

Plan d'expériences numérique sur un domaine non hypercubique

Pierre BARBILLON

MSTGA, 23/11/12

Travaux en collaboration avec Y. Auffray¹ et J.-M. Marin².



-
1. Dassault Aviation
 2. I3M, université Montpellier II

Contexte des expériences numériques/simulées

- Une expérience simulée est une évaluation **déterministe** d'un simulateur : fonction **boîte noire coûteuse** qui représente un modèle physique :

$$f : \mathbf{x} \in E \subset \mathbb{R}^d \mapsto \mathbb{R}.$$

- Incertitudes sur certaines entrées \Rightarrow modélisation par une variable aléatoire \mathbf{X} .

Problématiques applicatives

- prévisualisation et étude du comportement du modèle,
- optimisation,
- analyse de sensibilité,
- étude de la loi des sorties $Y = f(\mathbf{X})$ (moyenne, quantiles, probabilité d'un événement),
- problèmes inverses, reconstitution de la loi de \mathbf{X} conditionnellement à des observations de sorties,

sous la contrainte d'un nombre N d'appels à f limité.

Recours à un émulateur/métamodèle

Émulateur

- approximation de f en tout $\mathbf{x} \in E$,
- évaluation quasi-instantanée,
- (mesure de la confiance à accorder à l'approximation de $f(\mathbf{x})$).

Construction à partir d'un plan d'expérience (numérique)

$$D_n = \{\mathbf{x}_1, \dots, \mathbf{x}_n\} \quad (n \leq N),$$

où f est évaluée.

Problématiques méthodologiques

- Choix et construction de plans d'expérience,
- Choix et construction de métamodèles.

Réf. : [Fang et al. \(2006\)](#), [Koehler et Owen \(1996\)](#), [Santner et al. \(2003\)](#) .

Plan d'expérience numérique

Choisir le plan d'expérience en deux étapes :

- 1 Premier plan d'expérience remplissant l'espace (*space-filling design*).
- 2 Plan complété de manière adaptative pour répondre à un objectif précis : optimisation (Jones et al., 1998), estimation de probabilité de défaillance (Bect et al., 2011)...

Plan

- 1 Émulateur par interpolation à noyaux / Krigeage
 - Construction
 - Choix du plan d'expérience
- 2 Construction d'un plan maximin
 - Domaine Hypercubique
 - Domaine non hypercubique
- 3 Comparaison numérique sur un problème réel

Plan

- 1 Émulateur par interpolation à noyaux / Krigeage
 - Construction
 - Choix du plan d'expérience
- 2 Construction d'un plan maximin
 - Domaine Hypercubique
 - Domaine non hypercubique
- 3 Comparaison numérique sur un problème réel

Plan

- 1 Émulateur par interpolation à noyaux / Krigeage
 - Construction
 - Choix du plan d'expérience
- 2 Construction d'un plan maximin
 - Domaine Hypercubique
 - Domaine non hypercubique
- 3 Comparaison numérique sur un problème réel

A priori sur f

Sacks et al. (1989).

f réalisation d'un processus gaussien F :

$\forall \mathbf{x} \in E$,

$$F(\mathbf{x}) = \sum_{k=1}^Q \beta_k h_k(\mathbf{x}) + \zeta(\mathbf{x}) = H(\mathbf{x})^T \boldsymbol{\beta} + \zeta(\mathbf{x}),$$

où

- h_1, \dots, h_Q fonctions de régression et $\boldsymbol{\beta}$ vecteur paramètres,
- ζ processus gaussien centré de covariance donnée par

$$\text{Cov}(\zeta(\mathbf{x}), \zeta(\mathbf{x}')) = \sigma^2 K(\mathbf{x}, \mathbf{x}'),$$

avec K un noyau de corrélation.

- choix de $K \Rightarrow$ hypothèse sur la régularité de la fonction f .
- Paramètres $\boldsymbol{\beta}$, σ^2 et ceux du noyaux K supposés fixés.

Conditionnement

- $y_1 = f(\mathbf{x}_1), \dots, y_n = f(\mathbf{x}_n)$ évaluations de f aux points du plan d'expérience D_n .
- Processus F^{D_n} : Conditionnement de F à $F(\mathbf{x}_1) = y_1, \dots, F(\mathbf{x}_n) = y_n$.
Processus gaussien de moyenne $m_{D_n}(\mathbf{x})$ et de covariance $C_{D_n}(\mathbf{x}, \mathbf{x}')$ pour tous \mathbf{x}, \mathbf{x}' .

Pour tout $\mathbf{x} \in E$

- $m_{D_n}(\mathbf{x})$ approximation $f(\mathbf{x})$,
- $C_{D_n}(\mathbf{x}, \mathbf{x})$ mesure de la qualité de cette approximation
 - dépend de D_n ,
 - ne dépend pas des évaluations de f .

Pour tout $\mathbf{x}_i \in D_n$,

- $m_{D_n}(\mathbf{x}_i) = f(\mathbf{x}_i)$,
- $C_{D_n}(\mathbf{x}_i, \mathbf{x}_i) = 0$.

Plan

1 Émulateur par interpolation à noyaux / Krigeage

- Construction
- Choix du plan d'expérience

2 Construction d'un plan maximin

- Domaine Hypercubique
- Domaine non hypercubique

3 Comparaison numérique sur un problème réel

Contrôler $C_{D_n}(\mathbf{x}, \mathbf{x})$

Critères de choix de D_n : minimiser (Sacks et al., 1989)

$$\int_E C_{D_n}(\mathbf{x}, \mathbf{x}) \omega(\mathbf{x}) d\mathbf{x} \quad \text{ou} \quad \sup_{\mathbf{x} \in E} C_{D_n}(\mathbf{x}, \mathbf{x}).$$

Critères coûteux à évaluer.

Interpolateurs à noyaux

- K noyau défini positif $\Rightarrow \mathcal{H}_K$ RKHS associé,
- Si $f \in \mathcal{H}_K$, m_{D_n} interpole f sur D_n ($\|m_{D_n}\|_{\mathcal{H}_K}$ de norme minimale),
- $\forall \mathbf{x}_0$,

$$|f(\mathbf{x}_0) - m_{D_n}(\mathbf{x}_0)| \leq \|f\|_{\mathcal{H}_K} C_{D_n}(\mathbf{x}_0, \mathbf{x}_0).$$

D'après [Schaback \(1995\)](#), pour une grande classe de noyaux K ,

$$\sup_{\mathbf{x} \in E} C_{D_n}(\mathbf{x}, \mathbf{x}) \leq G_K(u(D_n)),$$

avec

- $G_K : h \mapsto G_K(h)$ croissante tendant vers 0 pour $h \rightarrow 0$,
- $u(D_n) = \sup_{\mathbf{z} \in E} \min_{1 \leq j \leq n} \|\mathbf{z} - \mathbf{x}_j\|$ avec $D_n = \{\mathbf{x}_1, \dots, \mathbf{x}_n\}$.

Majoration pour un noyau gaussien

Dans le cas d'un noyau gaussien $K_\theta(\mathbf{x}, \mathbf{x}') = e^{-\theta \|\mathbf{x} - \mathbf{x}'\|_2^2}$, la majoration suivante est donnée :

$$P_{D_n}(\mathbf{x}_0) \leq \exp\left(-\frac{\delta}{2u(D_n)^2}\right),$$

où δ dépend du θ du noyau d'où la majoration de l'erreur :

$$\sup_{\mathbf{x}_0} |f(\mathbf{x}_0) - m_{D_n}(\mathbf{x}_0)| \leq \|f\|_{\mathcal{H}_K} \exp\left(-\frac{\delta}{2u(D_n)^2}\right).$$

Du minimax au maximin

Définition (Johnson et al., 1990)

$D_n = \{\mathbf{x}_1, \dots, \mathbf{x}_n\}$ est **minimax** s'il minimise sur E^n :

$$u(D_n) = \sup_{\mathbf{z} \in E} \min_{1 \leq j \leq n} \|\mathbf{z} - \mathbf{x}_j\| .$$

- Plan d'expérience minimax \Rightarrow contrôle de C_{D_n} sur tout E .
- Mais critère **très difficile à calculer** si $\dim(E) \geq 3$.

Du minimax au maximin

Définition (Johnson et al., 1990)

$D_n = \{\mathbf{x}_1, \dots, \mathbf{x}_n\}$ est **maximin** s'il maximise sur E^n :

$$v(D_n) = \min_{1 \leq i, j \leq n} \|\mathbf{x}_i - \mathbf{x}_j\|.$$

D_n de plus petit indice si le nombre de paires de points à distance $v(D_n)$ est minimal.

Proposition

Si D_n maximin d'indice minimal,

$$\sup_{\mathbf{z} \in E} \min_{1 \leq j \leq n} \|\mathbf{z} - \mathbf{x}_j\| = u(D_n) \leq v(D_n) = \min_{1 \leq i, j \leq n} \|\mathbf{x}_i - \mathbf{x}_j\|.$$

Donc pour D_n maximin d'indice minimal,

$$\sup_{\mathbf{x} \in E} C_{D_n}(\mathbf{x}, \mathbf{x}) \leq G_K(u(D_n)) \leq G_K(v(D_n)).$$

Plan

- 1 Émulateur par interpolation à noyaux / Krigeage
 - Construction
 - Choix du plan d'expérience
- 2 Construction d'un plan maximin
 - Domaine Hypercubique
 - Domaine non hypercubique
- 3 Comparaison numérique sur un problème réel

Plan

- 1 Émulateur par interpolation à noyaux / Krigeage
 - Construction
 - Choix du plan d'expérience
- 2 Construction d'un plan maximin
 - **Domaine Hypercubique**
 - Domaine non hypercubique
- 3 Comparaison numérique sur un problème réel

Hypercube latin

$D = \{\mathbf{x}_1, \dots, \mathbf{x}_N\}$ est un hypercube latin s'il est construit ainsi

$$\mathbf{x}_i^j = \frac{\pi_j(i) - U_j^i}{N} \quad \forall 1 \leq i \leq N, \forall 1 \leq j \leq Q \quad (1)$$

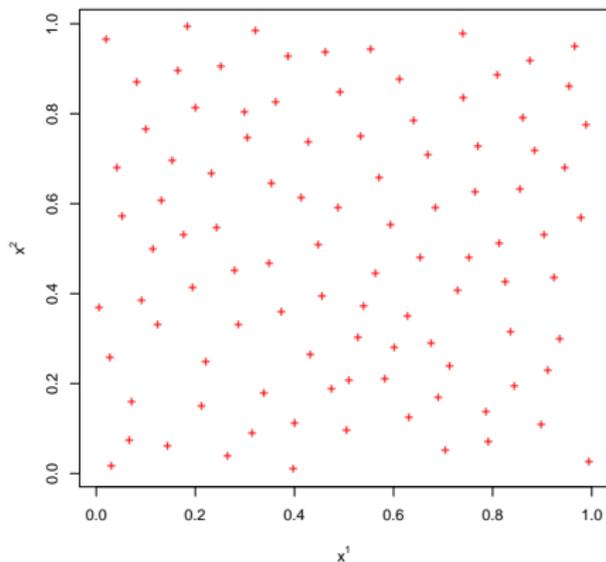
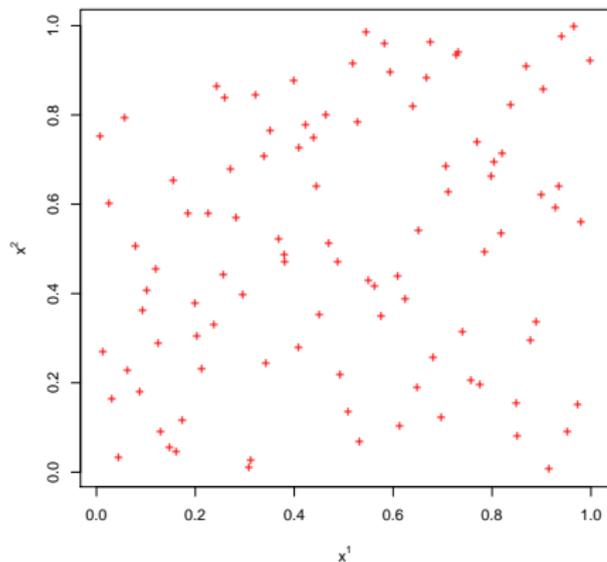
où

- les π_j sont des permutations aléatoires indépendantes des entiers de 1 à N ,
- les U_j^i sont des variables aléatoires indépendantes de loi $\mathcal{U}_{[0,1]}$ et indépendantes des π_j .

Intérêt : bonnes propriétés de projection sur chaque axe.

Domaine hypercubique : hypercube latin maximin

Morris & Mitchell (1995) ont proposé un algorithme de type recuit simulé afin d'obtenir un plan maximin dans la classe des hypercubes latins.



Plan

- 1 Émulateur par interpolation à noyaux / Krigeage
 - Construction
 - Choix du plan d'expérience
- 2 Construction d'un plan maximin
 - Domaine Hypercubique
 - **Domaine non hypercubique**
- 3 Comparaison numérique sur un problème réel

Dans un domaine non hypercubique

Hypothèse

E connu uniquement par une fonction indicatrice.

Algorithme de recuit simulé pour maximiser $D_n \mapsto v(D_n)$:

$$v(D_n = \{\mathbf{x}_1, \dots, \mathbf{x}_n\}) = \min_{1 \leq i, j \leq n} \|\mathbf{x}_i - \mathbf{x}_j\|,$$

dans l'espace des plans d'expérience candidats : $E^n \subset \mathbb{R}^{d \times n}$.

Initialisation de l'algorithme

- Plan d'expérience initial suivant une loi uniforme : $D_n^{(0)} = \{\mathbf{x}_1^{(0)}, \dots, \mathbf{x}_n^{(0)}\}$.
- Choix d'une température inverse initiale $\gamma_0 (= 1/T_0)$ et d'une variance initiale τ_0 .
- Choix d'un schéma de refroidissement inverse $t \mapsto \gamma_t$ et de variance $t \mapsto \tau_t$.

Algorithme : itération t

- 1 Choix de $(\mathbf{x}_i^{(t)}, \mathbf{x}_j^{(t)})$ dans $D_n^{(t)}$ selon une loi multinomiale avec probabilités $\propto \frac{1}{\|\mathbf{x}_i^{(t)} - \mathbf{x}_j^{(t)}\|}$.

- 2 Choix d'un des deux points avec probabilité $(\frac{1}{2})$: $\mathbf{x}_k^{(t)}$.

- 3 Point candidat obtenu par une marche aléatoire **contrainte** :

$$\mathbf{x}_k^{prop} \sim \mathcal{N}_d(\mathbf{x}_k^{(t)}, \tau_t \hat{\Sigma}) \mathbb{I}(\cdot \in E).$$

Plan d'expérience candidat :

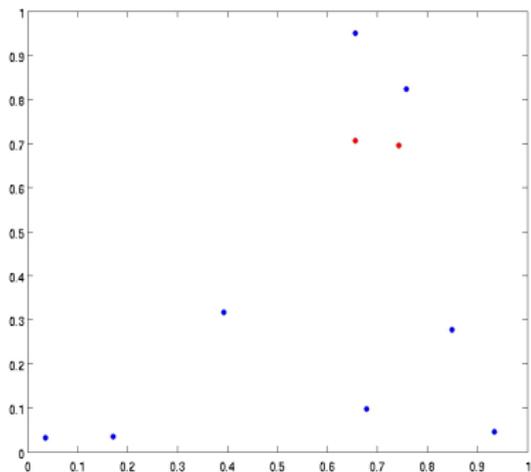
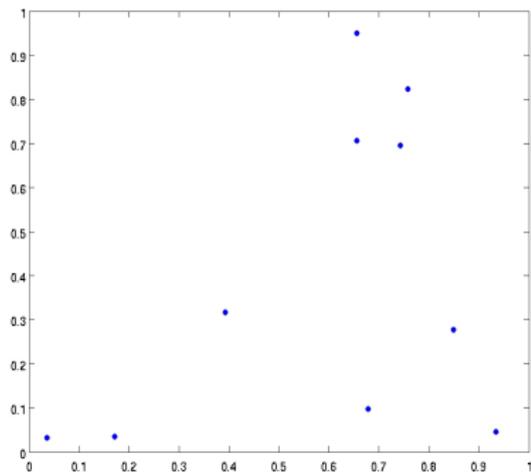
$$D_n^{prop} = \{\mathbf{x}_1^{(t)}, \dots, \mathbf{x}_{k-1}^{(t)}, \mathbf{x}_k^{prop}, \mathbf{x}_{k+1}^{(t)}, \dots, \mathbf{x}_n^{(t)}\}.$$

- 4 $D_n^{(t+1)} = D_n^{prop}$ avec probabilité

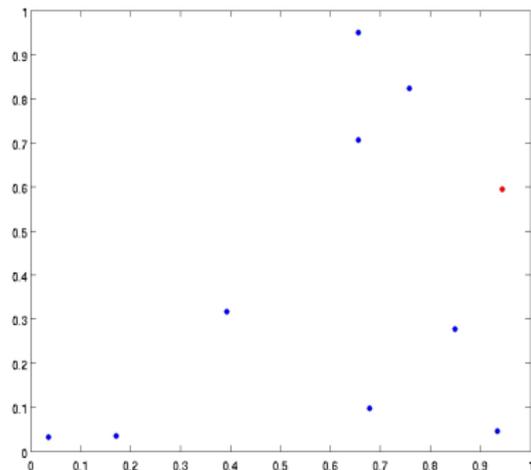
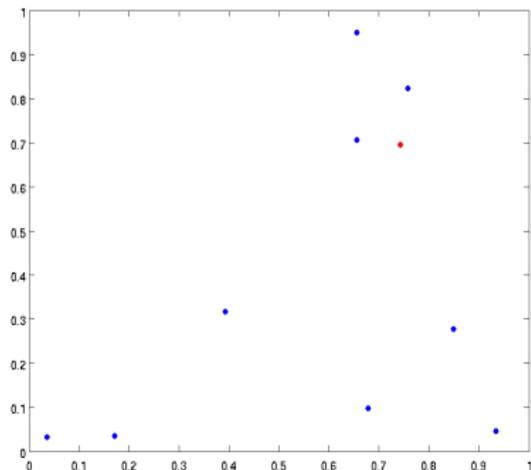
$$\min \left(1, \exp \left(\gamma_t (v(D_n^{prop}) - v(D_n^{(t)})) \right) \right).$$

sinon $D_n^{(t+1)} = D_n^{(t)}$.

Choix d'un couple

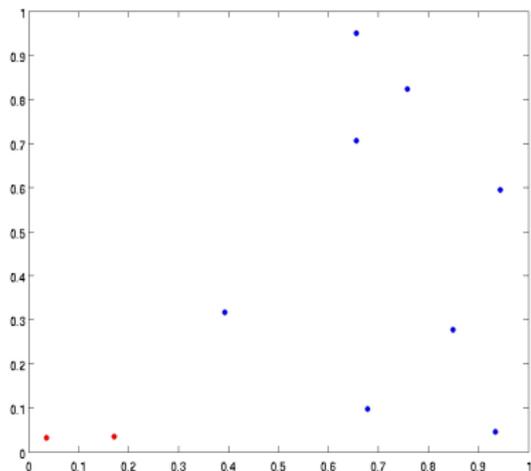
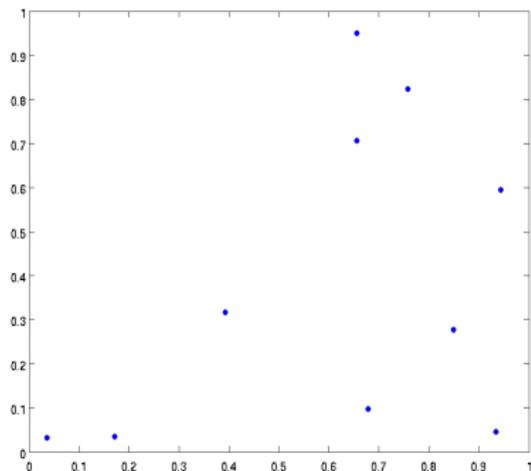


Marche aléatoire

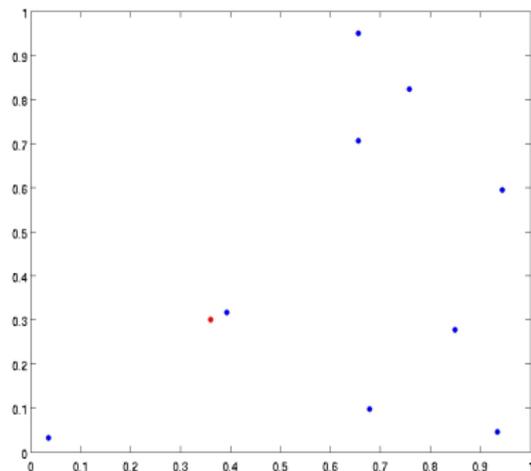
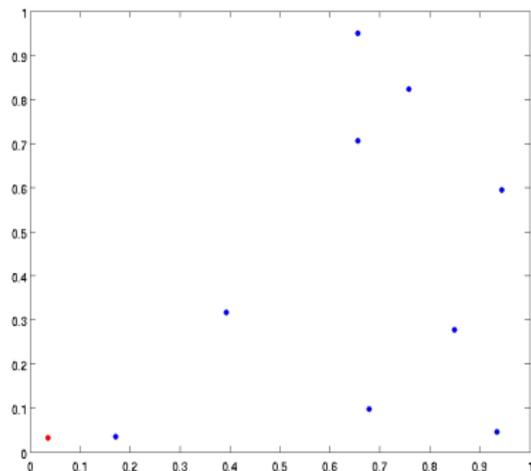


Le plan d'expérience proposé est accepté dans la mesure où la fonction v augmente.

Choix d'un couple



Marche aléatoire



Le plan d'expérience proposé est accepté avec probabilité :

$$\min \left(1, \exp \left(\gamma_t (v(D_n^{prop}) - v(D_n^{(t)})) \right) \right) .$$

Convergence du recuit simulé

Si on remplace l'étape 4 de l'algorithme par :

4bis $D_n^{(t+1)} = D_n^{prop}$ avec probabilité

$$\min \left(1, \exp \left(\gamma_t (v(D_n^{prop}) - v(D_n^{(t)})) \right) \frac{q_{\tau,k}(D_n^{prop}, D_n^{(t)})}{q_{\tau,k}(D_n^{(t)}, D_n^{prop})} \right) .$$

sinon $D_n^{(t+1)} = D_n^{(t)}$.

À température fixée $\gamma_t = \beta$:

- random scan Metropolis within Gibbs,
- chaîne irréductible si $\tau_t > 0$,
- chaîne invariante par rapport à la mesure de Gibbs :

$$\mu_\beta(dD_n) \propto \exp(\beta v(D_n)) \text{Leb}_E(dD_n) .$$

Convergence

En suivant [Bartoli et Del Moral \(2001\)](#),

Théorème

Si $(\tau_t)_{t \geq 0}$ t.q. $\tau_{\min} \leq \tau_t \leq \tau_{\max}$ et

$$\gamma_t = \frac{1}{C} \log(t + e), \quad C > n \cdot \text{osc}(v),$$

alors

$$\forall \epsilon > 0, \quad \lim_{t \rightarrow \infty} \mathbb{P}(D_n^{(t)} \in U^\epsilon) = 1$$

où $U^\epsilon = \{D; v(D) \geq \max_D(v(D)) - \epsilon\}$.

Mais marche aléatoire tronquée \Rightarrow Nécessité de calculer constante de normalisation pour q .

Ce qu'on a tout de même

Pour l'algorithme initial :

Proposition

Sous les mêmes conditions, l'espérance du temps de visite de

$U^\epsilon = \{D; v(D) \geq \max_D(v(D)) - \epsilon\}$ par $(D_n^{(t)})_{t \geq 0}$ est p.s. fini pour tout $\epsilon > 0$.

Heuristique pour l'initialisation et réglages

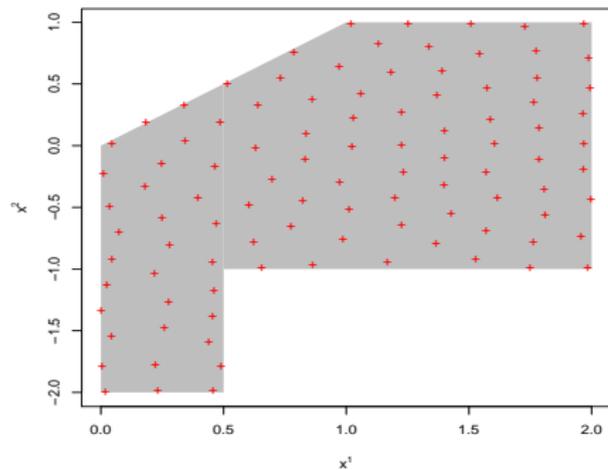
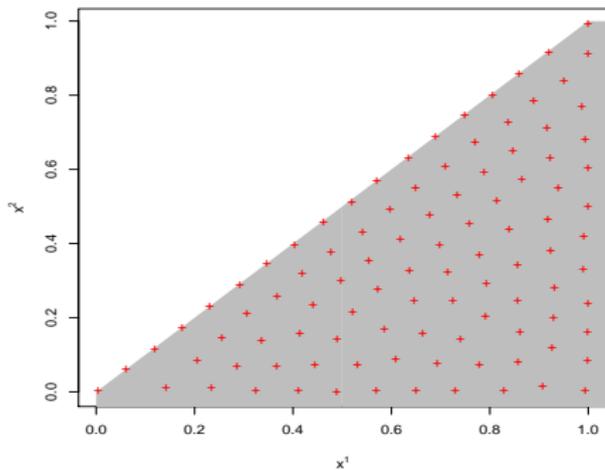
■ Initialisation

- $T_0 = 1/\gamma_0$ comme une fraction de la médiane de $v(D)$ pour D échantillonné aléatoirement de manière uniforme ;
- $\tau_0 = \text{Vol}(E)/n^{1/d}$.

■ Schéma de décroissance

- $\gamma_t = \gamma_0/\sqrt{t}$;
- $\tau_t = \tau_0$ fixe pour les premières itérations puis décroissance en \sqrt{t} .

2 exemples

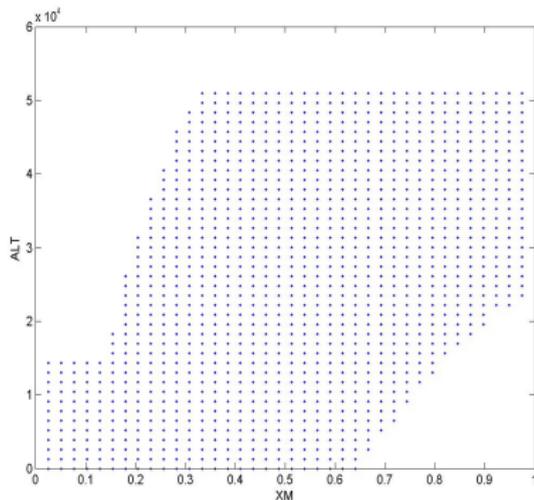


Plan

- 1 Émulateur par interpolation à noyaux / Krigeage
 - Construction
 - Choix du plan d'expérience
- 2 Construction d'un plan maximin
 - Domaine Hypercubique
 - Domaine non hypercubique
- 3 Comparaison numérique sur un problème réel

Émulateur d'un moteur d'avion

- f calcule le flux d'air sortant à partir de conditions de vols dans $E \subset \mathbb{R}^8$.
- Fonction indicatrice de E estimée : taux de mauvais classement 0.5%.



Comparaison

- 3 plans d'expérience de 1300 points chacun :
 - 1 Un plan d'expérience uniforme sur E ,
 - 2 Un plan d'expérience maximin obtenu grâce à 10^7 itérations de l'algorithme.
 - Schéma de refroidissement inverse : $\gamma_t = 1/T_0\sqrt{t}$;
 - schéma de variance de la marche aléatoire : constant pendant un quart des itérations puis $\tau_t = \tau_0/\sqrt{t - 10^7/4}$.
 - 3 Une suite de Sobol' (suite de faible discrédance) contrainte à être dans E .

- Un métamodèle construit par plan d'expérience avec

$$h_{1:Q} = 1, x_1, \dots, x_d, x_1^2, \dots, x_d^2 \text{ et } K(\mathbf{x}, \mathbf{x}') = \exp(-\theta\|\mathbf{x} - \mathbf{x}'\|_1).$$

Performances comparées sur un échantillon test de 1300 points

- Erreur moyenne relative (mER),

$$\frac{1}{1300} \sum_{i=1}^{1300} \left| \frac{f(\mathbf{z}_i) - m_{D_n}(\mathbf{z}_i)}{f(\mathbf{z}_i)} \right|,$$

- Erreur maximale relative (MaxER),

$$\max_{i=1, \dots, 1300} \left| \frac{f(\mathbf{z}_i) - m_{D_n}(\mathbf{z}_i)}{f(\mathbf{z}_i)} \right|,$$

- Erreur quadratique moyenne (EQM),

$$\frac{1}{1300} \sum_{i=1}^{1300} (f(\mathbf{z}_i) - m_{D_n}(\mathbf{z}_i))^2.$$

	mER	MaxER	EQM
Uniforme	0.49%	5.2 %	0.63
Maximin	0.47%	3.5%	0.56
Suite de Sobol'	0.46%	7.7%	0.62

TABLE: Comparaison de la qualité d'approximation sur trois plans d'expériences

Conclusions et perspectives

- Justification du critère de choix d'un plan d'expérience.
- Algorithme de construction convergent et utilisé en application.

Plans adaptatifs pour répondre à des objectifs précis :

- Optimisation,
- Estimation de la probabilité d'un événement rare,
- Problème inverse.

Références

- Auffray, Y., Barbillon, P., Marin, J.-M. (2011). Maximin Design on non-hypercube domain and Kernel Interpolation, *Statistics and Computing*.
- Auffray, Y., Barbillon, P., Marin, J.-M. (2011). Modèles réduits à partir d'expérience numériques, *Journal de la Société Française de Statistique*, 152(1), 89-102.
- Bartoli, N., Del Moral, P. (2001). *Simulation & algorithmes stochastiques*. Cépaduès.
- Bect, J., Ginsbourger, D., Li, L., Picheny V., Vazquez, E. (2011). Sequential design of computer experiments for the estimation of a probability of failure. *Statistics and Computing*.
- Jones, D. R., Schonlau, M., Welch., W. J.. Efficient Global Optimization of Expensive Black-Box Functions. *Journal of Global Optimization*, 13(4) :455–492, 1998
- Schaback, R. (1995). Error estimates and condition numbers for radial basis function interpolation. *Advances in Computational Mathematics*, 3 :251–264.