Examen multifidélité ModIA

28 mars 2025

```
[ ]: import numpy as np
     from numpy.random import Generator, PCG64
     import math
     from heateq import Exact, Simulateur
```

```
[ ]: seed = 213731490053398181466621250222036675538
     rng = Generator(PCG64(seed))
```

1 Description des simulateurs

On considère l'équation de la chaleur unidimensionnelle vue en cours, et la même distribution de probabilité pour $\mathbf{X}$. On rappelle que la précision de $\tilde{h}^{(K,Q)}$ est gouvernée par les entiers K et Q, et que le coût moyen d'évaluer $\tilde{h}^{(K,Q)}(\mathbf{X})$ est proportionnel à KQ.

On définit un simulateur haute fidélité $f = \tilde{h}^{(K=21,Q=60)}$ et un simulateur basse fidélité $g = \tilde{h}^{(K=3,Q=15)}$.

On note $w = c_g/c_f$, où c_f (respectivement, c_g) est le coût moyen d'évaluer $f(\mathbf{X})$ (respectivement, $g(\mathbf{X})$).

Dans la suite, on note $Y = f(\mathbf{X})$ et $Z = g(\mathbf{X})$.

Question 1 : Quelle est la valeur attendue de w ? Le vérifier expérimentalement.

```
[ ]: # Permet d'obtenir un n-échantillon du vecteur aléatoire d'entrée X
     def n_echantillon_X(n):
         return np.vstack(
             (
                 rng.uniform(-math.pi, math.pi, (3, n)),
                 rng.uniform(0.001, 0.009, (1, n)),
                 rng.uniform(-1., 1., (3, n))
             )
         )
```

```
[ ]: # Simulateurs haute et basse fidélité
     f = Simulateur(21, 60)
     g = Simulateur(3, 15)
```

```
[ ]: # Espérance exacte de la solution continue
     mu_exact = Exact().mu
```

2 Estimateur multifidélité (MF)

On considère un échantillon $\mathcal{X} = \{\mathbf{X}_1, \ldots, \mathbf{X}_N\}$.

On définit un nouvel estimateur multifidélité par $\bar{Y}^{\text{mf}}_{n,N}(\alpha) = \bar{Y}_n - \alpha(\bar{Z}_n - \bar{Z}_N)$, où $0 < n < N$ et

- $\bar{Y}_n = \dfrac{1}{n}\sum_{i=1}^{n} f(\mathbf{X}_i)$,
- $\bar{Z}_n = \dfrac{1}{n}\sum_{i=1}^{n} g(\mathbf{X}_i)$,
- $\bar{Z}_N = \dfrac{1}{N}\sum_{i=1}^{N} g(\mathbf{X}_i)$.

Question 2 : Montrer que le coût moyen d'évaluation de $\bar{Y}^{\text{mf}}_{n,N}(\alpha)$ est $c = n(c_f + \eta c_g)$, où $\eta = \dfrac{N}{n}$. En déduire le nombre d'évaluations haute fidélité équivalentes $\tilde{n}_f = c/c_f$, exprimée en fonction de η et w.

Question 3 : On note $V_{\tilde{n}_f} = \dfrac{\mathbb{V}[Y]}{\tilde{n}_f}$. Que représente $V_{\tilde{n}_f}$?

Question 4 : Montrer que

- $\mathbb{C}[\bar{Y}_n, \bar{Z}_n - \bar{Z}_N] = \dfrac{N-n}{nN}\mathbb{C}[Y, Z]$
- $\mathbb{V}[\bar{Z}_n - \bar{Z}_N] = \dfrac{N-n}{nN}\mathbb{V}[Z]$.

Question 5 : En déduire que la valeur optimale α^* de α qui minimise $\mathbb{V}[\bar{Y}^{\text{mf}}_{n,N}(\alpha)]$, est

$$\alpha^* = \frac{\mathbb{C}[Y, Z]}{\mathbb{V}[Z]}$$

et que la variance minimale est alors

$$\mathbb{V}[\bar{Y}^{\text{mf}}_{n,N}(\alpha^*)] = \frac{\mathbb{V}[Y]}{n}\left[1 - \frac{N-n}{N}\rho^2\right] = V_{\tilde{n}_f}(1 + \eta w)\left[1 - \left(1 - \frac{1}{\eta}\right)\rho^2\right]$$

où ρ est le coefficient de corrélation de Pearson entre Y et Z.

Question 6 : Comparer l'expression de $\mathbb{V}[\bar{Y}^{\text{mf}}_{n,N}(\alpha^*)]$ à celle de $\mathbb{V}[\bar{Y}^{\text{acv}}_{n,N}(\alpha^*)]$ vue en cours. En déduire, **sans faire de calculs** (uniquement à partir du cours), que la valeur optimale η^* de η qui minimise $\mathbb{V}[\bar{Y}^{\text{mf}}_{n,N}(\alpha^*)]$ est

$$\eta^* = \sqrt{\frac{\rho^2}{w(1-\rho^2)}}$$

et que, pour un budget $\tilde{n}_f$ donné, l'estimateur MF optimal est

$$\bar{Y}^{\text{mf}}_{n^*,N^*}(\alpha^*) = \bar{Y}_{n^*} - \alpha^*(\bar{Z}_{n^*} - \bar{Z}_{N^*})$$

où $n^* = \dfrac{\tilde{n}_f}{1 + \eta^* w}$ et $N^* = \eta^* n^*$.

Question 7 : Déduire du cours (toujours sans faire de calculs) que

$$\beta^{\mathrm{mf}} = \frac{\mathbb{V}[\bar{Y}^{\mathrm{mf}}_{n^*,N^*}(\alpha^*)]}{V_{\tilde{n}_f}} = (1+\eta^* w)^2(1-\rho^2).$$

Quelles conclusions peut-on tirer sur les estimateurs MF et ACV optimaux ?

3 Expérience pilote

Question 8 : Estimer α^*, η^* et ρ^2 à l'aide de `p = 10000` échantillons pilotes.

Question 9 : En déduire une estimation de $\beta^{\mathrm{mf}} = (1+\eta^* w)^2(1-\rho^2)$.

```
[ ]: p = 10000 # échantillons pilotes
     X = n_echantillon_X(p)
```

4 Étude numérique de l'estimateur MF

Question 10 : Utiliser la valeur de α^* estimée précédemment (avec les échantillons pilotes) pour contruire un estimateur MF de l'espérance de Y. Faire `nr = 1000` répétitions pour des budgets (en termes de nombre d'évaluations haute fidélité équivalentes) $\tilde{n}_f \in \{5; 10; 20; 50; 100; 200; 500; 1000\}$.

Question 11 : Pour chaque budget, estimer le rapport de variance β^{mf} à budget équivalent entre l'estimateur MF et l'estimateur Monte Carlo classique (haute fidélité). Ces estimations sont-elles conformes à la valeur de β^{mf} estimée précédemment (dans l'expérience pilote) ?

Question 12 : Tracer l'espérance et l'écart-type (sous forme de barres d'erreur) des estimateurs MF et Monte Carlo (haute fidélité) à budget équivalent en fonction de $\tilde{n}_f$. Sur un autre graphe (en échelle log-log), tracer l'évolution de la REQM (par rapport à l'espérance exacte du problème continu, `mu_exact`) des estimateurs en fonction de $\tilde{n}_f$.

```
[ ]: nr = 1000
     budgets = [5, 10, 20, 50, 100, 200, 500, 1000]
```