Surface de réponse : Plans pour l'approximation de codes numériques

Jean-Marc Azaïs

Modélisation Aléatoire à Finalité Industrielle et Appliquée, LSP
Intitut de Mathématiques Toulouse

MAFIA

Outline

Le Kriegeage

Un code peut être vu comme une fonction

$$Y = F(T)$$

T est de dimension p Pour simplifier : Y réel.
Le **kriegeage** consiste, à partir de n observations de (Y, X) à poser un modèle de régression de Y sur une base ϕ_j (polynômes, fourier, ondelettes) avec un "bruit coloré"

$$Y(t_i) = \sum_{j=1}^{p} \theta_j \phi_j(t_i) + Z(t_i),$$

où $Z(.)$ est un processus gaussien centré que l'on suppose le plus souvent stationnaire et isotrope pour simplifier l'identification.

Solutions bien connues : Si X est la matrice du modèle si Σ est à une constante σ^2 près la matrice de variance des erreurs,**alors** la prédiction au point t_0 (qui est nouveau) est

$$\widehat{Y}(t_0) = \sum_{j=1}^{p} \widehat{\theta}_j \phi_j(t_0)$$

où le vecteur $\widehat{\theta}$ est estimé par les equations des moindres carrés généralisés.

$$\widehat{\theta} = (X'\Sigma^{-1}X)^{-1}X'\Sigma^{-1}Y.$$

On peut trouver une description détaillée ainsi que l'approche bayesienne dans
Koehler and Owen (1996) "computer experiments".

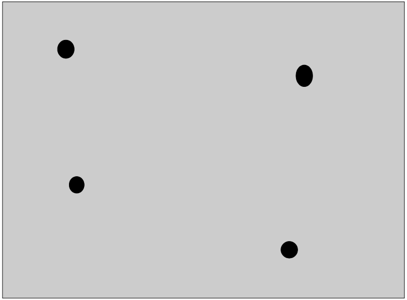
Qu'est qu'un bon plan pour un kriegeage?

Un critère utilisé est l'erreur moyenne
"integrate mean square error" IMSE
calculé sur un domaine par exemple rectangulaire. Une reference récente est
Sachs Shiller and Welch "design for computer experiments" Technometrics 1989.

Donne lieu à des calculs très importants et des optimisation numériques uniquement.

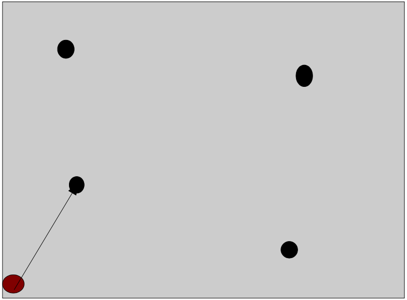
Plan minimax

On minimise la distance maximale d'un point quelconque du domaine au point le plus proche du plan d'expérience.



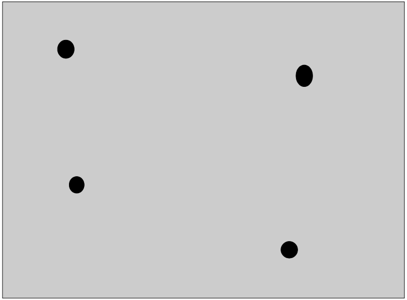
Critère minimax

On minimise la distance maximale d'un point quelconque du domaine au point le plus proche du plan d'expérience.



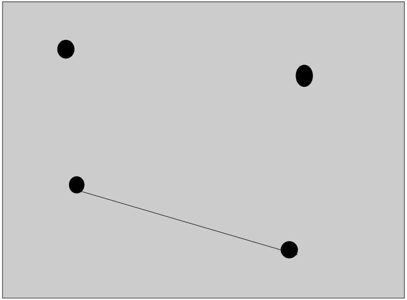
Plan maximin

On maximise la distance minimale entre deux points du plan d'expérience.



Critère maximin

On maximise la distance minimale entre deux points du plan d'expérience.



Pas de résultats théoriques

On peut donc utiliser des classes de plans

- des Hypercubes latins, McKay, Beckman and Conover (1979)
- Des "orthogonal arrays" Owen (1994) ou Tang (1993)
- Et mesurer a posteriori les performances.

Pas de résultats théoriques

On peut donc utiliser des classes de plans

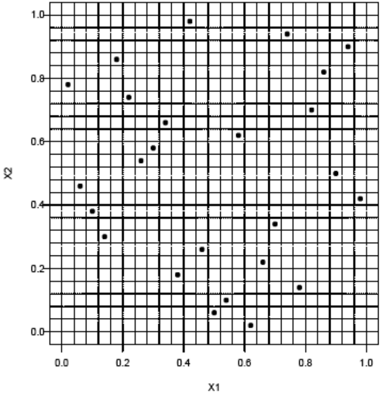
- des Hypercubes latins, McKay, Beckman and Conover (1979)
- Des "orthogonal arrays" Owen (1994) ou Tang (1993)
- Et mesurer a posteriori les performances.

Pas de résultats théoriques

On peut donc utiliser des classes de plans

- des Hypercubes latins, McKay, Beckman and Conover (1979)
- Des "orthogonal arrays" Owen (1994) ou Tang (1993)
- Et mesurer a posteriori les performances.

Hypercubes Latins



ure 13: 25 points of a Latin hypercube sample. The range of each input variable ma
titioned into 25 bins of equal width, drawn here with horizontal and vertical dotted lines
h such bin contains one of the points.

Orthogonal Array

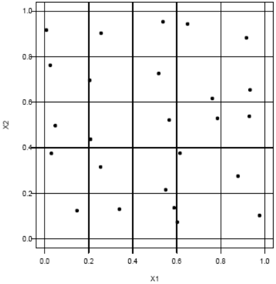


Figure 14: 25 points of a randomly centered randomized orthogonal array. For
five) variables that are plotted, there is one point in each reference square.

Principe : En dimension p
On découpe chaque coordonnée en d sous intervalles
On obtient d^p "petits-rectangle"
Dans cet ensemble de rectangle, vu comme un plan factoriel complet, on ne retient qu'une fraction correspondant à une fraction régulière de résolution r.

L'ensemble obtenu est de poids $r-1$: **les tables de contingence de $r-1$ coordonnées sont équilibrées.**

A la fin on tire au hasard dans chaque "petits-rectangle"
$r = 2$ **Correspond au cas de Hypercubes**

Qui peut le plus, peut le moins : les orthogonal arrays englobent les hypercubes.

Question ouverte

Si on retient le critère minimax ou maximin ou IMSE. Quels sont les bons plans parmi les "orthogonal array" ?

Quel est le bon choix de d et de r en fonction de p?

A même valeurs de d et de r, a t'on des choix meilleurs que d'autres?

La problématique

En général, en début d'étude, on dispose de beaucoup de facteurs dont on sait que **seule un faible partie est présente** et qu'a fortiori les inter-actions doubles ou plus sont absentes.

Le plan screening cherche à intégrer le plus de facteurs possible sous la contrainte

$$p < n$$

Plan de Plackett et Burman

	Pattern	X1	X2	X3	X4	X5	X6	X7	X8	X9	X10	X11
1	+++++++++++	+1	+1	+1	+1	+1	+1	+1	+1	+1	+1	+1
2	-+-+++—+-	-1	+1	-1	+1	+1	+1	-1	-1	-1	+1	-1
3	–+-+++—+	-1	-1	+1	-1	+1	+1	+1	-1	-1	-1	+1
4	+–+-+++—	+1	-1	-1	+1	-1	+1	+1	+1	-1	-1	-1
5	-+–+-+++–	-1	+1	-1	-1	+1	-1	+1	+1	+1	-1	-1
6	–+–+-+++-	-1	-1	+1	-1	-1	+1	-1	+1	+1	+1	-1
7	—+–+-+++	-1	-1	-1	+1	-1	-1	+1	-1	+1	+1	+1
8	+—+–+-++	+1	-1	-1	-1	+1	-1	-1	+1	-1	+1	+1
9	++—+–+-+	+1	+1	-1	-1	-1	+1	-1	-1	+1	-1	+1
10	+++—+–+-	+1	+1	+1	-1	-1	-1	+1	-1	-1	+1	-1
11	-+++—+–+	-1	+1	+1	+1	-1	-1	-1	+1	-1	-1	+1
12	+-+++—+–	+1	-1	+1	+1	+1	-1	-1	-1	+1	-1	-1

12 unités, 11 facteurs

Méthode de Dantzig pour modèles sporadiques

Dans un modèle linéaire **saturé**

$$Y = X\theta + \epsilon$$

ou le nombre p de paramètres est bien plus grand que le nombre d'observations n

a priori on est mort !

Mais si on sait que le vecteur θ est "sporadique" ou "creux" ou "sparse" : son nombre de coeffs non-nuls p_0 est inférieur à n

alors il y a un espoir

Si la matrice d'information $X'_T X_T$ de tout sous ensemble de taille inferieure à n a ses valeur propres λ_i qui vérifient

$$1 - S < \lambda_i < 1 + S \qquad |S| < 1$$

alors le θ qui minimise

$$\|\theta\|_1 = |\theta_1| + \cdots + |\theta_p| \qquad \text{sous la condition } |X' r\| \leq \sqrt{2 log(p)}$$

où r est le vecteur des résidus,
a,**avec grande proba**, les bons zeros.

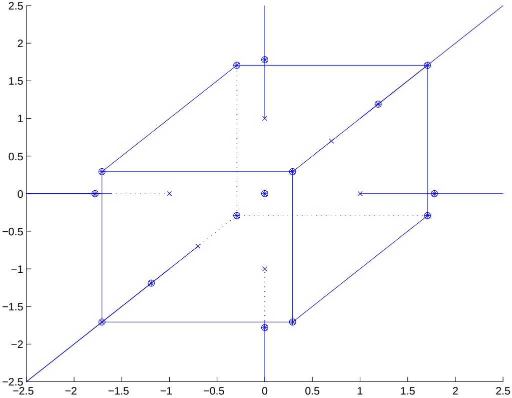
Candes Tao (2006)

La question ouverte

Dans leur article, Candes et Tao proposent un plan aléatoire basé sur des directions tirées au hasard de manière uniforme.

C'est un bonne stratégie pour de très grands dispositifs.

Souvent dans une logique de plan d'expériences ce n'est pas très satisfaisant. D'ou l'idée d'utiliser des plans isovariants et voir si on obtient les mêmes résultats.


2.5
2
1.5
1
0.5
0
−0.5
−1
−1.5
−2
−2.5
−2.5
−2
−1.5
−1
−0.5
0
0.5
1
1.5
2
2.5