

Codes numériques et Plans d'expériences

Journées "Incertitude et Simulation"

CEA DIF Bruyères-le-Châtel - 3 et 4 octobre 2007

Jean-Pierre Gauchi

INRA

Unité de Mathématiques et Informatique Appliquées de Jouy-en-Josas

- Le contexte
- La question posée
- Trois stratégies envisageables
 - Stratégie 1: Plans d'exploration de l'espace des entrées
 - Stratégie 2: Postulation d'un premier métamodèle: avec ou sans plan optimal
 - Stratégie 3: Plans séquentiels optimaux
- Conclusion
- Perspectives

- On dispose d'un **code numérique** qui calcule la valeur y_i (éventuellement multidimensionnelle) de la sortie (réponse) y d'un système physique **compliqué** pour la valeur **fixée** x_i (s -dimensionnelle) des s entrées (facteurs) : on suppose que la relation f qui relie la sortie aux entrées via des paramètres θ n'est pas analytique ou est **inconnue**.

- On dispose d'un **code numérique** qui calcule la valeur y_i (éventuellement multidimensionnelle) de la sortie (réponse) y d'un système physique **compliqué** pour la valeur **fixée** x_i (s -dimensionnelle) des s entrées (facteurs) : on suppose que la relation f qui relie la sortie aux entrées via des paramètres θ n'est pas analytique ou est **inconnue**.
- Le système est représenté par des **équations non linéaires, EDO, EDP, stochastiques**, etc...
⇒ l'obtention d'un seul résultat de calcul avec ce modèle "physique" prend beaucoup de temps

- On dispose d'un **code numérique** qui calcule la valeur y_i (éventuellement multidimensionnelle) de la sortie (réponse) y d'un système physique **compliqué** pour la valeur **fixée** x_i (s -dimensionnelle) des s entrées (facteurs) : on suppose que la relation f qui relie la sortie aux entrées via des paramètres θ n'est pas analytique ou est **inconnue**.
- Le système est représenté par des **équations non linéaires, EDO, EDP, stochastiques**, etc...
⇒ l'obtention d'un seul résultat de calcul avec ce modèle "physique" prend beaucoup de temps
- En conséquence tout ce qui nous intéresse devient **difficile à réaliser**:

- On dispose d'un **code numérique** qui calcule la valeur y_i (éventuellement multidimensionnelle) de la sortie (réponse) y d'un système physique **compliqué** pour la valeur **fixée** x_i (s -dimensionnelle) des s entrées (facteurs) : on suppose que la relation f qui relie la sortie aux entrées via des paramètres θ n'est pas analytique ou est **inconnue**.
- Le système est représenté par des **équations non linéaires, EDO, EDP, stochastiques**, etc...
⇒ l'obtention d'un seul résultat de calcul avec ce modèle "physique" prend beaucoup de temps
- En conséquence tout ce qui nous intéresse devient **difficile à réaliser**:
 - **Explorer** "correctement" l'espace des entrées.

- On dispose d'un **code numérique** qui calcule la valeur y_i (éventuellement multidimensionnelle) de la sortie (réponse) y d'un système physique **compliqué** pour la valeur **fixée** x_i (s -dimensionnelle) des s entrées (facteurs) : on suppose que la relation f qui relie la sortie aux entrées via des paramètres θ n'est pas analytique ou est **inconnue**.
- Le système est représenté par des **équations non linéaires, EDO, EDP, stochastiques**, etc...
⇒ l'obtention d'un seul résultat de calcul avec ce modèle "physique" prend beaucoup de temps
- En conséquence tout ce qui nous intéresse devient **difficile à réaliser**:
 - **Explorer** "correctement" l'espace des entrées.
 - Construire un **métamodèle** η de ce modèle physique en vue d'**optimisation ultérieure, d'ingénierie robuste, d'analyse de sensibilité, analyse probabiliste des incertitudes**, etc...

- Comment explorer l'espace des entrées χ selon que l'**on postule ou non** un métamodèle η qui soit d'approximation "raisonnable" du modèle "physique", qui ne soit toutefois pas trop compliqué pour permettre des calculs rapides, et qui permette in fine d'atteindre un ou plusieurs des objectifs précédents ?

- Comment explorer l'espace des entrées χ selon que l'**on postule ou non** un métamodèle η qui soit d'approximation "raisonnable" du modèle "physique", qui ne soit toutefois pas trop compliqué pour permettre des calculs rapides, et qui permette in fine d'atteindre un ou plusieurs des objectifs précédents ?
- On traitera ici cette question dans le contexte:

- Comment explorer l'espace des entrées χ selon que l'**on postule ou non** un métamodèle η qui soit d'approximation "raisonnable" du modèle "physique", qui ne soit toutefois pas trop compliqué pour permettre des calculs rapides, et qui permette in fine d'atteindre un ou plusieurs des objectifs précédents ?
- On traitera ici cette question dans le contexte:
 - Le temps de calcul unitaire d'une expérience numérique est important.

- Comment explorer l'espace des entrées χ selon que l'**on postule ou non** un métamodèle η qui soit d'approximation "raisonnable" du modèle "physique", qui ne soit toutefois pas trop compliqué pour permettre des calculs rapides, et qui permette in fine d'atteindre un ou plusieurs des objectifs précédents ?
- On traitera ici cette question dans le contexte:
 - Le temps de calcul unitaire d'une expérience numérique est important.
 - La dimension s de l'espace des entrées $\chi \subset \mathbb{R}^s$ est grande.

① Pas de base d'exemples et pas de métamodèle η

\implies On réalise quelques expériences numériques réparties "**au mieux**" dans l'espace χ des entrées, ce qui conduira à construire une première base d'exemples.

❶ Pas de base d'exemples et pas de métamodèle η

⇒ On réalise quelques expériences numériques réparties "**au mieux**" dans l'espace χ des entrées, ce qui conduira à construire une première base d'exemples.

❷ On dispose d'une base d'exemples et la démarche adoptée est non-séquentielle

⇒ On estime un métamodèle η en optimisant un critère de modèle, typiquement la minimisation du maximum ou de la moyenne sur χ de l'**erreur quadratique de prédiction**, avec ou sans plan optimal.

❶ Pas de base d'exemples et pas de métamodèle η

⇒ On réalise quelques expériences numériques réparties "**au mieux**" dans l'espace χ des entrées, ce qui conduira à construire une première base d'exemples.

❷ On dispose d'une base d'exemples et la démarche adoptée est non-séquentielle

⇒ On estime un métamodèle η en optimisant un critère de modèle, typiquement la minimisation du maximum ou de la moyenne sur χ de l'**erreur quadratique de prédiction**, avec ou sans plan optimal.

❸ On dispose d'une base d'exemples et la démarche adoptée est séquentielle

⇒ Après avoir estimé un premier métamodèle on choisit l'expérience suivante à l'endroit dans χ où l'**incertitude est maximale** si on sait calculer celle-ci (!?)

Stratégie 1: Plans d'exploration de l'espace des entrées

Difficultés intrinsèques aux grands espaces

- On souhaite répartir "**au mieux**" (?) les expériences dans χ .

Stratégie 1: Plans d'exploration de l'espace des entrées

Difficultés intrinsèques aux grands espaces

- On souhaite répartir "**au mieux**" (?) les expériences dans χ .
- Une répartition "idéale" serait une **répartition uniforme** de N expériences aux noeuds d'un maillage régulier serré sur χ (plan factoriel complet).
→ Solution non envisageable en général compte tenu de l'**explosion** du nombre $N = n^s$ où n est le nombre de niveaux choisi par entrée.

Stratégie 1: Plans d'exploration de l'espace des entrées

Difficultés intrinsèques aux grands espaces

- On souhaite répartir "**au mieux**" (?) les expériences dans χ .
- Une répartition "idéale" serait une **répartition uniforme** de N expériences aux noeuds d'un maillage régulier serré sur χ (plan factoriel complet).
→ Solution non envisageable en général compte tenu de l'**explosion** du nombre $N = n^s$ où n est le nombre de niveaux choisi par entrée.
- Un espace de grande dimension est vraiment grand et réserve des surprises !! Deux illustrations:

Stratégie 1: Plans d'exploration de l'espace des entrées

Difficultés intrinsèques aux grands espaces

- On souhaite répartir "**au mieux**" (?) les expériences dans χ .
- Une répartition "idéale" serait une **répartition uniforme** de N expériences aux noeuds d'un maillage régulier serré sur χ (plan factoriel complet).
→ Solution non envisageable en général compte tenu de l'**explosion** du nombre $N = n^s$ où n est le nombre de niveaux choisi par entrée.
- Un espace de grande dimension est vraiment grand et réserve des surprises !! Deux illustrations:
 - Pour un cube unité de dimension 20, son volume vaut 1 et la sphère qui s'y inscrit a un volume de $2.5 \times 10^{-8} \implies$ un échantillonnage aléatoire uniforme simple place les aléas dans les coins !

Stratégie 1: Plans d'exploration de l'espace des entrées

Difficultés intrinsèques aux grands espaces

- On souhaite répartir "**au mieux**" (?) les expériences dans χ .
- Une répartition "idéale" serait une **répartition uniforme** de N expériences aux noeuds d'un maillage régulier serré sur χ (plan factoriel complet).
→ Solution non envisageable en général compte tenu de l'**explosion** du nombre $N = n^s$ où n est le nombre de niveaux choisi par entrée.
- Un espace de grande dimension est vraiment grand et réserve des surprises !! Deux illustrations:
 - Pour un cube unité de dimension 20, son volume vaut 1 et la sphère qui s'y inscrit a un volume de $2.5 \times 10^{-8} \implies$ un échantillonnage aléatoire uniforme simple place les aléas dans les coins !
 - Pour approximer à environ 10% près la densité de probabilité d'une gaussienne de dimension 10 avec des noyaux gaussiens identiques, Silverman (1986) a montré qu'il fallait 10^6 noyaux (cependant avec des noyaux variables ce nombre diminue nettement, mais reste grand) \implies le remplissage de l'espace par des noyaux radiaux ne semble plus une démarche si pertinente !

Stratégie 1: Plans d'exploration de l'espace des entrées

Les "Latin Hypercube Sampling" (1)

- Aujourd'hui: **plusieurs familles de "space filling designs" (plans d'expériences de remplissage) co-existent** dont certaines sont liées à d'anciens dispositifs expérimentaux (notamment ceux pratiqués par les agronomes de la première moitié du $XX^{\text{ème}}$ siècle).

Stratégie 1: Plans d'exploration de l'espace des entrées

Les "Latin Hypercube Sampling" (1)

- Aujourd'hui: **plusieurs familles de "space filling designs" (plans d'expériences de remplissage) co-existent** dont certaines sont liées à d'anciens dispositifs expérimentaux (notamment ceux pratiqués par les agronomes de la première moitié du $XX^{\text{ème}}$ siècle).
- En 1979, McKay et al. proposent le **LHS** ("Latin Hypercube Sampling"), **l'échantillonnage par hypercube latin**.

Stratégie 1: Plans d'exploration de l'espace des entrées

Les "Latin Hypercube Sampling" (1)

- Aujourd'hui: **plusieurs familles de "space filling designs" (plans d'expériences de remplissage) co-existent** dont certaines sont liées à d'anciens dispositifs expérimentaux (notamment ceux pratiqués par les agronomes de la première moitié du $XX^{\text{ème}}$ siècle).
- En 1979, McKay et al. proposent le **LHS** ("Latin Hypercube Sampling"), **l'échantillonnage par hypercube latin**.
- Propriété majeure démontrée par les auteurs: **la variance** de la moyenne empirique des N points y_i du LHS est **inférieure** à la variance σ^2 / N de la moyenne obtenue à partir des N points y_i de l'échantillonnage aléatoire uniforme simple, si la fonction considérée (η) est **monotone** en chacun de ses arguments (entrées).

Stratégie 1: Plans d'exploration de l'espace des entrées

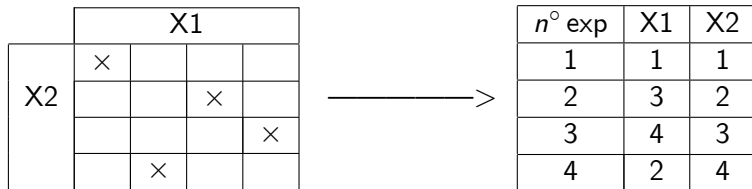
Les "Latin Hypercube Sampling" (1)

- Aujourd'hui: **plusieurs familles de "space filling designs" (plans d'expériences de remplissage) co-existent** dont certaines sont liées à d'anciens dispositifs expérimentaux (notamment ceux pratiqués par les agronomes de la première moitié du $XX^{\text{ème}}$ siècle).
- En 1979, McKay et al. proposent le **LHS** ("Latin Hypercube Sampling"), **l'échantillonnage par hypercube latin**.
- Propriété majeure démontrée par les auteurs: **la variance** de la moyenne empirique des N points y_i du LHS est **inférieure** à la variance σ^2 / N de la moyenne obtenue à partir des N points y_i de l'échantillonnage aléatoire uniforme simple, si la fonction considérée (η) est **monotone** en chacun de ses arguments (entrées).
- On peut considérer aussi le **LHD** ("Latin Hypercube Design"): le point n'est pas tiré aléatoirement dans la case, mais positionné au milieu.

Stratégie 1: Plans d'exploration de l'espace des entrées

Les "Latin Hypercube Sampling" (2)

Exemple: un LHD pour **2 facteurs à 4 niveaux par facteur**:

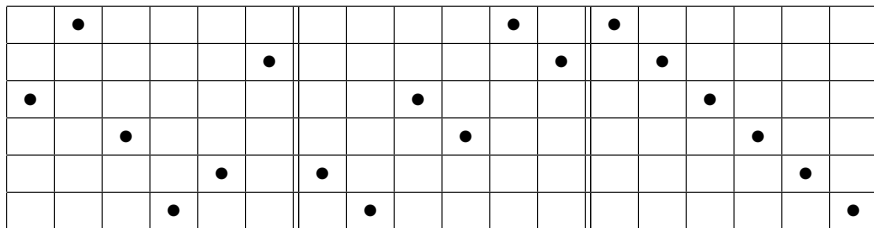


- Chaque facteur présente le **même** nombre n de niveaux
- Les marges sont **équilibrées**
- Chacun des niveaux est pris une fois et une seule par chaque facteur: chacune des colonnes du plan est donc une **permutation** de $\{1, 2, \dots, n\}$

Stratégie 1: Plans d'exploration de l'espace des entrées

Les "Latin Hypercube Sampling" (3)

Mais tous les LHD ne semblent pas convenir: exemple de trois LHD possibles pour **2 facteurs à 6 niveaux** par facteur:



↪ On peut les distinguer en calculant différents critères: **maximin**, **indépendance**, **discrédance**.

Stratégie 1: Plans d'exploration de l'espace des entrées

Critères

- Critère **maximin** (indicateur de remplissage) = distance D_{\min} entre les deux points les plus proches, **à maximiser**
- Critère d'**indépendance** des facteurs 2 à 2 = déterminant $|R|$ de la matrice des corrélations, **à maximiser** (il vaut 1 en l'absence de toute corrélation)
- Critère de **discrépance** = distance à une répartition uniforme, **à minimiser** (on utilise ici la discrèpance CL_2 , en anglais "centered L_2 discrepancy")

Stratégie 1: Plans d'exploration de l'espace des entrées

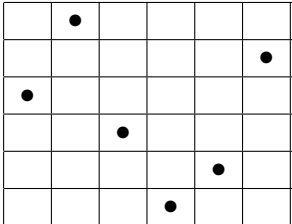
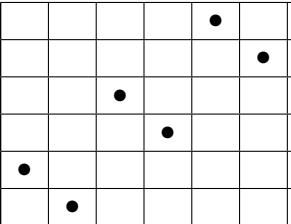
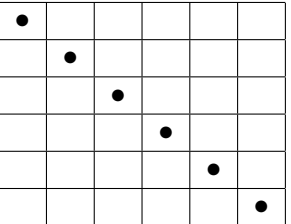
Discrépance

$$CL_2^2(\xi_N) = \left(\frac{13}{12}\right)^s - \frac{2}{N} \sum_{k=1}^N \prod_{j=1}^s \left[1 + \frac{1}{2} |x_{kj} - 0.5| - \frac{1}{2} |x_{kj} - 0.5|^2\right] \\ + \frac{1}{N^2} \sum_{k=1}^N \sum_{j=1}^N \prod_{i=1}^s \left[1 + \frac{1}{2} |x_{ki} - 0.5| + \frac{1}{2} |x_{ji} - 0.5| - \frac{1}{2} |x_{ki} - x_{ji}|\right]$$

Stratégie 1: Plans d'exploration de l'espace des entrées

Les "Latin Hypercube Sampling" (4)

Calculons ces critères pour nos trois LHD:

		
$D_{\min} = 1.414$	$D_{\min} = 1.414$	$D_{\min} = 1.414$
$ R = 0.934$	$ R = 0.313$	$ R = 0$
$CL_2 = 9.43 \times 10^{-3}$	$CL_2 = 14.3 \times 10^{-3}$	$CL_2 = 16.9 \times 10^{-3}$

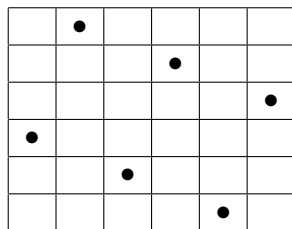
↪ En pratique: on tire un grand nombre de LHS et on les **trie** selon le ou les critères choisis comme prioritaires.

↪ Variantes: LHD **symétriques**, **orthogonaux**.

Stratégie 1: Plans d'exploration de l'espace des entrées

Le LHD symétrique

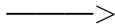
Un exemple à **6 expériences pour 2 facteurs**



$$D_{\min} = 2.24$$

$$|R| = 0.96$$

$$CL_2 = 8.14 \times 10^{-3}$$



⇒	1	3	⇐
→	2	6	←
⇒	3	2	⇐
⇒	4	5	⇐
→	5	1	←
⇒	6	4	⇐

→ $\Sigma = 7$

Stratégie 1: Plans d'exploration de l'espace des entrées

Le LHD orthogonal

Un exemple à **8 expériences pour 4 facteurs**

($D_{\min} = 6.48$; $|R| = 1$; $CL_2 = 24.1 \times 10^{-3}$)

5	3	1	7		+1	-2	-4	+3
6	5	2	1		+2	+1	-3	-4
7	1	6	4		+3	-4	+2	-1
8	7	5	6		+4	+3	+1	+2
1	2	4	3	→	-4	-3	-1	-2
2	8	3	5		-3	+4	-2	+1
3	4	7	8		-2	-1	+3	+4
4	6	8	2		-1	2	+4	-3

⇒ **indépendance** entre les estimateurs des effets linéaires et quadratiques, et aussi entre ceux des interactions et des effets linéaires.

↪ Problème: N **croît très vite** avec le nombre d'entrées.

Stratégie 1: Plans d'exploration de l'espace des entrées

Le plan "orthogonal array" (OA)

OA est une fraction régulière de résolution r prise dans un plan factoriel complet construit sur les n niveaux de s facteurs:

↔ Les marges de dimension $r - 1$ (force) sont **équilibrées**.

Un exemple d'OA (●) ($D_{\min} = 2$; $|R| = 1$; $CL_2 = 28.8 \times 10^{-3}$) et de LHD (×) ($D_{\min} = 1.414$; $|R| = 0.98$; $CL_2 = 4.3 \times 10^{-3}$) dérivé

							×	
	●	×		●			●	
				×				
×								
	●		×	●			●	
						×		
					×			
	●×			●			●	
								×

↔ Les $\langle , \rangle = 0$ pour OA, mais LHD explore mieux ! : Quel est l'objectif ?

Stratégie 1: Plans d'exploration de l'espace des entrées

Les plans minimax et maximin

Proposés par Johnson et al. (1990):

- **Minimax**: on minimise la distance maximale d'un point quelconque du domaine (un point candidat) au point le plus proche du plan d'expériences.
- **Maximin**: on maximise la distance minimale entre deux points du plan d'expériences: les plans maximin sont **plus faciles à construire** que les plans minimax.

Stratégie 1: Plans d'exploration de l'espace des entrées

Les plans Uniformes de Fang, 1990 (1)

Principe

↔ On cherche sur $C^s = [0, 1]^s$ le plan ξ_N de taille N fixée

$$\xi_N = \{x_1, \dots, x_N\}, x_i \in C^s, i = 1, \dots, N$$

tel que la différence $f(x_u) - \eta(x_u), \forall x_u \in C^s$ soit **la plus petite possible**, pour un résultat de calcul $y_u = f(x_u)$.

↔ On choisit un modèle simplissime (en stratégie 1): **le modèle de l'espérance** (densité uniforme sur C^s):

$$E(y) = \int_{C^s} f(x) dx$$

↔ On utilise l'estimateur de $E(y)$:

$$\hat{y}(\xi_N) = \frac{1}{N} \sum_{i=1}^N f(x_i)$$

↔ La variance de $\hat{y}(\xi_N)$, estimée à partir des N points, est utilisée pour comparer les plans.

Stratégie 1: Plans d'exploration de l'espace des entrées

Les plans Uniformes de Fang (2)

Construction

↪ On cherche à rendre **petite** la différence absolue

$$\Delta = |E(y) - \hat{y}(\xi_N)|$$

↪ L'**inégalité de Koksma-Hlawka** fournit une borne supérieure à Δ :

$$\Delta \leq V(f) D(\xi_N)$$

où $V(f)$ est la variation globale (en un certain sens) de la fonction (inconnue) f , et $D(\xi_N)$ est la **discrépance "star"** de ξ_N .

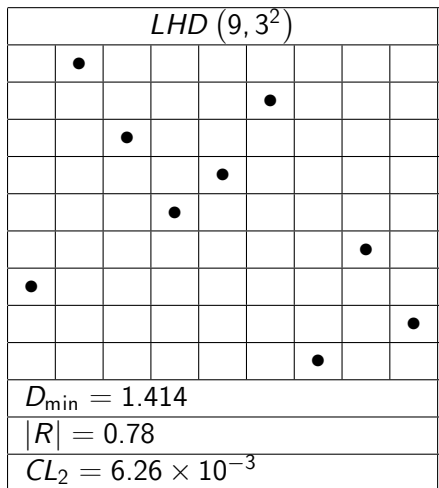
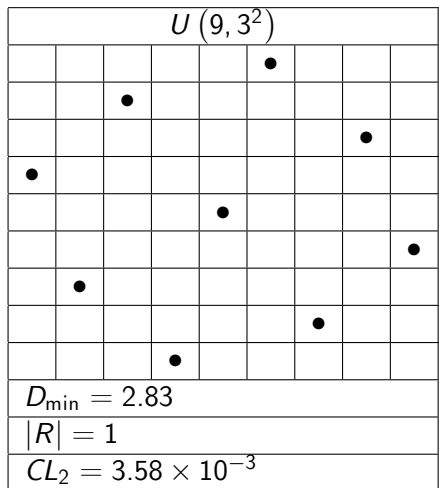
↪ $D(\xi_N)$, à **minimiser**, est donc une **mesure de non-uniformité** (\approx statistique de Kolmogorov-Smirnov).

↪ Plusieurs algorithmes de construction existent, des tables aussi sont disponibles sur le site de Fang.

Stratégie 1: Plans d'exploration de l'espace des entrées

Les plans Uniformes de Fang (3)

Exemple: **9 expériences pour 2 facteurs**



↪ Un Plan Uniforme est un LHD

Stratégie 1: Plans d'exploration de l'espace des entrées

Les LHD quasi-orthogonaux de Cioppa & Lucas (2007)

- Reprise des idées de Ye (1998) sur les LHD orthogonaux à bonnes propriétés mais peu économiques.

Stratégie 1: Plans d'exploration de l'espace des entrées

Les LHD quasi-orthogonaux de Cioppa & Lucas (2007)

- Reprise des idées de Ye (1998) sur les LHD orthogonaux à bonnes propriétés mais peu économiques.
- Algorithme puissant mis au point pour la recherche de ces plans.

Stratégie 1: Plans d'exploration de l'espace des entrées

Les LHD quasi-orthogonaux de Cioppa & Lucas (2007)

- Reprise des idées de Ye (1998) sur les LHD orthogonaux à bonnes propriétés mais peu économiques.
- Algorithme puissant mis au point pour la recherche de ces plans.
- Temps de calcul important mais **tables toutes prêtes à l'emploi** existent sur le site <http://harvest.nps.edu>

Stratégie 1: Plans d'exploration de l'espace des entrées

Les LHD quasi-orthogonaux de Cioppa & Lucas (2007)

- Reprise des idées de Ye (1998) sur les LHD orthogonaux à bonnes propriétés mais peu économiques.
- Algorithme puissant mis au point pour la recherche de ces plans.
- Temps de calcul important mais **tables toutes prêtes à l'emploi** existent sur le site <http://harvest.nps.edu>
- Les colonnes de ces plans sont très **faiblement corrélées**

Stratégie 1: Plans d'exploration de l'espace des entrées

Les LHD quasi-orthogonaux de Cioppa & Lucas (2007)

- Reprise des idées de Ye (1998) sur les LHD orthogonaux à bonnes propriétés mais peu économiques.
- Algorithme puissant mis au point pour la recherche de ces plans.
- Temps de calcul important mais **tables toutes prêtes à l'emploi** existent sur le site <http://harvest.nps.edu>
- Les colonnes de ces plans sont très **faiblement corrélées**
- Non corrélation entre **effets linéaires et quadratiques**

Stratégie 1: Plans d'exploration de l'espace des entrées

Les LHD quasi-orthogonaux de Cioppa & Lucas (2007)

- Reprise des idées de Ye (1998) sur les LHD orthogonaux à bonnes propriétés mais peu économiques.
- Algorithme puissant mis au point pour la recherche de ces plans.
- Temps de calcul important mais **tables toutes prêtes à l'emploi** existent sur le site <http://harvest.nps.edu>
- Les colonnes de ces plans sont très **faiblement corrélées**
- Non corrélation entre **effets linéaires et quadratiques**
- **Bon remplissage** de l'espace (D_{\min} élevée)

Stratégie 1: Plans d'exploration de l'espace des entrées

Les LHD quasi-orthogonaux de Cioppa & Lucas (2007)

- Reprise des idées de Ye (1998) sur les LHD orthogonaux à bonnes propriétés mais peu économiques.
- Algorithme puissant mis au point pour la recherche de ces plans.
- Temps de calcul important mais **tables toutes prêtes à l'emploi** existent sur le site <http://harvest.nps.edu>
- Les colonnes de ces plans sont très **faiblement corrélées**
- Non corrélation entre **effets linéaires et quadratiques**
- **Bon remplissage** de l'espace (D_{\min} élevée)
- Peu de recul encore par rapport à leur mise en pratique, mais recommandés par le praticien du calcul J. Leung (Renault/ Groupe Ingénierie Robuste).

Stratégie 2: Postulation d'un premier métamodèle

Sans recherche de plan optimal

- On dispose d'une **base initiale** conséquente d'exemples de qualité suffisante et la démarche adoptée est **non-séquentielle**.

Stratégie 2: Postulation d'un premier métamodèle

Sans recherche de plan optimal

- On dispose d'une **base initiale** conséquente d'exemples de qualité suffisante et la démarche adoptée est **non-séquentielle**.
- A partir de cette base, on cherche à estimer **le meilleur métamodèle η** parmi plusieurs possibles: **réseaux de neurone, RBF, ondelettes, krigeage, Fourier**, ...après avoir choisi un critère, typiquement parmi:

Stratégie 2: Postulation d'un premier métamodèle

Sans recherche de plan optimal

- On dispose d'une **base initiale** conséquente d'exemples de qualité suffisante et la démarche adoptée est **non-séquentielle**.
- A partir de cette base, on cherche à estimer **le meilleur métamodèle η** parmi plusieurs possibles: **réseaux de neurone, RBF, ondelettes, krigeage, Fourier**, ...après avoir choisi un critère, typiquement parmi:
 - minimisation du maximum ou de la moyenne sur χ de l'erreur quadratique de prédiction,

Stratégie 2: Postulation d'un premier métamodèle

Sans recherche de plan optimal

- On dispose d'une **base initiale** conséquente d'exemples de qualité suffisante et la démarche adoptée est **non-séquentielle**.
- A partir de cette base, on cherche à estimer **le meilleur métamodèle η** parmi plusieurs possibles: **réseaux de neurone, RBF, ondelettes, krigeage, Fourier**, ...après avoir choisi un critère, typiquement parmi:
 - minimisation du maximum ou de la moyenne sur χ de l'erreur quadratique de prédiction,
 - maximisation de l'entropie (\implies maximisation du déterminant de la matrice de corrélation en cas de modèle de Krigeage)

Stratégie 2: Postulation d'un premier métamodèle

Sans recherche de plan optimal

- On dispose d'une **base initiale** conséquente d'exemples de qualité suffisante et la démarche adoptée est **non-séquentielle**.
- A partir de cette base, on cherche à estimer **le meilleur métamodèle η** parmi plusieurs possibles: **réseaux de neurone, RBF, ondelettes, krigeage, Fourier**, ...après avoir choisi un critère, typiquement parmi:
 - minimisation du maximum ou de la moyenne sur χ de l'erreur quadratique de prédiction,
 - maximisation de l'entropie (\implies maximisation du déterminant de la matrice de corrélation en cas de modèle de Krigeage)
- On ne construit pas un plan optimal: on se contente de la base initiale pour calculer ces estimations et les comparer.

Stratégie 2: Postulation d'un premier métamodèle

Avec recherche de plan optimal

- On cherche un seul **plan optimal** par rapport à une forme choisie η avec valeurs a priori (nécessaires) pour les paramètres prises égales aux estimations obtenues avec les exemples de la base initiale.

Stratégie 2: Postulation d'un premier métamodèle

Avec recherche de plan optimal

- On cherche un seul **plan optimal** par rapport à une forme choisie η avec valeurs a priori (nécessaires) pour les paramètres prises égales aux estimations obtenues avec les exemples de la base initiale.
- On peut utiliser des critères de D – ou G –optimalité (Issanchou & Gauchi, 2006, 2007) **MAIS** il faut se souvenir que ces critères sont adaptés si le modèle postulé (η) est une bonne approximation du modèle inconnu f :

$$D - \text{optimalité} : \zeta_{N, N_S^*, \{r^*\}}^D = \underset{\zeta_{N, N_S, \{r\}}}{\text{Arg max}} \left[\det \left(\mathbf{M}_F \left(\zeta_{N, N_S, \{r\}}, \theta^* \right) \right) \right]$$

Stratégie 3: Plans séquentiels optimaux

Principe

- C'est la démarche **la plus riche** en termes de qualité d'information collectée et de propriétés statistiques quand un métamodèle est postulé.

Stratégie 3: Plans séquentiels optimaux

Principe

- C'est la démarche **la plus riche** en termes de qualité d'information collectée et de propriétés statistiques quand un métamodèle est postulé.
- Après avoir estimé un premier métamodèle **on choisit l'expérience suivante à l'endroit dans χ où l'incertitude de prédiction est maximale.**
⇒ ce qui suppose de savoir calculer cette incertitude en un point non encore exploré !

Stratégie 3: Plans séquentiels optimaux

Principe

- C'est la démarche **la plus riche** en termes de qualité d'information collectée et de propriétés statistiques quand un métamodèle est postulé.
- Après avoir estimé un premier métamodèle **on choisit l'expérience suivante à l'endroit dans χ où l'incertitude de prédiction est maximale.**
⇒ ce qui suppose de savoir calculer cette incertitude en un point non encore exploré !
- L. Pronzato a illustré de façon très claire (GdR MASCOT NUM, IFP, Solaize, mars 2007) comment les critères de D - et G -optimalité (locales) explorent très mal χ .

Stratégie 3: Plans séquentiels optimaux

Illustration de l'échec de la démarche usuelle (1)

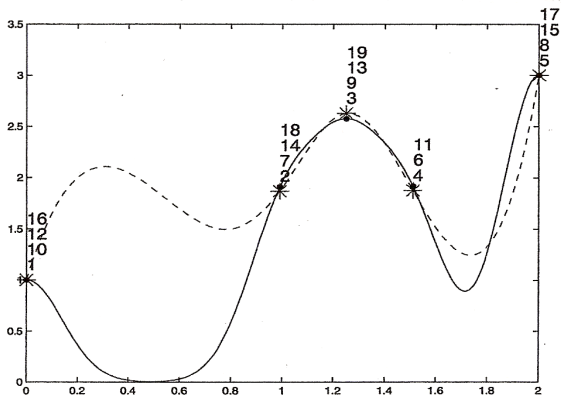
- Exemple proposé par L. Pronzato
- Soit un **modèle RBF** : $\eta(x) = \sum_{j=1}^p \theta_j K(x - N_j)$ avec $K(\cdot)$ noyau gaussien.
- C'est un modèle linéaire à p paramètres si les N_j sont fixés.
- Fixons p à 5 et à partir d'un plan D -optimal en 5 points x_1, \dots, x_5 cherchons séquentiellement les expériences suivantes en des endroits où la variance de prédiction est maximale (après avoir simulé un bruit).

⇒ On observe sur le graphique suivant qu'il n'y a pas d'exploration:
toutes les expériences restent concentrées en les mêmes 5 points D -optimaux du départ.

Stratégie 3: Plans séquentiels optimaux

Illustration de l'échec de la démarche usuelle (2)

Le graphique:



Stratégie 3: Plans séquentiels optimaux

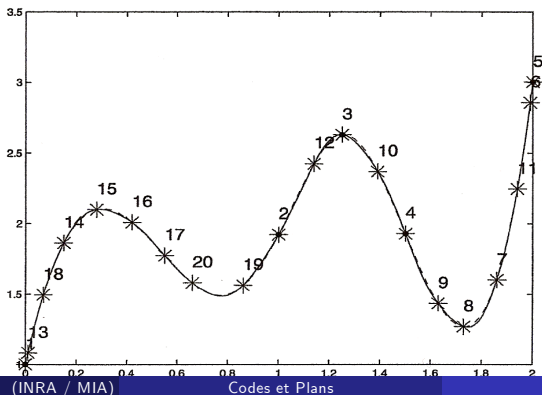
Résultats prometteurs avec la démarche basée sur LDR (1)

- **Méthode LDR** ("Learner Disagreement from experiment Resampling") proposée par Gazut, Martinez, Issanchou (2006).
- Principe:
 - **Apprentissage actif** en procédant à la construction de plusieurs modèles simplifiés (réseaux de neurone) à partir de **répliques bootstrap**.
 - Création d'une variance de prédiction par une approche de type "**Bagging**" ("Bootstrap Aggregating").
 - Les nouvelles expériences sont choisies de façon à **réduire** le maximum de cette variance de prédiction.

Stratégie 3: Plans séquentiels optimaux

Résultats prometteurs avec la démarche basée sur LDR (2)

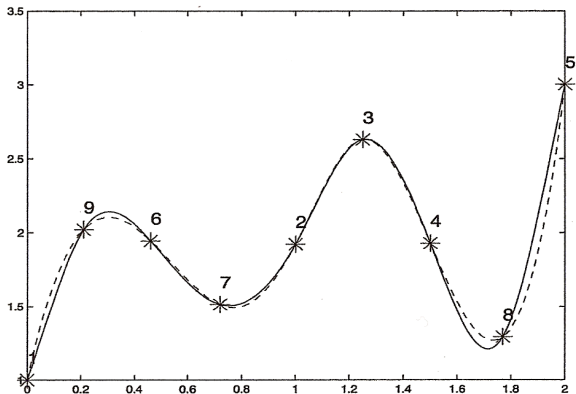
- L. Pronzato montre avec des graphiques l'efficacité d'une démarche séquentielle en choisissant l'expérience suivante par la **méthode LDR** mais avec un modèle RBF (le même que précédemment):



Stratégie 3: Plans séquentiels optimaux

Résultats prometteurs avec la démarche basée sur LDR (3)

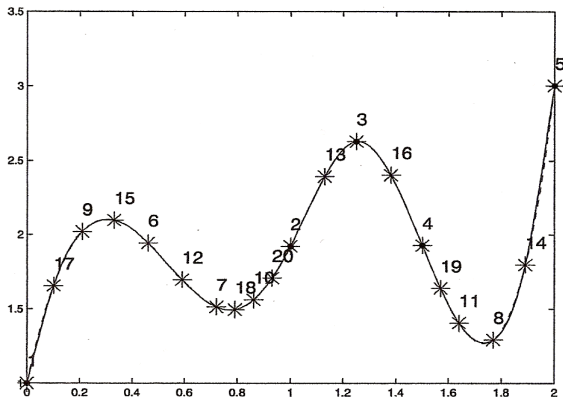
Le graphique avec **un modèle de krigeage à processus gaussien**:
l'expérience suivante est choisie en un endroit où la variance de krigeage est maximale:



Stratégie 3: Plans séquentiels optimaux

Résultats prometteurs avec la démarche basée sur LDR (4)

Le graphique pour le modèle de krigeage avec **20 points**:



Conclusion

- Jusqu'à **20-30 facteurs** on peut trouver des plans d'exploration avec des bonnes propriétés de remplissage et de faible corrélation (les plans uniformes ou les plans de Cioppa & Lucas par exemple).

- Jusqu'à **20-30 facteurs** on peut trouver des plans d'exploration avec des bonnes propriétés de remplissage et de faible corrélation (les plans uniformes ou les plans de Cioppa & Lucas par exemple).
- Pour **100 facteurs ou plus** difficile d'explorer χ
⇒ on se contente de sélectionner un sous-ensemble de sommets du cube: par exemple on peut trouver une fraction de résolution IV de 256 expériences pour 100 facteurs (à 2 niveaux) en moins de 3 secondes avec certains logiciels.

- Jusqu'à **20-30 facteurs** on peut trouver des plans d'exploration avec des bonnes propriétés de remplissage et de faible corrélation (les plans uniformes ou les plans de Cioppa & Lucas par exemple).
- Pour **100 facteurs ou plus** difficile d'explorer χ
⇒ on se contente de sélectionner un sous-ensemble de sommets du cube: par exemple on peut trouver une fraction de résolution IV de 256 expériences pour 100 facteurs (à 2 niveaux) en moins de 3 secondes avec certains logiciels.
- Entre **30 et 100 facteurs** de la recherche à faire pour trouver des plans d'exploration économiques à bonnes propriétés.

- Jusqu'à **20-30 facteurs** on peut trouver des plans d'exploration avec des bonnes propriétés de remplissage et de faible corrélation (les plans uniformes ou les plans de Cioppa & Lucas par exemple).
- Pour **100 facteurs ou plus** difficile d'explorer χ
⇒ on se contente de sélectionner un sous-ensemble de sommets du cube: par exemple on peut trouver une fraction de résolution IV de 256 expériences pour 100 facteurs (à 2 niveaux) en moins de 3 secondes avec certains logiciels.
- Entre **30 et 100 facteurs** de la recherche à faire pour trouver des plans d'exploration économiques à bonnes propriétés.
- Les plans uniformes semblent robustes par rapport aux mauvais choix de modèle (propriété utile en **ingénierie robuste**).

- Jusqu'à **20-30 facteurs** on peut trouver des plans d'exploration avec des bonnes propriétés de remplissage et de faible corrélation (les plans uniformes ou les plans de Cioppa & Lucas par exemple).
- Pour **100 facteurs ou plus** difficile d'explorer χ
⇒ on se contente de sélectionner un sous-ensemble de sommets du cube: par exemple on peut trouver une fraction de résolution IV de 256 expériences pour 100 facteurs (à 2 niveaux) en moins de 3 secondes avec certains logiciels.
- Entre **30 et 100 facteurs** de la recherche à faire pour trouver des plans d'exploration économiques à bonnes propriétés.
- Les plans uniformes semblent robustes par rapport aux mauvais choix de modèle (propriété utile en **ingénierie robuste**).
- L'utilisation des plans d'exploration et autres a été validée sur des **cas concrets industriels** par plusieurs auteurs.

- Beaucoup de problèmes restent ouverts, par exemple: **pourquoi LDR fonctionne ?**

- Beaucoup de problèmes restent ouverts, par exemple: **pourquoi LDR fonctionne ?**
- Des plans d'exploration en **stratégie non-séquentielle** qui soient également performants pour discriminer au mieux deux ou plusieurs métamodèles.

- Beaucoup de problèmes restent ouverts, par exemple: **pourquoi LDR fonctionne ?**
- Des plans d'exploration en **stratégie non-séquentielle** qui soient également performants pour discriminer au mieux deux ou plusieurs métamodèles.
- **Convergence** dans les démarches séquentielles.

- Beaucoup de problèmes restent ouverts, par exemple: **pourquoi LDR fonctionne ?**
- Des plans d'exploration en **stratégie non-séquentielle** qui soient également performants pour discriminer au mieux deux ou plusieurs métamodèles.
- **Convergence** dans les démarches séquentielles.
- Des stratégies (séquentielles) à imaginer pour être efficace **simultanément** en exploration et en optimisation.

- Beaucoup de problèmes restent ouverts, par exemple: **pourquoi LDR fonctionne ?**
- Des plans d'exploration en **stratégie non-séquentielle** qui soient également performants pour discriminer au mieux deux ou plusieurs métamodèles.
- **Convergence** dans les démarches séquentielles.
- Des stratégies (séquentielles) à imaginer pour être efficace **simultanément** en exploration et en optimisation.
- Des benchmarks **industriels** à construire pour comparer ces familles de plans.