



**HAL**  
open science

## Sous-ensembles maximaux cohérents multidimensionnels rendus faciles : illustration sur un problème d'estimation

Loïc Adam, Sébastien Destercke

### ► To cite this version:

Loïc Adam, Sébastien Destercke. Sous-ensembles maximaux cohérents multidimensionnels rendus faciles : illustration sur un problème d'estimation. 31èmes Rencontres Francophones sur la Logique Floue et ses Applications (LFA 2022), Oct 2022, Toulouse, France. hal-04336838

**HAL Id: hal-04336838**

**<https://hal.science/hal-04336838>**

Submitted on 11 Dec 2023

**HAL** is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

# Sous-ensembles maximaux cohérents multidimensionnels rendus faciles : illustration sur un problème d'estimation

Loïc Adam<sup>1</sup>

Sébastien Destercke<sup>1</sup>

<sup>1</sup> UMR CNRS 7253 Heudiasyc, Sorbonne Université, Université de Technologie de Compiègne  
CS 60319 - 60203 Compiègne cedex, France, {loic.adam, sebastien.destercke}@hds.utc.fr

## Résumé :

Fusionner des informations incertaines afin d'obtenir une information synthétique est une tâche difficile, notamment quand les informations sont inconsistantes entre elles. Une solution particulièrement intéressante pour résoudre ce conflit se focalise sur les sous-ensembles maximaux cohérents (MCS). Néanmoins, énumérer les différents MCSs est un problème combinatoire difficile, limitant leur utilisation en pratique. Dans ce papier, nous nous intéressons au cas où les différentes informations sont des ensembles ou des polytopes multi-dimensionnels. Bien que le problème reste difficile pour des polytopes génériques, nous montrons qu'il peut être résolu plus efficacement pour des hyperrectangles. Nous illustrons ensuite comment notre méthode peut être utilisée pour estimer des modèles linéaires en présence de valeur aberrantes ou de modèles incorrectement spécifiés.

## Mots-clés :

Fusion d'informations, sous-ensemble maximaux cohérents, régression linéaire, données imprécises.

## Abstract:

Fusing uncertain pieces of information to obtain a synthetic estimation when those are inconsistent is a difficult task. A particularly appealing solution to solve such conflict or inconsistency is to look at maximal coherent subsets of sources (MCS), and to concentrate on those. Yet, enumerating MCS is a difficult combinatorial task in general, making the use of MCS limited in practice. In this paper, we are interested in the case where the pieces of information are multi-dimensional sets or polytopes. While the problem remains difficult for general polytopes, we show that it can be solved more efficiently for hyperrectangles. We then illustrate how such an approach could be used to estimate linear models in the presence of outliers or in the presence of misspecified model.

## Keywords:

Information fusion, maximal coherent subsets, linear regression, imprecise data.

## 1 Introduction

Ce papier s'intéresse au problème de fusion d'informations [5]. L'objectif est de combiner les informations de différentes sources, pouvant être incomplètes ou incertaines, pour obtenir une nouvelle information synthétique avec moins d'incertitude et de redondance. Dans ce problème, limiter au minimum le conflit entre

plusieurs sources contradictoires est une tâche particulièrement difficile. Cela reste néanmoins une tâche souvent nécessaire. De plus, analyser les causes de la contradiction ou encore expliquer si possible l'apparition du conflit sont aussi des tâches importantes pour analyser la situation.

[6] passe en revue les différentes méthodes pour résoudre un conflit durant l'agrégation, par exemple en utilisant des données additionnelles comme la fiabilité. Néanmoins, obtenir ces données est parfois difficile ou impossible. Une approche pour gérer les contradictions sans information supplémentaire est basée sur les sous-ensembles maximaux cohérents (MCS), qui sont des groupes aussi grands que possible de sources consistantes. Les MCSs ont été étudiés dans le passé en logique [9] et dans des contextes numériques [3]. Ici, nous illustrons leur application aux problèmes d'estimation.

Détecter et énumérer les différents MCSs sont des problèmes d'optimisation difficiles. Il existe néanmoins des situations où le calcul est traitable, notamment en présence d'intervalles [4]. Dans ce papier, nous montrons qu'il est possible d'étendre cette exception aux hyperrectangles, pouvant être utilisés comme approximations.

Dans la Section 2, nous introduisons davantage les sous-ensembles maximaux cohérents et nous expliquons pourquoi il est difficile de tous les énumérer. Nous montrons ensuite dans la Section 3 qu'il est facile de le faire pour un ensemble d'hyperrectangles alignés avec les axes. Nous présenterons alors dans la section 4 comment approcher des polytopes par des hyperrectangles. Enfin, dans la Section 5 nous illustrons notre approche sur des problèmes de régression linéaire.

## 2 Sous-ensembles Maximaux Cohérents

Comme mentionné précédemment, les sous-ensembles cohérents maximaux sont en théorie une bonne solution pour gérer le conflit entre plusieurs sources d'information, visant à obtenir plusieurs ensembles disjoints tel que chaque ensemble contient le plus de sources cohérentes entre elles possible. Néanmoins, énumérer les différents MCSs est en général un problème combinatoire difficile.

### 2.1 Définition générale

Nous considérons un problème multi-sources avec un ensemble  $\mathcal{S} = \{S_1, \dots, S_N\}$  de sources d'information. Les informations que chaque source présente sont définies sur un espace d'information  $\mathcal{X}$  pour lesquels l'intersection  $\cap$  est bien définie. Un sous-ensemble cohérent maximal  $c \subseteq \{1, \dots, N\}$  est une liste d'indices de sources telle que  $\bigcap_{i \in c} S_i \neq \emptyset$  et, pour n'importe quel  $j \notin c$ , nous avons  $\bigcap_{i \in c} S_i \cap S_j = \emptyset$ . Cela équivaut à dire que le sous-ensemble de sources  $c$  est cohérent et maximal.

**Exemple 1.** Soit un ensemble de sources  $\mathcal{S} = \{S_1, \dots, S_4\}$  comme présenté sur la Figure 1. Spécifiquement dans cet exemple, les sources présentent leurs données sous forme d'intervalles.  $\{1, 2\}$  est un sous-ensemble cohérent, mais pas un sous-ensemble cohérent maximal, comme il est possible d'ajouter  $S_3$  tout en ayant une intersection non-vide.  $\{1, 2, 3\}$  est un sous-ensemble cohérent maximal, ainsi que  $\{1, 2, 4\}$ . Néanmoins,  $\{1, 2, 3, 4\}$  n'est pas un sous-ensemble cohérent, car  $S_4$  est en contradiction avec  $S_3$  (intersection vide entre les deux).

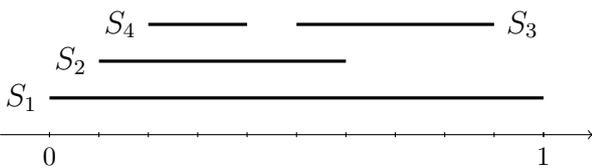


Figure 1 – Visualisation de sources  $S_i \in [0, \dots, 1]$  pas totalement cohérentes.

### 2.2 MCSs de polytopes

Nous nous intéressons aux polytopes convexes, qui délimitent par des hyperplans affines des ensembles convexes dans l'espace Euclidien en  $d$ -dimensions  $\mathbb{R}^d$ . Ces ensembles convexes particuliers, même en tant qu'approximation des informations, permettent d'utiliser les outils de la programmation mathématique et sont donc d'un grand intérêt pratique. Par la suite, nous représentons toutes les sources d'information  $S_i$  par des polytopes notés  $P_i$ . Les polytopes convexes ont deux représentations :

- La V-représentation, à partir des points extrêmes, donnant une enveloppe convexe.
- La H-représentation, à partir d'un système d'inéquations linéaires, donnant une intersection de demi-plans.

Déterminer l'intersection de deux polytopes définis par leurs V-représentation est un problème NP-difficile [12], alors que le même problème peut être résolu facilement si les polytopes sont définis par leurs H-représentations. Passer d'une représentation à une autre est un problème NP-complet<sup>1</sup>, avec des algorithmes efficaces pour des H-polytopes spécifiques [8].

Trouver un unique MCS quelconque en H-représentation est par conséquent facile : il suffit de rajouter des polytopes un par un, et nous avons un MCS dès qu'il n'est plus possible de rajouter un polytope sans avoir une intersection vide. Vérifier si un ensemble est un MCS est aussi facile : il suffit de vérifier que l'intersection des différents H-polytopes est non-vide et maximale. Néanmoins, énumérer tous les MCSs d'un ensemble de polytopes  $\mathcal{P} = \{P_1, \dots, P_N\}$  nécessite dans le pire cas de considérer tous les sous-ensembles de  $\mathcal{P}$ , soit  $2^N$  ensembles pour lesquels nous devons vérifier si l'intersection est non-vide (et maximale), ce qui est difficile pour un grand nombre de polytopes. Décider si un ensemble  $\mathcal{P}$  forme moins de  $k$  MCS est donc un problème de décision NP-complet.

1. Sinon les deux problèmes seraient faciles.

Le calcul est néanmoins traitable pour les intervalles, où il est possible d'énumérer les MCSs en temps polynomial [4]. Dans la Section 3, nous démontrons qu'il existe aussi un algorithme efficace pour énumérer les différents MCSs d'hyperrectangles alignés avec les axes.

### 3 Énumérer les MCSs d'hyperrectangles alignés avec les axes

Énumérer tous les MCSs d'un ensemble de polytopes est un problème difficile dans le cas général. Cependant, comme rappelé en Section 2, il existe des algorithmes polynomiaux dans le cas des intervalles. Nous montrons dans cette section que ces résultats peuvent être appliqués sur un ensemble  $\mathcal{H}$  d'hyperrectangles alignés avec les axes, afin de déterminer l'ensemble de MCSs  $\mathcal{C}_{\mathcal{H}}$

Nous notons par  $I_{H_i}^d \in \mathbb{R}$  la projection d'un hyperrectangle  $H_i$  sur la  $d^e$  dimension de l'espace  $\mathbb{R}^D$ . Nous avons alors une équivalence importante entre l'intersection des hyperrectangles et l'intersection de leurs projections :

**Proposition 1.** *Soit un ensemble  $\mathcal{H} = \{H_1, \dots, H_N\}$  d'hyperrectangles alignés avec les axes dans l'espace  $\mathbb{R}^D$  et leurs projections  $I_{H_i}^d \in \mathbb{R}$  sur les différentes dimensions  $d \in \{1, \dots, D\}$ , nous avons alors :*

$$\bigcap_{H_i \in \mathcal{H}} H_i \neq \emptyset \iff \bigcap_{\mathcal{H}} I_{H_i}^d \neq \emptyset \forall d \in \{1, \dots, D\}. \quad (1)$$

*Démonstration.* Pour voir l'équivalence, il est suffisant d'observer que  $\times_{d=1}^D \bigcap_{i=1}^N I_{H_i}^d = \bigcap_{H_i \in \mathcal{H}} H_i$ . En particulier, cela indique que n'importe quel point  $x \in \mathbb{R}^D$  tel que ses projections  $x^d \in \bigcap_{i=1}^N I_{H_i}^d$  pour tout  $d \in \{1, \dots, D\}$  seront aussi dans  $\bigcap_{H_i \in \mathcal{H}} H_i$ . Notons que cela n'est vrai que pour les hyperrectangles alignés.  $\square$

Le corollaire suivant, qui n'est que la négation de la Proposition 1, est utile pour les démonstrations ultérieures :

**Corollaire 1.**

$$\bigcap_{\mathcal{H}} H_i = \emptyset \iff \exists d \in \{1, \dots, D\}, \bigcap_{\mathcal{H}} I_{H_i}^d = \emptyset. \quad (2)$$

Dans la proposition suivante, nous montrons que les MCSs d'hyperrectangles alignés avec les axes peuvent être trouvés facilement en combinant les MCSs de leurs différentes projections, ce qui est faisable en temps polynomial comme les projections sont des intervalles.

**Proposition 2.** *Soit un ensemble  $\mathcal{H} = \{H_1, \dots, H_N\}$  d'hyperrectangles alignés avec les axes dans l'espace  $\mathbb{R}^D$ , son ensemble de MCSs  $\mathcal{C}_{\mathcal{H}}$  et les ensembles de MCSs  $\mathcal{C}_d$  de leurs projections sur la  $d^e$  dimension, nous avons alors :*

$$\mathcal{C}_{\mathcal{H}} = \{ \bigcap_{d=1}^D c_d \neq \emptyset \mid c_d \in \mathcal{C}_d \forall d \in \{1, \dots, D\} \}. \quad (3)$$

*Démonstration.* Montrons qu'il existe une double inclusion pour un MCS  $c_h \in \mathcal{C}_{\mathcal{H}}$  :

Nous commençons en démontrant qu'il existe  $c_i \in \{ \bigcap_{d=1}^D c_d \mid c_d \in \mathcal{C}_d \}$  tel que  $c_h \subseteq c_i$ , c.-à-d.  $c_i$  est une représentation externe de  $c_h$ . Comme nous avons  $\bigcap_{i \in c_h} H_i \neq \emptyset$  ( $c_h$  étant un MCS), la Proposition 1 nous indique que pour n'importe quelle dimension  $d$ , nous avons  $\bigcap_{i \in c_h} I_{H_i}^d \neq \emptyset$ , signifiant qu'il existe un MCS  $c_d \in \mathcal{C}_d$  tel que  $c_h \subseteq c_d$ . Comme cela est vrai pour tout  $d \in \{1, \dots, D\}$ , cela indique que  $c_h \subseteq \bigcap_{d=1}^D c_d$  pour toute collection de  $c_d$ , montrant la première inclusion.

Pour démontrer la deuxième inclusion, nous considérons un ensemble  $c_d, d \in \{1, \dots, D\}$  de MCS sur chaque dimension qui sont des approximations externes de  $c_h$  ( $c_h \subseteq c_d$ ) dont les existences sont validées par la première partie de la preuve. Nous voulons montrer que  $j \notin c_h$  implique  $j \notin \bigcap_{d=1}^D c_d$  et donc  $\bigcap_{d=1}^D c_d \subseteq c_h$ . Pour voir cela, nous considérons tout simplement l'ensemble des hyperrectangles  $H_k, k \in c_h \cup \{j\}$ . Dès lors, d'après le Corollaire 1, il existe une dimension  $d$  telle que  $\bigcap_{k \in c_h \cup \{j\}} I_{H_k}^d = \emptyset$ . Pourtant,  $\bigcap_{k \in c_h} I_{H_k}^d \neq \emptyset$  ( $c_h$  étant un MCS). Ceci montre que  $j \notin c_h$  implique  $j \notin \bigcap_{d=1}^D c_d$ .  $\square$

La Proposition 2 nous permet d’obtenir des MCSs assez facilement : nous commençons par projeter les hyperrectangles sur les différentes dimensions, afin d’obtenir des intervalles. Nous énumérons ensuite les MCSs sur chaque dimension, ce qui est de complexité polynomiale. Enfin, nous déterminons  $\mathcal{C}_{\mathcal{H}}$  comme l’ensemble des sources communes entre toutes les énumérations, c.-à-d. la combinaison conjonctive (intersection) des différentes sources. D’après la Proposition 2 et l’Équation (3), cela nous donne exactement l’ensemble des MCSs.

**Exemple 2.** Soit  $\mathcal{H} = \{H_1, \dots, H_5\}$  un ensemble de 5 rectangles alignés avec les axes. Nous cherchons à énumérer les MCSs de cet ensemble. Pour cela, nous déterminons leurs projections  $I_{H_j}^i, \forall i \in \{1, 2\}, \forall j \in \{1, \dots, 5\}$ , représentées sur la Figure 2. Nous commençons par déterminer les MCSs sur chaque axe. Sur le premier axe ( $\omega_1$ ) nous avons  $\mathcal{C}_1 = \{\{1, 2, 3\}, \{4, 5\}\}$ . Sur le deuxième axe ( $\omega_2$ ) nous avons  $\mathcal{C}_2 = \{\{2, 3, 4, 5\}, \{1, 4\}\}$ . Nous combinons alors ces sources d’information selon l’équation (3) et nous obtenons les ensembles suivants :  $\{2, 3\}, \{1\}, \{4, 5\}$ . Nous avons énuméré les MCSs de  $\mathcal{H}$ .

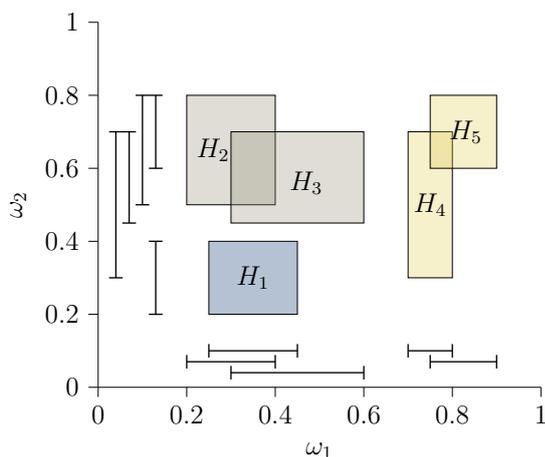


Figure 2 – Projection d’hyperrectangles sur les différents plans

Énumérer les MCSs d’hyperrectangles alignés avec les axes est donc facile. Néanmoins, il est

rare que des polytopes aient cette structure. Il est donc nécessaire d’approximer des polytopes convexes génériques, afin de pouvoir appliquer notre algorithme, comme nous allons le voir dans la Section 4.

## 4 Approximation de polytopes avec des hyperrectangles

Nous voulons approximer des polytopes convexes avec des hyperrectangles alignés avec les axes. Néanmoins, l’intersection des approximations n’a pas nécessairement les mêmes MCSs que l’intersection des polytopes initiaux. Il est donc nécessaire d’avoir les approximations les plus proches possibles de nos polytopes initiaux, voire de borner l’énumération des différents MCSs. Dans cette section, nous nous intéressons à l’approximation interne maximale, et à l’approximation externe minimale, et montrons en quoi elles sont intéressantes pour lister les différents MCSs.

### 4.1 Hyperrectangle inscrit d’hypervolume maximum aligné avec les axes

L’hyperrectangle inscrit d’hypervolume maximum aligné avec les axes (MVAIR) est un hyperrectangle  $H$  totalement contenu dans un polytope  $P$ , c.-à-d.  $H \subseteq P$ , dont on cherche à maximiser son hypervolume avec comme contrainte que ses arêtes doivent être parallèles aux axes (de coordonnées Cartésiennes). C’est une approximation interne de notre polytope qui est maximale au sens des contraintes. Un exemple est visible sur la Figure 3.

Dans le cas général, trouver un hyperrectangle inscrit d’hypervolume maximum nécessite de résoudre un problème d’optimisation non convexe avec un nombre exponentiel de contraintes [2]. Néanmoins, grâce à la contrainte d’alignement avec les axes, le problème d’optimisation devient convexe et est donc solvable relativement facilement.

Remarquons qu’un MVAIR n’est pas forcément unique : un simple rectangle avec  $l \neq L$  a un

nombre infini de MVAIR, obtenus par translation le long de la longueur.

## 4.2 Hyperrectangle englobant d'hypervolume minimum aligné avec les axes

L'hyperrectangle englobant d'hypervolume minimum aligné avec les axes (MVABH) est un hyperrectangle  $H$  contenant totalement un polytope  $P$ , c.-à-d.  $P \subseteq H$ , dont on cherche à minimiser son hypervolume avec comme contrainte que ses arêtes doivent être parallèles aux axes. C'est une approximation externe de notre polytope qui est minimale au sens des contraintes. Un exemple est visible sur la Figure 3.

La contrainte d'alignement avec les axes rend le problème d'optimisation trivial : il suffit de chercher le minimum et le maximum de notre polytope sur chaque dimension, par exemple en programmation linéaire, donnant les limites de notre hyperrectangle. Le problème devient plus compliqué dans le cas général. En effet, nous devons déterminer l'enveloppe convexe du polytope, et donc tout d'abord de passer d'une H-représentation vers une V-représentation, ce qui est un problème NP-complet comme nous l'avons remarqué dans la Section 2.

Contrairement au MVAIR, le MVABH est forcément unique, rendant l'approximation avec ce dernier beaucoup plus stable.

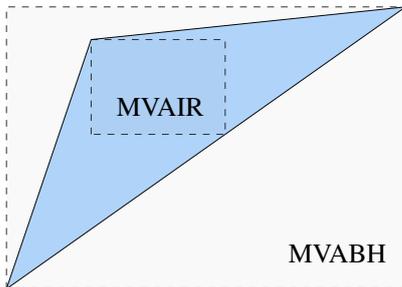


Figure 3 – MVAIR et MVABH d'un polytope

## 4.3 Délimitation des MCSs

À partir des approximations internes et externes de nos polytopes, il est possible de délimiter

l'ensemble des MCSs :

**Proposition 3.** Soit  $\mathcal{C}_P$  les MCSs de l'ensemble de polytopes  $\mathcal{P}$ ,  $\mathcal{C}_I$  et  $\mathcal{C}_O$  les MCSs des approximations internes  $\mathcal{I}$  et des approximations externes  $\mathcal{O}$  respectivement, avec  $I_i \subseteq P_i \subseteq O_i, \forall i$ , nous avons alors la relation suivante :

$$\forall c_P \in \mathcal{C}_P, \exists c_I \in \mathcal{C}_I, \exists c_O \in \mathcal{C}_O \quad (4) \\ \text{t.q. } c_I \subseteq c_P \subseteq c_O.$$

*Démonstration.* Nous prouvons tout d'abord qu'il existe  $c_I$  tel que  $c_I \subseteq c_P$ . Les  $I_i$  étant des approximations internes, il est impossible qu'un sur-ensemble de  $c_P$  forme un MCS, c.-à-d.  $\nexists c_P \subseteq c'$  tel que  $\bigcap_{i \in c'} I_i \neq \emptyset$ . Comme chaque  $I_i$  forme un ensemble consistant à lui seul, il existe forcément un ensemble  $c_I \subseteq c_P$  qui soit consistant et maximal pour les  $I_i$ .

La preuve est similaire pour les  $c_O$ , avec un argument inverse. □

**Proposition 4.** À partir des notions définies dans la Proposition 3, nous avons la délimitation suivante pour la cardinalité des ensembles :

$$|\mathcal{C}_O| \leq |\mathcal{C}_P| \leq |\mathcal{C}_I|. \quad (5)$$

*Démonstration.* Dans la preuve de la Proposition 3, nous avons montré que pour chaque  $c_P$ , il existe potentiellement plusieurs MCSs  $c_I$  incluant ce dernier. Cela suffit à montrer  $|\mathcal{C}_P| \leq |\mathcal{C}_I|$ .

Le même raisonnement peut être appliqué en remplaçant les approximations internes avec les polytopes, ainsi que les polytopes avec les approximations externes, concluant la preuve. □

Dans la Section 5, nous approximations des polytopes avec leurs représentations externes.

## 5 Illustration en estimation linéaire

Dans cette section, nous illustrons notre méthode sur deux petits problèmes de régression linéaire

imprécise, avec seulement deux dimensions. Notons que le but de cette section est seulement d'illustrer notre méthode, afin de montrer le potentiel de nos résultats sur de l'estimation d'un point de vue théorique logique - ensembliste.

Soit un ensemble de données  $\{y_i, x_i\}_{i=1}^N$  avec une unique entrée et une unique sortie, un modèle linéaire suppose que la relation entre la variable expliquée  $y_i$  et la variable explicative  $x_i$  est linéaire :

$$y_i = \beta_0 + \beta_1 x_i. \quad (6)$$

En général, les  $y_i$  sont considérées comme observées avec du bruit (normal)  $\epsilon$ , et on effectue une régression statistique. Dans notre cas, nous adoptons un point de vue logique/d'espace de version [10]. Nous présumons que les données prennent valeur dans un ensemble, et nous considérons que les valeurs précises sont consistantes avec un modèle linéaire.

### 5.1 Estimer les modèles linéaires possibles avec les MCSs en approximant

Dans les prochaines illustrations, nous considérons que nous observons des données imprécises  $R_i = ([y_i, \bar{y}_i], [x_i, \bar{x}_i])$ . Soient l'équation Cartésienne d'une ligne  $L = \{(x, y) \mid ax + by = c\}$  et deux points imprécis  $R_i$  et  $R_j$ , trouver toutes les lignes qui intersectionnent les deux rectangles est formalisé ainsi :

$$\mathcal{L}_{ij} = \{(a, b) \mid (L \cap H_i \neq \emptyset) \wedge (L \cap H_j \neq \emptyset)\}. \quad (7)$$

$n$  points peuvent être intersectionnés par une même ligne si et seulement si les  $\mathcal{L}_{ij}$  correspondants ( $\binom{n}{2}$  au total) ont une intersection non-vide, c.-à-d. que leurs indices appartiennent à un même sous-ensemble cohérent. Ce sous-ensemble est maximal si aucun autre point peut être intersectionné par la ligne.

Énumérer par recherche exhaustive les différents MCSs n'est pas possible, car les différents  $\mathcal{L}_{ij}$  sont des quadrilatères quelconques, en général

non-rectangles et énumérer les MCSs de polytopes quelconques est difficile comme nous l'avons mentionné dans la Section 2. Néanmoins, en approximant les différents  $\mathcal{L}_{ij}$  avec des MVABHs, ici des rectangles alignés avec les axes qui sont des approximations externes minimales (en respectant la contrainte), il est possible d'énumérer facilement les différents MCSs comme évoqué dans la Section 3.

Remarquons que nous n'utilisons pas les MVAIRs ici, car dans les prochaines illustrations les MVABHs sont suffisants pour avoir une bonne énumération des MCSs. De plus, déterminer les MVABHs  $H_{ij} = [\underline{a}, \bar{a}] \times [\underline{b}, \bar{b}]$  des  $\mathcal{L}_{ij}$  nécessite seulement pour chaque  $\mathcal{L}_{ij}$  de déterminer les valeurs minimales et maximales de  $a$  et  $b$ , ce qui est facile dans ce contexte, car maximiser (respectivement minimiser)  $a$  est équivalent à minimiser (respectivement maximiser)  $b$ . Par opposition, déterminer les MVAIRs est légèrement plus compliqué, car il faut connaître exactement la structure de  $\mathcal{L}_{ij}$  (qui est formé de quatre contraintes linéaires) pour résoudre le problème convexe correspondant.

### 5.2 Application

Nous commençons dans le cas où le modèle est en effet linéaire, mais où un point de données est aberrant (boîte 2), comme présenté sur la Figure 4. Comme nous pouvons le voir, le modèle de régression statistique, à cause du point aberrant, n'arrive pas retrouver le vrai modèle.

Par opposition, la Figure 5 montre 5 MCSs différents (et leurs intersections) obtenus en utilisant les approximations sous forme de rectangles alignés avec les axes. Nous pouvons voir que le plus grand MCS est  $c_3$  et il inclut les vrais paramètres, tandis que les autres MCSs sont plus petits et indiquent de possibles aberrations (en effet, l'observation 2 est dans tous les MCSs restants et pas dans le MCS de taille maximale). Notons qu'une approche de la sorte est différente des régressions imprécises standards, qui continuent d'utiliser des approches inspirées

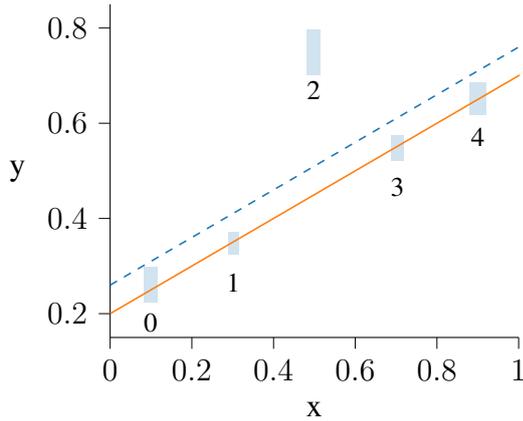


Figure 4 – Exemple d’un problème de régression linéaire imprécise. La ligne en pointillée correspond à une régression linéaire incluant le point aberrant. La ligne continue correspondant au vrai modèle

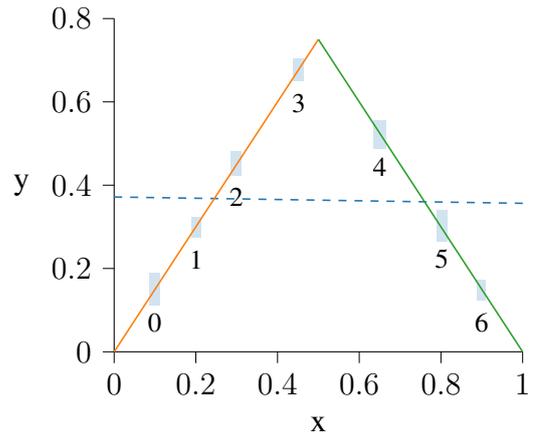


Figure 6 – Exemple d’un problème de régression linéaire par morceaux imprécise. La ligne en pointillés correspond à une régression linéaire classique. La ligne continue correspond aux deux vrais modèles.

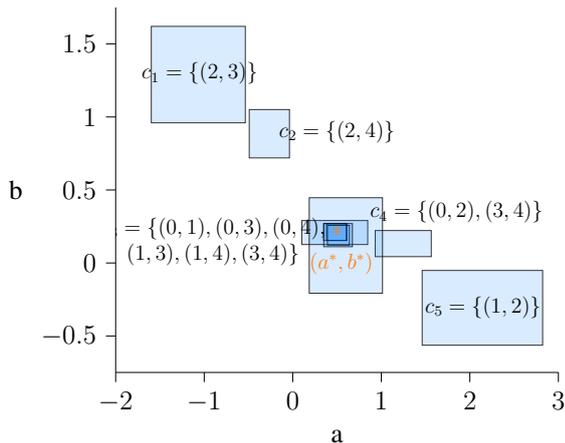


Figure 5 – Approximation des différents  $\mathcal{L}_{ij}$  de l’exemple vu sur la Figure 4

des moindres carrés (par exemple [7]).

La deuxième illustration s’intéresse à des observations non fausses (le modèle passe par toutes les observations), mais où le modèle supposé est faux, comme nous n’avons pas une régression linéaire, mais une régression linéaire par morceaux comme présenté sur la Figure 6. Nous avons deux partitions sur  $x$  :  $[0, 0.5]$  et  $[0.5, 1]$ . Les deux vrais modèles sont différents et une unique régression statistique linéaire sur le domaine entier de  $x$  s’ajuste mal aux données.

La Figure 7 montre ce qu’il se passe quand nous considérons une approche basée sur les MCSs. À la place d’avoir un unique MCS de grande taille et plusieurs petits, nous avons deux assez grands MCSs  $c_1$  and  $c_2$  avec plusieurs petits MCSs isolés. Les deux MCSs sont distincts et indiquent que l’erreur est dans le modèle, et que le modèle n’est valide que localement. Cela montre aussi qu’il n’est pas possible de déterminer à quel modèle l’observation 3 appartient, étant dans chaque MCS. Cette ambiguïté fait que  $c_1$  n’inclut pas les vrais paramètres  $(a_2^*, b_2^*)$  du deuxième modèle.

## 6 Conclusion

Les MCSs sont en théorie une notion très intéressante pour gérer des observations ou des informations conflictuelles, mais en pratique sont difficiles à énumérer. Nous avons montré que cette énumération est plus facile avec des hyperrectangles alignés avec les axes, car nous pouvons utiliser des algorithmes polynomiaux déjà connus sur les projections des hyperrectangles ayant valeur dans des intervalles. Nous avons illustré leur utilisation potentielle sur des problèmes d’estimation, en obtenant des infor-

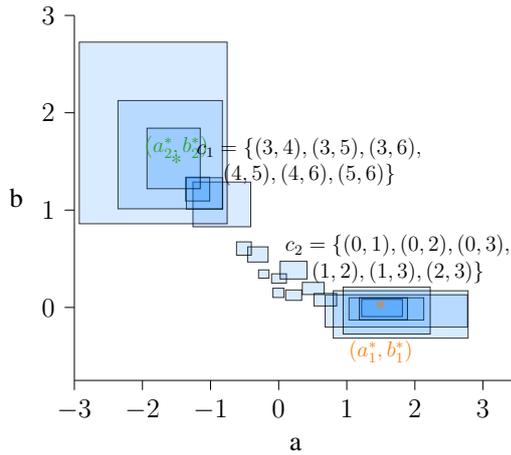


Figure 7 – Approximation des différents  $\mathcal{L}_{ij}$  de l'exemple vu sur la Figure 6

mations utiles sur les différents problèmes (observations erronées contre modèles erronés).

Dans le futur, nous prévoyons de comparer ces approches d'estimation avec des approches Bayésiennes dans le cas de modèles mal spécifiés, similairement à des travaux précédents pour résoudre des problèmes inverses [11]. Un autre cas où nos résultats peuvent être utiles est celui de la réparation d'informations inconsistantes en apprentissage de préférence [1].

#### Remerciements :

Ce travail a été réalisé dans le cadre du projet Pre-Serve, financé par la subvention ANR-18-CE23-0008 de l'Agence nationale de la recherche (ANR).

#### Références

- [1] L. Adam and S. Destercke. Possibilistic preference elicitation by minimax regret. In *Uncertainty in Artificial Intelligence*, pages 718–727. PMLR, 2021.
- [2] M. Behroozi. Largest inscribed rectangles in geometric convex sets. *arXiv preprint arXiv :1905.13246*, 2019.
- [3] S. Destercke, D. Dubois, and E. Chojnacki. Possibilistic information fusion using maximal coherent subsets. *IEEE Transactions on Fuzzy Systems*, 17(1) :79–92, 2008.

- [4] D. Dubois, H. Fargier, and H. Prade. Multiple-sources informations fusion—a practical inconsistency-tolerant approach. In *8th International Conference on Information Processing and Management of Uncertainty in Knowledge-based Systems (IPMU 2000)*. Consejo Superior de Investigaciones Científicas, 2000.
- [5] D. Dubois, W. Liu, J. Ma, and H. Prade. The basic principles of uncertain information fusion. an organised review of merging rules in different representation frameworks. *Information Fusion*, 32 :12–39, 2016.
- [6] D. Dubois, H. Prade, and R. Yager. Merging fuzzy information. In *Fuzzy sets in approximate reasoning and information systems*, pages 335–401. Springer, 1999.
- [7] M. B. Ferraro, R. Coppi, G. G. Rodríguez, and A. Colubi. A linear regression model for imprecise response. *International Journal of Approximate Reasoning*, 51(7) :759–770, 2010.
- [8] L. Khachiyan, E. Boros, K. Borys, V. Gurvich, and K. Elbassioni. Generating all vertices of a polyhedron is hard. In *Twentieth Anniversary Volume* :, pages 1–17. Springer, 2009.
- [9] R. Manor and N. Rescher. On inference from inconsistent premises. *Theory and Decision*, 1 :179–219, 1970.
- [10] T. M. Mitchell. Generalization as search. *Artificial intelligence*, 18(2) :203–226, 1982.
- [11] K. Shinde, P. Feissel, and S. Destercke. Dealing with inconsistent measurements in inverse problems : set-based approach. *International Journal for Uncertainty Quantification*, 11(3), 2021.
- [12] H. R. Tiwary. On the hardness of computing intersection, union and minkowski sum of polytopes. *Discrete & Computational Geometry*, 40(3) :469–479, 2008.