

Approches multifidélité
pour la propagation d'incertitudes en simulation numérique

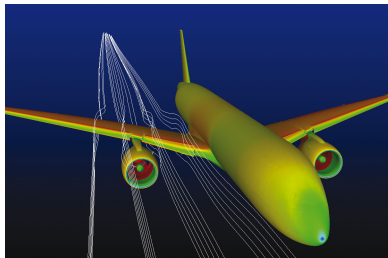
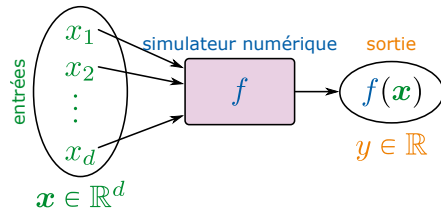
MODIA 2024-25

Paul Mycek
mycek@cerfacs.fr

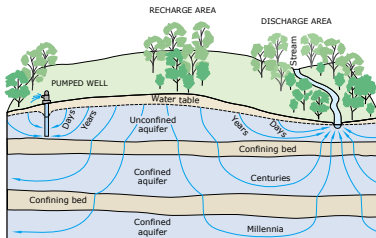


CENTRE EUROPÉEN DE RECHERCHE ET DE FORMATION AVANCÉE EN **CALCUL SCIENTIFIQUE**

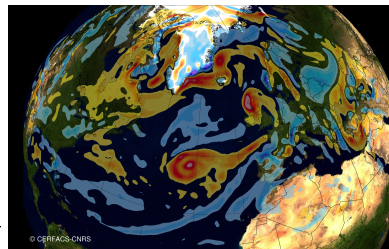
Simulation numérique



(a) f = écoulement autour d'un avion.



(b) f = écoulement souterrain.



(c) f = écoulement atmosphérique.

Exemple : modèle épidémiologique compartimental



Modèle SIR

$$\begin{cases} \frac{dS}{dt}(t) = -\beta I(t)S(t), & S(0) = S_0, \\ \frac{dR}{dt}(t) = \gamma I(t), & R(0) = R_0, \\ I(t) = 1 - S(t) - R(t), & I(0) = 1 - S_0 - R_0. \end{cases}$$

Discrétisation (Euler explicite)

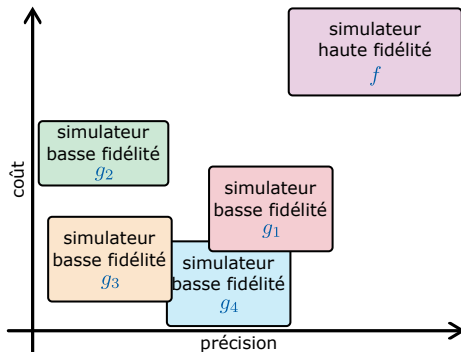
$$\begin{cases} \tilde{S}(t + \delta t) = \tilde{S}(t) - \delta t \beta \tilde{I}(t) \tilde{S}(t), & \tilde{S}(0) = S_0, \\ \tilde{R}(t + \delta t) = \tilde{R}(t) + \delta t \gamma \tilde{I}(t), & \tilde{R}(0) = R_0, \\ \tilde{I}(t + \delta t) = 1 - \tilde{S}(t + \delta t) - \tilde{R}(t + \delta t), & \tilde{I}(0) = 1 - S_0 - R_0. \end{cases}$$

Multifidélité

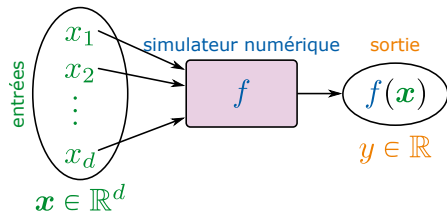
Simulateurs haute et basse fidélité

On suppose que l'on dispose

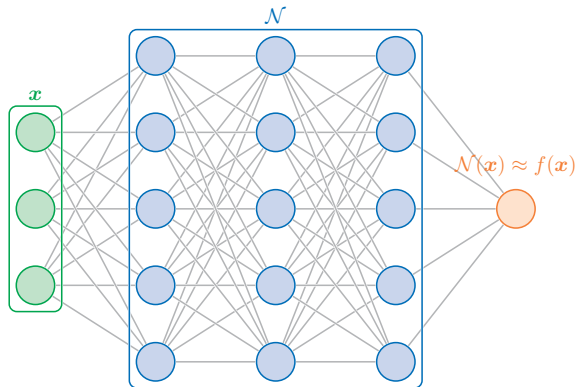
- d'un simulateur haute fidélité $f: \mathbb{R}^d \rightarrow \mathbb{R}$; et
- d'une famille de simulateurs basse fidélité $\{g_\ell: \mathbb{R}^d \rightarrow \mathbb{R}\}_{\ell=1,\dots,m}$.



Exemples de simulateurs/modèles basse fidélités I



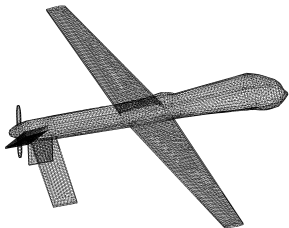
(a) Simulateur haute fidélité.



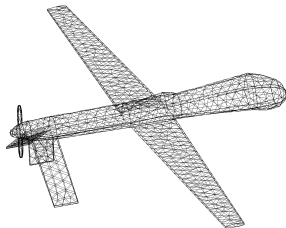
(b) Modèle basse fidélité.

Figure: Basse fidélité définie par un modèle de substitution (ex : réseau de neurones).

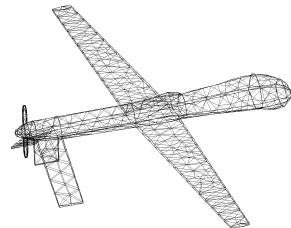
Exemples de simulateurs/modèles basse fidélités II



(a) Maillage fin.



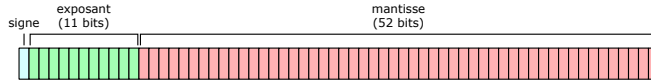
(b) Maillage moyen.



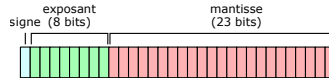
(c) Maillage grossier.

Figure: Différentes fidélités basées sur la discrétisation.

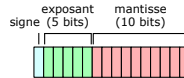
Exemples de simulateurs/modèles basse fidélités III



(a) *Double précision.*



(b) *Simple précision.*



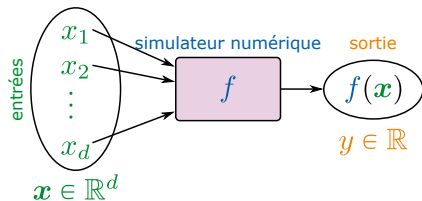
(c) *Demie précision.*

Figure: Différentes fidélités basées sur la précision de l'arithmétique flottante.

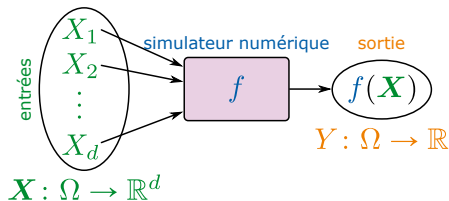
Autres exemples : physique dégradée, précision numérique dégradée, ...

Propagation des incertitudes dans un simulateur numérique: modèle probabiliste

(a) Contexte déterministe



(b) Contexte probabiliste



Modélisation probabiliste des incertitudes

- On considère un espace probabilisé $(\Omega, \mathcal{A}, \mathbb{P})$.
- On *modélise* les paramètres incertains par des v.a. $\mathbf{X} = (X_1, \dots, X_d)$ définies sur $(\Omega, \mathcal{A}, \mathbb{P})$.
- On s'intéresse à la v.a. $Y = f(\mathbf{X})$, qui représente l'incertitude sur la sortie.

Exemple: étant donnée (prescrite) l'incertitude sur \mathbf{X} , estimer $\mathbb{E}[Y]$, $\mathbb{V}[Y]$, $\mathbb{C}[X, Y]$, ...

Estimation statistique (rappels et définitions)

Objectif: estimer un paramètre statistique θ de la loi de la v.a. $Y = f(\mathbf{X})$.

Exemple: $\theta = \mathbb{E}[Y]$, $\theta = \mathbb{V}[Y]$, ...

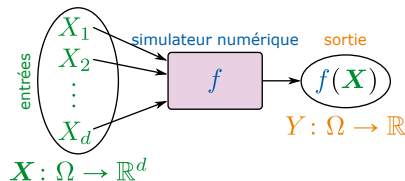
n -échantillon

On appelle n -échantillon de Y un n -uplet (Y_1, \dots, Y_n) de v.a. indépendantes et identiquement distribuées (i.i.d.) de même loi que Y .

Échantillon Monte Carlo de $Y = f(\mathbf{X})$

Soit $(\mathbf{X}_1, \dots, \mathbf{X}_n)$ un n -échantillon de \mathbf{X} .

Alors (Y_1, \dots, Y_n) , où $Y_i = f(\mathbf{X}_i)$, est un n -échantillon de Y .



Estimateur

Soit (Y_1, \dots, Y_n) un n -échantillon de Y et soit $\hat{\vartheta}_n: \mathbb{R}^n \rightarrow \mathbb{R}$.

Alors $\hat{\theta}_n = \hat{\vartheta}_n(Y_1, \dots, Y_n)$ est un *estimateur* de θ .

Estimation

Soit (y_1, \dots, y_n) une *réalisation* de (Y_1, \dots, Y_n) .

Alors $\hat{\vartheta}_n(y_1, \dots, y_n)$ est une *estimation* de θ .

Biais d'un estimateur

$$\text{Biais}(\hat{\theta}_n, \theta) = \mathbb{E}[\hat{\theta}_n - \theta] = \mathbb{E}[\hat{\theta}_n] - \theta.$$

- ▶ Si $\text{Biais}(\hat{\theta}_n, \theta) \neq 0$, on dit que $\hat{\theta}_n$ est un estimateur *biaisé* de θ .
- ▶ Sinon, on dit que c'est un estimateur *non-biaisé* (ou *sans biais*) de θ .

Erreur quadratique moyenne (EQM), et sa racine (REQM)

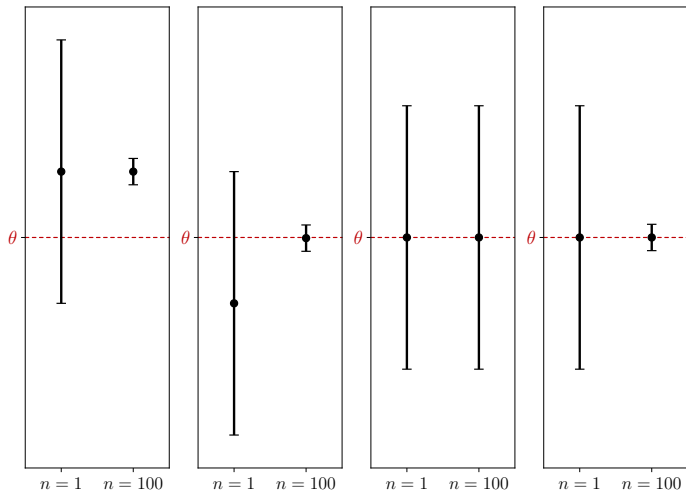
$$\text{EQM}(\hat{\theta}_n, \theta) = \mathbb{E}[(\hat{\theta}_n - \theta)^2], \quad \text{REQM}(\hat{\theta}_n, \theta) = \sqrt{\text{EQM}(\hat{\theta}_n, \theta)}.$$

Décomposition biais-variance

$$\text{EQM}(\hat{\theta}_n, \theta) = \text{Biais}(\hat{\theta}_n, \theta)^2 + \mathbb{V}[\hat{\theta}_n].$$

Discussion

Espérance et écart-type de différents estimateurs $\hat{\theta}_n$ de θ .



Estimation de l'espérance

Soit Y une v.a. réelle, d'espérance μ et de variance σ^2 .

Soit (Y_1, \dots, Y_n) un n -échantillon de Y .

Définition

On définit la *moyenne empirique* de Y par

$$\bar{Y}_n = \hat{\mu}_n(Y_1, \dots, Y_n) = \frac{1}{n} \sum_{i=1}^n Y_i.$$

Propriétés

- $\mathbb{E}[\bar{Y}_n] = \mu$, i.e. \bar{Y}_n est un estimateur **sans biais** de μ ;
- $\mathbb{V}[\bar{Y}_n] = \sigma^2/n$.

Discussion : biais et variance de simulateurs numériques.

Cas d'étude : équation de la chaleur

On considère le problème 1D suivant :

$$\begin{cases} \frac{\partial u(x, t; \mathbf{X})}{\partial t} - \nu(\mathbf{X}) \frac{\partial^2 u(x, t; \mathbf{X})}{\partial x^2} = 0, & x \in (0, 1), \quad t \in [0, T], \\ u(x, 0; \mathbf{X}) = u_0(x; \mathbf{X}) = \mathcal{G}(\mathbf{X})\mathcal{F}_1(x) + \mathcal{I}(\mathbf{X})\mathcal{F}_2(x), \\ u(0, t; \mathbf{X}) = u(1, t; \mathbf{X}) = 0, \end{cases}$$

où $\mathbf{X} = (X_1, \dots, X_7): \Omega \rightarrow \Xi \subset \mathbb{R}^7$, X_1, \dots, X_7 sont des v.a. **mutuellement indépendantes**.

Paramètres

$$\mathcal{F}_1(x) = \sin(\pi x),$$

$$\mathcal{F}_2(x) = \sin(2\pi x) + \sin(3\pi x) + 50[\sin(9\pi x) + \sin(21\pi x)],$$

$$\mathcal{G}(\mathbf{X}) = \frac{7}{2}[\sin(X_1) + 7\sin^2(X_2) + 0.1X_3^4 \sin(X_1)],$$

$$\mathcal{I}(\mathbf{X}) = 50(4|X_5| - 1)(4|X_6| - 1)(4|X_7| - 1),$$

$$\nu(\mathbf{X}) = X_4 > 0.$$

$$X_i \sim \mathcal{U}[-\pi, \pi], \text{ pour } i = 1, 2, 3,$$

$$X_4 \sim \mathcal{U}[0.001, 0.009],$$

$$X_i \sim \mathcal{U}[-1, 1], \text{ pour } i = 5, 6, 7.$$

Solution et quantité d'intérêt (sortie)

La solution s'écrit

$$u(x, t; \mathbf{X}) = \sum_{k=1}^{\infty} a_k(\mathbf{X}) \sin(k\pi x) \exp(-\nu(\mathbf{X})(k\pi)^2 t), \quad a_k(\mathbf{X}) \equiv \int_0^1 u_0(x; \mathbf{X}) \sin(k\pi x) \, dx.$$

On s'intéresse à la sortie scalaire

$$h(\mathbf{X}) = \int_0^1 u(x, t = 0.5; \mathbf{X}) \, dx,$$

dont on souhaite estimer l'espérance, $\mathbb{E}[h(\mathbf{X})] \approx 41.98$.

Discrétisation et définition des fidélités

Une approximation discrète de la sortie est obtenue en tronquant la série aux K premiers termes,

$$\tilde{u}^{(K,Q)}(x, t; \mathbf{X}) = \sum_{k=1}^K \tilde{a}_k^{(Q)}(\mathbf{X}) \sin(k\pi x) \exp(-\nu(\mathbf{X})(k\pi)^2 t),$$

et en approchant numériquement les intégrales par quadrature (ex : méthode des trapèzes) à Q points,

$$a_k(\mathbf{X}) \approx \tilde{a}_k^{(Q)}(\mathbf{X}) = \sum_{q=1}^Q \omega_q u_0(x_q; \mathbf{X}) \sin(k\pi x_q), \quad h(\mathbf{X}) \approx \tilde{h}^{(K,Q)}(\mathbf{X}) = \sum_{q=1}^Q \omega_q \tilde{u}^{(K,Q)}(x_q, t = 0.5; \mathbf{X}).$$

- **Objectif:** estimer $\mathbb{E}[\tilde{h}^{(K,Q)}(\mathbf{X})] \approx \mathbb{E}[h(\mathbf{X})]$.
- **Coût calculatoire moyen** d'évaluer $\tilde{h}^{(K,Q)}$ en \mathbf{X} : $\mathbb{E}[\text{cost}(\tilde{h}^{(K,Q)}(\mathbf{X}))] = \mathcal{O}(KQ)$.

Niveaux de fidélité

Le simulateur **haute fidélité** est défini par $f: \mathbf{X} \mapsto \tilde{h}^{(K=21, Q=100)}(\mathbf{X})$.

Le simulateur **basse fidélité** est défini par $g: \mathbf{X} \mapsto \tilde{h}^{(K=3, Q=20)}(\mathbf{X})$.

Discrétisation et définition des fidélités

Une approximation discrète de la sortie est obtenue en tronquant la série aux K premiers termes,

$$\tilde{u}^{(K,Q)}(x, t; \mathbf{X}) = \sum_{k=1}^K \tilde{a}_k^{(Q)}(\mathbf{X}) \sin(k\pi x) \exp(-\nu(\mathbf{X})(k\pi)^2 t),$$

et en approchant numériquement les intégrales par quadrature (ex : méthode des trapèzes) à Q points,

$$a_k(\mathbf{X}) \approx \tilde{a}_k^{(Q)}(\mathbf{X}) = \sum_{q=1}^Q \omega_q u_0(x_q; \mathbf{X}) \sin(k\pi x_q), \quad h(\mathbf{X}) \approx \tilde{h}^{(K,Q)}(\mathbf{X}) = \sum_{q=1}^Q \omega_q \tilde{u}^{(K,Q)}(x_q, t = 0.5; \mathbf{X}).$$

- **Objectif:** estimer $\mathbb{E}[\tilde{h}^{(K,Q)}(\mathbf{X})] \approx \mathbb{E}[h(\mathbf{X})]$.
- **Coût calculatoire moyen** d'évaluer $\tilde{h}^{(K,Q)}$ en \mathbf{X} : $\mathbb{E}[\text{cost}(\tilde{h}^{(K,Q)}(\mathbf{X}))] = \mathcal{O}(KQ)$.

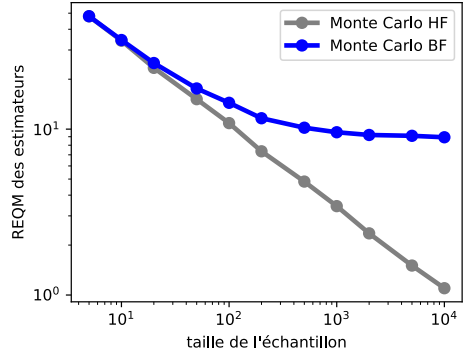
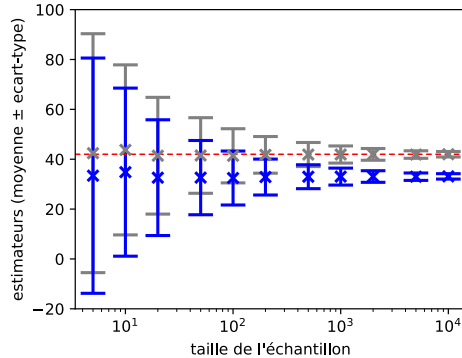
Niveaux de fidélité

Le simulateur **haute fidélité** est défini par $f: \mathbf{X} \mapsto \tilde{h}^{(K=21, Q=100)}(\mathbf{X})$.

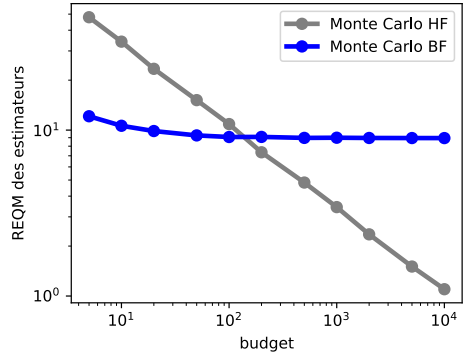
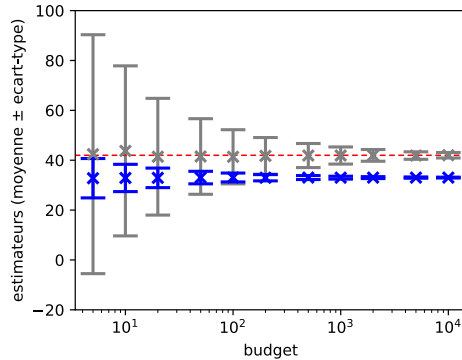
Le simulateur **basse fidélité** est défini par $g: \mathbf{X} \mapsto \tilde{h}^{(K=3, Q=20)}(\mathbf{X})$.

$$\implies w \equiv \mathbb{E}[\text{cost}(g(\mathbf{X}))] / \mathbb{E}[\text{cost}(f(\mathbf{X}))] \approx 1/35.$$

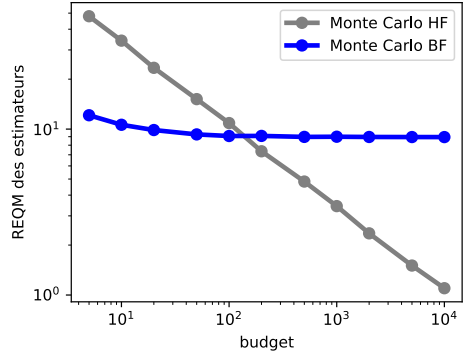
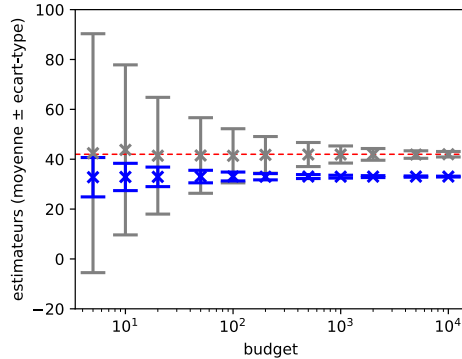
Estimateurs haute et basse fidélité



Estimateurs haute et basse fidélité



Estimateurs haute et basse fidélité



Question : comment combiner les estimateurs ?

Méthode à une variable de contrôle

Soit Y la v.a. dont on veut estimer l'espérance et soit Z une v.a. auxiliaire, appelée *variable de contrôle*, **dont on connaît l'espérance** μ_Z .

On s'intéresse à l'estimateur de la forme

$$\bar{Y}_n^{\text{cv}}(\alpha) = \bar{Y}_n - \alpha(\bar{Z}_n - \mu_Z),$$

et on cherche la valeur de α qui minimise $\mathbb{V}[\bar{Y}_n^{\text{cv}}(\alpha)]$.

Méthode à une variable de contrôle

Soit Y la v.a. dont on veut estimer l'espérance et soit Z une v.a. auxiliaire, appelée *variable de contrôle*, **dont on connaît l'espérance** μ_Z .

On s'intéresse à l'estimateur de la forme

$$\bar{Y}_n^{\text{cv}}(\alpha) = \bar{Y}_n - \alpha(\bar{Z}_n - \mu_Z),$$

et on cherche la valeur de α qui minimise $\mathbb{V}[\bar{Y}_n^{\text{cv}}(\alpha)]$.

Paramètre optimal

La valeur optimale de α est

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

Méthode à une variable de contrôle

Soit Y la v.a. dont on veut estimer l'espérance et soit Z une v.a. auxiliaire, appelée *variable de contrôle*, dont on connaît l'espérance μ_Z .

On s'intéresse à l'estimateur de la forme

$$\bar{Y}_n^{\text{cv}}(\alpha) = \bar{Y}_n - \alpha(\bar{Z}_n - \mu_Z),$$

et on cherche la valeur de α qui minimise $\mathbb{V}[\bar{Y}_n^{\text{cv}}(\alpha)]$.

Paramètre optimal

La valeur optimale de α est

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

Réduction de variance

La variance de l'estimateur optimal $\bar{Y}_n^{\text{cv}}(\alpha^*)$ est

$$\mathbb{V}[\bar{Y}_n^{\text{cv}}(\alpha^*)] = (1 - \rho^2)\mathbb{V}[\bar{Y}_n], \quad \rho^2 = \frac{\mathbb{C}[\bar{Y}_n, \bar{Z}_n]^2}{\mathbb{V}[\bar{Y}_n]\mathbb{V}[\bar{Z}_n]}.$$

Couplage (corrélation) par les paramètres d'entrée

Soit $Y = f(\mathbf{X})$ et $Z = g(\mathbf{X})$. On définit

$$\bar{Y}_n = \frac{1}{n} \sum_{i=1}^n f(\mathbf{X}_i), \quad \bar{Z}_n = \frac{1}{n} \sum_{i=1}^n g(\mathbf{X}_i)$$

à partir du **même échantillon** d'entrée $(\mathbf{X}_1, \dots, \mathbf{X}_n)$.

Alors

$$\mathbb{C}[\bar{Y}_n, \bar{Z}_n] = \frac{1}{n} \mathbb{C}[Y, Z].$$

Paramètre optimal et réduction de variance associée

$$\alpha^* = \frac{\mathbb{C}[\bar{Y}_n, \bar{Z}_n]}{\mathbb{V}[\bar{Z}_n]} = \frac{\mathbb{C}[Y, Z]}{\mathbb{V}[Z]} \quad \text{et} \quad \rho^2 = \frac{\mathbb{C}[\bar{Y}_n, \bar{Z}_n]^2}{\mathbb{V}[\bar{Y}_n] \mathbb{V}[\bar{Z}_n]} = \frac{\mathbb{C}[Y, Z]^2}{\mathbb{V}[Y] \mathbb{V}[Z]}.$$

Méthode à plusieurs variables de contrôle

On dispose de m variables de contrôle Z^1, \dots, Z^m .

On s'intéresse à l'estimateur de la forme

$$\bar{Y}_n^{\text{cv}}(\boldsymbol{\alpha}) = \bar{Y}_n - \sum_{k=1}^m \alpha_k (\bar{Z}_n^k - \mu^k) = \bar{Y}_n - (\bar{\mathbf{Z}}_n - \boldsymbol{\mu}_Z)^T \boldsymbol{\alpha},$$

et on cherche les valeurs des α_k qui minimisent $\mathbb{V}[\bar{Y}_n^{\text{cv}}(\boldsymbol{\alpha})] = \mathbb{V}[\bar{Y}_n] + \boldsymbol{\alpha}^T \mathbb{C}[\bar{\mathbf{Z}}_n, \bar{\mathbf{Z}}_n] \boldsymbol{\alpha} - 2\mathbb{C}[\bar{\mathbf{Z}}_n, \bar{Y}_n]^T \boldsymbol{\alpha}$.

Paramètre optimal

La valeur optimale de $\boldsymbol{\alpha}$ est $\boldsymbol{\alpha}^* = \mathbb{C}[\bar{\mathbf{Z}}_n, \bar{\mathbf{Z}}_n]^{-1} \mathbb{C}[\bar{\mathbf{Z}}_n, \bar{Y}_n] = \mathbb{C}[\mathbf{Z}, \mathbf{Z}]^{-1} \mathbb{C}[\mathbf{Z}, Y]$.

Réduction de variance

$$\begin{aligned} \mathbb{V}[\bar{Y}_n^{\text{cv}}(\boldsymbol{\alpha}^*)] &= (1 - R^2) \mathbb{V}[\bar{Y}_n], \quad R^2 = \frac{\mathbb{C}[\bar{\mathbf{Z}}_n, \bar{Y}_n]^T \mathbb{C}[\bar{\mathbf{Z}}_n, \bar{\mathbf{Z}}_n]^{-1} \mathbb{C}[\bar{\mathbf{Z}}_n, \bar{Y}_n]}{\mathbb{V}[\bar{Y}_n]} \\ &= \frac{\mathbb{C}[\mathbf{Z}, Y]^T \mathbb{C}[\mathbf{Z}, \mathbf{Z}]^{-1} \mathbb{C}[\mathbf{Z}, Y]}{\mathbb{V}[Y]} = \text{Corr}[\mathbf{Z}, Y]^T \text{Corr}[\mathbf{Z}, \mathbf{Z}]^{-1} \text{Corr}[\mathbf{Z}, Y]. \end{aligned}$$

Ajout de variables de contrôle supplémentaires

Lemme — inversion par bloc

Soit \mathbf{M} la matrice définie par bloc par $\mathbf{M} = \begin{bmatrix} \mathbf{A} & \mathbf{B} \\ \mathbf{C} & \mathbf{D} \end{bmatrix}$.

Si \mathbf{A} et \mathbf{M} sont inversibles, alors $\mathbf{S} = \mathbf{D} - \mathbf{CA}^{-1}\mathbf{B}$ est inversible et

$$\mathbf{M}^{-1} = \begin{bmatrix} \mathbf{A}^{-1} + \mathbf{A}^{-1}\mathbf{BS}^{-1}\mathbf{CA}^{-1} & -\mathbf{A}^{-1}\mathbf{BS}^{-1} \\ -\mathbf{S}^{-1}\mathbf{CA}^{-1} & \mathbf{S}^{-1} \end{bmatrix}.$$

Soient $\mathbf{R}_+ = \begin{bmatrix} \mathbf{R} & \mathbf{u} \\ \mathbf{u}^\top & 1 \end{bmatrix}$ et $\mathbf{r}_+ = \begin{bmatrix} \mathbf{r} \\ b \end{bmatrix}$, correspondant à l'ajout d'une variable Z^{m+1} .

Théorème

On suppose \mathbf{R} et \mathbf{R}_+ inversibles. Alors $R_+^2 = R^2 + \frac{(b - \mathbf{u}^\top \mathbf{R}^{-1} \mathbf{r})^2}{1 - \mathbf{u}^\top \mathbf{R}^{-1} \mathbf{u}} \geq R^2$.

De plus, $R_+^2 = R^2 \iff b = \mathbf{u}^\top \mathbf{R}^{-1} \mathbf{r}$.

L'approche des variables de contrôle soulève plusieurs questions essentielles :

1. Quand la méthode est-elle rentable ?
2. Comment calculer α^* ?
3. Que faire lorsque l'espérance des variables de contrôle n'est pas connue ?

Quand la méthode est-elle rentable ?

On considère l'estimateur optimal $\bar{Y}_n^{\text{cv}}(\alpha^*) = \bar{Y}_n - \alpha^*(\bar{Z}_n - \mu_Z)$.

Coût (moyen) de l'estimateur

- ▶ Soit $C_f(\mathbf{X})$ le coût (temps CPU) d'évaluer $f(\mathbf{X})$, et $c_f = \mathbb{E}[C_f(\mathbf{X})]$.
- ▶ Soit $C_g(\mathbf{X})$ le coût (temps CPU) d'évaluer $g(\mathbf{X})$, et $c_g = \mathbb{E}[C_g(\mathbf{X})] < c_f$.
- ▶ On note $w = c_g/c_f > 1$.

Le coût total moyen est $c = n(c_f + c_g)$ et on note $\tilde{n}_f = c/c_f = n(1 + w)$.

\tilde{n}_f représente le nombre d'échantillons *haute fidélité* équivalent à c .

Variance de l'estimateur

$$\mathbb{V}[\bar{Y}_n^{\text{cv}}(\alpha^*)] = \mathbb{V}[\bar{Y}_{\tilde{n}_f}](1 + w)(1 - \rho^2).$$

Il y a réduction de variance si et seulement si $\rho^2 > \frac{w}{1 + w}$.

Approximation de α^*

En pratique, α^* doit être approché par un estimateur $\hat{\alpha}^*$.

Estimateur indépendant à l'aide d'échantillons pilotes

$$\hat{\alpha}^* = \frac{\text{côv}_p(\tilde{Y}_1, \dots, \tilde{Y}_p; \tilde{Z}_1, \dots, \tilde{Z}_p)}{\text{vâr}_p(\tilde{Z}_1, \dots, \tilde{Z}_p)}, \quad \tilde{Y}_i = f(\tilde{\mathbf{X}}_i), \quad \tilde{Z}_i = g(\tilde{\mathbf{X}}_i),$$

où $(\tilde{\mathbf{X}}_1, \dots, \tilde{\mathbf{X}}_p)$ est un p -échantillon **indépendant** de $(\mathbf{X}_1, \dots, \mathbf{X}_n)$.

▷ Dans ce cas, $\bar{Y}_n^{\text{cv}}(\hat{\alpha}^*)$ reste **non-biaisé**.

Estimateur avec les mêmes échantillons que \bar{Y}_n et \bar{Z}_n

$$\hat{\alpha}^* = \frac{\text{côv}_n(Y_1, \dots, Y_n; Z_1, \dots, Z_n)}{\text{vâr}_n(Z_1, \dots, Z_n)}.$$

▷ Dans ce cas, $\bar{Y}_n^{\text{cv}}(\hat{\alpha}^*)$ est **biaisé**.

Dans tous les cas, cela induit une détérioration de la réduction de variance.