

Méthodes d'interpolation à noyaux pour l'approximation de fonctions type boîte noire coûteuses

Pierre Barbillon

▶ **To cite this version:**

Pierre Barbillon. Méthodes d'interpolation à noyaux pour l'approximation de fonctions type boîte noire coûteuses. Mathématiques [math]. Université Paris Sud - Paris XI, 2010. Français. NNT : . tel-00559502

HAL Id: tel-00559502

https://theses.hal.science/tel-00559502

Submitted on 25 Jan 2011

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.

N d'ordre : 10020

UNIVERSITÉ PARIS-SUD 11
FACULTÉ DES SCIENCES D'ORSAY

THÈSE

Présentée pour obtenir

LE GRADE DE DOCTEUR EN SCIENCES
DE L'UNIVERSITÉ PARIS-SUD 11

Spécialité : Mathématiques

par

Pierre Barbillon

Méthodes d'interpolation à noyaux pour l'approximation de fonctions type boîte noire coûteuses

Soutenue le 22 novembre 2010 après avis des rapporteurs

M. Olivier CAPPÉ
M. Pierre DRUILHET

devant la Commission d'examen composée de :

M. Yves AUFFRAY	(Co-directeur de thèse)
M. Olivier CAPPÉ	(Rapporteur)
M. Pierre DRUILHET	(Rapporteur)
M. Bertrand IOOSS	(Examinateur)
M. Jean-Michel MARIN	(Directeur de thèse)
M. Pascal MASSART	(Président du jury)

Remerciements

Je tiens tout d'abord à remercier Jean-Michel Marin d'avoir accepté d'encadrer mes travaux de thèse. Ses qualités professionnelles et humaines assourdissantes m'ont donné le goût de la recherche. J'espère que notre collaboration continuera puisqu'il est réellement stimulant d'échanger des idées et mener un projet en ta compagnie. Je garde de très bons souvenirs de mes excursions à Montpellier et j'en profite pour vous remercier toi, Carole, Lucas et Chloé de m'avoir reçu et régalé, entre autres, avec la cuisine du Gers.

Je suis très reconnaissant à Yves Auffray de m'avoir accueilli chez Dassault Aviation lors de mon stage de Master 2 et d'avoir suivi et contribué à l'élaboration de cette thèse. Ton soutien indéfectible et ton exigence ont été une aide précieuse au cours de ces trois années. Je suis désolé que les événements rares aient "ré-suscité" une dépendance tabagique.

Cette thèse et moi devons beaucoup à Pascal Massart. Il est à l'origine de ma rencontre avec Jean-Michel et Yves. Un grand merci pour m'avoir encouragé dans cette voie et pour avoir accepté de faire partie du jury de thèse.

Je remercie chaleureusement Gilles Celeux avec qui j'ai eu le privilège de travailler et de partager de nombreux cafés. Il a été aussi un soutien important au cours de ces trois années et toujours de bon conseil pour les pièces de théâtre.

J'adresse mes sincères remerciements à Olivier Cappé et Pierre Druilhet de m'avoir fait l'honneur d'accepter de rapporter mes travaux de thèse ainsi que de participer au jury. J'exprime également toute ma gratitude à Bertrand Iooss pour sa présence dans le jury.

Les conférences et les séminaires m'ont permis de rencontrer des statisticiens passionnants. J'en profite pour remercier Pierre Del Moral de nous avoir accordé de son temps pour discuter de la convergence de l'algorithme de recuit simulé.

Effectuer ma thèse au sein de l'équipe de probabilités et statistiques du laboratoire de mathématiques d'Orsay a été une expérience enrichissante et une chance. Je tiens notamment à saluer mes voisins du 440, Christine Keribin et Erwann Le Pennec. J'ai eu la chance de découvrir l'enseignement aux côtés de Odile Brandière, Farida Malek et Patrick Beau lors de mon service de moniteur. Toutes les tracasseries administratives ont été vaincues grâce à l'efficacité et à la patience de Valérie Lavigne et de Katia Evrat. Un grand merci à elles.
Je remercie également les deux directeurs successifs de l'école doctorale Pierre Pansu et David Harari ainsi que le conseiller aux thèses Frédéric Paulin d'avoir su créer une atmosphère conviviale et propice au bon déroulement d'une thèse.

En effectuant un ATER au MAP5 et à l'IUT de Paris Descartes, j'ai la possibilité de découvrir un nouveau laboratoire de recherche et de m'intégrer à une équipe sympathique. Je tiens notamment à remercier Guillaume Bordry, Servane Gey, Sylvie Hénaff, Mohamed Mellouk, Florence Muri, Elisabeth Ottenwaelter, Clarisse Pantin de la Guère, Jean-Michel Poggi et Adeline Samson pour leur chaleureux accueil et leur aide précieuse.

J'ai côtoyé au cours de mes années de thèse de nombreux doctorants dont mes compagnons du formidable bureau 227. Je remercie Cathy de m'avoir aiguillé et aidé à prendre mes marques. Je ne fus, j'en ai bien peur, d'aussi bon secours pour ceux arrivés ensuite. Je remercie Annalisa pour nos explorations des bars parisiens, Antoine pour m'avoir initié au Paris-Orsay à vélo, Jérôme pour ses bonbons, Maud pour m'avoir amené des tasses au MAP5, Nicolas pour avoir égaillé les pauses cafés, Patrick pour ses explications des règles du rugby, Shuai pour sa bonne humeur, Sourour pour sa gentillesse, Vincent pour m'avoir véhiculé lors des défaillances du RER B et un grand merci à tous pour m'avoir supporté.
Je tiens aussi à saluer les doctorants plus lointains dont Adeline, Benoît, Bertrand, Camille, Cyril, Jean-Patrick, Laure, Nathalie, Ramla, Robin, Sébastien et Wilson.

Merci à vous, amis strasbourgeois et messins (dans le sens recontrés à) de m'avoir accordé votre amitié et votre soutien dès lors qu'il fut nécessaire. Un merci particulier à Renaud avec qui j'ai commencé ma thèse en tant que colocataire.

Je dois beaucoup à ma famille qui m'a toujours soutenu et encouragé dans mes projets. Je remercie du fond du cœur mes parents et mes grands-parents. J'ai une pensée particulière pour ceux qui n'ont pas eu la patience d'attendre que je termine cette thèse.

Enfin, un énorme merci à Marianne qui m'a redonné confiance lorsque j'en avais le plus besoin. Merci pour ta tendresse auvergnate et pour ton côté *incontrôlable*.

Table des matières

1 Introduction **7**

2 État de l'art **11**
2.1 Métamodèles 11
2.1.1 Principe 11
2.1.2 Interpolateurs locaux 12
2.1.3 Techniques polynomiales 13
2.1.4 Splines 14
2.1.5 Interpolation à noyaux 15
2.1.6 Réseaux de neurones 16
2.1.7 Conclusion 17
2.2 Krigeage ou interpolation à noyaux 18
2.2.1 Modélisation par des processus gaussiens 18
2.2.2 Les noyaux 25
2.2.3 Interpolation à noyaux 27
2.2.4 Régularisation 31
2.2.5 Conclusion 32
2.3 Plans d'expérience numérique 33
2.3.1 Critères d'échantillonnage 34
2.3.2 Critères de distances entre les points 37
2.3.3 Plans d'expérience optimaux 39
2.3.4 Conclusion 41

3 Conditionally positive definite kernels **49**
3.1 Introduction 51
3.2 First definitions and notation 52
3.2.1 Measures with finite support 53
3.2.2 $\mathbb{P}$-unisolvent set 54
3.3 Bilinear forms induced by K 56
3.4 $\mathbb{P}$-conditionally positive definite kernel 60
3.4.1 $\mathbb{P}$-conditionally positive definite kernel 60
3.4.2 $\mathbb{P}$-Reproducing Kernel Semi-Hilbert Space 63
3.5 Interpolation in RKSHS 67
3.5.1 Preliminaries 67
3.5.2 Characterizations of interpolation in RKSHS 68
3.5.3 Lagrangian form of RKSHS interpolators 70

3.6 Regularized regression in RKSHS . . . 77
3.7 Discussion . . . 79

4 Maximin design **83**
4.1 Introduction . . . 84
4.2 Error bounds with kernel interpolations . . . 86
4.3 Computing maximin designs . . . 88
4.4 Variants of the algorithm . . . 96
4.5 Numerical illustrations . . . 98
4.6 Application to a simulator of an aircraft engine . . . 99

5 Non linear methods for inverse statistical problems **105**
5.1 Introduction . . . 106
5.2 The model and its linear identification . . . 107
5.3 Using a non linear approximation of the function H . . . 109
5.3.1 The SEM algorithm . . . 110
5.3.2 SEM with Kriging approximation of H . . . 112
5.4 Numerical experiments . . . 115
5.4.1 A flooding model . . . 115
5.4.2 A non linear example . . . 117
5.5 Discussion . . . 120

6 Estimation of rare events probabilities **125**
6.1 Introduction . . . 126
6.2 Bayesian estimator and credible interval . . . 128
6.3 Importance sampling . . . 132
6.4 Numerical experiments . . . 134
6.4.1 A toy example . . . 134
6.4.2 A real case study: release enveloppe clearance . . . 137
6.5 Dicussion . . . 142
6.6 Confidence bounds for the binomial distribution . . . 144

7 Discussion et perspectives **145**

Chapitre 1

Introduction

Contexte

Certaines expériences physiques ne sont pas réalisables de par leur coût, ou du fait de l'impossibilité de fixer les facteurs expérimentaux. Le cas échéant, il est toutefois envisageable, à partir d'un modèle mathématique décrivant le système physique étudié, d'avoir recours à une expérience simulée. La solution du modèle pour des conditions expérimentales choisies est alors déterminée par un code de calcul. La résolution n'est généralement pas analytique mais seulement numérique, ce qui est typiquement le cas lorsque le modèle contient des équations aux dérivées partielles. L'expérience simulée peut aussi être appelée in silicio ou in silico. Historiquement, les premières expériences simulées sont sans doute celles effectuées au laboratoire de Los Alamos pour étudier le comportement des armes nucléaires. Aujourd'hui, elles sont entre autres utilisées en fiabilité des structures, dans l'aéronautique, dans la sécurisation des réacteurs nucléaires, en science du climat... Une expérience simulée revient alors à l'évaluation d'une fonction f en une entrée vectorielle $\mathbf{x}$ qui représente les conditions expérimentales :

$$y = f(\mathbf{x}) \,,$$

la réponse ou sortie y est en général vectorielle. L'espace des entrées, noté E, est supposé être un espace compact inclus dans $\mathbb{R}^d$. La fonction f est déterministe ; répéter l'expérience au point $\mathbf{x}$ n'apporte aucune information supplémentaire. Elle est appelée fonction boîte noire car elle n'est pas connue explicitement. Le code servant à l'évaluer est soit non accessible, soit trop complexe pour appréhender le comportement de f. De plus, le calcul en un point $\mathbf{x}$ est souvent coûteux en temps. L'augmentation régulière de la capacité de calcul des ordinateurs ne réduit pas le problème car les modèles physiques se complexifient à la même vitesse.

Le vecteur des entrées $\mathbf{x}$ peut contenir deux types de variables :

- les variables de contrôle sont les variables d'intérêt, elles peuvent être fixées par l'ingénieur ou le scientifique afin de contrôler le système,
- les variables environnementales ne présentent pas un intérêt majeur, mais elles doivent être prises en compte puisqu'elles peuvent avoir un effet sur les sorties. Elles dépendent de l'environnement et ne peuvent être fixées dans une expérience physique. Elles sont aussi appelées variables de bruit.

Les entrées sont une source d'incertitude. En effet, il y a une incertitude sur le réglage des variables de contrôle et les variables environnementales souffrent d'une incertitude de mesure. La propagation des incertitudes des entrées aux sorties du modèle est une préoccupation importante. Il y a aussi une source d'incertitude due à l'écart entre le modèle et la réalité

physique, ce qu'il est impossible de mesurer puisque les expériences physiques ne seront pas réalisées. La principale difficulté dans l'utilisation des expériences simulées réside dans le fait que le code permettant l'évaluation de la fonction f est trop coûteux. Il n'est alors pas possible d'obtenir les sorties pour un grand nombre d'entrées ce qui est pourtant nécessaire pour explorer correctement le domaine expérimental et apprendre assez finement le comportement de f. C'est pourquoi il peut être utile d'avoir recours à une fonction qui approche la fonction f du modèle le plus précisément possible mais qui est d'évaluation quasi instantanée. Ainsi cette fonction permet de s'intéresser aux relations entre y et $\mathbf{x}$. Elle est notée $\hat{f}$. Elle est appelée un métamodèle (Kleijnen, 1987). Sa forme analytique et sa rapidité d'éxécution permettent de nombreuses applications (voir par exemple Koehler et Owen, 1996; Fang *et al.*, 2006) :

- Étude préliminaire et visualisation. Les graphiques demandent de nombreuses évaluations, ce qui est possible grâce au métamodèle. Les graphiques 3-D ou les animations permettent d'avoir une compréhension du modèle, des interactions entre les entrées et sorties et éventuellement de se rendre compte de l'existence d'extrema locaux.
- Prédiction et optimisation. Le métamodèle permet de proposer une valeur approchée de f en tous les points de E. Ainsi, il est, par exemple, possible d'approcher

$$\int_E f(\mathbf{x})d\mathbf{x}\,,$$

par

$$\int_E \hat{f}(\mathbf{x})d\mathbf{x}\,.$$

De plus, la forme analytique de certains métamodèles permettent une intégration formelle. Afin de déterminer un point $\mathbf{x}^*$ où f atteint un minimum global :

$$f(\mathbf{x}^*) = \min_{\mathbf{x}\in E} f(\mathbf{x})\,,$$

Jones *et al.* (1998) ont proposé un algorithme qui réduit sensiblement le nombre d'appels au code de calcul coûteux de f et qui s'appuie sur le métamodèle tout en contrôlant l'incertitude liée à son utilisation.

- Analyse de sensibilité. Le but est de quantifier et de classer les effets des entrées qui sont des variables aléatoires $\mathbf{X} = (X_1, \ldots, X_d)$ de lois connues. Sobol (1993) a proposé des indices qui quantifient la proportion de la variance des sorties $Y = f(\mathbf{X})$ expliquée par la variance des variables d'entrée X_i :

$$CR_i = \frac{\mathrm{Var}\mathbb{E}(Y|X_i)}{\mathrm{Var}(Y)}\,,$$

pour $i = 1, \ldots, d$. Ces indices sont calculés numériquement à l'aide de techniques d'intégration de Monte-Carlo. Ces méthodes demandent beaucoup d'évaluations de f, d'où l'utilité d'y substituer $\hat{f}$ (voir Marrel *et al.*, 2009, pour le calcul d'indices de Sobol à l'aide d'un certain type de métamodèles).

- Analyse probabiliste. Des questions de fiabilité et d'évaluation des risques industriels reposent sur les expériences simulées. Les entrées $\mathbf{X}$ sont supposées suivre une loi de probabilité connue de densité p et l'objectif est de prédire la probabilité que la sortie $Y = f(\mathbf{X})$ soit au-delà ou en deçà d'un seuil donné ρ (Haldar et Mahadevan, 2000). Par exemple, on s'intéresse à l'estimation de :

$$I = \mathbb{P}(Y > \rho) = \int \mathbb{1}_{\{\mathbf{x}:f(\mathbf{x})>\rho\}}(\mathbf{x})p(\mathbf{x})d\mathbf{x}\,.$$

Puisque f n'est pas sous forme analytique, il faut recourir à des techniques d'intégration numérique. Le métamodèle est alors utilisé pour proposer cet estimateur de la probabilité I :

$$\widehat{I} = \mathbb{P}(\hat{f}(\mathbf{X}) > \rho) = \int_E \mathbb{I}_{\{\mathbf{x}:\hat{f}(\mathbf{x})>\rho\}}(\mathbf{x}) p(\mathbf{x}) d\mathbf{x}\,.$$

L'approche statistique des expériences simulées implique deux parties :

- Le métamodèle est construit à partir d'un nombre restreint d'évaluations de la vraie fonction f en des points $\{\mathbf{x}_1, \ldots, \mathbf{x}_n\} = D \subset E$. Cet ensemble de points D est appelé un plan d'expérience. Ces plans d'"expérience numérique" sont différents de ceux utilisés dans le cadre d'expériences physiques (Fisher, 1971). En effet, la réplication d'une expérience en un même point n'a un sens que dans le cadre des expériences physiques.
- Les métamodèles doivent être très adaptatifs afin de tenir compte d'une non linéarité éventuelle et alors proposer de bonnes prédictions en tout point du domaine. Ce sont souvent des techniques venant des statistiques non paramétriques.

Organisation de la thèse

Dans un chapitre 2 introductif, nous présentons différents types de métamodèles et nous décrivons plus précisément ceux qui proviennent de méthodes d'interpolations à noyaux aussi connus sous le nom de krigeage. Dans cet état de l'art, nous traitons également de la construction de plans d'expérience numérique.

Les méthodes d'interpolation à noyaux sont au coeur de ce travail de thèse qui s'articule autour de quatre contributions principales.

Dans le chapitre 3, nous introduisons une définition plus générale de la notion de noyau conditionnellement défini positif que celle habituellement utilisée dans la littérature. Nous nous intéressons aux espaces associés à ce type de noyaux ainsi qu'à l'interpolation et la régression régularisée dans ces espaces. Cette définition permet une vraie généralisation du concept de noyau défini positif et des théorèmes associés.

Le chapitre 4 contient un algorithme de construction de plans d'expérience dans des domaines éventuellement non hypercubiques suivant un critère maximin pertinent pour le krigeage. Cet algorithme est fondé sur un recuit simulé. Sa convergence théorique est démontrée et des essais pratiques sont réalisés.

Dans le chapitre 5, nous traitons un problème statistique inverse. Nous utilisons un algorithme stochatisque EM dans lequel un métamodèle de krigeage est employé puisque le modèle liant les entrées aux sorties est un code boîte noire coûteux. Cette méthode est testée sur un modèle de crues fourni par EDF.

Enfin, toujours dans le contexte d'un modèle boîte noire coûteux, nous proposons une procédure d'échantillonnage préférentiel pour estimer et surtout majorer la probabilité de dépassement d'un seuil par les sorties du modèle dans la partie 6. Le dépassement du seuil est un événement rare et redouté. La loi instrumentale est définie grâce à un métamodèle de krigeage. Un estimateur bayésien de cette probabilité est également proposé. Ces méthodes sont testées sur des exemples jouets et un cas réel fourni par Dassault Aviation.

Chapitre 2

État de l'art

2.1 Métamodèles

2.1.1 Principe

Dans le domaine des expériences simulées, le modèle s'écrit

$$y = f(\mathbf{x})\,. \tag{2.1}$$

Le but est de proposer un estimateur de f qui l'approche de manière optimale (en un sens à définir) à partir d'un ensemble de données $\{(\mathbf{x}_i, y_i = f(\mathbf{x}_i)), 1 \leq i \leq n\}$. Cet ensemble est souvent appelé échantillon d'apprentissage car il sert à ajuster le métamodèle à l'information dont nous disposons sur f. L'ensemble de points noté $D = \{\mathbf{x}_1, \ldots, \mathbf{x}_n\}$ est le plan d'expérience. La construction d'un métamodèle peut être vue comme une régression sur des données non bruitées. Les concepts statistiques peuvent être appliqués et étendus au cadre des modélisations d'expériences simulées. Cependant, l'erreur commise en remplaçant f par $\hat{f}$ comporte uniquement un terme de biais. La variance est nulle car les sorties $y_i = f(\mathbf{x}_i)$ $(1 \leq i \leq n)$ sont déterministes. Le biais vient du fait que $\hat{f}$ appartient à un espace de fonctions ne contenant pas nécessairement f. Dans la suite, on pourra être amené à considérer f de manière aléatoire afin d'appliquer les outils de la régression statistique. Cela revient à transformer une partie de l'erreur de biais en une variance. On peut aussi construire le métamodèle comme un interpolateur de f aux points du plan d'expérience $D = \{\mathbf{x}_1, \ldots, \mathbf{x}_n\}$, c'est-à-dire construire $\hat{f}$ tel que

$$\hat{f}(\mathbf{x}_i) = y_i = f(\mathbf{x}_i),\ i = 1, \ldots, n\,.$$

Pour justifier le bien fondé de cet interpolateur, il peut être nécessaire de formuler des hypothèses sur f, telle sa régularité.

Il faut définir un critère mesurant la proximité entre le métamodèle proposé et la vraie fonction f. On peut mesurer l'erreur en moyenne quadratique intégrée (*IMSE : Integrated mean squared error*) entre eux

$$IMSE = \int_E (f(\mathbf{x}) - \hat{f}(\mathbf{x}))^2 d\mathbf{x}\,, \tag{2.2}$$

ou l'erreur pondérée ce qui permet d'introduire une information a priori sur les entrées

$$IWMSE = \int_E (f(\mathbf{x}) - \hat{f}(\mathbf{x}))^2 g(\mathbf{x}) d\mathbf{x}\,,$$

où g est une fonction de poids telle que $g \geq 0$ et $\int_E g(\mathbf{x})d\mathbf{x} = 1$. Cette démarche permet de tenir compte de la loi de probabilité des entrées. La fonction de poids g représente alors la fonction densité de cette loi. Le critère $IMSE$ est, à une constante près, le critère $IWMSE$ pour une loi uniforme sur E.
Ces quantités ne sont pas calculables directement puisqu'elles nécessitent un grand nombre d'appels à f. On peut cependant les estimer par validation croisée. Étant donné que la qualité du métamodèle dépend du nombre de points d'apprentissage, il est raisonnable d'utiliser une validation croisée dite "leave-one-out". C'est-à-dire que nous construisons n métamodèles $(\hat{f}_{-i})_{1\leq i\leq n}$ où $\hat{f}_{-i}$ est construit à l'aide des données $\{(\mathbf{x}_i, y_i), 1, \ldots, i-1, i+1, \ldots, n\}$. Ainsi $IWMSE$ est estimée par

$$\frac{1}{n}\sum_{i=1}^{n}(f(\mathbf{x}_i) - \hat{f}_{-i}(\mathbf{x}_i))^2 g(\mathbf{x}_i)\,. \tag{2.3}$$

2.1.2 Interpolateurs locaux

Un interpolateur simple est l'interpolateur des k-plus proches voisins. Pour k un entier fixé et $\mathbf{x} \in E$, on note $V_k(\mathbf{x})$ l'ensemble des k points de $\{\mathbf{x}_1, \ldots, \mathbf{x}_n\}$ les plus proches de $\mathbf{x}$ au sens de la distance euclidienne. L'interpolateur $\hat{f}$ est,

$$\begin{array}{ll} \forall \mathbf{x} \in E - \{\mathbf{x}_1, \ldots, \mathbf{x}_n\}, & \hat{f}(\mathbf{x}) = \frac{1}{k}\sum_{i:\mathbf{x}_i \in V_k(\mathbf{x})} y_i \\ \forall 1 \leq i \leq n, & \hat{f}(\mathbf{x}_i) = y_i \end{array}\,.$$

Cet interpolateur est discontinu tandis que f est généralement supposée relativement régulière. C'est pourquoi il est sensé d'utiliser un interpolateur des plus proches voisins pondéré par l'inverse des distances :

$$\begin{array}{ll} \forall \mathbf{x} \in E - \{\mathbf{x}_1, \ldots, \mathbf{x}_n\}, & \hat{f}(\mathbf{x}) = \sum_{i:\mathbf{x}_i \in V_k(\mathbf{x})} \frac{\|\mathbf{x}_i - \mathbf{x}\|_2^{-1}}{\sum_i \|\mathbf{x}_i - \mathbf{x}\|_2^{-1}} y_i \\ \forall 1 \leq i \leq n, & \hat{f}(\mathbf{x}_i) = y_i \end{array}\,.$$

Il est aussi possible (Hastie *et al.*, 2001, chap. 6) de pondérer les évaluations $(y_1, \ldots, y_n)$ à l'aide de noyaux :

$$\forall\ \mathbf{x} \in E - \{\mathbf{x}_1, \ldots, \mathbf{x}_n\},\ \hat{f}(\mathbf{x}) = \frac{\sum_{i=1}^{n} K_\lambda(\mathbf{x}, \mathbf{x}_i) y_i}{\sum_{i=1}^{n} K_\lambda(\mathbf{x}, \mathbf{x}_i)}\,.$$

La fonction K_λ est un noyau symétrique et λ est un paramètre qui détermine la largeur du voisinage au point considéré.
Soit K_λ un noyau d'Epanechnikov, on a

$$K_\lambda(\mathbf{x}, \mathbf{x}') = \begin{cases} \frac{3}{4}(1 - \frac{\|\mathbf{x}-\mathbf{x}'\|_2^2}{\lambda^2}) & \text{si } \frac{\|\mathbf{x}-\mathbf{x}'\|_2^2}{\lambda^2} \leq 1; \\ 0 & \text{sinon.} \end{cases}$$

Ce noyau K_λ est alors à support fini et λ permet de régler le rayon du support.
Le noyau gaussien est

$$K_\lambda(\mathbf{x}, \mathbf{x}') = \frac{1}{\sqrt{2\pi}\lambda} \exp\left(-\frac{\|\mathbf{x} - \mathbf{x}'\|_2^2}{2\lambda^2}\right),\ \forall t \in \mathbb{R}\,,$$

où λ est l'écart type.
Le paramètre λ permet de régler la régularité de l'estimateur. Dans le cas des k-plus proches voisins, k joue ce rôle de paramètre de régularisation. Ces paramètres peuvent être ajustés par une méthode de validation croisée comme décrit précédemment.
À l'aide des noyaux, il est possible de construire un métamodèle comme une régression polynomiale locale (Fan et Gijbels, 1996).

2.1.3 Techniques polynomiales

On cherche l'interpolateur sous la forme

$$\hat{f}(\mathbf{x}) = \sum_{j=1}^{L} \beta_j B_j(\mathbf{x}) \,,$$

où B_j $(1 \leq j \leq n)$ sont des fonctions polynomiales. Par exemple, on peut prendre $(B_j)_{1\leq j\leq L}$ comme une base de l'espace des polynômes de degré inférieur ou égal à deux. Ainsi par $\hat{f}$, on capturerait les interactions des deux premiers ordres. Les paramètres β_j $(1 \leq j \leq L)$ sont choisis par le critère des moindres carrés. C'est-à-dire qu'ils minimisent

$$\sum_{i=1}^{n} \left(y_i - \sum_{j=1}^{L} \beta_j B_j(\mathbf{x}_i) \right)^2 = \|\mathbf{y} - \mathbf{B}\boldsymbol{\beta}\|_2^2 \,.$$

On a noté, pour $D = \{\mathbf{x}_1, \ldots, \mathbf{x}_n\}$,

$$\mathbf{B}(D) = \begin{pmatrix} B_1(\mathbf{x}_1) & \cdots & B_p(\mathbf{x}_1) \\ \vdots & & \vdots \\ B_1(\mathbf{x}_n) & \cdots & B_p(\mathbf{x}_n) \end{pmatrix}, \; \boldsymbol{\beta} = \begin{pmatrix} \beta_1 \\ \vdots \\ \beta_p \end{pmatrix} . \tag{2.4}$$

Si $\mathbf{B}^T\mathbf{B}$ est inversible, on obtient

$$\boldsymbol{\beta} = (\mathbf{B}^T\mathbf{B})^{-1}\mathbf{B}^T\mathbf{y} \,. \tag{2.5}$$

Il est recommandé de prendre une base de fonctions orthogonales afin d'éviter les problèmes liés à la colinéarité des colonnes de la matrice $\mathbf{B}$. An et Owen (2001) proposent de construire les fonctions $B_j,\ 1 \leq j \leq L$, comme des produits tensoriels de polynômes univariés orthogonaux (e.g. les polynômes de Legendre, polynômes de Tchebychev). Supposons que l'approximation se fait sur le cube $[0,1]^d$, les polynômes orthogonaux univariés sont tels que $\phi_0(u) = 1$ pour $u \in [0,1]$ et pour $j \geq 1,\ \phi_j$ satisfait

$$\int_0^1 \phi_j(u) = 0, \; \int_0^1 \phi_j^2(u)du = 1, \text{ et } \int_0^1 \phi_j(u)\phi_k(u) = 0, \text{ pour } j \neq k \,.$$

En dimension d on obtient les fonctions de base, par tensorisation, $\mathbf{x} = (x_1, \ldots, x_d) \in [0,1]^d$,

$$\boldsymbol{\phi}_{r_1,\ldots,r_d}(\mathbf{x}) = \prod_{k=1}^{d} \phi_{r_k}(x_k) \,.$$

Le nombre de fonctions de base augmente fortement avec la dimension d des entrées. Le nombre n de données $\{(\mathbf{x}_i, y_i), 1 \leq i \leq n\}$ requis peut donc être trop important. On est alors obligé de se cantonner à des ordres polynomiaux peu élevés. Une sélection de modèle peut améliorer la prédiction en limitant le nombre L de fonctions de base.
La régression ridge (Hoerl et Kennard, 1970) permet d'introduire une régularisation. Elle consiste à minimiser un problème de moindres carrés pénalisés :

$$\|\mathbf{y} - \mathbf{B}\boldsymbol{\beta}\|_2^2 + \lambda\|\boldsymbol{\beta}\|_2^2 \,,$$

avec $\lambda > 0$. Une régularisation avec une norme L_1 est aussi possible :

$$\|\mathbf{y} - \mathbf{B}\boldsymbol{\beta}\|_2^2 + \lambda\|\boldsymbol{\beta}\|_1 \,.$$

Cette régularisation conduit à un vecteur $\boldsymbol{\beta}$ avec peu de termes non nuls. Elle est connue sous le nom de LASSO (voir Tibshirani, 1994). Si l'on souhaite obtenir un interpolateur de la fonction f, il faut prendre le nombre L de fonctions de base assez grand afin que l'équation

$$\mathbf{y} = \mathbf{B}\boldsymbol{\beta} \,, \tag{2.6}$$

ait une solution. Si $L \geq n$, la solution à ce problème n'est plus unique. Il est alors possible de chercher $\boldsymbol{\beta}$ qui minimise

$$\|\boldsymbol{\beta}\|^2 \,,$$

sous la contrainte (2.6) (Rao, 1973).

2.1.4 Splines

Nous traitons ici des splines de régression (Stone *et al.*, 1997) comme une extension des modèles polynomiaux. On tensorisera des fonctions splines à une dimension. Les fonctions splines présentées ici sont définies à l'aide de fonctions puissances. Il existe également la base des B-splines (De Boor, 1978). Pour des points $\kappa_1, \ldots, \kappa_L$ fixés appelés noeuds, on pose, pour $p \in \mathbb{N}^*$ et $u \in [0, 1]$,

$$\begin{gathered} S_0(u) = 1,\ S_1(u) = u, \ldots, S_p(u) = u^p, \\ S_{p+1}(u) = (u - \kappa_1)_+^p, \ldots, S_{p+L}(u) = (u - \kappa_L)_+^p. \end{gathered}$$

En dimension d, un produit tensoriel donne pour $\mathbf{x} = (x_1, \ldots, x_d) \in [0, 1]^d$,

$$B_{r_1,\ldots,r_d}(\mathbf{x}) = \prod_{k=1}^{d} S_{r_k}(x_k) \,.$$

L'ensemble des fonctions $\{B_{r_1,\ldots,r_d},\ 0 \leq r_k \leq p + L,\ k = 1, \ldots, d\}$ est une base d'un espace fonctionnel sur $\mathbf{x}$. Lorsque la dimension des entrées augmente, le nombre de fonctions de base croît exponentiellement. Friedman (1991) a proposé une méthode appelée *Multivariate Adaptive Regression Splines (MARS)* qui sélectionne les fonctions de base utilisées et les noeuds à partir des données. Pour $B_0, \ldots, B_M \in \{B_{r_1,\ldots,r_d},\ 0 \leq r_k \leq p + L,\ k = 1, \ldots, d\}$, on considère le modèle de spline de régression

$$f(\mathbf{x}) = \sum_{j=0}^{M} \beta_j B_j(\mathbf{x}) \,.$$

Si $M+1 \leq n$, le vecteur des coefficients $\boldsymbol{\beta} = (\beta_1, \ldots, \beta_M)$ est estimé par la formule des moindres carrés comme en régression polynomiale (2.5).
Il est également possible d'introduire une pénalisation pour des bases de splines (Eilers et Marx, 1996).
Les splines de lissage sont une autre façon de construire un métamodèle qui permet d'introduire un facteur de régularisation. Cette méthode sera évoquée dans la suite dans une présentation plus générale.

2.1.5 Interpolation à noyaux

On cherche $\hat{f}$ comme un interpolateur des points y_i en les points $\mathbf{x}_i$:

$$y_i = \hat{f}(\mathbf{x}_i),\ \forall 1 \leq i \leq n\,,$$

avec $\hat{f} \in \mathcal{H}_\mathcal{K}$, l'espace Hilbert de noyau reproduisant K (RKHS, Aronszajn, 1950). Le noyau K est supposé défini positif. Il peut être une fonction radiale de base (Powell, 1987). C'est-à-dire qu'il est de la forme pour $\mathbf{x}, \mathbf{x}' \in E$, $K(\mathbf{x}, \mathbf{x}') = R(\|\mathbf{x} - \mathbf{x}'\|)$ où $R : \mathbb{R} \to \mathbb{R}$. Le noyau gaussien $K(\mathbf{x}, \mathbf{x}') = \exp(-\theta\|\mathbf{x} - \mathbf{x}'\|_2^2)$, pour $\theta > 0$, en est un exemple. Si l'on fait l'hypothèse que $f \in \mathcal{H}_\mathcal{K}$, il existe alors un unique interpolateur de norme minimale dans $\mathcal{H}_K$ (Schaback, 1995a). Ainsi, $\hat{f}$ est l'interpolateur de f de norme minimale dans l'espace associé au noyau choisi et il est le projeté orthogonal de f sur l'espace engendré par $\{K_{\mathbf{x}_1}, \ldots, K_{\mathbf{x}_n}\}$ les fonctions partielles correspondant aux points du plan d'expérience ($i = 1, \ldots, n,\ K_{\mathbf{x}_i}(\mathbf{x}) = K(\mathbf{x}_i, \mathbf{x})$ pour tout $\mathbf{x} \in E$).
Schaback (2007) montre que les splines puissances présentées dans le paragraphe 2.1.4 peuvent être vues comme des noyaux dits conditionnellement définis positifs. Cela représente une classe plus générale de noyaux et les espaces fonctionnels associés ne sont plus hilbertiens. Dans ce cas, il y a une extension du théorème permettant de trouver l'interpolateur le plus lisse bien qu'il ne fournisse pas l'algorithme de recherche le plus efficace.

Une technique venant des géostatistiques, le krigeage (Cressie, 1993), consiste à modéliser une quantité d'intérêt variant suivant des données spatiales comme une réalisation d'un processus gaussien. La fonction de covariance de ce processus est définie à l'aide d'un noyau. Si l'on dispose d'observations $\{(\mathbf{x}_i, y_i), 1 \leq i \leq n\}$ on peut construire le meilleur prédicteur linéaire sans biais (BLUP) afin de prédire la valeur y_0 non observée correspondant au point $\mathbf{x}_0$. Sacks *et al.* (1989a) ont proposé d'utiliser cette modélisation pour la fonction boîte noire f déterministe et d'utiliser le BLUP comme métamodèle. Ce dernier interpole la fonction f aux points $\mathbf{x}_1, \ldots, \mathbf{x}_n$ et est égal à l'interpolateur de norme minimale dans le RKHS associé au noyau de covariance.

Une vision plus souple consiste à ne plus faire d'hypothèses sur la fonction f et à chercher une fonction $\hat{f}$ appartenant à $\mathcal{H}_K$ comme un compromis entre une proximité à f et la valeur de la norme $\|\hat{f}\|_{\mathcal{H}_K}$. Cela permet d'avoir une solution régularisée mais $\hat{f}$ n'est plus un interpolateur de f aux points $\mathbf{x}_1, \ldots, \mathbf{x}_n$. On cherche donc $\hat{f}$ comme une solution du problème :

$$\min_{g \in \mathcal{H}_K} \sum_{i=1}^{n} (y_i - g(\mathbf{x}_i))^2 + \lambda \|g\|_{\mathcal{H}_K}^2\,,$$

où λ est un réel strictement positif. Le théorème du représentant (Kimeldorf et Wahba, 1971) permet de garantir l'existence et l'unicité de la solution qui appartient aussi à l'espace en-

gendré par $\{K_{\mathbf{x}_1}, \ldots, K_{\mathbf{x}_n}\}$. Wahba (1990) traite le problème de la régularisation dans le cadre des splines de plaque mince. Celles-ci sont un type de splines de lissage et peuvent aussi être utilisées pour une problématique d'interpolation pure où elles correspondent à une fonction radiale de base définissant un noyau conditionnellement défini positif (Schaback, 1995b).

Les support vector machines (SVM) peuvent aussi être utilisées en régression. Cela permet de réduire le nombre de fonctions intervenant dans la construction du métamodèle (Hastie *et al.*, 2001) et d'obtenir un métamodèle parcimonieux. On cherche $\hat{f}$ de la forme

$$g = \sum_{i=1}^{n} \beta_i K(\mathbf{x}_i, .) + \beta_0 \,, \tag{2.7}$$

où K est un noyau défini positif. Les coefficients $(\beta_0, \beta_1, \ldots, \beta_n) = (\beta_0, \boldsymbol{\beta})$ sont choisis comme solution du problème de minimisation :

$$\sum_{i=1}^{n} V_\epsilon(y_i - g(\mathbf{x}_i)) + \lambda \|\boldsymbol{\beta}\|_2^2 \,,$$

où

$$V_\epsilon(t) = \begin{cases} 0 & \text{si } |t| < \epsilon, \\ |t| - \epsilon & \text{sinon.} \end{cases}$$

Le coefficient ϵ est strictement positif et la fonction V_ϵ est une mesure d'erreur dite ϵ-insensible (Vapnik, 1996). Ainsi seul un sous ensemble des coefficients de $\boldsymbol{\beta}$ est non nul du fait de la forme de la mesure d'erreur. Les vecteurs $\mathbf{x}_i$ associés aux β_i non nuls sont appelés vecteurs de support.

2.1.6 Réseaux de neurones

Nous présentons le modèle des perceptrons multi-couches (Bishop, 2006). Cela consiste en une régression à deux étapes. La première étape est la création de M unités cachées $z_1, \ldots, z_M$. Une unité cachée z_m est l'image par une fonction d'activation σ d'une combinaison linéaire des entrées $\mathbf{x} = (x_1, \ldots, x_d) \in E$:

$$z_m = \sigma(\alpha_0 + \boldsymbol{\alpha}_m^T \mathbf{x}), \ m = 1, \ldots, M \,.$$

On choisit en général la fonction sigmoïde $\sigma(v) = 1/(1 + \exp(-v))$ comme fonction d'activation. On peut aussi choisir la fonction radiale de base de type gaussien. La deuxième étape est une combinaison linéaire des unités cachées $\mathbf{z} = (z_1, \ldots, z_m)$. Cela donne le métamodèle

$$\hat{f}(\mathbf{x}) = \beta_0 + \boldsymbol{\beta}^T \mathbf{z}(\mathbf{x}) \,, \tag{2.8}$$

où $\mathbf{z} = (z_1, \ldots, z_M)$. Si la fonction σ est l'identité, on obtient simplement un modèle de régression linéaire. Ainsi, c'est une généralisation non linéaire du modèle linéaire. Cybenko (1989) a montré que toutes les fonctions continues bornées sont approchables avec une précision arbitraire par un réseau avec une couche cachée et utilisant des fonctions d'activation type sigmoïde.

Les coefficients $(\alpha_0, \boldsymbol{\alpha}, \beta_0, \boldsymbol{\beta}) = \boldsymbol{\theta}$ sont pris comme solutions d'un problème de minimisation de la somme des erreurs quadratiques sur l'échantillon d'apprentissage : pour g de la forme (2.8),

$$R(\boldsymbol{\theta}) = \sum_{i=1}^{n} (y_i - g(\mathbf{x}_i))^2 .$$

Il faut souvent pénaliser ce terme d'attache aux données afin de proposer un métamodèle plus lisse qui ne sera plus un interpolateur. On cherche alors les paramètres minimisant :

$$R(\boldsymbol{\theta}) + \lambda_1 \|\boldsymbol{\beta}\|_2^2 + \lambda_2 \|\boldsymbol{\alpha}\|_2^2 ,$$

où $\lambda_1 > 0$ et $\lambda_2 > 0$.

2.1.7 Conclusion

Ces techniques présentent de nombreux points communs et les méthodes employées pour les ajuster reposent sur les mêmes idées. Une fonction d'erreur quadratique sert à fixer les paramètres du métamodèle et une contrainte de régularisation est parfois introduite pour obtenir un métamodèle lisse. Il est possible de combiner des méthodes de construction d'un métamodèle. Par exemple, on peut effectuer une régression polynomiale sur nos données $\{(\mathbf{x}_i, y_i = f(\mathbf{x}_i)), 1 \leq i \leq n\}$ et interpoler ensuite les résidus par un interpolateur à noyaux. Les mêmes métamodèles peuvent être obtenus sous des interprétations différentes. Une approche bayésienne donne le même estimateur qu'une projection dans un sous espace fonctionnel d'un RKHS.
À partir d'un ensemble de données, il est possible de construire différents modèles et de les comparer sans faire de nouveaux appels au code grâce à la méthode de validation croisée présentée.

Dans cette thèse, nous nous intéressons aux métamodèles construits à l'aide de méthodes à noyaux. Ce sont des métamodèles souples qui englobent une grande diversité de fonctions. Leur interprétation bayésienne (krigeage) permet de modéliser l'incertitude introduite en remplaçant f par un métamodèle $\hat{f}$. Cette incertitude peut ensuite être incluse dans le cadre d'applications statistiques et il est possible d'en tirer parti. Sacks *et al.* (1989b) ont introduit cette méthode dans le contexte des expériences simuléees. Elle a ensuite été grandement utilisée (voir par exemple, Koehler et Owen, 1996; Santner *et al.*, 2003; Fang *et al.*, 2006). De plus, Simpson *et al.* (2001) ont testé de manière empirique des métamodèles sur différents exemples. Ils ont notamment comparé des modèles polynomiaux, des réseaux de neurones et de krigeage. Leurs recommandations sont :

1. Les modèles polynomiaux sont les plus simples à mettre en oeuvre. Dans des optiques d'exploration pour des codes de calculs déterministes, ils s'avèrent utiles si la fonction f est plutôt régulière et la dimension d'entrée d est faible. Les valeurs des coefficients sont interprétables et apportent une information intéressante.
2. Le krigeage est un très bon choix dans le cas où f est hautement non linéaire et son nombre de variables d'entrée est raisonnable (moins de 50).
3. Les perceptrons multi-couches sont intéressants dans le cas où la dimension des entrées est grande. Cependant, leur construction requiert un temps de calcul important.

2.2 Krigeage ou interpolation à noyaux

Dans cette partie, nous décrivons de manière plus précise les méthodes à noyaux qui ont été évoquées dans la section 2.1.5. Pour la simplicité de l'exposé, nous considérons la sortie y comme scalaire ($y \in \mathbb{R}$). Pour une sortie vectorielle, nous pouvons constuire un métamodèle par dimension. Nous ne traitons pas ici les sorties fonctionnelles (Carroll *et al.*, 1997), le cas où des valeurs des dérivées de f sont observées (Morris *et al.*, 1993) et les modèles de corrélations croisées entre les sorties (Kennedy et O'Hagan, 2000).

Nous commençons par définir le concept de base, à savoir un noyau défini positif (sur $\mathbb{R}^d$).

Définition 2.1. *Un noyau est une fonction symétrique $K : \mathbb{R}^d \times \mathbb{R}^d \to \mathbb{R}$, il est défini positif si $\forall m \in \mathbb{N},\ \forall (\lambda_1, \ldots, \lambda_m) \in \mathbb{R}^m,\ \forall (\mathbf{x}_1, \ldots, \mathbf{x}_m) \in \left(\mathbb{R}^d\right)^m$,*

$$\sum_{1 \leq l,m \leq N} \lambda_l \lambda_m K(\mathbf{x}_l, \mathbf{x}_m) \geq 0\,.$$

Nous pouvons donner une proposition importante pour éviter les confusions et faire une remarque qui permettra de fixer le vocabulaire.

Proposition 2.1. *Un noyau est défini positif si et seulement si, pour tout $n \in \mathbb{N}$, et pour tout ensemble de points $(\mathbf{x}_1, \ldots, \mathbf{x}_n) \in \mathbb{R}^d$, la matrice de Gram, $(K(\mathbf{x}_i, \mathbf{x}_j)_{1 \leq i,j \leq n})$ est positive.*

Remarque 2.1. *Si les matrices de Gram correspondant à un noyau K, pour des vecteurs de E distincts, sont toutes définies positives, le noyau K sera dit strictement défini positif. Cela garantit l'inversibilité des matrices de Gram.*

Nous présentons avant tout une vision statistique de la méthode d'interpolation en nous cantonnant aux noyaux définis positifs. Cette méthode consiste en une modélisation de f comme une réalisation d'un processus gaussien. Les propriétés et des exemples de noyaux utilisés pour définir la covariance des processus seront donnés dans la section 2.2.2. Ensuite, nous nous placerons dans le cadre des méthodes d'interpolation à noyaux et nous traiterons de la régression régularisée.

2.2.1 Modélisation par des processus gaussiens

Cette modélisation vient du krigeage qui a été à l'origine introduit par Krige (1951) dans son mémoire de master afin d'analyser des données minières. Ensuite, Matheron (1963) a proposé la méthode de krigeage gaussien pour modéliser les données spatiales en géostatistiques (voir aussi Cressie, 1993; Stein, 1999). Ce sont Sacks *et al.* (1989a) qui ont utilisé cette modélisation pour la construction de métamodèles dans le cadre des expériences simulées. Cette vision permet d'obtenir, en plus du métamodèle, un indicateur de l'incertitude accordée à une prédiction du métamodèle en un point donné. L'idée naturelle est de dire que si l'on connaît des évaluations de la fonction f aux points du plan d'expérience D, on dispose d'informations sur $f(\mathbf{x}_0)$ où $\mathbf{x}_0 \notin D$. On relie ces données à l'aide d'une modélisation par un processus gaussien :

$$\forall \mathbf{x} \in E,\ Y(\mathbf{x}) = \sum_{i=1}^{p} \beta_i h_i(\mathbf{x}) + Z(\mathbf{x}) = H(\mathbf{x})^T \boldsymbol{\beta} + Z(\mathbf{x})\,, \tag{2.9}$$

où

- $H(\mathbf{x}) = (h_1(\mathbf{x}), \ldots, h_p(\mathbf{x}))$ est un vecteur de fonctions de régression fixées,
- $\boldsymbol{\beta} = (\beta_1, \ldots, \beta_p)$ est un vecteur de paramètres,
- Z est un processus gaussien centré caractérisé par sa fonction de covariance $\mathrm{Cov}(Z(\mathbf{x}), Z(\mathbf{x}')) = \sigma^2 K_{\boldsymbol{\theta}}(\mathbf{x}, \mathbf{x}')$ où $K_{\boldsymbol{\theta}}$ est un noyau symétrique strictement défini positif tel que pour tout $\mathbf{x}$, $K_{\boldsymbol{\theta}}(\mathbf{x}, \mathbf{x}) = 1$ (on suppose ainsi que la variance est en tout point égale à σ^2). Ce noyau est la fonction d'autocorrélation de Y.

Les paramètres $\boldsymbol{\beta}, \sigma^2, \boldsymbol{\theta}$ sont inconnus. Certaines modélisations les supposent fixés tandis qu'en pratique ils sont préalablement estimés à partir des données $\{(\mathbf{x}_i, y_i), 1 \leq i \leq n\}$ par des méthodes décrites dans la suite. Ils peuvent aussi être considérés comme inconnus et leur estimation est alors prise en compte dans la construction du métamodèle ainsi que l'incertitude associée. Toutefois, cela peut mener à des métamodèles trop complexes. La famille de fonctions de covariance peut être choisie suivant la régularité supposée de la fonction inconnue f ou à l'aide d'une méthode de validation croisée qui comparerait plusieurs métamodèles construits avec des fonctions de covariance différentes.

Ce modèle conduit à faire l'hypothèse selon laquelle la fonction f est une réalisation du processus gaussien Y. En adoptant une vision bayésienne, on interpréte ce processus comme une loi a priori sur la fonction f.

À partir des évaluations $\{y_1, \ldots, y_n\}$ de la fonction f respectivement aux points du plan d'expérience $D = \{\mathbf{x}_1, \ldots, \mathbf{x}_n\}$, on construit le métamodèle $\hat{f}$. Pour un point $\mathbf{x}_0 \in E - D$, on s'intéresse à la distribution de $Y(\mathbf{x}_0)$ conditionnellement à $\{Y(\mathbf{x}_1) = y_1, \ldots, Y(\mathbf{x}_n) = y_n\} \equiv \{\mathbf{Y}_D = \mathbf{y}_D\}$. On note $\mathbf{Y}_D = (Y(\mathbf{x}_1), \ldots, Y(\mathbf{x}_n))$ et $\mathbf{y}_D = (y_1, \ldots, y_n)$. Les paramètres $\boldsymbol{\beta}, \sigma^2, \boldsymbol{\theta}$ sont ici fixés.

Proposition 2.2. *$Y(\mathbf{x}_0)$ conditionnellement aux évaluations $\mathbf{Y}_D = \mathbf{y}_D$ suit une loi normale $\mathcal{N}_1(\mu_{0|D}, \sigma^2_{0|D})$ où*

$$\begin{array}{l} \mu_{\mathbf{x}_0|D} = \mathbb{E}(Y(\mathbf{x}_0)|\mathbf{Y}_D = y_D) = H(\mathbf{x}_0)^T\boldsymbol{\beta} + \Sigma^T_{\mathbf{x}_0 D}\Sigma^{-1}_{DD}(y_D - H_D\boldsymbol{\beta}) \\ \sigma^2_{\mathbf{x}_0|D} = Var(Y(\mathbf{x}_0)|\mathbf{Y}_D = y_D) = \sigma^2\left(1 - \Sigma^T_{\mathbf{x}_0 D}\Sigma^{-1}_{DD}\Sigma_{\mathbf{x}_0 D}\right) \end{array}, \tag{2.10}$$

avec $H_D = (H(\mathbf{x}_1), \ldots, H(\mathbf{x}_n))^T$, $(\Sigma_{DD})_{1\leq i,j\leq n} = K_{\boldsymbol{\theta}}(\mathbf{x}_i, \mathbf{x}_j)$, la matrice de Gram correspondant aux points de D inversible car le noyau $K_{\boldsymbol{\theta}}$ est supposé strictement défini positif, et $\Sigma_{\mathbf{x}_0 D} = (K_{\boldsymbol{\theta}}(\mathbf{x}_i, \mathbf{x}_0))^T_{1\leq i\leq n}$.

La moyenne peut être utilisée afin de prédire, d'approcher la valeur $f(\mathbf{x}_0)$. Ainsi on peut choisir le métamodèle $\hat{f} : \mathbf{x}_0 \mapsto \mu_{\mathbf{x}_0|D}$. La variance au point $\mathbf{x}_0$ représente une variance de prédiction et elle décrit l'incertitude associée à la prédiction de $Y(\mathbf{x}_0)$ par $\hat{f}(\mathbf{x}_0)$. Ceci permet de donner un intervalle de confiance pour le métamodèle. En effet, on a

$$\frac{Y(\mathbf{x}_0) - \mu_{\mathbf{x}_0|D}}{\sqrt{\sigma^2_{\mathbf{x}_0|D}}} \sim \mathcal{N}(0, 1)\,. \tag{2.11}$$

Remarque 2.2. *Jones* et al. *(1998) proposent de valider le modèle (2.9) par une méthode de type validation croisée "leave-one-out" qui consiste à vérifier que, pour une grande majorité (99.7%) des $i = 1, \ldots, n$, on a*

$$\frac{f(\mathbf{x}_i) - \hat{f}_{-i}(\mathbf{x}_i)}{\sqrt{\sigma^2_{\mathbf{x}_0|D_{-i}}}} \in [-3, 3]\,,$$

où $\hat{f}_{-i}(\mathbf{x}_i)$ et $\sigma^2_{\mathbf{x}_0|D_{-i}}$ correspondent respectivement à la moyenne et à la variance a posteriori données par (2.10) et construites à partir du plan d'expérience $D_{-i} = \{\mathbf{x}_1, \ldots, \mathbf{x}_{i-1}, \mathbf{x}_{i+1}, \ldots, \mathbf{x}_n\}$.

Nous nous intéressons à présent aux propriétés du prédicteur.

Définition 2.2. *Un prédicteur de $Y(\mathbf{x}_0)$, noté $\hat{Y}(\mathbf{x}_0)$, est dit le meilleur prédicteur linéaire (BLP : Best Linear Predictor) s'il minimise*

$$MSE(\mathbf{x}_0) = \mathbb{E}\left(Y(\mathbf{x}_0) - \tilde{Y}(\mathbf{x}_0)\right)^2 ,$$

avec

$$\tilde{Y}(\mathbf{x}_0) = \lambda_0 + \boldsymbol{\lambda}^T \mathbf{Y}_D ,$$

où $\lambda_0 \in \mathbb{R}$ et $\boldsymbol{\lambda} \in \mathbb{R}^n$.

Proposition 2.3. *Le prédicteur*

$$\mathbf{x}_0 \mapsto H(\mathbf{x}_0)^T \boldsymbol{\beta} + \Sigma^T_{\mathbf{x}_0 D} \Sigma^{-1}_{DD} (Y_D - H_D \boldsymbol{\beta}) = \mathbb{E}(Y(\mathbf{x}_0)|\mathbf{Y}_D)$$

est le meilleur prédicteur linéaire de $Y(\mathbf{x}_0)$ sous les hypothèses du modèle (2.9) avec $\boldsymbol{\beta}, \sigma^2, \boldsymbol{\theta}$ fixés. Il est évidemment sans biais et son erreur quadratique moyenne de prédiction en $\mathbf{x}_0$ est égale à $MSE(\mathbf{x}_0) = \sigma^2_{\mathbf{x}_0|D}$ donnée par (2.10).

Remarque 2.3. *Si $\mathbf{x}_0 = \mathbf{x}_i \in D$,*

$$\begin{array}{l} \mu_{\mathbf{x}_0|D} = y_i \\ \sigma^2_{\mathbf{x}_0|D} = 0 \end{array} . \tag{2.12}$$

Remarque 2.4. *Il est possible de considérer le processus a posteriori noté Y^D qui reste gaussien, c'est-à-dire qui est conditionné aux observations. Sa moyenne en un point $\mathbf{x}_0$ est $\mu_{\mathbf{x}_0|D}$, sa variance est $\sigma^2_{\mathbf{x}_0|D}$ et sa covariance est :*

$$\forall \mathbf{x}, \mathbf{x}', \ Cov(Y^D_{\mathbf{x}}, Y^D_{\mathbf{x}'}) = \sigma^2 \left(K_{\boldsymbol{\theta}}(\mathbf{x}, \mathbf{x}') - \Sigma^T_{\mathbf{x}D} \Sigma^{-1}_{DD} \Sigma_{\mathbf{x}'D}\right) .$$

Il est contraint de passer par les valeurs observées aux points du plan d'expérience, c'est-à-dire :

$$\forall i = 1, \ldots, n, \ Y^D(\mathbf{x}_i) = y_i .$$

En pratique, le vecteur $\boldsymbol{\beta}$ est inconnu. On peut l'estimer en utilisant la méthode des moindres carrés généralisés qui correspond à la méthode du maximum de vraisemblance dans les hypothèses du modèle (2.9). Ceci donne

$$\hat{\boldsymbol{\beta}} = (H_D^T \Sigma^{-1}_{DD} H_D)^{-1} H_D^T \Sigma^{-1}_{DD} Y_D . \tag{2.13}$$

Finalement cela revient à appliquer une méthode de régression généralisée et à conditionner sur les résidus de la régression.
On obtient alors le prédicteur de $Y(\mathbf{x}_0)$:

$$\hat{Y}(\mathbf{x}_0) = H(\mathbf{x}_0)^T \hat{\boldsymbol{\beta}} + \Sigma^T_{\mathbf{x}_0 D} \Sigma^{-1}_{DD} (\mathbf{Y}_D - H_D \hat{\boldsymbol{\beta}}) \tag{2.14}$$

et l'erreur quadratique moyenne de prédiction est :

$$\begin{array}{l} \mathrm{Var}(\hat{Y}(\mathbf{x}_0)) = \mathbb{E}(\hat{Y}(\mathbf{x}_0) - Y(\mathbf{x}_0))^2 \\ = \sigma^2 (1 + u(\mathbf{x}_0)^T (H_D^T \Sigma^{-1}_{DD} H_D)^{-1} u(\mathbf{x}_0) - \Sigma^T_{\mathbf{x}_0 D} \Sigma^{-1}_{DD} \Sigma_{\mathbf{x}_0 D}) , \end{array} \tag{2.15}$$

où $u(\mathbf{x}_0) = (H_D^T \Sigma_{DD}^{-1} \Sigma_{\mathbf{x}_0 D} - H(\mathbf{x}_0))$.

Suivant que l'on suppose fixé ou non le vecteur de paramètre $\boldsymbol{\beta}$ dans la modélisation (2.9), l'erreur quadratique moyenne du prédicteur en un point $\mathbf{x}_0 \in E$, $MSE(\mathbf{x}_0)$ a des expressions différentes.

– Si $\boldsymbol{\beta}$ est fixé, l'erreur quadratique moyenne du prédicteur est :

$$MSE(\mathbf{x}_0) = \sigma^2 \left(1 - \Sigma_{\mathbf{x}_0 D}^T \Sigma_{DD}^{-1} \Sigma_{\mathbf{x}_0 D}\right) .$$

– Si $\boldsymbol{\beta}$ n'est pas fixé, l'erreur quadratique moyenne du prédicteur est :

$$MSE(\mathbf{x}_0) = \sigma^2 (1 + u(\mathbf{x}_0)^T (H_D^T \Sigma_{DD}^{-1} H_D)^{-1} u(\mathbf{x}_0) - \Sigma_{\mathbf{x}_0 D}^T \Sigma_{DD}^{-1} \Sigma_{\mathbf{x}_0 D}) ,$$

où $u(\mathbf{x}) = (H_D^T \Sigma_{DD}^{-1} \Sigma_{\mathbf{x}_0 D} - H(\mathbf{x}_0))$.

L'abréviation MSE pourra faire référence à l'une ou l'autre forme suivant l'hypothèse faite sur $\boldsymbol{\beta}$. Dans ces deux cas, l'erreur moyenne de prédiction est aussi la variance du prédicteur puisque celui-ci est sans biais.

Proposition 2.4. *$\hat{Y}(\mathbf{x}_0)$ est le meilleur prédicteur linéaire sans biais (BLUP : Best Linear Unbiased Predictor) de $Y(\mathbf{x}_0)$ sous les hypothèses du modèle (2.9) avec $\boldsymbol{\beta}$, σ^2 inconnus et la matrice de covariance supposée connue.*

La remarque 2.3 est aussi vérifiée par le prédicteur (2.14). Puisque $\boldsymbol{\beta}$ n'est plus fixé ici, la taille du modèle (2.9) est augmentée. En revanche, on ne considère plus que la classe des prédicteurs linéaires sans biais. Le paramètre σ^2 n'intervient pas dans l'expression du prédicteur et il n'est pas nécessaire de connaître sa valeur pour montrer que $\hat{Y}(\mathbf{x}_0)$ (2.14) est le BLUP. Par contre, il est nécessaire au calcul de l'erreur quadratique moyenne MSE. On peut alors l'estimer par son estimateur du maximum de vraisemblance :

$$\hat{\sigma}^2 = \frac{1}{n} (\mathbf{Y}_D - H_D \hat{\boldsymbol{\beta}})^T \Sigma_{DD}^{-1} (\mathbf{Y}_D - H_D \hat{\boldsymbol{\beta}}) . \tag{2.16}$$

On peut aussi adopter une approche bayésienne en utilisant des lois a priori sur $\boldsymbol{\beta}$ (Santner *et al.*, 2003). Pour $\sigma^2, \boldsymbol{\theta}$ fixés, nous avons

$$(Y(\mathbf{x}_0), Y_D | \boldsymbol{\beta}) \sim \mathcal{N}_{n+1} \left[\begin{pmatrix} h(\mathbf{x}_0)^T \\ H_D \end{pmatrix} \boldsymbol{\beta}, \sigma^2 \begin{pmatrix} 1 & \Sigma_{\mathbf{x}_0 D}^T \\ \Sigma_{\mathbf{x}_0 D} & \Sigma_{DD} \end{pmatrix} \right] . \tag{2.17}$$

Théorème 2.1. *Nous nous plaçons sous le modèle de (2.17).*

(i) Si

$$\boldsymbol{\beta} \sim \mathcal{N}_p(\mathbf{b}_0, \tau^2 V_0) ,$$

avec $\mathbf{b}_0$, τ^2, et V_0 fixés, alors $Y(\mathbf{x}_0)$ suit la loi conditionnée aux évaluations :

$$(Y(\mathbf{x}_0) | Y_D = y_D) \sim \mathcal{N}(\mu_{\mathbf{x}_0|D,(i)}, \sigma^2_{\mathbf{x}_0|D,(i)}) , \tag{2.18}$$

avec

$$\mu_{\mathbf{x}_0|D,(i)} = H(\mathbf{x}_0)^T \mu_{\boldsymbol{\beta}|D} + \Sigma_{\mathbf{x}_0 D}^T \Sigma_{DD}^{-1} (y_D - H_D \mu_{\boldsymbol{\beta}|D}) ,$$

où

$$\mu_{\beta|D} = \left(\frac{H_D^T \Sigma_{DD}^{-1} H_D}{\sigma^2} + \frac{V_0^{-1}}{\tau^2} \right)^{-1} \left(\frac{H_D^T \Sigma_{DD}^{-1} y_D}{\sigma^2} + \frac{V_0^{-1} \mathbf{b}_0}{\tau^2} \right) ,$$

et

$$\sigma^2_{\mathbf{x}_0|D,(i)} = \sigma^2 \left(1 - (H(\mathbf{x}_0)^T, \Sigma_{\mathbf{x}_0 D}^T) \begin{bmatrix} -\frac{\sigma^2}{\tau^2} V_0^{-1} & H_D^T \\ H_D & \Sigma_{DD} \end{bmatrix}^{-1} \begin{pmatrix} H(\mathbf{x}_0) \\ \Sigma_{\mathbf{x}_0 D} \end{pmatrix} \right) .$$

(ii) Si

$$\pi(\boldsymbol{\beta}) \propto 1 \text{ } \textit{(loi a priori de Laplace)}\,,$$

sur $\mathbb{R}^p$*, alors* $Y(\mathbf{x}_0)$ *suit la loi conditionnée aux évaluations*

$$(Y(\mathbf{x}_0)|Y_D = y_D) \sim \mathcal{N}(\mu_{\mathbf{x}_0|D,(ii)}, \sigma^2_{\mathbf{x}_0|D,(ii)})\,, \tag{2.19}$$

avec

$$\mu_{\mathbf{x}_0|D,(ii)} = H(\mathbf{x}_0)^T\boldsymbol{\beta}^* + \Sigma^T_{\mathbf{x}_0 D}\Sigma^{-1}_{DD}(y_D - H_D\boldsymbol{\beta}^*)\,,$$

où

$$\boldsymbol{\beta}^* = (H_D^T\Sigma^{-1}_{DD}H_D)^{-1}H_D^T\Sigma^{-1}_{DD}y_D\,,$$

et

$$\sigma^2_{\mathbf{x}_0|D,(ii)} = \sigma^2\left(1 - (H(\mathbf{x}_0)^T, \Sigma^T_{\mathbf{x}_0 D})\begin{bmatrix} 0 & H_D^T \\ H_D & \Sigma_{DD}\end{bmatrix}^{-1}\begin{pmatrix} H(\mathbf{x}_0) \\ \Sigma_{\mathbf{x}_0 D}\end{pmatrix}\right).$$

Dans le cas de la loi a priori de Laplace (ii), l'espérance de la loi a posteriori $\mu_{\mathbf{x}_0|D,(ii)}$ est égale à la prédiction par BLUP de $Y(\mathbf{x}_0)$. Par ailleurs, la variance du BLUP (2.15) est une autre écriture de la variance de la loi a posteriori $\sigma^2_{\mathbf{x}_0|D,(ii)}$.

Si on suppose également σ^2 aléatoire, on peut utiliser la décomposition a priori suivante

$$\pi(\boldsymbol{\beta}, \sigma^2) = \pi(\boldsymbol{\beta}|\sigma^2)\pi(\sigma^2)\,.$$

Nous nous intéressons à ces quatre combinaisons :

(1) $\beta|\sigma^2 \sim \mathcal{N}(\mathbf{b}_0, \sigma^2 V_0)$ et $\sigma^2 \sim c_0/\chi^2_{\nu_0}$ (loi inverse χ^2),

(2) $\beta|\sigma^2 \sim \mathcal{N}(\mathbf{b}_0, \sigma^2 V_0)$ et $\pi(\sigma^2) \propto 1/\sigma^2$ (loi a priori de Jeffrey),

(3) $\pi(\beta|\sigma^2) \propto 1$ et $\sigma^2 \sim c_0/\chi^2_{\nu_0}$,

(4) $\pi(\beta|\sigma^2) \propto 1$ et $\pi(\sigma^2) \propto 1/\sigma^2$.

Pour $\boldsymbol{\theta}$ fixé, rappelons que

$$(Y(\mathbf{x}_0), Y_D|\boldsymbol{\beta}, \sigma^2) \sim \mathcal{N}_{n+1}\left[\begin{pmatrix} h(\mathbf{x}_0)^T \\ H_D\end{pmatrix}\boldsymbol{\beta}, \sigma^2\begin{pmatrix} 1 & \Sigma^T_{\mathbf{x}_0 D} \\ \Sigma_{\mathbf{x}_0 D} & \Sigma_{DD}\end{pmatrix}\right]. \tag{2.20}$$

Théorème 2.2. *Si le vecteur* $(\boldsymbol{\beta}, \sigma^2)$ *suit une des lois a priori décrites ci-dessus et sous le modèle (2.20), nous avons*

$$Y_{\mathbf{x}_0}|\mathbf{Y}_D \sim T(\nu_i, \mu_i, \sigma_i^2)\,, \tag{2.21}$$

où $T_1(.,.,.)$ *représente la loi de Student non centrée, avec*

$$\nu_i = \begin{cases} n + \nu_0, & i = (1) \\ n, & i = (2) \\ n - p + \nu_0, & i = (3) \\ n - p, & i = (4)\end{cases},$$

$$\mu_i = \mu_i(\mathbf{x}_0) = \begin{cases} H(\mathbf{x}_0)^T\mu_{\boldsymbol{\beta}|D} + \Sigma^T_{\mathbf{x}_0 D}\Sigma^{-1}_{DD}(y_D - H_D\mu_{\boldsymbol{\beta}|D}), & i = (1)\ ou\ (2) \\ H(\mathbf{x}_0)^T\boldsymbol{\beta}^* + \Sigma^T_{\mathbf{x}_0 D}\Sigma^{-1}_{DD}(y_D - H_D\boldsymbol{\beta}^*), & i = (3)\ ou\ (4)\end{cases},$$

avec

$$\begin{array}{ll} \mu_{\boldsymbol{\beta}|D} & = \left(H_D^T\Sigma_{DD}^{-1}H_D + V_0^{-1}\right)^{-1}\left(H_D^T\Sigma_{DD}^{-1}y_D + V_0^{-1}\mathbf{b}_0\right) \\ \boldsymbol{\beta}^* & = (H_D^T\Sigma_{DD}^{-1}H_D)^{-1}H_D^T\Sigma_{DD}^{-1}y_D) \end{array} ,$$

et

$$\sigma_i^2 = \sigma_i^2(\mathbf{x}_0) = \frac{Q_i^2}{\nu_i}\left(1 - (H(\mathbf{x}_0)^T, \Sigma_{\mathbf{x}_0 D}^T)\begin{bmatrix} V_i & H_D^T \\ H_D & \Sigma_{DD}\end{bmatrix}^{-1}\begin{pmatrix} H(\mathbf{x}_0) \\ \Sigma_{\mathbf{x}_0 D}\end{pmatrix}\right), \tag{2.22}$$

avec

$$V_i = \begin{cases} -V_0^{-1}, & i = (1)\ ou\ (2) \\ 0, & i = (3)\ ou\ (4) \end{cases}$$

et

$$Q_i^2 = \begin{cases} c_0 + Q_2^2, & i = (1) \\ Q_4^2 + (\mathbf{b}_0 - \boldsymbol{\beta}^*)^T\left(V_0 + (H_D^T\Sigma_{DD}^{-1}H_D)^{-1}\right)^{-1}(\mathbf{b}_0 - \boldsymbol{\beta}^*), & i = (2) \\ c_0 + Q_4^2, & i = (3) \\ \mathbf{y}_D^T\left(\Sigma_{DD}^{-1} - \Sigma_{DD}^{-1}H_D\left(H_D^T\Sigma_{DD}^{-1}H_D\right)^{-1}H_D^T\Sigma_{DD}^{-1}\right)\mathbf{y}_D & i = (4) \end{cases} .$$

Les lois a priori de Laplace sur $\boldsymbol{\beta}$ nous mènent à la même prédiction que le BLUP. Le fait de supposer σ^2 aléatoire nous amène ici à avoir

$$\frac{Y(\mathbf{x}_0) - \mu_{\mathbf{x}_0|D}}{\sqrt{\sigma^2_{\mathbf{x}_0|D}}} \sim T_1(\nu_i, 0, 0)\,. \tag{2.23}$$

au lieu de (2.11). Ceci conduira à des intervalles de confiance plus larges sur la prédiction de $Y(\mathbf{x}_0)$.

Jusqu'à présent, tous les résultats donnés présupposaient le paramètre $\boldsymbol{\theta}$ de la fonction de covariance fixé et connu. En pratique, on utilise un estimateur noté $\hat{\boldsymbol{\theta}}$, ce qui donne lieu au prédicteur

$$\hat{Y}(\mathbf{x}_0) = H(\mathbf{x}_0)^T\hat{\boldsymbol{\beta}} + \hat{\Sigma}_{\mathbf{x}_0 D}^T\hat{\Sigma}_{DD}^{-1}(\mathbf{Y}_D - H_D\hat{\boldsymbol{\beta}})\,, \tag{2.24}$$

où $\hat{\Sigma}_{\mathbf{x}_0 D}$ et $\hat{\Sigma}_{DD}$ sont calculés avec le noyau $K_{\hat{\boldsymbol{\theta}}}$. Ce prédicteur est appelé EBLUP (*Empirical Best Linear Unbiaised Predictor*) bien qu'il ne soit ni linéaire ni sans biais du fait de l'estimation de $\boldsymbol{\theta}$.

Le paramètre $\boldsymbol{\theta}$ peut être estimé par une méthode du maximum de vraisemblance. La log-vraisemblance est à une constante près :

$$\ell(\boldsymbol{\beta}, \sigma^2, \boldsymbol{\theta}) = -\frac{n}{2}\log(\sigma^2) - \frac{1}{2}\log|R(\boldsymbol{\theta})| - \frac{1}{2\sigma^2}(\mathbf{y}_D - H_D\boldsymbol{\beta})^T\Sigma_{DD}^{-1}(\mathbf{y}_D - H_D\boldsymbol{\beta})\,. \tag{2.25}$$

À $\boldsymbol{\theta}$ donné, (2.13) et (2.16) donnent respectivement les estimateurs du maximum de vraisemblance de $\boldsymbol{\beta}$ et de σ^2. Étant donné que

$$\mathbb{E}\left(\frac{\partial^2\ell(\boldsymbol{\beta}, \sigma^2, \boldsymbol{\theta})}{\partial\boldsymbol{\beta}\partial\sigma^2}\right) = \mathbf{0} \text{ et } \mathbb{E}\left(\frac{\partial^2\ell(\boldsymbol{\beta}, \sigma^2, \boldsymbol{\theta})}{\partial\boldsymbol{\beta}\partial\boldsymbol{\theta}}\right) = \mathbf{0}\,,$$

la matrice d'information de Fisher est diagonale par blocs et l'estimateur du maximum de vraisemblance de $\boldsymbol{\beta}$ est donc asymptotiquement indépendant de celui de $(\sigma^2, \boldsymbol{\theta})$. Par ailleurs,

on peut estimer $\boldsymbol{\beta}$ et $(\sigma^2, \boldsymbol{\theta})$ séparément. Si on substitue dans la log vraisemblance les estimateurs de $\boldsymbol{\beta}$ et σ^2, on obtient :

$$\ell(\hat{\boldsymbol{\beta}}, \hat{\sigma}^2, \boldsymbol{\theta}) = -\frac{n}{2}\log(\hat{\sigma}^2) - \frac{1}{2}\log|R(\boldsymbol{\theta})| - \frac{1}{2}n\,.$$

Ainsi l'estimateur du maximum de vraisemblance de $\boldsymbol{\theta}$ minimise la fonction

$$(\hat{\sigma}^2(\boldsymbol{\theta}))^n |\Sigma_{DD}(\boldsymbol{\theta})|\,, \tag{2.26}$$

avec $\sigma^2(\boldsymbol{\theta})$ donné par (2.16). Le fait de pouvoir estimer séparément $\boldsymbol{\beta}$ et $(\sigma^2, \boldsymbol{\theta})$ incite à proposer un algorithme qui initialiserait $\boldsymbol{\beta}$ à l'estimateur des moindres carrés (non généralisés), chercherait un vecteur $(\sigma^2, \boldsymbol{\theta})$ minimisant (2.26) où $\boldsymbol{\beta}$ a été fixé, puis pour $\boldsymbol{\theta}$ obtenu calculerait $\hat{\boldsymbol{\beta}}$ par l'estimateur (2.13) et itérerait ces deux dernières étapes jusqu'à convergence des estimations.

Les programmes `GaSP` (Gaussian Stochastic Process, Welch *et al.*, 1992), `PErK` (Parametric EmpiRical Kriging Williams, 2001) et `DACE` (Design and Analysis of Computer Experiments Lophaven *et al.*, 2002b) permettent de calculer le EBLUP avec une estimation des paramètres par maximum de vraisemblance pour des fonctions de covariance généralement usitées.

Afin d'obtenir des estimateurs moins biaisés de σ^2 et $\boldsymbol{\theta}$, on peut utiliser la méthode du maximum de vraisemblance restreinte (Patterson et Thompson, 1971). Ainsi cela conduit à estimer σ^2 par :

$$\tilde{\sigma}^2 = \frac{1}{n-p}\left(\mathbf{y}_D - H_D\boldsymbol{\beta}^*\right)^T \Sigma_{DD}^{-1}\left(\mathbf{y}_D - H_D\boldsymbol{\beta}^*\right)\,.$$

L'estimateur du paramètre $\boldsymbol{\theta}$ est alors déterminé comme le minimiseur de

$$(\tilde{\sigma}^2(\boldsymbol{\theta}))^{n-p} |\Sigma_{DD}(\boldsymbol{\theta})|\,. \tag{2.27}$$

Li et Sudjianto (2005) ont proposé une autre approche qui consiste à pénaliser la vraisemblance pour empêcher d'obtenir des estimateurs de trop grande variance. Fang *et al.* (2006) comparent sur un exemple différentes pénalisations de la vraisemblance et montrent l'intérêt de celle-ci pour des plans d'expérience contenant peu de points.

Il y a aussi la méthode de validation croisée qui a la particularité de ne pas utiliser le modèle (2.9) pour donner une estimation des paramètres. En effet, pour $\boldsymbol{\theta}$ fixé, on considère les métamodèles $\hat{f}_{-i}(.,|\boldsymbol{\theta})$ comme des réalisations de BLUP construits à partir des données $\{(\mathbf{x}_j, y_j), 1 \leq j \leq n, j \neq i\}$ pour $i = 1, \ldots, n$. Et, on choisit $\boldsymbol{\theta}$ qui réalise le minimum de la quantité

$$\frac{1}{n}\sum_{i=1}^{n}(f(\mathbf{x}_i) - \hat{f}_{-i}(\mathbf{x}_i|\boldsymbol{\theta}))^2\,.$$

On peut éventuellement pondérer cette estimation de l'erreur quadratique intégrée comme dans (2.3). Pour construire les BLUP, on a utilisé les estimations du maximum de vraisemblance de $\boldsymbol{\beta}$ et σ^2 mais on pourrait aussi les estimer par cette méthode de validation croisée en construisant le prédicteur avec $\boldsymbol{\beta}$ fixé.

Dans le cas où le modèle (2.9) est posé conditionnellement au vecteur de paramètres $(\boldsymbol{\beta}, \sigma^2, \boldsymbol{\theta})$ pour lequel une loi a priori a été proposée, le meilleur prédicteur au sens de la minimisation de l'erreur quadratique est

$$\mathbb{E}(Y(\mathbf{x}_0)|\mathbf{Y}_D) = \mathbb{E}(\mathbb{E}(Y(\mathbf{x}_0)|\mathbf{Y}_D, \boldsymbol{\theta})|\mathbf{Y}_D)\,. \tag{2.28}$$

La loi a posteriori $\boldsymbol{\theta}|\mathbf{Y}_D$ n'est calculable sous une forme explicite (mais très complexe) que pour des lois a priori simples. Une solution consiste à s'intéresser uniquement au mode a posteriori de $[\boldsymbol{\theta}|\mathbf{Y}_D]$.

Remarque 2.5. *Pour la plupart des EBLUP proposés, la variance de prédiction (typiquement (2.15)) est estimée en insérant $\hat{\boldsymbol{\theta}}$ dans son expression sans tenir compte de l'incertitude liée à l'estimation de ce paramètre $\boldsymbol{\theta}$. Seule la méthodologie "totalement" bayésienne (2.28) l'intègre dans son calcul qui est néanmoins très lourd. Zimmerman et Cressie (1992) ont montré que la variance de prédiction où l'on a inséré $\hat{\boldsymbol{\theta}}$ sous-estime la variance de prédiction effective et ont proposé une correction. Néanmoins, Prasad et Rao (1990) ont prouvé que l'erreur était asymptotiquement négligeable dans des cas de modèles linéaires généralisés. Une méthode de bootstrap est proposée par den Hertog* et al. *(2006) pour estimer la variance de krigeage en tenant compte de l'estimation de $\boldsymbol{\theta}$.*

Santner *et al.* (2003) proposent une étude de différents EBLUP obtenus par les différentes méthodes d'estimation des paramètres et recommandent de privilégier les prédicteurs dont les paramètres ont été estimés par une méthode du maximum de vraisemblance restreint ou non.

2.2.2 Les noyaux

On a supposé que le processus Y dans la modélisation (2.9) avait une variance égale à σ^2 en tout point $\mathbf{x} \in E$. On suppose ici que le processus Y est stationnaire au second ordre. Ceci implique que le processus Z centré est fortement stationnaire. En particulier, on suppose que la covariance s'écrit, $\forall \mathbf{x}, \mathbf{x}' \in E$,

$$\mathrm{Cov}(Y(\mathbf{x}), Y(\mathbf{x}')) = \mathrm{Cov}(Z(\mathbf{x}), Z(\mathbf{x}')) = \sigma^2 K_{\boldsymbol{\theta}}(\mathbf{x}, \mathbf{x}') = \sigma^2 C_{\boldsymbol{\theta}}(\mathbf{x} - \mathbf{x}')\,, \tag{2.29}$$

où $C_{\boldsymbol{\theta}} : E \subset \mathbb{R}^d \to \mathbb{R}$ telle que $C_{\boldsymbol{\theta}}(0) = 1$. Nécessairement, $C_{\boldsymbol{\theta}}(\mathbf{x} - \mathbf{x}') = C_{\boldsymbol{\theta}}(\mathbf{x}' - \mathbf{x})$ car $K_{\boldsymbol{\theta}}$ est symétrique. La fonction $C_{\boldsymbol{\theta}}$ est construite comme un produit de fonctions de corrélation univariées. Le vecteur de paramètre $\boldsymbol{\theta}$ est décomposé ainsi $\boldsymbol{\theta} = (\theta_1, \ldots, \theta_d, \nu)$. À chaque dimension des variables d'entrée correspond un θ_j $(j = 1, \ldots, d)$ qui est un facteur d'échelle. Le paramètre ν est un paramètre qui sert en général à régler la régularité du processus. La fonction de corrélation $C_{\boldsymbol{\theta}}$ se décompose,

$$C_{\boldsymbol{\theta}}(\mathbf{x} - \mathbf{x}') = \prod_{j=1}^{d} c((\theta_j, \nu), |x_j - x'_j|)\,.$$

Le modèle est dit isotrope si $\theta_1 = \ldots = \theta_d$. Ce choix permet de réduire le nombre de paramètres à estimer si peu d'observations sont disponibles. Sinon un modèle anisotrope est privilégié.

Le choix d'un type de fonction de corrélation entraîne un a priori sur la régularité de la fonction f. Les fonctions de régression généralement utilisées sont en général des polynômes

donc infiniment différentiables. La régularité du processus Y dépend alors de la fonction de corrélation $C_{\boldsymbol{\theta}}$ de Z. Des notions naturelles sont la continuité et la différentiabilité en moyenne quadratique (Adler, 1981).

Définition 2.3. *Soit Z un processus aléatoire stationnaire sur E admettant des moments d'ordre deux. Z est dit continu en moyenne quadratique si*

$$\lim_{\mathbf{x}\to 0} \mathbb{E}\left((Z(\mathbf{x}) - Z(0))^2\right) = 0\,.$$

Écrire la continuité en 0 implique la continuité sur E entier par stationnarité du processus. On peut remarquer que

$$\mathbb{E}\left((Z(\mathbf{x}) - Z(0))^2\right) = 2(C_{\boldsymbol{\theta}}(0) - C_{\boldsymbol{\theta}}(\mathbf{x} - 0))\,,$$

d'où la proposition suivante.

Proposition 2.5. *Le processus aléatoire stationnaire Z est continu en moyenne quadratique si sa fonction de covariance $C_{\boldsymbol{\theta}}$ est continue en* 0.

La différentiabilité en moyenne quadratique se définit à l'aide de la proposition suivante.

Proposition 2.6. *Si Z est un processus stationnaire tel que les dérivées $\partial^2 K_{\boldsymbol{\theta}}(\mathbf{x}, \mathbf{x}')/\partial x_k \partial x'_l = \partial^2 C_{\boldsymbol{\theta}}(\mathbf{x} - \mathbf{x}')/\partial x_k \partial x'_l$ existent et sont finies au point $(0, 0)$, alors la limite*

$$\partial_i Z(\mathbf{x}) = \lim_{\mathbf{h}\to 0} \frac{Z(\mathbf{x} + \mathbf{h}) - Z(\mathbf{x})}{\mathbf{h}}\,,$$

existe, et $\partial_i Z(\mathbf{x})$ est appelée la dérivée en moyenne quadratique de $Z(\mathbf{x})$. Le processus Z a alors une dérivée partielle en moyenne quadratique. La fonction de covariance de $\partial_i Z$ est alors donnée par $\partial^2 K_{\boldsymbol{\theta}}(\mathbf{x}, \mathbf{x}')/\partial x_k \partial x'_l$.

Les différentielles d'ordre supérieure peuvent ensuite être obtenues de manière itérative. On peut donner des propriétés sur la régularité des trajectoires d'un processus à partir de la régularité de sa fonction de covariance dans le cadre des processus gaussiens. Adler (1981) montre un théorème qui lie la continuité des trajectoires d'un processus gaussien à la vitesse de convergence de $C_{\boldsymbol{\theta}}(\mathbf{x})$ pour $\mathbf{x} \to 0$.

Théorème 2.3. *Si Z est un processus stationnaire de fonction de corrélation C qui vérifie*

$$1 - C_{\boldsymbol{\theta}}(\mathbf{x}) \leq \frac{a}{|\log(\|\mathbf{x}\|_2)|^{1+\epsilon}},\ \forall \|\mathbf{x}\|_2 < \delta\,, \tag{2.30}$$

où $a > 0$, $\epsilon > 0$ et $\delta < 1$, alors Z a des trajectoires presque surement continues.

La proposition 2.6 donne la fonction de covariance des dérivées partielles du processus en moyenne quadratique. Il suffit que la fonction de covariance $\partial^2 K_{\boldsymbol{\theta}}(\mathbf{x}, \mathbf{x}')/\partial x_k \partial x'_l$ de la dérivée partielle $\partial_i Z$ vérifie la condition (2.30) du théorème précédent.

Nous présentons quelques exemples de fonctions de noyaux de covariance qui sont généralement utilisées et sont souvent incorporées aux programmes. Nous donnons juste la fonction c univariée. Nous notons $\mathbf{x} = (x_1, \ldots, x_d),\ \mathbf{x}' = (x'_1, \ldots, x'_d) \in E$.

Noyaux de type exponentiel

$$c((\theta_j, \nu, x_j - x'_j) = \exp(-\theta_j |x_j - x'_j|^\nu)\,, \tag{2.31}$$

pour $0 < \nu \leq 2$. Pour $\nu = 1$, le noyau est dit exponentiel. Pour $\nu = 2$, le noyau est dit gaussien. Sinon on parle de noyau exponentiel généralisé. Il n'est différentiable en moyenne quadratique que dans le cas gaussien. Il est même infiniment différentiable, ce qui donne lieu à un processus très lisse. Pour toute valeur $0 < \nu \leq 2$, le processus est continu en moyenne quadratique et les trajectoires sont presque surement continues.

Noyaux cubiques

$$c((\theta_j, x_j - x'_j) = \begin{cases} 1 - 6\left(\frac{x_j - x'j}{\theta_j}\right)^2 + 6\left(\frac{|x_j - x'_j|}{\theta_j}\right)^3, & |x_j - x'_j| \leq \theta_j/2 \\ 2\left(1 - \frac{|x_j - x'_j|}{\theta_j}\right)^3, & \theta_j/2 < |x_j - x'_j| \leq \theta_j \\ 0, & \theta_j < |x_j - x'_j| \end{cases},$$

pour $\theta > 0$. Les trajectoires des processus ayant ce noyau de covariance sont continues et différentiables une fois. Ce type de noyau conduit aux interpolateurs par des fonctions splines cubiques. D'autres fonctions de corrélation de ce type sont données par Mitchell *et al.* (1990).

Noyaux de Matérn

$$c((\theta_j, \nu, x_j - x'_j) = \frac{(\theta_j |x_j - x'_j|)^\nu}{\Gamma(\nu) 2^{\nu-1}} J_\nu(\theta_j |x_j - x'_j|) \tag{2.32}$$

où on doit avoir $\theta_j \in (0, \infty)$ et $\nu \in (-1, \infty)$. J_ν est une fonction de Bessel modifiée d'ordre ν. Le processus associé sera m fois différentiable en moyenne quadratique si et seulement si $\nu > m$ et la régularité presque sure des trajectoires est de l'ordre de $(\lceil \nu \rceil - 1)$ ($\lceil . \rceil$ désigne la partie entière supérieure). Ainsi la régularité du processus est gouvernée par le paramètre ν et θ_j contrôle l'échelle de corrélation.

L'avantage de la modélisation par un processsus gaussien est de donner lieu à une estimation des paramètres liés au noyau par maximum de vraisemblance qui se révèle souvent efficace. Parmi les noyaux testés, Santner *et al.* (2003) conseillent d'utiliser les noyaux exponentiels généralisés ou de Matérn. Toutefois, ces derniers sont plus lourds à calculer et ne sont pas toujours sous forme explicite.

2.2.3 Interpolation à noyaux

Dans cette partie, le paramètre $\boldsymbol{\theta}$ ne sera plus mentionné en indice. Nous commençons par définir l'espace fonctionnel dans lequel nous travaillons.

Définition 2.4. *Soit $\mathcal{H}$ un espace de Hilbert fonctionnel sur l'ensemble E de produit scalaire $(.,.)_\mathcal{H}$. Le noyau $K : E \times E \to \mathbb{R}$ est appelé noyau reproduisant si*

1. *pour tout $\mathbf{x} \in E$, les fonctions $K_\mathbf{x} : \mathbf{x}' \mapsto K(\mathbf{x}, \mathbf{x}')$ appartiennent à $\mathcal{H}$,*

2. pour tous $\mathbf{x} \in E$ *et* $f \in \mathcal{H}$*, la propriété de reproduction est vraie :*

$$(f, K_{\mathbf{x}})_{\mathcal{H}} = f(\mathbf{x})\,. \tag{2.33}$$

Si un noyau reproduisant K *existe,* $\mathcal{H}$ *est appelé un espace de Hilbert à noyau reproduisant (RKHS, Reproducing Kernel Hilbert Space).*

Le théorème d'Aronszajn (1950) donne les propriétés du noyau reproduisant et permet d'associer à un noyau K défini positif un espace hilbertien.

Théorème 2.4 (Aronszajn)**.**

- *Si un noyau reproduisant existe, il est unique.*
- *Un noyau reproduisant existe si et seulement si, pour tout* $\mathbf{x} \in E$*, les applications*

$$\begin{aligned} \mathcal{H} &\to \mathbb{R} \\ f &\mapsto f(\mathbf{x})\,, \end{aligned}$$

  sont continues.
- *Le noyau reproduisant est défini positif.*
- *Réciproquement, si* K *est un noyau défini positif, il existe un espace noté* $\mathcal{H}_K$ *qui est un RKHS de noyau reproduisant* K*.*

Cet espace correspond au complété de l'espace engendré par les fonctions partielles $\mathbf{x}' \mapsto K_{\mathbf{x}}(\mathbf{x}') = K(\mathbf{x}, \mathbf{x}')$ pour $\mathbf{x} \in E$, pour lequel on a défini le produit scalaire :

$$(K_{\mathbf{x}}, K_{\mathbf{x}'}) = K(\mathbf{x}, \mathbf{x}')\,.$$

Schaback (2007) nomme cet espace l'espace natif (Native space). Il sera noté $\mathcal{H}_K$ dans la suite. L'application

$$\begin{aligned} \Psi :\ & E \to \mathcal{H}_K \\ & \mathbf{x} \mapsto K_{\mathbf{x}}\,, \end{aligned} \tag{2.34}$$

est appelée application de modélisation ("feature map") puisqu'elle permet d'associer à un élément de E un élément de $\mathcal{H}_K$ qui est l'espace de modélisation ("feature space"). Ces dénominations sont utilisées principalement en théorie de l'apprentissage. Avec ces notations pour $\mathbf{x}, \mathbf{x}' \in E$, le produit scalaire des images respectives de $\mathbf{x}$ et $\mathbf{x}'$ dans l'espace $\mathcal{H}_K$ est donné par $K(\mathbf{x}, \mathbf{x}') = (\Psi(\mathbf{x}), \Psi(\mathbf{x}'))_{\mathcal{H}_K}$.

Les noyaux utilisés sont souvent invariants par translation. Comme dans la partie précédente,

$$\forall \mathbf{x}, \mathbf{x}',\ K(\mathbf{x}, \mathbf{x}') = C(\mathbf{x} - \mathbf{x}')\,. \tag{2.35}$$

Les fonctions radiales de base $R(\|\mathbf{x} - \mathbf{x}'\|)$ où $R : \mathbb{R} \to \mathbb{R}$, vérifient bien évidemment cette propriété. La norme $\|.\|$ sur E utilisée n'est pas forcément la norme euclidienne. On peut la modifier afin de tenir compte de l'anisotropie comme dans la tensorisation (2.2.2). Wendland (2005); Schaback (2007) présentent différentes méthodes pour construire les noyaux. Les noyaux de Mercer peuvent être utilisés.

Définition 2.5. *Le noyau* $K : E \times E \to \mathbb{R}$ *est un noyau de Mercer s'il est continu, défini sur un espace* E *compact et si*

$$\int_{E \times E} K(\mathbf{x}, \mathbf{x}') f(\mathbf{x}) f(\mathbf{x}') d\mathbf{x} d\mathbf{x}' \geq 0\,,$$

pour toute fonction $f : E \to \mathbb{R}$ *continue.*

Les noyaux de Mercer sont définis positifs (Schölkopf et Smola, 2001). Le théorème de Mercer (1909) est une première étape pour donner une expression explicite de l'espace et de l'application de modélisation associés au noyau de Mercer K. Nous notons $L_2(E) = \{f : E \to \mathbb{R} : \int_E |f(\mathbf{x})|^2 d\mathbf{x} < \infty\}$ et $l^2 = \{(a_j)_{j\in\mathbb{N}^*} \in \mathbb{R}^{\mathbb{N}^*} : \sum_{j\in\mathbb{N}^*} |a_j|^2 < \infty\}$.

Théorème 2.5 (Mercer)**.** *Soit l'opérateur linéaire de $L_2(E)$ défini par*

$$\forall f \in L_2(E),\ (L_K f)(.) = \int_E K(\mathbf{x}, .) f(\mathbf{x}) d\mathbf{x}\,.$$

Si $\lambda_1, \lambda_2 \ldots$ sont les valeurs propres de L_K données dans l'ordre décroissant et $\phi_1, \phi_2, \ldots$ sont les fonctions propres correspondantes. Alors pour presque tous $\mathbf{x}, \mathbf{x}' \in E$,

$$K(\mathbf{x}, \mathbf{x}') = \sum_{j\in\mathbb{N}^*} \lambda_j \phi_j(\mathbf{x}) \phi_j(\mathbf{x}') = (\Phi(\mathbf{x}), \Phi(\mathbf{x}'))_{\ell_2}\,,$$

avec $\Phi : E \to \ell^2$ défini par $\Phi(\mathbf{x}) = \left(\sqrt{\lambda_j}\phi_j(\mathbf{x})\right)_{j\in\mathbb{N}^*}$.

Il est possible de donner une expression explicite de l'espace et de l'application de modélisation. On a alors pour K un noyau de Mercer, avec les notations du théorème précédent,

$$\mathcal{H}_K = \left\{ g \in L_2(E) : g = \sum_j a_j \phi_j,\ \text{avec} \sum_{j:\lambda_j>0} a_i^2/\lambda_j < \infty \right\},$$

avec le produit scalaire, pour $g = \sum_j a_j \phi_j,\ h = \sum_j b_j \phi_j$,

$$(g, h)_{\mathcal{H}_K} = \sum_{j:\lambda_j>0} \frac{a_j b_j}{\lambda_j}\,.$$

L'application de modélisation $\Psi : E \to \mathcal{H}_K$ est définie ainsi pour $\mathbf{x} \in E$,

$$\Psi(\mathbf{x})(.) = \sum_{j\in\mathbb{N}^*} \lambda_j \phi_j(\mathbf{x}) \phi_j(.)\,.$$

On a alors bien la relation

$$\forall \mathbf{x}, \mathbf{x}' \in E,\ K(\mathbf{x}, \mathbf{x}') = (\Psi(\mathbf{x}), \Psi(\mathbf{x}'))_{\mathcal{H}_K}\,.$$

Le théorème de Mercer ne s'applique plus si le noyau est considéré comme défini sur $\mathbb{R}^d$ entier. Une technique consiste à utiliser les transformées de Fourier si le noyau est invariant par translation. Pour la fonction C associée au noyau K par la relation (2.35), on note $\mathcal{F}C$ sa transformée de Fourier si elle existe. Le théorème suivant permet alors d'expliciter le RKHS associé.

Théorème 2.6. *Soit K un noyau défini positif sur $\mathbb{R}^d \times \mathbb{R}^d$ tel que la fonction C associée appartienne à $L_1(\mathbb{R}^d)$ ainsi que sa transformée de Fourier $\mathcal{F}C$. Le sous-espace $\mathcal{H}_K$ de $L_2(\mathbb{R}^d)$, composé des fonctions g continues et dans $L_1(\mathbb{R}^d)$ qui vérifient :*

$$\|g\|_{\mathcal{H}_K} = \frac{1}{(2\pi)^d} \int_{\mathbb{R}^d} \frac{|\mathcal{F}g(\omega)|^2}{\mathcal{F}C(\omega)} d\omega < \infty\,.$$

et équipé du produit scalaire :

$$(g, h)_{\mathcal{H}_K} = \frac{1}{(2\pi)^d} \int_{\mathbb{R}^d} \frac{\mathcal{F}g(\omega)\mathcal{F}h(\omega)^*}{\mathcal{F}C(\omega)} d\omega\,,$$

où a^* est le complexe conjugué de a, est un RKHS de noyau reproduisant K.

Par exemple, le noyau gaussien défini par

$$K_\theta(\mathbf{x}, \mathbf{x}') = C_\theta(\mathbf{x} - \mathbf{x}') = \exp(-\theta\|\mathbf{x} - \mathbf{x}'\|^2)\,,$$

pour $\theta > 0$ est associé à l'espace de fonctions

$$\mathcal{H}_{K_\theta} = \left\{ g \in L_1(\mathbb{R}^d) : \int |\mathcal{F}g(\omega)|^2 \exp\left(\frac{\|\omega\|^2}{4\theta}\right) < \infty \right\}.$$

Les fonctions dans ce RKHS sont infiniment différentiables avec toutes les dérivées dans $L_2(\mathbb{R}^d)$. On a les inclusions suivantes pour $0 < \theta < \tau$, $\mathcal{H}_{K_\theta} \subset \mathcal{H}_{K_\tau}$ (Vert et Vert, 2006).

Nous supposons dans cette partie que $f \in \mathcal{H}_K$ et comme dans la partie 2.2.1 que f a été évaluée aux points du plan d'expérience $D = \{\mathbf{x}_1, \ldots, \mathbf{x}_n\} \subset E$. Soit $S_D(f)$ la projection orthogonale de f sur le sous espace de $\mathcal{H}_K$, $\mathcal{H}_K(D) = \text{Vect}\{K_{\mathbf{x}_1}, \ldots, K_{\mathbf{x}_n}\}$. La propriété suivante indique que cette projection est l'interpolateur de plus petite norme dans le RKHS $\mathcal{H}_K$ et en donne une écriture lagrangienne.

Proposition 2.7.

1. *$S_D(f)$ est l'interpolateur de f aux points de D, de norme minimale. Ceci signifie que $S_D(f)$ est solution du problème :*

$$\begin{cases} \min_{g \in \mathcal{H}_K} \|g\|_{\mathcal{H}_K} \\ g(x_i) = f(\mathbf{x}_i),\ i = 1, \ldots, n \end{cases}.$$

2. *L'interpolateur $S_D(f)$ peut s'écrire, pour $\mathbf{x}_0 \in E$,*

$$S_D(f)(\mathbf{x}_0) = \sum_{i=1}^{n} f(\mathbf{x}_i) u_i(\mathbf{x}_0)\,,$$

où les fonctions $u_i : E \to \mathbb{R}$, pour $i = 1, \ldots, n$, appartiennent à $\mathcal{H}_K(D)$. En gardant les mêmes notations que celles introduites dans la proposition 2.2, et en posant $U(\mathbf{x}) = (u_1(\mathbf{x}), \ldots, u_n(\mathbf{x}))$, ce vecteur vérifie, pour tout $\mathbf{x} \in E$:

$$\Sigma_{\mathbf{x}_0 D} = \Sigma_{DD} U(\mathbf{x})\,.$$

Il est possible de contrôler l'erreur ponctuelle commise par l'interpolateur en tout $\mathbf{x}_0 \in E$, en utilisant la propriété de reproduction et en appliquant l'inégalité de Cauchy Schwarz,

$$\begin{aligned} |S_D(f)(\mathbf{x}_0) - f(\mathbf{x}_0)| &= |(f, K_{\mathbf{x}_0} - \sum_{i=1}^{n} u_i(\mathbf{x}_0) K_{\mathbf{x}_i})_{\mathcal{H}_K}| \\ &\leq \|f\|_{\mathcal{H}_K} \|K_{\mathbf{x}_0} - \sum_{i=1}^{n} u_i(\mathbf{x}_0) K_{\mathbf{x}_i}\|_{\mathcal{H}_K} \end{aligned} \tag{2.36}$$

On note

$$P_D(\mathbf{x}_0) = \|K_{\mathbf{x}_0} - \sum_{i=1}^{n} u_i(\mathbf{x}_0) K_{\mathbf{x}_i}\|_{\mathcal{H}_K}\,. \tag{2.37}$$

Schaback (1995b) nomme P_D fonction puissance et donne des majorants qui sont fonction d'un critère d'espacement des points dans le plan d'expérience D dans le cas de noyaux K usuels. Ce résultat sera utilisé pour justifier un choix de plan d'expérience dans la partie 4.

Nous pouvons constater que l'interpolation à noyaux conduit au même métamodèle que la modélisation par un processus gaussien. En effet, l'interpolateur $S_D(f - H(\mathbf{x}_0)\boldsymbol{\beta})$ correspond à la partie noyau de la moyenne a posteriori $\mu_{\mathbf{x}_0|D}$ (2.10) si l'on suppose que $f - H(\mathbf{x}_0)\boldsymbol{\beta} \in \mathcal{H}_K$ et que Σ_{DD} inversible. De plus, la fonction P_D (2.37) est égale à la variance a posteriori $\sigma^2_{\mathbf{x}_0|D}$ (2.10). Des fonctions de régression comme en krigeage peuvent être introduites mais la présentation est plus technique et est incluse dans l'interpolation avec des noyaux conditionnellement définis positifs (voir la section 3). Dans ce cas, on a toujours la même expression pour le métamodèle et la fonction puissance est égale à la variance du BLUP (2.15).

Toutefois, Driscoll (1973) montre qu'un processus gaussien dont le noyau de covariance est strictement défini positif et continu a presque toutes ses réalisations qui n'appartiennent pas au RKHS $\mathcal{H}_K$. Bien que la méthode d'interpolation à noyaux et la modélisation par un processus gaussien conduisent au même métamodèle, les hypothèses posées sur f dans chacun des cas s'excluent. Dans une modélisation par un processus gaussien de noyau de covariance K, on suppose que f en est une réalisation donc f n'appartient presque surement pas à l'espace $\mathcal{H}_K$. Cependant, Driscoll (1973) propose un théorème qui donne une condition nécessaire et suffisante sur un autre noyau S pour que les trajectoires d'un processus gaussien de noyau K appartiennent presque surement au RKHS $\mathcal{H}_S$.

2.2.4 Régularisation

Bien que le modèle des expériences simulées que nous considérons (2.1) ne souffre d'aucun bruit de mesure, l'interpolation exacte, si l'on a de nombreuses données (de l'ordre de $n = 1000$), peut rencontrer des problèmes numériques. Schaback (1995b) a formulé une sorte de principe d'incertitude qui indique qu'il n'est pas possible d'avoir à la fois une erreur d'interpolation faible et une bonne stabilité de l'interpolateur par rapport aux données $\mathbf{y}_D$. Ainsi, il peut être intéressant de chercher $\hat{f}$ comme la solution du problème régularisé suivant

$$\min_{g \in \mathcal{H}_K} \sum_{i=1}^{n} (y_i - g(\mathbf{x}_i))^2 + \lambda \|g\|^2_{\mathcal{H}_K}\,, \tag{2.38}$$

où λ est un réel strictement positif.

La solution est explicite et appartient à $\mathcal{H}_K(D)$ par application du théorème du représentant (Kimeldorf et Wahba, 1971).

Proposition 2.8. *$\hat{f}$ s'écrit*

$$\hat{f}(x) = \sum_{i=1}^{n} a_i K(\mathbf{x}_i, \mathbf{x})\,, \tag{2.39}$$

où $\mathbf{a} = (a_1, \ldots, a_n)$ est l'unique solution du système linéaire

$$(\lambda I_n + \Sigma_{DD})\mathbf{a} = y_D, \tag{2.40}$$

où I_n est la matrice identité en dimension n.

Nous pouvons remarquer que nous n'avons aucun problème d'inversion dans (2.40) même si le noyau K n'est pas strictement défini positif.

Remarque 2.6.
- *$\hat{f}$ ainsi obtenu approche f mais ne l'interpole pas aux points $\mathbf{x}_i$ du plan d'expérience.*
- *$\hat{f}$ s'écrit aussi sous la forme langrangienne*

$$\forall \mathbf{x}_0 \in E,\ \hat{f}(\mathbf{x}_0) = \sum_{i=1}^{n} v_i(\mathbf{x}_0) f(\mathbf{x}_i) = V(\mathbf{x}_0)^T y_d\,, \tag{2.41}$$

avec $V(\mathbf{x}_0) = (\Sigma_{DD} + \lambda I_n)^{-1}\Sigma_{\mathbf{x}_0 D}$.

Ce type de régularisation est aussi présent en krigeage. On parle généralement d'effet pépite. Cela revient à considérer que les observations du modèle (2.9) sont perturbées par un bruit blanc additif :

$$\forall \mathbf{x} \in E,\ Y(\mathbf{x}) = \sum_{i=1}^{p} \beta_i h_i(\mathbf{x}) + Z(\mathbf{x}) = H(\mathbf{x})^T \boldsymbol{\beta} + Z(\mathbf{x}) + \epsilon(\mathbf{x})\,, \tag{2.42}$$

où $\epsilon \sim \mathcal{N}(0, \tau^2)$ est indépendant de Z et de $\epsilon(\mathbf{x}')$ pour tout $\mathbf{x}' \in E$. Même pour deux observations successives au point $\mathbf{x}$, les bruits blancs sont supposés indépendants. Le BLUP associé est alors pour $\mathbf{x}_0 \in E$,

$$\hat{Y}(\mathbf{x}_0) = H(\mathbf{x}_0)^T \tilde{\boldsymbol{\beta}} + \Sigma_{\mathbf{x}_0 D}^T \left(\Sigma_{DD} + \tau^2 I_n\right)^{-1} \left(\mathbf{Y}_D - H_D \tilde{\boldsymbol{\beta}}\right),$$

avec

$$\tilde{\boldsymbol{\beta}} = (H_D^T \left(\Sigma_{DD} + \tau^2 I_n\right)^{-1} H_D)^{-1} H_D^T \left(\Sigma_{DD} + \tau^2 I_n\right)^{-1} Y_D\,.$$

L'expression de la vraisemblance en est alors modifiée. La matrice $\left(\Sigma_{DD} + \tau^2 I_n\right)^{-1}$ remplace Σ_{DD}^{-1} ce qui donne des estimations par maximum de vraisemblance plus stables car la matrice est mieux conditionnée. Cette régularisation est présente dans des cas non bruités pour compenser les arrondis machine et pour faciliter l'inversion des matrices de Gram dans l'algorithme d'estimation des paramètres $(\sigma^2, \boldsymbol{\theta})$. Par exemple, la boîte à outils DACE (Lophaven *et al.*, 2002a) effectue une régularisation systématique de l'ordre de $n \times eps$ où $eps = 2^{-52}$ est la précision machine de Matlab.
Cette démarche nous conduit au même métamodèle que la résolution du problème (2.38) si on fixe $\lambda = \tau^2/n$.

2.2.5 Conclusion

Dans cette partie, nous avons présenté les résultats dans un contexte assez simple. Nous n'avons considéré que des modélisations par processus gaussiens alors que certains résultats peuvent être étendus à d'autres familles de processus (Santner *et al.*, 2003). Ces modèles sont envisagés comme des a priori bayésiens sur la fonction f inconnue qui est considérée comme une réalisation du processus. Pillai *et al.* (2007) distinguent cette manière de fixer directement une loi a priori pour f de la méthode qui introduit une loi a priori sur les mesures aléatoires de E ce qui induit une loi a priori sur un espace de fonctions définies par un modèle à noyau.

Les résultats concernant l'interpolation par noyaux n'ont été énoncés que dans le cadre de noyaux définis positifs. Les travaux de Schaback (2007); Wendland (2005) sont plus généraux car ils traitent de noyaux conditionnellement définis positifs. Le krigeage intrinsèque (Matheron, 1973) qui consiste à faire des hypothèses de stationnarité plus générales que la stationnarité au second ordre permet l'utilisation de noyaux conditionnellement définis positifs. Dans la partie 3, nous proposons une généralisation de la définition de noyau conditionnellement défini positif couramment employée et nous donnons des résultats concernant l'interpolation et la régression régularisée à partir de cette définition. Les parties suivantes utiliseront la modélisation par processus gaussien. Dans les applications, la vision statistique (krigeage) sera privilégiée.

2.3 Plans d'expérience numérique

Dans les parties précédentes, le plan d'expérience à partir duquel le métamodèle est construit était supposé donné. À présent, la question du choix de ce plan se pose. Les appels au code de calcul de f étant coûteux, il faut obtenir le maximum d'informations sur f à l'aide de n évaluations de cette fonction. Le nombre n est vu comme un budget alloué à la construction du métamodèle. On garde la notation d'un plan d'expérience $D = \{\mathbf{x}_1, \ldots, \mathbf{x}_n\} \subset E^n$. Nous rappelons que f est déterministe, les réplications dans les plans d'expérience numérique n'apportent aucune information supplémentaire. Il semble sensé de choisir un plan d'expérience qui comporte des points bien répartis dans le domaine E. Un tel plan sera dit exploratoire. Il faut alors définir un critère mathématique au sens duquel le plan sera exploratoire. Plusieurs types de critère sont envisageables, il est possible de s'attacher aux propriétés d'échantillonnage des points de D, aux relations entre les points ou/et aux performances de D vis-à-vis d'un certain critère calculé sous les hypothèses de la modélisation par processus gaussien (2.9). La loi uniforme sert souvent de base aux échantillonnages proposés. Si la loi des entrées n'est pas uniforme, il est possible pour certaines méthodes de construire le plan d'expérience en fonction de cette loi. Cependant, suivant les objectifs visés, cela n'est pas forcément préférable. Il peut être important/utile d'avoir échantillonné dans des zones du domaine E de faible probabilité. Certains plans exploratoires décrits dans la suite ont de bonnes propriétés de projection qui permettent d'éviter la redondance de l'information en cas de projection des points sur un sous espace de E de dimension inférieure lorsque des variables d'entrée s'avèrent inutiles. Le choix d'une grille pourrait passer pour un choix naturel de plan d'expérience à $n = q^d$ points où $q \in \mathbb{N}$. Cependant, si l'on conclut après étude du métamodèle qu'une des variables d'entrée n'a aucune influence sur la sortie y, on projettera sur l'espace de dimension $d-1$. Les projections coïncideront et le plan d'expérience ne comportera plus que q^{d-1} points. Par exemple, la figure 2.1 représente une grille à 16 points pour un domaine $E = [0,1]^2$. Si la variable d'entrée qui correspond aux ordonnées peut être éliminée, le plan se résumera aux seuls 4 points sur l'axe des abscisses.

Dans cette partie, le domaine E sera souvent supposé hypercubique, c'est-à-dire que pour chaque variable d'entrée nous aurons une borne inférieure et une borne supérieure qui définissent le domaine.

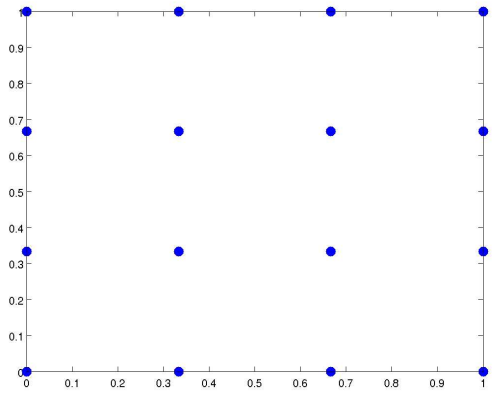


FIG. 2.1 – Grille à 16 points dans $E = [0,1]^2$

2.3.1 Critères d'échantillonnage

Pour juger de la qualité d'un plan d'expérience, on peut s'intéresser à l'estimateur de l'intégrale

$$\mathbb{E}(f(\mathbf{X})) = \frac{1}{\mathrm{vol}(E)} \int_E f(\mathbf{x}) d\mathbf{x}\,, \tag{2.43}$$

qui est l'espérance de la variable aléatoire $f(\mathbf{X})$ pour $\mathbf{X}$ suivant une loi uniforme sur E. À partir des évaluations de f aux points de D, on peut proposer l'estimation suivante de $\mathbb{E}(f(\mathbf{X}))$,

$$\bar{y}(D) = \frac{1}{n} \sum_{i=1}^{n} f(\mathbf{x}_i)\,.$$

On est amené à chercher le plan d'expérience D pour que cet estimateur soit optimal.

D'un point de vue statistique, le plan D est optimal si la variance de l'estimateur $\bar{Y}(D)$ est minimale. On est amené ici à considérer $D = \{\mathbf{X}_1, \ldots, \mathbf{X}_n\}$ comme un échantillon aléatoire. L'estimateur $\bar{Y}(D)$ est bien une variable aléatoire. On se compare à un plan d'expérience D qui est un n échantillon d'une variable aléatoire de loi uniforme dans E. Dans ce cas, la variance de $\bar{Y}(D)$ est égale à $\mathrm{Var}(f(\mathbf{X}))/n$. McKay *et al.* (1979) ont proposé l'échantillonnage en hypercube latin. Si on suppose que $E = [0,1]^d$, on constuit le plan $D = \{\mathbf{X}_1, \ldots, \mathbf{X}_n\}$ en hypercube latin en prenant pour $i = 1, \ldots, n$, $j = 1, \ldots, d$,

$$X_{ij} = \frac{\pi_j(i) - U_j^i}{n}\,, \tag{2.44}$$

où pour $i = 1, \ldots, n$, $\mathbf{X}_i = (X_{i1}, \ldots, X_{id})^T$, les π_j sont des permutations aléatoires indépendantes des entiers de $\{1, \ldots, n\}$, et les U_j^i sont i.i.d. de loi uniforme $\mathcal{U}[0,1]$ indépendantes des π_j. S'il a été construit ainsi le plan d'expérience sera appelé LHD (Latin Hypercube Design). Un LHD a des projections bien réparties sur chaque axe. La projection sur un axe d'un LHD à n points comporte un point et un seul dans l'intervalle $\left[\frac{k}{n}; \frac{k+1}{n}\right]$ pour $k = 0, \ldots, n-1$. Il est aussi possible de centrer les points de l'hypercube latin dans les intervalles en remplaçant

les U_j^i par la valeur 0.5. McKay *et al.* (1979) montrent que si D_1 est un n-échantillon de loi uniforme dans E, D_2 est un LHD, et si f est monotone suivant toutes ses variables d'entrée, alors

$$\mathrm{Var}(\bar{Y}(D_1)) \geq \mathrm{Var}(\bar{Y}(D_2))\,.$$

Sans hypothèses de monotonie sur f, Stein (1987) montre que la variance de $\bar{Y}(D_2)$ est asymptotiquement plus petite que la variance de $\bar{Y}(D_1)$ si le moment d'ordre 2 de $f(\mathbf{X})$ existe. La figure 2.2 fournit deux exemples de LHD et montre que l'appartenance à la classe des LHD

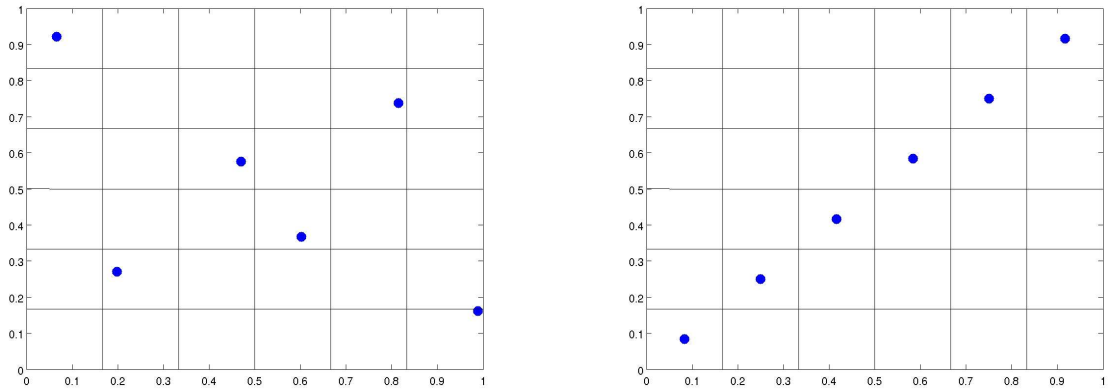


FIG. 2.2 – Deux exemples de LHD à 6 points dans $E = [0,1]^2$

n'est pas forcément une propriété suffisante pour obtenir un plan d'expérience exploratoire. Dans la suite, nous évoquerons des plans d'expérience définis comme réalisant l'optimum d'un critère dans la classe des LHD.

La construction de LHD a été proposée sous l'hypothèse que le vecteur de variables d'entrée $\mathbf{X}$ suivait une loi uniforme dans l'hypercube $[0,1]^d$ avec toutes ses composantes indépendantes. Un hypercube latin peut être construit pour d'autres lois sur les composantes de $\mathbf{X}$ à partir des quantiles des fonctions de répartition. Stein (1987) donne une méthode pour construire des hypercubes latins dans le cas où les composantes de $\mathbf{X}$ ne sont plus indépendantes.

Owen (1992); Tang (1993) ont proposé d'utiliser les tableaux orthogonaux aléatoires (randomized orthogonal arrays) qui sont une généralisation des LHD. Un tableau orthogonal A est une matrice $n \times d$ d'entiers $0 \leq A_{ij} \leq b-1$. Ce tableau a la force $t \leq d$, si dans toutes les sous matrices de A de taille $n \times t$, toutes les b^t lignes possibles apparaissent le même nombre de fois. Dans ce cas, nécessairement $n = \lambda b^t$. Le tableau A est associé au plan d'expérience $D = \{\mathbf{x}_1, \ldots, \mathbf{x}_n\}$ dans $E = [0,1]^d$, en écrivant pour tout $i = 1, \ldots, n$ et pour tout $j = 1, \ldots, d$,

$$x_{ij} = \frac{A_{ij} + 0.5}{b}\,.$$

On appellera indifféremment A ou D tableau orthogonal. Le plan d'expérience $D = \{\mathbf{X}_1, \ldots, \mathbf{X}_n\}$ est un tableau orthogonal aléatoire s'il est obtenu ainsi,

$$X_{ij} = \frac{\pi_j(A_{ij}) + U_j^i}{b}\,,$$

ou la version centrée,

$$X_{ij} = \frac{\pi_j(A_{ij}) + 0.5}{b},$$

où les π_j sont des permutations aléatoires indépendantes des entiers $\{0, \ldots, b-1\}$ et les U_j^i sont i.i.d. de loi uniforme $\mathcal{U}[0,1]$ indépendantes des π_j. Un LHD correspond à un tableau orthogonal aléatoire de force $t = 1$, avec $\lambda = 1$. Un LHD garantit des projections équi-réparties sur les sous espaces de dimension 1 c'est-à-dire les axes, tandis qu'un tableau orthogonal aléatoire de force t garantit des projections équi-réparties sur les sous espaces de dimension inférieure ou égale à t. Toutefois du fait de la relation $n = \lambda s^t$, les tableaux orthogonaux n'existent pas pour toutes les valeurs de n et ne sont utilisables que pour de petites valeurs de s et t. D'autres méthodes qui visent à réduire la variance de l'estimateur de la moyenne (2.43) sont décrites par Koehler et Owen (1996).

D'un point de vue déterministe, l'inégalité de Koksma-Hlawka (Niederreiter, 1992) donne,

$$|\mathbb{E}(f(\mathbf{X}) - \bar{y}(D))| \leq TV(f)Dis(D), \tag{2.45}$$

où $Dis(D)$ est la discrépance étoile ou discrépance en norme L_∞ de D et $TV(f)$ est la variation totale de f au sens de Hardy et Krause. La discrépance étoile est une mesure d'uniformité utilisée dans les méthodes de quasi Monte-Carlo et est la statistique du test d'adéquation de Kolmogorov-Smirnov pour une loi uniforme. Cette discrépance ne dépend que de D et ce dernier sera optimal s'il la minimise. On note F_D la fonction de répartition empirique associée à l'échantillon qu'est le plan d'expérience D : $\forall \mathbf{x} = (x_1, \ldots, x_d) \in E$,

$$F_D = \frac{1}{n} \sum_{i=1}^{n} \mathbb{I}_{\{x_{i1} \leq x_1, \ldots, x_{in} \leq x_n\}}.$$

Pour F la fonction de répartition d'une loi uniforme sur E, la discrépance étoile Dis est ainsi définie

$$Dis(D) = \|F_D - F\|_{L_\infty} = \sup_{\mathbf{x} \in E} |F_D(\mathbf{x}) - F(\mathbf{x})|.$$

Remarque 2.7. *Si $d = 1$, le plan de discrépance minimale à n points dans le domaine $[0,1]$ est l'hypercube latin centré :*

$$D = \left\{ \frac{1}{2n}, \frac{3}{2n}, \ldots, \frac{2n-1}{2n} \right\}.$$

Une conjecture en théorie des nombres donne pour tout ensemble de n points D_n,

$$Dis(D_n) \geq c(d) \frac{(\log n)^{d-1}}{n},$$

où c ne dépend que de la dimension d. Ainsi si on a une suite de plan d'expérience D_n qui ont une discrépance de l'ordre de $n^{-1}(\log n)^{d-1}$ pour $n \to \infty$, on considère que ces plans sont uniformément répartis pour de grandes valeurs de n. En comparaison, si D_n est généré par un tirage de Monte-Carlo, sa discrépance sera de l'ordre de $n^{-1/2}$. La méthode des bons points de réseau (good-lattice-point ; Fang et Wang, 1994) permet d'obtenir des suites de plans avec une discrépance d'ordre souhaitée $n^{-1}(\log n)^{d-1}$.

Santner *et al.* (2003) illustrent par un exemple (exemple 5.7) que le choix de la discrépance étoile comme unique critère pour attester de la qualité d'un plan n'est pas suffisant. Fang *et al.* (2000) imposent dans leur définition de plan uniforme que la matrice correspondant au plan soit de rang égal à la dimension d. La discrépance étoile étant d'évaluation coûteuse, ils introduisent la discréprance L_2, avec les mêmes notations que précédemment,

$$Dis_2(D) = \|F_D - F\|_{L_2} .$$

Elle a une expression analytique. Cependant, elle a le défaut d'avoir de mauvaises propriétés de projection et elle n'est pas invariante aux rotations du plan d'expérience. Ils proposent alors d'autres mesures de discrépance L_2, la discrépance L_2 symétrique, la discrépance L_2 centrée et la discrépance L_2 modifiée. Ces discrépances ont de meilleures propriétés et satisfont toutes une inégalité de type Koksma-Hlawka (2.45). Fang *et al.* (2000) donnent deux algorithmes pour obtenir des plans d'expérience minimisant une discrépance donnée parmi l'ensemble des tableaux orthogonaux centrés de force 1. Leurs essais numériques les mènent à conjecturer que le plan qui est optimal parmi les tableaux orthogonaux de force 1 au sens d'une des trois discrépances L_2 symétrique, L_2 centrée ou L_2 modifiée est en fait un tableau orthogonal de force 2 si un tel tableau existe pour n et d donnés. Ainsi les algorithmes de recherche de plans uniformes sont un moyen d'obtenir tableaux orthogonaux de force 2. Ils ont l'intuition que tout tableau orthogonal de force 2 est optimal au sens d'une certaine discrépance. D'autres formes de discrépances et des algorithmes pour obtenir des plans uniformes sont disponibles dans le livre de Fang *et al.* (2006). Des suites à discrépance faible comme les suites de Halton (Rafajlowicz et Schwabe, 2006) ou de Sobol (Bratley et Fox, 1988) peuvent aussi être utilisées pour former un plan d'expérience. Les suites de faibles discrépances présentent de bonnes propriétés de projection et sont robustes aux spécifications du modèle de régression comme cela est montré par Wiens (1991).

2.3.2 Critères de distances entre les points

Johnson *et al.* (1990) ont proposé des critères fondés sur la distance entre les points du plan d'expérience pour juger de sa qualité. La distance utilisée est presque toujours la distance euclidienne.

Définition 2.6. *Un plan d'expérience $D = \{\mathbf{x}_1, \ldots, \mathbf{x}_n\} \subset E^n$ est dit minimax s'il minimise*

$$h_D = \sup_{\mathbf{x} \in E} \min_{1 \leq i \leq n} \|\mathbf{x} - \mathbf{x}_i\|_2 . \tag{2.46}$$

Un plan d'expérience $D = \{\mathbf{x}_1, \ldots, \mathbf{x}_n\} \subset E^n$ est dit maximin s'il maximise

$$\delta_D = \min_{1 \leq i,j \leq n} \|\mathbf{x}_i - \mathbf{x}_j\|_2 . \tag{2.47}$$

Puisque E est un espace compact et les fonctions $D \mapsto h_D$ et $D \mapsto \delta_D$ sont continues, l'existence des plans minimax et maximin est garantie. Il n'y a cependant pas unicité de la solution. Parmi plusieurs plans maximin, sont privilégiés ceux de plus petit indice c'est-à-dire ceux qui ont un nombre minimal de paires réalisant la distance δ_D. Un plan minimax assure qu'en tout point du domaine E, on ne sera jamais trop loin d'un point du plan d'expérience D tandis qu'un plan maximin espace de manière optimale ses points afin d'éviter les réplications. La figure 2.3 illustre ces définitions. Un plan minimax aura tendance à placer ses points à l'intérieur de E et un plan maximin les placera sur les bords de E. Si $E = [0,1]^d$ et $n \leq 2^d$, les points des plans maximin seront situés sur les coins de E c'est-à-dire qu'ils appartiendront à $\{0,1\}^d$.

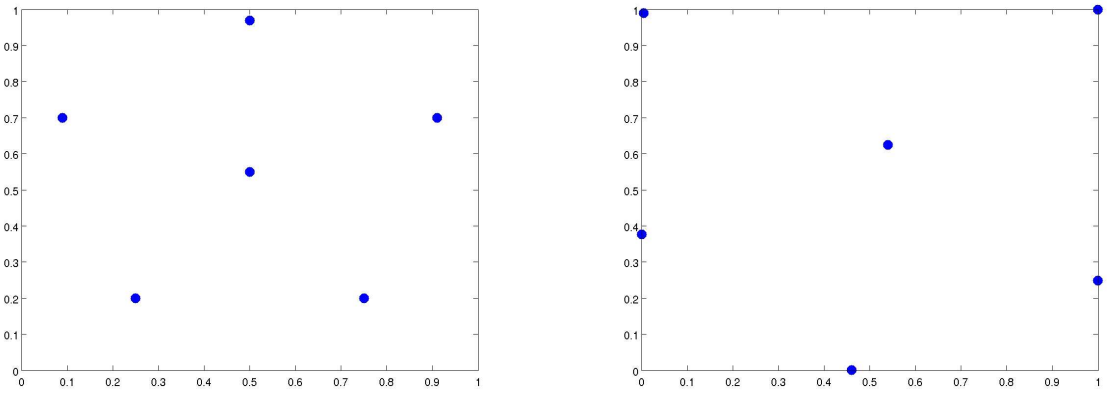


FIG. 2.3 – Plans d'expérience minimax et maximin à 6 points dans $E = [0,1]^2$

Exemples Si $E = [0,1]$,

- un plan minimax à n points est $D = \{\frac{1}{2n}, \frac{3}{2n}, \ldots, \frac{2n-1}{2n}\}$ et la distance minimax est $h_D = \frac{1}{2n}$;
- un plan maximin à n points est $D = \{0, \frac{1}{n-1}, \frac{2}{n-1}, \ldots, \frac{n-2}{n-1}, 1\}$, sa distance maximin est $\delta_D = \frac{1}{n-1}$ et son indice est $n-1$.

Dans le cadre de métamodèles construits comme des interpolateurs à noyaux, Schaback (1995b) propose des majorations de la fonction puissance P_D (2.37) qui intervient dans la borne de l'erreur d'interpolation en un point (2.36). Ces majorations sont données pour des noyaux couramment utilisés et sont de la forme, pour $\mathbf{x}_0 \in E$,

$$P_D(\mathbf{x}_0) \leq G_K(h_D)\,, \tag{2.48}$$

où G_K est une fonction croissante de h_D que l'on reconnaît comme étant la distance minimax. Ce paramètre h_D est aussi appelé distance de remplissage. La fonction G_K tend vers 0 pour h_D tendant vers 0, la vitesse dépend de la régularité du noyau (et donc de f puisqu'il est supposé dans le cadre de l'interpolation à noyaux que $f \in \mathcal{H}_K$). Par exemple, pour un noyau gaussien, $K(\mathbf{x}, \mathbf{x}') = \exp\left(-\theta\|\mathbf{x} - \mathbf{x}'\|_2^2\right)$, la fonction G_K correspondante est sous la forme $G_K(h_D) = e^{-c/h^2}$ où $c > 0$ est une constante qui dépend uniquement du paramètre θ. Marchi et Schaback (2008) traitent aussi de la stabilité des interpolations réalisées à partir d'un plan d'expérience D. Pour ce faire, ils s'intéressent à la valeur propre minimale de la matrice de Gram Σ_{DD} dont dépend son conditionnement. Une valeur propre minimale trop petite induira un grand conditionnement donc une instabilité des solutions. Pour certains noyaux, ils proposent la minoration suivante de la plus petite valeur propre de Σ_{DD} notée $\lambda_{\Sigma_{DD}}$:

$$\lambda_{\Sigma_{DD}} \geq L(\delta_D)\,,$$

où δ_D correspond au critère optimisé pour obtenir des plans maximin et est aussi appelée distance de séparation du plan d'expérience D. La fonction $L : \mathbb{R}_+ \to \mathbb{R}$ tend vers 0 en 0^+. Pour un noyau gaussien, $L(q) = q^{-d} \exp\left(-\frac{c}{q^2}\right)$. Ainsi un plan d'expérience maximin permet aussi une stabilité numérique de l'interpolateur.

Morris et Mitchell (1995) proposent de chercher un plan d'expérience maximin dans la classe des hypercubes latins dans le cas d'un domaine E hypercubique. On peut ainsi associer une propriété de dispersion des points dans le domaine grâce au critère maximin à de bonnes propriétés de projection grâce à l'échantillonnage en hypercube latin. Le plan d'expérience présenté par la figure 2.4 a été obtenu grâce à cet algorithme. Il est à comparer aux LHDs donnés en exemple par la figure 2.2. Joseph et Hung (2008) proposent un algorithme pour

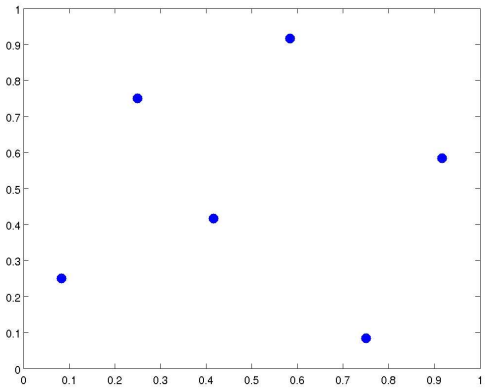


FIG. 2.4 – LHD maximin à 6 points dans $E = [0, 1]^2$

obtenir un plan d'expérience en hypercube latin qui optimise un critère multi-objectif. Il s'agit d'écarter les points à l'aide d'un critère maximin et de minimiser la corrélation entre les variables. Cette approche peut se justifier dans le cadre statistique de la modélisation par processus gaussien parce que la dispersion des points limite la variance de prédiction et que la minimisation de la corrélation des variables permet une bonne estimation des paramètres de régression (le vecteur $\boldsymbol{\beta}$).

2.3.3 Plans d'expérience optimaux

Il est possible de chercher, pour une modélisation fixée, un plan d'expérience optimal vis-à-vis de celle-ci. Cependant, il faut être assez vigilant au fait que les plans optimaux ne sont pas forcément robustes à une mauvaise spécification du modèle.

Dans le cadre du modèle linéaire, les réponses $\mathbf{y} = (y_1, \ldots, y_n)^T$ sont associées respectivement aux points du plan d'expérience $D = (\mathbf{x}_1, \ldots, \mathbf{x}_n)$ par la relation supposée :

$$y_i = \sum_{j=1}^{L} \beta_j B_j(\mathbf{x}_i) + \epsilon_i, \ i = 1, \ldots, n\,,$$

avec les fonctions B_j fixées et les ϵ_i tels que $\mathbb{E}(\epsilon_i) = 0$, $\mathrm{Var}(\epsilon_i) = \sigma^2$, pour $i = 1, \ldots, n$ et $\mathrm{Cov}(\epsilon_i, \epsilon_k) = 0$ pour $k \neq i$. Ce modèle peut s'écrire de façon matricielle en reprenant les notations introduites par (2.4) :

$$\mathbf{y} = \mathbf{B}^T\boldsymbol{\beta} + \boldsymbol{\epsilon}\,,$$

où $\boldsymbol{\epsilon} = (\epsilon_1, \ldots, \epsilon_n)^T$. On note $\mathbf{M} = \frac{1}{n}\mathbf{B}^T\mathbf{B}$ la matrice d'information qui dépend du plan d'expérience D choisi. On peut montrer que la matrice de variance-covariance du vecteur de

paramètres $\boldsymbol{\beta}$ estimé par la méthode des moindres carrés est

$$\mathrm{Var}(\hat{\boldsymbol{\beta}}) = \frac{\sigma^2}{n}\mathbf{M}^{-1}\,.$$

À partir de cette covariance, il est possible de définir plusieurs notions d'optimalité telles que :
- la D-optimalité qui consiste à chercher D qui maximise le déterminant de $\mathbf{M}$;
- la A-optimalité qui consiste à chercher D qui minimise la trace de $\mathbf{M}^{-1}$;
- la E-optimalité qui consiste à chercher D qui minimise la plus grande valeur propre de $\mathbf{M}^{-1}$.

La D-optimalité revient à minimiser le volume de la région de confiance pour le paramètre $\boldsymbol{\beta}$. La A-optimalité est équivalente à minimiser la somme des variances des estimateurs $\hat{\beta}_1, \ldots, \hat{\beta}_L$ tandis que la E-optimalité minimise la variance du contraste $\mathbf{a}^T\boldsymbol{\beta}$ (pour $\mathbf{a} \in \mathbb{R}^L$ tel que $\mathbf{a}^T\mathbf{a} = 1$) le moins bien estimé (voir Atkinson et Donev, 1992).
Bien que les expériences simulées n'entrent pas dans le cadre du modèle linéaire car elles ne ne souffrent d'aucun bruit de mesure, ces notions d'optimalité peuvent être utilisées pour les métamodèles polynomiaux ou de splines.

Dans le cadre de la modélisation par processus gaussien décrite en 2.2.1, Currin *et al.* (1991) proposent d'utiliser un plan d'expérience qui maximise l'entropie a priori ce qui implique minimiser l'entropie a posteriori.

Définition 2.7. *Un plan d'expérience D est dit d'entropie maximale s'il maximise*

$$\mathbb{E}\left(-\log p(\mathbf{Y}_D)\right)\,,$$

où $p(.)$ est la densité du vecteur aléatoire $\mathbf{Y}_D$ sous les hypothèses du modèle (2.9).

Si l'on se place sous les hypothèses du théorème 2.1 avec la loi a priori de type (i) c'est-à-dire telle que $\boldsymbol{\beta} \sim \mathcal{N}_p(\mathbf{b}_0, \tau^2 V_0)$, le critère de plan d'expérience d'entropie maximale est équivalent à chercher un plan D qui maximise le produit de déterminants :

$$|\Sigma_{DD}| \cdot |H_D^T \Sigma_{DD}^{-1} H_D + \tau^{-2} V_0^{-1}|\,.$$

Si la loi a priori sur $\boldsymbol{\beta}$ est diffuse ($\tau^2 \to \infty$), cela revient à maximiser :

$$|\Sigma_{DD}| \cdot |H_D^T \Sigma_{DD}^{-1} H_D|\,,$$

et si $\boldsymbol{\beta}$ est considéré fixé, un plan est d'entropie maximale s'il maximise $|\Sigma_{DD}|$. Currin *et al.* (1991) nomment ce critère D-optimalité pour cette modélisation. Il dépend à travers Σ_{DD} du choix du noyau de covariance K et du vecteur de paramètres $\boldsymbol{\theta}$ qui est en général estimé grâce aux données obtenues à partir du plan d'expérience. Koehler et Owen (1996) présentent différents plans d'entropie maximale en fonction du vecteur $\boldsymbol{\theta}$ choisi. Le critère maximin (2.47) a été proposé par Johnson *et al.* (1990) pour construire des plans d'expérience d'entropie maximale dans le cas asymptotique où la corrélation entre les sites devient très faible.

Une idée naturelle pour choisir un plan d'expérience pour une modélisation par processus gaussien (2.9) est décrite par Sacks *et al.* (1989b). Il s'agit de choisir le plan qui minimise l'erreur quadratique intégrée $IMSE$. Si l'on considère $\boldsymbol{\beta}$ inconnu dans la modélisation et que

l'on reprend les notations de la partie 2.2.1, l'erreur quadratique moyenne du BLUP en un point $\mathbf{x}_0 \in E$ s'écrit sous la forme :

$$MSE(\mathbf{x}_0) = \sigma^2 \left(1 - (H(\mathbf{x}_0)^T, \Sigma_{\mathbf{x}_0 D}^T) \begin{bmatrix} 0 & H_D^T \\ H_D & \Sigma_{DD} \end{bmatrix}^{-1} \begin{pmatrix} H(\mathbf{x}_0) \\ \Sigma_{\mathbf{x}_0 D} \end{pmatrix} \right) .$$

Nous cherchons un plan qui minimise l'erreur quadratique moyenne intégrée par rapport à une fonction de poids g. Son expression est :

$$\sigma^2 \left(1 - \text{trace} \left(\begin{bmatrix} 0 & H_D^T \\ H_D & \Sigma_{DD} \end{bmatrix}^{-1} \int \begin{pmatrix} h(\mathbf{x})h(\mathbf{x})^T & h(\mathbf{x})\Sigma_{\mathbf{x}D}^T \\ \Sigma_{\mathbf{x}D}h(\mathbf{x})^T & \Sigma_{\mathbf{x}D}\Sigma_{\mathbf{x}D}^T \end{pmatrix} g(\mathbf{x}) d\mathbf{x} \right) \right) .$$

Comme pour les plans d'entropie maximale, les plan optimaux au sens de $IMSE$ dépendent du noyau choisi pour modéliser la covariance et du vecteur de paramètres $\boldsymbol{\theta}$. Cependant, Sacks *et al.* (1989b) ont montré que le plan était assez robuste à un mauvais choix de $\boldsymbol{\theta}$. Zhu et Zhang (2006) proposent de chercher un plan d'expérience qui permet une bonne prédiction par le BLUP avec des paramètres $\boldsymbol{\theta}$ estimés. Pour ce faire, ils utilisent un critère qui consiste à modifier l'erreur quadratique moyenne intégrée pour qu'elle prenne en compte l'estimation des paramètres. Un autre critère d'optimalité aussi fondé sur l'erreur quadratique moyenne du prédicteur est de chercher un plan d'expérience qui minimise :

$$\max_{\mathbf{x} \in E} MSE(\mathbf{x}) .$$

Nous avons fait le lien dans la partie 2.2.3 entre la fonction puissance P_D associée au plan d'expérience D servant à majorer l'erreur ponctuelle dans le cadre de l'interpolation à noyaux et l'erreur quadratique moyenne de prédiction. Ces deux quantités sont égales. Ainsi on dispose d'une majoration de l'erreur quadratique MSE si le noyau utilisé pour modéliser la covariance correspond aux fonctions radiales de base pour lesquelles Schaback (1995b) donne des bornes de P_D. En reprenant les notations de (2.48), on a alors pour tout $\mathbf{x}_0 \in E$:

$$MSE(\mathbf{x}_0) \leq G_K(h_D) ,$$

où h_D correspond à la distance de remplissage définie par l'équation (2.46). Cette borne est indépendante du point $\mathbf{x}_0$, c'est par conséquent une borne de $\max_{\mathbf{x} \in E} MSE(\mathbf{x})$. Ainsi un plan minimax et a fortiori un plan maximin garantissent un contrôle uniforme sur E de l'erreur quadratique moyenne du prédicteur.

2.3.4 Conclusion

Nous nous sommes concentrés dans cette partie sur la présentation de différents critères pour choisir un plan d'expérience numérique. Toutefois, il est nécessaire d'avoir recours à des algorithmes souvent stochastiques d'optimisation pour construire certains types de plans notamment les plans minimisant une discrépance, les plans minimax, maximin, et les plans optimaux présentés dans la partie précédente. Il peut être intéressant d'optimiser le critère dans une classe donnée de plans d'expérience tels les plans en hypercube latin (LHD). Jin *et al.* (2005) proposent un algorithme de recherche dans la classe des LHD, de plans minimisant la discrépance L_2 centrée, de plans de type maximin et de plans maximisant l'entropie.

Nous renvoyons aussi à Fang *et al.* (2006) pour des algorithmes de construction de plans d'expérience. Dans la majorité des cas, le domaine des entrées E est supposé hypercubique. Il est nécessaire de faire cette hypothèse par exemple pour que la notion de LHD ait un sens. Cependant, ce n'est pas toujours le cas. Dans le chapître 4, nous proposerons un algorithme de recherche de plans d'expérience maximin dans un domaine non hypercubique.
Suivant les objectifs visés par la construction du métamodèle, des stratégies séquentielles d'enrichissement du plan d'expérience sont possibles. C'est notamment le cas de l'algorithme proposé par Jones *et al.* (1998) pour trouver le maximum d'une fonction boîte noire coûteuse. Certains types de plans d'expérience peuvent être enrichis séquentiellement sans en détruire la structure. C'est le cas des suites quasi uniforme de Sobol. Par contre, il est très difficile voire impossible de conserver la structure d'hypercube latin, la propriété maximin, ou des conditions d'optimalité lors de l'ajout d'un point à un plan.

Bibliographie

Adler, R. J. (1981). *The Geometry of Random Fields*. John Wiley & Sons Inc.

An, J. et Owen, A. B. (2001). Quasi-regression. *Journal of Complexity*, 17:588–607.

Aronszajn, N. (1950). Theory of reproducing kernel. *Transactions of American Mathematical Society*, 68(3):337–404.

Atkinson, A. et Donev, A. (1992). *Optimum Experimental Designs*. Oxford Science Publications, Oxford.

Bishop, C. M. (2006). *Pattern Recognition and Machine Learning*. Springer, New-York.

Bratley, P. et Fox, B. L. (1988). Algorithm 659 : Implementing sobol's quasirandom sequence generator. *ACM Trans. Math. Softw.*, 14(1):88–100.

Carroll, R. J., Chen, R., Li, T. H., Newton, H. J., Schmiediche, H., Wang, N. et George, E. I. (1997). Modeling ozone exposure in harris county. *Journal of the American Statistical Association*, 92:392–413.

Cressie, N. (1993). *Statistics for Spatial Data*. Wiley, New York.

Currin, C., Mitchell, T., Morris, M. et Ylvisaker, D. (1991). Bayesian prediction of deterministic functions, with applications to the design and analysis of computer experiments. *Journal of the American Statistical Association*, 86(416):953–963.

Cybenko, G. (1989). Approximation by superposition of a sigmoidal function. *Mathematics of Control, Signals and Systems*, 2:303–314.

De Boor, C. (1978). *A Practical Guide to Splines*. Springer-Verlag, New-York.

den Hertog, D., Kleijnen, J. P. C. et Siem, A. Y. D. (2006). The correct kriging variance estimated by bootstrapping. *Journal of the Operational Research Society*, 57(4):400–409.

Driscoll, M. F. (1973). The reproducing kernel hilbert space structure of the sample paths of a gaussian process. *Zeitschrift für Wahrscheinlichkeitstheorie und verwandte Gebiete*, 26:309–316.

Eilers, P. H. C. et Marx, B. D. (1996). Flexible smoothing with b-splines and penalties. *Statistical Science*, 11:89–121.

Fan, J. et Gijbels, I. (1996). *Local Polynomial Modelling and Its Applications*. Chapman & Hall, London.

Fang, K.-T., Li, R. et Sudjianto, A. (2006). *Design and Modeling for Computer Experiments*. Computer Science and Data Analysis. Chapman & Hall/CRC.

Fang, K. T., Lin, D. K. J., Winker, P. et Zhang, Y. (2000). Uniform design : Theory and application. *Technometrics*, 42(3):237–248.

Fang, K. T. et Wang, Y. (1994). *Number-Theoretic Methods in Statistics*. Chapman & Hall, London.

Fisher, R. (1971). *The Design of Experiments*. Macmillan, 9$^{\text{th}}$ édition.

Friedman, J. H. (1991). Multivariate adaptive regression splines. *Annals of Statistics*, 19:1–141.

Haldar, A. et Mahadevan, S. (2000). *Reliability Assessment Using Stochastic Finite Element Analysis*. John Wiley & Sons, New York.

Hastie, T., Tibshirani, R. et Friedman, J. (2001). *The Elements of Statistical Learning*. Springer Series in Statistics. Springer, New York.

Hoerl, A. E. et Kennard, R. W. (1970). Ridge regression : Biased estimation for nonorthogonal problems. *Technometrics*, 12:55–67.

Jin, R., Chen, W. et Sudjianto, A. (2005). An efficient algorithm for constructing optimal design of computer experiments. *Journal of Statistical Planning and Inference*, 134(1):268 – 287.

Johnson, M. E., Moore, L. M. et Ylvisaker, D. (1990). Minimax and maximin distance designs. *Journal of Statistical Planning and Inference*, 26(2):131 – 148.

Jones, D. R., Schonlau, M. et Welch, W. J. (1998). Efficient global optimization of expensive black-box functions. *Journal of Global Optimization*, 13(4):455–492.

Joseph, V. R. et Hung, Y. (2008). Orthogonal-maximin latin hypercube designs. *Statistica Sinica*, 18:171–186.

Kennedy, M. C. et O'Hagan, A. (2000). Predicting the output from a complex computer code when fast approximations are available. *Biometrika*, 87:1–13.

Kimeldorf, G. et Wahba, G. (1971). Some results on tchebycheffian spline functions. *Journal of Mathematical Analysis and Applications*, 33(1):82–95.

Kleijnen, J. P. C. (1987). *Statistical Tools for Simulation Practitioners*. Marcel Decker, New York.

Koehler, J. R. et Owen, A. B. (1996). Computer experiments. *In Design and analysis of experiments*, volume 13 de *Handbook of Statistics*, pages 261–308. North Holland, Amsterdam.

Krige, D. (1951). A statistical approach to some mine valuations and allied problems at the witwatersrand. Mémoire de D.E.A., University of Witwatersrand.

Li, R. et Sudjianto, A. (2005). Analysis of computer experiments using penalized likelihood in gaussian kriging models. *Technometrics*, 47:111–120.

Lophaven, N., Nielsen, H. et Sondergaard, J. (2002a). Aspects of the matlab toolbox dace. Rapport technique IMM-REP-2002-13, Informatics and Mathematical Modelling, DTU. Available as http ://www.imm.dtu.dk/»hbn/publ/TR0213.ps.

Lophaven, N., Nielsen, H. et Sondergaard, J. (2002b). Dace, a matlab kriging toolbox. Rapport technique IMM-TR-2002-12, DTU. Available to : http ://www2.imm.dtu.dk/ hbn/dace/dace.pdf.

Marchi, S. D. et Schaback, R. (2008). Stability of kernel-based interpolation. *In Advances in Computational Mathematics*.

Marrel, A., Iooss, B., Laurent, B. et Roustant, O. (2009). Calculations of sobol indices for the gaussian process metamodel. *Reliability engineering & systems safety*, 94(3):742–751.

Matheron, G. (1963). Principles of geostatistics. *Economic Geology*, 58:1246–1266.

Matheron, G. (1973). The intrinsic random functions and their applications. *Advances in Applied Probability*, 5(3):439–468.

McKay, M. D., Beckman, R. J. et Conover, W. J. (1979). A comparison of three methods for selecting values of input variables in the analysis of output from a computer code. *Technometrics*, 21(2):239–245.

Mercer, J. (1909). Functions of positive and negative type, and their connection with the theory of integral equations. *Philosophical Transactions of the Royal Society of London. Series A*, 209:415–446.

Mitchell, T., Morris, M. et Ylvisaker, D. (1990). Existence of smoothed stationary processes on an interval. *Stochastic Processes and Their Application*, 35(109-119).

Morris, M. D. et Mitchell, T. J. (1995). Exploratory designs for computer experiments. *Journal of Statistical Planning and Inference*, 43:381–402.

Morris, M. D., Mitchell, T. J. et Ylvisaker, D. (1993). Bayesian design and analysis of computer experiments : Use of derivatives in surface prediction. *Technometrics*, 35:243–255.

Niederreiter, H. (1992). *Random Number Generation and Quasi-Monte Carlo Methods*. SAIM, Philadelphia.

Owen, A. B. (1992). Randomly orthogonal arrays for computer experiments, integration and visualization. *Statistica Sinica*, 2:439–452.

Patterson, H. D. et Thompson, R. (1971). Recovery of inter-block information when block sizes are unequal. *Biometrika*, 58:545–554.

Pillai, N. S., Wu, Q., Liang, F., Mukherjee, S. et Wolpert, R. L. (2007). Characterizing the function space for bayesian kernel models. *Journal of Machine Learning Research*, 8:1769–1797.

Powell, M. J. D. (1987). Radial basis functions for multivariable interpolation : a review. *In* Mason, J. C. et Cox, M. G., éditeurs : *Algorithm for Approximation*, pages 143–167. Clarendon Press, Oxford.

Prasad, N. G. N. et Rao, J. N. K. (1990). The estimation of the mean squared error of small-area estimators. *Journal of the American Statistical Association*, 85:163–171.

Rafajlowicz, E. et Schwabe, R. (2006). Halton and hammersley sequences in multivariate nonparametric regression. *Statistics & Probability Letters*, 76(8):803–812.

Rao, C. R. (1973). *Linear Statistical Inference and its Applications*. Wiley, New-York.

Sacks, J., Schiller, S. B., Mitchell, T. J. et Wynn, H. P. (1989a). Design and analysis of computer experiments (with discussion). *Statistica Sinica*, 4:409–435.

Sacks, J., Schiller, S. B. et Welch, W. J. (1989b). Designs for computer experiments. *Technometrics*, 31(1):41–47.

Santner, T. J., Williams, B. et Notz, W. (2003). *The Design and Analysis of Computer Experiments*. Springer-Verlag.

Schaback, R. (1995a). Comparison of radial basis function interpolants. *In In Multivariate Approximation. From CAGD to Wavelets*, pages 293–305. World Scientific.

Schaback, R. (1995b). Error estimates and condition numbers for radial basis function interpolation. *Advances in Computational Mathematics*, 3:251–264.

Schaback, R. (2007). Kernel-based meshless methods. Rapport technique, Institute for Numerical and Applied Mathematics, Georg-August-University Goettingen.

Schölkopf, B. et Smola, A. J. (2001). *Learning with Kernels : Support Vector Machines, Regularization, Optimization, and Beyond*. MIT Press, Cambridge, MA, USA.

Simpson, T. W., Peplinski, J. D., Koch, P. N. et Allen, J. K. (2001). Metamodels for computer-based engineering design : survey and recommendations. *Engineering with Computers*, 17:129–150.

Sobol, I. M. (1993). Sensitivity analysis for nonlinear mathematical models. *Mathematical Modeling and Computational Experiment*, 1:407–414.

Stein, M. L. (1987). Large sample properties of simulations using latin hypercube sampling. *Technometrics*, 29:143–151.

Stein, M. L. (1999). *Interpolation of Spatial Data : Some Theory for Kriging*. Springer, New York.

Stone, C. J., Hansen, M. H., Kooperberg, C. et Truong, Y. K. (1997). Polynomial splines and their tensor products in extended linear modeling. *Annals of Statistics*, 25:1371–1470.

Tang, Y. (1993). Orthogonal array-based latin hypercubes. *Journal of the American Statistical Association*, 88:1392–1397.

Tibshirani, R. (1994). Regression shrinkage and selection via the lasso. *Journal of the Royal Statistical Society, Series B*, 58:267–288.

Vapnik, V. N. (1996). *The Nature of Statistical Learning Theory*. Springer-Verlag, New-York.

Vert, R. et Vert, J.-P. (2006). Consistency and convergence rates of one-class svms and related algorithms. *J. Mach. Learn. Res.*, 7:817–854.

Wahba, G. (1990). *Spline models for observational data*, volume 59 de *CBMS-NSF Regional Conference Series in Applied Mathematics*. Society for Industrial and Applied Mathematics (SIAM), Philadelphia, PA.

Welch, W. J., Buck, R. J., Sack, J., Wynn, H. P., Mitchell, T. J. et Morris, M. D. (1992). Screening, predicting, and computer experiments. *Technometrics*, 34:15–25.

Wendland, H. (2005). *Scattered data approximation*, volume 17 de *Cambridge Monographs on Applied and Computational Mathematics*. Cambridge University Press, Cambridge.

Wiens, D. P. (1991). Designs for approximately linear regression : two optimality properties of uniform designs. *Statistics & Probability Letters*, 12(3):217–221.

Williams, B. J. (2001). Perk-parametric empirical kriging with examples. Rapport technique 678, Department of Statistics, The Ohio State University.

Zhu, Z. et Zhang, H. (2006). Spatial sampling design under the infill asymptotic framework. *Environmetrics*, 17:337.

Zimmerman, D. L. et Cressie, N. A. (1992). Mean squared prediction error in the spatial linear model with estimated covariance parameters. *Annals of the Institute of Statistical Mathematics*, 44:27–43.

Chapitre 3

Conditionally positive definite kernels : theoretical contribution, application to interpolation and approximation

Résumé

Dans la présentation des méthodes à noyaux de la partie 2.2, les noyaux utilisés sont supposés définis positifs. Nous nous intéressons dans ce chapitre au cas plus général des noyaux conditionnellement définis positifs.
Le théorème 2.4 d'Aronszajn établit qu'à tout noyau défini positif K, est associé un espace de Hilbert à noyau reproduisant (RKHS) de noyau reproduisant K. Le cadre des espaces natifs présenté habituellement dans la littérature ne permet pas une généralisation complète de ce théorème puisque la définition de noyau conditionnellement défini positif qui y est considérée est trop restrictive. Nous proposons une définition plus générale et plus naturelle à partir de laquelle est démontrée une véritable généralisation du théorème d'Aronszajn. Celle-ci établit qu'à chaque couple $(K, \mathbb{P})$ tel que $\mathbb{P}$ est un espace vectoriel fonctionnel de dimension finie et K est un noyau $\mathbb{P}$-conditionnellement défini positif, il existe un unique espace semi-hilbertien de fonctions $\mathcal{H}_{K,\mathbb{P}}$ (RKSHS) satisfaisant une propriété de reproduction généralisée.
Nous montrons pour cette nouvelle définition, que la proposition 2.7 est généralisable et que l'interpolateur obtenu correspond au métamodèle de krigeage (2.14) tout comme dans le cadre des espaces natifs. Nous montrons également que la solution d'un problème de régression régularisée est identifiable dans un RKSHS ce qui généralise la proposition 2.8.

Mots clés : Noyaux (conditionnellement) définis positifs, RKHS, Espace natif, Krigeage, Interpolation à noyaux, Régression régularisée.

Ce chapitre est issu d'une collaboration avec Yves Auffray. Il est disponible au format rapport de recherche INRIA à l'adresse http://hal.inria.fr/inria-00359944/fr/.

Abstract

Since Aronszajn (1950), it is well known that a functional Hilbert space, called Reproducing Kernel Hilbert Space (RKHS), can be associated to any positive definite kernel K. This correspondance is the basis of many useful algorithms. In the more general context of conditionally positive definite kernels the *native spaces* are the usual theoretical framework. However, the definition of *conditionally positive definite* used in that framework is not adapted to extend the results of the positive definite case. We propose a more natural and general definition from which we state a full generalization of Aronszajn's theorem. It states that for every couple $(K, \mathbb{P})$ such that $\mathbb{P}$ is a finite-dimensional vector space of functions and K is a $\mathbb{P}$-conditionally definite positive kernel, there is a unique functional semi-Hilbert space $\mathcal{H}_{K,\mathbb{P}}$ satisfying a generalized reproducing property.
Eventually, we verify that this tool, as native spaces, leads to the same interpolation operator than the one provided by the kriging method and that, using *representer theorem*, we can identify the solution of a regularized regression problem in $\mathcal{H}_{K,\mathbb{P}}$.

Keywords: (Conditionally) positive definite kernels, RKHS, Native space, Kriging, Kernel interpolation, Regularized regression.

3.1 Introduction

Conditionally positive definite kernels arise in many contexts including approximation function algorithms (Wahba, 1990), surface reconstruction (Wendland, 2005; Schaback, 2007), numerical analysis of fluid-structure interactions (Wendland, 2006), computer experiment (Koehler and Owen, 1996; Vazquez, 2005), geostatistics (Cressie, 1993; Wackernagel, 2003). They are intended to generalize the well known positive definite kernel case. As far as we know, the current mostly used and refered to theoretical framework in conditionally positive definite kernel context, is the *native spaces* theory which was firstly developed by Schaback (1997) and more recently by Wendland (2005).

In our opinion, conditionally positive definite kernel definition in the native spaces theory as given by Schaback (1997) and Wendland (2005), is not the natural generalization of the positive definite one. We think that the word *definite* in "conditionally positive definite" has not been interpreted in its full genuine meaning by these authors (see below the first remark following Aronszajn's theorem). As a result, the native space theory does not fully contain the positive definite case: for example, it rules out positive definite kernels defining a finite dimensional reproducing kernel Hilbert space. Moreover, the geometrical simplicity of the positive definite case is lost.

In this paper, we first aim at giving general theoretical foundations to conditionally positive definite kernels used to interpolate or to approximate functions. We want these foundations to fully contain the positive definite case.
In the positive definite kernel case the key property is Aronszajn's theorem, which we recall here.
Let $K : E \times E \mapsto \mathbb{R}$ be a positive definite kernel: that is K is symmetric and satisfies the following property

$$\forall (\lambda_1, \mathbf{x}_1) \dots (\lambda_N, \mathbf{x}_N) \in \mathbb{R} \times E, \sum_{1 \leq l,m \leq N} \lambda_l \lambda_m K(\mathbf{x}_l, \mathbf{x}_m) \geq 0 \,.$$

For any $\mathbf{x} \in E$ let us denote by $K_{\mathbf{x}}$ the partial function $\mathbf{x}' \in E \mapsto K(\mathbf{x}, \mathbf{x}') \in \mathbb{R}$.
Let $\mathcal{F}_K$ be the vector space of (finite) linear combinations of functions taken in $\{K_{\mathbf{x}}, \mathbf{x} \in E\}$. It is easy to see that the formula

$$< \sum_{l=1}^{L} \lambda_l K_{\mathbf{x}_l}, \sum_{m=1}^{M} \mu_m K_{\mathbf{x}'_m} >_{\mathcal{F}_K} = \sum_{l=1}^{L} \sum_{m=1}^{M} \lambda_l \mu_m K(\mathbf{x}_l, \mathbf{x}'_m)$$

defines a symmetric, positive, bilinear form on $\mathcal{F}_K$. Now Aronszajn's theorem (Aronszajn, 1950) reads as

Theorem 3.1 (Aronszajn)**.**

1. $<,>_{\mathcal{F}_K}$*, as a bilinear form, is positive* **definite**.
2. *There is a unique Hilbert space of real functions defined on* E*,* $\mathcal{H}_K$*, called Reproducing Kernel Hilbert Space (RKHS) of kernel* K *such that*
   - $(\mathcal{F}_K, <,>_{\mathcal{F}_K})$ *is a prehilbertian subspace of* $\mathcal{H}_K$*,*
   - *the following reproducing property is satisfied*

$$\forall f \in \mathcal{H}_K, \ \mathbf{x} \in E, \ f(\mathbf{x}) =< f, K_{\mathbf{x}} >_{\mathcal{H}_K} \,. \tag{3.1}$$

Let us make several remarks, in the light of that theorem.
First of all, the word *definite* in *positive definite kernels* relates to the positive definiteness of $<,>_{\mathcal{F}_K}$, as stated by point 1 of Theorem 3.1, and not to the positive definiteness of matrices

$$(K(\mathbf{x}_l, \mathbf{x}_m))_{1\leq l,m\leq N}, N \in \mathbb{N}, (\mathbf{x}_1, \ldots, \mathbf{x}_N) \in E^N$$

which are not necessary positive definite.
Secondly, let $\mathbf{X} = \{\mathbf{x}_1, \ldots, \mathbf{x}_N\} \subset E$ be a set of points, the reproducing property (3.1), leads to a simple and useful charaterization of the orthogonal projection $S_{K,\mathbf{X}}(f)$ of any $f \in \mathcal{H}_K$ on $\mathcal{F}_K(\mathbf{X})$, the subspace of $\mathcal{F}_K$ spanned by $K_{\mathbf{x}_1}, \ldots, K_{\mathbf{x}_N}$: it is the interpolation of f at the points of $\mathbf{X}$ with minimal $\mathcal{H}_K$-norm.
At last, as an easy consequence of the previous fact, the well known *representer theorem* (Kimeldorf and Wahba, 1971) applied here in a regularized regression context, is stated as follows:
let $(\mathbf{x}_1, \mathbf{y}_1), \ldots, (\mathbf{x}_N, \mathbf{y}_N) \in E \times \mathbb{R}$ and $\lambda > 0$,
any solution of

$$\min_{f\in\mathcal{H}_K} \sum_{k=1}^{N} (\mathbf{y}_k - f(\mathbf{x}_k))^2 + \lambda ||f||_{\mathcal{H}_K}$$

lies in $\mathcal{F}_K(\mathbf{X})$.

The main result of our work has exactly the same form as Aronszajn's theorem:

- K, instead of being positive definite, will be what we will call, after a detailed justification, $\mathbb{P}$-conditionally positive definite, where $\mathbb{P}$ is a finite-dimensional vector space of real functions defined on E.
- The RKHS$\mathcal{H}_K$ will be replaced by a $\mathbb{P}$-dependent semi-Hilbert space of functions, satisfying a generalized reproducing property and leading to the acronym ($\mathbb{P}$-)RKSHS.
- Aronszajn's theorem is recovered for $\mathbb{P} = \{0\}$.

This paper is organised as follows. Section 3.2 introduces the mathematical objects and notations we need. Section 3.3 details the relations between these objects leading to a summing up, simple commutative diagram. Section 3.4 is the core of the paper. There we formulate "our" conditionally positive definite definition, state and prove Aronszajn's theorem analog for conditionally positive definite context. Sections 3.5 and 3.6 are devoted to applications. We first state and prove a generalized interpolation result in the spirit of the second remark following Aronszajn's theorem and the useful Lagrange formulation of these interpolations. Besides, we revisit the regularized regression problem in the context of our conditionally positive definite kernels: the representer theorem is verified and an explicit solution of the regularized regression problem is given.

3.2 First definitions and notation

In this paper, we will denote by

- E an arbitrary set and $\mathbb{R}^E$ the vector space of real functions defined on E;

- $\mathbb{P} \subset \mathbb{R}^E$ a n dimensional vector space;

- $K : E \times E \to \mathbb{R}$, our generic *kernel*, which is assumed to be, at least, symmetric and $\mathbb{P}$-conditionally positive:

Definition 3.1 ($\mathbb{P}$-conditionally positive kernel)**.** *The kernel K is $\mathbb{P}$-conditionally positive if the following property is satisfied:*

$$\sum_{1 \leq k,l \leq L} \lambda_l \lambda_k K(\mathbf{x}_l, \mathbf{x}_k) \geq 0\,,$$

for all $L \in \mathbb{N}$, $\mathbf{x}_1, \ldots, \mathbf{x}_L \in E$, $\lambda_1, \ldots, \lambda_L \in \mathbb{R}$ such that

$$\forall p \in \mathbb{P}, \sum_{l=1}^{L} \lambda_l p(\mathbf{x}_l) = 0\,.$$

- $K_{\mathbf{x}}$, for $\mathbf{x} \in E$, the partial function $\mathbf{x}' \in E \mapsto K(\mathbf{x}, \mathbf{x}')$.

3.2.1 Measures with finite support

Let us set:

- $\delta_{\mathbf{x}}$ the Dirac measure concentrated at $\mathbf{x}$, for any $\mathbf{x} \in E$.

- $\mathcal{M}$ the set of real measures on E with finite support:

$$\boldsymbol{\mu} \in \mathcal{M} \;\Leftrightarrow \begin{cases} \boldsymbol{\mu} \text{ is the null measure on } E \\ \text{or} \\ \exists \mathbf{x}_1, \ldots, \mathbf{x}_N \in E \text{ pairwise distinct, and} \\ \mu_1, \ldots, \mu_N \in (\mathbb{R} - \{0\}),\ \boldsymbol{\mu} = \sum_{k=1}^{N} \mu_k \delta_{\mathbf{x}_k} \end{cases}.$$

  $\mathcal{M}$ is obviously a real vector space a base of which is $\{\delta_{\mathbf{x}} : \mathbf{x} \in E\}$.

- $\boldsymbol{\mu}(f) = \sum_{k=1}^{N} \mu_k f(\mathbf{x}_k)$ the integral of any $f \in \mathbb{R}^E$ against any $\boldsymbol{\mu} = \sum_{k=1}^{N} \mu_k \delta_{\mathbf{x}_k} \in \mathcal{M}$.

- $\mathcal{M}_{\mathbb{P}}$ the subspace of measures lying in $\mathcal{M}$ vanishing on $\mathbb{P}$:

$$\boldsymbol{\mu} \in \mathcal{M}_{\mathbb{P}} \;\Leftrightarrow \boldsymbol{\mu}(p) = 0,\ \forall p \in \mathbb{P}\,.$$

- If we are given $\mathbf{X} \subset E$,

  - $\mathcal{M}(\mathbf{X}) = \{\boldsymbol{\lambda} = \sum_{l=1}^{L} \lambda_l \delta_{\mathbf{x}_l} : (\lambda_1, \mathbf{x}_1), \ldots, (\lambda_L, \mathbf{x}_L) \in \mathbb{R} \times \mathbf{X}\}$,
  - $\mathcal{M}_{\mathbb{P}}(\mathbf{X}) = \mathcal{M}(\mathbf{X}) \cap \mathcal{M}_{\mathbb{P}}$.

3.2.2 $\mathbb{P}$-unisolvent set

Definition 3.2. *$\mathbf{X} = \{\mathbf{x}_1, \ldots, \mathbf{x}_N\} \subset E$ is said to be $\mathbb{P}$-unisolvent if the linear application*

$$L^{\mathbf{X}} : p \in \mathbb{P} \mapsto (p(\mathbf{x}_1), \ldots, p(\mathbf{x}_N)) \in \mathbb{R}^N$$

is injective, or equivalently, if the only $p \in \mathbb{P}$ which vanishes on every $\mathbf{x} \in \mathbf{X}$ is $0 \in \mathbb{P}$.

In this paper, we will always assume that $\mathbb{P}$ is such that $\mathbb{P}$-unisolvent sets exist.

Recalling that $\dim(\mathbb{P}) = n$, elementary arguments lead to:

Lemma 3.1. *A $\mathbb{P}$-unisolvent set is minimal if and only if it contains exactly n elements.*

Now, let

$$\boldsymbol{\Xi} = \{\boldsymbol{\xi}_1, \ldots, \boldsymbol{\xi}_n\}$$

be a minimal $\mathbb{P}$-unisolvent set.
Since $L^{\boldsymbol{\Xi}}$ is a bijection, the relations

$$\left\{ \begin{array}{rcll} h_k^{\boldsymbol{\Xi}}(\boldsymbol{\xi}_j) & = & 1 & \text{if } j = k \\ h_k^{\boldsymbol{\Xi}}(\boldsymbol{\xi}_j) & = & 0 & \text{otherwise} \end{array} \right. , \; k = 1, \ldots, n \,,$$

which are equivalent to

$$L_{\mathbb{P}}^{\boldsymbol{\Xi}}(h_k^{\boldsymbol{\Xi}}) = \mathbf{e}_k, k = 1, \ldots, n \,,$$

where $\mathbf{e}_k$ is the kth vector of the $\mathbb{R}^n$ canonical basis, define a $\mathbb{P}$ basis $(h_1^{\boldsymbol{\Xi}}, \ldots, h_n^{\boldsymbol{\Xi}})$.
Let us then define

$$\boldsymbol{\pi}^{\boldsymbol{\Xi}} : f \in \mathbb{R}^E \mapsto \sum_{k=1}^{n} f(\boldsymbol{\xi}_k) h_k^{\boldsymbol{\Xi}} \in \mathbb{P} \,.$$

This immediately follows:

Proposition 3.1. *$\boldsymbol{\pi}^{\boldsymbol{\Xi}}$ is a projector on $\mathbb{P}$, and, for all $f \in \mathbb{R}^E$, $\boldsymbol{\pi}^{\boldsymbol{\Xi}}(f)$ interpolates f on $\boldsymbol{\Xi}$.*

For any element $\mathbf{x}$ of E, let us introduce

$$\delta_{\mathbf{x}}^{\boldsymbol{\Xi}} = \delta_{\mathbf{x}} - \sum_{i=1}^{n} h_i^{\boldsymbol{\Xi}}(\mathbf{x}) \delta_{\boldsymbol{\xi}_i}.$$

Obviously:

- $\delta_{\boldsymbol{\xi}_k}^{\boldsymbol{\Xi}} = 0, \; k = 1, \ldots, n,$
- $\delta_{\mathbf{x}}^{\boldsymbol{\Xi}} \in \mathcal{M}_{\mathbb{P}}$, since

$$\delta_{\mathbf{x}}^{\boldsymbol{\Xi}}(h_k^{\boldsymbol{\Xi}}) = \delta_{\mathbf{x}}(h_k^{\boldsymbol{\Xi}}) - \sum_{i=1}^{n} h_i^{\boldsymbol{\Xi}}(\mathbf{x}) \delta_{\boldsymbol{\xi}_i}(h_k^{\boldsymbol{\Xi}}) = h_k^{\boldsymbol{\Xi}}(\mathbf{x}) - \sum_{i=1}^{n} h_i^{\boldsymbol{\Xi}}(\mathbf{x}) h_k^{\boldsymbol{\Xi}}(\boldsymbol{\xi}_i) = h_k^{\boldsymbol{\Xi}}(\mathbf{x}) - h_k^{\boldsymbol{\Xi}}(\mathbf{x}) = 0 \,.$$

We then establish this technical proposition that will be useful in the sequel.

Proposition 3.2. *Let* $\boldsymbol{\Xi} = \{\boldsymbol{\xi}_1, \dots, \boldsymbol{\xi}_n\}$ *be any minimal* $\mathbb{P}$*-unisolvent set. Every* $\boldsymbol{\lambda} = \sum_{l=1}^{L} \lambda_l \delta_{\mathbf{x}_l} \in \mathcal{M}$ *has the alternative form:*

$$\boldsymbol{\lambda} = \sum_{l=1}^{L} \lambda_l \delta_{\mathbf{x}_l}^{\boldsymbol{\Xi}} + \sum_{k=1}^{n} \boldsymbol{\lambda}(h_k^{\boldsymbol{\Xi}}) \delta_{\boldsymbol{\xi}_k} . \tag{3.2}$$

As a consequence,

- $\mathcal{M}_{\mathbb{P}}(\boldsymbol{\Xi}) = \{0\}$,
- *for any* $\mathbf{X} \subset E$*, such that* $\boldsymbol{\Xi} \subset \mathbf{X}$, $\{\delta_{\mathbf{x}}^{\boldsymbol{\Xi}} : \mathbf{x} \in \mathbf{X} - \boldsymbol{\Xi}\}$ *is a* $\mathcal{M}_{\mathbb{P}}(\mathbf{X})$*-basis.*

Proof
We readily have:

$$\begin{aligned}\sum_{l=1}^{L} \lambda_l \delta_{\mathbf{x}_l}^{\boldsymbol{\Xi}} &= \sum_{l=1}^{L} \lambda_l (\delta_{\mathbf{x}_l} - \sum_{k=1}^{n} h_k^{\boldsymbol{\Xi}}(\mathbf{x}_l) \delta_{\boldsymbol{\xi}_k}) = \sum_{l=1}^{L} \lambda_l \delta_{\mathbf{x}_l} - \sum_{l=1}^{L} \sum_{k=1}^{n} \lambda_l h_k^{\boldsymbol{\Xi}}(\mathbf{x}_l) \delta_{\boldsymbol{\xi}_k} \\ &= \boldsymbol{\lambda} - \sum_{k=1}^{n} \left[\sum_{l=1}^{L} \lambda_l h_k^{\boldsymbol{\Xi}}(\mathbf{x}_l) \right] \delta_{\boldsymbol{\xi}_k} = \boldsymbol{\lambda} - \sum_{k=1}^{n} \boldsymbol{\lambda}(h_k^{\boldsymbol{\Xi}}) \delta_{\boldsymbol{\xi}_k} ,\end{aligned}$$

hence (3.2). $\mathcal{M}_{\mathbb{P}}(\boldsymbol{\Xi}) = \{0\}$ follows immediately.
Let $\mathbf{X}$ be a subset of E which contains $\boldsymbol{\Xi}$.
Any $\boldsymbol{\lambda} = \sum_{i=1}^{N} \lambda_i \delta_{\mathbf{x}_i} \in \mathcal{M}_{\mathbb{P}}(\mathbf{X})$ can be written, using (3.2): $\boldsymbol{\lambda} = \sum_{l=1}^{L} \lambda_l \delta_{\mathbf{x}_l}^{\boldsymbol{\Xi}}$. Thus, since $\delta_{\mathbf{x}}^{\boldsymbol{\Xi}} \in \mathcal{M}_{\mathbb{P}}(\mathbf{X}), \mathbf{x} \in \mathbf{X}$, $\{\delta_{\mathbf{x}}^{\boldsymbol{\Xi}} : \mathbf{x} \in \mathbf{X} - \boldsymbol{\Xi}\}$ spans $\mathcal{M}_{\mathbb{P}}(\mathbf{X})$.
Moreover, $\{\delta_{\mathbf{x}}^{\boldsymbol{\Xi}} : \mathbf{x} \in (\mathbf{X} - \boldsymbol{\Xi})\}$ are linearly independent.
Indeed, let $\mathbf{x}_1, \dots, \mathbf{x}_N$ be N pairwise distinct elements of $\mathbf{X} - \boldsymbol{\Xi}$. For $(\alpha_1, \dots, \alpha_N) \in \mathbb{R}^N$ we have from (3.2):

$$\sum_{k=1}^{N} \alpha_k \delta_{\mathbf{x}_k}^{\boldsymbol{\Xi}} = \sum_{k=1}^{N} \alpha_k \delta_{\mathbf{x}_k} - \sum_{i=1}^{n} \left[\sum_{k=1}^{N} \alpha_k h_i^{\boldsymbol{\Xi}}(\mathbf{x}_k) \right] \delta_{\boldsymbol{\xi}_i} .$$

But $\mathbf{x}_1, \dots, \mathbf{x}_N, \boldsymbol{\xi}_1, \dots, \boldsymbol{\xi}_n$ are distinct, thus $\delta_{\mathbf{x}_1}, \dots, \delta_{\mathbf{x}_N}, \delta_{\boldsymbol{\xi}_1}, \dots, \delta_{\boldsymbol{\xi}_n}$ are linearly independent and

$$\sum_{k=1}^{N} \alpha_k \delta_{\mathbf{x}_k}^{\boldsymbol{\Xi}} = 0 \Rightarrow \sum_{k=1}^{N} \alpha_k \delta_{\mathbf{x}_k} - \sum_{i=1}^{n} \left[\sum_{k=1}^{N} \alpha_k h_i^{\boldsymbol{\Xi}}(\mathbf{x}_k) \right] \delta_{\boldsymbol{\xi}_i} = 0 \Rightarrow \alpha_k = 0, k = 1, \dots, N .$$

□
Let us now define:

$$\Phi^{\boldsymbol{\Xi}} : \boldsymbol{\mu} \in \mathcal{M} \mapsto \boldsymbol{\mu} - \sum_{k=1}^{n} \boldsymbol{\mu}(h_k^{\boldsymbol{\Xi}}) \delta_{\boldsymbol{\xi}_k} \in \mathcal{M}.$$

The following facts are obvious:

- $\Phi^{\boldsymbol{\Xi}}(\sum_{i=1}^{N} \lambda_i \delta_{\mathbf{x}_i}) = \sum_{i=1}^{N} \lambda_i \delta_{\mathbf{x}_i}^{\boldsymbol{\Xi}}$,
- the relation (3.2) can be rephrased as

$$\boldsymbol{\lambda} = \Phi^{\boldsymbol{\Xi}}(\boldsymbol{\lambda}) + \sum_{k=1}^{n} \boldsymbol{\lambda}(h_k^{\boldsymbol{\Xi}}) \delta_{\boldsymbol{\xi}_k} ,$$

- $\Phi^{\boldsymbol{\Xi}}$ is a projection on $\mathcal{M}_{\mathbb{P}}$.

3.3 Bilinear forms induced by K

Let $\boldsymbol{\mu} = \sum_{m=1}^{M} \mu_m \delta_{\mathbf{x}_m}$ and $\boldsymbol{\lambda} = \sum_{l=1}^{L} \lambda_l \delta_{\mathbf{z}_l}$ be two measures taken in $\mathcal{M}$.
The formula

$$< \boldsymbol{\mu}, \boldsymbol{\lambda} >_{\mathcal{M},K} = \sum_{m=1}^{M} \sum_{l=1}^{L} \mu_m \lambda_l K(\mathbf{x}_m, \mathbf{z}_l)$$

defines a symmetric bilinear form $<,>_{\mathcal{M},K}$ on $\mathcal{M}$.
$\mathbb{P}$-conditional positiveness of K means that the restriction of $<,>_{\mathcal{M},K}$ to $\mathcal{M}_{\mathbb{P}}$ is positive.

Kernel K also induces a natural linear application

$$F_K : \boldsymbol{\mu} = \sum_{m=1}^{M} \mu_m \delta_{\mathbf{x}_m} \in \mathcal{M} \mapsto \sum_{m=1}^{M} \mu_m K_{\mathbf{x}_m} \in \mathbb{R}^E.$$

For any $\mathbf{X} \subset E$, let us then set

$$\mathcal{F}_K(\mathbf{X}) = F_K(\mathcal{M}(\mathbf{X}))$$

and

$$\mathcal{F}_{K,\mathbb{P}}(\mathbf{X}) = F_K(\mathcal{M}_{\mathbb{P}}(\mathbf{X}))\,,$$

which will be merely denoted $\mathcal{F}_K$ and $\mathcal{F}_{K,\mathbb{P}}$ when $\mathbf{X} = E$.

Using F_K, we can carry the bilinear structure from $\mathcal{M}$ to $\mathcal{F}_K$:

Proposition 3.3. *Let f, g be functions in $\mathcal{F}_K$ and $\boldsymbol{\lambda}, \boldsymbol{\mu} \in \mathcal{M}$ such that $f = F_K(\boldsymbol{\lambda})$ and $g = F_K(\boldsymbol{\mu})$.*
The formula

$$< f, g >_{\mathcal{F}_K} = < \boldsymbol{\lambda}, \boldsymbol{\mu} >_{\mathcal{M},K}$$

only depends on f and g, and not on the particular choice of $\boldsymbol{\lambda}, \boldsymbol{\mu}$.
Thus it defines a symmetric bilinear form on $\mathcal{F}_K$ whose restriction to $\mathcal{F}_{K,\mathbb{P}}$ is positive.
This reproducing formula holds for any $g \in \mathcal{F}_K$ and $\mathbf{x} \in E$:

$$< K_{\mathbf{x}}, g >_{\mathcal{F}_K} = g(\mathbf{x}). \tag{3.3}$$

Proof
Let us start with

Lemma 3.2. *For every $\boldsymbol{\lambda}, \boldsymbol{\mu} \in \mathcal{M}$,*

$$< \boldsymbol{\lambda}, \boldsymbol{\mu} >_{\mathcal{M},K} = \boldsymbol{\lambda}(F_K(\boldsymbol{\mu})). \tag{3.4}$$

Proof
Let $\boldsymbol{\lambda} = \sum_{l=1}^{L} \lambda_l \delta_{\mathbf{x}_l}$ and $\boldsymbol{\mu} = \sum_{m=1}^{M} \mu_m \delta_{\mathbf{z}_m}$ be the expressions of $\boldsymbol{\lambda}$ and $\boldsymbol{\mu}$ in the $\mathcal{M}$ basis $\{\delta_{\mathbf{x}} : \mathbf{x} \in E\}$.
We readily have:

$$\boldsymbol{\lambda}(F_K(\boldsymbol{\mu})) = \sum_{l=1}^{L} \lambda_l \delta_{\mathbf{x}_l}(\sum_{m=1}^{M} \mu_m K_{\mathbf{z}_m}) = \sum_{l=1}^{L} \sum_{m=1}^{M} \lambda_l \mu_m K(\mathbf{x}_l, \mathbf{z}_m) = < \boldsymbol{\lambda}, \boldsymbol{\mu} >_{\mathcal{M},K}\ .$$

□
From (3.4) we have

$$\begin{aligned} < \boldsymbol{\lambda}, \boldsymbol{\mu} >_{\mathcal{M},K} &= \boldsymbol{\lambda}(F_K(\boldsymbol{\mu})) = \boldsymbol{\lambda}(g) \\ &= \boldsymbol{\mu}(F_K(\boldsymbol{\lambda})) = \boldsymbol{\mu}(f) \end{aligned}$$

and $< f, g >_{\mathcal{F}_K}$ only depends on f and g.

Now, since the restriction of $<,>_{\mathcal{M},K}$ to $\mathcal{M}_{\mathbb{P}}$ is positive, taking $f = F_K(\boldsymbol{\lambda}) \in \mathcal{F}_{K,\mathbb{P}}$ with $\boldsymbol{\lambda} \in \mathcal{M}_{\mathbb{P}}$ leads to:

$$0 \leq < \boldsymbol{\lambda}, \boldsymbol{\lambda} >_{\mathcal{M},K} = < f, f >_{\mathcal{F}_K}$$

and the restriction of $<,>_{\mathcal{F}_K}$ to $\mathcal{F}_{K,\mathbb{P}}$ is positive.
Applied to $g = F_K(\boldsymbol{\mu})$ and $f = K_{\mathbf{x}} = F_K(\delta_{\mathbf{x}})$, (3.4) leads to the reproducing formula:

$$< K_{\mathbf{x}}, g >_{\mathcal{F}_K} = < \delta_{\mathbf{x}}, \boldsymbol{\mu} >_{\mathcal{M},K} = \delta_{\mathbf{x}}(F_K(\boldsymbol{\mu})) = g(\mathbf{x}) \, .$$

□

From K and a minimal $\mathbb{P}$-unisolvent set $\boldsymbol{\Xi}$, we introduce the new kernel $K^{\boldsymbol{\Xi}}$:

$$K^{\boldsymbol{\Xi}} : (\mathbf{x}, \mathbf{x}') \in E^2 \mapsto < \delta_{\mathbf{x}}^{\boldsymbol{\Xi}}, \delta_{\mathbf{x}'}^{\boldsymbol{\Xi}} >_{\mathcal{M},K} \, .$$

This simple calculation:

$$\sum_{1 \leq i,j \leq N} \lambda_i \lambda_j K^{\boldsymbol{\Xi}}(\mathbf{x}_i, \mathbf{x}_j) = \sum_{1 \leq i,j \leq N} \lambda_i \lambda_j < \delta_{\mathbf{x}_i}^{\boldsymbol{\Xi}}, \delta_{\mathbf{x}_j}^{\boldsymbol{\Xi}} >_{\mathcal{M},K} = < \sum_{i=1}^{N} \lambda_i \delta_{\mathbf{x}_i}^{\boldsymbol{\Xi}}, \sum_{i=1}^{N} \lambda_i \delta_{\mathbf{x}_i}^{\boldsymbol{\Xi}} >_{\mathcal{M},K} \geq 0$$

leads to

Proposition 3.4. *$K^{\boldsymbol{\Xi}}$ is a (unconditionally) positive kernel.*

We now sum up the main relations between bilinear structures induced by a conditionally positive kernel we met up to this point. This summary consists in the following commutative diagram:

$$\begin{array}{ccc} (\mathcal{M}_{\mathbb{P}}(\mathbf{X}), <,>_{\mathcal{M},K}) & \xleftarrow{\Phi^{\boldsymbol{\Xi}}} & \left(\mathcal{M}(\mathbf{X}), <,>_{\mathcal{M},K^{\boldsymbol{\Xi}}}\right) \\ \Big\downarrow F_K & \searrow^{F_K^{\boldsymbol{\Xi}}} & \Big\downarrow F_{K^{\boldsymbol{\Xi}}} \\ (\mathcal{F}_{K,\mathbb{P}}(\mathbf{X}), <,>_{\mathcal{F}_K}) & \xrightarrow{\mathrm{Id} - \boldsymbol{\pi}^{\boldsymbol{\Xi}}} & \left(\mathcal{F}_{K^{\boldsymbol{\Xi}}}(\mathbf{X}), <,>_{\mathcal{F}_{K^{\boldsymbol{\Xi}}}}\right) \end{array} \tag{3.5}$$

where

- $\mathbf{X}$ is any subset of E;
- $\boldsymbol{\Xi} \subset \mathbf{X}$ is a minimal $\mathbb{P}$-unisolvent set;
- $F_{K^{\boldsymbol{\Xi}}}$ and $\mathcal{F}_{K^{\boldsymbol{\Xi}}}$ are the analogs of F_K and $\mathcal{F}_K$ with $K^{\boldsymbol{\Xi}}$ in place of K;
- $F_K^{\boldsymbol{\Xi}} : \mathcal{M} \mapsto \mathbb{R}^E$ is specified by

$$F_K^{\boldsymbol{\Xi}}(\boldsymbol{\lambda}) : \mathbf{x} \mapsto < \boldsymbol{\lambda}, \delta_{\mathbf{x}}^{\boldsymbol{\Xi}} >_{\mathcal{M},K} .$$

The diagram (3.5) must be read with the following conventions:

- Any arrow between two bilinear structures is a morphism for them.
- Any two oriented paths from one structure to another lead to the same composite mapping: e.g. $F_{K^{\boldsymbol{\Xi}}} = F_K^{\boldsymbol{\Xi}} \circ \Phi^{\boldsymbol{\Xi}}$.

The "mapping" part of that diagram is the immediate consequence of

Proposition 3.5. *For all* $\boldsymbol{\lambda} \in \mathcal{M}$,

$$F_K(\boldsymbol{\lambda}) = \boldsymbol{\pi}^{\boldsymbol{\Xi}}(F_K(\boldsymbol{\lambda})) + F_K^{\boldsymbol{\Xi}}(\boldsymbol{\lambda}), \tag{R_1}$$

$$F_{K^{\boldsymbol{\Xi}}}(\boldsymbol{\lambda}) = F_K^{\boldsymbol{\Xi}}(\Phi^{\boldsymbol{\Xi}}(\boldsymbol{\lambda})). \tag{R_2}$$

Proof
(R_1) follows from this:

$$\begin{aligned}
F_K^{\boldsymbol{\Xi}}(\boldsymbol{\lambda})(\mathbf{x}) &= < \boldsymbol{\lambda}, \delta_{\mathbf{x}}^{\boldsymbol{\Xi}} >_{\mathcal{M},K} \\
&= < \boldsymbol{\lambda}, \delta_{\mathbf{x}} - \sum_{i=1}^{n} h_i^{\boldsymbol{\Xi}}(\mathbf{x}) \delta_{\boldsymbol{\xi}_i} >_{\mathcal{M},K} \\
&= < \boldsymbol{\lambda}, \delta_{\mathbf{x}} >_{\mathcal{M},K} - \sum_{i=1}^{n} h_i^{\boldsymbol{\Xi}}(\mathbf{x}) < \boldsymbol{\lambda}, \delta_{\boldsymbol{\xi}_i} >_{\mathcal{M},K} \\
&= F_K(\boldsymbol{\lambda})(\mathbf{x}) - \sum_{i=1}^{n} h_i^{\boldsymbol{\Xi}}(\mathbf{x}) F_K(\boldsymbol{\lambda})(\boldsymbol{\xi}_i) \\
&= F_K(\boldsymbol{\lambda})(\mathbf{x}) - \boldsymbol{\pi}^{\boldsymbol{\Xi}}(F_K(\boldsymbol{\lambda}))(\mathbf{x}).
\end{aligned}$$

(R_2) comes from: if $\boldsymbol{\lambda} = \sum_{l=1}^{L} \lambda_l \delta_{\mathbf{x}_l}$

$$\begin{aligned}
F_{K^{\boldsymbol{\Xi}}}(\boldsymbol{\lambda})(\mathbf{x}) &= < \boldsymbol{\lambda}, \delta_{\mathbf{x}} >_{\mathcal{M},K^{\boldsymbol{\Xi}}} = \sum_{l=1}^{L} \lambda_l K^{\boldsymbol{\Xi}}(\mathbf{x}_l, \mathbf{x}) = \sum_{l=1}^{L} \lambda_l < \delta_{\mathbf{x}_l}^{\boldsymbol{\Xi}}, \delta_{\mathbf{x}}^{\boldsymbol{\Xi}} >_{\mathcal{M},K} \\
&= < \sum_{l=1}^{L} \lambda_l \delta_{\mathbf{x}_l}^{\boldsymbol{\Xi}}, \delta_{\mathbf{x}}^{\boldsymbol{\Xi}} >_{\mathcal{M},K} = F_K^{\boldsymbol{\Xi}}(\sum_{l=1}^{L} \lambda_l \delta_{\mathbf{x}_l}^{\boldsymbol{\Xi}})(\mathbf{x}) = F_K^{\boldsymbol{\Xi}}(\Phi^{\boldsymbol{\Xi}}(\boldsymbol{\lambda}))(\mathbf{x}).
\end{aligned}$$

□
The morphism part of (3.5) is easily verified from:

Proposition 3.6.

1. Φ^{Ξ} *is a morphism between* $(\mathcal{M}(\mathbf{X}), <,>_{\mathcal{M},K^{\Xi}})$ *and* $(\mathcal{M}_{\mathbb{P}}(\mathbf{X}), <,>_{\mathcal{M},K})$.

2. $\mathrm{Id} - \boldsymbol{\pi}^{\Xi}$ *is a morphism between* $(\mathcal{F}_{K,\mathbb{P}}(\mathbf{X}), <,>_{\mathcal{F}_K})$ *and* $(\mathcal{F}_{K^{\Xi}}(\mathcal{M}(\mathbf{X})), <,>_{\mathcal{F}_{K^{\Xi}}})$.

Proof

1. $< \Phi^{\Xi}(\delta_{\mathbf{x}}), \Phi^{\Xi}(\delta_{\mathbf{x}'}) >_{\mathcal{M},K} = < \delta^{\Xi}_{\mathbf{x}}, \delta^{\Xi}_{\mathbf{x}'} >_{\mathcal{M},K} = K^{\Xi}(\mathbf{x},\mathbf{x}') = < \delta_{\mathbf{x}}, \delta_{\mathbf{x}'} >_{\mathcal{M},K^{\Xi}}$ leads immediately to
$$< \Phi^{\Xi}(\boldsymbol{\lambda}), \Phi^{\Xi}(\boldsymbol{\mu}) >_{\mathcal{M},K} = < \boldsymbol{\lambda}, \boldsymbol{\mu} >_{\mathcal{M},K^{\Xi}}, \tag{3.6}$$
for any $\boldsymbol{\lambda}, \boldsymbol{\mu} \in \mathcal{M}$.

2. Let f and g be two functions in $\mathcal{F}_{K,\mathbb{P}}(\mathbf{X})$: there exists $\boldsymbol{\lambda}, \boldsymbol{\mu} \in \mathcal{M}_{\mathbb{P}}(\mathbf{X})$ such that $f = F_K(\boldsymbol{\lambda})$ and $g = F_K(\boldsymbol{\mu})$.
Recalling that $\boldsymbol{\lambda}, \boldsymbol{\mu} \in \mathcal{M}_{\mathbb{P}}(\mathbf{X}) \Rightarrow \Phi^{\Xi}(\boldsymbol{\lambda}) = \boldsymbol{\lambda},\ \Phi^{\Xi}(\boldsymbol{\mu}) = \boldsymbol{\mu}$, we actually have
$$f = F_K(\Phi^{\Xi}(\boldsymbol{\lambda})) \text{ and } g = F_K(\Phi^{\Xi}(\boldsymbol{\mu}))\,.$$
From Proposition 3.5 it follows
$$f - \boldsymbol{\pi}^{\Xi}(f) = F_{K^{\Xi}}(\Phi^{\Xi}(\boldsymbol{\lambda})) = F_{K^{\Xi}}(\boldsymbol{\lambda})$$
and
$$g - \boldsymbol{\pi}^{\Xi}(g) = F_{K^{\Xi}}(\Phi^{\Xi}(\boldsymbol{\mu})) = F_{K^{\Xi}}(\boldsymbol{\mu})$$
leading to:
$$< f - \boldsymbol{\pi}^{\Xi}(f), g - \boldsymbol{\pi}^{\Xi}(g) >_{\mathcal{F}_{K^{\Xi}}} = < F_{K^{\Xi}}(\boldsymbol{\lambda}), F_{K^{\Xi}}(\boldsymbol{\mu}) >_{\mathcal{F}_{K^{\Xi}}} = < \boldsymbol{\lambda}, \boldsymbol{\mu} >_{\mathcal{M},K^{\Xi}}\ .$$
But F_K definition directly gives:
$$< f, g >_{\mathcal{F}_K} = < F_K(\boldsymbol{\lambda}), F_K(\boldsymbol{\mu}) >_{\mathcal{F}_K} = < \boldsymbol{\lambda}, \boldsymbol{\mu} >_{\mathcal{M},K} = < \Phi^{\Xi}(\boldsymbol{\lambda}), \Phi^{\Xi}(\boldsymbol{\mu}) >_{\mathcal{M},K}\,,$$
then, with (3.6):
$$< f, g >_{\mathcal{F}_K} = < \boldsymbol{\lambda}, \boldsymbol{\mu} >_{\mathcal{M},K^{\Xi}}\ .$$
Hence
$$< f, g >_{\mathcal{F}_K} = < f - \boldsymbol{\pi}^{\Xi}(f), g - \boldsymbol{\pi}^{\Xi}(g) >_{\mathcal{F}_{K^{\Xi}}}\ .$$

□

Remark 3.1. *These consequences of diagram (3.5) will be often used in the sequel:*

$$\forall f \in \mathbb{P} + \mathcal{F}_{K,\mathbb{P}}(\mathbf{X}), f - \boldsymbol{\pi}^{\Xi}(f) \in \mathcal{F}_{K^{\Xi}}(\mathbf{X}), \tag{3.7}$$

$$\mathbb{P} + \mathcal{F}_{K,\mathbb{P}}(\mathbf{X}) = \mathbb{P} + \mathcal{F}_{K^{\Xi}}(\mathbf{X}). \tag{3.8}$$

Indeed, if $f \in \mathbb{P} + \mathcal{F}_{K,\mathbb{P}}(\mathbf{X})$, we can write $f = p + g$ with $p \in \mathbb{P}$ and $g \in \mathcal{F}_{K,\mathbb{P}}(\mathbf{X})$.
So, $f - \boldsymbol{\pi}^{\boldsymbol{\Xi}}(f) = p + g - p - \boldsymbol{\pi}^{\boldsymbol{\Xi}}(g) = g - \boldsymbol{\pi}^{\boldsymbol{\Xi}}(g)$, and diagram (3.5) gives

$$g - \boldsymbol{\pi}^{\boldsymbol{\Xi}}(g) \in \mathcal{F}_{K^{\boldsymbol{\Xi}}}(\mathbf{X}),$$

hence (3.7).
Now (3.7) implies

$$\mathbb{P} + \mathcal{F}_{K,\mathbb{P}}(\mathbf{X}) \subset \mathbb{P} + \mathcal{F}_{K^{\boldsymbol{\Xi}}}(\mathbf{X}).$$

Moreover, since $F_{K^{\boldsymbol{\Xi}}}$ is an onto mapping from $\mathcal{M}(\mathbf{X})$ on $\mathcal{F}_{K^{\boldsymbol{\Xi}}}(\mathbf{X})$, so is the mapping $\mathrm{Id} - \boldsymbol{\pi}^{\boldsymbol{\Xi}}$ between $\mathcal{F}_{K,\mathbb{P}}(\mathbf{X})$, hence

$$\mathbb{P} + \mathcal{F}_{K^{\boldsymbol{\Xi}}}(\mathbf{X}) \subset \mathbb{P} + \mathcal{F}_{K,\mathbb{P}}(\mathbf{X}).$$

Thus (3.8) holds.

3.4 $\mathbb{P}$-conditionally positive definite kernel

3.4.1 $\mathbb{P}$-conditionally positive definite kernel

We know from Proposition 3.3 that, K being $\mathbb{P}$-conditionally positive, $<,>_{\mathcal{F}_K}$ is a positive symmetric bilinear form on $\mathcal{F}_{K,\mathbb{P}}$.
Here is a characterization of couples $(K, \mathbb{P})$ which leads to the positive definiteness of $<,>_{\mathcal{F}_K}$ on $\mathcal{F}_{K,\mathbb{P}}$.

Proposition 3.7. *For any $f \in \mathcal{F}_{K,\mathbb{P}}$,*

$$< f, f >_{\mathcal{F}_K} = 0 \Leftrightarrow f \in \mathbb{P}.$$

Hence $<,>_{\mathcal{F}_K}$ is positive definite on $\mathcal{F}_{K,\mathbb{P}}$ if and only if $\mathbb{P} \cap \mathcal{F}_{K,\mathbb{P}} = \{0\}$.

Proof
Let us first set this well known property:

Lemma 3.3. *If R is a positive kernel, then $<,>_{\mathcal{F}_R}$ is positive definite.*

Proof
Let $g \in \mathcal{F}_R$.
Reproducing property (3.3) and Cauchy-Schwarz inequality leads to:

$$|g(\mathbf{x})| = | < R_{\mathbf{x}}, g >_{\mathcal{F}_R} | \leq \sqrt{< g, g >_{\mathcal{F}_R}} \sqrt{< R_{\mathbf{x}}, R_{\mathbf{x}} >_{\mathcal{F}_R}}.$$

Hence, $< g, g >_{\mathcal{F}_R} = 0 \Rightarrow \forall \mathbf{x} \in E, g(\mathbf{x}) = 0 \Rightarrow g = 0$.
□
Now, let $f \in \mathcal{F}_{K,\mathbb{P}}$ and $\boldsymbol{\lambda} \in \mathcal{M}_{\mathbb{P}}$ be such that $f = F_K(\boldsymbol{\lambda})$.
From (3.4), we get:

$$< f, f >_{\mathcal{F}_{K,\mathbb{P}}} = < \boldsymbol{\lambda}, \boldsymbol{\lambda} >_{\mathcal{M},K} = \boldsymbol{\lambda}(F_K(\boldsymbol{\lambda})). \tag{3.9}$$

Since $\boldsymbol{\lambda} \in \mathcal{M}_{\mathbb{P}}$, it follows that $\Phi^{\boldsymbol{\Xi}}(\boldsymbol{\lambda}) = \boldsymbol{\lambda}$ and, then, diagram (3.5) implies

$$F_K(\boldsymbol{\lambda}) = \boldsymbol{\pi}^{\boldsymbol{\Xi}}(F_K(\boldsymbol{\lambda})) + F_{K^{\boldsymbol{\Xi}}}(\boldsymbol{\lambda}). \tag{3.10}$$

Applying $\boldsymbol{\lambda}$ to both terms of (3.10), leads to

$$\boldsymbol{\lambda}(F_K(\boldsymbol{\lambda})) = \boldsymbol{\lambda}(F_{K^{\boldsymbol{\Xi}}}(\boldsymbol{\lambda}))\,,$$

since $\boldsymbol{\lambda} \in \mathcal{M}_{\mathbb{P}}$ implies that $\boldsymbol{\lambda}(\boldsymbol{\pi}^{\boldsymbol{\Xi}}(F_K(\boldsymbol{\lambda}))) = 0$.
Equality (3.9) then becomes

$$< f, f >_{\mathcal{F}_{K,\mathbb{P}}} = \boldsymbol{\lambda}(F_{K^{\boldsymbol{\Xi}}}(\boldsymbol{\lambda})) =< F_{K^{\boldsymbol{\Xi}}}(\boldsymbol{\lambda}), F_{K^{\boldsymbol{\Xi}}}(\boldsymbol{\lambda}) >_{\mathcal{F}_{K^{\boldsymbol{\Xi}}}}\ .$$

Hence

$$< f, f >_{\mathcal{F}_{K,\mathbb{P}}} = 0 \Leftrightarrow < F_{K^{\boldsymbol{\Xi}}}(\boldsymbol{\lambda}), F_{K^{\boldsymbol{\Xi}}}(\boldsymbol{\lambda}) >_{\mathcal{F}_{K^{\boldsymbol{\Xi}}}} = 0\,.$$

which, with Lemma 3.3 applied to $K^{\boldsymbol{\Xi}}$, leads to

$$< f, f >_{\mathcal{F}_{K,\mathbb{P}}} = 0 \Leftrightarrow F_{K^{\boldsymbol{\Xi}}}(\boldsymbol{\lambda}) = 0\,.$$

Eventually, from (3.10)

$$< f, f >_{\mathcal{F}_{K,\mathbb{P}}} = 0 \Leftrightarrow f = \boldsymbol{\pi}^{\boldsymbol{\Xi}}(f) \Leftrightarrow f \in \mathbb{P}\,.$$

□
We are naturally led to the following definition:

Definition 3.3 ($\mathbb{P}$-conditionally positive definite kernel)**.** *A $\mathbb{P}$-conditionally positive kernel K is said $\mathbb{P}$-conditionally positive* **definite** *if*

$$\mathbb{P} \cap \mathcal{F}_{K,\mathbb{P}} = \{0\}.$$

In other words: K is $\mathbb{P}$-conditionally positive definite if and only if $<,>_{\mathcal{F}_K}$ is a positive definite symmetric bilinear form on $\mathcal{F}_{K,\mathbb{P}}$.

Here are three particular cases where $\mathbb{P} \cap \mathcal{F}_{K,\mathbb{P}} = \{0\}$ and consequently where K is $\mathbb{P}$-conditionally positive definite.

1. $\mathbb{P} = \{0\}$.
   It is the classical case of positive definite kernel. There is no differences between positive kernel and positive definite kernel.

2. More generally, whatever $\mathbb{P}$ is, if K is positive then it is $\mathbb{P}$-conditionally positive definite. Indeed, let f be in $\mathbb{P} \cap \mathcal{F}_{K,\mathbb{P}}$. Since $f \in \mathcal{F}_{K,\mathbb{P}}$, there exists $\boldsymbol{\lambda} \in \mathcal{M}_{\mathbb{P}}$ such that $f = F_K(\boldsymbol{\lambda})$. We have, using (3.4)

   $$< f, f >_{\mathcal{F}_K} =< \boldsymbol{\lambda}, \boldsymbol{\lambda} >_{\mathcal{M},K} = \boldsymbol{\lambda}(F_K(\boldsymbol{\lambda})) = \boldsymbol{\lambda}(f) = 0\,,$$

   since $f \in \mathbb{P}$.
   But, K positive implies that $<,>_{\mathcal{F}_K}$ is positive definite (see Lemma 3.3), hence

   $$< f, f >_{\mathcal{F}_K} = 0 \Rightarrow f = 0\,.$$

3. The following condition is the $\mathbb{P}$-conditionally positive definite kernel definition given by Wendland (2005): for all $L \in \mathbb{N}$, and every $\mathbf{x}_1, \ldots, \mathbf{x}_L$ **pairwise distinct**

$$\forall (\lambda_1, \ldots, \lambda_L) \in \mathbb{R}^L \begin{cases} \sum_{1 \leq k,l \leq L} \lambda_l \lambda_k K(\mathbf{x}_l, \mathbf{x}_k) = 0 \\ \text{et} \\ \sum_{l=1}^{L} \lambda_l p(\mathbf{x}_l) = 0, \forall p \in \mathbb{P} \end{cases} \Rightarrow \lambda_l = 0, l = 1, \ldots, L. \quad (3.11)$$

Indeed, suppose $K, \mathbb{P}$ are satisfying (3.11) and let f be in $\mathbb{P} \cap \mathcal{F}_{K,\mathbb{P}}$.
On the one hand $f \in \mathcal{F}_{K,\mathbb{P}}$. Hence there exists $\boldsymbol{\mu} \in \mathcal{M}_{\mathbb{P}}$ such that $f = F_K(\boldsymbol{\mu})$.
On the other hand $f \in \mathbb{P}$, thus $\boldsymbol{\mu}(f) = 0$.
Combining these two facts we get

$$\boldsymbol{\mu}(F_K(\boldsymbol{\mu})) = 0. \quad (3.12)$$

Let us now write $\boldsymbol{\mu} = \sum_{m=1}^{M} \mu_k \delta_{\mathbf{x}_k}$, with $\{\mathbf{x}_1, \ldots, \mathbf{x}_M\}$ pairwise distinct. Relation (3.12) becomes:

$$\sum_{m,l} \mu_l \mu_m K(\mathbf{x}_l, \mathbf{x}_m) = 0\,.$$

Since $\boldsymbol{\mu} \in \mathcal{M}_{\mathbb{P}}$, it follows from (3.11) that $\mu_k = 0, k = 1, \ldots, M$, then $\boldsymbol{\mu} = 0$ and eventually $f = 0$.

Let us notice that condition (3.11) cannot be satisfied if $\mathcal{F}_{K,\mathbb{P}}$ is a finite-dimensional vector space, and E infinite.
Indeed, suppose $\mathcal{F}_{K,\mathbb{P}}$ is a finite-dimensional vector space.
Let $\boldsymbol{\Xi} = \{\boldsymbol{\xi}_1, \ldots, \boldsymbol{\xi}_n\}$ be a minimal $\mathbb{P}$-unisolvent set. There exists $\mathbf{x}_1, \ldots, \mathbf{x}_L \in E - \boldsymbol{\Xi}$ pairwise distinct such that $F_K(\delta^{\boldsymbol{\Xi}}_{\mathbf{x}_1}), \ldots, F_K(\delta^{\boldsymbol{\Xi}}_{\mathbf{x}_L})$, which all are in $\mathcal{F}_{K,\mathbb{P}}$, are linearly dependent:

$$\exists (\lambda_1, \ldots, \lambda_L) \neq 0 \in \mathbb{R}^L \text{such that } \sum_{l=1}^{L} \lambda_l F_K(\delta^{\boldsymbol{\Xi}}_{\mathbf{x}_l}) = 0. \quad (3.13)$$

Hence

$$0 =< F_K(\sum_{l=1}^{L} \lambda_l \delta^{\boldsymbol{\Xi}}_{\mathbf{x}_l}), F_K(\sum_{l=1}^{L} \lambda_l \delta^{\boldsymbol{\Xi}}_{\mathbf{x}_l}) >_{\mathcal{F}_K} =< \sum_{l=1}^{L} \lambda_l \delta^{\boldsymbol{\Xi}}_{\mathbf{x}_l}, \sum_{l=1}^{L} \lambda_l \delta^{\boldsymbol{\Xi}}_{\mathbf{x}_l} >_{\mathcal{M},K}\,. \quad (3.14)$$

Since $\sum_{l=1}^{L} \lambda_l \delta^{\boldsymbol{\Xi}}_{\mathbf{x}_l} = \sum_{l=1}^{L} \lambda_l \delta_{\mathbf{x}_l} - \sum_{k=1}^{n} \left[\sum_{l=1}^{L} \lambda_l h^{\boldsymbol{\Xi}}_k(\mathbf{x}_l)\right] \delta_{\boldsymbol{\xi}_k}$ we can write

$$\sum_{l=1}^{L} \lambda_l \delta^{\boldsymbol{\Xi}}_{\mathbf{x}_l} = \sum_{l=1}^{L+n} \lambda_l \delta_{\mathbf{x}_l}\,,$$

where $\lambda_{L+k} = -\sum_{l=1}^{L} \lambda_l h^{\boldsymbol{\Xi}}_k(\mathbf{x}_l)$ and $\mathbf{x}_{L+k} = \boldsymbol{\xi}_k,\ k = 1, \ldots n$. And (3.14) becomes:

$$0 =< \sum_{l=1}^{L+n} \lambda_l \delta_{\mathbf{x}_l}, \sum_{l=1}^{L+n} \lambda_l \delta_{\mathbf{x}_l} >_{\mathcal{M},K} = \sum_{1 \leq i,j \leq L+n} \lambda_i \lambda_j K(\mathbf{x}_i, \mathbf{x}_j)\,.$$

If condition (3.11) were satisfied, recalling $\sum_{l=1}^{L+n} \lambda_l \delta_{\mathbf{x}_l} \in \mathcal{M}_{\mathbb{P}}$, this last equality would imply: $\lambda_i = 0, i = 1, \ldots, L+n$, which conflicts with (3.13).

3.4.2 $\mathbb{P}$-Reproducing Kernel Semi-Hilbert Space

Here is the main result of our study:

Theorem 3.2. *Assume K is a $\mathbb{P}$-conditionally positive definite kernel.*
There is a unique semi-Hilbert space of real functions defined on E, $\left(\mathcal{H}_{K,\mathbb{P}}, <,>_{\mathcal{H}_{K,\mathbb{P}}}\right)$ such that

1. *$\left(\mathcal{F}_{K,\mathbb{P}}, <,>_{\mathcal{F}_{K,\mathbb{P}}}\right)$ is a pre-hilbertian subspace of $\left(\mathcal{H}_{K,\mathbb{P}}, <,>_{\mathcal{H}_{K,\mathbb{P}}}\right)$,*
2. *$\mathbb{P} \subset \mathcal{H}_{K,\mathbb{P}}$ is the null space of $<,>_{\mathcal{H}_{K,\mathbb{P}}}$,*
3. *for all $\boldsymbol{\Xi}$, minimal $\mathbb{P}$-unisolvent set, the following reproducing property is satisfied:*

$$\forall f \in \mathcal{H}_{K,\mathbb{P}},\ \mathbf{x} \in E,\ f(\mathbf{x}) = \boldsymbol{\pi}^{\boldsymbol{\Xi}}(f)(\mathbf{x}) + < f, F_K(\delta_{\mathbf{x}}^{\boldsymbol{\Xi}}) >_{\mathcal{H}_{K,\mathbb{P}}} . \tag{3.15}$$

We call $\left(\mathcal{H}_{K,\mathbb{P}}, <,>_{\mathcal{H}_{K,\mathbb{P}}}\right)$ the $\mathbb{P}$-reproducing kernel semi-Hilbert space ($\mathbb{P}$-RKSHS) associated with $(K, \mathbb{P})$.

By a semi-Hilbert space, we mean:

Definition 3.4. *A vector space $\mathcal{L}$ equipped with a symmetric positive bilinear form $<,>_{\mathcal{L}}$ is semi-hilbertian if, $\mathcal{K}$ being the null subspace[1] of $(\mathcal{L}, <,>_{\mathcal{L}})$, the quotient space $\mathcal{L}/\mathcal{K}$ endowed with the bilinear form induced by $<,>_{\mathcal{L}}$ is a Hilbert space.*

As a byproduct useful result, we will also get

Proposition 3.8. *Any choice of a minimal $\mathbb{P}$-unisolvent set $\boldsymbol{\Xi}$, leads to the direct sum decomposition:*

$$\mathcal{H}_{K,\mathbb{P}} = \mathbb{P} \bigoplus \mathcal{H}_{K^{\boldsymbol{\Xi}}} ,$$

with $\boldsymbol{\pi}^{\boldsymbol{\Xi}}$ and $\left(\mathrm{Id}_{\mathcal{H}_{K,\mathbb{P}}} - \boldsymbol{\pi}^{\boldsymbol{\Xi}}\right)$ as associated projectors.
Moreover,

$$< f, g >_{\mathcal{H}_{K,\mathbb{P}}} = < f - \boldsymbol{\pi}^{\boldsymbol{\Xi}}(f), g - \boldsymbol{\pi}^{\boldsymbol{\Xi}}(g) >_{\mathcal{H}_{K^{\boldsymbol{\Xi}}}} .$$

Remark 3.2. *Since $< f, F_K(\delta_{\mathbf{x}}^{\boldsymbol{\Xi}}) >_{\mathcal{H}_{K,\mathbb{P}}} = < f, F_K(\delta_{\mathbf{x}}^{\boldsymbol{\Xi}}) - \boldsymbol{\pi}^{\boldsymbol{\Xi}}(F_K(\delta_{\mathbf{x}}^{\boldsymbol{\Xi}})) >_{\mathcal{H}_{K,\mathbb{P}}} = < f, K_{\mathbf{x}}^{\boldsymbol{\Xi}} >_{\mathcal{H}_{K,\mathbb{P}}}$, the reproducing formula (3.15) can be written:*

$$\forall f \in \mathcal{H}_{K,\mathbb{P}},\ \mathbf{x} \in E,\ f(\mathbf{x}) = \boldsymbol{\pi}^{\boldsymbol{\Xi}}(f)(\mathbf{x}) + < f, K_{\mathbf{x}}^{\boldsymbol{\Xi}} >_{\mathcal{H}_{K,\mathbb{P}}} .$$

Positive definite case

Suppose K is positive and $\mathbb{P} = \{0\}$. Kernel K is also positive definite according to definition 3.3.
Theorem 3.2 reduces to Aronszajn's

Theorem 3.3. *There is a unique Hilbert space of real functions $(\mathcal{H}_K, <,>_{\mathcal{H}_K})$ such that:*

1. *$(\mathcal{F}_K, <,>_{\mathcal{F}_K})$ is a pre-Hilbert subspace of $(\mathcal{H}_K, <,>_{\mathcal{H}_K})$,*

[1] $\mathcal{K} = \{\mathbf{u} \in \mathcal{L} :< \mathbf{u}, \mathbf{v} >_{\mathcal{L}} = 0, \forall \mathbf{v} \in \mathcal{L}\} = \{\mathbf{u} \in \mathcal{L} :< \mathbf{u}, \mathbf{u} >_{\mathcal{L}} = 0\}$

2. the following reproducing property is satisfied

$$\forall f \in \mathcal{H}_K,\ \mathbf{x} \in E,\ f(\mathbf{x}) =< f, K_{\mathbf{x}} >_{\mathcal{H}_K} . \tag{3.16}$$

$\mathcal{H}_K$ is the reproducing kernel Hilbert space (RKHS) with reproducing kernel K.

Proof
Existence
$<,>_{\mathcal{F}_K}$ being positive definite on $\mathcal{F}_K$, let $(\mathcal{H}, <, >_{\mathcal{H}})$ be the Hilbert completion of $(\mathcal{F}_K, <, >_{\mathcal{F}_K})$.

Lemma 3.4. *The mapping*

$$R : h \in \mathcal{H} \mapsto \{\mathbf{x} \mapsto < h, K_{\mathbf{x}} >_{\mathcal{H}}\} \in \mathbb{R}^E$$

is an injection.

Proof
The set $\{K_{\mathbf{x}} : \mathbf{x} \in E\}$ is total in $\mathcal{H}$, since it spans $\mathcal{F}_K$ which is dense in $\mathcal{H}$.
Hence

$$R(h) = 0 \Leftrightarrow < h, K_{\mathbf{x}} >_{\mathcal{H}} = 0,\ \forall \mathbf{x} \in E \Leftrightarrow h = 0 .$$

□
Let $\mathcal{H}_K = R(\mathcal{H})$, be equipped with the following inner product:

$$< R(h_1), R(h_2) >_{\mathcal{H}_K} =< h_1, h_2 >_{\mathcal{H}} .$$

$(\mathcal{H}_K, <, >_{\mathcal{H}_K})$ is a Hilbert space as isomorphic image of $\mathcal{H}$.
It satisfies the required properties:

1. $R(K_{\mathbf{x}}) = K_{\mathbf{x}}$ as shown by

   $$R(K_{\mathbf{z}})(\mathbf{x}) =< K_{\mathbf{z}}, K_{\mathbf{x}} >_{\mathcal{H}} =< K_{\mathbf{z}}, K_{\mathbf{x}} >_{\mathcal{F}_K} = K(\mathbf{z}, \mathbf{x}) = K_{\mathbf{z}}(\mathbf{x}) ,$$

   implies $R(f) = f$ for any $f \in \mathcal{F}_K$.
   Hence $\mathcal{F}_K \subset \mathcal{H}_K$ which leads readily to first property.

2. Let f be any function in $\mathcal{H}_K$, and $h \in \mathcal{H}$ be such that $R(h) = f$. We have:

   $$< f, K_{\mathbf{x}} >_{\mathcal{H}_K} =< R(h), R(K_{\mathbf{x}}) >_{\mathcal{H}_K} =< h, K_{\mathbf{x}} >_{\mathcal{H}} = R(h)(\mathbf{x}) = f(\mathbf{x}) .$$

Unicity
It comes from this fact:

Lemma 3.5. *If $\mathcal{H}$ is an Hilbert space of functions satisfying the specifications of Theorem 3.3, then $\{K_{\mathbf{x}} : \mathbf{x} \in E\}$ is a total set in $\mathcal{H}$.*

Proof
Let $h \in \mathcal{H}$ be such that

$$\forall \mathbf{x} \in E,\ < h, K_{\mathbf{x}} >_{\mathcal{H}} = 0 .$$

From the reproduction property (3.16) it follows:

$$\forall \mathbf{x} \in E,\ h(\mathbf{x}) = 0 ,$$

hence $h = 0$.
□
Now let $\mathcal{H}$ and $\mathcal{H}'$ be two Hilbert spaces of real functions defined on E, satisfying Theorem 3.3 properties.
From Lemma 3.5, they both contain $(\mathcal{F}_K, <,>_{\mathcal{F}_K})$ as dense subspace.
The identity on $\mathcal{F}_K$ can be then extended as an isometry

$$I : \mathcal{H} \mapsto \mathcal{H}' .$$

Hence

$$\forall h \in \mathcal{H}, \mathbf{x} \in E, \ < h, K_{\mathbf{x}} >_{\mathcal{H}} = < I(h), K_{\mathbf{x}} >_{\mathcal{H}'}$$

or

$$\forall h \in \mathcal{H}, \mathbf{x} \in E \ h(\mathbf{x}) = I(h)(\mathbf{x}) ,$$

which means $\forall h \in \mathcal{H}, h = I(h)$
□

General case: existence

Let $\boldsymbol{\Xi}$ be a minimal $\mathbb{P}$-unisolvent set. Theorem 3.3 can be applied to $K^{\boldsymbol{\Xi}}$.
Observe that any function f of its RKHS satisfies:

$$f(\boldsymbol{\xi}) = 0, \ \forall \boldsymbol{\xi} \in \boldsymbol{\Xi} . \tag{3.17}$$

Indeed, (3.17) is true for $f = K^{\boldsymbol{\Xi}}_{\mathbf{x}}$ since

$$f(\boldsymbol{\xi}) = K^{\boldsymbol{\Xi}}_{\mathbf{x}}(\boldsymbol{\xi}) = K^{\boldsymbol{\Xi}}(\mathbf{x}, \boldsymbol{\xi}) = < \delta^{\boldsymbol{\Xi}}_{\mathbf{x}}, \delta^{\boldsymbol{\Xi}}_{\boldsymbol{\xi}} >_{\mathcal{M},K}$$

and $\delta^{\boldsymbol{\Xi}}_{\boldsymbol{\xi}} = 0$.
Hence it is true for any $f \in \mathcal{F}_{K^{\boldsymbol{\Xi}}}$.
From Lemma 3.5, $\mathcal{F}_{K^{\boldsymbol{\Xi}}}$ is dense in $\mathcal{H}_{K^{\boldsymbol{\Xi}}}$. So any $f \in \mathcal{H}_{K^{\boldsymbol{\Xi}}}$ can be written as limit of a sequence $(f_k)_{k\in\mathbb{N}}$ of functions of $\mathcal{F}_{K^{\boldsymbol{\Xi}}}$.
Then for any $\boldsymbol{\xi} \in \boldsymbol{\Xi}$, we have:

$$f(\boldsymbol{\xi}) = < f, K^{\boldsymbol{\Xi}}_{\boldsymbol{\xi}} >_{\mathcal{H}_{K^{\boldsymbol{\Xi}}}} = \lim_{k\to\infty} < f_k, K^{\boldsymbol{\Xi}}_{\boldsymbol{\xi}} >_{\mathcal{H}_{K^{\boldsymbol{\Xi}}}} = \lim_{k\to\infty} f_k(\boldsymbol{\xi}) = 0.$$

Hence (3.17) follows.
An immediate consequence of (3.17) is:

Proposition 3.9. *The sum $\mathcal{N} = \mathbb{P} + \mathcal{H}_{K^{\boldsymbol{\Xi}}}$ is direct.*
Moreover, $\boldsymbol{\pi}^{\boldsymbol{\Xi}}$ and $\mathrm{Id} - \boldsymbol{\pi}^{\boldsymbol{\Xi}}$ restricted to $\mathcal{N}$ are the associated projections of this direct sum decomposition.

Moreover:

Proposition 3.10 (existence)**.** *$\mathcal{N} = \mathbb{P} \oplus \mathcal{H}_{K^{\boldsymbol{\Xi}}}$ with the following bilinear form*

$$<,>_{\mathcal{N}} : (p_1 + h_1, p_2 + h_2) \in [\mathbb{P} \oplus \mathcal{H}_{K^{\boldsymbol{\Xi}}}]^2 \mapsto < h_1, h_2 >_{\mathcal{H}_{K^{\boldsymbol{\Xi}}}}$$

satisfies the properties required by Theorem 3.2.

Proof
Equipped with the form induced by $<,>_{\mathcal{N}}$, $\mathcal{N}/\mathbb{P}$ is obviously isomorphic to $\left(\mathcal{H}_{K^{\Xi}}, <,>_{\mathcal{H}_{K^{\Xi}}}\right)$: $(\mathcal{N}, <,>_{\mathcal{N}})$ is semi-hilbertian and its null space is $\mathbb{P}$.
From (3.8) we know that

$$\mathbb{P} + \mathcal{F}_{K,\mathbb{P}}(\mathbf{X}) = \mathbb{P} + \mathcal{F}_{K^{\Xi}}(\mathbf{X})$$

hence $\mathcal{F}_{K,\mathbb{P}}(\mathbf{X}) \subset \mathbb{P} + \mathcal{F}_{K^{\Xi}}(\mathbf{X}) \subset \mathcal{N}$.
From diagram (3.5), it comes

$$\forall f, g \in \mathcal{F}_{K,\mathbb{P}},\ < f, g >_{\mathcal{F}_{K,\mathbb{P}}} = < (\mathrm{Id} - \boldsymbol{\pi}^{\Xi})(f), (\mathrm{Id} - \boldsymbol{\pi}^{\Xi})(g) >_{\mathcal{F}_{K^{\Xi}}} = < f, g >_{\mathcal{N}} \ .$$

Hence $(\mathcal{F}_{K,\mathbb{P}}, <,>_{\mathcal{F}_{K,\mathbb{P}}})$ is a pre-hilbertian subspace of $\mathcal{N}$.
Let us now prove reproducing formula (3.15).
Let f be in $\mathcal{N}$ and $\Xi' = \{\boldsymbol{\xi}'_1, \ldots, \boldsymbol{\xi}'_n\}$ be a minimal $\mathbb{P}$-unisolvent set.
Observe first that:

$$< f, F_K(\delta_{\mathbf{x}}^{\Xi'}) >_{\mathcal{N}} = < f - \boldsymbol{\pi}^{\Xi}(f), F_K(\delta_{\mathbf{x}}^{\Xi'}) - \boldsymbol{\pi}^{\Xi}(F_K(\delta_{\mathbf{x}}^{\Xi'})) >_{\mathcal{H}_{K^{\Xi}}} \ .$$

From diagram (3.5), we get

$$F_K(\delta_{\mathbf{x}}^{\Xi'}) - \boldsymbol{\pi}^{\Xi}(F_K(\delta_{\mathbf{x}}^{\Xi'})) = (\mathrm{Id} - \boldsymbol{\pi}^{\Xi})(F_K(\delta_{\mathbf{x}}^{\Xi'})) = F_{K^{\Xi}}(\delta_{\mathbf{x}}^{\Xi'})$$

and, since $\delta_{\mathbf{x}}^{\Xi'} = \delta_{\mathbf{x}} - \sum_{k=1}^{n} h_k^{\Xi'}(\mathbf{x})\delta_{\boldsymbol{\xi}'_k}$:

$$F_K(\delta_{\mathbf{x}}^{\Xi'}) - \boldsymbol{\pi}^{\Xi}(F_K(\delta_{\mathbf{x}}^{\Xi'})) = K_{\mathbf{x}}^{\Xi} - \sum_{k=1}^{n} h_k^{\Xi'}(\mathbf{x}) K_{\boldsymbol{\xi}'_k}^{\Xi} \ .$$

Hence, applying twice reproducing formula in $\mathcal{H}_{K^{\Xi}}$,

$$\begin{aligned}
< f, F_K(\delta_{\mathbf{x}}^{\Xi'}) >_{\mathcal{N}} &= < f - \boldsymbol{\pi}^{\Xi}(f), K_{\mathbf{x}}^{\Xi} - \sum_{k=1}^{n} h_k^{\Xi'}(\mathbf{x}) K_{\boldsymbol{\xi}'_k}^{\Xi} >_{\mathcal{H}_{K^{\Xi}}} \\
&= f(\mathbf{x}) - \boldsymbol{\pi}^{\Xi}(f)(\mathbf{x}) - \sum_{k=1}^{n} h_k^{\Xi'}(\mathbf{x})(f(\boldsymbol{\xi}'_k) - \boldsymbol{\pi}^{\Xi}(f)(\boldsymbol{\xi}'_k)) \\
&= f(\mathbf{x}) - \boldsymbol{\pi}^{\Xi}(f)(\mathbf{x}) - \boldsymbol{\pi}^{\Xi'}\left[f - \boldsymbol{\pi}^{\Xi}(f)\right](\mathbf{x}) \\
&= f(\mathbf{x}) - \boldsymbol{\pi}^{\Xi}(f)(\mathbf{x}) - \boldsymbol{\pi}^{\Xi'}(f)(\mathbf{x}) + \boldsymbol{\pi}^{\Xi'}(\boldsymbol{\pi}^{\Xi}(f)(\mathbf{x}))
\end{aligned}$$

and eventually, as $\boldsymbol{\pi}^{\Xi'} \circ \boldsymbol{\pi}^{\Xi} = \boldsymbol{\pi}^{\Xi}$,

$$< f, F_K(\delta_{\mathbf{x}}^{\Xi'}) >_{\mathcal{N}} = f(\mathbf{x}) - \boldsymbol{\pi}^{\Xi'}(f)(\mathbf{x}) \,,$$

which is the reproducing formula (3.15).
□

General case: unicity

Lemma 3.6. *Let $\mathcal{N} \subset \mathbb{R}^E$ be satisfying properties of Theorem 3.2 and $\boldsymbol{\Xi} \subset E$ be a minimal $\mathbb{P}$-unisolvent set.*
Let us set $\overline{f}$ for the modulo $\mathbb{P}$ class of any $f \in \mathcal{N}$.
$\{\overline{K_{\mathbf{x}}^{\boldsymbol{\Xi}}} : \mathbf{x} \in E\}$ is a total set in the Hilbert space $\mathcal{N}/\mathbb{P}$.

Proof
Let $h \in \mathcal{N}$ be such that $\forall \mathbf{x} \in E,\ < \overline{h}, \overline{K_{\mathbf{x}}^{\boldsymbol{\Xi}}} >_{\mathcal{N}/\mathbb{P}} = 0$.
As $< \overline{h}, \overline{K_{\mathbf{x}}^{\boldsymbol{\Xi}}} >_{\mathcal{N}/\mathbb{P}} = < h, K_{\mathbf{x}}^{\boldsymbol{\Xi}} >_{\mathcal{N}}$, h satisfies

$$\forall \mathbf{x} \in E, < h, K_{\mathbf{x}}^{\boldsymbol{\Xi}} >_{\mathcal{N}} = 0\,.$$

From reproducing property (3.15), we get, $\mathbf{x} \in E$:

$$h(\mathbf{x}) = \boldsymbol{\pi}^{\boldsymbol{\Xi}}(h)(\mathbf{x}) + < h, K_{\mathbf{x}}^{\boldsymbol{\Xi}} >_{\mathcal{N}} = \boldsymbol{\pi}^{\boldsymbol{\Xi}}(h)(\mathbf{x})$$

That is

$$h \in \mathbb{P} \text{ thus } \overline{h} = 0\,.$$

□
Suppose now that two spaces $\mathcal{N}, \mathcal{N}'$ satisfy theorem 3.2 specifications.
Let $\boldsymbol{\Xi}$ be a $\mathbb{P}$-unisolvent minimal set. From Lemma 3.6, it follows that both of $\mathcal{N}/\mathbb{P}$ and $\mathcal{N}'/\mathbb{P}$ contain $\mathcal{F}_{K^{\boldsymbol{\Xi}}}/\mathbb{P}$ as a dense subspace.
Hence identity function on $\mathcal{F}_{K^{\boldsymbol{\Xi}}}/\mathbb{P}$ can be extended by an isometry

$$I : \mathcal{N}/\mathbb{P} \mapsto \mathcal{N}'/\mathbb{P}\,.$$

Thus, for any $\mathbf{x} \in E$ and $h \in \mathcal{N}$, applying again reproducing formula (3.15):

$$\begin{aligned} h(\mathbf{x}) &= \boldsymbol{\pi}^{\boldsymbol{\Xi}}(h)(\mathbf{x}) + < h, K_{\mathbf{x}}^{\boldsymbol{\Xi}} >_{\mathcal{N}} \\ &= \boldsymbol{\pi}^{\boldsymbol{\Xi}}(h)(\mathbf{x}) + < \overline{h}, \overline{K_{\mathbf{x}}^{\boldsymbol{\Xi}}} >_{\mathcal{N}/\mathbb{P}} \\ &= \boldsymbol{\pi}^{\boldsymbol{\Xi}}(h)(\mathbf{x}) + < I(\overline{h}), I(\overline{K_{\mathbf{x}}^{\boldsymbol{\Xi}}}) >_{\mathcal{N}'/\mathbb{P}} \\ &= \boldsymbol{\pi}^{\boldsymbol{\Xi}}(h)(\mathbf{x}) + < h', K_{\mathbf{x}}^{\boldsymbol{\Xi}} >_{\mathcal{N}'} \\ &= \boldsymbol{\pi}^{\boldsymbol{\Xi}}(h)(\mathbf{x}) - \boldsymbol{\pi}^{\boldsymbol{\Xi}}(h')(\mathbf{x}) + h'(\mathbf{x})\,, \end{aligned}$$

where $h' \in \mathcal{N}'$ is a class representant of $I(\overline{h})$.
So $h \in \mathcal{N}'$. □

3.5 Interpolation in RKSHS

3.5.1 Preliminaries

In this section, we assume that

- $\mathbb{P}$ a finite dimensional vector space of functions;
- K is a $\mathbb{P}$-conditionally positive definite kernel;

- $\mathbf{X} = \{\mathbf{x}_1, \ldots, \mathbf{x}_N\} \subset E$ is a $\mathbb{P}$-unisolvent set.

If we are given a minimal $\mathbb{P}$-unisolvent set $\boldsymbol{\Xi}$, we know that $\mathcal{F}_{K^{\boldsymbol{\Xi}}}(\mathbf{X})$ is a (finite dimensional) vector subspace of Hilbert space $\mathcal{H}_{K^{\boldsymbol{\Xi}}}$. By $[\mathcal{F}_{K^{\boldsymbol{\Xi}}}(\mathbf{X})]^{\perp}$ is denoted the orthogonal complement of $\mathcal{F}_{K^{\boldsymbol{\Xi}}}(\mathbf{X})$ in $\mathcal{H}_{K^{\boldsymbol{\Xi}}}$.

3.5.2 Characterizations of interpolation in RKSHS

Let f be a function in $\mathcal{H}_{K,\mathbb{P}}$ that we only know on $\mathbf{X}$. We want to interpolate f in a reasonable way just using $f(\mathbf{x}_1), \ldots, f(\mathbf{x}_N)$ and K.
We start with a geometrical characterization of interpolation in $\mathcal{H}_{K,\mathbb{P}}$.

Proposition 3.11. *Let $\boldsymbol{\Xi} \subset \mathbf{X}$ be a minimal $\mathbb{P}$-unisolvent set.*
For every $f, g \in \mathcal{H}_{K,\mathbb{P}}$,

$$g \text{ interpolates } f \text{ on } \mathbf{X} \text{ if and only if } f - g \in [\mathcal{F}_{K^{\boldsymbol{\Xi}}}(\mathbf{X})]^{\perp}.$$

Proof
Applying reproducing property (3.15) to $f - g$, we get, for any $\mathbf{x} \in E$,

$$\begin{aligned} f(\mathbf{x}) - g(\mathbf{x}) &= \boldsymbol{\pi}^{\boldsymbol{\Xi}}(f-g)(\mathbf{x}) + < f - g, F_K(\delta_{\mathbf{x}}^{\boldsymbol{\Xi}}) >_{\mathcal{H}_{K,\mathbb{P}}} \\ &= \boldsymbol{\pi}^{\boldsymbol{\Xi}}(f-g)(\mathbf{x}) + < f - g - \boldsymbol{\pi}^{\boldsymbol{\Xi}}(f-g), F_K(\delta_{\mathbf{x}}^{\boldsymbol{\Xi}}) - \boldsymbol{\pi}^{\boldsymbol{\Xi}}(F_K(\delta_{\mathbf{x}}^{\boldsymbol{\Xi}})) >_{\mathcal{H}_{K^{\boldsymbol{\Xi}}}} . \end{aligned}$$

From diagram (3.5) comes:

$$F_K(\delta_{\mathbf{x}}^{\boldsymbol{\Xi}}) - \boldsymbol{\pi}^{\boldsymbol{\Xi}}(F_K(\delta_{\mathbf{x}}^{\boldsymbol{\Xi}})) = K_{\mathbf{x}}^{\boldsymbol{\Xi}} .$$

Hence

$$f(\mathbf{x}) - g(\mathbf{x}) = \boldsymbol{\pi}^{\boldsymbol{\Xi}}(f-g)(\mathbf{x}) + < f - g - \boldsymbol{\pi}^{\boldsymbol{\Xi}}(f-g), K_{\mathbf{x}}^{\boldsymbol{\Xi}} >_{\mathcal{H}_{K^{\boldsymbol{\Xi}}}} . \tag{3.18}$$

Suppose that g interpolates f on $\mathbf{X}$: $\forall \mathbf{x} \in \mathbf{X}, f(\mathbf{x}) = g(\mathbf{x})$.
Then, specifically,

$$\forall \boldsymbol{\xi} \in \boldsymbol{\Xi},\ f(\boldsymbol{\xi}) = g(\boldsymbol{\xi}) ,$$

which means

$$\boldsymbol{\pi}^{\boldsymbol{\Xi}}(f) = \boldsymbol{\pi}^{\boldsymbol{\Xi}}(g) ,$$

and implies:

- $f - g \in \mathcal{H}_{K^{\boldsymbol{\Xi}}}$, from Proposition 3.8,
- $0 =< f - g, K_{\mathbf{x}}^{\boldsymbol{\Xi}} >_{\mathcal{H}_{K^{\boldsymbol{\Xi}}}}, \forall \mathbf{x} \in \mathbf{X}$ from (3.18).

Hence, $f - g \in [\mathcal{F}_{K^{\boldsymbol{\Xi}}}(\mathbf{X})]^{\perp}$.
Conversely, if $f - g \in [\mathcal{F}_{K^{\boldsymbol{\Xi}}}(\mathbf{X})]^{\perp}$, then

$$\forall \mathbf{x} \in \mathbf{X},\ < f - g, K_{\mathbf{x}}^{\boldsymbol{\Xi}} >_{\mathcal{H}_{K^{\boldsymbol{\Xi}}}} = 0 ,$$

which reads, by reproducing property in $\mathcal{H}_{K^{\boldsymbol{\Xi}}}$,

$$\forall \mathbf{x} \in \mathbf{X},\ f(\mathbf{x}) - g(\mathbf{x}) = 0 .$$

□
From that proposition we draw this useful property

Corollary 3.1. *Any function f in $\mathbb{P}+\mathcal{F}_{K,\mathbb{P}}(\mathbf{X})$ is uniquely defined by its value on $\mathbf{X}$.*

Proof
Suppose that $f, g \in \mathbb{P}+\mathcal{F}_{K,\mathbb{P}}(\mathbf{X})$ coincide on $\mathbf{X}$.
From Proposition 3.11 we know that $f-g \in [\mathcal{F}_{K^{\boldsymbol{\Xi}}}(\mathbf{X})]^{\perp}$.
And, according to (3.7) applied to $f-g$, we have $f-g \in \mathcal{F}_{K^{\boldsymbol{\Xi}}}(\mathbf{X})$.
Hence $f=g$.
□
We now state the main result about interpolation: among all the interpolators lying in $\mathcal{H}_{K,\mathbb{P}}$ of any function $f \in \mathcal{H}_{K,\mathbb{P}}$ on $\mathbf{X}$, the *best* one belongs to $\mathbb{P}+\mathcal{F}_{K,\mathbb{P}}(\mathbf{X})$. That comes out from:

Proposition 3.12. *Let f be in $\mathcal{H}_{K,\mathbb{P}}$.*
If $\mathbf{X}$ is $\mathbb{P}$-unisolvent,

1. *The following problem*
$$\min_{g \in \mathbb{P}+\mathcal{F}_{K,\mathbb{P}}(\mathbf{X})} ||f-g||_{\mathcal{H}_{K,\mathbb{P}}} \tag{3.19}$$
has a unique solution which interpolates f on $\mathbf{X}$. Let $S_{K,\mathbb{P},\mathbf{X}}(f)$ denote this interpolator.

2. *Given a minimal $\mathbb{P}$-unisolvent set $\boldsymbol{\Xi} \subset \mathbf{X}$,*
$$S_{K,\mathbb{P},\mathbf{X}}(f) = \boldsymbol{\pi}^{\boldsymbol{\Xi}}(f) + S_{K^{\boldsymbol{\Xi}},\mathbf{X}}(f-\boldsymbol{\pi}^{\boldsymbol{\Xi}}(f))\,, \tag{3.20}$$
where $S_{K^{\boldsymbol{\Xi}},\mathbf{X}} : \mathcal{H}_{K^{\boldsymbol{\Xi}}} \mapsto \mathcal{F}_{K^{\boldsymbol{\Xi}}}(\mathbf{X})$ denotes the orthogonal projector on $\mathcal{F}_{K^{\boldsymbol{\Xi}}}(\mathbf{X})$.

3. *$S_{K,\mathbb{P},\mathbf{X}}(f)$ is the interpolator of f on $\mathbf{X}$ with minimal semi-norm.*

Proof
Let $\boldsymbol{\Xi}$ be any $\mathbb{P}$-unisolvent set and g be defined as
$$g = \boldsymbol{\pi}^{\boldsymbol{\Xi}}(f) + S_{K^{\boldsymbol{\Xi}},\mathbf{X}}(f-\boldsymbol{\pi}^{\boldsymbol{\Xi}}(f))\,,$$
which is meaningful since $f-\boldsymbol{\pi}^{\boldsymbol{\Xi}}(f) \in \mathcal{H}_{K^{\boldsymbol{\Xi}}}$.
We have, $S_{K^{\boldsymbol{\Xi}},\mathbf{X}}$ being the orthogonal projection on $\mathcal{F}_{K^{\boldsymbol{\Xi}}}(\mathbf{X})$:
$$f-g = f-\boldsymbol{\pi}^{\boldsymbol{\Xi}}(f) - S_{K^{\boldsymbol{\Xi}},\mathbf{X}}(f-\boldsymbol{\pi}^{\boldsymbol{\Xi}}(f)) \in [\mathcal{F}_{K^{\boldsymbol{\Xi}}}(\mathbf{X})]^{\perp}\ .$$
Hence, from Proposition 3.11 it follows that g interpolates f on $\mathbf{X}$.
Besides, by construction g lies in $\mathbb{P}+\mathcal{F}_{K^{\boldsymbol{\Xi}}}(\mathbf{X})$ and, recalling (3.8):
$$\mathbb{P}+\mathcal{F}_{K,\mathbb{P}}(\mathbf{X}) = \mathbb{P}+\mathcal{F}_{K^{\boldsymbol{\Xi}}}(\mathbf{X})\,,$$
g lies $\mathbb{P}+\mathcal{F}_{K,\mathbb{P}}(\mathbf{X})$.
Now, let us recall this easy fact, for two any functions φ_1, φ_2, belonging to $\mathcal{H}_{K,\mathbb{P}}$:
$$||\varphi_1-\varphi_2||^2_{\mathcal{H}_{K,\mathbb{P}}} = ||\varphi_1-g+g-\varphi_2||^2_{\mathcal{H}_{K,\mathbb{P}}} = ||\varphi_1-g||^2_{\mathcal{H}_{K,\mathbb{P}}} + ||g-\varphi_2||^2_{\mathcal{H}_{K,\mathbb{P}}} + 2 < \varphi_1-g, g-\varphi_2 >_{\mathcal{H}_{K,\mathbb{P}}}\ . \tag{3.21}$$
Let $h \in \mathbb{P}+\mathcal{F}_{K,\mathbb{P}}(\mathbf{X}) = \mathbb{P}+\mathcal{F}_{K^{\boldsymbol{\Xi}}}(\mathbf{X})$.
Applying (3.21) to $\varphi_1 = f$ and $\varphi_2 = h$ leads to
$$||f-h||^2_{\mathcal{H}_{K,\mathbb{P}}} = ||f-g||^2_{\mathcal{H}_{K,\mathbb{P}}} + ||g-h||^2_{\mathcal{H}_{K,\mathbb{P}}} + 2 < f-g, g-h >_{\mathcal{H}_{K,\mathbb{P}}}\ . \tag{3.22}$$

Since

$$g - h - \boldsymbol{\pi}^{\boldsymbol{\Xi}}(g-h) \in \mathcal{F}_{K^{\boldsymbol{\Xi}}}(\mathbf{X})\,,$$

and, $f - g \in [\mathcal{F}_{K^{\boldsymbol{\Xi}}}(\mathbf{X})]^{\perp}$, then

$$< f-g, g-h >_{\mathcal{H}_{K,\mathbb{P}}} = < f-g, g-h-\boldsymbol{\pi}^{\boldsymbol{\Xi}}(g-h) >_{\mathcal{H}_{K^{\boldsymbol{\Xi}}}} = 0\,.$$

Thus, relation (3.22) gives:

$$||f-h||^2_{\mathcal{H}_{K,\mathbb{P}}} = ||f-g||^2_{\mathcal{H}_{K,\mathbb{P}}} + ||g-h||^2_{\mathcal{H}_{K,\mathbb{P}}}.$$

That shows that g is a solution of problem (3.19). By corollary 3.1 there is no other interpolant of f in $\mathbb{P} + \mathcal{F}_{K,\mathbb{P}}(\mathbf{X})$.
Let now $h \in \mathcal{H}_{K,\mathbb{P}}$ be an other interpolator of f on $\mathbf{X}$. Let us apply (3.21) to $\varphi_1 = h$, and $\varphi_2 = 0$:

$$||h||^2_{\mathcal{H}_{K,\mathbb{P}}} = ||h-g||^2_{\mathcal{H}_{K,\mathbb{P}}} + ||g||^2_{\mathcal{H}_{K,\mathbb{P}}} + 2 < h-g, g >_{\mathcal{H}_{K,\mathbb{P}}}\,. \tag{3.23}$$

Since h interpolates f on $\mathbf{X}$, it also interpolates g on $\mathbf{X}$. Proposition 3.11 tells us that

$$h - g \in [\mathcal{F}_{K^{\boldsymbol{\Xi}}}(\mathbf{X})]^{\perp}\,.$$

Hence, since $g - \boldsymbol{\pi}^{\boldsymbol{\Xi}}(g) \in \mathcal{F}_{K^{\boldsymbol{\Xi}}}(\mathbf{X})$

$$< h-g, g >_{\mathcal{H}_{K,\mathbb{P}}} = < h-g, g-\boldsymbol{\pi}^{\boldsymbol{\Xi}}(g) >_{\mathcal{H}_{K^{\boldsymbol{\Xi}}}} = 0\,.$$

Relation (3.23) becomes

$$||h||^2_{\mathcal{H}_{K,\mathbb{P}}} = ||h-g||^2_{\mathcal{H}_{K,\mathbb{P}}} + ||g||^2_{\mathcal{H}_{K,\mathbb{P}}}.$$

Thus, $||h||_{\mathcal{H}_{K,\mathbb{P}}} \geq ||g||_{\mathcal{H}_{K,\mathbb{P}}}$.
Moreover, $||h||_{\mathcal{H}_{K,\mathbb{P}}} = ||g||_{\mathcal{H}_{K,\mathbb{P}}}$ only when $||h-g||_{\mathcal{H}_{K,\mathbb{P}}} = 0$. Since h interpolates g on $\mathbf{X}$, hence on $\boldsymbol{\Xi}$, we have $\boldsymbol{\pi}^{\boldsymbol{\Xi}}(h-g) = 0$ and

$$||h-g||_{\mathcal{H}_{K,\mathbb{P}}} = 0 \Leftrightarrow ||h-g-\boldsymbol{\pi}^{\boldsymbol{\Xi}}(h-g)||_{\mathcal{H}_{K^{\boldsymbol{\Xi}}}} = 0 \Leftrightarrow ||h-g||_{\mathcal{H}_{K^{\boldsymbol{\Xi}}}} = 0 \Leftrightarrow h = g\,.$$

□

3.5.3 Lagrangian form of RKSHS interpolators

We now want to set in our framework, the formulation known as Lagrangian formulation (Schaback, 2007; Wendland, 2005), which is much better for error analysis.
We first introduce a useful tool.

Free $\mathbb{P}$-unisolvent set

Definition 3.5. *Any $\mathbb{P}$-unisolvent set $\mathbf{Z}$ which does not possess a strict $\mathbb{P}$-unisolvent subset $\mathbf{Y}$ satisfying*

$$\mathbb{P} + \mathcal{F}_{K,\mathbb{P}}(\mathbf{Z}) = \mathbb{P} + \mathcal{F}_{K,\mathbb{P}}(\mathbf{Y})\,,$$

*will be called a (K-)***free** *$\mathbb{P}$-unisolvent set.*

We will state two characterizations of freeness.
The first one is:

Lemma 3.7. *A $\mathbb{P}$-unisolvent set $\mathbf{Z}$ is a free $\mathbb{P}$-unisolvent set if and only if*

$$\dim(\mathbb{P} + \mathcal{F}_{K,\mathbb{P}}(\mathbf{Z})) = \text{Cardinal}(\mathbf{Z})\,.$$

Proof
Suppose that $\mathbf{Z}$ is free.
If $\mathbf{Z}$ is a minimal $\mathbb{P}$-unisolvent set, we have $\text{Cardinal}(\mathbf{Z}) = n$ and $\mathcal{M}_{\mathbb{P}}(\mathbf{Z}) = \{0\}$. Hence $\mathbb{P} + \mathcal{F}_{K,\mathbb{P}}(\mathbf{Z}) = \mathbb{P}$ and

$$\dim(\mathbb{P} + \mathcal{F}_{K,\mathbb{P}}(\mathbf{Z})) = \text{Cardinal}(\mathbf{Z})\,.$$

Now, if $\mathbf{Z}$ is not a minimal $\mathbb{P}$-unisolvent set, it strictly contains a minimal $\mathbb{P}$-unisolvent set $\mathbf{\Xi}$.
Let us first show that $K_{\mathbf{z}}^{\mathbf{\Xi}}, \mathbf{z} \in \mathbf{Z} - \mathbf{\Xi}$ is a $\mathcal{F}_{K^{\mathbf{\Xi}}}(\mathbf{Z})$-basis.
Otherwise there would be $\mathbf{z}_0 \in \mathbf{Z} - \mathbf{\Xi}$ such that, setting $\mathbf{Z}' = \mathbf{Z} - \{\mathbf{z}_0\}$, $K_{\mathbf{z}}^{\mathbf{\Xi}}, \mathbf{z} \in \mathbf{Z}' - \mathbf{\Xi}$ spans $\mathcal{F}_{K^{\mathbf{\Xi}}}(\mathbf{Z})$. Hence we would have $\mathbb{P} + \mathcal{F}_{K^{\mathbf{\Xi}}}(\mathbf{Z}') = \mathbb{P} + \mathcal{F}_{K^{\mathbf{\Xi}}}(\mathbf{Z})$ or equivalently

$$\mathbb{P} + \mathcal{F}_{K,\mathbb{P}}(\mathbf{Z}') = \mathbb{P} + \mathcal{F}_{K,\mathbb{P}}(\mathbf{Z})\,,$$

which, since $\mathbf{Z}'$, containing $\mathbf{\Xi}$, is a $\mathbb{P}$- unisolvent set, conflicts with $\mathbf{Z}$ beeing free.
Therefore $\dim(\mathbb{P} + \mathcal{F}_{K,\mathbb{P}}(\mathbf{Z})) = \dim(\mathbb{P} + \mathcal{F}_{K^{\mathbf{\Xi}}}(\mathbf{Z})) = n + \text{Cardinal}(\mathbf{Z}) - n = \text{Cardinal}(\mathbf{Z})$.
Conversely assume that $\dim(\mathbb{P} + \mathcal{F}_{K,\mathbb{P}}(\mathbf{Z})) = \text{Cardinal}(\mathbf{Z})$.
If $\mathbf{Z}$ were not free. There would exist $\mathbf{Z}'$, a $\mathbb{P}$-unisolvent strict subset of $\mathbf{Z}$, verifying

$$\mathbb{P} + \mathcal{F}_{K,\mathbb{P}}(\mathbf{Z}') = \mathbb{P} + \mathcal{F}_{K,\mathbb{P}}(\mathbf{Z})\,.$$

Thus, since $K_{\mathbf{z}}^{\mathbf{\Xi}}, \mathbf{z} \in \mathbf{Z}' - \mathbf{\Xi}$ spans $\mathcal{F}_{K^{\mathbf{\Xi}}}(\mathbf{Z}')$:

$$\begin{aligned}\dim(\mathbb{P} + \mathcal{F}_{K,\mathbb{P}}(\mathbf{Z})) &= \dim(\mathbb{P} + \mathcal{F}_{K,\mathbb{P}}(\mathbf{Z}')) \\ &= \dim(\mathbb{P}) + \dim(\mathcal{F}_{K^{\mathbf{\Xi}}}(\mathbf{Z}')) \\ &< n + \text{Cardinal}(\mathbf{Z}) - n = \text{Cardinal}(\mathbf{Z})\,,\end{aligned}$$

which conflicts with the hypothesis $\dim(\mathbb{P} + \mathcal{F}_{K,\mathbb{P}}(\mathbf{Z})) = \text{Cardinal}(\mathbf{Z})$. □
In order to state our second freeness characterization, we need some more definitions.

Definition 3.6. *Let $\mathcal{P} = (p_1, \ldots, p_n)$ be a $\mathbb{P}$-basis.*
To any finite set $\mathbf{Z} = \{\mathbf{z}_1, \ldots, \mathbf{z}_M\} \subset E$ we define the matrix

$$\mathbf{Q}_{\mathcal{P},\mathbf{Z}} = \begin{pmatrix} \mathbf{K_Z} & \mathbf{P_Z} \\ \mathbf{P_Z}^T & 0 \end{pmatrix},$$

where

- $\mathbf{P_Z} = \begin{pmatrix} p_1(\mathbf{z}_1) & \ldots & p_n(\mathbf{z}_1) \\ \vdots & \ldots & \vdots \\ p_1(\mathbf{z}_M) & \ldots & p_n(\mathbf{z}_M) \end{pmatrix},$
- $\mathbf{K_Z} = \begin{pmatrix} K(\mathbf{z}_1, \mathbf{z}_1) & \ldots & K(\mathbf{z}_1, \mathbf{z}_M) \\ \vdots & \ldots & \vdots \\ K(\mathbf{z}_M, \mathbf{z}_1) & \ldots & K(\mathbf{z}_M, \mathbf{z}_M) \end{pmatrix}.$

If $\mathbf{Q}_{\mathcal{P},\mathbf{Z}}$ is non degenerate we have this helpful construction.

Lemma 3.8. *$\mathcal{P}$ and $\mathbf{Z}$ being as in definition 3.6, if $\mathbf{Q}_{\mathcal{P},\mathbf{Z}}$ is non degenerate then the application $\mathcal{R}_{\mathcal{P},\mathbf{Z}}$ defined as*

$$\mathcal{R}_{\mathcal{P},\mathbf{Z}} : \mathbf{w} \in \mathbb{R}^M \mapsto \sum_{i=1}^{n} \alpha_i(\mathbf{w}) p_i + \sum_{j=1}^{M} \gamma_j(\mathbf{w}) K_{\mathbf{z}_j} ,$$

where, for any $\mathbf{w} \in \mathbb{R}^M$, $\boldsymbol{\alpha}(\mathbf{w}) = \begin{pmatrix} \alpha_1(\mathbf{w}) \\ \vdots \\ \alpha_n(\mathbf{w}) \end{pmatrix} \in \mathbb{R}^n$, $\boldsymbol{\gamma}(\mathbf{w}) = \begin{pmatrix} \gamma_1(\mathbf{w}) \\ \vdots \\ \gamma_M(\mathbf{w}) \end{pmatrix} \in \mathbb{R}^M$ are such that $\begin{pmatrix} \boldsymbol{\gamma}(\mathbf{w}) \\ \boldsymbol{\alpha}(\mathbf{w}) \end{pmatrix} = \mathbf{Q}_{\mathcal{P},\mathbf{Z}}^{-1} \begin{pmatrix} \mathbf{w} \\ 0 \end{pmatrix}$ is a linear isomorphism between $\mathbb{R}^M$ and $\mathbb{P} + \mathcal{F}_{K,\mathbb{P}}(\mathbf{Z})$.

Proof
For any $\mathbf{w} \in \mathbb{R}^M$,

$$\begin{pmatrix} \boldsymbol{\gamma}(\mathbf{w}) \\ \boldsymbol{\alpha}(\mathbf{w}) \end{pmatrix} = \mathbf{Q}_{\mathcal{P},\mathbf{Z}}^{-1} \begin{pmatrix} \mathbf{w} \\ 0 \end{pmatrix}$$

is equivalently rephrased as

$$(\boldsymbol{\alpha}(\mathbf{w}), \boldsymbol{\gamma}(\mathbf{w})) \text{ is the unique solution of } \begin{cases} \mathbf{K}_{\mathbf{Z}}\boldsymbol{\gamma} + \mathbf{P}_{\mathbf{Z}}\boldsymbol{\alpha} = \mathbf{w} \\ {\mathbf{P}_{\mathbf{Z}}}^T\boldsymbol{\gamma} = 0 \end{cases} .$$

The second equation tells us that $\mathcal{R}_{\mathcal{P},\mathbf{Z}}(\mathbf{w}) = \sum_{i=1}^{n} \alpha_i(\mathbf{w}) p_i + \sum_{j=1}^{M} \gamma_j(\mathbf{w}) K_{\mathbf{z}_j} \in \mathbb{P} + \mathcal{F}_{K,\mathbb{P}}(\mathbf{Z})$.
Moreover, $\mathcal{R}_{\mathcal{P},\mathbf{Z}}$ is an onto application since if $g = \sum_{i=1}^{n} \alpha_i p_i + \sum_{j=1}^{M} \gamma_j K_{\mathbf{z}_j} \in \mathbb{P} + \mathcal{F}_{K,\mathbb{P}}(\mathbf{Z})$, let $g_{\mathbf{Z}}$ be the vector whose coordinates are the values taken by g on $\mathbf{Z}$, we have

$$\begin{cases} \mathbf{K}_{\mathbf{Z}}\boldsymbol{\gamma} + \mathbf{P}_{\mathbf{Z}}\boldsymbol{\alpha} = g_{\mathbf{Z}} \\ {\mathbf{P}_{\mathbf{Z}}}^T\boldsymbol{\gamma} = 0 \end{cases} ,$$

which means $\mathcal{R}_{\mathcal{P},\mathbf{Z}}(g_{\mathbf{Z}}) = g$.
Lastly, $\mathcal{R}_{\mathcal{P},\mathbf{Z}}$ is injective, since, according to corollary 3.1, $\mathcal{R}_{\mathcal{P},\mathbf{Z}}(\mathbf{w})$ as a function of $\mathbb{P}+\mathcal{F}_{K,\mathbb{P}}(\mathbf{Z})$ is uniquely defined by its values on $\mathbf{Z}$, which are the coordinates of $\mathbf{w}$.
Therafter $\mathcal{R}_{\mathcal{P},\mathbf{Z}}$ is a bijection from $\mathbb{R}^M$ to $\mathbb{P} + \mathcal{F}_{K,\mathbb{P}}(\mathbf{Z})$. □

We can then state our second freeness characterization

Lemma 3.9. *Let $\mathcal{P}$ be a $\mathbb{P}$-basis.*
A $\mathbb{P}$-unisolvent set $\mathbf{Z}$ is free if and only if $\mathbf{Q}_{\mathcal{P},\mathbf{Z}}$ is non degenerate.

Proof
Let us denote $M = \text{Cardinal}(\mathbf{Z})$.
Suppose that $\mathbf{Q}_{\mathcal{P},\mathbf{Z}}$ is degenerate: let $(\boldsymbol{\gamma}, \boldsymbol{\alpha}) \neq (0,0) \in \mathbb{R}^n \times \mathbb{R}^M$ such that $\mathbf{Q}_{\mathcal{P},\mathbf{Z}} \begin{pmatrix} \boldsymbol{\gamma} \\ \boldsymbol{\alpha} \end{pmatrix} = 0$
i.e

$$\begin{cases} \mathbf{K}_{\mathbf{Z}}\boldsymbol{\gamma} + \mathbf{P}_{\mathbf{Z}}\boldsymbol{\alpha} = 0 \\ {\mathbf{P}_{\mathbf{Z}}}^T\boldsymbol{\gamma} = 0 \end{cases} . \tag{3.24}$$

The function $f = \sum_{i=1}^{n} \alpha_i p_i + \sum_{j=1}^{M} \gamma_j K_{\mathbf{z}_j}$ is in $\mathbb{P} + \mathcal{F}_{K,\mathbb{P}}(\mathbf{Z})$ since the second equation of (3.24) implies $\sum_{j=1}^{M} \gamma_j \delta_{\mathbf{z}_j} \in \mathcal{M}_{\mathbb{P}}(\mathbf{Z})$. The first equation tells us that f is null on $\mathbf{Z}$, and

actually everywhere from corollary 3.1 of Proposition 3.11.
Now

$$f = 0 \Leftrightarrow \sum_{i=1}^{n} \alpha_i p_i = -\sum_{j=1}^{M} \gamma_j K_{\mathbf{z}_j} .$$

But that implies

$$\left\{ \begin{array}{l} \sum_{i=1}^{n} \alpha_i p_i = 0 \\ \sum_{j=1}^{M} \gamma_j K_{\mathbf{z}_j} = 0 \end{array} \right. , \tag{3.25}$$

since, K being $\mathbb{P}$-conditionally positive definite, we have $\mathbb{P} \cap \mathcal{F}_{K,\mathbb{P}}(\mathbf{Z}) = \{0\}$.
First equation of (3.25) gives $\alpha_i = 0, i = 1, \ldots, n$, since $\mathcal{P} = \{p_1, \ldots, p_n\}$ is a $\mathbb{P}$-basis.
Hence

$$\sum_{j=1}^{M} \gamma_j K_{\mathbf{z}_j} = 0 , \tag{3.26}$$

with at least one of the $\gamma_j, j = 1, \ldots, M$ being different of 0.
Notice that consequently, $\mathbf{Z}$ which is $\mathbb{P}$-unisolvent cannot be a **minimal** $\mathbb{P}$-unisolvent set: if it were then $\mathcal{M}_{\mathbb{P}}(\mathbf{Z}) = \{0\}$ and therefore $\sum_{j=1}^{M} \gamma_j \delta_{\mathbf{z}_j} = 0$. That would imply that $\gamma_j = 0, j = 1, \ldots, M$.
So $\mathbf{Z}$ contains a minimal $\mathbb{P}$-unisolvent set $\mathbf{\Xi}$ which is a strict subset. Observe, now, that at least one l of $\{1, \ldots, M\}$ is such that $\mathbf{z}_l \in \mathbf{Z} - \mathbf{\Xi}$ and $\gamma_l \neq 0$: otherwise, $\sum_{j=1}^{M} \gamma_j \delta_{\mathbf{z}_j}$ would belong to $\mathcal{M}_{\mathbb{P}}(\mathbf{\Xi})$ which reduces to $\{0\}$ and therefore $\gamma_j = 0, j = 1, \ldots, M$ would be implied.
Thus, there is j, say $j = 1$, such that $\mathbf{Z}' = \mathbf{Z} - \{\mathbf{z}_1\}$ is $\mathbb{P}$-unisolvent and $\gamma_1 \neq 0$.
Let us now show that $\mathbb{P} + \mathcal{F}_{K,\mathbb{P}}(\mathbf{Z}) = \mathbb{P} + \mathcal{F}_{K,\mathbb{P}}(\mathbf{Z}')$.
Thanks to (3.26), every $g = \sum_{i=1}^{n} \beta_i p_i + \sum_{j=1}^{M} \rho_j K_{\mathbf{z}_j}$ in $\mathbb{P} + \mathcal{F}_{K,\mathbb{P}}(\mathbf{Z})$ can be written

$$g = \sum_{i=1}^{n} \beta_i p_i + \sum_{j=2}^{M} (\rho_j - \rho_1 \frac{\gamma_j}{\gamma_1}) K_{\mathbf{z}_j} .$$

To show that $g \in \mathbb{P} + \mathcal{F}_{K,\mathbb{P}}(\mathbf{Z}')$ we just have to verify that: $\sum_{j=2}^{M} (\rho_j - \rho_1 \frac{\gamma_j}{\gamma_1}) \delta_{\mathbf{z}_j} \in \mathcal{M}_{\mathbb{P}}$. Indeed, for $p \in \mathbb{P}$

$$\begin{aligned} \sum_{j=2}^{M} (\rho_j - \rho_1 \frac{\gamma_j}{\gamma_1}) p(\mathbf{z}_j) &= \sum_{j=2}^{M} \rho_j p(\mathbf{z}_j) - \frac{\rho_1}{\gamma_1} \sum_{j=2}^{M} \gamma_j p(\mathbf{z}_j) \\ &= -\rho_1 p(\mathbf{z}_1) + \frac{\rho_1}{\gamma_1} \gamma_1 p(\mathbf{z}_1) \\ &= 0 , \end{aligned}$$

where we used

$$\sum_{j=1}^{M} \rho_j \delta_{\mathbf{z}_j} \in \mathcal{M}_{\mathbb{P}} \Rightarrow \sum_{j=1}^{M} \rho_j p(\mathbf{z}_j) = 0 \Rightarrow \sum_{j=2}^{M} \rho_j p(\mathbf{z}_j) = -\rho_1 p(\mathbf{z}_1)$$

and

$$\sum_{j=1}^{M} \gamma_j \delta_{\mathbf{z}_j} \in \mathcal{M}_{\mathbb{P}} \Rightarrow \sum_{j=1}^{M} \gamma_j p(\mathbf{z}_j) = 0 \Rightarrow \sum_{j=2}^{M} \gamma_j p(\mathbf{z}_j) = -\gamma_1 p(\mathbf{z}_1) .$$

Conversely, if $\mathbf{Q}_{\mathcal{P},\mathbf{Z}}$ is not degenerate, then from Lemma 3.8, we know that $\mathcal{R}_{\mathcal{P},\mathbf{Z}}$ is a linear isomorphism between $\mathbb{R}^M$ and $\mathbb{P} + \mathcal{F}_{K,\mathbb{P}}(\mathbf{Z})$ and we thus have:

$$\dim(\mathbb{P} + \mathcal{F}_{K,\mathbb{P}}(\mathbf{Z})) = \text{Cardinal}(\mathbf{Z}),$$

which from Lemma 3.7 implies that $\mathbf{Z}$ is free. □

Lagrangian formulation

Proposition 3.13 (Lagrangian formulation)**.** *Let $\mathbf{X}$ be a $\mathbb{P}$-unisolvent set. For any free $\mathbb{P}$-unisolvent set $\mathbf{X}' = \{\mathbf{x}'_1, \ldots, \mathbf{x}'_{N'}\} \subset \mathbf{X}$ satisfying*

$$\mathbb{P} + \mathcal{F}_{K,\mathbb{P}}(\mathbf{X}) = \mathbb{P} + \mathcal{F}_{K,\mathbb{P}}(\mathbf{X}'), \tag{3.27}$$

the following relations uniquely define $u_1, \ldots, u_{N'}$ in $\mathbb{P} + \mathcal{F}_{K,\mathbb{P}}(\mathbf{X})$,

$$u_k(\mathbf{x}'_l) = \delta_{k,l}, \forall k, l \in \{1, \ldots, N'\}. \tag{3.28}$$

Moreover $u_1, \ldots, u_{N'}$ is a $[\mathbb{P} + \mathcal{F}_{K,\mathbb{P}}(\mathbf{X})]$-basis, and every $g \in \mathbb{P} + \mathcal{F}_{K,\mathbb{P}}(\mathbf{X})$ can be written

$$g = \sum_{k=1}^{N'} g(\mathbf{x}'_k) u_k. \tag{3.29}$$

Consequently,

$$\forall f \in \mathcal{H}_{K,\mathbb{P}},\ S_{K,\mathbb{P},\mathbf{X}}(f) = \sum_{k=1}^{N'} f(\mathbf{x}'_k) u_k. \tag{3.30}$$

Proof

Let $\mathbf{X}' = \{\mathbf{x}'_1, \ldots, \mathbf{x}'_{N'}\}$ be a free $\mathbb{P}$-unisolvent subset of $\mathbf{X}$, such that $\mathbb{P} + \mathcal{F}_{K,\mathbb{P}}(\mathbf{X}) = \mathbb{P} + \mathcal{F}_{K,\mathbb{P}}(\mathbf{X}')$ and $\mathcal{P}$ be a $\mathbb{P}$-basis.
The application $\mathcal{R}_{\mathcal{P},\mathbf{X}'}$ defined in Lemma 3.8 is a linear isomorphism between $\mathbb{R}^{N'}$ and $\mathbb{P} + \mathcal{F}_K(\mathbf{X}') = \mathbb{P} + \mathcal{F}_K(\mathbf{X})$.
Let $\mathbf{e}_1, \ldots, \mathbf{e}_{N'}$ be the canonical $\mathbb{R}^{N'} - basis$.
Let us set:

$$u_j = \mathcal{R}_{\mathcal{P},\mathbf{X}'}(\mathbf{e}_j),$$

where $u_1, \ldots, u_{N'}$ satisfies (3.28).
Indeed, $u_k = \mathcal{R}_{\mathcal{P},\mathbf{X}'}(\mathbf{e}_k)$ means:

$$u_k = \sum_{i=1}^{n} \alpha_i^{(k)} p_i + \sum_{j=1}^{N'} \gamma_j^{(k)} K_{\mathbf{x}'_j},$$

where $\boldsymbol{\alpha}^{(k)} = \begin{pmatrix} \alpha_1^{(k)} \\ \vdots \\ \alpha_n^{(k)} \end{pmatrix}$, $\boldsymbol{\gamma}^{(k)} = \begin{pmatrix} \gamma_1^{(k)} \\ \vdots \\ \gamma_{N'}^{(k)} \end{pmatrix}$ is the unique solution of

$$\begin{cases} \mathbf{K}_{\mathbf{X}'}\boldsymbol{\gamma} + \mathbf{P}_{\mathbf{X}'}\boldsymbol{\alpha} = \mathbf{e}_k \\ {\mathbf{P}_{\mathbf{X}'}}^T\boldsymbol{\gamma} = 0 \end{cases}. \tag{3.31}$$

The first equation reads exactly: $u_k(\mathbf{z}_l) = \begin{cases} 0 \text{ if } k \neq l \\ 1 \text{ if } k = l \end{cases}$ which is (3.28).
Now, satisfying (3.28), $u_1, \dots, u_{N'}$ are obviously linearly independant, and, since

$$N' = \dim(\mathbb{P} + \mathcal{F}_{K,\mathbb{P}}(\mathbf{X}')) = \dim(\mathbb{P} + \mathcal{F}_{K,\mathbb{P}}(\mathbf{X}))\,,$$

$u_1, \dots, u_{N'}$ is a $[\mathbb{P} + \mathcal{F}_{K,\mathbb{P}}(\mathbf{X})]$-basis.
Every $g \in \mathbb{P} + \mathcal{F}_{K,\mathbb{P}}(\mathbf{X})$ can thus be written $g = \sum_{i=1}^{N'} \alpha_i u_i$, and by (3.28) we get

$$g(\mathbf{x}'_j) = \sum_{i=1}^{N'} \alpha_i u_i(\mathbf{x}'_j) = \alpha_j\,,$$

hence (3.29).
And (3.30) follows immediately.
Unicity of $u_1, \dots, u_{N'}$ satisfying (3.28) is immediate, since any other $v_1, \dots, v_{N'}$ satisfying (3.28) would verify

$$v_j = \sum_{i=1}^{N'} v_i(\mathbf{x}'_j) u_i = u_j\,.$$

□
To conclude this section devoted to interpolation, let us make several remarks

1. The preceding proof gives a direct method to compute $(u_1, \dots, u_{N'})$: we only have to solve (3.31), that is to compute the inverse of $\mathbf{Q}_{\mathcal{P},\mathbf{X}'}$.

2. In the native spaces and kriging literature (Schaback, 2007; Wendland, 2005), we find this relation:
$$\begin{cases} \mathbf{K}_{\mathbf{X}'}\ \mathbf{u}(\mathbf{x}) & + & \mathbf{P}_{\mathbf{X}'}\ \mathbf{v}(\mathbf{x}) & = & \mathbf{k}_{\mathbf{X}'}(\mathbf{x}) \\ {\mathbf{P}_{\mathbf{X}'}}^T\ \mathbf{u}(\mathbf{x}) & & & = & \mathbf{p}(\mathbf{x}) \end{cases}, \tag{3.32}$$
satisfied by $\mathbf{k}_{\mathbf{X}'}(\mathbf{x}) = \begin{pmatrix} K_{\mathbf{x}'_1}(\mathbf{x}) \\ \vdots \\ K_{\mathbf{x}'_{N'}}(\mathbf{x}) \end{pmatrix}$, $\mathbf{p}(\mathbf{x}) = \begin{pmatrix} p_1(\mathbf{x}) \\ \vdots \\ p_n(\mathbf{x}) \end{pmatrix}$, $\mathbf{u}(\mathbf{x}) = \begin{pmatrix} u_1(\mathbf{x}) \\ \vdots \\ u_{N'}(\mathbf{x}) \end{pmatrix}$ and a vector $\mathbf{v}(\mathbf{x}) = \begin{pmatrix} v_1(\mathbf{x}) \\ \vdots \\ v_{N'}(\mathbf{x}) \end{pmatrix} \in \mathbb{R}^n$.
The solution of (3.32) in $\mathbf{u}(\mathbf{x}), \mathbf{v}(\mathbf{x})$ leads to $u_1(\mathbf{x}), \dots, u_{N'}(\mathbf{x})$.
Let us see why there exists $\mathbf{v}(\mathbf{x}) \in \mathbb{R}^n$ such that (3.32) is verified.
Firstly, each of $(p_1, \dots, p_n)$ in $\mathbb{P}$ belongs to $\mathbb{P} + \mathcal{F}_{K,\mathbb{P}}(\mathbf{X}')$. Thus

$$p_i(\mathbf{x}) = \sum_{k=1}^{N'} p_i(\mathbf{x}_k) u_k(\mathbf{x}), i = 1, \dots, n\,,$$

which is the second equation of (3.32).
Then, recall that

$$u_k(\mathbf{x}) = \sum_{i=1}^{n} \alpha_i^{(k)} p_i(\mathbf{x}) + \sum_{j=1}^{N'} \gamma_j^{(k)} K_{\mathbf{x}'_j}(\mathbf{x}) = \begin{pmatrix} {\boldsymbol{\gamma}^{(k)}}^T & {\boldsymbol{\alpha}^{(k)}}^T \end{pmatrix} \begin{pmatrix} \mathbf{k}_{\mathbf{X}'}(\mathbf{x}) \\ \mathbf{p}(\mathbf{x}) \end{pmatrix},$$

where $\boldsymbol{\alpha}^{(k)} = \begin{pmatrix} \alpha_1^{(k)} \\ \vdots \\ \alpha_n^{(k)} \end{pmatrix}$, $\boldsymbol{\gamma}^{(k)} = \begin{pmatrix} \gamma_1^{(k)} \\ \vdots \\ \gamma_{N'}^{(k)} \end{pmatrix}$ are given by

$\begin{pmatrix} \boldsymbol{\gamma}^{(k)} \\ \boldsymbol{\alpha}^{(k)} \end{pmatrix} = \mathbf{Q}_{\mathcal{P},\mathbf{X}'}^{-1} \begin{pmatrix} \mathbf{e}_k \\ 0 \end{pmatrix}$ so that:

$$\mathbf{u}(\mathbf{x}) = \begin{pmatrix} {\boldsymbol{\gamma}^{(1)}}^T & {\boldsymbol{\alpha}^{(1)}}^T \\ \vdots & \vdots \\ {\boldsymbol{\gamma}^{(N')}}^T & {\boldsymbol{\alpha}^{(N')}}^T \end{pmatrix} \begin{pmatrix} \mathbf{k}_{\mathbf{X}'}(\mathbf{x}) \\ \mathbf{p}(\mathbf{x}) \end{pmatrix} = \begin{pmatrix} \mathrm{Id}_{N'} & 0 \end{pmatrix} \mathbf{Q}_{\mathcal{P},\mathbf{X}'}^{-1} \begin{pmatrix} \mathbf{k}_{\mathbf{X}'}(\mathbf{x}) \\ \mathbf{p}(\mathbf{x}) \end{pmatrix}$$

and

$$\mathbf{K}_{\mathbf{X}'}\mathbf{u}(\mathbf{x}) = \mathbf{K}_{\mathbf{X}'} \begin{pmatrix} \mathrm{Id}_{N'} & 0 \end{pmatrix} \mathbf{Q}_{\mathcal{P},\mathbf{X}'}^{-1} \begin{pmatrix} \mathbf{k}_{\mathbf{X}'}(\mathbf{x}) \\ \mathbf{p}(\mathbf{x}) \end{pmatrix} = \begin{pmatrix} \mathbf{K}_{\mathbf{X}'} & 0 \end{pmatrix} \mathbf{Q}_{\mathcal{P},\mathbf{X}'}^{-1} \begin{pmatrix} \mathbf{k}_{\mathbf{X}'}(\mathbf{x}) \\ \mathbf{p}(\mathbf{x}) \end{pmatrix}.$$

Now it is readily seen that

$$\begin{pmatrix} \mathbf{K}_{\mathbf{X}'} & 0 \end{pmatrix} \mathbf{Q}_{\mathcal{P},\mathbf{X}'}^{-1} = \begin{pmatrix} \mathrm{Id}_{N'} & 0 \end{pmatrix} - \mathbf{P}_{\mathbf{X}'}\mathbf{M}, \tag{3.33}$$

for a $(n \times (N' + n))$ well chosen matrix $\mathbf{M}$, leading to

$$\mathbf{K}_{\mathbf{X}'}\mathbf{u}(\mathbf{x}) + \mathbf{P}_{\mathbf{X}'}\mathbf{M} \begin{pmatrix} \mathbf{k}_{\mathbf{X}'}(\mathbf{x}) \\ \mathbf{p}(\mathbf{x}) \end{pmatrix} = \mathbf{k}_{\mathbf{X}'}(\mathbf{x})$$

which is the first equation of (3.32) with $\mathbf{v}(\mathbf{x}) = \mathbf{M} \begin{pmatrix} \mathbf{k}_{\mathbf{X}'}(\mathbf{x}) \\ \mathbf{p}(\mathbf{x}) \end{pmatrix}$.

Regarding (3.33), let us denote $\mathbf{A}, \mathbf{B}, \mathbf{C}$ and $\mathbf{D}$ those matrices of respective dimensions $N' \times N'$, $N' \times n$, $n \times N'$ and $n \times n$ such that

$$\mathbf{Q}_{\mathcal{P},\mathbf{X}'}^{-1} = \begin{pmatrix} \mathbf{A} & \mathbf{B} \\ \mathbf{C} & \mathbf{D} \end{pmatrix}.$$

We have

$$\begin{pmatrix} \mathbf{K}_{\mathbf{X}'} & 0 \end{pmatrix} \mathbf{Q}_{\mathcal{P},\mathbf{X}'}^{-1} = \begin{pmatrix} \mathbf{K}_{\mathbf{X}'}\mathbf{A} & \mathbf{K}_{\mathbf{X}'}\mathbf{B} \end{pmatrix}.$$

Using $\mathbf{Q}_{\mathcal{P},\mathbf{X}'}\mathbf{Q}_{\mathcal{P},\mathbf{X}'}^{-1} = \mathrm{Id}_{N'+n}$, we get

$$\mathbf{K}_{\mathbf{X}'}\mathbf{A} = \mathrm{Id}_{N'} - \mathbf{P}_{\mathbf{X}'}\mathbf{C} \text{ and } \mathbf{K}_{\mathbf{X}'}\mathbf{B} = -\mathbf{P}_{\mathbf{X}'}\mathbf{D},$$

and eventually

$$\begin{pmatrix} \mathbf{K}_{\mathbf{X}'} & 0 \end{pmatrix} \mathbf{Q}_{\mathcal{P},\mathbf{X}'}^{-1} = \begin{pmatrix} \mathrm{Id}_{N'} & 0 \end{pmatrix} - \mathbf{P}_{\mathbf{X}'} \begin{pmatrix} \mathbf{C} & \mathbf{D} \end{pmatrix}.$$

3. Kriging ((Koehler and Owen, 1996; Vazquez, 2005) is very popular in computer experiments and geostatistics. Let us recall how that technique is linked to interpolation. Kriging aims at approximating a function $f \in \mathbb{R}^E$ only known on a *design* $\mathbf{X} = \{\mathbf{x}_1, \ldots, \mathbf{x}_N\} \subset E$. In its simplest form, it postulates that f is a realization of a gaussian process F whose parameter lies in E:

$$F(\mathbf{x}) = \sum_{i=1}^{n} \beta_i p_i(\mathbf{x}) + Z(\mathbf{x}), \tag{3.34}$$

where $(p_1, \ldots, p_n)$ is a basis of a vector space of functions $\mathbb{P} \subset \mathbb{R}^E$ and Z is a centered gaussian process. Then it consists in approximating $f(\mathbf{x})$ by the best linear unbiased predictor (BLUP).
Now, it is readily seen that the BLUP depends on F only through this centered gaussian process whose parameter is in $\mathcal{M}_{\mathbb{P}}$:
$F_{\mathbb{P}}(\sum_{m=1}^{M} \mu_m \delta_{\mathbf{x}_k}) = \sum_{m=1}^{M} \mu_m F(\mathbf{x}_i)$.
Hence the idea of *intrinsic* kriging: forget the model (3.34) and start, instead, with G, a centered gaussian process whose parameter is in $\mathcal{M}_{\mathbb{P}}$ and whose covariance is specified by a $\mathbb{P}$-conditionally positive definite kernel K, then solve the BLUP equations with G in place of $F_{\mathbb{P}}$.
This method leads exactly to the same equations than those that are to be solved to get the interpolator $S_{K,\mathbb{P},\mathbf{X}}(f)$.

4. Observing that $\delta_{\mathbf{x}} - \sum_{k=1}^{N'} u_k(\mathbf{x}) \delta_{\mathbf{x}'_k} \in \mathcal{M}_{\mathbb{P}}$ we rediscover this error estimation

$$
\begin{aligned}
|f(\mathbf{x}) - S_{K,\mathbb{P},\mathbf{X}}(f)(\mathbf{x})| &= \left| \left[\delta_{\mathbf{x}} - \sum_{k=1}^{N'} u_k(\mathbf{x}) \delta_{\mathbf{x}'_k} \right] (f) \right| \\
&= | < F_K(\delta_{\mathbf{x}} - \sum_{k=1}^{N'} u_k(\mathbf{x}) \delta_{\mathbf{x}'_k}), f >_{\mathcal{H}_{K,\mathbb{P}}} \\
&= | < K_{\mathbf{x}} - \sum_{k=1}^{N'} u_k(\mathbf{x}) K_{\mathbf{x}'_k}, f >_{\mathcal{H}_{K,\mathbb{P}}} | \\
&\leq ||f||_{\mathcal{H}_{K,\mathbb{P}}} \, ||K_{\mathbf{x}} - \sum_{k=1}^{N'} u_k(\mathbf{x}) K_{\mathbf{x}'_k}||_{\mathcal{H}_{K,\mathbb{P}}}.
\end{aligned}
$$

3.6 Regularized regression in RKSHS

As in the previous section, it is assumed that $\mathbb{P}$ denotes a finite dimensional vector space of functions and that K is a $\mathbb{P}$-conditionally positive definite kernel. Furthermore, suppose that, besides the "design" $\mathbf{X}$, we are given values $\mathbf{y}_1, \ldots, \mathbf{y}_N \in \mathbb{R}$. For $\mathbb{P}$ a finite dimensional vector space and K a $\mathbb{P}$-conditionally positive definite kernel, we want now to solve the following regularized regression problem:

$$
\min_{f \in \mathcal{H}_{K,\mathbb{P}}} \sum_{k=1}^{N} (\mathbf{y}_k - f(\mathbf{x}_k))^2 + \lambda ||f||^2_{\mathcal{H}_{K,\mathbb{P}}}, \tag{3.35}
$$

where λ is a strictly positive real.
The representer theorem is true in $\mathbb{P}$-RKSHS:

Theorem 3.4. *Any solution of (3.35) lies in* $\mathbb{P} + \mathcal{F}_{K,\mathbb{P}}(\mathbf{X})$.

Proof
Let $f \in \mathcal{H}_{K,\mathbb{P}}$ be a solution of problem (3.35).
By Proposition 3.12, $g = S_{K,\mathbb{P},\mathbf{X}}(f)$ belongs to $\mathbb{P} + \mathcal{F}_{K,\mathbb{P}}(\mathbf{X})$ and interpolates f on $\mathbf{X}$, hence

$$
\sum_{k=1}^{N} (\mathbf{y}_k - f(\mathbf{x}_k))^2 = \sum_{k=1}^{N} (\mathbf{y}_k - g(\mathbf{x}_k))^2 .
$$

Moreover, if f and g were distinct, the same proposition 3.12 would imply:

$$||g||_{\mathcal{H}_{K,\mathbb{P}}} < ||f||_{\mathcal{H}_{K,\mathbb{P}}}\,,$$

thus,

$$\sum_{k=1}^{N}(\mathbf{y}_k - g(\mathbf{x}_k))^2 + \lambda||g||^2_{\mathcal{H}_{K,\mathbb{P}}} < \sum_{k=1}^{N}(\mathbf{y}_k - f(\mathbf{x}_k))^2 + \lambda||f||^2_{\mathcal{H}_{K,\mathbb{P}}}\,,$$

which contradicts the fact that f is a solution of (3.35).
□
Explicit solution of (3.35) is given by:

Proposition 3.14. *Let* $\mathbf{X}$ *be a free* $\mathbb{P}$*-unisolvent set.*

The solution of (3.35) is $f = \sum_{i=1}^{n}\alpha_i p_i + \sum_{j=1}^{N}\gamma_j K_{\mathbf{x}_j}$ *with* $\boldsymbol{\alpha} = \begin{pmatrix}\alpha_1\\ \vdots\\ \alpha_n\end{pmatrix} \in \mathbb{R}^n, \boldsymbol{\gamma} = \begin{pmatrix}\gamma_1\\ \vdots\\ \gamma_M\end{pmatrix} \in \mathbb{R}^N$ *given by*

$$\begin{cases} \boldsymbol{\gamma} = (\mathbf{K_X} + \boldsymbol{\lambda}\mathrm{Id}_N)^{-1}(\mathbf{Y} - \mathbf{P_X}\boldsymbol{\alpha}) \\ \boldsymbol{\alpha} = \left[\mathbf{P_X}^T(\mathbf{K_X} + \boldsymbol{\lambda}\mathrm{Id}_N)^{-1}\mathbf{P_X}\right]^{-1}\mathbf{P_X}^T(\mathbf{K_X} + \boldsymbol{\lambda}\mathrm{Id}_N)^{-1}\mathbf{Y} \end{cases}, \tag{3.36}$$

where $\mathbf{Y} = \begin{pmatrix}y_1\\ \vdots\\ y_N\end{pmatrix}$.

Proof
From Theorem 3.4, we know that the solution is to be searched in $\mathbb{P} + \mathcal{F}_{K,\mathbb{P}}(\mathbf{X})$.
The function $g = \sum_{i=1}^{n}\alpha_i^{(0)}p_i + \sum_{j=1}^{N}\gamma_j^{(0)}K_{\mathbf{x}_j} \in \mathbb{P} + \mathcal{F}_{K,\mathbb{P}}(\mathbf{X})$ is solution of (3.35) if and only if $\boldsymbol{\alpha}^{(0)} = \begin{pmatrix}\alpha_1^{(0)}\\ \vdots\\ \alpha_n^{(0)}\end{pmatrix} \in \mathbb{R}^n, \boldsymbol{\gamma}^{(0)} = \begin{pmatrix}\gamma_1^{(0)}\\ \vdots\\ \gamma_M^{(0)}\end{pmatrix} \in \mathbb{R}^N$ is solution of

$$\min\{J(\boldsymbol{\alpha},\boldsymbol{\gamma}) : \boldsymbol{\alpha} \in \mathbb{R}^n, \boldsymbol{\gamma} \in \mathbb{R}^N, \mathbf{P_X}^T\boldsymbol{\gamma} = 0\}\,, \tag{3.37}$$

where

$$J(\boldsymbol{\alpha},\boldsymbol{\gamma}) = ||\mathbf{Y} - (\mathbf{K_X}\boldsymbol{\gamma} + \mathbf{P_X}\boldsymbol{\alpha})||^2_{\mathbb{R}^N} + \lambda\boldsymbol{\gamma}^T\mathbf{K_X}\boldsymbol{\gamma}\,.$$

To solve (3.37) let us form the Lagrangian

$$L(\boldsymbol{\alpha},\boldsymbol{\gamma},\boldsymbol{\mu}) = J(\boldsymbol{\alpha},\boldsymbol{\gamma}) + < \mathbf{P_X}^T, \boldsymbol{\mu} >_{\mathbb{R}^n}\ .$$

A solution of (3.37) satisfies the following first order conditions:

$$\begin{cases} 2\mathbf{P_X}^T\left[\mathbf{K_X}\boldsymbol{\gamma} + \mathbf{P_X}\boldsymbol{\alpha} - \mathbf{Y}\right] = 0 \\ 2\mathbf{K_X}\left[\mathbf{K_X}\boldsymbol{\gamma} + \mathbf{P_X}\boldsymbol{\alpha} - \mathbf{Y}\right] + 2\boldsymbol{\lambda}\mathbf{K_X}\boldsymbol{\gamma} + \mathbf{P_X}\boldsymbol{\mu} = 0 \\ \mathbf{P_X}^T\boldsymbol{\gamma} = 0 \end{cases}\ . \tag{3.38}$$

Rewriting first equation as $\left\{ \begin{array}{l} \mathbf{K_X}\boldsymbol{\gamma} + \mathbf{P_X}\boldsymbol{\alpha} - \mathbf{Y} = \mathbf{e} \\ {\mathbf{P_X}}^T\mathbf{e} = 0 \end{array} \right.$, (3.38) becomes

$$\left\{ \begin{array}{l} \mathbf{K_X}\boldsymbol{\gamma} + \mathbf{P_X}\boldsymbol{\alpha} - \mathbf{Y} = \mathbf{e} \\ \mathbf{K_X}\left[\mathbf{e} + \boldsymbol{\lambda}\boldsymbol{\gamma}\right] + \mathbf{P_X}(\frac{1}{2}\boldsymbol{\mu}) = 0 \\ {\mathbf{P_X}}^T\mathbf{e} = 0 \\ {\mathbf{P_X}}^T\boldsymbol{\gamma} = 0 \end{array} \right. . \tag{3.39}$$

From the three last equations we then draw:

$$\left\{ \begin{array}{l} \mathbf{K_X}\left[\mathbf{e} + \boldsymbol{\lambda}\boldsymbol{\gamma}\right] + \mathbf{P_X}(\frac{1}{2}\boldsymbol{\mu}) = 0 \\ {\mathbf{P_X}}^T\left[\mathbf{e} + \boldsymbol{\lambda}\boldsymbol{\gamma}\right] = 0 \end{array} \right. . \tag{3.40}$$

Since $\mathbf{X}$ is free, Lemma 3.9 implies that $\left(\begin{array}{cc} \mathbf{K_X} & \mathbf{P_X} \\ {\mathbf{P_X}}^T & 0 \end{array} \right)$ is non degenerate. Hence (3.40) gives

$$\left\{ \begin{array}{l} \mathbf{e} + \boldsymbol{\lambda}\boldsymbol{\gamma} = 0 \\ \boldsymbol{\mu} = 0 \end{array} \right. \Rightarrow \left\{ \begin{array}{l} \mathbf{e} = -\boldsymbol{\lambda}\boldsymbol{\gamma} \\ \boldsymbol{\mu} = 0 \end{array} \right. ,$$

and, used in (3.39)

$$\left\{ \begin{array}{l} \boldsymbol{\gamma} = (\mathbf{K_X} + \boldsymbol{\lambda}\mathrm{Id}_N)^{-1}(\mathbf{Y} - \mathbf{P_X}\boldsymbol{\alpha}) \\ {\mathbf{P_X}}^T(\mathbf{K_X} + \boldsymbol{\lambda}\mathrm{Id}_N)^{-1}\mathbf{P_X}\boldsymbol{\alpha} = {\mathbf{P_X}}^T(\mathbf{K_X} + \boldsymbol{\lambda}\mathrm{Id}_N)^{-1}\mathbf{Y} \end{array} \right. . \tag{3.41}$$

Notice, then, that ${\mathbf{P_X}}^T(\mathbf{K_X} + \boldsymbol{\lambda}\mathrm{Id}_N)^{-1}\mathbf{P_X}$ is a symmetric positive definite matrix. Indeed, $(\mathbf{K_X} + \boldsymbol{\lambda}\mathrm{Id}_N)^{-1}$ is obviously a symmetric positive definite matrix so that

$$\mathbf{a}^T{\mathbf{P_X}}^T(\mathbf{K_X} + \boldsymbol{\lambda}\mathrm{Id}_N)^{-1}\mathbf{P_X}\mathbf{a} = 0 \Leftrightarrow \mathbf{P_X}\mathbf{a} = 0\,,$$

which implies $\mathbf{a} = 0$ since $\mathbf{X}$ is $\mathbb{P}$-unisolvent.
Hence, eventually, (3.41) leads to

$$\left\{ \begin{array}{l} \boldsymbol{\gamma} = (\mathbf{K_X} + \boldsymbol{\lambda}\mathrm{Id}_N)^{-1}(\mathbf{Y} - \mathbf{P_X}\boldsymbol{\alpha}) \\ \boldsymbol{\alpha} = \left[{\mathbf{P_X}}^T(\mathbf{K_X} + \boldsymbol{\lambda}\mathrm{Id}_N)^{-1}\mathbf{P_X}\right]^{-1}{\mathbf{P_X}}^T(\mathbf{K_X} + \boldsymbol{\lambda}\mathrm{Id}_N)^{-1}\mathbf{Y} \end{array} \right. .$$

□
The solution (3.36) is formally the same as the one proposed by Wahba (1990) in the context of thin-plate splines on $\mathbb{R}^d$ which is known to correspond to this $\mathbb{P}$-conditionally positive definite kernel:

$$K(\mathbf{x}, \mathbf{x}') = (-1)^{k+1}||\mathbf{x} - \mathbf{x}'||^{2k}_{\mathbb{R}^d}\log(||\mathbf{x} - \mathbf{x}'||_{\mathbb{R}^d})\,,$$

where $\mathbb{P}$ is the set of the d-variate polynomials of degree less than $k+1$.

3.7 Discussion

In this paper we propose a new definition of the *conditionally positive definite kernel* which, generalizing the usual one, leads to a full extension of the results of the positive definite case. The core of our work is an Aronszajn's theorem analog which links any conditionally positive definite kernel to a functional semi-Hilbert space (RKSHS), generalizing RKHS for positive definite kernel.

We show that the useful interpolation operator still works and specifically can be computed in this generalized context. As another benchmark test we state the explicit solution of a regularized regression problem, which we recognize to be formally identical to the one stated by Wahba (1990), in the context of thin-plate splines.

Bibliography

Aronszajn, N. (1950). Theory of reproducing kernel. *Transactions of American Mathematical Society*, 68(3):337–404.

Cressie, N. (1993). *Statistics for Spatial Data*. Wiley, New York.

Kimeldorf, G. and Wahba, G. (1971). Some results on tchebycheffian spline functions. *Journal of Mathematical Analysis and Applications*, 33(1):82–95.

Koehler, J. R. and Owen, A. B. (1996). Computer experiments. In *Design and analysis of experiments*, volume 13 of *Handbook of Statistics*, pages 261–308. North Holland, Amsterdam.

Schaback, R. (1997). Native hilbert spaces for radial basis functions i. In *New Developments in Approximation Theory, number 132 in International Series of Numerical Mathematics*, pages 255–282. Birkhauser Verlag.

Schaback, R. (2007). Kernel-based meshless methods. Technical report, Institute for Numerical and Applied Mathematics, Georg-August-University Goettingen.

Vazquez, E. (2005). *Modélisation comportementale de systèmes non-linéaires multivariables par méthodes à noyaux et applications*. PhD thesis, Université Paris-sud.

Wackernagel, H. (2003). *Multivariate Geostatistics: An Introduction with Applications*. Springer.

Wahba, G. (1990). *Spline models for observational data*, volume 59 of *CBMS-NSF Regional Conference Series in Applied Mathematics*. Society for Industrial and Applied Mathematics (SIAM), Philadelphia, PA.

Wendland, H. (2005). *Scattered data approximation*, volume 17 of *Cambridge Monographs on Applied and Computational Mathematics*. Cambridge University Press, Cambridge.

Wendland, H. (2006). Spatial coupling in aeroelasticity by meshless kernel-based methods. In Wesseling, P., Onate, E., and Periaux, J., editors, *ECCOMAS CFD*, Egmond aan Zee, The Netherlands.

Chapitre 4

Maximin design on non hypercube domain and kernel interpolation

Résumé

Dans la partie 2.3, différentes constructions de plans d'expérience numérique exploratoires sont proposées. Les plans d'expérience MAXIMIN (2.47) fondés sur un critère de dispersion entre les points sont justifiés lorsqu'ils servent à la construction d'un interpolateur à noyaux comme métamodèle. Ils sont justifiés du point de vue de l'interpolation dans les RKHS (voir la partie 2.3.2) et du point de vue du krigeage (voir la partie 2.3.3).
Le domaine E sur lequel nous souhaitons approcher une fonction boîte noire est souvent supposé hypercubique. Le cas échéant, la stratégie standard consiste à chercher un plan d'expérience MAXIMIN dans la classe des hypercubes latins. Cependant, si ce n'est pas le cas, l'échantillonnage en hypercube latin ne fait plus sens et n'est pas forcément possible.
Nous proposons alors un algorithme de recherche de plans d'expérience MAXIMIN dans des domaines non nécessairement hypercubiques. Cet algorithme repose sur un recuit simulé dont nous montrons la convergence théorique. Finalement, nous montrons numériquement pour un modèle décrivant un moteur d'avion, le gain possible à construire un interpolateur à noyaux à partir d'un plan d'expérience maximin.

Mots clés : Expériences simulées, Interpolation à noyaux, Krigeage, Plans d'expérience MAXIMIN, Recuit simulé.

Ce chapitre est issu d'une collaboration avec Yves Auffray et Jean-Michel Marin. Il a été soumis pour publication.

Abstract

In the paradigm of computer experiments, the choice of an experimental design is an important issue. When no information is available about the black-box function to be approximated, an exploratory design have to be used. In this context, two dispersion criteria are usually considered: the MINIMAX and the MAXIMIN ones. In the case of a hypercube domain, a standard strategy consists of taking the MAXIMIN design within the class of latin hypercube designs. However, in a non hypercube context, it does not make sense to use the latin hypercube strategy. Moreover, whatever the design is, the black-box function is typically approximated thanks to kernel interpolation. Here, we first provide a theoretical justification to the MAXIMIN criterion with respect to kernel interpolations. Then, we propose simulated annealing algorithms to determine MAXIMIN designs in any bounded connected domain. We prove the convergence of the different schemes. Finally, the methodology is applied on a challenging real example where the black-blox function describes the behaviour of an aircraft engine.

Keywords: Computer experiments, Kernel interpolation, Kriging, MAXIMIN designs, Simulated annealing.

4.1 Introduction

A function $f : E \to \mathbb{R}$ is said to be a black-box function if f is only known through a time consuming code. It is assumed that E is enclosed in a known bounded set of $\mathbb{R}^d$. E is not necessarily a hypercube domain or explicit. E can be given by an indicator function only. In order to deal with some concerns such as pre-visualization, prediction, optimization and

probabilistic analysis which depend on f, an approximation of f is usually used. This is the paradigm of computer experiments (Santner et al., 2003; Fang et al., 2006) where the unknown function f is deterministic. The approximation of f can be obtained thanks to a kernel interpolation method (Schaback, 1995, 2007) also known as kriging (Matheron, 1963). Due to its flexibility and its good properties in high-dimension's case, kriging is one of the most used approximation method by the computer experiments community. For more details on kriging, one can see for instance: Cressie (1993); Laslett (1994); Stein (1999, 2002); Li and Sudjianto (2005); Joseph (2006); den Hertog et al. (2006).

The kernel interpolation methodology needs the choice of a kernel K (kernel satisfying some conditions detailed below) and a design $\mathbf{X} = \{\mathbf{x}_1, \dots \mathbf{x}_N\}$ where the function f is to be evaluated $\{f(\mathbf{x}_1), \dots, f(\mathbf{x}_N)\}$. As it is well-known, a space of functions $\mathcal{H}_K$ is associated to K. If it is assumed that f lies in $\mathcal{H}_K$, the interpolator of f on $\mathbf{X}$, denoted by $s_{K,\mathbf{X}}(f)$, can be used to approximate f. In this deterministic paradigm (the function f is not random), there are essentially two main kinds of properties that a design can have (Koehler and Owen, 1996):

- projection properties such as Latin hypercube designs McKay et al. (1979);
- exploratory properties which are warranted by criteria such as:
  - MINIMAX which means that the design has to minimize

$$h_{\mathbf{X}} = \sup_{\mathbf{y} \in E} \min_{1 \leq i \leq N} \|\mathbf{y} - \mathbf{x}_i\|, \tag{4.1}$$

  - MAXIMIN which means that the design has to maximize

$$\delta_{\mathbf{X}} = \min_{1 \leq i,j \leq N} \|\mathbf{x}_i - \mathbf{x}_j\| . \tag{4.2}$$

    Moreover, between two designs $\mathbf{X}_1$ and $\mathbf{X}_2$ such that $\delta_{\mathbf{X}_1} = \delta_{\mathbf{X}_2}$, using the MAXIMIN criterion, we choose the design for which the number of pairs of points with distance equal to $\delta_{\mathbf{X}_1}$ is minimal.
  - the integrated mean square error (IMSE) criterion (Sacks et al., 1989).

For others criteria, one can see (Bursztyn and Steinberg, 2006).

For kernels defined by radial basis functions, Schaback (1995) and Madych and Nelson (1992) have shown that the MININAX criterion $h_{\mathbf{X}}$ explicitly intervenes in an upper bound on the point-wise error between f and $s_{K,\mathbf{X}}(f)$. The upper bound has the form $G(h_{\mathbf{X}})$ where G is an increasing function $\mathbb{R}_+ \to \mathbb{R}_+$. Here, we generalize this result to the case of MAXIMIN designs.

MINIMAX and IMSE criteria are costly to evaluate and, typically, the MAXIMIN criterion is privileged. In the case where E is a hypercubic set, Morris and Mitchell (1995) provided an algorithm based on simulated annealing to obtain a design very close to a MAXIMIN Latin hypercube designs, (the criterion optimized is not exactly the MAXIMIN one). For the two-dimensional case, van Dam et al. (2007) derived explicit constructions for MAXIMIN Latin hypercube designs when the distance measure is L_∞ or L_1. For the L_2 distance measure, they obtained MAXIMIN Latin hypercube designs for $N \leq 70$.

In the case where E is not hypercubic but only enclosed in a hypercubic set, projection properties are not sensible. Only exploratory properties are to be focused on. In the case of an explicit constrained subset of $[0,1]^d$, Stinstra et al. (2003) proposed an algorithm based on the use of NLP solvers. Here, we propose some algorithms to achieve a MAXIMIN design for general (even not explicit) non hypercubic domains. Our schemes are based on simulated annealing. Our proposals are not heuristic, we study the convergence properties of all our schemes.

Recall that the simulated annealing algorithm aims at finding a global extremum of a function by using a Markovian kernel which is the composition of an exploratory kernel and an acceptance step depending on a temperature which decreases during the iterations. It is based on the Metropolis Hasting-algorithm (Chib and Greenberg, 1995). At a fixed temperature, the Markov chain tends to a stationnary distribution which is the Gibbs measure. As the temperature decreases, the Gibbs measure concentrates on the global extremum of the function (Bartoli and Del Moral, 2001). Hence, the simulated annealing algorithm provides a Markov chain which tends to concentrate on a global extremum of the function to be optimized with high probability when the number of iterations tends to infinity.

The paper is organized as follows, in Section 2 the kernel interpolation method is described and a theoretical justification of the MINIMAX and MAXIMIN criteria is provided thanks to the pointwise error bound between the interpolator and the function f. Then, in Section 3 the simulated annealing algorithm is presented. A proof of convergence is given. Section 4 deals with the case where E is not explicit and can only be known by an indicator function. Two variants of the algorithm are proposed and their theoretical properties are stated. In Section 5, the algorithms are tried on some examples and practical issues are discussed. Finally, in a last Section, the methodology is applied on a real example for which the domain is not an hypercube.

4.2 Error bounds with kernel interpolations

A kernel is a symmetric function $K : E \times E \to \mathbb{R}$ where E is the input space which is assumed to be bounded. The kernel has to be at least conditionally positive definite to be used in kernel interpolation. For the sake of simplicity, kernel interpolation is presented for positive definite kernels only. $\mathbb{R}^E$ denotes the space of functions from E to $\mathbb{R}$.

Définition 4.1. *A kernel K is definite positive if*

$$\forall(\lambda_1, \mathbf{x}_1)\dots(\lambda_N, \mathbf{x}_N) \in \mathbb{R} \times E, \sum_{1\leq l,m\leq N} \lambda_l \lambda_m K(\mathbf{x}_l, \mathbf{x}_m) \geq 0\,.$$

For any $\mathbf{x} \in E$, let $K_{\mathbf{x}}$ denote the partial function $\mathbf{x}' \in E \mapsto K(\mathbf{x}, \mathbf{x}') \in \mathbb{R}$. The linear combinations of functions taken in $\{K_{\mathbf{x}}, \mathbf{x} \in E\}$ span a functional pre-Hilbert space $\mathcal{F}_K$ where

$$< \sum_{l=1}^{L} \lambda_l K_{\mathbf{x}_l}, \sum_{m=1}^{M} \mu_m K_{\mathbf{x}'_m} >_{\mathcal{F}_K} = \sum_{m=1}^{M} \sum_{l=1}^{L} \lambda_l \mu_m K(\mathbf{x}_l, \mathbf{x}'_m)$$

is the scalar product. Aronszajn's theorem states that there exists a unique space $\mathcal{H}_K$ which is a completion of $\mathcal{F}_K$ where the following reproducing property holds

$$\forall f \in \mathcal{H}_K,\ \mathbf{x} \in E,\ f(\mathbf{x}) =< f, K_{\mathbf{x}} >_{\mathcal{H}_K}\ .$$

$\mathcal{H}_K$ is called a Reproducing Kernel Hilbert Space (RKHS).
Let us denote by $s_{K,\mathbf{X}}(f)$ the orthogonal projection of f on $\mathcal{H}_K(\mathbf{X}) = \text{span}\{K_{\mathbf{x}_1}, \ldots, K_{\mathbf{x}_N}\}$ (f is assumed to be in $\mathcal{H}_K$; $\mathbf{X} = \{\mathbf{x}_1, \ldots, \mathbf{x}_N\}$ and K are given).

Lemma 4.1. *$s_{K,\mathbf{X}}(f)$ interpolates f on $\mathbf{X}$. Among the interpolators of f on $\mathbf{X}$, $s_{K,\mathbf{X}}(f)$ has the smallest norm: $s_{K,\mathbf{X}}(f)$ is the solution of the following problem*

$$\begin{cases} \min_{g \in \mathcal{H}_K} \|g\|_{\mathcal{H}_K} \\ g(\mathbf{x}_k) = f(\mathbf{x}_k), \ k = 1, \ldots N \end{cases} .$$

This interpolator is also known in the kriging literature (Cressie, 1993; Stein, 2002) as the best linear unbiased predictor. It has a Lagrangian formulation.

Lemma 4.2. *For any $\mathbf{x} \in E$,*

$$s_{K,\mathbf{X}}(f)(\mathbf{x}) = \sum_{i=1}^{N} u_i(\mathbf{x}) f(\mathbf{x}_i)$$

where the functions $(u_i : E \to \mathbb{R}) \in \mathcal{H}_K(\mathbf{X})$ are such that, $\forall\ 1 \leq i \leq N$,

$$\begin{cases} u_i(\mathbf{x}_i) = 1 \\ u_i(\mathbf{x}_k) = 0 \quad \text{if } k \neq i \end{cases} ,$$

and

$$K[\mathbf{X}, \mathbf{x}] = K[\mathbf{X}, \mathbf{X}] U(\mathbf{x}) \,,$$

where $U(\mathbf{x}) = \begin{pmatrix} u_1(\mathbf{x}) \\ \vdots \\ u_N(\mathbf{x}) \end{pmatrix}$, $K[\mathbf{X}, \mathbf{x}] = \begin{pmatrix} K(\mathbf{x}_1, \mathbf{x}) \\ \vdots \\ K(\mathbf{x}_N, \mathbf{x}) \end{pmatrix}$ *and $K[\mathbf{X}, \mathbf{X}]$ is such that* $(K[\mathbf{X}, \mathbf{X}])_{1 \leq i,j \leq N} = K(\mathbf{x}_i, \mathbf{x}_j)$.

Hence, the pointwise error can be bounded from above, $\forall \mathbf{x} \in E$

$$|f(\mathbf{x}) - s_{K,\mathbf{X}}(f)(\mathbf{x})| = | < f, K_{\mathbf{x}} - \sum_{i=1}^{N} u_i(\mathbf{x}) K_{\mathbf{x}_i} >_{\mathcal{H}_K} | \leq \|f\|_{\mathcal{H}_K} \|K_{\mathbf{x}} - \sum_{i=1}^{N} u_i(\mathbf{x}) K_{\mathbf{x}_i}\|_{\mathcal{H}_K}.$$

Let $P_{\mathbf{X}}(\mathbf{x}) = \|K_{\mathbf{x}} - \sum_{i=1}^{N} u_i(\mathbf{x}) K_{\mathbf{x}_i}\|_{\mathcal{H}_K}$. $P_{\mathbf{X}}$ depends only on the kernel K and on the design $\mathbf{X}$. From a Kriging point of view, it is the mean squared error. When it is integrated on the domain E, it gives the Integrated Mean Squared Error (IMSE). IMSE can be used as an exploratory criterion for a design. However, it depends on the kernel and it is costly to compute.
For some kernels K defined by radial basis functions, ie $K(\mathbf{x}, \mathbf{x}') = \phi(\mathbf{x} - \mathbf{x}')$ with $\phi : E \to \mathbb{R}$, Schaback (1995) provides the following upper bound on $P_{\mathbf{X}}(\mathbf{x})$:

$$P_{\mathbf{X}}(\mathbf{x}) \leq G_K(h_{\mathbf{X}}) \,.$$

The quantity $h_{\mathbf{X}} = \sup_{\mathbf{y} \in E} \min_{1 \leq i \leq N} \|\mathbf{y} - \mathbf{x}_i\|$ is associated to the MINIMAX criterion. G_K is an increasing function, obviously depending on the kernel. The smoother the kernel K, the faster $G_K(h)$ tends to 0 for $h \xrightarrow{>} 0$. For instance, the gaussian kernel is defined by $K(\mathbf{x}, \mathbf{x}') = e^{-\theta \|\mathbf{x} - \mathbf{x}'\|^2}$ where θ is a real positive parameter; in that case, $G_K(h) = Ce^{-\delta/h^2}$

where C and δ are constants depending on θ. The kernel is not fixed when the design is chosen, the purpose is then to find a design $\mathbf{X}$ with a low $h_{\mathbf{X}}$. That clearly justifies the MINIMAX criterion (4.1). The next proposition ensures a bound on the pointwise interpolation error thanks to a MAXIMIN design (4.2).

Proposition 4.1. *If* $\mathbf{X}$ *is a* MAXIMIN *design,* E *is enclosed in the union of the balls of center* $\mathbf{x}_i$ *and of radius* $\delta_{\mathbf{X}} = \min_{1 \leq i,j \leq N} \|\mathbf{x}_i - \mathbf{x}_j\|$.

Proof

This proposition is proved by contradiction: let $\mathbf{X}$ be a MAXIMIN design and let us suppose that there exists a point $\mathbf{x}_0 \in E$ such that $\|\mathbf{x}_0 - \mathbf{x}_i\| > \delta_{\mathbf{X}}$ for all $\mathbf{x}_i \in \mathbf{X}$.
Let $(\mathbf{x}_{i_0}, \mathbf{x}_{j_0}) \in \mathbf{X}^2$ be a pair of points such that $\|\mathbf{x}_{i_0} - \mathbf{x}_{j_0}\| = \delta_{\mathbf{X}}$ and construct the design $\mathbf{X}' = \{\mathbf{x}_1 \dots \mathbf{x}_{i_0-1}, \mathbf{x}_0, \mathbf{x}_{i_0+1} \dots \mathbf{x}_N\}$ where the point $\mathbf{x}_{i_0}$ is replaced by the point $\mathbf{x}_0$.
$\delta_{\mathbf{X}'} \geq \delta_{\mathbf{X}}$ and, in the case $\delta_{\mathbf{X}'} = \delta_{\mathbf{X}}$, $\mathbf{X}'$ is better than $\mathbf{X}$ with respect to the MAXIMIN criterion because the $\mathbf{X}'$ contains less pairs of points for which the distance is equal to $\delta_{\mathbf{X}}$.
Thus, there is a contradiction because $\mathbf{X}$ is not a MAXIMIN design. Hence, any $\mathbf{x} \in E$ is such that $\|\mathbf{x} - \mathbf{x}_i\| \leq \delta_{\mathbf{X}}$ for all $\mathbf{x}_i \in \mathbf{X}$. □

As a consequence of this proposition, if $\mathbf{X}$ is a MAXIMIN design,

$$|f(\mathbf{x}) - s_{K,\mathbf{X}}(f)(\mathbf{x})| \leq \|f\|_{\mathcal{H}_K} G_K(\delta_X).$$

This result justifies theoretically the use of MAXIMIN designs when a kernel interpolation is used as an approximation of f. Besides it proves that the interpolation done thanks to a MAXIMIN design is consistent.

4.3 Computing maximin designs

In this Section, we propose an algorithm to provide a MAXIMIN design with N points in any set E enclosed in a bounded set. It is based on a simulated annealing method. It aims at finding the global minimum of the function $U : E^N \to \mathbb{R}_+$, $U(\mathbf{X}) = \text{diam}(E) - \delta_{\mathbf{X}}$ where $\text{diam}(E)$ is the diameter of the set E ($\text{diam} = \max_{\mathbf{x},\mathbf{x}' \in E} \|\mathbf{x} - \mathbf{x}'\|$). It is obvious that to minimize U is equivalent to maximize $\delta : \mathbf{X} \mapsto \delta_{\mathbf{X}}$.

The initialization step consists of simulating uniformly a lot of points in the domain E and of calculating the corresponding empirical covariance matrix denoted by Σ. At the end of the initialization step, we randomly keep N points, denoted by $\mathbf{X}^{(0)} = \{\mathbf{x}_1^{(0)}, \dots, \mathbf{x}_N^{(0)}\}$. Then, we propose to iterate the following steps, for $t = 1, \dots$:

Algorithm 4.1.

1. *A pair of points* $(\mathbf{x}_i^{(t)}, \mathbf{x}_j^{(t)})$ *is drawn in* $\mathbf{X}^{(t)}$ *according to a multinomial distribution with probabilities proportional to* $1/(\|\mathbf{x}_i - \mathbf{x}_j\| + \alpha)$ *;*

2. *One of the two points is chosen with probability* $\frac{1}{2}$*, it is denoted by* $\mathbf{x}_k^{(t)}$ *;*

3. *A constraint gaussian random walk is used to propose a new point :*

$$\mathbf{x}_k^{prop} \sim \mathcal{N}_d(\mathbf{x}_k^{(t)}, \tau\Sigma)\mathbb{I}_E(.)\,,$$

   The proposed design is denoted by $\mathbf{X}^{prop} = \{\mathbf{x}_1^{(t)}, \ldots, \mathbf{x}_{k-1}^{(t)}, \mathbf{x}_k^{prop}, \mathbf{x}_{k+1}^{(t)}, \ldots, \mathbf{x}_N^{(t)}\}$*;*

4. $\mathbf{X}^{(t+1)} = \mathbf{X}^{prop}$ *with probability*

$$\min\left(1, \exp\left(-\beta_t(U(\mathbf{X}^{prop}) - U(\mathbf{X}^{(t)}))\right)\frac{q_\tau(\mathbf{X}^{prop}, \mathbf{X}^{(t)})}{q_\tau(\mathbf{X}^{(t)}, \mathbf{X}^{prop})}\right)\,,$$

   otherwise $\mathbf{X}^{(t+1)} = \mathbf{X}^{(t)}$*.*

The idea behind this proposal is to force the pairs of points which are very close to be more distant. $\beta : t \mapsto \beta_t$ is an inverse cooling schedule (ie β_t is an increasing positive sequence and $\lim_{t\to\infty} \beta_t = \infty$) which is chosen in order to ensure the convergence of the algorithm. $q_\tau(\mathbf{X}, .)$ is the probability density function of the proposal kernel $Q_\tau(\mathbf{X}, d\mathbf{Y})$ where $\mathbf{X} \in E^N$ is the current state, $d\mathbf{Y}$ is an infinitesimal neighborhood of the state $\mathbf{Y}$.
τ is a variance parameter which is allowed to change during the iterations but, at each iteration, τ is such that $\tau_0 \geq \tau \geq \tau_{min}$. $\alpha > 0$ is a very small integer which prevents the denominator of $1/(\|\mathbf{x}_i - \mathbf{x}_j\| + \alpha)$ to vanish.

In order to explicit the proposal kernel $Q_\tau(\mathbf{X}, d\mathbf{Y})$, let us introduce some notations:

- $d_{i,j}^{\mathbf{X}} = 1/(\|\mathbf{x}_i - \mathbf{x}_j\| + \alpha)$,
- $D^{\mathbf{X}} = \sum_{k,l:k<l} d_{k,l}^{\mathbf{X}}$,
- $\phi(.|\mu, S)$ denotes the gaussian pdf with mean μ and covariance matrix S,
- $G_{\mu,S} = \int_E \phi(\mathbf{y}|\mu, S)d\mathbf{y}$ denotes the normalization constant associated to $\phi(.|\mu, S)$ on the domain E,
- $\delta_{\mathbf{x}}$ the Dirac mass on $\mathbf{x}$.

The density of the proposal reads as, for $\mathbf{X} \in E^N,\ \mathbf{Y} \in (\mathbb{R}^d)^N$,

$$q_\tau(\mathbf{X}, \mathbf{Y}) = \sum_{i=1}^{N} \phi(\mathbf{y}_i|\mathbf{x}_i, \tau\Sigma))G_{\mathbf{x}_i,\tau\Sigma}^{-1}\left(\sum_{j:j\neq i}\frac{1}{2}\frac{d_{i,j}^{\mathbf{X}}}{D^{\mathbf{X}}}\right)\left(\prod_{j:j\neq i}\delta_{\mathbf{x}_j}(\mathbf{y}_j)\right)\mathbb{I}_{\{\mathbf{y}_i\in E\}}.$$

In order to show the convergence of the previous algorithm, some lemmas are introduced.

Lemma 4.3. *For all* $\mathbf{X} \in E^N$, $q_\tau(\mathbf{X}, .) \geq q_{min} > 0$ *and* $q_\tau(\mathbf{X}, .) \leq q_{max}$, $Q_\tau(\mathbf{X}, .)$*-almost everywhere on* E^N.

Proof

The fact that $q_\tau(\mathbf{X}, .) \leq q_{max}$ is true since the normalization constants are lower-bounded, the gaussian densities are uniformly bounded since $\tau_0 \geq \tau \geq \tau_{min} > 0$ and all the other terms can be upper bounded by 1.
The other assertion is only true $Q_\tau(\mathbf{X}, .)$-almost everywhere on E^N. It means that the lower bound on $q_\tau(\mathbf{X}, \mathbf{Y})$ is given when $\mathbf{X}$ and $\mathbf{Y}$ have at least $N-1$ points in common and are both in E^N.

The following lower bounds are used:

- $G^{-1}_{\mathbf{x}_i, \tau\Sigma} \geq 1$,
- $\sum_{j:j\neq i} \frac{1}{2} \frac{d^{\mathbf{X}}_{i,j}}{D^{\mathbf{X}}} \geq \frac{(\text{diam}(E)+\alpha)^{-1}}{N\alpha^{-1}}$,
- $\phi(\mathbf{y}|\mathbf{x}, \tau\Sigma) \geq \frac{1}{(2\pi)^{d/2}|\Sigma|^{1/2}\tau_0^{d/2}} \exp\left(-\frac{1}{2}\tau_{min}^{-1}\text{diam}(E)^2\xi\right)$ where ξ is the largest eigenvalue of Σ^{-1}.

$q_{min} > 0$ is found by multiplying these expressions and it is a lower bound of $q_\tau(\mathbf{X}, \mathbf{Y})$ which does not depend on τ and on the states if $\mathbf{X} \in E^N$ and $\mathbf{Y} \in E^N$ have at least $N-1$ points in common. □

Let us denote by $(\tau_t)_{t\geq 1}$ the values of τ used during the iterations of the algorithm. This lemma gives that, for a sequence of N proposal kernels $(Q_{\tau_1}, \ldots, Q_{\tau_N})$, it is possible to reach any state $\mathbf{Y} \in E^N$ from any state $\mathbf{X} \in E^N$. Indeed, at each transition the density is lower bounded by q_{min} and at each transition one of the N points is moved. Hence, we get the following lemma:

Lemma 4.4. *If* τ_t *is such that* $\tau_0 \leq \tau_t \leq \tau_{min}$, $\forall t \geq 1$, *there exists* $\epsilon > 0$ *such that for all* $A \in \mathbb{B}(E^N)$ *(Borelian subset of* E^N*), and for all* $\mathbf{X} \in E^N$,

$$(Q_{\tau_1} \cdots Q_{\tau_N})(\mathbf{X}, A) \geq \epsilon\lambda(A)/\lambda(E^N). \tag{4.3}$$

where λ *denotes the Lebesgue measure on the compact set* E^N *(*$\lambda(d\mathbf{X}) = \mathbb{I}_{E^N}(\mathbf{X})Leb(d\mathbf{X})$ *where Leb is the Lebesgue measure on* $(\mathbb{R}^d)^N$*).*

According to the previous comments $\epsilon = q_{min}^N$ suits.

Then, the Hasting-Metropolis (HM) kernel is focused on. It is the global kernel which describes an iteration of the algorithm. It obviously depends on the parameters β and τ. It reads as,

$$K_{\beta,\tau}(\mathbf{X}, d\mathbf{Y}) = a_{\beta,\tau}(\mathbf{X}, \mathbf{Y})Q_\tau(\mathbf{X}, d\mathbf{Y}) + \left(1 - \int_{E^N} a_{\beta,\tau}(\mathbf{X}, \mathbf{Z})Q_\tau(\mathbf{X}, d\mathbf{Z})\right)\delta_{\mathbf{X}}(d\mathbf{Y})$$

where $a_{\beta,\tau}(\mathbf{X}, \mathbf{Y}) = \mathbb{I}_{E^N - \{\mathbf{X}\}}(\mathbf{Y})\left(1 \wedge \frac{\mu_\beta(\mathbf{Y})q_\tau(\mathbf{Y},\mathbf{X})}{\mu_\beta(\mathbf{X})q_\tau(\mathbf{X},\mathbf{Y})}\right)$. μ_β is the target distribution when β is fixed. In a simulated annealing algorithm, the target distribution is the Gibbs measure, ie $\mu_\beta(d\mathbf{X}) = \exp(-\beta U(\mathbf{X}))Z_\beta^{-1}\lambda(d\mathbf{X})$ where $Z_\beta = \int e^{-\beta U(\mathbf{Y})}\lambda(d\mathbf{Y})$.

Lemma 4.5. *μ_β is $K_{\beta,\tau}$-reversible for all τ, β. It implies that μ_β is $K_{\beta,\tau}$-invariant.*

Proof

If $\mathbf{X} \neq \mathbf{Y}$, we have $\mu_\beta(\mathbf{X})q_\tau(\mathbf{X},\mathbf{Y})a_{\beta,\tau}(\mathbf{X},\mathbf{Y}) = \mu_\beta(\mathbf{Y})q_\tau(\mathbf{Y},\mathbf{X})a_{\beta,\tau}(\mathbf{Y},\mathbf{X})$. Indeed, if $\mu_\beta(\mathbf{Y})q_\tau(\mathbf{Y},\mathbf{X}) > \mu_\beta(\mathbf{X})q_\tau(\mathbf{X},\mathbf{Y})$, $a_{\beta,\tau}(\mathbf{X},\mathbf{Y}) = 1$ and $a_{\beta,\tau}(\mathbf{Y},\mathbf{X}) = \frac{\mu_\beta(\mathbf{Y})q_\tau(\mathbf{Y},\mathbf{X})}{\mu_\beta(\mathbf{X})q_\tau(\mathbf{X},\mathbf{Y})}$. The other case is done by symmetry in $\mathbf{X}$ and $\mathbf{Y}$.
Let $\bar{b}_{\beta,\tau}(\mathbf{X}) = 1 - \int_E a_{\beta,\tau}(\mathbf{X},\mathbf{Z})Q_\tau(\mathbf{X},d\mathbf{Z})$,
we have $\mu_\beta(\mathbf{X})\bar{b}_{\beta,\tau}(\mathbf{X})\delta_{\mathbf{X}}(d\mathbf{Y}) = \mu_\beta(\mathbf{Y})\bar{b}_{\beta,\tau}(\mathbf{Y})\delta_{\mathbf{Y}}(d\mathbf{X})$.
Indeed, this measure is non-zero only in the case $\mathbf{X} = \mathbf{Y}$. Therefore,

$$\mu_\beta(d\mathbf{X})K_{\beta,\tau}(\mathbf{X},d\mathbf{Y}) = \mu_\beta(d\mathbf{Y})K_{\beta,\tau}(\mathbf{Y},d\mathbf{X})\,.$$

□

We will show the convergence of our algorithm following the proof given in Bartoli and Del Moral (2001). Some adaptations are necessary since our proposal kernel depends on a variance parameter τ and since there is no reversible measure for the proposal kernel Q_τ. For those reasons, the ratio between the proposal densities has to be in the acceptance rate $a_{\beta,\tau}$ since it makes μ_β $K_{\beta,\tau}$-reversible and then invariant.

In Bartoli and Del Moral (2001), the reversibility of K_β was shown thanks to the reversibility of Q. That is why the ratio of the proposal does not intervene in the acceptance rate of their algorithm.

The next lemma states that when β is large, the target distribution μ_β concentrates on the minima of the function U. $U : E^N \to \mathbb{R}_+$ is lower bounded with respect to λ (the Lebesgue measure on the compact set E^N). We use the following notation
$m = \sup_a[a; \lambda(\{\mathbf{X}; U(\mathbf{X}) < a\}) = 0]$, by definition $\lambda(\{\mathbf{X}; U(\mathbf{X}) < m\}) = 0$.
Moreover, for all $\epsilon > 0$, we define $U_\lambda^\epsilon = \{\mathbf{X}; U(\mathbf{X}) \leq m + \epsilon\}$ which is clearly such that $\lambda(U_\lambda^\epsilon) > 0$ and $U_\lambda^{\epsilon,c} = \{\mathbf{X};\ U(\mathbf{X}) > m + \epsilon\}$.

Lemma 4.6.

$$\forall \epsilon > 0, \quad \lim_{\beta\to\infty} \mu_\beta(U_\lambda^\epsilon) = 1\,.$$

Proof

If $\mathbf{X} \in U_\lambda^\epsilon$, then $e^{-\beta(U(\mathbf{X})-(m+\epsilon))} \geq 1$ and

$$\begin{aligned}
\lambda(e^{-\beta(U-(m+\epsilon))}) &= \int e^{-\beta(U(\mathbf{X})-(m+\epsilon))}\lambda(d\mathbf{X}) \\
&\geq \int \mathbb{I}_{U_\lambda^\epsilon}(\mathbf{X})e^{-\beta(U(\mathbf{X})-(m+\epsilon))}\lambda(d\mathbf{X}) \\
&\geq \lambda(U_\lambda^\epsilon)\,.
\end{aligned}$$

Then,

$$\begin{aligned}\lambda(\mathbb{I}_{U_\lambda^{\epsilon,c}} e^{-\beta(U(\mathbf{X})-(m+\epsilon))}) &= Z_\beta^{-1} \int_{U_\lambda^{\epsilon,c}} e^{-\beta(U(\mathbf{X})} \lambda(d\mathbf{X}) Z_\beta e^{\beta(m+\epsilon)} \\ &= \mu_\beta(U_\lambda^{\epsilon,c}) \int e^{-\beta(U(\mathbf{X})-(m+\epsilon))} \lambda(d\mathbf{X}) \\ &\geq \mu_\beta(U_\lambda^{\epsilon,c}) \lambda(U_\lambda^{\epsilon}).\end{aligned}$$

As a consequence,

$$\mu_\beta(U_\lambda^{\epsilon,c}) \leq \frac{1}{\lambda(U_\lambda^{\epsilon})} \lambda(\mathbb{I}_{U_\lambda^{\epsilon,c}} e^{-\beta(U(\mathbf{X})-(m+\epsilon))}) \,.$$

Dominated convergence Theorem can be applied to the integral on the right-hand side since the function is bounded by 1 which is integrable on the compact set E^N. Thus,

$$\lim_{\beta\to\infty} \lambda\left(\mathbb{I}_{U_\lambda^{\epsilon,c}} e^{-\beta(U(\mathbf{X})-(m+\epsilon))}\right) = 0 \,.$$

And then for any $\epsilon > 0$,

$$\lim_{\beta\to\infty} \mu_\beta(U_\lambda^{\epsilon,c}) = 0 \Rightarrow \lim_{\beta\to\infty} \mu_\beta(U_\lambda^{\epsilon}) = 1.$$

□

The distribution of the Markov chain associated to an inverse cooling schedule $t \mapsto \beta(t)$ and to a variance schedule $t \mapsto \tau(t)$ is denoted η_n. According to the previous results, we have $\eta_{n+1} = \eta_n K_{\beta(n),\tau(n)}$ and $\mu_{\beta(n)} = \mu_{\beta(n)} K_{\beta(n),\tau(n)}$. The aim is to prove that $\lim_{n\to\infty} \|\eta_n - \mu_{\beta(n)}\| = 0$ where $\|.\|$ is the distance in total variation.

Lemma 4.7. *If at each iteration of the algorithm $\tau_0 \geq \tau \geq \tau_{min}$, then $\forall \beta > 0$ and $(\mathbf{X}, A) \in E^N \times \mathbb{B}(E^N)$*

$$K_{\beta,\tau}(\mathbf{X}, A) \geq e^{-\beta osc(U)} \frac{q_{min}}{q_{max}} Q_\tau(\mathbf{X}, A)$$

where $osc(U)$ is the smallest positive number h such that for all $\mathbf{X}$, $\mathbf{Y}$ in E^N, $U(\mathbf{Y}) - U(\mathbf{X}) \leq h$.

Proof

By definition of $osc(U)$, for all $\mathbf{X} \in E^N$ and for all $Q_\tau(\mathbf{X}, .)$-almost everywhere $\mathbf{Y} \in E^N$, the following inequalities hold

$$\mathbb{I}_{E^N - \{\mathbf{X}\}}(\mathbf{Y}) \geq a_{\beta,\tau}(\mathbf{X}, \mathbf{Y}) \geq \mathbb{I}_{E^N - \{\mathbf{X}\}}(\mathbf{Y}) e^{-\beta osc(U)} \frac{q_{min}}{q_{max}} \,.$$

According to the upper-bound of $a_{\beta,\tau}(\mathbf{X}, \mathbf{Y})$, it is shown that

$$\begin{aligned}\left(1 - \int_{E^N} a_{\beta,\tau}(\mathbf{X}, \mathbf{Z}) Q_\tau(\mathbf{X}, d\mathbf{Z})\right) &\geq 1 - Q_\tau(\mathbf{X}, E^N - \{\mathbf{X}\}) = Q_\tau(\mathbf{X}, \{\mathbf{X}\}) \\ &\geq e^{-\beta osc(U)} Q_\tau(\mathbf{X}, \{\mathbf{X}\}) \frac{q_{min}}{q_{max}}.\end{aligned}$$

Thus for $(\mathbf{X}, A) \in E^N \times \mathbb{B}(E^N)$, $K_{\beta,\tau}(\mathbf{X}, A) \geq e^{-\beta osc(U)} \frac{q_{min}}{q_{max}} Q_\tau(\mathbf{X}, A)$. $\square$

By this lemma, for $p \geq 1$, for all non-decreasing sequence $0 \leq \beta_1 \leq \ldots \leq \beta_p$ and for a sequence $(\tau_i)_{1 \leq i \leq p}$ such $\tau_0 \geq \tau_i \geq \tau_{min}$, we have $\forall (\mathbf{X}, A) \in E^N \times \mathbb{B}(E^N)$

$$\begin{aligned}(K_{\beta_1,\tau_1} \cdots K_{\beta_p,\tau_p})(\mathbf{X}, A) &\geq e^{-(\beta_1+\cdots+\beta_p)osc(U)} \left(\frac{q_{min}}{q_{max}}\right)^p (Q_{\tau_1} \cdots Q_{\tau_p})(\mathbf{X}, A) \\ &\geq e^{-p\beta_p osc(U)} \left(\frac{q_{min}}{q_{max}}\right)^p (Q_{\tau_1} \cdots Q_{\tau_p})(\mathbf{X}, A)\end{aligned}$$

For $p = N$ and thanks to the condition (4.3), there is an upper-bound to the Dobrushin coefficient,

$$a(K_{\beta_1,\tau_1} \cdots K_{\beta_N,\tau_N}) \geq \epsilon \left(\frac{q_{min}}{q_{max}}\right)^N e^{-N\beta_N osc(U)}.$$

As a consequence, an application of Dobrushin Theorem states that for all μ_1, μ_2 probability measures on E^N,

$$\begin{aligned}\|\mu_1 K_{\beta_1,\tau_1} \cdots K_{\beta_N,\tau_N} - \mu_2 K_{\beta_1,\tau_1} \cdots K_{\beta_N,\tau_N}\| &\leq \left(1 - \epsilon \left(\frac{q_{min}}{q_{max}}\right)^N e^{-N\beta_N osc(U)}\right) \|\mu_1 - \mu_2\| \\ &= \left(1 - \varepsilon e^{-N\beta_N osc(U)}\right) \|\mu_1 - \mu_2\|\end{aligned}$$

where $\varepsilon = \epsilon \left(\frac{q_{min}}{q_{max}}\right)^N$.

Lemma 4.8. *For all probability measure on E^N and for all function $U : E^N \to \mathbb{R}_+$ such that $\lambda(U) > 0$, these notations are used*

$$\mu_U(d\mathbf{X}) = Z_U^{-1} e^{-U(\mathbf{X})} \lambda(d\mathbf{X}) \quad \textit{where } Z_U = \lambda(e^{-U}).$$

If U_1 and U_2 are two functions such that $\lambda(U_1)$ and $\lambda(U_2) > 0$ then,

$$\|\mu_{U_1} - \mu_{U_2}\| \leq osc(U_1 - U_2).$$

Proof

We have $\mu_{U_1} = Z_{U_1}^{-1} Z_{\frac{U_1+U_2}{2}} e^{\frac{U_2-U_1}{2}} \mu_{\frac{U_1+U_2}{2}}$ and
$Z_{U_1}^{-1} Z_{\frac{U_1+U_2}{2}} = \mu_{U_1}\left(\exp\left(-\frac{U_2-U_1}{2}\right)\right) \geq \exp\left(-\frac{1}{2} osc(U_2 - U_1)\right)$. Therefore,

$$\forall A \in \mathbb{B}(E^N),\ \mu_{U_1}(A) \geq \mu_{\frac{U_1+U_2}{2}}(A) \exp\left(-\frac{1}{2} osc(U_2 - U_1)\right).$$

As $osc(U_2 - U_1) = osc(U_1 - U_2)$,

$$\forall A \in \mathbb{B}(E^N),\ \mu_{U_2}(A) \geq \mu_{\frac{U_1+U_2}{2}}(A) \exp\left(-\frac{1}{2} osc(U_2 - U_1)\right).$$

Thanks to Dobrushin Theorem, we get

$$\|\mu_{U_1} - \mu_{U_2}\| \leq 1 - \exp\left(-\frac{1}{2} osc(U_2 - U_1)\right) \leq osc(U_1 - U_2).$$

□
If this lemma is applied to $U_1 = \beta_1 U$ and $U_2 = \beta_2 U$, $0 < \beta_1 < \beta_2$, then an upper bound is obtained on the Gibbs measures:

$$\|\mu_{\beta_1} - \mu_{\beta_2}\| \leq (\beta_2 - \beta_1) osc(U).$$

The next lemma is useful in order to choose the funtion $n \mapsto \beta_n$.

Lemma 4.9. *Let $I_n,\ a_n,\ b_n,\ n \geq 0$ be three sequences of positive numbers such that $\forall n \geq 1$ $I_n \leq (1-a_n)I_{n-1}+b_n$. If a_n and b_n are such that $\lim_{n\to\infty} \frac{b_n}{a_n} = 0$ and $\lim_{n\to\infty} \prod_{p=1}^{n}(1-a_p) = 0$ then,*

$$\lim_{n\to\infty} I_n = 0\,.$$

Proof

According to the assumptions, for all $\epsilon > 0$ there exists an integer $n(\epsilon) \geq 1$ such that

$$\forall n \geq n(\epsilon), \quad b_n \geq \epsilon a_n, \quad \prod_{p=1}^{n}(1 - a_p) \leq \epsilon\,.$$

As a consequence for all these $n \geq n(\epsilon)$, it holds that

$$\begin{aligned} I_n - \epsilon &\leq (1 - a_n)I_{n-1} - \epsilon(1 - a_n) = (1 - a_n)(I_{n-1} - \epsilon) \\ &\leq \left(\prod_{p=1}^{n}(1 - a_p)\right)(I_0 - \epsilon). \end{aligned}$$

It implies that for all $n \geq n(\epsilon)$,

$$0 \leq I_n \leq \epsilon + \epsilon(I_0 + \epsilon) \leq \epsilon(1 + \epsilon + |I_0|)$$

which ends the proof. □
The convergence Theorem can now be stated.

Theorem 4.1. *If the sequence $(\tau_n)_{n\geq 0}$ is such that $\forall n \geq 0,\ \tau_0 \geq \tau_n \geq \tau_{min} > 0$ and if*

$$\beta_n = \frac{1}{C}\log(n + e), \quad C > N osc(U)\,,$$

we get

$$\forall \epsilon > 0, \quad \lim_{n\to\infty} \mathbb{P}_\eta(\mathbf{X}_n \in U_\lambda^\epsilon) = 1$$

where $U_\lambda^\epsilon = \{\mathbf{X} \in E; U(\mathbf{X}) \leq m + \epsilon\}$ and $\{\mathbf{X}_n; n \geq 0\}$ denotes the random sequence we get from the simulated annealing algorithm with an initial probability distribution η on E^N.

Proof

For any non-decreasing sequence,

$$0 \leq \beta_1 \leq \ldots \leq \beta_p$$

and for every probability distribution η on E^N, it is first noticed that

$$\|\eta K_{\beta_1,\tau_1}\cdots K_{\beta_N,\tau_N} - \mu_{\beta_{N+1}}\| \leq \|\eta K_{\beta_1,\tau_1}\cdots K_{\beta_N,\tau_N} - \mu_{\beta_1}K_{\beta_1,\tau_1}\cdots K_{\beta_N,\tau_N}\| + \|\mu_{\beta_1}K_{\beta_1,\tau_1}\cdots K_{\beta_N,\tau_N} - \mu_{\beta_{N+1}}\|.$$

Thanks to the remark following the lemma 4.7, for $\varepsilon > 0$, it holds that

$$\|\eta K_{\beta_1,\tau_1}\cdots K_{\beta_N,\tau_N} - \mu_{\beta_1}K_{\beta_1,\tau_1}\cdots K_{\beta_N,\tau_N}\| \leq \left(1 - \varepsilon e^{-N\beta_N osc(U)}\right)\|\eta - \mu_{\beta_1}\|. \tag{4.4}$$

For the second term, the following decomposition is used

$$\begin{aligned}\mu_{\beta_1}K_{\beta_1,\tau_1}\cdots K_{\beta_N,\tau_N} - \mu_{\beta_{N+1}} &= \sum_{k=1}^{N}\left(\mu_{\beta_k}K_{\beta_k,\tau_k}\cdots K_{\beta_N,\tau_N} - \mu_{\beta_1}K_{\beta_{k+1},\tau_{k+1}}\cdots K_{\beta_N,\tau_N}\right)\\ &= \sum_{k=1}^{p}(\mu_{\beta_k} - \mu_{\beta_{k+1}})K_{\beta_{k+1},\tau_{k+1}}\cdots K_{\beta_N,\tau_N}\end{aligned}$$

with the convention $K_{\beta_{N+1},\tau_{N+1}}\cdots K_{\beta_N,\tau_N} = Id$. The last equality comes from the equation $\mu_{\beta_k}K_{\beta_k,\tau_k} = \mu_{\beta_k}$ for all $k \geq 1$ given by lemma 4.5. By using the triangular equality, it is deduced that

$$\begin{aligned}\|\mu_{\beta_1}K_{\beta_1,\tau_1}\cdots K_{\beta_N,\tau_N} - \mu_{\beta_{N+1}}\| &\leq \sum_{k=1}^{p}\|\mu_{\beta_k} - \mu_{\beta_{k+1}}\|b(K_{\beta_{k+1},\tau_{k+1}}\cdots K_{\beta_N,\tau_N})\\ &\leq \sum_{k=1}^{p}\|\mu_{\beta_k} - \mu_{\beta_{k+1}}\|,\end{aligned}$$

where b is the contraction coefficient (by Dobrushin Theorem $a(K) + b(K) = 1$).
An application of the lemma 4.8 gives

$$\|\mu_{\beta_1}K_{\beta_1,\tau_1}\cdots K_{\beta_N,\tau_N} - \mu_{\beta_{N+1}}\| \leq osc(U)\sum_{k=1}^{N}(\beta_{k+1} - \beta_k) = (\beta_{N+1} - \beta_1)osc(U). \tag{4.5}$$

By combining (4.4) and (4.5), it is deduced that

$$\|\eta K_{\beta_1,\tau_1}\cdots K_{\beta_N,\tau_N} - \mu_{\beta_{N+1}}\| \leq \left(1 - \varepsilon e^{-N\beta_N osc(U)}\right)\|\eta - \mu_{\beta_1}\| + (\beta_{N+1} - \beta_1)osc(U).$$

Instead of $(\beta_1,\ldots,\beta_N)$ and η, we take $(\beta_{kN},\ldots,\beta_{(k+1)N})$ and η_{kN}:

$$I_{k+1} = \|\eta_{kN}K_{\beta_{kN},\tau_{KN}}\cdots K_{\beta_{(k+1)N},\tau_{(k+1)N}} - \mu_{\beta_{(k+1)N}}\| = \|\eta_{(k+1)N} - \mu_{\beta_{(k+1)N}}\|.$$

By the previous upper bound, the recursive inequalities are stated

$$\begin{aligned}I_{k+1} &\leq \left(1 - \varepsilon e^{-N\beta_{N(k+1)-1}osc(U)}\right)I_k + (\beta_{(k+1)N} - \beta_{kN})osc(U)\\ &\leq \left(1 - \varepsilon((k+1)N + e)^{-N\frac{osc(U)}{C}}\right)I_k + \frac{osc(U)}{C}\log\left(1 + \frac{N+e}{kN+e}\right).\end{aligned}$$

Thanks to the inequality, $\log(1+|x|) \leq |x|$, it holds that $I_{k+1} \leq (1 - a_{k+1})I_k + b_{k+1}$ where $a_{k+1} = \frac{\varepsilon}{((k+1)N+e)^{p\frac{osc(U)}{C}}}$ and $b_{k+1} = \frac{osc(U)}{C}\frac{p+e}{Np+e}$. In order to apply the lemma 4.9, it is to be

checked that if $C > Nosc(U)$ then $\frac{b_{k+1}}{a_{k+1}} = \frac{osc(U)(p+e)}{\varepsilon C}\frac{(k+1)N+e}{kN+e}\frac{1}{((k+1)N+e)^{1-N\frac{osc(U)}{C}}} \to 0$ when $k \to \infty$, and $\prod_{p=1}^{n}(1-a_p) \leq \exp\left(-\sum_{p=1}^{N}\frac{\varepsilon}{(pN+e)^{N\frac{osc(U)}{C}}}\right) \to 0$ when $k \to \infty$.
Hence, $\lim_{k\to\infty}\|\eta_{kN} - \mu_{kN}\| = 0$ and, thus, thanks to lemma 4.6,

$$\forall \epsilon > 0, \quad \lim_{n\to\infty} \eta_n(U_\lambda^\epsilon) = \lim_{n\to\infty} \mathbb{P}_\eta(\mathbf{X}_n \in U_\lambda^\epsilon) = 1\,.$$

□

4.4 Variants of the algorithm

In the case where E is not explicit, the normalization constant $G_{m,S}$ of a gaussian distribution with mean m and covariance matrix S cannot be computed. Hence, the ratio of densities of proposal kernels is not tractable. In that case, we first propose to use as a proposal an unconstrained gaussian random walk. The steps 3 and 4 of Algorithm 4.1 are modified.

Algorithm 4.2. *The first steps until step 3 are the same.*
Step 3 is replaced with

3bis. A gaussian random walk is used to propose a new point :

$$\mathbf{x}_k^{prop} \sim \mathcal{N}_d(\mathbf{x}_k^{(t)}, \tau\Sigma)\,.$$

And step 4 is replaced with

4bis. If $\mathbf{X}^{prop} \in E^N$, $\mathbf{X}^{(t+1)} = \mathbf{X}^{prop}$ *with probability*

$$\min\left(1, \exp\left(-\beta_t(U(\mathbf{X}^{prop} - U(\mathbf{X}^{(t)}))\right)\frac{\tilde{q}_\tau(\mathbf{X}^{prop}, \mathbf{X}^{(t)})}{\tilde{q}_\tau(\mathbf{X}^{(t)}, \mathbf{X}^{prop})}\right)\,,$$

otherwise $\mathbf{X}^{(t+1)} = \mathbf{X}^{(t)}$.

In the last step, $\tilde{q}_\tau(\mathbf{X}, .)$ stands for the density of the proposal kernel where the gaussian random walk is not constraint to remain in the domain E. For any $\mathbf{X} \in E^N$, $\mathbf{Y} \in (\mathbb{R}^d)^N$,

$$\tilde{q}_\tau(\mathbf{X}, \mathbf{Y}) = \sum_{i=1}^{N}\phi(\mathbf{y}_i|\mathbf{x}_i, \tau\Sigma))\left(\sum_{j:j\neq i}\frac{1}{2}\frac{d_{i,j}^{\mathbf{X}}}{D^{\mathbf{X}}}\right)\left(\prod_{j:j\neq i}\delta_{\mathbf{x}_j}(\mathbf{y}_j)\right)\,.$$

Since a lemma similar to lemma 4.3 can be proved for the kernel $\tilde{Q}_\tau$ (corresponding to the density $\tilde{q}_\tau$), theorem 4.1 still applies to it. Hence, there is also a convergence result for Algorithm 4.2.

However, since a point can be proposed outside of the domain E, this algorithm can suffer from a lack of efficiency. Another solution is to use the first algorithm without the ratio of densities of proposal kernels.

Algorithm 4.3. *The first steps until step 4 are the same than in Algorithm 4.1. Step 4 is replaced with*

4ter. $\mathbf{X}^{(t+1)} = \mathbf{X}^{prop}$ *with probability*

$$\min\left(1, \exp\left(-\beta_t(U(\mathbf{X}^{prop} - U(\mathbf{X}^{(t)}))\right)\right),$$

otherwise $\mathbf{X}^{(t+1)} = \mathbf{X}^{(t)}$.

As it is not possible to find a reversible measure for the kernel Q, the previous convergence proof does not apply here.
However, since the best design ever found during the iterations is saved, the following lemma provides a theoretical guarantee for this algorithm.

Lemma 4.10. *For any $\epsilon > 0$, if, $\forall n \in \mathbb{N}$, $\beta_n \leq \frac{1}{C}\log(n+e)$ with $C > Nosc(U)$ the expected time until the first visit in U_λ^ϵ is finite.*

Proof

The expected time until the first visit in U_λ^ϵ is equal to

$$\sum_{k=1}^{\infty} k\mathbb{P}(\mathbf{X}_1, \ldots, \mathbf{X}_k \notin U_\lambda^\epsilon | \mathbf{X}_0 \notin U_\lambda^\epsilon) \times \mathbb{P}(\mathbf{X}_{k+1} \in U_\lambda^\epsilon | \mathbf{X}_0, \ldots, \mathbf{X}_k \notin U_\lambda^\epsilon).$$

The aim is to find an upper bound in order to show that it is finite. The second probability in the argument of the series is limited from above with one. The first probability in the argument of the series is the probability of never visiting U_λ^ϵ in the first k steps.
It can also be written as:

$$\mathbb{P}(\mathbf{X}_1, \ldots, \mathbf{X}_N \notin U_\lambda^\epsilon | \mathbf{X}_0 \notin U_\lambda^\epsilon) \times \cdots \times \mathbb{P}(\mathbf{X}_{\lfloor k/N \rfloor N}, \ldots, \mathbf{X}_k \notin U_\lambda^\epsilon | \mathbf{X}_{\lfloor k/N \rfloor N-1}, \ldots, \mathbf{X}_0 \notin U_\lambda^\epsilon).$$

Thanks to lemmas 4.3 and 4.7, if δ denotes q_{min}, it holds that

$$\mathbb{P}(\text{at least one visit in } U_\lambda^\epsilon \text{ in the first } N\text{steps}) \geq \mathbb{P}(\mathbf{X}_N \in U_\lambda^\epsilon) \geq \delta^N \lambda(U_\lambda^\epsilon) \exp(-\beta_N Nosc(U)).$$

Indeed,

$$\begin{aligned}
\mathbb{P}(\mathbf{X}_N \in U_\lambda^\epsilon) &= (K_{\beta_1,\tau_1} \cdots K_{\beta_N,\tau_N})(\mathbf{X}_0, U_\lambda^\epsilon) \\
&\geq e^{-(\beta_1+\cdots+\beta_N)osc(U)}(Q_{\tau_1} \cdots Q_{\tau_N})(\mathbf{X}_0, U_\lambda^\epsilon) \\
&\geq e^{-N\beta_N osc(U)} q_{min}^N \lambda(U_\lambda^\epsilon)
\end{aligned}$$

Thus,

$$\mathbb{P}(\mathbf{X}_1, \ldots, \mathbf{X}_N \notin U_\lambda^\epsilon | \mathbf{X}_0 \notin U_\lambda^\epsilon) \leq 1 - \delta^N \lambda(U_\lambda^\epsilon) \exp(-\beta_N Nosc(U)).$$

And in a similar way,

$$\mathbb{P}(\mathbf{X}_{iN+1}, \ldots, \mathbf{X}_{(i+1)N} \notin U_\lambda^\epsilon | \mathbf{X}_0, \ldots, \mathbf{X}_{iN} \notin U_\lambda^\epsilon) \leq 1 - \delta^N \lambda(U_\lambda^\epsilon) \exp(-\beta_{N(i+1)} Nosc(U)).$$

Hence, the expected time before the first visit in U_λ^ϵ can be bounded from above by

$$\sum_{k=1}^{\infty} k \prod_{i=1}^{\lfloor k/N \rfloor} \left(1 - \delta^N \lambda(U_\lambda^\epsilon) \exp(-\beta_{N(i+1)} Nosc(U))\right) .$$

As $\log(1-2x) < -x$ if $0 < x < 1/2$, the previous sum is bounded by

$$\sum_{k=1}^{\infty} k \exp\left(- \sum_{i=1}^{\lfloor k/N \rfloor} \frac{\delta^N \lambda(U_\lambda^\epsilon)}{2} \exp(-\beta_{N(i+1)} Nosc(U))\right) .$$

If β_k is chosen such that $\beta(n) = \frac{1}{C}\log(n+e), \quad (C > Nosc(U))$, the sum becomes

$$\sum_{k=1}^{\infty} k \exp\left(-\frac{\delta^N \lambda(U_\lambda^\epsilon)}{2} \sum_{i=1}^{\lfloor k/N \rfloor} \left(\frac{1}{(i+1)N}\right)^{Nosc(u)/C}\right) ,$$

which can be bounded above by

$$\sum_{k=1}^{\infty} k \exp\left(-\frac{\delta^N \lambda(U_\lambda^\epsilon)}{2} \lfloor k/N \rfloor \left(\frac{1}{(k+N)}\right)^{Nosc(u)/C}\right) ,$$

which is a convergent series. □

Since the best design ever found during the iterations is saved, this lemma means that a design reaching a neighborhood U_λ^ϵ of a global maximum of $\delta_{\mathbf{X}}$ can be achieved in a finite number of iterations almost surely. However, this kind of result can be obtained with any algorithm producing a Markov chain which well visits the space of states even if the temperature is fixed.

4.5 Numerical illustrations

The three algorithms are tested on three different toy cases: a design with 100 points in $[0,1]^2$, a design with 250 points in $[0,1]^5$ and a design with 400 points in $[0,1]^8$. In these hypercubic cases, the normalization constants can be computed and Algorithm 4.1 can be used. In each case, 100 calls are made to one million iterations of each algorithm. The inverse cooling schedule is $\beta_n = 1/T_0 \log(n)$ and the variance schedule is $\tau_n = \tau_0/\sqrt{n}$.

In order to choose T_0, a lot of designs with N points can be drawn uniformly in E. Then, a median of $\delta_{\mathbf{X}}$, the minimum distance between pair of points in these designs, is computed. Thus, it is a mean to access to an order of magnitude of $\delta_{\mathbf{X}}$ when $\mathbf{X}$ is uniformly distributed. A fraction of this value is a good choice for T_0 according to our tries. Note that it is much lower than the one required in the convergence theorem.

The parameter τ_0 can be chosen from an analogy with a grid. For example in $[0,1]^2$, a grid of 100 points has 10 points on each line and 10 points on each column, thus it could make sense to divide by 10 the matrix $\hat{\Sigma}$ which is nearly the covariance matrix of an uniform distribution in $[0,1]^2$. As a consequence, τ_0 is taken as $\tau_0 = \mathrm{Vol}(E)/N^{1/d}$ where $\mathrm{Vol}(E)$ is the volume of E or an upper bound of this volume.

Figures 4.1, 4.2 and 4.3 present the results. For each algorithm, it is given the boxplots of the best solutions to the maximization of $\delta_{\mathbf{X}}$ over one million iterations (boxplots are

constructed using 100 replicates). Algorithms 4.1 and 4.3 give the best results. Algorithm 4.2 suffers from the fact that the proposal can be outside of the domain.

Other cooling schedules than the ones which have theoretical guarantees can be tried. It seems that they can lead to satisfying results which are even better than the ones obtained with the log schedule. Since the results depend too much on the examples, it is quite hard to state a general rule. However, a schedule $\beta_n = 1/T_0\sqrt{n}$ is robust to a bad choice in T_0 and a schedule $\tau_n = \tau_0/\sqrt{n}$ performs quite well.

4.6 Application to a simulator of an aircraft engine

The behaviour of an aircraft engine is described by a numerical code. A run of the code determines if the given flight conditions are acceptable and, provided they are, computes the corresponding outputs. The function which associates the outputs to the flight conditions is denoted by f. It is accessible only through runs of the code. It is a black box function and a run is quite burdensome. We have to compute an approximation of f. The acceptable flight conditions represent the domain of definition of f, denoted by E. Outside E, the code cannot provide outputs since the conditions are physically impossible or the code encounters convergence failures. E is not explicit, as explained above we have to run the code to know if the flight conditions are acceptable. Therefore, we need to estimate E (the indicator function associated to E). This is not our goal here. E is included in a known hypercube (lower and upper bounds are available on each of these variables). Using other prior information and some calls to f, a binary classification tree has been built to determine an estimate of E. This method works quite well and leads to a misclassification error rate around 0.5%. The resulting domain is not an hypercube.

In the following case study, only the flow rate output is focused on. The flight conditions are described by ten variables such as altitude, speed, temperature, humidity... A variable selection procedure has shown that only $d = 8$ input variables are useful for prediction. Hence, the considered function to be approximated is $f : E \subset \mathbb{R}^d \to \mathbb{R}$.

A MAXIMIN design is drawn thanks to 10^7 iterations of Algorithm 4.3. The initial temperature T_0 and the initial variance τ_0 were chosen as described in the previous section. The inverse cooling shedule was $\beta_n = 1/T_0\sqrt{n}$ and the variance schedule was constant during the first quarter of iterations and then $\tau_n = \tau_0/\sqrt{n - 10^7/4}$.

Approximations of the function f are made by kernel interpolations on three different designs: the MAXIMIN design that was computed, a design whose points follow an uniform distribution on E and a desing which is obtained by truncating a Latin hypercube design defined on the hypercube domain containing E. The kernel interpolations are computed by the Matlab toolbox DACE (Lophaven et al., 2002). The regression functions are chosen as the polynomials with degree smaller than or equal to two and the kernel is a generalized exponential kernel:

$$K(\mathbf{x}, \mathbf{x}') = \exp\left(-\sum_{j=1}^{d} \theta_j |x^{(j)} - x'^{(j)}|^{\nu}\right),$$

where $x^{(j)}, x'^{(j)}, j = 1, \ldots, d$ are respectively the j^{th} coordinates of $\mathbf{x}, \mathbf{x}'$ and $\theta_1, \ldots, \theta_d, \nu$ are parameters which are estimated using the usual maximum likelihood estimators. The three designs are sets of $1,300$ points which are included in the domain E according to the estimated indicator function. The function f is computed at the points of the designs. Some points

have to be removed from the designs since the code indicates that they are not in E (recall that the designs were built thanks to an estimate of E).

Table 4.1 provides the performances of kernel interpolations according to the designs. The performances are evaluated on another set whose the $1,300$ points generated according an uniform distribution on E and on which the function f is also computed. If $\hat{f}$ denotes a kernel interpolator and $\{\mathbf{z}_1, \ldots, \mathbf{z}_{1300}\}$ is the set of test points, those quantities are reported:

- the Mean Relative Error (MRE),

$$\frac{1}{1300} \sum_{i=1}^{1300} \left| \frac{f(\mathbf{z}_i) - \hat{f}(\mathbf{z}_i)}{f(\mathbf{z}_i)} \right| ,$$

- the Maximum Relative Error (MaxRE),

$$\max_{i=1,\ldots,1300} \left| \frac{f(\mathbf{z}_i) - \hat{f}(\mathbf{z}_i)}{f(\mathbf{z}_i)} \right| ,$$

- the Mean Squared Error (MSE),

$$\frac{1}{1300} \sum_{i=1}^{1300} \Big(f(\mathbf{z}_i) - \hat{f}(\mathbf{z}_i) \Big)^2 .$$

Table 4.1 also contains the number of points which are actually in E and the minimal distance $\delta_{\mathbf{X}}$ between the pairs of points of the designs. To compute these distances, the designs were translated into the hypercube $[0, 1]^8$.

	mRE	MaxRE	MSE	Nb of Points	$\delta_{\mathbf{X}}$
Uniform	0.49%	5.2%	0.63	1284	0.15
LHS	0.48%	6.9%	0.73	1275	0.14
MAXIMIN	0.47%	3.5%	0.56	1249	0.33

Table 4.1: Comparison of performances of kernel interpolation on the different designs

The MAXIMIN design makes the kernel interpolation more efficient especially according to the MaxRE criterion. As it was shown, the kernel interpolation accuracy depends sharply on the spreading out of the points of the design. Thus, the MAXIMIN design which ensures that any point of E is not far from the points of the design leads to the best performances.

Acknowledgements

The authors are grateful to Pierre Del Moral for very helpful discussions on the convergence properties of the algorithms. This work has been supported by the Agence Nationale de la Recherche (ANR, 212, rue de Bercy 75012 Paris) through the 2009-2012 project Big'MC.

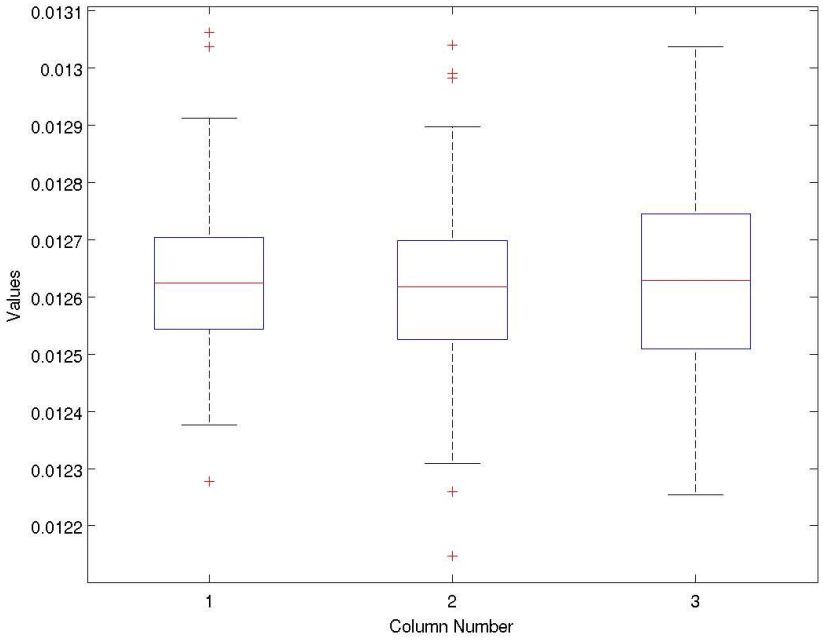


Figure 4.1: Case of a design of 100 points in $[0,1]^2$

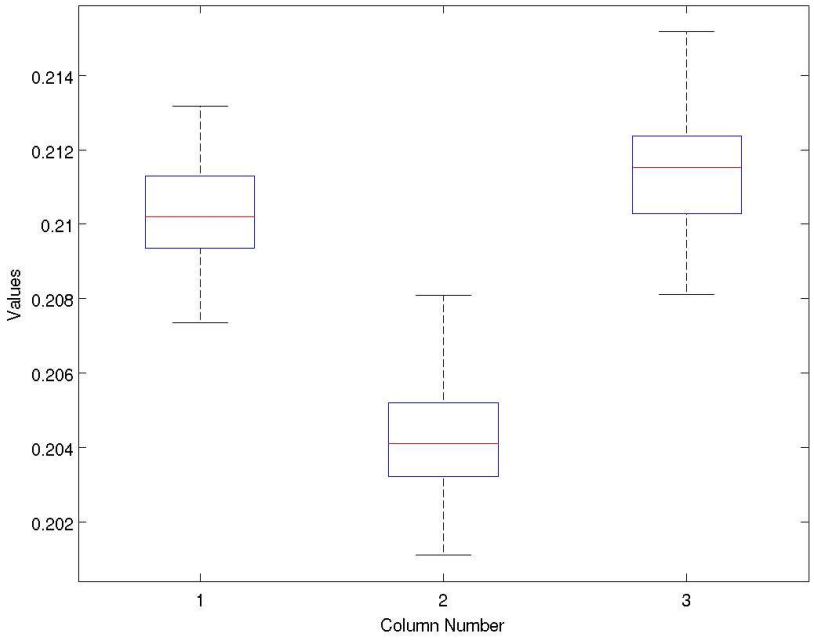


Figure 4.2: Case of a design of 250 points in $[0,1]^5$

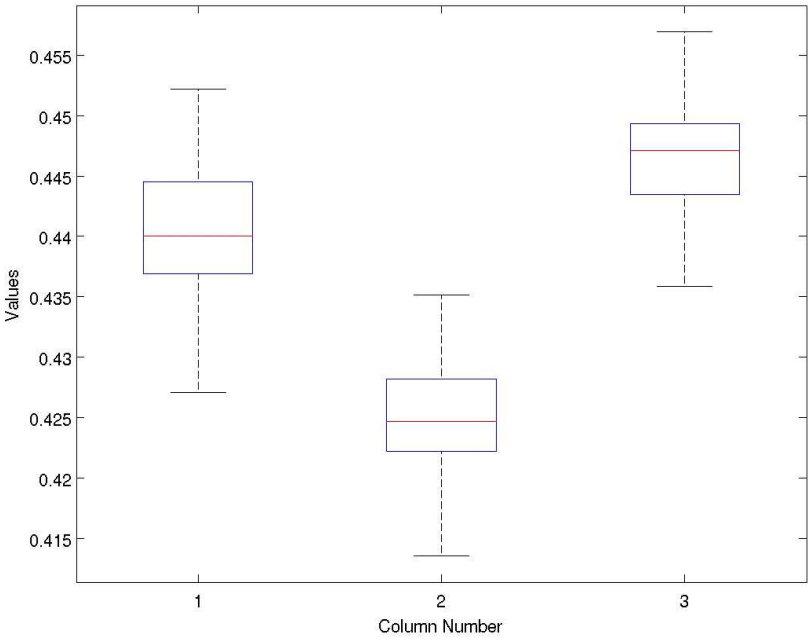


Figure 4.3: Case of a design of 400 points in $[0,1]^8$

Bibliography

Bartoli, N. and Del Moral, P. (2001). *Simulation & algorithmes stochastiques*. Cépaduès.

Bursztyn, D. and Steinberg, D. M. (2006). Comparison of designs for computer experiments. *J. Statist. Plann. Inference*, 136(3):1103–1119.

Chib, S. and Greenberg, E. (1995). Understanding the metropolis-hastings algorithm. *The American Statistician*, 49(4):327–335.

Cressie, N. (1993). *Statistics for Spatial Data*. Wiley, New York.

den Hertog, D., Kleijnen, J. P. C., and Siem, A. Y. D. (2006). The correct kriging variance estimated by bootstrapping. *Journal of the Operational Research Society*, 57(4):400–409.

Fang, K.-T., Li, R., and Sudjianto, A. (2006). *Design and Modeling for Computer Experiments*. Computer Science and Data Analysis. Chapman & Hall/CRC.

Joseph, V. R. (2006). Limit kriging. *Technometrics*, 48(4):458–466.

Koehler, J. R. and Owen, A. B. (1996). Computer experiments. In *Design and analysis of experiments*, volume 13 of *Handbook of Statistics*, pages 261–308. North Holland, Amsterdam.

Laslett, G. M. (1994). Kriging and splines: an empirical comparison of their predictive performance in some applications. *J. Amer. Statist. Assoc.*, 89(426):391–409. With comments and a rejoinder by the author.

Li, R. and Sudjianto, A. (2005). Analysis of computer experiments using penalized likelihood in gaussian kriging models. *Technometrics*, 47:111–120.

Lophaven, N., Nielsen, H., and Sondergaard, J. (2002). Dace, a matlab kriging toolbox. Technical Report IMM-TR-2002-12, DTU. Available to : http://www2.imm.dtu.dk/ hbn/dace/dace.pdf.

Madych, W. R. and Nelson, S. A. (1992). Bounds on multivariate polynomials and exponential error estimates for multiquadric interpolation. *Journal of Approximation Theory*, pages 94–114.

Matheron, G. (1963). Principles of geostatistics. *Economic Geology*, 58:1246–1266.

McKay, M. D., Beckman, R. J., and Conover, W. J. (1979). A comparison of three methods for selecting values of input variables in the analysis of output from a computer code. *Technometrics*, 21(2):239–245.

Morris, M. D. and Mitchell, T. J. (1995). Exploratory designs for computer experiments. *Journal of Statistical Planning and Inference*, 43:381–402.

Sacks, J., Schiller, S. B., Mitchell, T. J., and Wynn, H. P. (1989). Design and analysis of computer experiments (with discussion). *Statistica Sinica*, 4:409–435.

Santner, T. J., B., W., and W., N. (2003). *The Design and Analysis of Computer Experiments*. Springer-Verlag.

Schaback, R. (1995). Error estimates and condition numbers for radial basis function interpolation. *Advances in Computational Mathematics*, 3:251–264.

Schaback, R. (2007). Kernel-based meshless methods. Technical report, Institute for Numerical and Applied Mathematics, Georg-August-University Goettingen.

Stein, M. L. (1999). *Interpolation of Spatial Data: Some Theory for Kriging*. Springer, New York.

Stein, M. L. (2002). The screening effect in kriging. *Ann. Statist.*, 30(1):298–323.

Stinstra, E., den Hertog, D., Stehouwer, P., and Vestjens, A. (2003). Constrained maximin designs for computer experiments. *Technometrics*, 45(4):340–346.

van Dam, E. R., Husslage, B., den Hertog, D., and Melissen, H. (2007). Maximin Latin Hypercube Designs in Two Dimensions. *Operations Research*, 55:158–169.

Chapitre 5

Non linear methods for inverse statistical problems

Résumé

Nous proposons une application des métamodèles de krigeage à un problème statistique inverse où il s'agit d'appréhender les incertitudes sur les entrées d'un modèle physique. Elles sont modélisées par une loi de probabilité multivariée et représentent la variabilité intrinsèque des entrées. Il s'agit d'identifier cette loi à partir d'observations des sorties du modèle. Afin de se limiter à un nombre d'appels raisonnable au code de calcul (souvent coûteux) du modèle physique dans l'algorithme d'inversion, une méthodologie faisant intervenir une approximation non linéaire par un métamodèle de krigeage et un algorithme EM stochastique est présentée. Elle est comparée à une méthode utilisant une approximation linéaire itérative sur la base de jeux de données simulées provenant d'un modèle de crues simplifié mais réaliste. Les cas où cette approche non linéaire est préférable seront mis en lumière.

Mots clés : Modélisation des incertitudes, Approximation non linéaire, Krigeage, Algorithme stochastique.

Ce chapitre est issu d'une collaboration avec Agnès Grimaud, Gilles Celeux, Yannick Lefebvre et Étienne de Rocquigny. Il paraitra au format article dans Computational Statistics and Data Analysis 55 (2011) p 132-142.

Abstract

In the uncertainty treatment framework considered, the intrinsic variability of the inputs of a physical simulation model is modelled by a multivariate probability distribution. The objective is to identify this probability distribution - the dispersion of which is independent of the sample size since intrinsic variability is at stake - based on observation of some model outputs. Moreover, in order to limit to a reasonable level the number of (usually burdensome) physical model runs inside the inversion algorithm, a non linear approximation methodology making use of Kriging and stochastic EM algorithm is presented. It is compared with iterated linear approximation on the basis of numerical experiments on simulated data sets coming from a simplified but realistic modelling of a dyke overflow. Situations where this non linear approach is to be preferred to linearisation are highlighted.

Keywords: Uncertainty Modelling, Non linear Approximation, Kriging, Stochastic Algorithm.

5.1 Introduction

Probabilistic uncertainty treatment is gaining fast growing interest in the industrial field, as reviewed by Rocquigny (de). In the energy sector, such uncertainty analyses are for instance carried out in environmental studies (flood protection, effluent control, etc.), or in nuclear safety studies involving large scientific computing (thermo-hydraulics, mechanics, neutronics etc.). Besides the uncertainty propagation challenges when dealing with complex and high CPU-time demanding physical models, one of the key issues regards the quantification of the sources of uncertainties. The problem is to choose reliable statistical models for the input

variables such as uncertain physical properties of the materials or industrial process or natural random phenomena (wind, flood, temperature, etc.).

A key difficulty, traditionally encountered at this stage, is linked to the highly-limited sampling information directly available on uncertain input variables. An industrial case-study can largely benefit from integrate indirect information such as data on other more easily observable parameters linked to the uncertain variable of interest by a physical model. It demands methods using of probabilistic inverse methods since the recovering of indirect information involves generally the inversion of a physical model. Roughly speaking, this inversion transforms the information into a virtual sample of the variable of interest, before applying to it standard statistical estimation. Yet, it is mandatory to limit to a reasonable level the number of (usually large CPU-time consuming) physical model runs inside the inverse algorithms.

As in Celeux et al. (2010), this paper concentrates on the situation where there is an *irreducible* uncertainty or variability in the input parameters of a physical model. Observations are modelled with a vector of physical variables y that are connected to uncertain inputs x through a deterministic (and supposedly well-known) physical model $y = H(x, d)$. As a clear difference to classical parameter identification x is not supposed to have a fixed, albeit unknown physical value: It will be modelled by a random variable taking different realisations for each observation. The purpose is thus to estimate its probability distribution function instead of its point value. On the other hand, d stands for fixed inputs. A key difficulty is that the time needed to compute the physical function H is huge since H is often the result of a complex code. Thus, it is desirable or necessary to limit the number of calls to the H function. For this very reason, Celeux et al. (2010) investigated efficient estimation algorithms based on a linearisation of the model around a fixed value x_0 to estimate the parameters distributions in this context. But, the linearisation method has some drawbacks associated to the approximation error induced and to the potential difficulty in choosing an adequate linearisation point before identification. In this paper, we propose an alternative solution avoiding the linearisation of H by using a non linear approximation of the function H obtained through Kriging. The paper is organised as follows. In Section 5.2, the model is stated and the linear procedure of Celeux et al. (2010) is summarised. In Section 5.3, a stochastic procedure using a non linear approximation of H is presented. Section 5.4 is devoted to the presentation of numerical experiments for comparing the two approaches. A short discussion section ends the paper.

5.2 The model and its linear identification

The considered model takes the form

$$Y_i = H(X_i, d_i) + U_i, \quad 1 \leq i \leq n\,, \tag{5.1}$$

with the following features

- (Y_i) in $\mathbb{R}^p$ denotes the vector data,
- H denotes a known function from $\mathbb{R}^{(q+q_2)}$ to $\mathbb{R}^p$. The function H can be typically regarded as a "black box" and getting the output $H(x, d)$ from any input (x, d) is quite expensive. To ensure the identifiability of model (5.1), H is assumed to be injective.

- (X_i) in $\mathbb{R}^q$ denotes non observed random data, assumed to be independent and identically distributed (i.i.d.) with a Gaussian distribution $\mathcal{N}(\mu, C)$.
- (d_i) denotes observed variables related to the experimental conditions, with dimension q_2.
- (U_i) denotes measurement-model errors, assumed i.i.d. with distribution $\mathcal{N}(0, R)$, R being known or unknown. Variables (X_i) and (U_i) are assumed to be independent.

The aim is to estimate the parameters (μ, C, R) from the data $(Y_i, d_i), i = 1, \ldots, n$. Since the (X_i) are not observed, a good estimation of this missing structure model would require to compute the function H a lot of times. But, as written above, computing values of H is quite expensive.

A linearised method In order to limit to a reasonable amount of computation the number of calls to the function H to estimate the model parameters, a linear approximation of the model defined in (5.1) has been investigated in Celeux et al. (2010). In this approach the function H is linearised around a fixed value x_0 (chosen from expert informations). The approximated model is

$$Y_i = H(x_0, d_i) + J_H(x_0, d_i)(X_i - x_0) + U_i, \quad 1 \leq i \leq n\,, \tag{5.2}$$

where $J_H(x_0, d_i)$ is the Jacobian matrix of the function H in x_0, with dimension $p \times q$.

In the following, for simplicity, the variance matrix R is assumed to be known to sketch the approach of Celeux et al. (2010). First the linear model (5.2) is supposed to be identifiable. It is ensured if and only if $\text{rank}(\mathbf{J_H}) = q$ with $\mathbf{J_H} = (J_H(x_0, d_1), ..., J_H(x_0, d_n))^T$.

The data (X_i) being non observed, the estimation problem is a missing data structure problem that can be solved with an EM-type algorithm (Dempster et al., 1977). The EM algorithm alternates two steps at iteration $(k + 1)$:

- E step (Expectation): It consists of computing $Q(\theta, \theta^{(k)}) = E[L(\theta, \mathbf{Z})|\mathbf{Y}, \theta^{(k)}]$ where L is the completed loglikelihood.
- M step (Maximisation): $\theta^{(k+1)} = \arg\max_{\theta \in \Theta} Q(\theta, \theta^{(k)})$.

In the present context updating formulas for $\mu^{(k)}$ and $C^{(k)}$ in the M step are closed form.

A variant devoted to accelerate the EM algorithm, which is known to often encounter slow convergence situations, is the ECME (Expectation-Conditional Maximisation Either) algorithm of Liu and Rubin (1994). The M-step is replaced by CME-steps (Conditional Maximisation Either), maximising conditionally to some parameters, the Q-function or the actual observed loglikelihood, $\ln(L(\theta))$.

To compute $\theta^{(k+1)} = (\mu^{(k+1)}, C^{(k+1)})$ for model (5.2), the iteration $(k + 1)$ of ECME is as follows: the E-step is the same as in EM and the M-step is replaced with two steps. The first CME step, to update the variance matrix C, is similar to the M step of EM with μ fixed to $\mu^{(k)}$. The second CME step, to update the parameter μ, maximises the incomplete-data loglikelihood over μ, assuming $C = C^{(k+1)}$ (see also De Crecy, 1996). Introducing the notation: $h_i = H(x_0, d_i)$; $J_i = J_H(x_0, d_i)$,

$$A_i^{(k)} = Y_i - h_i - J_i(\mu^{(k)} - x_0), B_i^{(k)} = C^{(k)} J_i^T \text{ and } V_i^{(k)} = J_i C^{(k)} J_i^T + R\,,$$

the ECME updating equations for model (5.2) are

$$C^{(k+1)} = C^{(k)} + \frac{1}{n}\sum_{i=1}^{n}\left[(B_i^{(k)}(V_i^{(k)})^{-1}A_i^{(k)})(B_i^{(k)}(V_i^{(k)})^{-1}A_i^{(k)})^T - B_i^{(k)}(V_i^{(k)})^{-1}(B_i^{(k)})^T\right],$$

$$\mu^{(k+1)} - x_0 = \left(\sum_{i=1}^{n} J_i^T(V_i^{(k+1)})^{-1}J_i\right)^{-1}\left(\sum_{i=1}^{n} J_i^T(V_i^{(k+1)})^{-1}(Y_i - h_i)\right).$$

The EM and ECME algorithms have shown to work well in practice (Celeux et al., 2010). But the linearisation approach could be sensitive to the linearisation point x_0. To reduce its influence, a simple solution is to use an iterative linearisation of the physical model H, as now described:

- Initial Step: Starting from an initial linearisation point: $x_{\text{lin}} = x_0$; $(H(x_0, d_i))_i$ and $(J_H(x_0, d_i))_i$ are computed. Then the ECME algorithm, initiated at $\theta_{init} = (x_0, C_0)$, is run leading to the estimate $\hat{\theta}^{(1)}$.

- Step $l+1$: Let $x_{\text{lin}} = \hat{\mu}^{(l)}$. Then $(H(x_{\text{lin}}, d_i))$ and $(J_H(x_{\text{lin}}, d_i))$ are computed and the ECME algorithm initiated with $\theta_{init} = \hat{\theta}^{(l)}$, leads to the estimate $\hat{\theta}^{(l+1)}$.

This algorithm is run until some stopping criterion, as $\max_j \left(\frac{|\theta_j^{(l+1)} - \theta_j^{(l)}|}{|\theta_j^{(l)}|}\right) \leq \varepsilon$ with some fixed ε, is satisfied.

Remark: In the general case where the experimental conditions d_i vary throughout the sample, changing the linearisation point requires n calls of H for $H(x_{\text{lin}}, d_i)$ plus $n \times q \times a$ calls of H for $J_H(x_{\text{lin}}, d_i)$ through finite differences where $a = 1$ to say $a = 5$ according to the roughness of H. This iterate linearisation is expected to perform well when the function H is not highly non linear. Otherwise alternative non linear approximations of H could be required.

5.3 Using a non linear approximation of the function H

In some cases, linear approximation of the function H could be unsatisfactory. But in such cases, the E and M steps in EM and ECME algorithms are difficult to implement. For instance, the conditional expectation function Q is not closed form. A possible answer is to use a stochastic version of the EM algorithm such as the SEM algorithm (Celeux and Diebolt, 1985, 1987) or the SAEM algorithm (Delyon et al., 1999). However these algorithms which require to simulate the missing x_i according to their current conditional distribution at each iteration, need to call H some thousand times which is far too CPU time consuming. In practice, to save CPU running time, the number of calls to the function H is constrained to be smaller than a maximum value $N_{\max}$. Therefore, we propose a method coupling the SEM algorithm with a non linear approximation of H. Its principle is as follows: A set of points $D = \{(x_1, d_1), ..., (x_{N_{\max}}, d_{N_{\max}})\}$ with size $N_{\max}$ is chosen. Then H is computed at each point of D and will be not called again in the algorithm. Whenever H has to be computed at a point (x, d), the true value $H(x, d)$ is replaced by an approximation $\hat{H}(x, d)$, obtained with a barycentric interpolation or Kriging.

The considered model is the model (5.1):

$$Y_i = H(X_i, d_i) + U_i, \quad 1 \leq i \leq n.$$

In this section, the variance matrix R of the measurement model error can be assumed known or not. The aim is to estimate the parameter $\theta = (\mu, C, (R))$.

5.3.1 The SEM algorithm

The Stochastic EM (SEM) algorithm incorporates a simulation step between the E and M steps. Its $(k+1)$th iteration involves three steps:

- E step: Computation of the conditional density $p(.|\mathbf{Y}; \theta^{(k)})$ of $\mathbf{X}^{(k)}$, $\theta^{(k)}$ being the current fit of parameter θ.
- S step (Stochastic): It is a Restoration step: a completed sample $\mathbf{Z}^{(k)} = (\mathbf{Y}, \mathbf{X}^{(k)})$ is generated by drawing $\mathbf{X}^{(k)}$ from the conditional density $p(.|\mathbf{Y}; \theta^{(k)})$.
- M step: The updated estimate $\theta^{(k+1)}$ is the maximum likelihood estimate computed on the basis of $\mathbf{Z}^{(k)}$.

This SEM algorithm generates an irreducible Markov chain whose stationary distribution is concentrated around maximum likelihood estimate of θ (see Nielsen, 2000). To derive pointwise estimates from SEM, a warm-up step of length ℓ is required to reach the stationary regime of the generated Markov chain, then mean $\sum_{k=\ell+1}^{L} \theta^{(k)}$ is computed with L large enough to get an estimate of θ.

The SEM algorithm is now described for the model (5.1). The first task is to calculate the completed loglikelihood $L(\theta, \mathbf{Z}) = \ln p(\mathbf{Y}, \mathbf{X}; \theta)$: We have $p(\mathbf{Y}, \mathbf{X}; \theta) = p(\mathbf{Y}|\mathbf{X}, \theta) p(\mathbf{X}; \theta)$ with

$$p(\mathbf{Y}|\mathbf{X}, \theta) = (2\pi)^{-\frac{nd}{2}} |R|^{-\frac{n}{2}} \exp\left(-\frac{1}{2} \sum_{i=1}^{n} (Y_i - H(X_i, d_i))^T R^{-1} (Y_i - H(X_i, d_i)) \right)$$

and

$$p(\mathbf{X}; \theta) = (2\pi)^{-\frac{nq}{2}} |C|^{-\frac{n}{2}} \exp\left(-\frac{1}{2} \sum_{i=1}^{n} (X_i - \mu)^T C^{-1} (X_i - \mu) \right).$$

Thus

$$\ln p(\mathbf{Y}, \mathbf{X}; \theta) = -\frac{n}{2} \ln(|R|) - \frac{1}{2} \sum_{i=1}^{n} (Y_i - H(X_i, d_i))^T R^{-1} (Y_i - H(X_i, d_i))$$

$$-\frac{n}{2} \ln(|C|) - \frac{1}{2} \sum_{i=1}^{n} (X_i - \mu)^T C^{-1} (X_i - \mu) + Cst\,.$$

And, $\theta^{(k+1)}$ is obtained by solving the likelihood equations

$$\frac{\partial}{\partial R} \ln p(\mathbf{Y}, \mathbf{X}^{(k)}; \theta) = \frac{\partial}{\partial \mu} \ln p(\mathbf{Y}, \mathbf{X}^{(k)}; \theta) = \frac{\partial}{\partial C} \ln p(\mathbf{Y}, \mathbf{X}^{(k)}; \theta) = 0.$$

This leads to the closed form formulas

$$R^{(k+1)} = \frac{1}{n}\sum_{i=1}^{n}(Y_i - H(X_i^{(k)}, d_i))(Y_i - H(X_i^{(k)}, d_i))^T ,$$

$$\mu^{(k+1)} = \frac{1}{n}\sum_{i=1}^{n} X_i^{(k)}$$

and

$$C^{(k+1)} = \frac{1}{n}\sum_{i=1}^{n}(X_i^{(k)} - \mu^{(k+1)})(X_i^{(k)} - \mu^{(k+1)})^T .$$

For model (5.1), the simulation step of SEM induces a difficulty since the conditional distribution of $(\mathbf{X}|\mathbf{Y}, \theta)$ is not directly available. A MCMC (Markov Chain Monte Carlo) algorithm is needed to perform the S step. At iteration k, the S step consists of m iterations of a Hastings-Metropolis algorithm. For $i = 1, ..., n$

- Let $X_{i,0} = X_i^{(k-1)}$.
- For $s = 1, ..., m$
  1. Generate $\tilde{X}_{i,s}$ from the proposal distribution $q_{\theta_k}(X_{i,s-1}, .)$.
  2. $X_{i,s} = \tilde{X}_{i,s}$ with probability

$$\alpha(X_{i,s-1}, \tilde{X}_{i,s}) = \min\left(1, \frac{p(\tilde{X}_{i,s}|Y_i; \theta^{(k)}) q_{\theta_k}(\tilde{X}_{i,s}, X_{i,s-1})}{p(X_{i,s-1}|Y_i; \theta^{(k)}) q_{\theta_k}(X_{i,s-1}, \tilde{X}_{i,s})}\right)$$

  and $X_{i,s} = X_{i,s-1}$ with probability $1 - \alpha(X_{i,s-1}, \tilde{X}_{i,s})$.
- $X_i^{(k)} = X_{i,m}$.

Several proposal distributions taking into account assumptions made on the (X_i) distribution may be used. Here three proposals are alternately considered (see Kuhn and Lavielle, 2004):

1. q_{θ_k} is the "prior" distribution of X_i at iteration k, that is the Gaussian distribution $\mathcal{N}(\mu_k, C_k)$. Then

$$\alpha(X_{i,s-1}, \tilde{X}_{i,s}) = \min\left(1, \frac{p(Y_i|\tilde{X}_{i,s}; \theta^{(k)})}{p(Y_i|X_{i,s-1}; \theta^{(k)})}\right) .$$

2. q_{θ_k} is the multidimensional random walk with dimension q: $\mathcal{N}(X_{i,s-1}, \kappa C_k)$. Then

$$\alpha(X_{i,s-1}, \tilde{X}_{i,s}) = \min\left(1, \frac{p(Y_i, \tilde{X}_{i,s}; \theta^{(k)})}{p(Y_i, X_{i,s-1}; \theta^{(k)})}\right) .$$

3. q_{θ_k} is the succession of q unidimensional Gaussian random walks $\mathcal{N}(X_{i,s-1}(l), \kappa C_k(l, l))$: each component of X is successively updated.

At iteration k, the S step consists of running m_1 iterations with proposal 1, m_2 iterations with proposal 2 and m_3 iterations with proposal 3, with $m_1 + m_2 + m_3 = m$. In proposals 2 and 3, κ has to be chosen between 0 and 1. It is tuned so that the first iterations of the S step have acceptance rates between 0.3 and 0.6 to ensure that the Hasting-Metropolis chain well explores the possible values of X_i. In the following simulations, $\kappa = 0.1$ suits and $m_1 = 100$, $m_2 = 0$ and $m_3 = 100$ are set.

To compute the acceptance probabilities α, the function H is called m times, for each i. Hence for each iteration of the SEM algorithm, H is to be computed nm times. But recall that the number of calls to H is limited to at most $N_{\max}$. It means that in most situations the above described SEM algorithm is infeasible. To cope with this difficulty, we propose to first compute H on a set of $N_{\max}$ points. Then, H is replaced by an approximation $\hat{H}$ built from the $N_{\max}$ evaluations, in the SEM algorithm.

5.3.2 SEM with Kriging approximation of H

In this section, $H(z)$ could denote $H_i(z) = H(z, d_i)$ where $z \in \mathbb{R}^q$ as well as $H(z) = H(z_1, z_2)$ where $(z_1, z_2) \in \mathbb{R}^q \times \mathbb{R}^{q_2}$. That is to say that an approximation is made for each d_i (thus for each H_i) or a single approximation of H is made. This point is further discussed in the Remark (iii) of Section 5.3.2. It is considered that $z \in \mathbb{R}^Q$ where $Q = q$ or $Q = q + q_2$.

The approximation $\hat{H}$ could be a barycentric approximation derived from $N_{\max}$ exact values $H(z_1), \ldots, H(z_{N_{\max}})$ of H. The approximation is, for $z \notin D = \{z_1, \ldots, z_{N_{\max}}\}$

$$\hat{H}(z) = \sum_{j \in V_k(z)} \frac{\|z_j - z\|^{-1}}{\sum_j \|z_j - z\|^{-1}} H(z_j)\,,$$

where $V_k(z)$ is the subset of the k nearest neighbours of z in D, for a fixed k. Preliminary numerical experiments (not reported here) show that this simple barycentric method could be not efficient enough and that Kriging, that is now described, is to be preferred.

Kriging (see Currin et al., 1991; Koehler and Owen, 1996) is a method devoted to approximate a function $H : \Omega \mapsto \mathbb{R}$ where the input set $\Omega \subset \mathbb{R}^Q$ is a bounded hypercube. Our approximation will be warranted only on Ω. With no loss of generality, it is assumed that $\Omega = [0,\ 1]^Q$ for the clarity of exposition. An approximation is computed for each of the p outputs of the model.

Choosing a design

The first concern is to select the set of points D where the function H is computed. This set will be called the design and has to be chosen carefully since the number of calls to H is limited to $N_{\max}$. In order to get an exploratory design, a Latin Hypercube Sampling (LHS)-*maximin* strategy is used. A design $D = \{z_1, \ldots z_N\} \subset \Omega \subset [0,\ 1]^Q$ is a LHS (McKay et al., 1979) if it is constructed as follows

$$z_i^j = \frac{\pi_j(i) - U_j^i}{N} \quad \forall\ 1 \leq i \leq N,\ \forall\ 1 \leq j \leq Q\,, \tag{5.3}$$

where π_j are independent uniform random permutations of the integers 1 through N, and the U_j^i are independent $\mathcal{U}_{[0,1]}$ random variables independent of the π_js. A LHS guarantees good projection properties. The sample points are stratified on each of Q input axis and the

projection of the design on any axis is well scattered. Therefore, it takes into account the variability for all dimensions.

Then, in order to have good exploratory properties which means that the points are spread in the input set, the design D is said to be chosen to be *maximin*. A design D is *maximin* if the distance between the sites is maximum: D has to maximise

$$\delta_D = \min_{z_i, z_j \in D} \|z_i - z_j\| , \tag{5.4}$$

and the number of pair of points (z_{i_0}, z_{j_0}) such that $\|z_{i_0} - z_{j_0}\| = \delta_D$ has to be minimal. $\max_D \delta_D$ is called the *maximin* distance. Morris and Mitchell (1995) provide a stochastic algorithm based on simulated annealing which aims at finding an optimal design according to the *maximin* property (5.4) within the class of LHS designs (5.3): the Latin hypercube sampling and the *maximin* property ensure that the provided design is well spread in the domain of interest.

Kriging predictor

It is assumed that $D = \{z_1, \dots z_N\}$ is a LHS-*maximin* design. The function H can be seen as the realisation of a Gaussian process Y

$$Y(z) = \sum_{i=1}^{P} \beta_i f_i(z) + G(z) = F(z)^T \boldsymbol{\beta} + G(z) . \tag{5.5}$$

In this setting, the f_i are known regression functions, the β_i are unknown parameters to be estimated and G is a centered Gaussian process characterised by its covariance function $\mathrm{cov}(G(s), G(t)) = \sigma^2 K_\theta(s,t)$ where K_θ is a symmetric positive definite kernel such that for all s, $K_\theta(s,s) = K_\theta(0,0) = 1$. The choice of the parameter θ allows us to tune the regularity of the process G. For instance in the case of a Gaussian kernel where $\theta \in \mathbb{R}_+$ and $K_\theta(r,s) = e^{-\theta\|r-s\|^2}$, the larger θ is, the smoother the process G is. For the sake of simplicity, the particular isotropic Gaussian kernel has been presented here. More general kernels can be found in Koehler and Owen (1996).
Therefore, the distribution of $Y_D = \{Y(z_1), \dots Y(z_N)\}$ is

$$p(Y_D) = \mathcal{N}(F_D\boldsymbol{\beta}, \sigma^2 \Sigma_{DD}) ,$$

where $F_D = (F(z_1) \dots F(z_N))^T$ and $(\Sigma_{DD})_{1 \le i,j \le N} = K_\theta(z_i, z_j) = \mathrm{corr}(Y(z_i), Y(z_j))$. The conditional process knowing the vector Y_D, is a Gaussian process. The distribution of $Y(z_0)$, given Y_D, is $\mathcal{N}(\mu_{z_0|D}, \sigma_{z_0 z_0|D})$, with

$$\begin{aligned} \mu_{z_0|D} &= E(Y(z_0)|Y_D) = F(z_0)^T\boldsymbol{\beta} + \Sigma_{z_0 D}^T \Sigma_{DD}^{-1}(Y_D - F_D\boldsymbol{\beta}) , \\ \sigma_{z_0 z_0|D} &= \mathrm{var}(Y(z_0)|Y_D) = \sigma^2 \left(1 - \Sigma_{z_0 D}^T \Sigma_{DD}^{-1} \Sigma_{z_0 D}\right) , \end{aligned}$$

where $\Sigma_{z_0 D} = (K_\theta(z_1, z_0), \dots K_\theta(z_N, z_0))^T$. The conditional mean $\mu_{z_0|D}$ can be used as a predictor of $H(z_0)$. Furthermore, $\boldsymbol{\beta}$, θ and σ^2 are estimated by maximising the likelihood. It leads to

$$\begin{aligned} \hat{\boldsymbol{\beta}} &= (F_D^T \Sigma_{DD}^{-1} F_D)^{-1} F_D^T \Sigma_{DD}^{-1} Y_D , \\ \hat{\sigma}^2 &= \frac{1}{N}(Y_D - F_D\hat{\boldsymbol{\beta}})^T \Sigma_{DD}^{-1}(Y_D - F_D\hat{\boldsymbol{\beta}}). \end{aligned}$$

Those estimators depend on θ via Σ_{DD}. The maximisation in θ is not explicit and is made by minimising

$$\psi(\theta) = |\Sigma_{DD}|^{\frac{1}{N}}\sigma^2(\theta).$$

The Matlab toolbox DACE (Lophaven et al., 2002) is used to compute all this parameters and to solve the optimisation problem in θ.

As a result, for all $z_0 \in \Omega$, the Kriging predictor of H is

$$\hat{H}(z_0) = F(z_0)^T\hat{\boldsymbol{\beta}} + \hat{\Sigma}_{z_0D}^T\hat{\Sigma}_{DD}^{-1}(Y_D - F_D\hat{\boldsymbol{\beta}}), \tag{5.6}$$

where $\hat{\Sigma}$ stands for $\Sigma(\hat{\theta}, \hat{\sigma}^2)$. Moreover, this predictor is exact for any $z_0 = z_i$, and it is the best linear unbiased predictor of $Y(z_0)$ for all $z_0 \in \Omega$. A fully Bayesian method as described in Santner et al. (2003) is possible. In this framework, a Gaussian prior distribution is set on the parameters $(\beta_i)_{1\leq i\leq p}$. If the prior distribution is diffuse enough, the posterior mean of $Y(z_0)$ (hence the predictor) tends to be the same than the maximum likelihood conditional mean of the Gaussian process.

Practical figures

(i) The choice of the input set Ω is sensitive. Ω has to be large enough to contain with a high probability the values of the random variable X and not too large in order to be efficient since the quality of Kriging depends on the design points concentration. The choice of Ω may rely on expert judgement. In practice, maximal plausible ranges for the x values are often known on a physical basis and are expected to be conservative though those ranges may exceed the likeliest (say 95%) range of true variability as the point in inverse statistical problems is precisely to identify the distribution. To prevent wrong results due to a poor approximation of H outside Ω, either the MCMC simulations are constrained to remain inside Ω or H is approximated thanks to a barycentric method outside Ω. It can lead to an adaptive scheme adapting the size of the domain according to early identification stages.

(ii) In order to compare Kriging to a barycentric approximation, a two class cross validation method is used. The design D is randomly split in two equal parts M times: $D = (D_1^{(i)}, D_2^{(i)})_{1\leq i\leq M}$. Then, for each i, the estimator is computed on the first part $D_1^{(i)}$, denoting $\hat{H}_{D_1^{(i)}}$, the relative prediction error computed on the other part $D_2^{(i)}$ is

$$ER(D_2^{(i)}|D_1^{(i)}) = \frac{2}{N}\sum_{z_j\in D_2^{(i)}}\left|\frac{H(z_j) - \hat{H}_{D_1^{(i)}}(z_j)}{H(z_j)}\right|.$$

Permuting the role of $D_1^{(i)}$ and $D_2^{(i)}$ leads to the error approximation

$$ER^{MC} = \frac{1}{2M}\sum_{i=1}^{M}\left(ER(D_2^{(i)}|D_1^{(i)}) + ER(D_1^{(i)}|D_2^{(i)})\right).$$

This Monte Carlo half sampling strategy is also a mean to choose the regression functions $(f_i)_{1\leq i\leq P}$ and the positive kernel K_θ for the Kriging predictor. Three spaces of regression

functions are usually chosen: the space generated by constant functions, the space of polynomials with degree smaller than or equal to one and the space of polynomials with degree smaller than or equal to two. The covariance function is chosen among the ones presented by Koehler and Owen (1996).

(iii) In order to decide if a single approximation suffices or if the approximation is to be made for each $H_i(= H(., d_i))$ the *maximin* distances (5.4) of the two strategies can be compared. For example, assuming there are ten different d_i for each i, these two strategies are respectively

1. Take a *maximin* design with 1000 points in $[0,1]^3$,
2. Take 10 (one for each d_i) *maximin* designs with 100 points in $[0,1]^2$.

For the first strategy, the *maximin* distance is denoted δ_1. If $\{z_1, \ldots, z_{1000}\}$ is a *maximin* design and $r = \delta_1/2$, the balls $(\mathcal{B}(z_i, r))_{1\leq i\leq 1000}$ are non-intersecting. Furthermore, the disjoint union $\bigsqcup_{1\leq i\leq 1000} \mathcal{B}(z_i, r)$ is included in the cube $[-r, 1+r]^3$ since the centers $(z_i)_{1\leq i\leq 1000}$ are in the cube $[0,1]^3$. Hence, by comparing volumes, the following inequality holds:

$$\begin{aligned} 1000\frac{4\pi}{3}r^3 &\leq (1+2r)^3 \\ r &\leq \left(\left(1000\frac{4}{3}\pi\right)^{1/3} - 2\right)^{-1} \approx 0.07\,. \end{aligned}$$

Thus, $\delta_1 = 2r \leq 0.14$.
Now, the minimal distance between the points of a regular grid of 100 points is $\frac{10}{9}$. Thus $\delta_2 \geq \frac{10}{9}$ where δ_2 is the *maximin* distance corresponding to the second strategy. As a consequence of $\delta_1 < \delta_2$, the first strategy is to be preferred in this case since it leads to a better concentration of the design points favouring a good behaviour of the Kriging predictor (Schaback, 2007).

The unknown parameters can be then estimated with the SEM algorithm defined in Section 5.3.1 where the Kriging approximation of H is used instead of H.

5.4 Numerical experiments

5.4.1 A flooding model

The model is related to the risk of dyke overflow during a flood event. It is a truly physics-based hydrodynamic model - even though quite simplified, as resulting from the well-known St-Venant equations in the one-dimensional case with a steady and uniform flow - that has been used as a benchmark in Rocquigny (de) or in Pasanisi et al. (2009). The available model computes the water level at the dyke position (Z_c) and the speed of the river (V) with respect to the observed flow of the river upstream of the dyke (Q), and non observed quantities: The river bed level at the dyke position (Z_v), and the value of Strickler coefficient K_s measuring the friction of the river bed, which is assumed to be homogeneous in this simplified model. Thus

$$\begin{pmatrix} Z_c \\ V \end{pmatrix} = H(Z_v, K_s; Q) + U \text{ with}$$

$$H(Z_v, K_s; Q) = \begin{pmatrix} Z_v + \left(\frac{\sqrt{L}}{B}\right)^{3/5} Q^{3/5} K_s^{-3/5} (Z_m - Z_v)^{-3/10} \\ B^{-2/5} L^{-3/10} Q^{2/5} K_s^{3/5} (Z_m - Z_v)^{3/10} \end{pmatrix},$$

where the values of the section length L and its width B are given and assumed to be fixed ($L = 5000, B = 300$). The river bed level beyond upstream (Z_m) has to be fixed to his mean value 55 in order to ensure identifiability. The ECME and SEM algorithms are used in the case where:

- Q follows a Gumbel distribution with mode $a = 1013$ and scaling parameter $b = 458$. (Cumulative distribution function $F(q) = 1 - \exp[-\exp((q-a)/b)]$).
- K_s follows a normal distribution with mean $\mu_{K_s} = 30$ and standard deviation $\sigma_{K_s} = 7.5$.
- Z_v follows a normal distribution with mean $\mu_{Z_v} = 50$ and standard deviation $\sigma_{Z_v} = 1$.

The goal is to estimate properly the parameters of the normal distributions of the data K_s and Z_v which are not observed, while flow values Q are assumed to be measurable: indeed, while such flood flows are generally unpredictable, upstream hydrological observations generally issue credible estimates. The ECME algorithm is used with iterative linearisations of the function H. The SEM algorithm is used in the case where the real model H is computed and in the case where H is replaced by a Kriging approximation $\hat{H}$. They are called respectively "full SEM" and "Kriging SEM". One hundred samples of $n = 50$ observations have been drawn to compare the parameter estimates given by these three algorithms. These estimates are compared to the ones obtained by maximising the completed likelihood if the non observed data were available.

The domain Ω where the Kriging approximation $\hat{H}$ of H is built, is chosen as $\Omega = [1,\ 65] \times [40,\ 54.9] \times [\min(Q_{obs}),\ \max(Q_{obs})]$ where $\min(Q_{obs}),\ \max(Q_{obs})$ are respectively the minimum and the maximum of the observations of Q. A smaller domain was early taken which have led to unsatisfying estimates with Kriging SEM. For the Kriging predictor, the regression functions are set to be linear and the kernel to be Gaussian i.e. $K_\theta(z, z') = \exp(-\theta\|z - z'\|_2^2)$. The initial values have been chosen as follows: for K_s, mean $\mu_{K_s}^{(0)} = 40$ and standard deviation $\sigma_{K_s}^{(0)} = 15$; for Z_v, mean $\mu_{Z_v}^{(0)} = 47$ and standard deviation $\sigma_{Z_v}^{(0)} = 3$. Different sets of initial values were used. However, only the results corresponding to that set were reported as those initial values are pretty far form the true ones and that the other runs lead to the same results. In ECME, the initial linearisation point is chosen to be $\mu^{(0)} = (\mu_{K_s}^{(0)}, \mu_{Z_v}^{(0)})$. The variance matrix of U is fixed to $R = \begin{pmatrix} 10^{-5} & 0 \\ 0 & 10^{-5} \end{pmatrix}$, and is supposed to be known.

Smooth histograms are plotted for the four parameters to be estimated in Figure 5.1. Table 5.1 provides the mean and the standard error of the 100 computed estimates. All the methods give similar results. The model is simple and a local linear approximation of H is efficient, that is why the linearisation in ECME perform well. ECME algorithm needs between five and ten iterations of the linearisation process until the stopping criterion (set to 10^{-15})

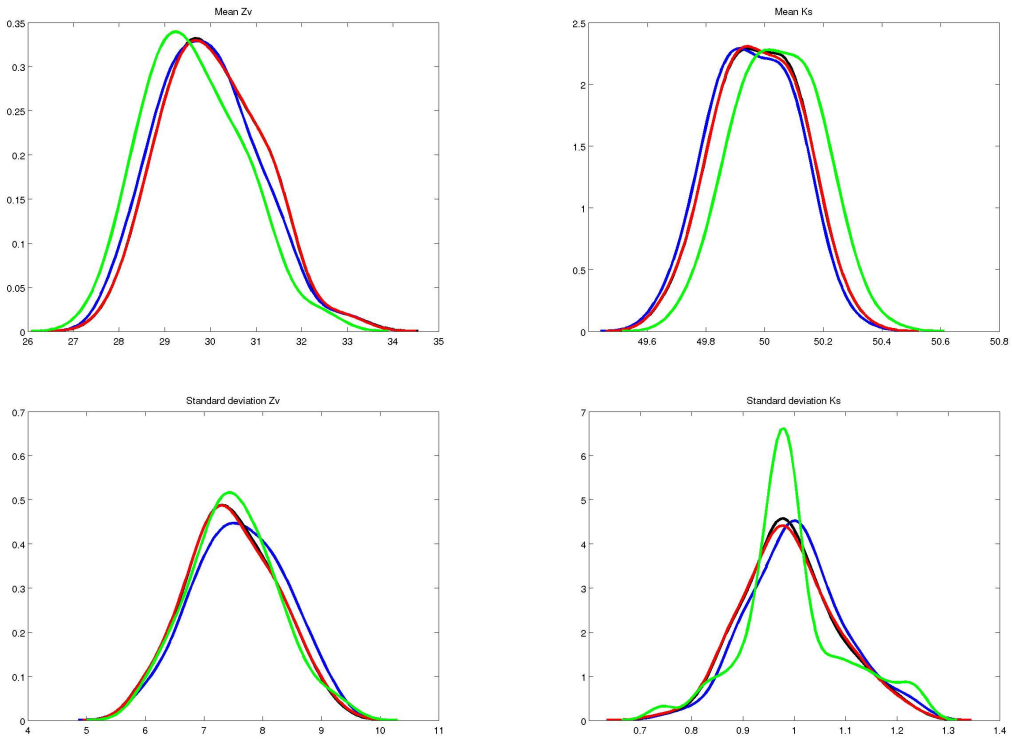


Figure 5.1: Smooth histograms for the four parameters to be estimated in the flooding example. The red line stands for the maximum likelihood estimates from the complete data, the black line for full SEM, the blue line for Kriging SEM and the green line for ECME. The black and the red lines are frequently superposed and the blue line is close to them.

is reached. For each iteration, $3n(= 150)$ calls to H are necessary. While only 100 calls to H are necessary to have a Kriging approximation with the Kriging SEM. The full SEM could not have worked if H were a real expensive black-box function since 50 iterations of S step are run. Each S step includes 200 iterations of the Hasting-Metropolis algorithm where H is to be evaluated for all the $n = 50$ points of the sample. Hence, $50 \cdot 200 \cdot 50 = 500000$ calls to the H function were required with the full SEM algorithm.

5.4.2 A non linear example

We have built an example to illustrate a problem which can occur when the function H cannot be locally linearly approximated. The model function is taken to be $H : [0,1]^3 \to \mathbb{R}^2$,

$$\begin{aligned}(y_1, y_2) &= H(x_1, x_2, d) \\ &= \begin{pmatrix} 5x_1(2x_2-1)^2 + x_2\cos(\pi(1-x_1)) + x_1|x_2 - 0.4| \\ (d+1)\sin(\pi(x_2 - 0.5)) \end{pmatrix}^T .\end{aligned}$$

The (y_1, y_2) are observed values depending on d which is observed and follows an uniform distribution on $[0,1]$ and on non observed values: x_1 following a normal distribution with mean 0.4 and standard deviation $\sqrt{2}/10$ and x_2 following a normal distribution with mean

Parameters	m_{K_s}	m_{Z_v}	σ_{K_s}	σ_{Z_v}	Numbers of calls to H
M.L. from complete data					N/A
Mean estimate	30.06	49.98	7.48	0.99	
Standard error	1.07	0.14	0.74	0.09	
ECME					between 750 and 1500
Mean estimate	29.63	50.04	7.50	1.01	
Standard error	1.06	0.14	0.74	0.12	
Full SEM					500 000
Mean estimate	30.06	49.98	7.48	0.99	
Standard error	1.07	0.14	0.74	0.09	
Kriging SEM					100
Mean estimate	29.92	49.96	7.61	1.00	
Standard error	1.09	0.14	0.76	0.09	

Table 5.1: Mean and standard error of the 100 computed estimates for the flooding example.

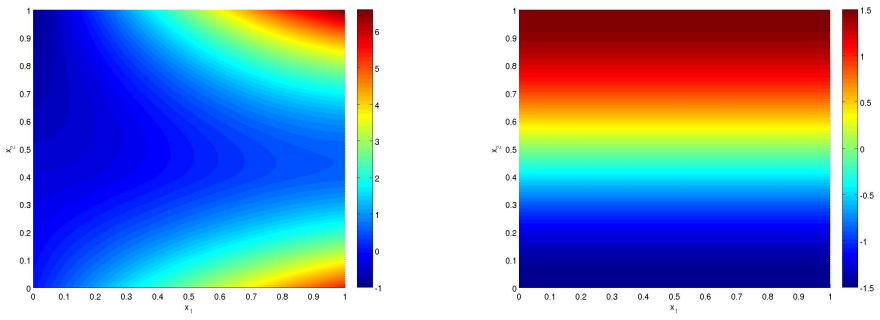


Figure 5.2: Colormaps corresponding to y_1 (on the left handside) and y_2 (on the right handside), where d is set to 0.5.

0.5 and standard deviation $\sqrt{2}/10$. As in the previous example, 100 samples of size $n = 50$ have been drawn to assess the estimation performances of each method: maximum likelihood estimator from the complete data, ECME with iterative linearisations of the function H, full SEM, Kriging SEM. $N_{\max} = 100$ evaluations of the function H have been used to obtain the Kriging approximation. Between six and ten linearisations were considered for the ECME algorithm. The domain where the Kriging approximation is done is $\Omega = [0,\ 1]^3$. In the Kriging predictor, the regression functions are set to be polynomials with degree equal or less than 2 and the kernel to be exponential i.e. $K_\theta(z, z') = \exp(-\theta\|z - z'\|_1)$. The initial values have been chosen as follows: for x_1, mean 0.2 and standard deviation 0.2; for x_2, mean 0.2 and standard deviation $2\sqrt{2}/10$. Those initial values were chosen far enough from the true values to show that the method coupling linearisations and ECME algorithm can be misleading. The histogram plots of all the methods are displayed in Figure 5.3.

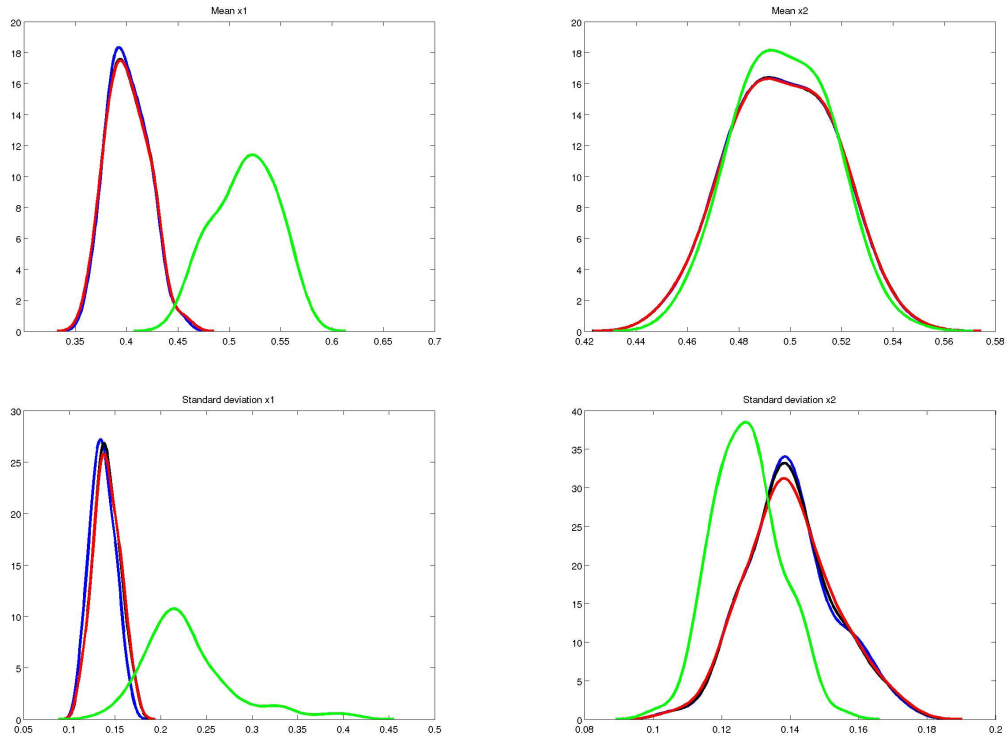


Figure 5.3: Smooth histograms for the four parameters to be estimated in the highly non linear example. The red line stands for the maximum likelihood estimates from the complete data, the black line for full SEM, the blue line for Kriging SEM and the green line for ECME. The black and the red lines are frequently superposed and the blue line is closed to them.

Four typically spurious estimates (over 100) given by ECME method were not taken into account into the plots and in Table 5.2 which summarised the results. As it is apparent from Figure 5.3, the ECME algorithm cannot estimate reasonably well the parameters of the unobserved variable x_1. A linear approximation of H in a neighbourhood of the mean of (x_1, x_2) (i.e. $(0.4, 0.5)$) performs poorly and ECME is misleading. The Kriging approximation is much more flexible than a linear approximation. Thus, a design of $N_{\max} = 100$ points is

Parameters	m_{x_1}	m_{x_2}	σ_{x_1}	σ_{x_2}	Numbers of calls to H
M.L. from complete data					N/A
Mean estimate	0.40	0.50	0.14	0.14	
Standard error	0.020	0.020	0.014	0.013	
ECME					between 750 and 1500
Mean estimate	0.52	0.50	0.23	0.13	
Standard error	0.030	0.020	0.050	0.010	
Full SEM					500 000
Mean estimate	0.40	0.50	0.14	0.14	
Standard error	0.020	0.020	0.014	0.013	
Kriging SEM					100
Mean estimate	0.40	0.50	0.14	0.14	
Standard error	0.019	0.020	0.013	0.013	

Table 5.2: Mean and standard error of the 100 computed estimates for the highly non linear example.

enough to get an approximation of H on $[0,1]^3$ leading to reasonable estimates with Kriging SEM. When the Jacobian matrices are computed at the different linearisation points, noticing a change in a sign of one of the coefficients can be a hint to think that the linear approximation would be misleading in ECME algorithm.

5.5 Discussion

A non linear method has been presented as an alternative to a linear method described in Celeux and Diebolt (1985) to solve an inverse problem occurring often in an industrial context. The function H governing the model is supposed to be highly non linear and only known for a limited number of points because it is the output of an expensive black-box. To identify such a model, a non linear method based on a Stochastic EM (SEM) algorithm has been proposed. But, since the model function H cannot been made available for a large number of points, it is approximated by Kriging in order to simulate the non observed variables conditionally to the observed variables resulting in an approximated SEM algorithm, the so-called Kriging SEM algorithm. In this paper, examples have been studied to assess the error made when H is replaced by a Kriging approximation. No matter which method is used with the flooding model, where function H can be reasonably linearised, the estimators behave almost like the ideal maximum likelihood estimator based on the complete data. But, it can be noticed that Kriging SEM needs less exact values of H to be computed (namely, a design of 100 points to approximate the model function gives good results with Kriging SEM) than ECME algorithm (at least 750 exact values of H are needed in the case where five iterated linearisations are enough). Furthermore, in ECME algorithm, the number of linearisations until the stopping criterion is reached is unknown a priori. Hence at the beginning of ECME algorithm, the number of needed calls to the model function H is not determined and this situation is somewhat uncomfortable. The second considered example where H was highly non linear illustrates that the linearisations at work in ECME can be misleading while SEM algorithm with a Kriging approximation continue to provide reasonable estimates. Linearisations are

actually harmful if H has locally highly non linear behaviours. Although the provided examples deal with a function H from $\mathbb{R}^{2+1}$ to $\mathbb{R}^2$, the method mixing SEM algorithm and Kriging can still perform well as long as the model (5.1) is identifiable and the Kriging approximation is close enough of the true function H. Kriging can give a reasonable approximation until about ten dimensions for the inputs (Fang et al., 2006).

An important and difficult issue is assessing the results: Has the algorithm converged? Are the estimates satisfactory? Unfortunately, there is no well-grounded criteria to answer those questions. Only experts can say if the estimates seem realistic. In particular, expert knowledge is required to decide which method is safer. Moreover, in the case where Kriging SEM could be recommended, experts are supposed to determine the domain where Kriging approximation is to be made and to propose a reasonable number of calls to the model function. As mentioned by a reviewer, if the experts have no idea of a reasonable number of calls, a solution could be to estimate the parameters for a growing number of initial calls to the model, until stabilisation of the estimator. Furthermore, the motivation for identifying the input probability distribution has to be kept in mind. This distribution is generally required for a further risk analysis: it will be propagated through a (possibly different) physical model to control the risk level of a key decision variable. Therefore, the sensitivity of this final variable as a function of this probability distribution would have to be taken into account in industrial applications in order to assess fairly the differences between the inversion algorithms investigated in the paper.

Bibliography

Celeux, G. and Diebolt, J. (1985). The SEM algorithm: a probabilistic teacher algorithm derived from the em algorithm for the mixture problem. *Computational Statistics Quaterly*, 2:73–82.

Celeux, G. and Diebolt, J. (1987). A probabilistic teacher algorithm for iterative maximum likelihood estimation. In *Classification and related methods of Data Analysis*, pages 617–623, Amsterdam, North Holland.

Celeux, G., Grimaud, A., Lefebvre, Y., and De Rocquigny, E. (2010). Identifying variability in multivariate systems through linearised inverse methods. *Inverse Problems In Science & Engineering*, 18(3):401–415.

Currin, C., Mitchell, T., Morris, M., and Ylvisaker, D. (1991). Bayesian prediction of deterministic functions, with applications to the design and analysis of computer experiments. *Journal of the American Statistical Association*, 86(416):953–963.

De Crecy, A. (1996). Determination of the uncertainties of the constitutive relationships in the cathare 2 code. In *Proceedings of the 4th ASME/JSME International Conference on Nuclear Engineering*.

Delyon, B., Lavielle, M., and Moulines, E. (1999). Convergence of a stochastic approximation version of the EM algorithm. *Annals of Statistics*, 27:94–128.

Dempster, E. J., Laird, N. M., and Rubin, D. B. (1977). Maximum likelihood from incomplete data via EM algorithm. *Annals of the Royal Statistical Society, Series B*, 39:1–38.

Fang, K.-T., Li, R., and Sudjianto, A. (2006). *Design and Modeling for Computer Experiments*. Computer Science and Data Analysis. Chapman & Hall/CRC.

Koehler, J. R. and Owen, A. B. (1996). Computer experiments. In *Design and analysis of experiments*, volume 13 of *Handbook of Statistics*, pages 261–308. North Holland, Amsterdam.

Kuhn, E. and Lavielle, M. (2004). Coupling a stochastic approximation version of EM with a MCMC procedure. *ESAIM P& S*, 8:115–131.

Liu, C. and Rubin, D. B. (1994). The ECME algorithm: a simple extension of EM and ECM with faster monotone convergence. *Biometrika*, 81:633–648.

Lophaven, N., Nielsen, H., and Sondergaard, J. (2002). Dace, a matlab kriging toolbox. Technical Report IMM-TR-2002-12, DTU. Available to : http://www2.imm.dtu.dk/ hbn/dace/dace.pdf.

McKay, M. D., Beckman, R. J., and Conover, W. J. (1979). A comparison of three methods for selecting values of input variables in the analysis of output from a computer code. *Technometrics*, 21(2):239–245.

Morris, M. D. and Mitchell, T. J. (1995). Exploratory designs for computer experiments. *Journal of Statistical Planning and Inference*, 43:381–402.

Nielsen, S. F. (2000). The stochastic EM algorithm: estimation and asymptotic results. *Bernoulli*, 6:457–489.

Pasanisi, A., Rocquigny (de), E., Bousquet, N., and Parent, E. (2009). Some useful features of the bayesian setting while dealing with uncertainties in industrial practice. In *Proceedings of the ESREL 2009 Conference*, volume 3, pages 1795–1802.

Rocquigny (de), E. (2009). Structural reliability under monotony: Properties of form, simulation or response surface methods and a new class of monotonous reliability methods (mrm). *Structural Safety*, 31(5):363–374.

Rocquigny (de), E., Devictor, N., and Tarantola, S., editors (2008). *Uncertainty in industrial practice, a guide to quantitative uncertainty management*. Wiley.

Santner, T. J., Williams, B., and Notz, W. (2003). *The Design and Analysis of Computer Experiments*. Springer-Verlag.

Schaback, R. (2007). Kernel-based meshless methods. Technical report, Institute for Numerical and Applied Mathematics, Georg-August-University Goettingen.

Chapitre 6

Estimation of rare events probabilities in computer experiments

Résumé

Nous proposons à présent une application à l'estimation de la probabilité d'événements rares. Ces événements dépendent des sorties d'un modèle physique dont les entrées sont aléatoires. Le modèle physique étant seulement connu à travers une fonction boîte-noire coûteuse f, le nombre d'évaluations possibles est limité. Ainsi un estimateur de Monte-Carlo naïf ne pourra donner une estimation fine et une borne de confiance précise sur cette probabilité. Notre but étant de garantir la fiabilité d'un système, nous nous devons d'obtenir une telle borne.
Nous proposons alors deux stratégies qui sont une estimation bayésienne et une méthode d'échantillonnage préférentiel. Elles reposent sur un métamodèle de krigeage qui revient à considérer une loi a priori sur la fonction f. Comme cela a été vu dans la partie 2.2.1, la loi a posteriori s'obtient à partir d'évaluations de f aux points d'un plan d'expérience. À partir de celle-ci, la stratégie bayésienne propose un estimateur et des bornes de crédibilité sur la probabilité de l'événement rare. La stratégie d'échantillonnage préférentiel utilise une loi instrumentale définie à l'aide du métamodèle. Les hypothèses bayésiennes sur f sont nécessaires pour assurer une borne de confiance sur cette probabilité.
Finalement, ces deux stratégies sont testées sur un exemple jouet et un cas pratique concernant l'estimation de la probabilité de collision entre un emport et l'avion l'ayant largué, est traité par une combinaison astucieuse des deux stratégies.

Mots clés : expériences simulées, événements rares, krigeage, échantillonnage préférentiel, Estimation bayésienne, fiabilité en aéronautique militaire.

Ce chapitre est issu d'une collaboration avec Yves Auffray et Jean-Michel Marin. Il a été soumis pour publication.

Abstract

We are interested in estimating the probability of rare events in the context of computer experiments. These rare events depends on the output of a physical model with random input variables. Since the model is only known through an expensive black box function, a crude Monte Carlo estimator does not perform well. We then propose two strategies to cope with this difficulty: a Bayesian estimate and an importance sampling method. Both methods relies on Kriging metamodeling. They are able to achieve sharp upper confidence bounds on the rare event probability.
These methods are applied to a toy example and a real case study which consists of finding an upper bound of the probability that the trajectory of an airborne load collides the aircraft that has released it.
Keywords: computer experiments, rare events, Kriging, importance sampling, Bayesian estimates, risk assessment with fighter aircraft.

Keywords: computer experiments, rare events, Kriging, importance sampling, Bayesian estimates, risk assessment with fighter aircraft.

6.1 Introduction

Rare events are a major concern in reliability of complex systems (Heidelberg, 1995; Shahabuddin, 1995). We focus here on rare events depending on computer experiments. A

computer experiment (Welch et al., 1992; Koehler and Owen, 1996) consists of an evaluation of a black box function which describes a physical model,

$$y = f(\mathbf{x}) \,, \tag{6.1}$$

where $y \in \mathbb{R}$ and $\mathbf{x} \in E$ where E is a compact subset of $\mathbb{R}$. The code which computes f is expensive since the model is complex. We assume that no more than N calls to f are possible. The input $\mathbf{x}$ are measured with a lack of precision and some variables are uncontrollabe. Both sources of uncertainties are modeled by a random distribution on E. Let $\mathbf{X}$ be the random variable. Our goal is to estimate the probability:

$$\pi_\rho = \mathbb{P}(f(\mathbf{X}) < \rho)) = \mathbb{P}(\mathbf{X} \in R_\rho) = \mathbb{P}_{\mathbf{X}}(R_\rho) \,,$$

where R_ρ is a subset of E defined by $R_\rho = \{\mathbf{x} : f(\mathbf{x}) < \rho\}$ and $\rho \in \mathbb{R}$ is a given threshold.

A crude Monte Carlo scheme leads to the following estimator of π_ρ:

$$\hat{\pi}_{\rho,N} = \frac{\Gamma(f, \mathbf{X}_{1:N}, \rho)}{N} \,, \tag{6.2}$$

where $\Gamma(f, \mathbf{X}_{1:N}, \rho)$ is defined by

$$\Gamma(f, \mathbf{X}_{1:N}, \rho) = \sum_{i=1}^{N} \mathbb{I}_{]-\infty,\rho[}(f(\mathbf{X}_i)) \,, \tag{6.3}$$

and $\mathbf{X}_{1:N} = (\mathbf{X}_1, \ldots, \mathbf{X}_N)$ is a N-sample of random variables with the same distribution than $\mathbf{X}$. Its expectation and its variance are:

$$\mathbb{E}(\hat{\pi}_{\rho,N}) = \mathbb{P}(\mathbf{X} \in R_\rho) = \pi_\rho \,, \quad \mathbb{V}(\hat{\pi}_{\rho,N}) = \frac{1}{N}\pi_\rho(1 - \pi_\rho) \,.$$

Hence, its relative error is $\dfrac{(\mathbb{V}(\hat{\pi}_{\rho,N}))^{1/2}}{\mathbb{E}(\hat{\pi}_{\rho,N})} \approx (\pi_\rho N)^{-1/2}$ when $\pi_\rho << \frac{1}{N}$. Therefore, the relative error can be very large. Furthermore, since $\Gamma(f, \mathbf{X}_{1:N}, \rho)$ follows a binomial distribution with parameters N and π_ρ, an exact confidence upper bound on π_ρ:

$$\mathbb{P}(\pi_\rho \leq b(\Gamma(f, \mathbf{X}_{1:N}, \rho), N, \alpha)) \geq 1 - \alpha \,,$$

is available as it is explained in Appendix 6.6. In the case where $\Gamma(f, \mathbf{X}_{1:N}, \rho) = 0$ which happens with probability $(1 - \pi_\rho)^N$, the $(1 - \alpha)$-confidence interval is $[0, 1 - (\alpha)^{1/N}]$. As an example, if the realization of $\Gamma(f, \mathbf{X}_{1:N}, \rho)$ is equal to 0, an upper confidence upper bound at level 0.9, $\pi_\rho \leq 10^{-5}$ can be warranted only if more than 230,000 calls to f were performed. When the purpose is to assess the reliability of a system under the constrain of a limited number of calls to f, there is a need for a sharper upper bound on π_ρ. Several ways to improve the precision of estimation have been proposed in the literature.

Since Monte Carlo estimation works better for frequent event, the first idea is to change the crude scheme in such a manner that the event becomes less rare. It is what importance sampling and splitting methods schemes try to achieve.

For example L'Ecuyer et al. (2007) showed that randomized quasi-Monte Carlo can be used jointly with splitting and/or importance sampling. By analysing a rare event as a cascade of intermediate less rare events, Del Moral and Garnier (2005) developped a genealogical particle

system approach to explore the space of inputs E. Cérou and Guyader (2007a,b) proposed an adaptive multilevel splitting also based on particle systems. An adaptive directional sampling method is presented by Munoz Zuniga et al. (2010) to accelerate the Monte Carlo simulation method. These methods can still need too many calls to f and the importance distribution is hard to set for an importance sampling method.

A general approach in computer experiments is to make use of a metamodel which is a fast computing function that approximates f. It has to be built on the basis of data $\{f(\mathbf{x}_1), \cdots, f(\mathbf{x}_n)\}$ which are evaluations of f at points of a well chosen design $D_n = \{\mathbf{x}_1, \cdots, \mathbf{x}_n\}$. The bet is that these n evaluations will allow to build more accurate estimators and bounds on the probability of the target event.
Kriging is such a metamodeling tool, one can see Santner et al. (2003) and more recently Li and Sudjianto (2005); Joseph (2006); Bingham et al. (2006). The function f is seen as a realization of a Gaussian process which is a Bayesian prior.
The related posterior distribution is computed conditionally to the data. It is still a Gaussian process whose mean can be used as a prediction of f everywhere on E and the variance as a pointwise measure of the accuracy of the prediction.
By using this mean and this variance, Oakley (2004) has developped a sequential method to estime quantiles and Vazquez and Bect (2009) a sequential method to estimate the probability of a rare event. Cannamela et al. (2008) have proposed some sampling strategies based only on a reduced model which is a coarse approximation of f (no information about the accuracy of prediction are given), to estimate quantiles.

Two approaches are investigated in that paper. Both rely on the hypothesis that f is a realization of a Gaussian process F independent of $\mathbf{X}$. As a consequence, π_ρ is a realization of the random variable:

$$\Pi_\rho = \mathbb{E}(\mathbb{I}_{]-\infty,\rho[}(F(\mathbf{X}))|F)\,.$$

The first approach consists of focusing on the posterior distribution of Π_ρ which depends on the posterior distribution of f given its computed evaluations. We show that a Bayesian estimator of Π_ρ can be computed and a credible bound is reachable by simulating Gaussian processes to obtain realizations of Π_ρ.
The other approach is an importance sampling method whose the importance distribution is based on the metamodel.

The paper is organized as follows: Section 6.2 describes the posterior distribution of the Gaussian process and how to obtain an estimator and a credible interval on Π_ρ. Section 6.3 presents the importance sampling method and the confidence upper bound which is provided with a high probability. Finally in Section 6.4, these methods are used on a toy example to ensure that they perform well and a solution to a real aeronautical case study about the risk that the trajectory of an airborne load collides the aircraft that has released it, is proposed.

6.2 Bayesian estimator and credible interval

The first step for Kriging metamodeling is to choose a design $D_n = \{\mathbf{x}_1, \ldots, \mathbf{x}_n\}$ of numerical experiments (one can see Morris and Mitchell (1995); Koehler and Owen (1996) and more recently Fang et al. (2006); Mease and Bingham (2006); Dette and Pepelyshev (2010)). Let $y_{D_n} = (y_1 = f(\mathbf{x}_1), \ldots, y_n = f(\mathbf{x}_n))$ be the evaluations of f on D_n. Let us start from a statistical model consisting of Gaussian processes $F_{\boldsymbol{\beta},\sigma,\boldsymbol{\theta}}$ whose the expressions are given by:

for $\mathbf{x} \in E$,

$$F_{\boldsymbol{\beta},\sigma,\boldsymbol{\theta}}(\mathbf{x}) = \sum_{k=1}^{L} \beta_j h_j(\mathbf{x}) + \zeta(\mathbf{x}) = H(\mathbf{x})^T \boldsymbol{\beta} + \zeta(\mathbf{x}) \,, \tag{6.4}$$

where

- $h_1, \ldots, h_L$ are regression functions, and $\boldsymbol{\beta} = (\beta_1, \ldots, \beta_L)$ is a vector of parameters,
- ζ is a centered Gaussian process with covariance

  $$\mathrm{Cov}(\zeta(\mathbf{x}), \zeta(\mathbf{x}')) = \sigma^2 K_{\boldsymbol{\theta}}(\mathbf{x}, \mathbf{x}') \,,$$

  where $K_{\boldsymbol{\theta}}$ is a correlation function depending on some parameters $\boldsymbol{\theta}$ (for details about kernels, see Koehler and Owen, 1996).

The maximum likelihood estimates $\hat{\boldsymbol{\beta}}, \hat{\sigma}, \hat{\boldsymbol{\theta}}$ of $\boldsymbol{\beta}, \sigma, \boldsymbol{\theta}$ are computed on the basis of the observations. Then, the Bayesian prior on f is chosen to be $F = F_{\hat{\boldsymbol{\beta}},\hat{\sigma},\hat{\boldsymbol{\theta}}}$ and the process F is assumed independent of $\mathbf{X}$. We denote F^{D_n} the process F conditionally to $F(\mathbf{x}_1) = y_1, \ldots, F(\mathbf{x}_n) = y_n$, in short $Y_{D_n} = y_{D_n}$. The process F^{D_n} is still a Gaussian process (see Santner et al., 2003) with

- mean: $\forall \mathbf{x}$,

  $$m_{D_n}(\mathbf{x}) = H(\mathbf{x})^T \hat{\boldsymbol{\beta}} + \Sigma_{\mathbf{x}D_n}^T \Sigma_{D_nD_n}^{-1} (y_{D_n} - H_{D_n} \hat{\boldsymbol{\beta}}) \,, \tag{6.5}$$

- covariance: $\forall \mathbf{x}, \mathbf{x}'$,

  $$K_{D_n}(\mathbf{x}, \mathbf{x}') = \hat{\sigma}^2 (K_{\hat{\boldsymbol{\theta}}}(\mathbf{x}, \mathbf{x}') - \Sigma_{\mathbf{x}D_n}^T \Sigma_{D_nD_n}^{-1} \Sigma_{\mathbf{x}'D_n}) \,, \tag{6.6}$$

where

$$(\Sigma_{D_nD_n})_{1 \le i,j \le n} = K_{\hat{\boldsymbol{\theta}}}(\mathbf{x}_i, \mathbf{x}_j) \text{ and } \Sigma_{\mathbf{x}D_n} = \left(K_{\hat{\boldsymbol{\theta}}}(\mathbf{x}, \mathbf{x}_i)\right)_{1 \le i \le n}^T \,.$$

In this approach the conditioning on the data regard the parameters as fixed although they are estimated.

The Bayesian prior distribution $\mathbb{P}_F$ on f leads to a Bayesian prior distribution on Π_ρ. Our goal is to use the distribution of the posterior process F^{D_n} conditionally to the observation of Y_{D_n}, to learn about the posterior distribution of Π_ρ. The random variable $\Pi_\rho^{D_n}$ is defined by:

$$\Pi_\rho^{D_n} = \mathbb{E}(\mathbb{I}_{]-\infty,\rho[}(F^{D_n}(\mathbf{X}))|F^{D_n}) \,. \tag{6.7}$$

Its distribution is the posterior distribution of Π_ρ conditionally to $Y_{D_n} = y_{D_n}$, as the following useful lemma states.

Lemma 6.1. *For all measurable function $g : \mathbb{R} \mapsto \mathbb{R}$,*

$$\mathbb{E}(g(\Pi_\rho^{D_n})) = \mathbb{E}(g(\mathbb{E}(\mathbb{I}_{]-\infty,\rho[}(F^{D_n}(\mathbf{X}))|F^{D_n}))) \,.$$

Proof

$$\begin{aligned}
\mathbb{E}(g(\Pi_\rho^{D_n})) &= \mathbb{E}(g(\Pi_\rho)|Y_{D_n} = y_{D_n}) \\
&= \mathbb{E}(g(\mathbb{E}(\mathbb{I}_{]-\infty,\rho[}(F(\mathbf{X}))|F))|Y_{D_n} = y_{D_n}) \\
&= \int_{\mathbb{R}^E} g(\mathbb{E}(\mathbb{I}_{]-\infty,\rho[}(F(\mathbf{X}))|F = \varphi)) \mathbb{P}_{F|Y_{D_n}=y_{D_n}}(d\varphi) \\
&= \int_{\mathbb{R}^E} g(\mathbb{E}(\mathbb{I}_{]-\infty,\rho[}(F(\mathbf{X}))|F = \varphi)) \mathbb{P}_{F^{D_n}}(d\varphi) \,.
\end{aligned}$$

Since $\mathbf{X}$ and F are independent,

$$\mathbb{E}(\mathbb{I}_{]-\infty,\rho[}(F(\mathbf{X}))|F=\varphi) = \mathbb{E}(\mathbb{I}_{]-\infty,\rho[}(\varphi(\mathbf{X}))) .$$

Hence,

$$\begin{aligned}\mathbb{E}(g(\Pi_\rho^{D_n})) &= \int_{\mathbb{R}^E} g(\mathbb{E}(\mathbb{I}_{]-\infty,\rho[}(\varphi(\mathbf{X}))))\mathbb{P}_{F^{D_n}}(d\varphi)\\ &= \int_{\mathbb{R}^E} g(\mathbb{E}(\mathbb{I}_{]-\infty,\rho[}(F^{D_n}(\mathbf{X}))|F^{D_n}=\varphi))\mathbb{P}_{F^{D_n}}(d\varphi)\\ &= \mathbb{E}(g(\mathbb{E}(\mathbb{I}_{]-\infty,\rho[}(F^{D_n}(\mathbf{X}))|F^{D_n}))) .\end{aligned}$$

□
The mean and the variance of $\Pi_\rho^{D_n}$ are, then, given by:

Proposition 6.1.

$$\mathbb{E}(\Pi_\rho^{D_n}) = \int_E \mathbb{E}(\mathbb{I}_{]-\infty,\rho[}(F^{D_n}(\mathbf{x})))\mathbb{P}_{\mathbf{X}}(d\mathbf{x}) = \mathbb{E}\left(\Phi\left(\frac{\rho - m_{D_n}(\mathbf{X})}{\sqrt{K_{D_n}(\mathbf{X},\mathbf{X})}}\right)\right) , \tag{6.8}$$

where Φ is the cumulative distribution function of a centered reduced Gaussian random variable.

$$\mathbb{V}(\Pi_\rho^{D_n}) = \int_{E\times E} \mathbb{C}\mathrm{ov}(\mathbb{I}_{]-\infty,\rho[}(F^{D_n}(\mathbf{x})), \mathbb{I}_{]-\infty,\rho[}(F^{D_n}(\mathbf{x}'))\mathbb{P}_{\mathbf{X}}\times\mathbb{P}_{\mathbf{X}}(d\mathbf{x}, d\mathbf{x}') . \tag{6.9}$$

Proof
From Lemma 6.1, it comes

$$\begin{aligned}\mathbb{E}(\Pi_\rho^{D_n}) &= \mathbb{E}(\mathbb{E}(\mathbb{I}_{]-\infty,\rho[}(F^{D_n}(\mathbf{X}))|F^{D_n}))\\ &= \mathbb{E}(\mathbb{I}_{]-\infty,\rho[}(F^{D_n}(\mathbf{X}))) = \int_E \mathbb{E}(\mathbb{I}_{]-\infty,\rho[}(F^{D_n}(\mathbf{x})))\mathbb{P}_{\mathbf{X}}(d\mathbf{x}) .\end{aligned}$$

Since $F^{D_n}(\mathbf{x})$ follows Gaussian distribution with mean $m_{D_n}(\mathbf{x})$ and variance $K_{D_n}(\mathbf{x},\mathbf{x})$,

$$\mathbb{E}(\Pi_\rho^{D_n}) = \mathbb{E}\left(\Phi\left(\frac{\rho - m_{D_n}(\mathbf{X})}{\sqrt{K_{D_n}(\mathbf{X},\mathbf{X})}}\right)\right) .$$

Then, $\mathbb{E}((\Pi_\rho^{D_n})^2)$ is computed by using again Lemma 6.1 and the independence of $\mathbf{X}$ and F^{D_n}:

$$\begin{aligned}\mathbb{E}((\Pi_\rho^{D_n})^2) &= \mathbb{E}(\left[\mathbb{E}(\mathbb{I}_{]-\infty,\rho[}(F^{D_n}(\mathbf{X}))|F^{D_n}\right]^2)\\ &= \int_{\mathbb{R}^E} \left[\mathbb{E}(\mathbb{I}_{]-\infty,\rho[}(\varphi(\mathbf{X})))\right]^2 \mathbb{P}_{F^{D_n}}(d\varphi)\\ &= \int_{\mathbb{R}^E}\int_E \mathbb{I}_{]-\infty,\rho[}(\varphi(\mathbf{x}))\mathbb{P}_{\mathbf{X}}(d\mathbf{x}) \int_E \mathbb{I}_{]-\infty,\rho[}(\varphi(\mathbf{x}))\mathbb{P}_{\mathbf{X}}(d\mathbf{x})\mathbb{P}_{F^{D_n}}(d\varphi)\\ &= \int_{E^2} \mathbb{E}(\mathbb{I}_{]-\infty,\rho[}(F^{D_n}(\mathbf{x}))\mathbb{I}_{]-\infty,\rho[}(F^{D_n}(\mathbf{x}')))\mathbb{P}_{\mathbf{X}}\times\mathbb{P}_{\mathbf{X}}(d\mathbf{x}, d\mathbf{x}') .\end{aligned}$$

As, it also holds:

$$\mathbb{E}(\Pi_\rho^{D_n})^2 = \int_{E^2} \mathbb{E}(\mathbb{I}_{]-\infty,\rho[}(F^{D_n}(\mathbf{x})))\mathbb{E}(\mathbb{I}_{]-\infty,\rho[}(F^{D_n}(\mathbf{x}')))\mathbb{P}_{\mathbf{X}} \times \mathbb{P}_{\mathbf{X}}(d\mathbf{x}, d\mathbf{x}')\,,$$

we get

$$\begin{aligned}\mathbb{V}(\Pi_\rho^{D_n}) &= \mathbb{E}((\Pi_\rho^{D_n})^2) - \mathbb{E}(\Pi_\rho^{D_n})^2 \\ &= \int_{E^2} \mathbb{C}\mathrm{ov}(\mathbb{I}_{]-\infty,\rho[}(F^{D_n}(\mathbf{x})), \mathbb{I}_{]-\infty,\rho[}(F^{D_n}(\mathbf{x}'))\mathbb{P}_{\mathbf{X}} \times \mathbb{P}_{\mathbf{X}}(d\mathbf{x}, d\mathbf{x}')\,.\end{aligned}$$

□

A numerical Monte Carlo integration can be used to compute the posterior mean and variance since they do not need more calls to f. However, the computation time requested by a massive Monte Carlo integration, especially for $\mathbb{V}(\Pi_\rho^{D_n})$, can be very long as it is noticed in the examples.

The mean and the variance of $\Pi_\rho^{D_n}$ can be used to obtain credible bounds. As a consequence of Markov inequality, it holds, for any $\alpha \in [0,1]$,

$$\mathbb{P}\left(\Pi_\rho^{D_n} \leq \frac{\mathbb{E}(\Pi_\rho^{D_n})}{\alpha}\right) \geq 1-\alpha\,. \tag{6.10}$$

Likewise, Chebychev inequality gives, for any $\alpha \in [0,1]$,

$$\mathbb{P}\left(\Pi_\rho^{D_n} \leq \mathbb{E}(\Pi_\rho^{D_n}) + \sqrt{\frac{\mathbb{V}(\Pi_\rho^{D_n})}{\alpha}}\right) \geq 1-\alpha\,. \tag{6.11}$$

The quantiles of $\Pi_\rho^{D_n}$ are exactly the upper bounds that are sought. They can be reached through massive simulation of $\Pi_\rho^{D_n}$. For example, the following algorithm provides realizations of $\Pi_\rho^{D_n}$. It relies on a discretization of the Gaussian process to be simulated.

Algorithm 6.1.

1. **Simulate a realization of a Gaussian process:** *A realization of the vector of points $\tilde{y} = (y_{\tilde{\mathbf{x}}_i})_{1\leq i\leq \tilde{n}}$ is drawn according to the distribution F^{D_n} of the Gaussian process. The points $\tilde{\mathbf{x}}_1, \ldots, \tilde{\mathbf{x}}_{\tilde{n}}$ can be a grid in E.*

2. **Reconstruction of the realization:** *By a Kriging method, the points $\tilde{y} \cup y_{D_n}$ are interpolated. This interpolation is considered as a realization of F^{D_n} on E.*

3. **Numerical integration:** *The realization π_ρ corresponding to the realization of the Gaussian process is hence computed using a massive Monte Carlo integration with respect to the distribution of $\mathbf{X}$.*

Using a lot of iterations, it is possible to obtain an approximation of the cumulative distribution function of $\Pi_\rho^{D_n}$ which gives estimates of quantiles. Thus, a credible interval on Π_ρ is constructed. A constant $a \in [0,1]$ is found such that:

$$\mathbb{P}(\Pi_\rho^{D_n} < a) \geq 1-\alpha\,.$$

This approach can suffer of an error due to the spatial discretization needed at step 1 of the algorithm.

6.3 Importance sampling

As it was explained in Section 6.1, the major drawback of the crude Monte Carlo scheme is the high level of uncertainty when it is used for estimating the probability of a rare event. Importance sampling is a way to tackle this problem. The basic idea is to change the distribution to make the target event more frequent. We aim at sampling according to the importance distribution:

$$\mathbb{P}_{\mathbf{Z}} : A \subset E \mapsto \mathbb{P}_{\mathbf{X}}(A|\hat{R}_\rho)\,,$$

where $\hat{R}_\rho \subset E$ is to be designed close to $R_\rho = \{\mathbf{x} \in E : f(\mathbf{x}) < \rho\}$. Thanks to n calls to the metamodel, a set $\hat{R}_\rho$ can be chosen as follows:

$$\hat{R}_\rho = \hat{R}_{\rho,\kappa} = \left\{\mathbf{x} : m_{D_n}(\mathbf{x}) < \rho + \kappa\sqrt{K_{D_n}(\mathbf{x},\mathbf{x})}\right\}\,, \tag{6.12}$$

where κ is fixed such that "$\{\mathbf{x} : F(\mathbf{x}) < \rho\} \subset \hat{R}_{\rho,\kappa}$ with a good confidence level". In other words, if $\mathbf{x}$ is such that $f(\mathbf{x}) < \rho$, we want $\mathbf{x}$ to be in $\hat{R}_{\rho,\kappa}$. We recall that the posterior mean $m_{D_n}(\mathbf{x})$ is an approximation of $f(\mathbf{x})$ and $\kappa\sqrt{K_{D_n}(\mathbf{x},\mathbf{x})}$ has been added to take into account the uncertainty of the approximation.

A set of m points, $\mathbf{Z}_{1:m} = (\mathbf{Z}_1, \ldots, \mathbf{Z}_m)$, is drawn to be an i.i.d. sample following the importance distribution. The corresponding importance sampling estimator of π_ρ is

$$\frac{\mathbb{P}_{\mathbf{X}}(\hat{R}_\rho)}{m}\Gamma(f, \mathbf{Z}_{1:m}) = \frac{\mathbb{P}_{\mathbf{X}}(\hat{R}_\rho)}{m}\sum_{k=1}^{m}\mathbb{I}_{]-\infty,\rho[}(f(\mathbf{Z}_k))\,. \tag{6.13}$$

The probability $\mathbb{P}_{\mathbf{X}}(\hat{R}_\rho)$ is computable by a Monte Carlo integration since it does not depend on f; yet, m more calls to f are necessary to compute $\mathbb{I}_{]-\infty,\rho[}(f(\mathbf{Z}_k))$. This estimator is only unbiased provided that $R_\rho \subset \hat{R}_\rho$. Nevertheless, it is an unbiased estimator of $\mathbb{E}_{\mathbf{X}}(\mathbb{I}_{]-\infty,\rho[}(f(\mathbf{X}))\mathbb{I}_{\hat{R}_\rho}(\mathbf{X}))$. Since $\Gamma(f, \mathbf{Z}_{1:m})$ follows a binomial distribution $\mathcal{B}\left(m, \frac{\mathbb{E}(\mathbb{I}_{]-\infty,\rho[}(f(\mathbf{X}))\mathbb{I}_{\hat{R}_\rho}(\mathbf{X}))}{\mathbb{P}_{\mathbf{X}}(\hat{R}_\rho)}\right)$, for any $\alpha \in]0;1[$, the following confidence upper bound holds:

$$\mathbb{P}\left(\mathbb{E}(\mathbb{I}_{]-\infty,\rho[}(f(\mathbf{X}))\mathbb{I}_{\hat{R}_\rho}(\mathbf{X})) \leq b(\Gamma(f, \mathbf{Z}_{1:m}, \rho), m, \alpha)\mathbb{P}_{\mathbf{X}}(\hat{R}_\rho)\right) > 1 - \alpha\,, \tag{6.14}$$

by using the bound described in Appendix 6.6. This is an upper bound on π_ρ only if the estimator (6.13) is unbiased i.e. only if $R_\rho \subset \hat{R}_\rho$. As it is noticed in the decomposition:

$$\pi_\rho = \mathbb{E}(\mathbb{I}_{]-\infty,\rho[}(f(\mathbf{X}))) = \mathbb{E}(\mathbb{I}_{]-\infty,\rho[}(f(\mathbf{X}))\mathbb{I}_{\hat{R}_\rho}(\mathbf{X})) + \mathbb{E}(\mathbb{I}_{]-\infty,\rho[}(f(\mathbf{X}))(1 - \mathbb{I}_{\hat{R}_\rho}(\mathbf{X})))\,,$$

the second term on the right-hand side which is the opposite of the bias has to be controlled. That is why the random variable

$$\Pi_\rho^{D_n} = \mathbb{E}(\mathbb{I}_{]-\infty,\rho[}(F^{D_n}(\mathbf{X}))|F^{D_n})\,,$$

whose a realisation is π_ρ, is considered.
Similarly to the previous decomposition, it holds

$$\Pi_\rho^{D_n} = \mathbb{E}(\mathbb{I}_{]-\infty,\rho[}(F^{D_n}(\mathbf{X}))\mathbb{I}_{\hat{R}_\rho}(\mathbf{X})|F^{D_n}) + \mathbb{E}(\mathbb{I}_{]-\infty,\rho[}(F^{D_n}(\mathbf{X}))(1 - \mathbb{I}_{\hat{R}_\rho}(\mathbf{X}))|F^{D_n})\,. \tag{6.15}$$

A bound on $\mathbb{E}(\mathbb{I}_{]-\infty,\rho[}(F^{D_n}(\mathbf{X}))\mathbb{I}_{\hat{R}_\rho}(\mathbf{X})|F^{D_n})$ comes from (6.14).

Proposition 6.2. *For $\alpha \in]0,1[$, it holds*

$$\mathbb{P}(\left(\mathbb{E}(\mathbb{I}_{]-\infty,\rho[}(F^{D_n}(\mathbf{X}))\mathbb{I}_{\hat{R}_\rho}(\mathbf{X})|F^{D_n}) \leq \mathbf{b}\ \mathbb{P}_{\mathbf{X}}(\hat{R}_\rho)\right) \geq 1-\alpha\,, \tag{6.16}$$

where $\mathbf{b}$ *stands for* $b(\Gamma(F^{D_n}, \mathbf{Z}_{1:m}, \rho), m, \alpha)$.

Proof
Let φ be any realisation of F^{D_n}.
As in (6.14), we have

$$\mathbb{P}\left(\mathbb{E}(\mathbb{I}_{]-\infty,\rho[}(\varphi(\mathbf{X}))\mathbb{I}_{\hat{R}_\rho}(\mathbf{X})) \leq b(\Gamma(\varphi, \mathbf{Z}_{1:m}, \rho), m, \alpha)\mathbb{P}_{\mathbf{X}}(\hat{R}_\rho)\right) \geq 1-\alpha\,.$$

Thus, since this result holds for any realisation of F^{D_n},

$$\mathbb{P}\left(\mathbb{E}(\mathbb{I}_{]-\infty,\rho[}(F^{D_n}(\mathbf{X}))\mathbb{I}_{\hat{R}_\rho}(\mathbf{X})|F^{D_n}) \leq b(\Gamma(F^{D_n}, \mathbf{Z}_{1:m}, \rho), m, \alpha)\mathbb{P}_{\mathbf{X}}(\hat{R}_\rho)\right) \geq 1-\alpha\,.$$

□
The next proposition states an upper bound for the second term in (6.15).

Proposition 6.3. *For $\beta \in]0,1[$, it holds*

$$\mathbb{P}\left(\mathbb{E}(\mathbb{I}_{]-\infty,\rho[}(F^{D_n}(\mathbf{X}))(1-\mathbb{I}_{\hat{R}_\rho}(\mathbf{X}))|F^{D_n}) \leq \frac{\mathbf{c}}{\beta}\right) \geq 1-\beta\,,$$

where $\mathbf{c} = \mathbb{E}\left(\Phi\left(\frac{\rho - m_{D_n}(\mathbf{X})}{\sqrt{K_{D_n}(\mathbf{X},\mathbf{X})}}\right)(1-\mathbb{I}_{\hat{R}_\rho}(\mathbf{X}))\right)$.

Proof
The mean of $\mathbb{E}(\mathbb{I}_{]-\infty,\rho[}(F^{D_n}(\mathbf{X}))(1-\mathbb{I}_{\hat{R}_\rho}(\mathbf{X}))|F^{D_n})$ can be computed in the same fashion than the mean of $\Pi_\rho^{D_n}$ in Proposition 6.1. It gives

$$\mathbb{E}\left(\mathbb{E}(\mathbb{I}_{]-\infty,\rho[}(F^{D_n}(\mathbf{X}))(1-\mathbb{I}_{\hat{R}_\rho}(\mathbf{X}))|F^{D_n})\right) = \mathbb{E}\left(\Phi\left(\frac{\rho - m_{D_n}(\mathbf{X})}{\sqrt{K_{D_n}(\mathbf{X},\mathbf{X})}}\right)(1-\mathbb{I}_{\hat{R}_\rho}(\mathbf{X}))\right)\,.$$

Then, Markov inequality is applied which completes the proof. □
Finally, by gathering the results of Proposition 6.2 and Proposition 6.3, a stochastic upper bound is found on $\Pi_\rho^{D_n}$.

Proposition 6.4. *For $\alpha, \beta \in]0,1[$ such that $\alpha+\beta<1$, it holds*

$$\mathbb{P}\left(\Pi_\rho^{D_n} \leq \mathbf{b}\mathbb{P}_{\mathbf{X}}(\hat{R}_\rho) + \frac{\mathbf{c}}{\beta}\right) \geq 1-(\alpha+\beta)\,, \tag{6.17}$$

where $\mathbf{b}$ *and* $\mathbf{c}$ *have been defined above.*

The proof is obvious.
If $\hat{R}_\rho$ is chosen as proposed in (6.12), the bound $\mathbf{c}$ is:

$$\mathbf{c} = c(\kappa) = \mathbb{E}\left(\Phi\left(\frac{\rho - m_{D_n}(\mathbf{X})}{\sqrt{K_{D_n}(\mathbf{X},\mathbf{X})}}\right)\mathbb{I}_{]-\infty,-\kappa[}\left(\frac{\rho - m_{D_n}(\mathbf{X})}{\sqrt{K_{D_n}(\mathbf{X},\mathbf{X})}}\right)\right)\,.$$

6.4 Numerical experiments

6.4.1 A toy example

The function $f : [-10, 10]^2 \to \mathbb{R}_+$ is assumed to describe a physical model:

$$f(x_1, x_2) = -\frac{\sin(x_1)}{x_1} - \frac{\sin(x_2 + 2)}{x_2 + 2} + 2\,.$$

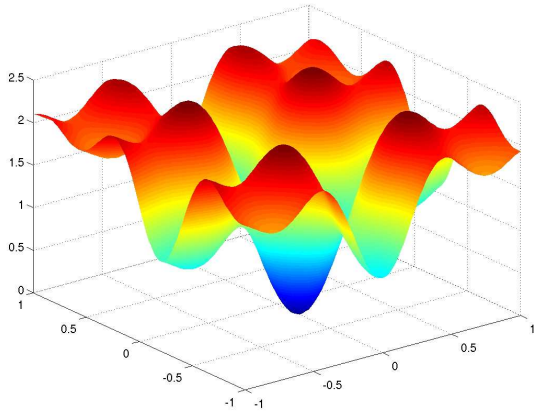


Figure 6.1: The function f

The input vector $\mathbf{X}$ is supposed to have an uniform distribution on $[-10, 10]^2$. The threshold is set to $\rho = 0.01$ which corresponds to the probability $\mathbb{P}_{\mathbf{X}}\left(f(\mathbf{X}) < \rho\right) = 4.72 \cdot 10^{-4}$. This probability was computed thanks to a massive Monte Carlo integration. In the case where only $N = 100$ calls to f are available, the two strategies are tested. A maximin design with 100 points for the Bayesian strategy and one with 50 points for importance sampling strategy are computed thanks to a simulated annealing algorithm. Kriging metamodels are built with an intercept as the regression function and a Gaussian correlation function is chosen as the correlation function of the Gaussian process ζ i.e. $\forall \mathbf{x} \in E$, $h(\mathbf{x}) = 1$ and $\forall \mathbf{x}, \mathbf{x}' \in E$, $K(\mathbf{x}, \mathbf{x}) = \exp\left(-\theta \|\mathbf{x} - \mathbf{x}'\|^2\right)$ are set for the model given by equation (6.4). The Bayesian estimate of π_ρ is $4.63 \cdot 10^{-4}$. It was computed by a Monte Carlo integration on a 10^7-sample using the result of Proposition 6.1. Yet, we were not able to determine the posterior variance in a reasonable time. The importance sampling estimate of π_ρ, constructed on a 50-sample, is $6.13 \cdot 10^{-4}$. The probability $\mathbb{P}_{\mathbf{X}}(\hat{R}_{\rho,\kappa})$ (and also the bound on the bias, given in Proposition 6.3) was also computed by a Monte Carlo integration on a 10^7-sample and $\kappa = 3$ has been set.

Then, the stochastic bounds on $\Pi_\rho^{D_n}$ are focused on. A thousand iterations of Algorithm 6.1 where the points $\tilde{\mathbf{x}}_1, \ldots, \tilde{\mathbf{x}}_{\tilde{n}}$ have been chosen to be a grid of one hundred points in $[-1, 1]^2$ and the numerical integration at step 3 is performed with a 10^5-sample, are done. In order to prevent the covariance matrix of the posterior process to be ill-conditioned the identity matrix multiplied by a small coefficient (here 10^{-5}) is added. It is a regularization of the solution known as a nugget effect in the Kriging literature. The estimates of the posterior quantiles are $1.2 \cdot 10^{-3}$ at level 90% and $2.1 \cdot 10^{-3}$ at level 98%. The bounds found with importance

sampling are $1.5 \cdot 10^{-3}$ at level 90% ($\alpha = \beta = 5\%$) and $2.1 \cdot 10^{-3}$ at level 98% ($\alpha = \beta = 1\%$). If a crude Monte Carlo scheme is used here with only $N = 100$ calls, the estimator is equal to 0 with probability greater than 0.95 and in this case, the upper confidence bounds are 0.023 and 0.038 respectively at levels 90% and 98%.

There are sources of variability on the estimators and the bounds due to the choice in the designs. Indeed, the designs are computed to be maximin by using a finite number of iterations of a simulated annealing algorithm. Moreover, there exist symmetries within the class of maximin designs. Concerning the importance sampling strategy, the sampling which gives $\mathbf{Z}_{1:m}$ induces variability.

In order to test the sensitivity of the estimators and the bounds to these sources of variability, each of the two strategies as described just above, are repeated one hundred times. Figure 6.2 displays a boxplot of one hundred estimates obtained with the Bayesian method on the left-hand side and a boxplot of one hundred estimates obtained with the importance sampling method on the right-hand side.

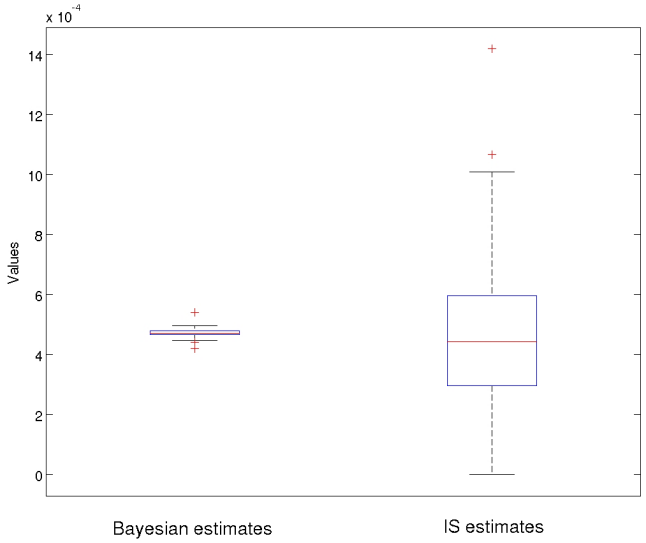


Figure 6.2: Estimates of π_ρ

Figure 6.3 displays the boxplots concerning bounds at level 90% and at level 98% given by the Bayesian method (left-hand side) and the importance sampling method (right-hand side). Table 6.1 summarizes the estimates and Table 6.2 summarizes the bounds.

These results show that the Bayesian method is very reliable for estimating π_ρ while the importance sampling method provides the sharpest upper bounds. The Bayesian method suffers from the fact that the posterior quantiles are estimated thanks to an algorithm which relies on a discretization of the space and is burdensome which implies a limited number of possible iterations. The importance sampling methods which splits into two terms the probability to bound is much more efficient. As these methods depend on the Kriging model hypothesis (6.4), a leave-one-out cross validation as proposed by Jones et al. (1998) can be performed to check if this hypothesis is sensible. It consists of building n metamodels with

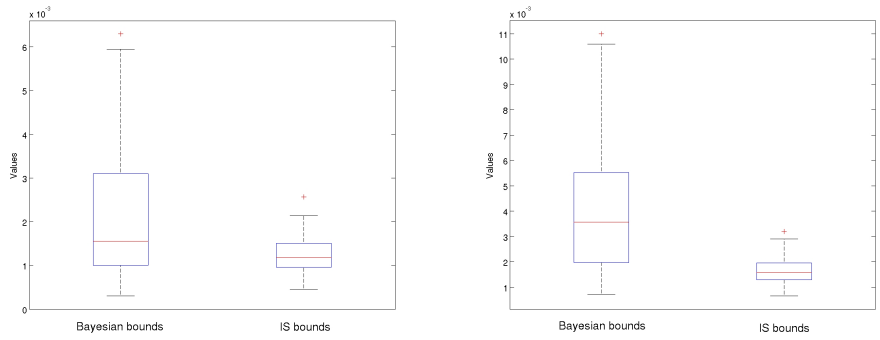


Figure 6.3: Bounds on π_ρ at level 90% (left) and at level 98% (right)

posterior mean and variance denoted respectively by $m_{D_n^{-i}}$ and $\sigma^2_{D_n^{-i}}$, from designs

$$D_n^{-i} = \{\mathbf{x}_1, \ldots, \mathbf{x}_{i-1}, \mathbf{x}_{i+1}, \ldots, \mathbf{x}_n\},$$

where $i = 1, \ldots, n$.
Then, the values

$$\frac{|f(\mathbf{x}_i) - m_{D_n^{-i}}(\mathbf{x}_i)|}{\sigma^2_{D_n^{-i}}(\mathbf{x}_i)}, \tag{6.18}$$

are computed. If something like 99.7% of them lies in the interval $[-3, 3]$, the Kriging hypothesis is not rejected. In our toy example, all of the tests which were made give that all these values are in $[-2, 2]$.

	Bayesian estimates	IS estimates
Minimum	4.19	0
Maximum	5.40	14
Mean	4.72	4.72

Table 6.1: Estimates of π_ρ multiplied by 10^{-4}

	Bayesian bounds		IS bounds	
	90%	98%	90%	98%
Minimum	3	7	4.5	6.5
Maximum	63	110	26	32
Mean	20	39	12	16

Table 6.2: Bounds on π_ρ multiplied by 10^{-4}

6.4.2 A real case study: release enveloppe clearance

Context

When releasing an airborne load, a critical issue is the risk that its trajectory could collide the aircraft. The behaviour of such a load after release depends on many variables. Some are under control of the crew: mach, altitude, load factor etc. We call them *controlled variables* and note C their variation domain. The others are *uncontrolled variables*: let E be the set of their possible values. The *release enveloppe clearance* problem consists of exploring the set C to find a subset where the release is safe, whatever the uncontrolled variables are. To investigate this problem, we can use a simulator which computes the trajectory of the carriage when the values of all the variables are given. Moreover, for $\mathbf{x}_C \in C$ and $\mathbf{x} \in E$, besides the trajectory $\tau(\mathbf{x}_C, \mathbf{x})$, the program delivers a *dangerousness score* $f(\mathbf{x}_C, \mathbf{x})$ to be interpreted as an "algebraic distance": a negative value characterizes a collision trajectory.
To assess the safety of release at a given point of C, we suppose that the values of the uncontrolled variables are realizations of a random variable $\mathbf{X} \in E$ that can be simulated. Therefore, for a given value $\mathbf{x}_C \in C$, and $\rho \geq 0$ the ρ-collision risk is the probability

$$\pi_\rho(\mathbf{x}_C) = \mathbb{P}(f(\mathbf{x}_C, \mathbf{X}) < \rho)\,.$$

We do not aim at estimating accurately this risk.
We would rather classify the points into three categories: according to the position of 0-risk $\pi_0(\mathbf{x}_C)$ with respect to the two markers 10^{-5} and 10^{-2}, $\mathbf{x}_C$ is said to be

1. **totally safe** if $\pi_0(\mathbf{x}_C) \leq 10^{-5}$,
2. **relatively safe** if $10^{-5} < \pi_0(\mathbf{x}_C) < 10^{-2}$,
3. **unsafe** if $\pi_0(\mathbf{x}_C) \geq 10^{-2}$.

In this example, there are 5 controlled and 26 uncontrolled variables, so that $C \subset \mathbb{R}^5, E \subset \mathbb{R}^{26}$. From budget point of view, experts consider that a set of about 400 representative points of C are enough to cover consistently the domain C. On the other hand, the computation of 800000 trajectories takes about 4 days which is considered reasonable. On the basis of these indications, the maximum amount of available calls to the simulator is $N = 2000$ per point.

Estimation strategy

Our estimation strategy which applies iteratively to each point of the set of representative points, has two steps each of which uses half of the calls budget: $m = n = \frac{N}{2} = 1000$. Let $\mathbf{x}_C \in C$ be the current point of interest that we suppose fixed. For any $\mathbf{x} \in E$, $f(\mathbf{x}) = f(\mathbf{x}_C, \mathbf{x})$ is set, recovering the notation introduced in the first part of the paper.

1. At the first stage, a gaussian process is built as explained in (6.2), on the basis of evaluations $f(\mathbf{x}_1), \cdots, f(\mathbf{x}_n) \in \mathbb{R}^n$ of f on $D_n = (\mathbf{x}_1, \cdots, \mathbf{x}_n)$. From Proposition 6.1, we know that π_ρ is a realization of the random variable $\Pi_\rho^{D_n}$ whose mean

$$\mathbb{E}(\Pi_\rho^{D_n}) = \mathbb{E}\left(\Phi\left(\frac{\rho - m_{D_n}(\mathbf{X})}{\sqrt{K_{D_n}(\mathbf{X}, \mathbf{X})}}\right)\right)\,,$$

can be computed accurately.
As stated by (6.10), applying Markov inequality gives, for any $\alpha \in]0;1[$,

$$\mathbb{P}\left(\Pi_\rho^{D_n} \leq \frac{\mathbb{E}(\Pi_\rho^{D_n})}{\alpha}\right) \geq 1-\alpha\,.$$

According to the value of $\mathbb{E}(\Pi_\rho^{D_n})$, we, then, take the following decisions:

- if $\mathbb{E}(\Pi_\rho^{D_n}) \leq \frac{1}{2}10^{-10}$ which leads by (6.10) to $\mathbb{P}\left(\Pi_\rho^{D_n} \leq \frac{10^{-5}}{2}\right) \geq 1-\frac{10^{-5}}{2}$, we qualify the current point $\mathbf{x}_C \in C$ as totally safe,
- if $\mathbb{E}(\Pi_\rho^{D_n}) \geq 10^{-2}$, we conservatively classify $\mathbf{x}_C$ as unsafe,
- if $\frac{1}{2}10^{-10} < \mathbb{E}(\Pi_\rho^{D_n}) < 10^{-2}$ we use a second stage procedure to refine the risk assessment.

2. A million-sample $\mathbf{x}_1, \cdots, \mathbf{x}_M$ of $\mathbf{X}$ is drawn from which we tune κ in such a way that $m = 1000$ of these million elements of E are in $\hat{R}_{\rho,\kappa}$. The resulting points $\mathbf{z}_1, \cdots, \mathbf{z}_m$ are a m-sample $\mathbf{z}_{1:m}$ of realizations of the random variable $\mathbf{Z}$ which follows the importance distribution,

$$\mathbb{P}_{\mathbf{Z}} : A \mapsto \mathbb{P}_{\mathbf{X}}(A|\hat{R}_{\rho,\kappa})\,.$$

By using m calls to the simulator, $\Gamma(f, \mathbf{z}_{1:m}, \rho)$ is computed. Drawn from Proposition 6.4 with setting $\alpha = \beta$, we obtain the bound

$$b(\Gamma(f, \mathbf{z}_{1:m}, \rho), m, \alpha)\mathbb{P}_{\mathbf{X}}(\hat{R}_{\rho,\kappa}) + \frac{c(\kappa)}{\alpha}\,,$$

which is a decreasing function of α.
Let define $\alpha_0 = \min\{\alpha : b(\Gamma(f, \mathbf{z}_{1:m}, \rho), m, \alpha)\mathbb{P}_{\mathbf{X}}(\hat{R}_{\rho,\kappa}) + \frac{c(\kappa)}{\alpha} \leq 2\alpha\}$. For such an α_0, Proposition 6.4 states:

$$\mathbb{P}\left(\Pi_\rho^{D_n} \leq b(\Gamma(F^{D_n}, \mathbf{Z}_{1:m}, \rho), m, \alpha_0)\mathbb{P}_{\mathbf{X}}(\hat{R}_\rho) + \frac{c(\kappa)}{\alpha_0}\right) \geq 1-2\alpha_0\,,$$

which provides $2\alpha_0$ as a $1-2\alpha_0$ confidence upper bound on π_ρ.

Experiments

Three points of C have been tested. Of these cases the first one is known to be a null 0-risk point, while the third one is very unsafe and the second one is in-between.

For benchmarking purpose, besides the simulator calls budget required for the estimation process described in 6.4.2, a 10000-samples of $f(\mathbf{x}_E, \mathbf{X})$ has been computed for each of the three examples. For each case, we began by estimating a Gaussian process on the basis of f-values computed on the points of a 1000 points maximin latin hypercube design $D_n = \{\mathbf{x}_1, \cdots, \mathbf{x}_n\}$. Figures 6.4, 6.5 and 6.6 show the predictive performance of the processes when applied to the benchmark points. These points, which appear in red, are sorted according to their process mean values while the blue curves mark the predicted 3 standard deviation positions around the means. As it appears rather clearly, the dispersion of ther eal values is underestimated by the model: they overflow the blue zone with a frequency ($\sim 5\%$) higher

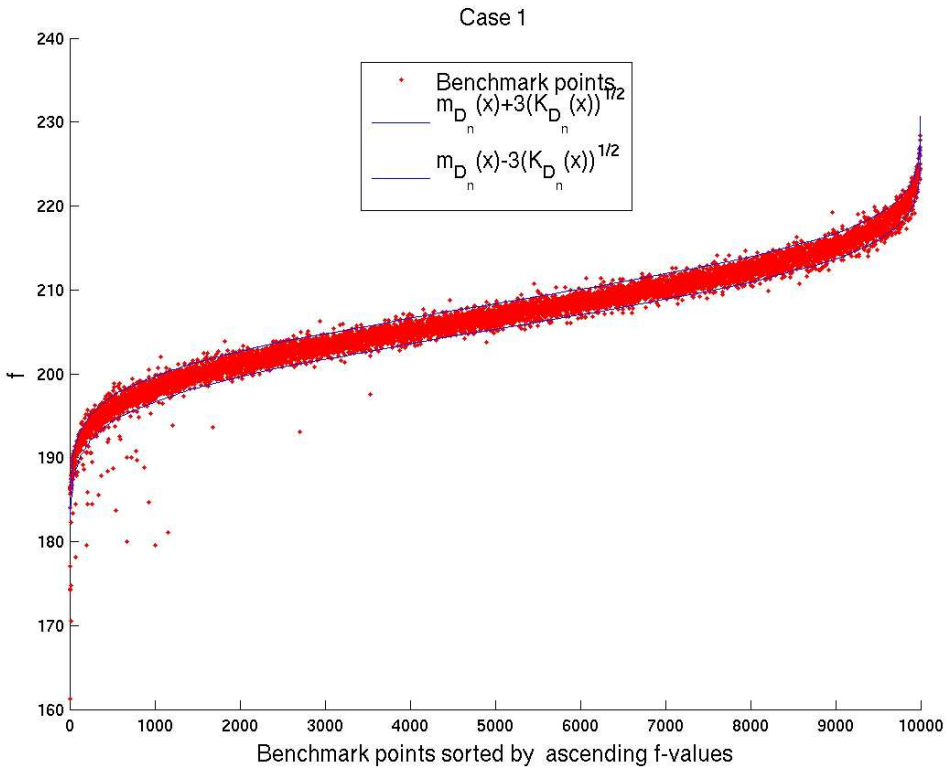


Figure 6.4: Prediction performance case 1

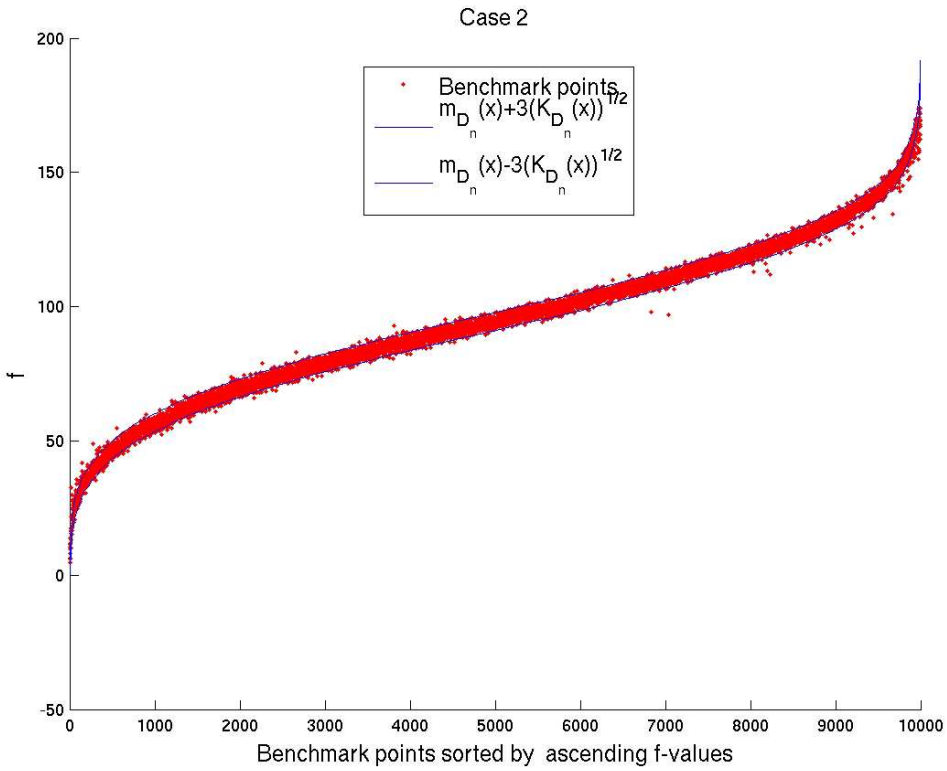


Figure 6.5: Prediction performance case 2

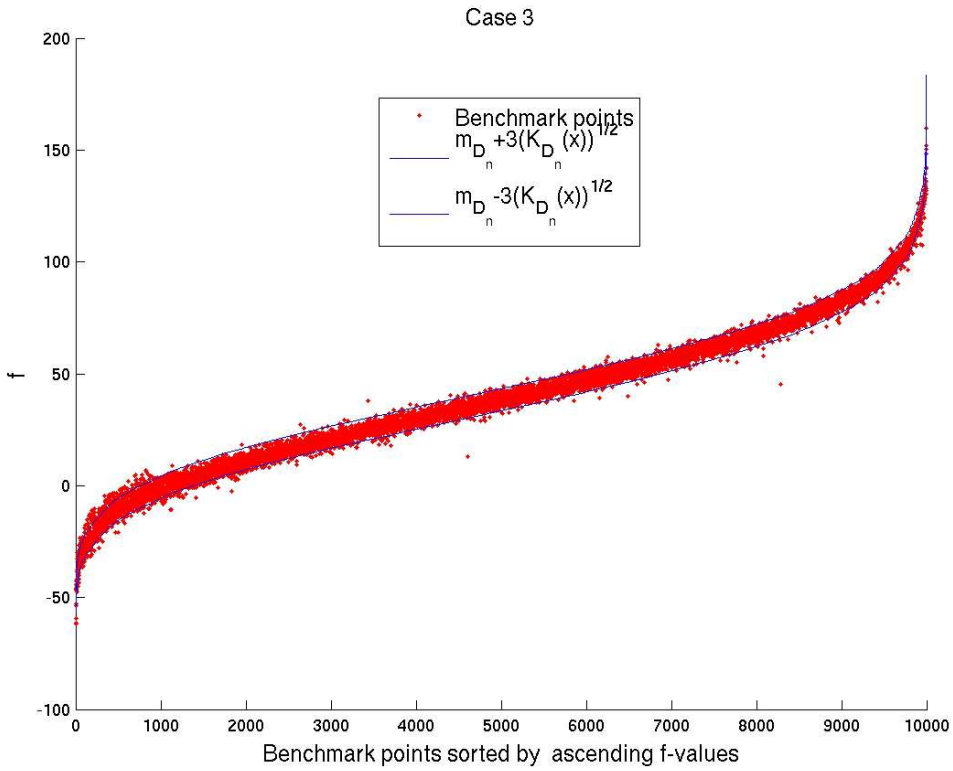


Figure 6.6: Prediction performance case 3

than expected (0.27%). The worse case is the first one, for which large deviations appear for benchmark points with low values of f. In order to obtain bounds from (6.10), we then computed $\mathbb{E}(\Pi_0^{D_n})$ using (6.8):

- In the first case, the massive Monte Carlo procedure leads to a numerically null evaluation of $\mathbb{E}(\Pi_0^{D_n})$ and, as a consequence, to the classification of the related C point as totally safe.
- In the second example, $\mathbb{E}(\Pi_0^{D_n})$ beeing evaluated at $1.68\ 10^{-4}$, we need to proceed the second step.
- $\mathbb{E}(\Pi_0^{D_n}) = 0.103$ in case 3 which is consistent with the 90% confidence interval $[0.0999; 0.1101]$, obtained on benchmark data.

We now applied the procedure second stage to refine collision probability estimation: the obtained confidence upper bound is $1.2\ 10^{-5}$ at confidence level $1 - 1.2\ 10^{-5}$. The benchmark data do not show collision case: a 90% confidence upper bound is $2.3\ 10^{-4}$.

6.5 Dicussion

In this paper, two methods were proposed to estimate and to bound the probability of a rare event which depends on an expensive black-box function. They are both based on a Kriging hypothesis which induces a random interpretation of the probability to estimate. That is why the Bayesian context is natural in this problem and leads to a very accurate estimator. As it is hard to reach the posterior quantiles, it does not achieve as tight upper bounds as the importance sampling method does. The importance sampling method relies on a split in the possible calls to f. We have proposed to use half of the calls to compute a metamodel and half of the calls to draw a sample according to the importance distribution; yet, other way of splitting can be investigated.

As it was noticed on the toy example, there is a variability due to the choice in the design. To reduce it, some points can be added where uncertainties on the prediction of the metamodel are high ($K_{D_n}(\mathbf{x}, \mathbf{x})$ is large) and the probability that f is smaller than ρ is high. It can consist of adding sequentially points of $\hat{R}_{\rho,\kappa}$ where the variance of prediction is the largest, as in Vazquez and Bect (2009) and in Ranjan et al. (2008); Picheny et al. (2010) for contour estimation.

We have dealt with a cross validation method to assess the Kriging hypothesis. However, in the case where the cross validation leads to reconsider this hypothesis, a solution is to extend the confidence interval on the prediction by tuning at hand the parameter σ^2 in equation (6.4). In Bayesian words, it can be called using a less informative prior distribution on f.

We have not manage to compute the posterior variance (given by Proposition 6.1) by using a massive Monte Carlo integration in our examples since it is very small. However, other rare events methods can be investigated since the variance does not depend anymore on f.

Bibliography

Bingham, D., Hengartner, N., Higdon, D., and Kenny, Q. Y. (2006). Variable Selection for Gaussian Process Models in Computer Experiments. *Technometrics*, 48(4):478–490.

Cannamela, C., Garnier, J., and Iooss, B. (2008). Controlled stratification for quantile estimation. *The annals of applied statistics*, 2(4):1554–1580.

Cérou, F. and Guyader, A. (2007a). Adaptive multilevel splitting for rare event analysis. *Stoch. Anal. Appl.*, 25(2):417–443.

Cérou, F. and Guyader, A. (2007b). Adaptive particle techniques and rare event estimation. In *Conference Oxford sur les méthodes de Monte Carlo séquentielles*, volume 19 of *ESAIM Proc.*, pages 65–72. EDP Sci., Les Ulis.

Del Moral, P. and Garnier, J. (2005). Genealogical particle analysis of rare events. *Ann. Appl. Probab.*, 15(4):2496–2534.

Dette, H. and Pepelyshev, A. (2010). Generalized Latin Hypercube Design for Computer Experiments. *Technometrics*, 52(4):421–429.

Fang, K.-T., Li, R., and Sudjianto, A. (2006). *Design and Modeling for Computer Experiments*. Computer Science and Data Analysis. Chapman & Hall/CRC.

Heidelberg, P. (1995). Fast simulation of rare events in queuing and reliability models. *ACM Transactions on Modeling and Computer Simulation*, 5:43–85.

Jones, D. R., Schonlau, M., and Welch, W. J. (1998). Efficient global optimization of expensive black-box functions. *Journal of Global Optimization*, 13(4):455–492.

Joseph, V. R. (2006). Limit kriging. *Technometrics*, 48(4):458–466.

Koehler, J. R. and Owen, A. B. (1996). Computer experiments. In *Design and analysis of experiments*, volume 13 of *Handbook of Statistics*, pages 261–308. North Holland, Amsterdam.

L'Ecuyer, P., Demers, V., and Tuffin, B. (2007). Rare events, splitting, and quasi-monte carlo. *ACM Trans. Model. Comput. Simul.*, 17(2):9.

Li, R. and Sudjianto, A. (2005). Analysis of computer experiments using penalized likelihood in gaussian kriging models. *Technometrics*, 47:111–120.

Mease, D. and Bingham, D. (2006). Latin Hyperrectangle Sampling for Computer Experiments. *Technometrics*, 48(4):467–477.

Morris, M. D. and Mitchell, T. J. (1995). Exploratory designs for computer experiments. *Journal of Statistical Planning and Inference*, 43:381–402.

Munoz Zuniga, M., Garnier, J., Remy, E., and de Rocquigny, E. (2010). Adaptative directional stratification for controlled estimation of the probability of a rare event. Technical report.

Oakley, J. (2004). Estimating percentiles of uncertain computer codes outputs. *Applied Statistics*, 53:83–93.

Picheny, V., Ginsbourger, D., Roustant, O., and Haftka, R. (2010). Adaptive designs of experiments for accurate approximation of a target region. *Journal of Mechanical Design*, 132(7).

Ranjan, P., Bingham, D., and Michailidis, G. (2008). Sequential Experiment Design for Contour Estimation From Complex Computer Codes. *Technometrics*, 50(4):527–541.

Santner, T. J., Williams, B., and Notz, W. (2003). *The Design and Analysis of Computer Experiments*. Springer-Verlag.

Shahabuddin, P. (1995). Rare event simulation in stochastic models. In *WSC '95: Proceedings of the 27th conference on Winter simulation*, pages 178–185, Washington, DC, USA. IEEE Computer Society.

Vazquez, E. and Bect, J. (2009). A sequential bayesian algorithm to estimate a probability of failure. In Elsevier, editor, *15th IFAC Symposium on System Identification (SYSID 2009)*.

Welch, W. J., Buck, R. J., Sack, J., Wynn, H. P., Mitchell, T. J., and Morris, M. D. (1992). Screening, predicting, and computer experiments. *Technometrics*, 34:15–25.

6.6 Confidence bounds for the binomial distribution

Let T be a random variable which follows a binomial distribution with parameters N and p. For a real number $\alpha \in [0,1]$, the upper confidence bound b on p:

$$\mathbb{P}_T(p \leq b(T,N,\alpha)) \geq 1-\alpha$$

is such that:

$$\begin{cases} b = 1 & \text{if } T = N \\ b \text{ is the solution of equation } \sum_{k=0}^{T} \binom{N}{k} b^k (1-b)^{N-k} = \alpha & \text{otherwise} \end{cases}.$$

This upper bound is not in closed form but easily computable.

Chapitre 7

Discussion et perspectives

Nos travaux de thèse concernent principalement le traitement statistique des expériences simulées. La propagation de l'incertitude des entrées du modèle physique aux sorties doit être appréhendée. La majeure difficulté est que le modèle est accessible uniquement par le biais d'une fonction type boîte-noire coûteuse. L'idée clef pour pallier le nombre limité d'appels disponibles à cette fonction, est la construction d'un métamodèle d'évaluation quasi-instantanée l'approchant. Nous avons concentré notre étude sur les interpolateurs à noyaux. Ceux-ci sont interprétés dans le cadre purement fonctionnnel de la théorie de l'approximation (Schaback, 1995; Wendland, 2005) et aussi de manière statistique (Koehler et Owen, 1996; Santner *et al.*, 2003). Ces deux visions sont liées et suivant les cas, l'une ou l'autre est privilégiée. En théorie de l'approximation les bornes d'erreurs données par Madych et Nelson (1992); Schaback (1995) justifient que les plans d'expérience maximin sont adaptés aux interpolateurs à noyaux. Dans la partie 2.3.3, nous avons vu que ces bornes avaient aussi un sens pour l'interprétation statistique. Il est alors intéressant de pouvoir construire de tels plans, ce que nous proposons de faire grâce à un algorithme de recuit simulé dans le chapitre 4. Les applications au problème statistique inverse dans le chapitre 5 et à l'estimation de la probabilité d'événements rares dans le chapitre 6 que nous avons traitées utilisent la vision statistique car elle permet de prendre en compte l'incertitude venant de l'utilisation du métamodèle.

Dans le chapitre 5, l'approximation du modèle physique H par le métamodèle de krigeage $\hat{H}$ (5.6) revient à considérer cette modélisation du problème

$$Y_i = \hat{H}(X_i, d_i) + U_i, \quad 1 \leq i \leq n\,,$$

au lieu de celle définie par l'équation (5.1). Une manière de prendre en compte l'incertitude liée à l'approximation de H par $\hat{H}$ est d'écrire :

$$Y_i = \hat{H}(X_i, d_i) + (H(X_i, d_i) - \hat{H}(X_i, d_i)) + U_i, \quad 1 \leq i \leq n\,.$$

Nous notons respectivement H_j et $\hat{H}_j$ la $j^{\text{ème}}$ sortie du modèle et du métamodèle pour $j = 1, \ldots, p$. Nous rappelons que p est la dimension des vecteurs de sorties Y_i. Le vecteur $E = (E_1, \ldots, E_p)$ de longueur $p \cdot n$ avec

$$E_j = (H_j(X_1, d_1) - \hat{H}_j(X_1, d_1), \ldots, H_j(X_n, d_n) - \hat{H}_j(X_n, d_n)) = (E_{j1}, \ldots, E_{jn})\,,$$

pour $j = 1, \ldots, p$, est considéré comme la réalisation d'un vecteur aléatoire d'après les hypothèses de krigeage (5.5). Il est raisonnable de supposer que ce vecteur est indépendant des

variables aléatoires $U_1, \dots, U_n$ décrivant les erreurs de mesure. Les vecteurs $E_1, \dots, E_p$ sont mutuellement indépendants conditionnellement aux variables $X_1, \dots, X_n, d_1, \dots, d_n$ puisque l'on a construit un métamodèle par sortie sans hypothèse de corrélations entre elles. Conditionnellement aux variables d'entrées non observées $X_1, \dots, X_n$ et observées $d_1, \dots, d_n$, aux points du plan d'expérience noté D et aux réalisations Y_D des processus gaussiens en ces points, le vecteur E suit une loi gaussienne de moyenne le vecteur nul (car les $\hat{H}_j$ sont les moyennes a posteriori des processus) et sa matrice de covariance est diagonale par blocs. Un bloc correspond à la matrice de covariance d'un vecteur E_j. À j fixé, celle-ci est donnée par

$$\mathrm{Cov}(E_{jk}, E_{jl}) = \sigma^2(K(z_k, z_l) + u(z_k)^T(H_D^T\Sigma_{DD}^{-1}H_D)^{-1}u(z_l) - \Sigma_{z_kD}^T\Sigma_{DD}^{-1}\Sigma_{z_lD}),$$

où $u(z) = (F_D^T\Sigma_{DD}^{-1}\Sigma_{zD} - F(z))$ et $z_k = (X_k, d_k)$ pour $k, l = 1, \dots, n$.
L'approximation de H est alors bien prise en compte dans le calcul de la vraisemblance. Cependant, vu que les erreurs ne sont pas indépendantes, il est nécessaire de calculer la vraisemblance pour le jeu entier de données $(\mathbf{X}, \mathbf{Y})$ et non juste la vraisemblance du point (X_i, Y_i), à chaque itération de l'algorithme MCMC pour l'étape S de l'algorithme d'inversion présenté dans la partie 5.3.1. Cela mène à des calculs plus lourds. Des chaînes MCMC qui permettraient de faire varier plusieurs X_i à la fois pourraient être envisagées.
Les premiers essais numériques effectués sont encourageants mais, dans certains cas, l'algorithme peut rencontrer des difficultés de convergence. En effet, des problèmes d'identifiabilité apparaissent si l'incertitude due à l'approximation est trop grande par rapport aux erreurs de mesure. Le cas échéant, ce serait un indicateur sensé pour décider d'enrichir le plan d'expérience initial. Les points ajoutés doivent réduire l'incertitude d'approximation dans les zones où la vraisemblance des données non observées est forte. Une loi a priori sur les données non observées permettraient également de s'extraire des problèmes d'identifiabilité. Cette loi serait fondée sur les connaissances d'experts du modèle physique.

Dans le chapitre 6, une loi a priori est placée sur la fonction f décrivant le modèle physique. Cela revient à poser l'hypothèse que f est la réalisation d'un certain processus gaussien. Par une méthode de validation croisée (voir la partie 6.4.1), cette hypothèse est testée. S'il s'avère qu'elle est trop optimiste, c'est-à-dire que le nombre de points tels que

$$\frac{|f(\mathbf{x}_i) - m_{D_n^{-i}}(\mathbf{x}_i)|}{\sigma^2_{D_n^{-i}}(\mathbf{x}_i)} \notin [-3; 3],$$

est trop important, il faudrait envisager une manière de calibrer la loi a priori sur f pour rendre valide cette hypothèse. Cela reposerait essentiellement sur le réglage "manuel" du paramètre σ^2 dans le modèle 6.4.
La stratégie d'échantillonnage préférentiel a été employée en consacrant une moitié du budget d'appels à f pour construire le métamodèle et l'autre moitié pour former l'estimateur. On pourrait étudier une répartition optimale de ce budget. Le paramètre κ a été fixé pour définir

$$\hat{R}_{\rho,\kappa} = \left\{\mathbf{x} : m_{D_n}(\mathbf{x}) < \rho + \kappa\sqrt{\sigma^2_{D_n}(\mathbf{x})}\right\},$$

qui "contienne $\{\mathbf{x} : F(\mathbf{x}) < \rho\}$ avec un bon niveau de confiance". On pourrait aussi se pencher sur un réglage optimal de ce paramètre.
Un plan d'expérience adaptatif rendrait les stratégies d'estimation plus stables. Un premier métamodèle serait construit à partir d'un plan d'expérience exploratoire. Ensuite, le plan

serait enrichi en ajoutant des points appartenant à $\hat{R}_{\rho,\kappa}$ où l'incertitude est forte.

D'autres applications peuvent tirer parti de la vision statistique (Rutherford, 2006). Une question intéressante est de relier cette modélisation bayésienne à une hypothèse sur l'appartenance de f à un certain espace de fonction. Dans le chapitre 3, nous avons étudié les liens entre les noyaux et les espaces fonctionnels. Cependant, il reste à établir clairement le rapport entre les hypothèses, f est la réalisation d'un processus gaussien et f appartient à un espace de fonctions de telle régularité (Pillai *et al.*, 2007).

Si la dimension des entrées est grande, typiquement supérieure à 50, le nombre de points du plan d'expérience doit être conséquent pour permettre une approximation de bonne qualité. Il est alors coûteux de choisir un plan d'expérience par le biais d'algorithmes de recuit simulé (Morris et Mitchell (1995), chapitre 4) ainsi il faut envisager d'autres stratégies. De plus, le métamodèle d'interpolateur à noyaux est lourd à calculer. Construire un interpolateur à noyaux de manière locale est une solution à explorer dans ce cas.

Bibliographie

Koehler, J. R. et Owen, A. B. (1996). Computer experiments. *In Design and analysis of experiments*, volume 13 de *Handbook of Statistics*, pages 261–308. North Holland, Amsterdam.

Madych, W. R. et Nelson, S. A. (1992). Bounds on multivariate polynomials and exponential error estimates for multiquadric interpolation. *Journal of Approximation Theory*, pages 94–114.

Morris, M. D. et Mitchell, T. J. (1995). Exploratory designs for computer experiments. *Journal of Statistical Planning and Inference*, 43:381–402.

Pillai, N. S., Wu, Q., Liang, F., Mukherjee, S. et Wolpert, R. L. (2007). Characterizing the function space for bayesian kernel models. *Journal of Machine Learning Research*, 8:1769–1797.

Rutherford, B. (2006). A response-modeling alternative to surrogate models for support in computational analyses. *Reliability Engineering & System Safety*, 91(10-11):1322–1330.

Santner, T. J., B., W. et W., N. (2003). *The Design and Analysis of Computer Experiments*. Springer-Verlag.

Schaback, R. (1995). Error estimates and condition numbers for radial basis function interpolation. *Advances in Computational Mathematics*, 3:251–264.

Wendland, H. (2005). *Scattered data approximation*, volume 17 de *Cambridge Monographs on Applied and Computational Mathematics*. Cambridge University Press, Cambridge.

Résumé

Cette thèse se place dans le cadre des expériences simulées auxquelles on a recours lorsque des expériences physiques ne sont pas réalisables. Une expérience simulée consiste à évaluer une fonction déterministe type boîte-noire coûteuse qui décrit un modèle physique. Les entrées de ce modèle, entachées d'incertitude, forment un vecteur aléatoire. Cela implique que les sorties que nous souhaitons étudier sont aléatoires. Une technique standard pour rendre possibles de nombreux traitements statistiques, est de remplacer la fonction type boîte-noire par un métamodèle d'évaluation quasi-instantanée l'approchant.
Nous nous concentrons plus particulièrement sur les métamodèles d'interpolateurs à noyaux dont nous étudions la construction et l'utilisation. Dans ce cadre, une première contribution est la proposition d'une définition plus générale de noyau conditionnellement positif qui permet une vraie généralisation du concept de noyau défini positif et des théorèmes associés. Nous donnons ensuite, dans une deuxième contribution, un algorithme de construction de plans d'expérience dans des domaines éventuellement non hypercubiques suivant un critère maximin pertinent pour ces métamodèles. Dans une troisième contribution, nous traitons un problème statistique inverse en utilisant un métamodèle d'interpolateurs à noyaux dans un algorithme stochastique EM puisque le modèle liant les entrées aux sorties est de type boîte-noire coûteux. Enfin, nous proposons aussi, dans la dernière contribution, l'utilisation d'un tel métamodèle pour développer deux stratégies d'estimation et de majoration de probabilités d'événements rares dépendant d'une fonction type boîte-noire coûteuse.
Mots-clés : Expériences simulées, métamodèles, interpolation à noyaux, krigeage, plans d'expériences numériques, problème statistique inverse, événements rares.

Abstract

This work is in the field of computer experiment which is the natural context when physical experiments are impracticable. A computer experiment consists of an evaluation of an expensive black-box function which describes a physical model. The input variables are treated as a random vector since they suffer from uncertainties. This implies that the outputs of the model which are focused on, are random. In order to make statistical analyses tractable, the black-box function can be replaced with a metamodel which approximates it and is fast to compute.
We especially focus on metamodeling with kernel interpolation and the use of these metamodels. In this context, the first contribution consists of proposing a more general definition of a conditionally positive definite kernel which allows a full generalization of the concept of positive definite kernel and its associated theorems. We provide, in a second contribution, an algorithm to obtain numerical designs of experiments according to a maximin criterion which is sensible for these metamodels. In a third contribution, an inverse statistical problem is treated by using a kernel interpolation metamodel into a stochastic EM algorithm since the outputs depend on the inputs through an expensive black-box model. In the last contribution, we propose two strategies relying also on such a metamodel to estimate and to upper bound the probability of rare events based on the outputs of an expensive black-box function.
Keywords : Computer experiments, metamodeling, kernel interpolation, Kriging, numerical designs of experiment, inverse statistical problem, rare events.