Planification d'expériences pour l'optimisation de simulateur numérique

Barry Adama[1 2 3]
Doctorant en Mathématiques Appliquées

[1]Institut Mathématique de Toulouse & [2]INRIA Grenoble & [3]IFP Énergies Nouvelles

RJCAF - Décembre 2022

Plan

Outline

Présentation

- Nom & Prénom : Barry Adama
- Sujet de thèse : Plans d'expériences pour la calibration et la validation de simulateur numérique.
- Un des axes de recherche du Consortium Industrie Recherche pour l'Optimisation et la QUantification d'incertitude pour les données Onéreuses (**CIROQU0**) proposé par l'IFPEN.

Contexte

- $f : \mathcal{X} \subseteq \mathbf{R}^d \to \mathbf{R}, d \in \mathbb{N}$
- f coûteuse :
  $\Rightarrow$ nombre d'appel limité
- Forme analytique inconnue et ∇f indisponible
- Exemples :
  1. Conception d'une voilure d'avion (Forester et al. 2008)
  2. Optimisation de la forme des conduits d'un moteur de voiture (Villemonteix et al. 2007)
  3. Etc.

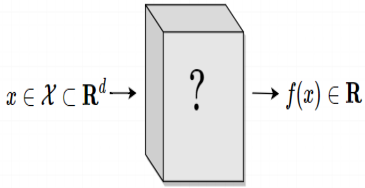


Figure: Simulateur numérique

Objectif

Objectif

Estimer $x^* = argmin_{x \in \mathcal{X}} f(x)$ et $m^* = f(x^*)$

- **Cadre statistique:** approche bayésienne

  f est considérée comme la réalisation d'un processus aléatoire.

Définitions

- **Simulation** : évaluation du simulateur.
- **Expérience** : point de simulation.
- **Planification d'expériences** : choix des expériences à réaliser en fonction des précédents.
- **Méta-modèle** : modèle de substitution du simulateur.

Outline

Processus Gaussiens (1/3)

Processus Gaussien

Un processus gaussien Y est une collection de variables aléatoires telle que la distribution conjointe de chaque sous-ensemble fini est gaussienne multivariée, on note :

$$Y \sim \mathbf{GP}(m, k),$$

où $m(x)$ et $k(x, x^{'})$ représentent respectivement la fonction moyenne et la fonction de covariance.

Fonction moyenne et de covariance

1. La fonction moyenne $m : x \longmapsto m(x)$
2. La fonction de covariance $k : (x, x^{'}) \longmapsto k(x, x^{'})$

Processus Gaussiens (2/3) : Choix de la fonction de covariance

- Un processus gaussien est dit stationnaire si sa fonction moyenne est constante et sa fonction de covariance est telle que :

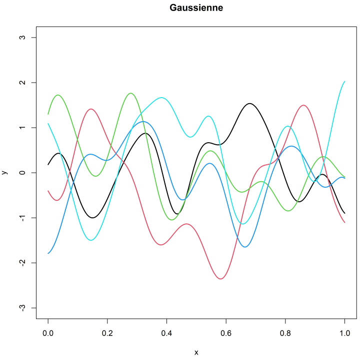
$$\forall x, x^{'} : k(x, x^{'}) = k_{\sigma,l}(\| x - x^{'} \|) = k_{\sigma,l}(d),$$

où σ et l sont les paramètres de la fonction de covariance.

Nom	Expression
Exponentielle	$\sigma^2 e^{-\|d\|/l}$
Gaussienne	$\sigma^2 e^{-d^2/l^2}$
Matern 3/2	$\sigma^2(1 + \sqrt{6}\frac{\|d\|}{l})e^{-\sqrt{6}\|d\|/l}$
Matern 5/2	$\sigma^2(1 + \sqrt{10}\frac{\|d\|}{l} + \sqrt{10}\frac{\|d\|^2}{l^2})e^{-\sqrt{10}\|d\|/l}$

Table: Exemples de fonctions de covariance

Processus Gaussiens (3/3) : exemples de réalisations



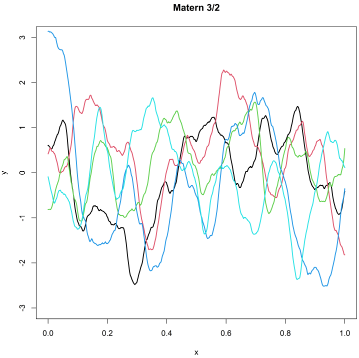


Figure: Fonction de covariance gaussienne avec $\sigma^2 = 1$ et $l = 0.1$.

Figure: Fonction de covariance Matern 3/2 avec $\sigma^2 = 1$ et $l = 0.1$.

Méta-modèle de Processus Gaussien

Méta-modèle

On suppose que f est la réalisation d'un processus Gaussien :

$$Y \sim \mathbf{GP}(m_\beta, \sigma^2 C_\psi), \tag{1}$$

où :

- $m_\beta(x) = h(x)^T \beta$, avec $h(x) = \big(h_1(x), \ldots, h_s(x)\big)^T$ est le vecteur des fonctions de régression connues,
- $\beta \in \mathbf{R}^s$ est le vecteur des paramètres,
- C_ψ : fonction de corrélation et ψ est le vecteur d'hyperparamètres.

Méta-modèle de Processus Gaussien

On dispose de :

- $D_k = \{x_1, \ldots, x_k\}$: le plan d'expériences numériques.
- $f(D_k) = \{f(x_1), \ldots, f(x_k)\}$: les évaluations du simulateur.

A posteriori :

$$Y_k = Y \mid f(D_k) \sim \mathbf{GP}(\mu^k_{\beta,\psi}, V^k_{\psi,\sigma}),$$

où μ^k_β et $V_{\psi,\sigma}$ représentent la fonction moyenne et la covariance à posteriori.

Méta-modèle de Processus Gaussien

- La moyenne a posteriori :

$$\mu^k_{\beta,\psi}(x) = m_\beta(x) + \Sigma_\psi(x, D_k)^T \Sigma_\psi(D_k)^{-1}(y(D_k) - m_\beta(x))$$

- La covariance a posteriori :

$$V^k_{\psi,\sigma}(x, x^{'}) = \sigma^2 \Big[C_\psi(x, x^{'}) - \Sigma_\psi(x, D_k)^T \Sigma_\psi(D_k)^{-1} \Sigma_\psi(x^{'}, D_k)\Big],$$

où :

$$\Sigma_\psi(D_k) = \Big(C_\psi(x_i, x_j)\Big)_{1\leq i,j\leq k} \quad , \Sigma_\psi(x, D_k) \quad = \Big(C_\psi(x, x_i)\Big)_{1\leq i\leq k}$$

et $x, x^{'} \in \mathcal{X}$.

Outline

Algorithme général

Algorithme général

1. Choisir le plan initial de taille n_0: un plan de remplissage de l'espace $\mathcal{X}$
2. Pour n allant de n_0 à n_{budget} :
   - Ajuster un méta-modèle $\hat{f}$ sur $\{(x_i, f(x_i)\}_{i=1,\dots,n}$
   - Utiliser ce méta-modèle pour choisir le prochain point x_{n+1} selon un critère C_n
3. Calculer $\hat{x}^* = \mathrm{argmin}_{x\in\mathcal{X}} \hat{f}(x)$

Critères de choix de x_{n+1}

$$x_{n+1} = argmax_{x \in \mathcal{X}} \quad C_n(x)$$

Quelques critères :

1. Amelioration moyenne (**Excepted Improvement**) :

$$\mathbf{EI}_n(x) = \mathbf{E}\big[(m_n - f(x))\mathbf{1}_{f(x) \leq m_n} \mid f(D_n)\big],$$

   où $m_n = min_{i=1,\ldots,n} f(x_i)$

2. Borne de confiance (**Upper Confidence Bound**) :

$$\mathbf{UCB}_n(x, \lambda) = \mu_n(x) + \lambda \sigma_n(x),$$

   où $\mu_n(x) = \mathbf{E}(f(x) \mid f(D_n))$, $\sigma_n^2 = \mathbf{V}\mathrm{ar}(f(x) \mid f(D_n))$ et $\lambda \in \mathbf{R}$.

3. Probabilité d'amélioration (**Probability of Improvement**):

$$\mathbf{PI}_n(x) = \mathbf{E}\big[\mathbf{1}_{f(x) \leq m_k} \mid f(D_n)\big]$$

Algorithme d'optimisation globale

Les étapes de l'algorithme sont les suivantes :

1. Initialisation :
   - Choisir un plan initial $D_0 = \{x_1, \ldots, x_{n_0}\}$ de taille n_0
   - Évaluer le simulateur en D_0 et calculer $m_0 = \min\{f(x_1), \ldots, f(x_{n_0})\}$
   - Entraîner le méta-modèle sur $\big(D_0, f(D_0)\big)$.
2. Pour $k = n_0 + 1, \ldots, n_{budget}$, faire :
   - Déterminer

$$x_k = argmax_{x \in \mathcal{X}} EI_k(x) = \mathbf{E}\big[(m_k - f(x))\mathbf{1}_{f(x) \leq m_k} \mid f(D_k)\big]$$

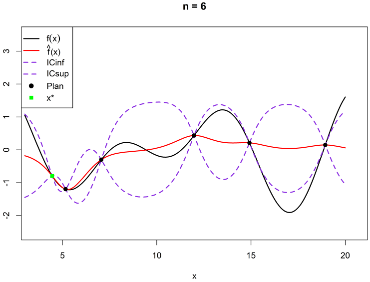
   - Enrichir le plan d'expériences $D_k = D_{k-1} \bigcup \{x_k\}$ et calculer $f(x_k)$
   - Mettre à jour le méta-modèle sur $\big(D_k, f(D_k)\big)$
   - Mettre à jour $m_k = min\{f(x_1), \ldots, f(x_k)\}$
3. Retourner $\hat{x}^* = argmin_{x \in \mathcal{X}} \hat{f}(x)$; $\hat{f}(.)$ est le prédicteur donné par le méta-modèle.

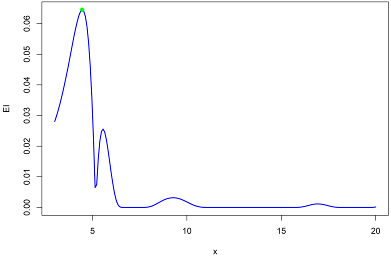
Outline

Application

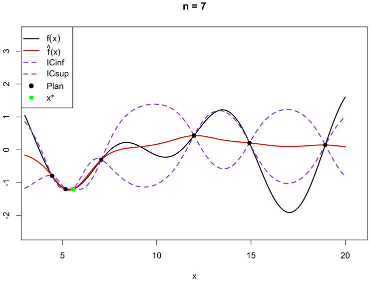
- Fonction de test : $f : x \in [3, 20] \longmapsto \sin(x) + \sin(\frac{2}{3}x) \in \mathbf{R}$
- Taille budget $n_{budget} = 20$ et taille du plan initiale $n_0 = 5$
- Fonction de covariance : *Matern* 5/2
- D_0 : Plan de remplissage Latin Hypercube maximin de taille n_0
- Critère utilisé : **Excepted Improvement**

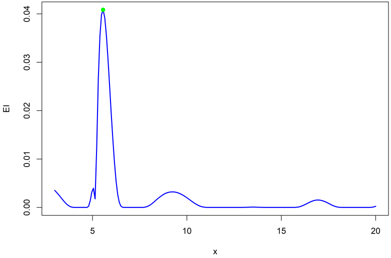
Application : illustration



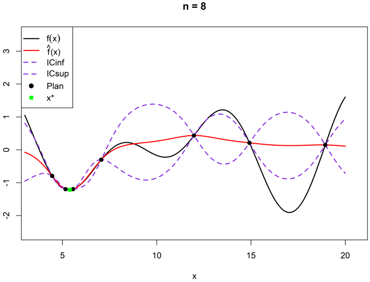


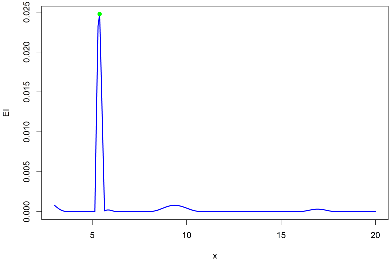
Application : illustration



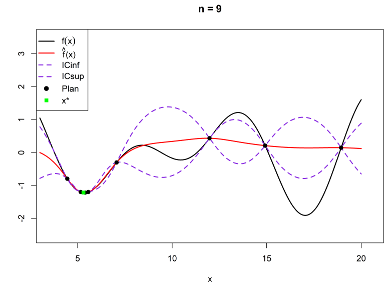


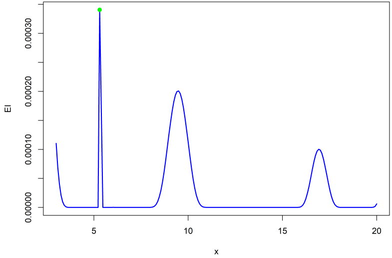
Application : illustration



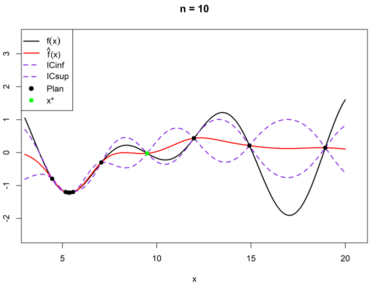


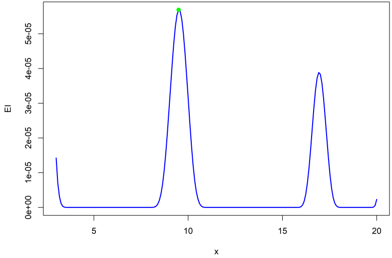
Application : illustration



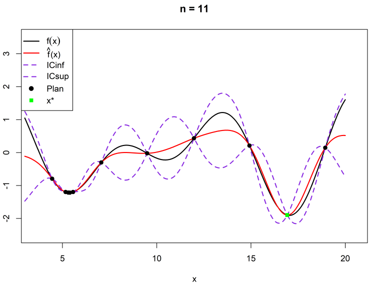


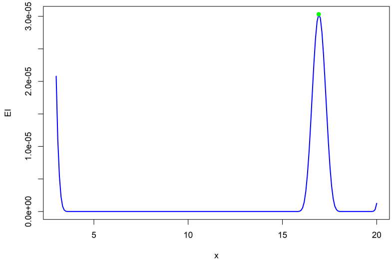
Application : illustration



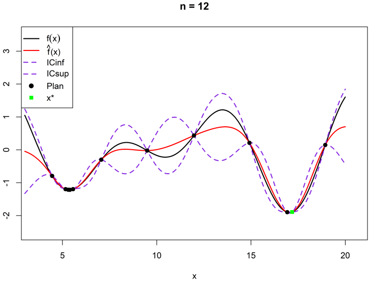


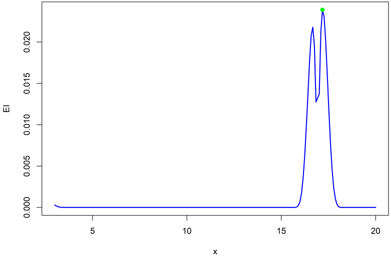
Application : illustration



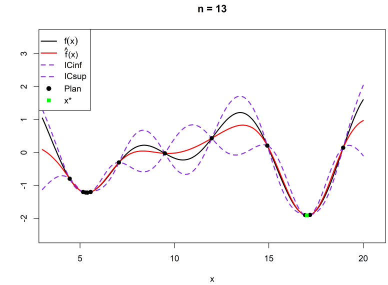


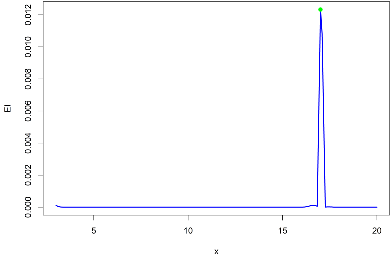
Application : illustration



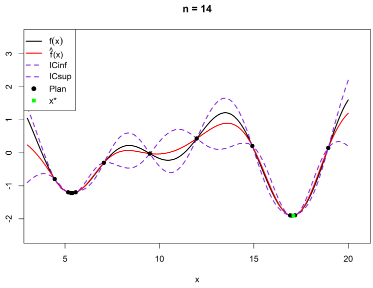


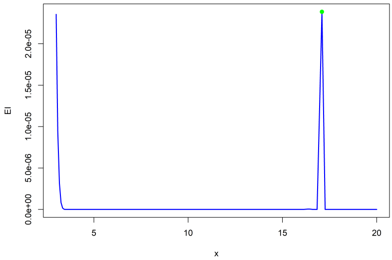
Application : illustration



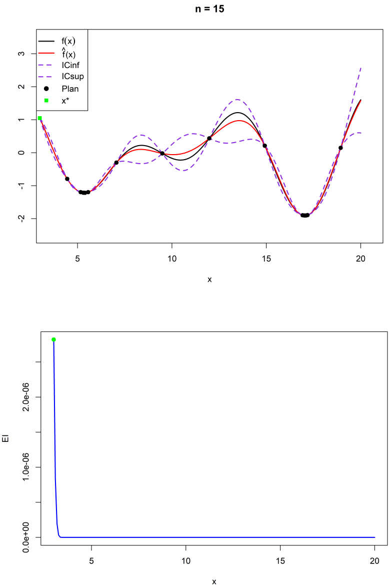


Application : illustration

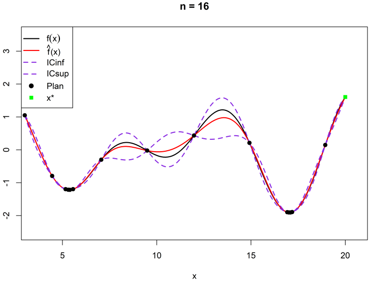


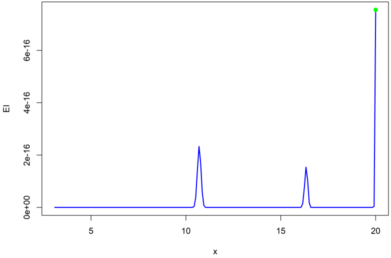


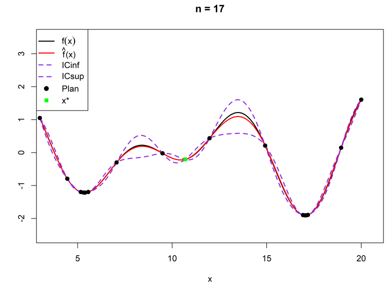
Application : illustration

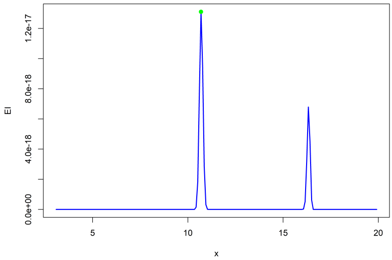


Application : illustration

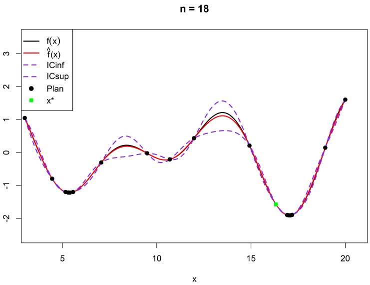


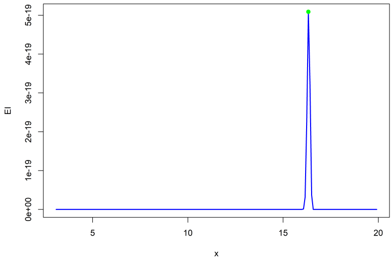


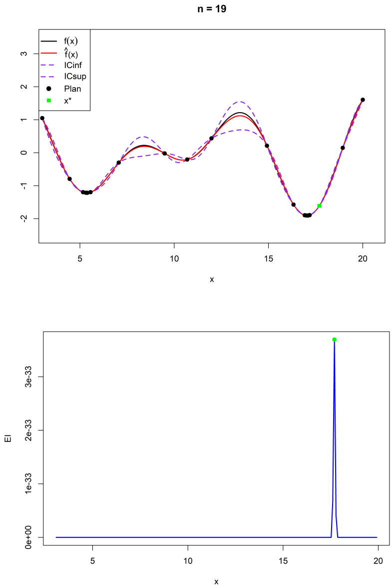

n = 17
f(x)
f̂(x)
ICinf
ICsup
Plan
x*
3
2
1
0
-1
-2
5
10
15
20
x


EI
1.2e-17
8.0e-18
4.0e-18
0.0e+00
5
10
15
20
x

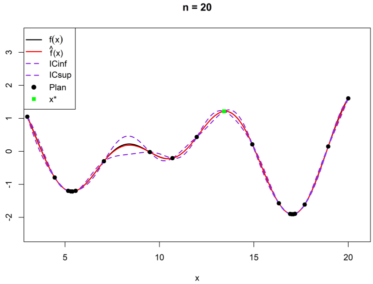
Application : illustration

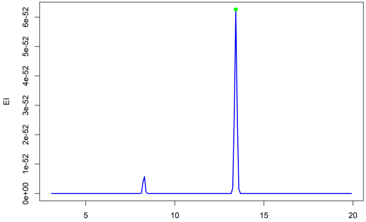





n = 19
f(x)
f̂(x)
ICinf
ICsup
Plan
x*
3
2
1
0
-1
-2
5
10
15
20
x
EI
3e-33
2e-33
1e-33
0e+00
5
10
15
20
x

Finalement : $\hat{x}^* = 17.039 = x^*$ et $\hat{f}(\hat{x^*}) = -1.9059 = f(x^*)$





Outline

Conclusion

Synthèse :

- La modélisation par Processus Gaussien offre un cadre propice à la planification d'expériences.
- La stratégie basée sur le critère d'acquisition **EI** améliore au fur et à mesure la connaissance du minimum en choisissant les expériences qui apporte le plus d'informations pour son estimation.
- Pour chaque objectif particulier des critères adaptés peuvent être définis à l'aide du méta-modèle de Processus Gaussien.

Merci pour votre attention !!!

Références I

Peter I. Frazier, *A tutorial on bayesian optimization*, 2018.

D. Krige, *A statistical approach to some basic mine valuation problems on the witwatersrand, by d.g. krige, published in the journal, december 1951 : introduction by the author*, Journal of The South African Institute of Mining and Metallurgy **52** (1951), 201–203.

Carl Edward Rasmussen and Christopher K. I. Williams, *Gaussian processes for machine learning.*, Adaptive computation and machine learning, MIT Press, 2006.

Jerome Sacks, Susannah Schiller, and William Welch, *Designs for computer experiments*, Technometrics **31** (1989), 41–47.

Références II

Julien Villemonteix, Emmanuel Vazquez, and Eric Walter, *Identification of expensive-to-simulate parametric models using kriging and stepwise uncertainty reduction*, 2007 46th IEEE Conference on Decision and Control, IEEE, 2007, pp. 5505–5510.