

$$\hat{\lambda}(x_0 | \phi_W) = \sum_{x_1, x_2 \in W}^{\neq} w(x_1, x_2; x_0), \quad (10)$$

et on veut déterminer l'estimateur de cette forme qui minimise l'erreur quadratique parmi les estimateurs sans biais. On fixe $x_0 \notin W$ et pour simplifier les notations, on écrira

- $\hat{\lambda}$ pour $\hat{\lambda}(x_0 | \phi_W)$,

- et $w(x_1, x_2)$ pour $w(x_1, x_2; x_0)$.

On peut remarquer que

$$\begin{aligned}\hat{\lambda} &= \frac{1}{2} \sum_{x_1, x_2 \in \phi_W}^{\neq} w(x_1, x_2) + \frac{1}{2} \sum_{x_1, x_2 \in \phi_W}^{\neq} w(x_2, x_1) \\ &= \frac{1}{2} \sum_{x_1, x_2 \in \phi_W}^{\neq} (w(x_1, x_2) + w(x_2, x_1)).\end{aligned}$$

Pour une fonction w donnée, on définit p par $p(x_1, x_2) = w(x_1, x_2) + w(x_2, x_1)$ qui a l'avantage d'être symétrique. Le calcul ci-dessus montre qu'on cherche à déterminer une fonction p qui peut être obtenue à partir de différentes fonctions w .

On pourrait aussi décomposer w en la somme d'une fonction symétrique et une fonction antisymétrique. La partie antisymétrique de w n'a alors aucune influence sur $\hat{\lambda}$ donc on veut juste estimer la partie symétrique de w qui est p (à un facteur 2 près).

On va donc déterminer la fonction de poids p associée au meilleur estimateur sans biais de la forme (10). On a

$$\mathbb{E}[\hat{\lambda}] = \int_{W^2} w(x_1, x_2) \lambda(x_1) \lambda(x_2) g_2(x_1, x_2) dx_1 dx_2$$

donc pour avoir un estimateur non-biaisé, la condition suivante est nécessaire

$$\int_{W^2} w(x_1, x_2) \lambda(x_1) \lambda(x_2) g_2(x_1, x_2) dx_1 dx_2 = \lambda(x_0). \quad (11)$$

Pour calculer l'erreur quadratique de notre estimateur on calcule les quantités

$$\begin{aligned}\mathbb{E}[\hat{\lambda}^2] &= \int_{W^2} w^2(x_1, x_2) \lambda(x_1) \lambda(x_2) g_2(x_1, x_2) dx_1 dx_2 \\ &+ \int_{W^2} w(x_1, x_2) w(x_2, x_1) \lambda(x_1) \lambda(x_2) g_2(x_1, x_2) dx_1 dx_2 \\ &+ \int_{W^3} w(x_1, x_2) w(x_1, x_3) \lambda(x_1) \lambda(x_2) \lambda(x_3) g_3(x_1, x_2, x_3) dx_1 dx_2 dx_3 \\ &+ \int_{W^3} w(x_1, x_2) w(x_3, x_2) \lambda(x_1) \lambda(x_2) \lambda(x_3) g_3(x_1, x_2, x_3) dx_1 dx_2 dx_3 \\ &+ 2 \int_{W^3} w(x_1, x_2) w(x_3, x_1) \lambda(x_1) \lambda(x_2) \lambda(x_3) g_3(x_1, x_2, x_3) dx_1 dx_2 dx_3 \\ &+ \int_{W^4} w(x_1, x_2) w(x_3, x_4) \lambda(x_1) \lambda(x_2) \lambda(x_3) \lambda(x_4) g_4(x_1, x_2, x_3, x_4) dx_1 dx_2 dx_3 dx_4\end{aligned}$$

et

$$\mathbb{E}[\hat{\lambda} \lambda] = \lambda(x_0) \int_{W^2} w(x_1, x_2) \lambda(x_1) \lambda(x_2) g_3(x_0, x_1, x_2) dx_1 dx_2.$$

Minimiser l'erreur quadratique revient à minimiser la fonctionnelle suivante

$$\begin{aligned}
F(w) &= \mathbb{E} [\hat{\lambda}^2] - 2\mathbb{E}[\hat{\lambda}\lambda] \\
&= \int_{W^2} w^2(x_1, x_2) \lambda(x_1) \lambda(x_2) g_2(x_1, x_2) dx_1 dx_2 \\
&+ \int_{W^2} w(x_1, x_2) w(x_2, x_1) \lambda(x_1) \lambda(x_2) g_2(x_1, x_2) dx_1 dx_2 \\
&+ \int_{W^3} w(x_1, x_2) w(x_1, x_3) \lambda(x_1) \lambda(x_2) \lambda(x_3) g_3(x_1, x_2, x_3) dx_1 dx_2 dx_3 \\
&+ \int_{W^3} w(x_1, x_2) w(x_3, x_2) \lambda(x_1) \lambda(x_2) \lambda(x_3) g_3(x_1, x_2, x_3) dx_1 dx_2 dx_3 \\
&+ 2 \int_{W^3} w(x_1, x_2) w(x_3, x_1) \lambda(x_1) \lambda(x_2) \lambda(x_3) g_3(x_1, x_2, x_3) dx_1 dx_2 dx_3 \\
&+ \int_{W^4} w(x_1, x_2) w(x_3, x_4) \lambda(x_1) \lambda(x_2) \lambda(x_3) \lambda(x_4) g_4(x_1, x_2, x_3, x_4) dx_1 dx_2 dx_3 dx_4 \\
&- 2\lambda(x_0) \int_{W^2} w(x_1, x_2) \lambda(x_1) \lambda(x_2) g_3(x_0, x_1, x_2) dx_1 dx_2.
\end{aligned}$$

On utilise la méthode des multiplicateurs de Lagrange pour minimiser F conditionnellement à (11). On note

$$T(w) = F(w) + \mu \left(\int_{W^2} w(x_1, x_2) \lambda(x_1) \lambda(x_2) g_2(x_1, x_2) dx_1 dx_2 - \lambda(x_0) \right)$$

où μ est une constante telle que notre estimateur soit sans biais. On note $\alpha = w + \epsilon$ et on a

$$\begin{aligned}
T(\alpha) - T(w) &\approx 2 \int_{W^2} w(x_1, x_2) \epsilon(x_1, x_2) \lambda(x_1) \lambda(x_2) g_2(x_1, x_2) dx_1 dx_2 \\
&+ 2 \int_{W^2} w(x_1, x_2) \epsilon(x_2, x_1) \lambda(x_1) \lambda(x_2) g_2(x_1, x_2) dx_1 dx_2 \\
&+ 2 \int_{W^3} w(x_1, x_2) \epsilon(x_1, x_3) \lambda(x_1) \lambda(x_2) \lambda(x_3) g_3(x_1, x_2, x_3) dx_1 dx_2 dx_3 \\
&+ 2 \int_{W^3} w(x_1, x_2) \epsilon(x_3, x_2) \lambda(x_1) \lambda(x_2) \lambda(x_3) g_3(x_1, x_2, x_3) dx_1 dx_2 dx_3 \\
&+ 2 \int_{W^3} w(x_1, x_2) \epsilon(x_3, x_1) \lambda(x_1) \lambda(x_2) \lambda(x_3) g_3(x_1, x_2, x_3) dx_1 dx_2 dx_3 \\
&+ 2 \int_{W^3} w(x_1, x_2) \epsilon(x_2, x_3) \lambda(x_1) \lambda(x_2) \lambda(x_3) g_3(x_1, x_2, x_3) dx_1 dx_2 dx_3 \\
&+ 2 \int_{W^4} w(x_1, x_2) \epsilon(x_3, x_4) \lambda(x_1) \lambda(x_2) \lambda(x_3) \lambda(x_4) g_4(x_1, x_2, x_3, x_4) dx_1 dx_2 dx_3 dx_4 \\
&- 2\lambda(x_0) \int_{W^2} \epsilon(x_1, x_2) \lambda(x_1) \lambda(x_2) g_3(x_0, x_1, x_2) dx_1 dx_2 \\
&+ \mu \int_{W^2} \epsilon(x_1, x_2) \lambda(x_1) \lambda(x_2) g_2(x_1, x_2) dx_1 dx_2.
\end{aligned}$$

Soit

$$T(\alpha) - T(w) \approx 2 \int_{W^2} \epsilon(x_1, x_2) h(x_1, x_2) \lambda(x_1) \lambda(x_2) dx_1 dx_2$$

où

$$\begin{aligned} h(x_1, x_2) &= (w(x_1, x_2) + w(x_2, x_1)) g_2(x_1, x_2) \\ &+ \int_W (w(x_1, x_3) + w(x_3, x_2) + w(x_3, x_1) + w(x_2, x_3)) \lambda(x_3) g_3(x_1, x_2, x_3) dx_3 \\ &+ \int_{W^2} w(x_3, x_4) \lambda(x_3) \lambda(x_4) g_4(x_1, x_2, x_3, x_4) dx_3 dx_4 \\ &- \lambda(x_0) g_3(x_0, x_1, x_2) \\ &+ \frac{\mu}{2} g_2(x_1, x_2). \end{aligned}$$

On peut réécrire h en utilisant p plutôt que w pour les raisons évoquées plus tôt. On a

$$\begin{aligned} h(x_1, x_2) &= p(x_1, x_2) g_2(x_1, x_2) + \int_W (p(x_1, x_3) + p(x_2, x_3)) \lambda(x_3) g_3(x_1, x_2, x_3) dx_3 \\ &+ \frac{1}{2} \int_{W^2} p(x_3, x_4) \lambda(x_3) \lambda(x_4) g_4(x_1, x_2, x_3, x_4) dx_3 dx_4 \\ &- \lambda(x_0) g_3(x_0, x_1, x_2) + \frac{\mu}{2} g_2(x_1, x_2). \end{aligned}$$

On en déduit que w est un extremum de T si et seulement si

$$\forall x_1, x_2 \in W, \lambda(x_1) \lambda(x_2) h(x_1, x_2) = 0.$$

On intègre pour obtenir la valeur de μ sachant(11) d'où

$$\begin{aligned} 0 &= \int_{W^2} \lambda(x_1) \lambda(x_2) h(x_1, x_2) dx_1 dx_2 \\ &= 2\lambda(x_0) + \int_{W^3} \lambda(x_1) \lambda(x_2) \lambda(x_3) (p(x_1, x_3) + p(x_2, x_3)) g_3(x_1, x_2, x_3) dx_1 dx_2 dx_3 \\ &+ \frac{1}{2} \int_{W^4} \lambda(x_1) \lambda(x_2) \lambda(x_3) \lambda(x_4) p(x_1, x_2) g_4(x_1, x_2, x_3, x_4) dx_1 dx_2 dx_3 dx_4 \\ &- \lambda(x_0) \int_{W^2} \lambda(x_1) \lambda(x_2) g_3(x_1, x_2, x_0) dx_1 dx_2 \\ &+ \frac{\mu}{2} \int_{W^2} \lambda(x_1) \lambda(x_2) g_2(x_1, x_2) dx_1 dx_2. \end{aligned}$$

On a alors

$$\begin{aligned}
h(x_1, x_2) = & p(x_1, x_2)g_2(x_1, x_2) \\
& + \int_W (p(x_1, x_3) + p(x_2, x_3))\lambda(x_3)g_3(x_1, x_2, x_3)dx_3 \\
& + \frac{1}{2} \int_{W^2} p(x_3, x_4)\lambda(x_3)\lambda(x_4)g_4(x_1, x_2, x_3, x_4)dx_3dx_4 \\
& - \lambda(x_0)g_3(x_0, x_1, x_2) \\
& - \frac{g_2(x_1, x_2)}{\alpha^{(2)}(W^2)} \int_{W^3} \lambda(x_1)\lambda(x_2)\lambda(x_3)(p(x_1, x_3) + p(x_2, x_3))g_3(x_1, x_2, x_3)dx_1dx_2dx_3 \\
& - \frac{1}{2} \frac{g_2(x_1, x_2)}{\alpha^{(2)}(W^2)} \int_{W^4} \lambda(x_1)\lambda(x_2)\lambda(x_3)\lambda(x_4)p(x_3, x_4)g_4(x_1, x_2, x_3, x_4)dx_1dx_2dx_3dx_4 \\
& - \frac{\lambda(x_0)g_2(x_1, x_2)}{\alpha^{(2)}(W^2)} \left[2 - \int_{W^2} \lambda(x_1)\lambda(x_2)g_3(x_1, x_2, x_0)dx_1dx_2 \right].
\end{aligned}$$

Finalement, on a l'équation intégrale suivante

$$\begin{aligned}
\lambda(x_1)\lambda(x_2) \Big[& p(x_1, x_2)g_2(x_1, x_2) + \int_W (p(x_1, x_3) + p(x_2, x_3))\lambda(x_3)g_3(x_1, x_2, x_3)dx_3 \\
& + \int_{W^2} \lambda(x_3)\lambda(x_4)p(x_3, x_4)K_4(x_1, x_2, x_3, x_4)dx_3dx_4 \\
& - f(x_1, x_2; x_0) \Big] = 0
\end{aligned}$$

où

$$\begin{aligned}
K_4(x_1, x_2, x_3, x_4) = & \frac{1}{2}g_4(x_1, x_2, x_3, x_4) - \frac{g_2(x_1, x_2)}{2\alpha^{(2)}(W^2)} \int_{W^2} \lambda(x_1)\lambda(x_2)g_4(x_1, x_2, x_3, x_4)dx_1dx_2 \\
& - \frac{2g_2(x_1, x_2)}{\alpha^{(2)}(W^2)} \int_W \lambda(x_1)g_3(x_1, x_3, x_4)dx_1,
\end{aligned}$$

et

$$\begin{aligned}
f(x_1, x_2; x_0) = & \lambda(x_0) \left(g_3(x_0, x_1, x_2) + \frac{2g_2(x_1, x_2)}{\alpha^{(2)}(W^2)} \right. \\
& \left. - \frac{g_2(x_1, x_2)}{\alpha^{(2)}(W^2)} \int_{W^2} \lambda(x_1)\lambda(x_2)g_3(x_0, x_1, x_2)dx_1dx_2 \right).
\end{aligned}$$

B Résolution par la méthode des éléments finis

Dans cette partie, on détaille les calculs pour appliquer la méthode de résolution approchée des éléments finis aux équations obtenus dans la partie précédente. En pratique, j'ai utilisé par la suite *FreeFem++* pour faire des prédictions.

B.1 Processus à marque continue

On rappelle qu'on veut résoudre une équation intégrale de la forme

$$\lambda(x, m) \left\{ w(x, m) + \int_{\mathbb{R} \times W} w(y, n) K(x, m, y, n) dndy - f(x, m; x_0, m_0) \right\} = 0.$$

On note V_h l'espace d'approximation, sous-espace de dimension finie de $L^2(W \times \mathbb{M})$. On note $N = \dim(V_h)$ sa dimension et $\{\phi_i\}_{i=1, \dots, N}$ une base de V_h . On projette les différentes fonctions sur l'espace d'approximation par

$$w(x, m) \approx \sum_{i=1}^N \omega_i \phi_i(x, m)$$

et

$$K(x, m, x', m') \approx \sum_{1 \leq i, j \leq N} K_{i,j} \phi_i(x, m) \phi_j(x', m').$$

Fixons $l \in \{1, \dots, N\}$. L'équation devient

$$\begin{aligned} \sum_{i=1}^N \omega_i \Lambda(\mathbb{1}_{W \times \mathbb{M}} \phi_i \phi_l) + \sum_{i=1}^N \omega_i \sum_{1 \leq j_1, j_2 \leq N} K_{j_1, j_2} \Lambda(\mathbb{1}_{W \times \mathbb{M}} \phi_l \phi_{j_1}) \Lambda(\mathbb{1}_{W \times \mathbb{M}} \phi_i \phi_{j_2}) \\ = \int_{W \times \mathbb{M}} \lambda(x, m) \phi_l(x, m) f(x, m; x_0, m_0) dx dm. \end{aligned}$$

On définit la matrice de masses M par

$$M_{i,j} = \Lambda(\mathbb{1}_{W \times \mathbb{M}} \phi_i \phi_j).$$

On a alors

$$(M\Omega)_l + (MKM\Omega)_l = F_l,$$

où $F_l = \int_{W \times \mathbb{M}} \lambda(x, m) \phi_l(x, m) f(x, m; x_0, m_0) dx dm$. Cela est vrai pour tout $l \in \{1, \dots, N\}$. On a finalement l'équation matricielle

$$(M + MKM)\Omega = F$$

et il faut alors inverser (si possible) la matrice $M + MKM$ pour obtenir Ω . On peut bien voir que Ω et F dépendent de (x_0, m_0) mais pas M et K donc il n'y a qu'une seule inversion de matrice à réaliser quelque soit le nombre de prédictions à faire.

B.2 Processus multitype

On rappelle l'équation (3) vérifiée par les poids $(w_k)_k$ de notre estimateur. Elle est de la forme

$$\lambda_k(x) \cdot \left\{ w_k(x) + \sum_{k'=1}^M \int_{W_{k'}} w_{k'}(y) \lambda_{k'}(y) K_{k,k'}(x,y) dy - f(x; x_0) \right\} = 0, \forall x \in W_k, \forall k \in \{1, \dots, M\}.$$

Soit $k \in \{1, \dots, M\}$. Soit V_k l'espace des fonctions de base associé à la fenêtre W_k . On note $N_k = \dim(V_k)$ et $\{\phi_{ik}\}_{i=1, \dots, N_k}$ est une base de V_k . On définit :

- le vecteur $\Omega_k = (\omega_{ik})_i$ par

$$w_k(x) \approx \sum_{j=1}^{N_k} \omega_{jk} \phi_{jk}(x),$$

- la matrice des masses (symétrique) M_k par

$$(M_k)_{ij} \approx \Lambda_k(\mathbb{1}_{W_k} \phi_{ik} \phi_{jk}),$$

- la matrice $\kappa^{k,k'}$ par

$$K_{k,k'}(x,y) \approx \sum_{i_1=1}^{N_k} \sum_{i_2=1}^{N_{k'}} \kappa_{i_1, i_2}^{k,k'} \phi_{i_1 k}(x) \phi_{i_2 k'}(y),$$

- et le vecteur F_k par

$$(F_k)_l = \int_{W_k} \lambda_k(x) \phi_{lk}(x) f_k(x; x_0) dx.$$

On fixe $l \in \{1, \dots, N_k\}$. En intégrant contre ϕ_{lk} , l'équation devient

$$\begin{aligned} \sum_{j=1}^{N_k} \omega_{jk} \Lambda_k(\mathbb{1}_{W_k} \phi_{jk} \phi_{lk}) + \sum_{k'=1}^M \sum_{i=1}^{N_{k'}} \omega_{ik'} \int_{W_k \times W_{k'}} \lambda_k(x) \phi_{lk}(x) \lambda_{k'}(y) \phi_{ik'}(y) K_{k,k'}(x,y) dx dy \\ = \int_{W_k} \lambda_k(x) \phi_{lk}(x) f(x; x_0) dx. \end{aligned}$$

soit

$$(M_k)_{l \bullet} \Omega_k + \sum_{k'=1}^M \sum_{i=1}^{N_{k'}} \omega_{ik'} \sum_{j_1=1}^{N_k} \sum_{j_2=1}^{N_{k'}} \kappa_{j_1, j_2}^{k,k'} \Lambda(\mathbb{1}_{W_k} \phi_{lk} \phi_{j_1 k}) \Lambda_{k'}(\mathbb{1}_{W_{k'}} \phi_{ik'} \phi_{j_2 k'}) = (F_k)_l.$$

On peut écrire plus simplement

$$(M_k)_{l\bullet}\Omega_k + \sum_{k'=1}^M (M_k)_{l\bullet}\kappa^{k,k'} M_{k'}\Omega_{k'} = (F_k)_l.$$

On a cette égalité pour tout $l \in \{1, \dots, N_k\}$ donc on a l'égalité matricielle

$$M_k\Omega_k + \sum_{k'=1}^M M_k\kappa^{k,k'} M_{k'}\Omega_{k'} = F_k.$$

Finalement, on a M équations matricielles de cette forme pour M matrices inconnues $\Omega_1, \dots, \Omega_M$. On peut réécrire les M équations matricielles en une seule équation matricielle en considérant des matrices par blocs, i.e.

$$\text{Diag}(M)\Omega + \text{Diag}(M)\kappa\text{Diag}(M)\Omega = F,$$

où

$$\begin{aligned} \text{Diag}(M) &= \begin{pmatrix} M_1 & 0 & \dots & 0 & 0 \\ 0 & M_2 & \dots & 0 & 0 \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & \dots & M_{M-1} & 0 \\ 0 & 0 & \dots & 0 & M_M \end{pmatrix} && \in (\mathbb{R}^{N_1+\dots+N_M})^{\otimes 2}, \\ \kappa &= \begin{pmatrix} \kappa^{1,1} & \dots & \kappa^{1,M} \\ \vdots & \ddots & \vdots \\ \kappa^{M,1} & \dots & \kappa^{M,M} \end{pmatrix} && \in (\mathbb{R}^{N_1+\dots+N_M})^{\otimes 2}, \\ \Omega &= \begin{pmatrix} \Omega_1 \\ \vdots \\ \Omega_M \end{pmatrix} && \in \mathbb{R}^{N_1+\dots+N_M}, \\ F &= \begin{pmatrix} F_1 \\ \vdots \\ F_M \end{pmatrix} && \in \mathbb{R}^{N_1+\dots+N_M}. \end{aligned}$$

On peut finalement obtenir les fonctions de poids par

$$\Omega = AL^{-1}F$$

où

$$AL = \begin{pmatrix} M_1 + M_1\kappa^{1,1}M_1 & M_1\kappa^{1,2}M_2 & \dots & M_1\kappa^{1,M}M_M \\ M_2\kappa^{2,1}M_1 & M_2 + M_2\kappa^{2,2}M_2 & \dots & M_2\kappa^{2,M}M_M \\ \vdots & \vdots & \ddots & \vdots \\ M_M\kappa^{M,1}M_1 & M_M\kappa^{M,2}M_2 & \dots & M_M + M_M\kappa^{M,M}M_M \end{pmatrix}.$$

On rappelle qu'on a également la condition suivante

$$\begin{aligned} \frac{1}{\Lambda_{k_1}(W_{k_1})} & \left[\lambda_{k_0}(x_0) \int_{W_{k_1}} \lambda_{k_1}(x) (g_{k_1,k_0}(x, x_0) - 1) dx \right. \\ & - \sum_{k'=1}^M \int_{W_{k_1} \times W_{k'}} w_{k'}(x') \lambda_{k_1}(x) \lambda_{k'}(x') (g_{k_1,k'}(x, x') - 1) dx dx' \\ & \left. - \int_W w_{k_1}(x) \lambda_{k_1}(x) dx \right] = \frac{1}{\Lambda_{k_2}(W_{k_2})} \left[\lambda_{k_0}(x_0) \int_{W_{k_2}} \lambda_{k_2}(x) (g_{k_2,k_0}(x, x_0) - 1) dx \right. \\ & - \sum_{k'=1}^M \int_{W_{k_2} \times W_{k'}} w_{k'}(x') \lambda_{k_2}(x) \lambda_{k'}(x') (g_{k_2,k'}(x, x') - 1) dx dx' \\ & \left. - \int_{W_{k_2}} w_{k_2}(x) \lambda_{k_2}(x) dx \right] \end{aligned}$$

pour tous $k_1, k_2 \in \{1, \dots, M\}$. Malheureusement, cette condition est resté inutilisée en pratique.

B.3 Nouvel estimateur

On rappelle l'équation qui est sous la forme

$$\begin{aligned} \lambda(x_1) \lambda(x_2) & \left[p(x_1, x_2) g_2(x_1, x_2) + \int_W (p(x_1, x_3) + p(x_2, x_3)) \lambda(x_3) g_3(x_1, x_2, x_3) dx_3 \right. \\ & + \int_{W^2} \lambda(x_3) \lambda(x_4) p(x_3, x_4) K_4(x_1, x_2, x_3, x_4) dx_3 dx_4 \\ & \left. - f(x_1, x_2; x_0) \right] = 0. \end{aligned}$$

Soit V_h l'espace d'approximation, sous-espace de dimension finie de $L^2(W)$. On note $N = \dim(V_h)$ et $\{\phi_i\}_{i=1, \dots, N}$ une base de V_h .

On écrit l'approximation

des poids

$$p(x_1, x_2) \approx \sum_{1 \leq i_1, i_2 \leq N} P_{i_1, i_2} \phi_{i_1}(x_1) \phi_{i_2}(x_2),$$

de la fonctions de corrélation de paire

$$g_2(x_1, x_2) \approx \sum_{1 \leq i_1, i_2 \leq N} D_{i_1, i_2} \phi_{i_1}(x_1) \phi_{i_2}(x_2),$$

de la fonction de corrélation d'ordre 3

$$g_3(x_1, x_2, x_3) \approx \sum_{1 \leq i_1, i_2, i_3 \leq N} T_{i_1, i_2, i_3} \phi_{i_1}(x_1) \phi_{i_2}(x_2) \phi_{i_3}(x_3),$$

et du noyau

$$K_4(x_1, x_2, x_3, x_4) \approx \sum_{1 \leq i_1, i_2, i_3, i_4 \leq N} Q_{i_1, i_2, i_3, i_4} \phi_{i_1}(x_1) \phi_{i_2}(x_2) \phi_{i_3}(x_3) \phi_{i_4}(x_4).$$

On fixe $l_1, l_2 \in \{1, \dots, N\}$. En intégrant contre $\phi_{l_1}(x_1) \phi_{l_2}(x_2)$, l'équation devient

$$\begin{aligned} & \sum_{1 \leq i_1, i_2, j_1, j_2 \leq N} P_{j_1, j_2} D_{i_1, i_2} \Lambda(\mathbb{1}_W \phi_{j_1} \phi_{l_1} \phi_{i_1}) \Lambda(\mathbb{1}_W \phi_{j_2} \phi_{l_2} \phi_{i_2}) \\ & + \sum_{1 \leq i_1, i_2, i_3, j_1, j_2 \leq N} P_{j_1, j_2} T_{i_1, i_2, i_3} \Lambda(\mathbb{1}_W \phi_{l_1} \phi_{j_1} \phi_{i_1}) \Lambda(\mathbb{1}_W \phi_{l_2} \phi_{i_2}) \Lambda(\mathbb{1}_W \phi_{j_2} \phi_{i_3}) \\ & + \sum_{1 \leq i_1, i_2, i_3, j_1, j_2 \leq N} P_{j_1, j_2} T_{i_1, i_2, i_3} \Lambda(\mathbb{1}_W \phi_{l_1} \phi_{i_1}) \Lambda(\mathbb{1}_W \phi_{l_2} \phi_{j_1} \phi_{i_2}) \Lambda(\mathbb{1}_W \phi_{j_2} \phi_{i_3}) \\ & + \sum_{1 \leq i_1, i_2, i_3, i_4, j_1, j_2 \leq N} P_{j_1, j_2} Q_{i_1, i_2, i_3, i_4} \Lambda(\mathbb{1}_W \phi_{l_1} \phi_{i_1}) \Lambda(\mathbb{1}_W \phi_{l_2} \phi_{i_2}) \Lambda(\mathbb{1}_W \phi_{i_3} \phi_{j_1}) \Lambda(\mathbb{1}_W \phi_{i_4} \phi_{j_2}) \\ & = \int_{W^2} \lambda(x_1) \lambda(x_2) \phi_{l_1}(x_1) \phi_{l_2}(x_2) f(x_1, x_2; x_0) dx_1 dx_2. \end{aligned}$$

On définit les “matrices de masses”

$$M_{i,j}^{(2)} = \Lambda(\mathbb{1}_W \phi_i \phi_j),$$

et

$$M_{i,j,k}^{(3)} = \Lambda(\mathbb{1}_W \phi_i \phi_j \phi_k).$$

On définit également la matrice F par

$$F_{l_1, l_2} = \int_{W^2} \lambda(x_1) \lambda(x_2) \phi_{l_1}(x_1) \phi_{l_2}(x_2) f(x_1, x_2; x_0) dx_1 dx_2.$$

On peut remarquer que tous les tenseurs définis précédemment sont symétriques, i.e. ils ont la même valeur lorsqu'on permute les indices. On peut finalement réécrire

$$\begin{aligned} F_{l_1, l_2} &= \sum_{1 \leq i_1, i_2, j_1, j_2 \leq N} P_{j_1, j_2} D_{i_1, i_2} M_{i_1, j_1, l_1}^{(3)} M_{i_2, j_2, l_2}^{(3)} \\ &+ \sum_{1 \leq i_1, i_2, i_3, j_1, j_2 \leq N} P_{j_1, j_2} T_{i_1, i_2, i_3} M_{l_1, i_1, j_1}^{(3)} M_{l_2, i_2}^{(2)} M_{i_3, j_2}^{(2)} \\ &+ \sum_{1 \leq i_1, i_2, i_3, j_1, j_2 \leq N} P_{j_1, j_2} T_{i_1, i_2, i_3} M_{l_1, i_1}^{(2)} M_{i_2, j_1, l_2}^{(3)} M_{i_3, j_2}^{(2)} \\ &+ \sum_{1 \leq i_1, i_2, i_3, i_4, j_1, j_2 \leq N} P_{j_1, j_2} Q_{i_1, i_2, i_3, i_4} M_{l_1, i_1}^{(2)} M_{l_2, i_2}^{(2)} M_{i_3, j_1}^{(2)} M_{i_4, j_2}^{(2)}. \end{aligned}$$

En définissant les bonnes notations, on obtient

$$\begin{aligned}
F_{l_1, l_2} &= \sum_{1 \leq j_1, j_2 \leq N} D(M_{j_1, l_1}^{(3)}, M_{j_2, l_2}^{(3)}) P_{j_1, j_2} \\
&+ \sum_{1 \leq j_1, j_2 \leq N} T(M_{l_1, j_1}^{(3)}, M_{l_2}^{(2)}, M_{j_2}^{(2)}) P_{j_1, j_2} \\
&+ \sum_{1 \leq j_1, j_2 \leq N} T(M_{l_1}^{(2)}, M_{j_1, l_2}^{(3)}, M_{j_2}^{(2)}) P_{j_1, j_2} \\
&+ \sum_{1 \leq j_1, j_2 \leq N} Q(M_{l_1}^{(2)}, M_{l_2}^{(2)}, M_{j_1}^{(2)}, M_{j_2}^{(2)}) P_{j_1, j_2}.
\end{aligned}$$

On note $L = (l_1, l_2)$ et $J = (j_1, j_2)$. Toujours avec les bonnes notations, on a

$$\begin{aligned}
&\left(D \left(M_{j_1, l_1}^{(3)}, M_{j_2, l_2}^{(3)} \right) + T \left(M_{l_1, j_1}^{(3)}, M_{l_2}^{(2)}, M_{j_2}^{(2)} \right) + T \left(M_{l_1}^{(2)}, M_{j_1, l_2}^{(3)}, M_{j_2}^{(2)} \right) \right. \\
&\quad \left. + Q \left(M_{j_1}^{(2)}, M_{l_1}^{(2)}, M_{j_2}^{(2)}, M_{l_2}^{(2)} \right) \right)_{L \bullet} P = F_L.
\end{aligned}$$

Cela reste valable pour tout L donc avec

$$A = \left(D \left(M_{j_1, l_1}^{(3)}, M_{j_2, l_2}^{(3)} \right) + T \left(M_{l_1, j_1}^{(3)}, M_{l_2}^{(2)}, M_{j_2}^{(2)} \right) + T \left(M_{l_1}^{(2)}, M_{j_1, l_2}^{(3)}, M_{j_2}^{(2)} \right) + Q \left(M_{j_1}^{(2)}, M_{l_1}^{(2)}, M_{j_2}^{(2)}, M_{l_2}^{(2)} \right) \right)_{L, J},$$

on a

$$A \times P = F.$$

C Équivalence

Dans la partie A, on admettait l'existence d'une fonction de poids minimisant l'erreur quadratique sous la contrainte de non-biais puis on déterminait quelle équation vérifiait cette fonction de poids, si elle existe. Or, en pratique on va directement résoudre l'équation intégrale mais il est possible que cela donne une autre fonction de poids qui ne donne pas le BLUP. En s'inspirant de ce qui est fait par Mrkvička et Molchanov [5], on prouve que si l'équation intégrale admet une solution alors c'est un minimum global sur l'ensemble d'intérêt et réciproquement. On ne présente ici que le cas d'un processus à marque continue mais on peut suivre la même démarche pour obtenir le même résultat dans les autres cas.

On rappelle qu'on observe un processus à marque continue sur une fenêtre W (borélien de mesure finie). Par hypothèse on a

$$\alpha^{(2)}(d(x_1, m_1, x_2, m_2)) = \lambda(x_1, m_1) \lambda(x_2, m_2) g_2(x_1, m_1, x_2, m_2) dx_1 dm_1 dx_2 dm_2$$

avec g_2 continues. On note

$$L^2(W \times \mathbb{R}, \mathcal{B}(W \times \mathbb{R}), \Lambda) = \{w \in \mathcal{B}(W \times \mathbb{R}) : \Lambda(w^2) < \infty\}$$

et pour $w \in L^2(W \times \mathbb{R}, \mathcal{B}(W \times \mathbb{R}), \Lambda)$ on note

$$\hat{\lambda}_w = \sum_{[x,m] \in \phi_W}^{\neq} w(x,m)$$

pour x_0, m_0 fixé. On définit la fonctionnelle J_ϕ par

$$J_\phi(w) = \mathbb{E} \left[(\hat{\lambda}_w - \lambda(x_0|\phi_W))^2 \right] - \mathbb{E} [\lambda(x_0|\phi_W)^2] \geq -\mathbb{E} [\lambda(x_0|\phi_W)^2]. \quad (12)$$

On a

$$\begin{aligned} J_\phi(w) &= \mathbb{E} [\hat{\lambda}_w^2] - 2\mathbb{E} [\hat{\lambda}_w \lambda(x_0|\phi_W)] \\ &= \int_{W \times \mathbb{R}} w^2(x,m) \lambda(x,m) dx dm \\ &\quad + \int_{W^2 \times \mathbb{R}^2} w(x,m) w(y,n) \lambda(x,m) \lambda(y,n) g(x,m,y,n) dx dm dy dn \\ &\quad - 2\lambda(x_0, m_0) \int_{W \times \mathbb{R}} w(x,m) \lambda(x,m) g(x,m, x_0, m_0) dx dm \\ &= B_\phi(w) + L_\phi(w), \end{aligned}$$

où $B_\phi = \mathbb{E} [\hat{\lambda}_w^2]$ et $L_\phi = -2 [\hat{\lambda}_w \lambda(x_0|\phi_W)]$ sont respectivement la partie bilinéaire et la partie linéaire de J_ϕ . On note alors

$$\mathcal{F}_\phi = \left\{ w \in L^2(W \times \mathbb{R}) : |B_\phi(w)| < \infty \text{ et } |L_\phi(w)| < \infty \right\}.$$

De plus, on définit

$$\mathcal{G}_\phi = \left\{ w \in \mathcal{F}_\phi : \mathbb{E} [\hat{\lambda}_w] = \int_{W \times \mathbb{R}} w(x,m) \lambda(x,m) dx dm = \lambda(x_0, m_0) \right\}$$

et

$$\mathcal{H}_\phi = \left\{ w \in \mathcal{F}_\phi : \int_{W \times \mathbb{R}} w(x,m) \lambda(x,m) dx dm = 0 \right\}.$$

Minimiser $\mathbb{E} [(\hat{\lambda}_w - \lambda(x_0|\phi_W))^2]$ revient à minimiser $J_\phi(w)$. On remarque que B_ϕ est une forme bilinéaire définie positive, donc B_ϕ est convexe sur \mathcal{F}_ϕ et L_ϕ est linéaire donc également convexe. Finalement, J_ϕ est convexe sur \mathcal{F}_ϕ et J_ϕ est minorée par (12). On sait alors que tout minimum local de J_ϕ sur \mathcal{F}_ϕ est un minimum global.

Lemme 1. *La dérivée directionnelle $\delta J_\phi(w, v)$ de $J_\phi(w)$ en $w \in \mathcal{F}_\phi$ et dans la direction v existe, est finie et est donnée par*

$$\begin{aligned}\delta J_\phi(w, v) &= 2 \int_{W \times \mathbb{R}} w(x, m) v(x, m) \lambda(x, m) dx dm \\ &\quad + 2 \int_{W^2 \times \mathbb{R}^2} v(x, m) w(y, n) \lambda(x, m) \lambda(y, n) g(x, m, y, n) dx dm dy dn \\ &\quad - 2\lambda(x_0) \int_{W \times \mathbb{R}} v(x, m) \lambda(x, m) g(x_0, m_0, x, m) dx dm.\end{aligned}$$

La définition de dérivée directionnelle (voir [8]) donne

$$\begin{aligned}\delta J_\phi(w, v) &= \frac{\partial}{\partial \epsilon} J_\phi(w + \epsilon v)|_{\epsilon=0} \\ &= 2\mathbb{E} \left[\sum_{[x, m], [y, n] \in \phi_W} w(x, m) v(y, n) \right] \\ &\quad - 2\lambda(x_0) \int_{W \times \mathbb{R}} v(x, m) \lambda(x, m) g(x_0, m_0, x, m) dx dm \\ &= 2 \left(\mathbb{E} [\hat{\lambda}_w \hat{\lambda}_v] - \mathbb{E} [\hat{\lambda}_v \lambda] \right)\end{aligned}$$

il reste à prouver que cette quantité est finie. Pour $w, v \in \mathcal{F}_\phi$, on a

$$\begin{aligned}\frac{1}{2} |\delta J_\phi(w, v)| &\leq \left(\mathbb{E} [\hat{\lambda}_w^2] \mathbb{E} [\hat{\lambda}_v^2] \right)^{1/2} + |\mathbb{E} [\hat{\lambda}_v \lambda]| \\ &= (B_\phi(w) B_\phi(v))^{1/2} + \frac{1}{2} |L_\phi(v)| < \infty.\end{aligned}$$

Théorème 2. *$J_\phi(w)$, $w \in \mathcal{G}_\phi$, est minimal pour $w = w_{\min}$ si et seulement si*

$$\begin{aligned}\lambda(x, m) \left\{ w_{\min}(x, m) + \int_{W \times \mathbb{R}} w_{\min}(y, n) \lambda(y, n) g(x, m, y, n) dy dn \right. \\ \left. - \lambda(x_0) g(x_0, m_0, x, m) - C \right\} = 0\end{aligned}$$

pour presque tout $x \in W$, où C est une constante telle que $\int_{W \times \mathbb{R}} w(x, m) \lambda(x, m) dx dm = \lambda(x_0, m_0)$.

Il faut montrer que $J_\phi(w)$ atteint son minimum sur \mathcal{G}_ϕ en $w = w_{\min}$ si et seulement si $\delta J_\phi(w_{\min}, v) = 0$ pour tout $v \in \mathcal{H}_\phi$. On prouve cela par analyse-synthèse.

Analyse.

Soit $A_+ \subset W \times \mathbb{R}$ un borélien tel que $\Lambda(A_+) = \Lambda(W \times \mathbb{R})/2 = \Lambda_g(W)/2$. On définit

également $A_- = W \times \mathbb{R} \setminus A_+$. On fixe $v \in \mathcal{H}_\phi$ définie par $v(x, m) = \mathbb{1}_{A_+}(x, m) - \mathbb{1}_{A_-}(x, m)$. On a alors,

$$\begin{aligned} 0 &= \frac{1}{2} \delta J_\phi(w_{\min}, v) = \int_{A_+} \lambda(x, m) w_{\min}(x, m) dx dm - \int_{A_-} \lambda(x, m) w_{\min}(x, m) dx dm \\ &\quad + \int_{A_+} \int_{W \times \mathbb{R}} \lambda(x, m) w_{\min}(x, m) \lambda(y, n) g(x, m, y, n) dy dn dx dm \\ &\quad - \int_{A_-} \int_{W \times \mathbb{R}} \lambda(x, m) w_{\min}(x, m) \lambda(y, n) g(x, m, y, n) dy dn dx dm \\ &\quad - \lambda(x_0) \left[\int_{A_+} \lambda(x, m) g(x_0, m_0, x, m) dx dm - \int_{A_-} \lambda(x, m) g(x_0, m_0, x, m) dx dm \right] \\ &= \int_{A_+} \lambda(x, m) h(x, m) dx dm - \int_{A_-} \lambda(x, m) h(x, m) dx dm \end{aligned}$$

où

$$\begin{aligned} h(x, m) &= w_{\min}(x, m) + \int_{W \times \mathbb{R}} w_{\min}(x, m) \lambda(y, n) g(x, m, y, n) dy dn \\ &\quad - \lambda(x_0) g(x_0, m_0, x, m). \end{aligned}$$

L'égalité reste vraie pour tout borélien A_+ tel que $\Lambda(A_+) = \Lambda(W \times \mathbb{R})/2$ et donc il existe une constante C tel que $h(x, m) = C$ pour presque tout $x, m \in W \times \mathbb{R}$ par rapport à la mesure Λ . Synthèse.

On a pour $v \in \mathcal{F}_\phi$

$$\delta J_\phi(w, v) = 2 \int_{W \times \mathbb{R}} \lambda(x, m) h(x, m) v(x, m) dx dm.$$

Prenons w tel que pour presque tout $x, m \in W \times \mathbb{R}$, $h(x, m) = C$ presque partout par rapport à Λ , où C est la constante obtenue dans la partie précédente (de manière unique). Alors on a

$$\delta J_\phi(w, v) = 2C \int_{W \times \mathbb{R}} \lambda(x, m) v(x, m) dx dm = 0,$$

car $v \in \mathcal{F}_\phi$. Toutes les dérivées directionnelles de J_ϕ sont nulles en w st un extremum de J_ϕ . La convexité de J_ϕ donne bien que w en est un minimiseur sur \mathcal{G}_ϕ .

On obtient la valeur de C en intégrant $\lambda(x, m) h(x, m)$ sur $W \times \mathbb{R}$, sachant $\int_{W \times \mathbb{R}} w_{\min}(x, m) \lambda(x, m) dx dm = \lambda(x_0, m_0)$. On retrouve exactement l'équation obtenue en partie A avec la méthode des multiplicateurs de Lagrange.

D Existence minimiseur

On regarde uniquement le cas d'un processus à marque continue. On suppose comme précédemment que $\Lambda(W \times \mathbb{R}) > 0$. On définit la mesure positive β sur $W \times \mathbb{R}$, muni de la tribu borélienne, par

$$\beta(A) = \mu^{(2)}(A \cap (W \times \mathbb{R}) \times A \cap (W \times \mathbb{R})) = \mathbb{E} \left[\left(\sum_{[x,m] \in \phi_W} \mathbb{1}_A(x,m) \right)^2 \right].$$

L'espace $L^2(\beta, W \times \mathbb{R})$ est un espace de Banach sur \mathbb{R} donc complet.

- On note E l'espace des variables aléatoires réelles qui possèdent un moment d'ordre 2 et qui sont mesurables par rapport à $\sigma(\phi_W)$. On munit E du produit scalaire

$$\langle, \rangle : \left\{ \begin{array}{ll} E \times E & \longrightarrow \mathbb{R} \\ (X,Y) & \longmapsto \mathbb{E}[XY] \end{array} \right. ,$$

qui en fait un espace de Banach.

- On définit l'application linéaire \mathcal{F} par

$$\mathcal{F} : \left\{ \begin{array}{ll} L^2(\beta, W \times \mathbb{R}) & \longrightarrow E \\ w & \longmapsto \sum_{[x,m] \in \phi_W} w(x,m) \end{array} \right. .$$

\mathcal{F} est une isométrie. On définit également l'application linéaire

$$e : \left\{ \begin{array}{ll} E & \longrightarrow \mathbb{R} \\ X & \longmapsto \mathbb{E}[X] \end{array} \right. .$$

- On note

$$C = e^{-1}(\{\lambda(x_0, m_0)\}) \cap \mathcal{F}(L^2(\beta, W \times \mathbb{R})),$$

et

$$C_0 = \mathcal{F}^{-1}(C).$$

On peut vérifier que

- E est un espace préhilbertien réel,
- C est un ensemble convexe par linéarité de L et de \mathcal{F} ,
- C est non vide (pour $w = \lambda(x_0, m_0)/\Lambda(W \times \mathbb{R})$, $\mathcal{F}(w) \in C$),

- et C est fermé car c'est l'intersection de deux fermés. En effet, $e^{-1}(\{\lambda(x_0, m_0)\})$ est un fermé en tant qu'image réciproque d'un fermé par une application continue. De plus, \mathcal{F} est une application linéaire d'un espace L^2 vers un espace L^2 donc $\mathcal{F}(L^2(\beta, W \times \mathbb{R}))$ est fermé.

Par le théorème de projection sur un convexe fermé, il existe une (unique) application P_C de E dans C qui à une variable aléatoire $X \in E$ associe la variable aléatoire $P_C(X)$ de C , tel que la distance de X à C soit égale à celle de X à $P_C(X)$. On pose $X_0 = \lambda(x_0, m_0 | \phi_W) \in E$, alors la variable aléatoire $P_C(X_0)$ est l'unique point vérifiant

$$\forall Y \in C, \|X_0 - P_C(X_0)\| \leq \|X_0 - Y\|$$

donc il existe $w^* \in L^2(\beta, W \times \mathbb{R})$ tel que $P_C(X_0) = \mathcal{F}(w^*)$ et on a

$$\forall v \in C_0, \mathbb{E} \left[(\lambda(x_0, m_0 | \phi_W) - \mathcal{F}(w^*))^2 \right] \leq \mathbb{E} \left[(\lambda(x_0, m_0 | \phi_W) - \mathcal{F}(v))^2 \right].$$