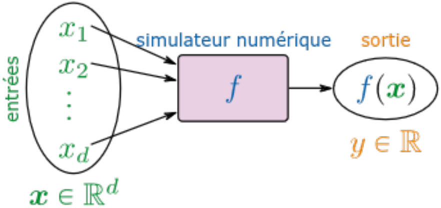
Approches multifidélité
pour la propagation d'incertitudes en simulation numérique

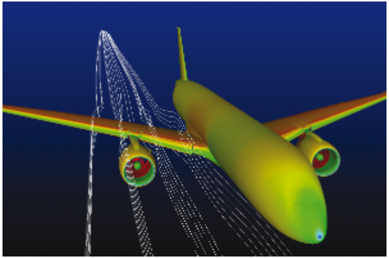
MODIA 2023-24

Paul Mycek
mycek@cerfacs.fr

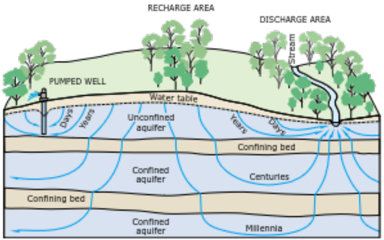
Cerfacs (Toulouse)

Simulation numérique

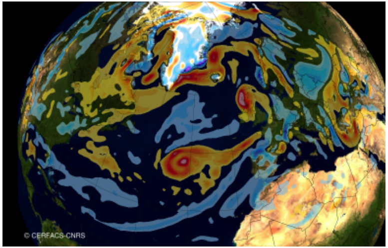




(a) f = écoulement autour d'un avion.



(b) f = écoulement souterrain.



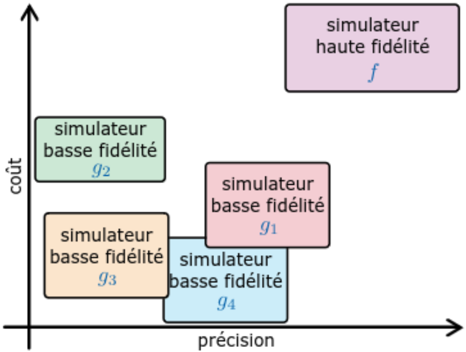
(c) f = écoulement atmosphérique.

Multifidélité

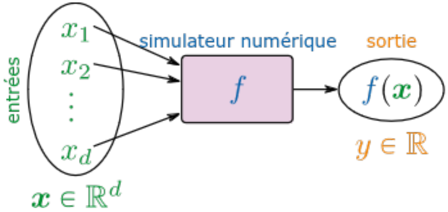
Simulateurs haute et basse fidélité

On suppose que l'on dispose

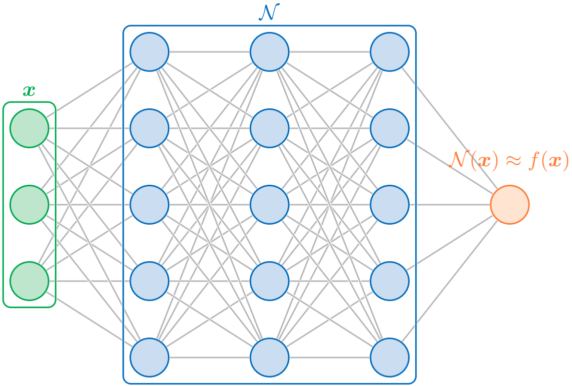
- d'un simulateur haute fidélité $f\colon \mathbb{R}^d \to \mathbb{R}$; et
- d'une famille de simulateurs basse fidélité $\{g_\ell\colon \mathbb{R}^d \to \mathbb{R}\}_{\ell=1,\dots,m}$.



Exemples de simulateurs/modèles basse fidélités I



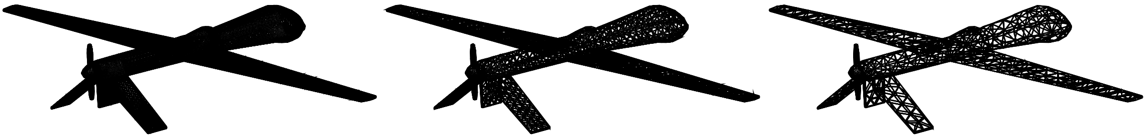
(a) Simulateur haute fidélité.



(b) Modèle basse fidélité.

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

Exemples de simulateurs/modèles basse fidélités II



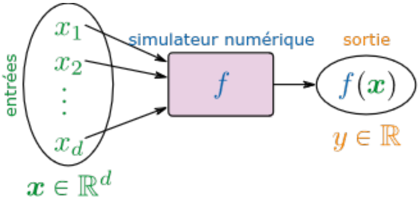
(a) Maillage fin. (b) Maillage moyen. (c) Maillage grossier.

Figure: Différentes fidélités basées sur la discrétisation.

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

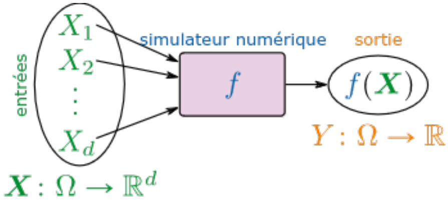
Propagation des incertitudes dans un simulateur numérique: modèle probabiliste

(a) Contexte déterministe



$\rightsquigarrow$

(b) Contexte probabiliste



Modélisation probabiliste des incertitudes

- On modélise les paramètres incertains par des v.a. $\boldsymbol{X} = (X_1, \ldots, X_d)$ que l'on sait échantillonner.
- On s'intéresse à la v.a. $Y = f(\boldsymbol{X})$, qui représente l'incertitude sur la sortie.

Exemple: étant donnée (prescrite) l'incertitude sur $\boldsymbol{X}$, 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 d'une v.a. Y, définie sur un espace probabilisé $(\Omega, \mathcal{A}, \mathbb{P})$.

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

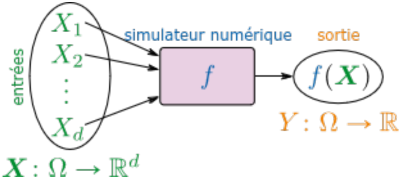
n-échantillon

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 $Y = f(\boldsymbol{X})$

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

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



Estimateur

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

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

Qualité d'un estimateur

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)

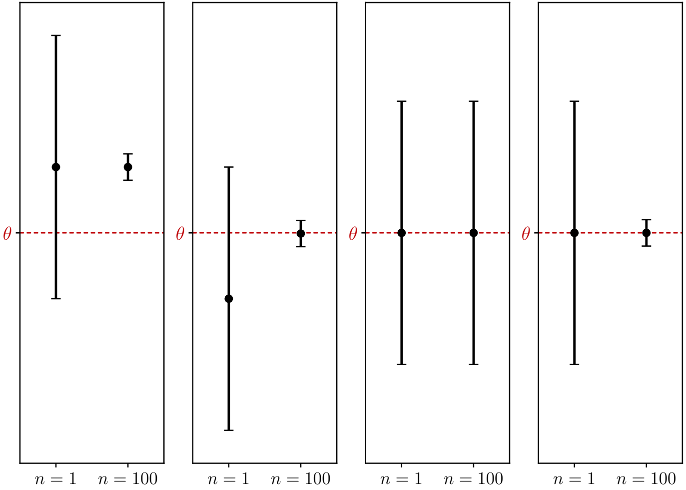
$$\text{EQM}(\hat{\theta}_n, \theta) = \mathbb{E}[(\hat{\theta}_n - \theta)^2], \qquad \text{REQM}(\hat{\theta}_n, \theta) = \sqrt{\text{EQM}(\hat{\theta}_n, \theta)}.$$

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 de θ.



Estimation de l'espérance

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

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

Définition

On définit la *moyenne empirique* de Y par

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

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 :

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

où $\boldsymbol{X} = (X_1, \ldots, X_7)\colon \Omega \to \Xi \subset \mathbb{R}^7$, $X_1, \ldots, X_7$ sont des v.a. **mutuellement indépendantes**.

Paramètres

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

$\mathcal{F}_2(x) = \sin(2\pi x) + \sin(3\pi x) + 50[\sin(9\pi x) + \sin(21\pi x)],$ $\qquad X_i \sim \mathcal{U}[-\pi, \pi]$, pour $i = 1,2,3,$

$\mathcal{G}(\boldsymbol{X}) = \dfrac{7}{2}[\sin(X_1) + 7\sin^2(X_2) + 0.1X_3^4\sin(X_1)],$ $\qquad X_4 \sim \mathcal{U}[0.001, 0.009],$

$\mathcal{I}(\boldsymbol{X}) = 50(4|X_5| - 1)(4|X_6| - 1)(4|X_7| - 1),$ $\qquad X_i \sim \mathcal{U}[-1, 1]$, pour $i = 5,6,7.$

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

Solution et quantité d'intérêt (sortie)

La solution s'écrit

$$u(x,t;\boldsymbol{X}) = \sum_{k=1}^{\infty} a_k(\boldsymbol{X}) \sin(k\pi x) \exp\left(-\nu(\boldsymbol{X})(k\pi)^2 t\right), \qquad a_k(\boldsymbol{X}) \equiv \int_0^1 u_0(x;\boldsymbol{X}) \sin(k\pi x)\,\mathsf{d}x.$$

On s'intéresse à la sortie scalaire

$$h(\boldsymbol{X}) = \int_0^1 u(x, t=0.5; \boldsymbol{X})\,\mathsf{d}x,$$

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

Discrétisation et définition des fidélités

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

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

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

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

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

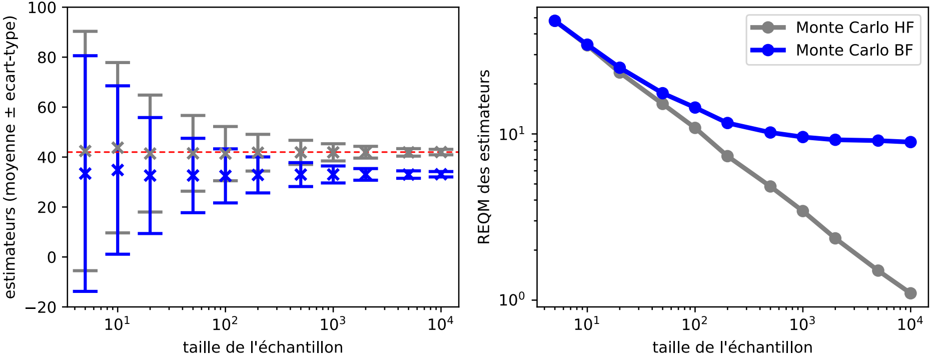
Niveaux de fidélité

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

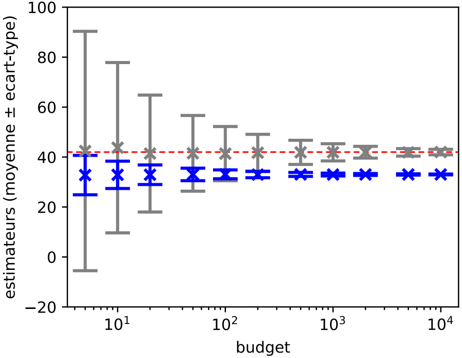
Le simulateur **basse fidélité** est défini par $g\colon \boldsymbol{X} \mapsto \tilde{h}^{(K=3,Q=20)}(\boldsymbol{X})$.

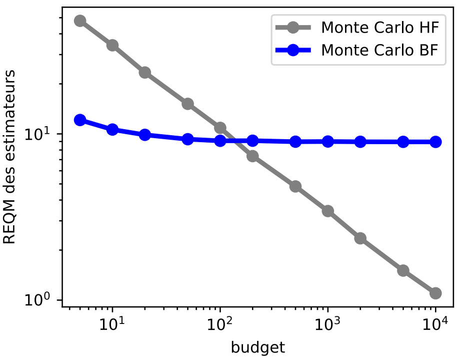
$$\implies w \equiv \mathbb{E}[\mathrm{cost}(g(\boldsymbol{X}))]/\mathbb{E}[\mathrm{cost}(f(\boldsymbol{X}))] \approx 1/35.$$

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

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

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

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 $\bar{Y}_n^{\mathsf{cv}}(\alpha^*)$ est

$$\mathbb{V}[\bar{Y}_n^{\mathsf{cv}}(\alpha^*)] = (1 - \rho^2)\mathbb{V}[\bar{Y}_n].$$

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

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

et on cherche les valeurs des α_k qui minimisent $\mathbb{V}[\bar{Y}_n^{\mathsf{cv}}(\boldsymbol{\alpha})]$,

$$\mathbb{V}[\bar{Y}_n^{\mathsf{cv}}(\boldsymbol{\alpha})] = \mathbb{V}[\bar{Y}_n] + \boldsymbol{\alpha}^{\mathsf{T}} \mathbb{C}[\bar{\boldsymbol{Z}}_n, \bar{\boldsymbol{Z}}_n] \boldsymbol{\alpha} - 2\mathbb{C}[\bar{\boldsymbol{Z}}_n, \bar{Y}_n]^{\mathsf{T}} \boldsymbol{\alpha}.$$

Paramètre optimal

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

Réduction de variance

$$\mathbb{V}[\bar{Y}_n^{\mathsf{cv}}(\boldsymbol{\alpha}^*)] = (1 - R^2)\mathbb{V}[\bar{Y}_n], \qquad R^2 = \frac{\mathbb{C}[\bar{\boldsymbol{Z}}_n, \bar{Y}_n]^{\mathsf{T}} \mathbb{C}[\bar{\boldsymbol{Z}}_n, \bar{\boldsymbol{Z}}_n]^{-1} \mathbb{C}[\bar{\boldsymbol{Z}}_n, \bar{Y}_n]}{\mathbb{V}[\bar{Y}_n]}.$$

Couplage (corrélation) par les paramètres d'entrée

Si $Y = f(\boldsymbol{X})$ et $Z^k = g_k(\boldsymbol{X})$, alors

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

Paramètre optimal et réduction de variance associée

$$\boldsymbol{\alpha}^* = \mathbb{C}[\boldsymbol{Z}, \boldsymbol{Z}]^{-1}\mathbb{C}[\boldsymbol{Z}, Y], \quad \text{et } R^2 = \frac{\mathbb{C}[\boldsymbol{Z}, Y]^\mathsf{T}\mathbb{C}[\boldsymbol{Z}, \boldsymbol{Z}]^{-1}\mathbb{C}[\boldsymbol{Z}, Y]}{\mathbb{V}[Y]} = \mathrm{Corr}[\boldsymbol{Z}, Y]^\mathsf{T}\,\mathrm{Corr}[\boldsymbol{Z}, \boldsymbol{Z}]^{-1}\,\mathrm{Corr}[\boldsymbol{Z}, Y].$$

Si $m = 1$ (une seule variable de contrôle), on a $\alpha^* = \dfrac{\mathbb{C}[Y, Z]}{\mathbb{V}[Z]}$ et $R^2 = \dfrac{\mathbb{C}[Y, Z]^2}{\mathbb{V}[Y]\mathbb{V}[Z]} = \mathrm{Corr}[Y, Z]^2 = \rho^2$.

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 $\mathbf{S} = \mathbf{D} - \mathbf{C}\mathbf{A}^{-1}\mathbf{B}$ est inversible et

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

Soient $\mathbf{R}_+ = \begin{bmatrix} \mathbf{R} & \boldsymbol{u} \\ \boldsymbol{u}^\mathsf{T} & 1 \end{bmatrix}$ et $\boldsymbol{r}_+ = \begin{bmatrix} \boldsymbol{r} \\ b \end{bmatrix}$, correspondant à l'ajout d'une variable Z^{m+1}.

Théorème

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

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

Application dans un contexte de simulation numérique

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 $C_f(\boldsymbol{X})$ le coût (temps CPU) d'évaluer $f(\boldsymbol{X})$, et $c_f = \mathbb{E}[C_f(\boldsymbol{X})]$.
- Soit $C_g(\boldsymbol{X})$ le coût (temps CPU) d'évaluer $g(\boldsymbol{X})$, et $c_g = \mathbb{E}[C_g(\boldsymbol{X})] < c_f$.
- 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 $\rho^2 > \dfrac{w}{1+w}$.

Approximation de α^*

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

Estimateur indépendant à l'aide d'échantillons pilotes

$$\hat{\alpha}^* = \frac{\hat{\mathrm{cov}}_p(\tilde{Y}_1, \ldots, \tilde{Y}_p; \tilde{Z}_1, \ldots, \tilde{Z}_p)}{\hat{\mathrm{var}}_p(\tilde{Z}_1, \ldots, \tilde{Z}_p)}, \qquad \tilde{Y}_i = f(\tilde{\boldsymbol{X}}_i), \quad \tilde{Z}_i = g(\tilde{\boldsymbol{X}}_i),$$

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

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

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

$$\hat{\alpha}^* = \frac{\hat{\mathrm{cov}}_n(Y_1, \ldots, Y_n; Z_1, \ldots, Z_n)}{\hat{\mathrm{var}}_n(Z_1, \ldots, Z_n)}.$$

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

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

Espérance des variable de contrôle inconnue

Si μ_Z n'est pas connue, on peut l'estimer avec $N > n$ échantillons.

Estimateur par variable de contrôle approchée

On approche $\bar{Y}_n^{\mathsf{cv}}(\alpha)$ par $\bar{Y}_{n,N}^{\mathsf{acv}}(\alpha) = \bar{Y}_n - \alpha(\bar{Z}_n - \bar{Z}_N)$.

La valeur optimale de α est

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

Nombre d'échantillons haute fidélité équivalent

On note $\eta = N/n > 1$. On a $c = n(c_f + \eta c_g)$ et $\tilde{n}_f = c/c_f = n(1 + \eta w)$.

Variance de l'estimateur

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

Quand la méthode approchée est-elle rentable ?

Ratio η optimal

La valeur η^* qui minimise la variance est $\eta^* = \sqrt{\dfrac{\rho^2}{w(1-\rho^2)}}$.

De plus,

$$\eta^* > 1 \iff \rho^2 > \frac{w}{1+w}.$$

Variance de l'estimateur

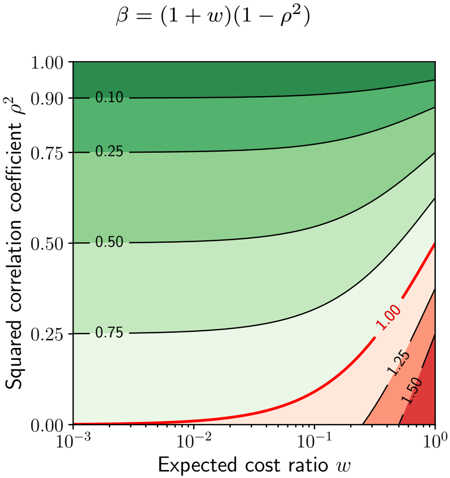
On a

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

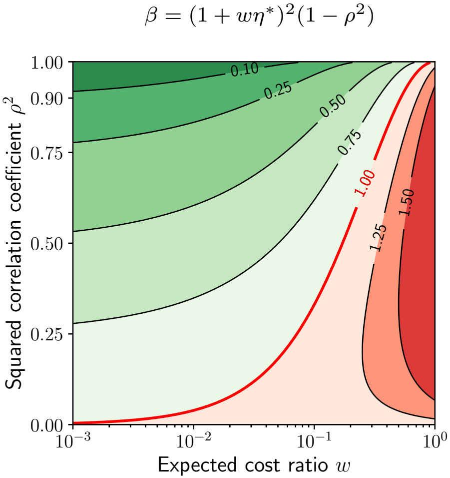
Il y a réduction de variance si et seulement si $\rho^2 > \dfrac{4w}{(1+w)^2}$.

Dans ce cas, on a bien $\eta^* > 1$.

Régions de réduction de variance ($\beta < 1$)



(a) Espérance connue (CV)



(b) Espérance inconnue (ACV)