

QUELQUES PRECISIONS SUR LES ERREURS

NICOLAS BOULEAU*

*Exposé au colloquium de mathématiques de l'Université de Strasbourg
le 7 mars 2006*

Plan

- I. La dichotomie des petites erreurs
- II. Calculs d'erreurs intrinsèques
- III. Calcul d'erreur symétrique et complet
- IV. Les quatre opérateurs de biais
- V. Statistique et erreurs

*ENPC 28 rue des Saints Pères, 75007 Paris; e-mail : bouleau@enpc.fr

I. La dichotomie des petites erreurs.

Il y a deux sortes de petites erreurs qui ne suivent pas le même calcul différentiel.

Considérons deux chercheurs, disons deux mathématiciens appliqués, qui s'appliquent à faire de la simulation rigoureusement. Ils sont l'un et l'autre experts avec toutes les méthodes d'inversion et de rejet de sorte qu'ils savent simuler n'importe quelle loi de probabilité pourvu qu'ils puissent tirer un réel au hasard sur l'intervalle $[0, 1]$.

Pour cela bien sûr, ils disposent des méthodes de congruence, statistiquement excellentes, mais ils veulent être rigoureux :

- le premier tire les digits binaires à pile ou face
- le second utilise l'urne de Polya.

Examinons les biais et les variances.

Dans le premier cas $x = 0, a_1 a_2 a_3 \dots$ et on approche

$$x = \sum_{k=1}^{\infty} \frac{a_k}{2^k} \quad \text{par} \quad x_n = \sum_{k=1}^n \frac{a_k}{2^k} \quad a_k \in \{0, 1\}$$

On a

$$b_n = \mathbb{E}[(x - x_n) | \mathcal{F}_n] = \sum_{n+1}^{\infty} \frac{1/2}{2^k} = \frac{1}{2^{n+1}}$$

$$v_n = \mathbb{E}[(x - x_n)^2 | \mathcal{F}_n] = \frac{1}{3} \frac{1}{4^n}.$$

Dans le second cas, je rappelle le principe de l'urne de Polya. Il y a initialement une boule blanche et une boule noire dans une urne et chaque fois qu'on tire une boule on la remet avec une autre de la même couleur.

Après n tirages la proportion de boules blanches dans l'urne s'écrit

$$X_{n+1}(n+3) = X_n(n+2) + 1_{\{U_{n+1} \leq X_n\}}$$

où U_{n+1} est uniformément répartie sur $[0, 1]$ indépendante de $\mathcal{F}_n = \sigma(X_0, \dots, X_n)$.

Autrement dit

$$X_{n+1} = X_n + \frac{1}{n+3} (1_{\{U_{n+1} \leq X_n\}} - X_n).$$

On voit que X_n est une martingale qui converge p.s. et dans L^p , $p \in [1, \infty[$, vers X_∞ et il n'est pas difficile de montrer que lorsque la composition initiale de l'urne est une blanche et une noire, X_∞ est uniformément répartie sur $[0, 1]$.

On a pour le biais et la variance

$$b_n = \mathbb{E}[X_\infty - X_n | \mathcal{F}_n] = 0$$

$$v_n = \mathbb{E}[(X_\infty - X_n)^2 | \mathcal{F}_n] \quad \text{avec} \quad \mathbb{E}[v_n] = \frac{1}{6n} + o(1/n)$$

Nous voyons que dans le premier cas les variances sont beaucoup plus petites que les biais, et dans le second cas les biais sont plus petits que les variances.

Comment ceci va-t-il se propager dans les calculs de simulation de nos deux modélisateurs ?

Faisons un petit développement de Taylor sur une fonction \mathcal{C}^3 à dérivées bornées.

$$f(X) - f(X_n) = (X - X_n)f'(X_n) + \frac{1}{2}(X - X_n)^2 f''(X_n) + \frac{1}{6}(X - X_n)^3 f'''(X_n + \theta(X - X_n))$$

$$\begin{aligned} \mathbb{E}[f(X_\infty) - f(X_n) | \mathcal{F}_n] &= b_n f'(X_n) + \frac{1}{2} v_n f''(X_n) + o(v_n) \\ \mathbb{E}[(f(X_\infty) - f(X_n))^2 | \mathcal{F}_n] &= v_n f'^2(X_n) + o(v_n) \end{aligned}$$

Nous pouvons dégager trois cas

1/ Si la variance est négligeable devant le biais, $v_n \ll b_n$, ce qui est le cas pour le premier chercheur, le terme prépondérant asymptotiquement pour le biais est le premier terme, $\mathbb{E}[(f(X_\infty) - f(X_n))^2 | \mathcal{F}_n]$ est négligeable devant $\mathbb{E}[f(X_\infty) - f(X_n) | \mathcal{F}_n]$ et la situation va se perpétuer. On n'a besoin que de la formule

$$\mathbb{E}[f(X_\infty) - f(X_n) | \mathcal{F}_n] = b_n f'(X_n) + o(b_n)$$

2/ Si la variance est du même ordre de grandeur que le biais, cette situation va se maintenir.

3/ Si le biais est négligeable devant la variance, le terme prépondérant du biais devient $\frac{1}{2} v_n f''(X_n)$ et on retombe dans le cas 2/ où biais et variances ont le même ordre de grandeur.

Nous voyons que notre premier modélisateur peut se contenter d'un calcul d'erreur basé sur les dérivées premières. Alors que notre second modélisateur, celui qui utilise l'urne de Polya, doit faire un calcul d'erreur pour les biais et les variances

- le calcul d'erreur sur les variances est du premier ordre
- celui sur les biais est du second ordre et utilise celui sur les variances.

Dans la pratique, en général, on n'est pas maître de la nature des erreurs. Dans une modélisation les erreurs sur les données sont exogènes, on ne sait pas très bien d'où elles viennent. Il est donc prudent de faire comme si on était dans le second cas, notamment pour tenir compte de la nature aléatoire des erreurs par les non-linéarités du modèle.

II. Calculs d'erreurs intrinsèques.

Dans le calcul d'erreur sur les biais et les variances, le calcul sur les variances ne fait pas intervenir les biais, il est normal de commencer par lui, et ceci nous ramène à Gauss au début du XIXème siècle.

Le calcul d'erreur de Gauss sur les variances et covariances.

Douze ans après sa démonstration justifiant, sous certaines hypothèses, que les erreurs suivent une loi normale, Gauss s'intéresse à la propagation des erreurs (*Theoria combinationis* 1821). Etant donnée une grandeur $U = F(X_1, X_2, \dots)$ fonction d'autres grandeurs X_1, X_2, \dots , il pose le problème de calculer l'erreur quadratique sur U connaissant les erreurs quadratiques $\sigma_1^2, \sigma_2^2, \dots$ sur X_1, X_2, \dots , ces erreurs étant supposées petites et indépendantes.

Sa réponse est la suivante :

$$(1) \quad \sigma_U^2 = \left(\frac{\partial F}{\partial X_1}\right)^2 \sigma_1^2 + \left(\frac{\partial F}{\partial X_2}\right)^2 \sigma_2^2 + \dots$$

et pour une autre fonction $V = G(X_1, X_2, \dots)$ il donne la covariance de l'erreur sur U et V :

$$(2) \quad cov_{UV} = \frac{\partial F}{\partial X_1} \frac{\partial G}{\partial X_1} \sigma_1^2 + \frac{\partial F}{\partial X_2} \frac{\partial G}{\partial X_2} \sigma_2^2 + \dots$$

Gauss n'a pas étudié la propagation des biais qui, comme nous le verrons, est plus délicate.

Les errements des manuels : les formules "laides".

Malgré les travaux de Gauss, on a enseigné durant tout le XIXème siècle et encore au XXème siècle dans les livres de physique et de mathématiques des formules très ambiguës du genre

$$(3) \quad \Delta U = \left| \frac{\partial F}{\partial X_1} \right| \Delta X_1 + \left| \frac{\partial F}{\partial X_2} \right| \Delta X_2 + \dots$$

et encore dans le cours de *Mathématiques générales* 1947 de Vessiot et Montel, où les $\Delta X_1, \Delta X_2, \Delta U$, sont des valeurs absolues de l'erreur estimée, (3) étant parfois justifiée (J. Taylor *Incertitudes et analyses des erreurs dans les mesures physiques*, Dunod 2000) par l'inégalité

$$(4) \quad \sqrt{\left(\frac{\partial F}{\partial X_1}\right)^2 \sigma_1^2 + \dots + \left(\frac{\partial F}{\partial X_k}\right)^2 \sigma_k^2} \leq \left| \frac{\partial F}{\partial X_1} \right| \sigma_1 + \dots + \left| \frac{\partial F}{\partial X_k} \right| \sigma_k.$$

qui rendrait inutile les formules de Gauss (1) et (2). Indépendamment du fait que dans (3) on ne sait pas très bien ce que sont $\Delta X_1, \Delta X_2, \Delta U$ ces formules sont *moches*.

Ce n'est pas une question de goût et d'esthétique, c'est une affaire de symbolique et de concepts : avec (3) $|\Delta U|$ dépend en général de la manière d'écrire la fonction F . En composant deux applications linéaires à valeurs \mathbb{R}^2 on voit que l'application identique augmente les erreurs ... impossible de travailler proprement dans ces conditions.

Ceci ne se produit pas avec le calcul de Gauss. Pour le comprendre on peut introduire l'opérateur différentiel

$$L = \frac{1}{2} \sigma_1^2 \frac{\partial^2}{\partial X_1^2} + \frac{1}{2} \sigma_2^2 \frac{\partial^2}{\partial X_2^2} + \dots$$

et remarquer que (1) s'écrit

$$\sigma_U^2 = LF^2 - 2FLF.$$

La cohérence est alors liée au transport d'un opérateur différentiel par une fonction. Si u et v désignent des applications régulières injectives et si on note $\theta_u L$ l'opérateur $\varphi \mapsto L(\varphi \circ u) \circ u^{-1}$, on a $\theta_{v \circ u} L = \theta_v(\theta_u L)$. Nous allons préciser cela.

Géométrisation.

Les erreurs sur X_1, X_2, \dots peuvent ne pas être supposées indépendantes et peuvent dépendre des valeurs de X_1, X_2, \dots : on se donne un champ de matrices symétriques positives $\sigma_{ij}(x_1, x_2, \dots)$ sur \mathbb{R}^d représentant les variances et covariances conditionnelles des erreurs sachant les valeurs x_1, x_2, \dots de X_1, X_2, \dots , et le calcul d'erreur s'écrit

$$(5) \quad \sigma_F^2 = \sum_{ij} \frac{\partial F}{\partial X_i}(x_1, x_2, \dots) \frac{\partial F}{\partial X_j}(x_1, x_2, \dots) \sigma_{ij}(x_1, x_2, \dots)$$

Pour géométriser le calcul d'erreur, c'est-à-dire trouver un langage pour les biais et les variances qui ne dépende que des objets mathématiques et non de leur écriture, nous procéderons en deux temps.

D'abord nous raisonnerons sur \mathbb{R}^d en nous y prenant de telle sorte que si

$$F = f \circ g = k \circ h$$

l'erreur ne dépende que de F . Puis nous étendrons facilement cela au cas où F est à valeurs dans une variété.

a) Soit F une fonction régulière, disons \mathcal{C}^∞ , de mettons deux variables x et y , considérons l'accroissement de F entre (x, y) et $(x + dx, y + dy)$. Ici dx et dy sont des accroissements arbitraires.

La formule de Taylor donne

$$\Delta F = F(x+dx, y+dy) - F(x, y) = P_1(dx, dy) + \dots + \frac{1}{n!} P_n(dx, dy) + (|dx| + |dy|)^n o(|dx| + |dy|)$$

où

$$P_k(dx, dy) = \sum_{p+q=k} \frac{k!}{p!q!} \frac{\partial^j F}{\partial^p x \partial^q y} dx^p dy^q$$

est un polynôme homogène de degré k en dx, dy à coefficients dépendant de x, y qu'on appellera différentielle d'ordre k de F et qu'on notera $d^k F$.

De la formule

$$\Delta F = dF + \frac{1}{2} d^2 F + \dots + \frac{1}{n!} d^n F + (|dx| + |dy|)^n o(|dx| + |dy|)$$

on déduit les calculs différentiels qui nous intéressent :

si $x = f(u, v, w)$ et $y = g(u, v, w)$ et si $Z(u, v, w) = F(f(u, v, w), g(u, v, w))$ on a

$$(6) \quad \left\{ \begin{array}{l} dZ = P_1(du, dv, dw) = \frac{\partial F}{\partial x} dx + \frac{\partial F}{\partial y} dy \\ d^2Z = P_2(du, dv, dw) = d^2F \text{ (termes quadratiques en } dx, dy) \\ \quad + \frac{\partial F}{\partial x} d^2x + \frac{\partial F}{\partial y} d^2y \text{ (termes linéaires en } d^2x, d^2y \text{ quadratiques en } du, dv, dw) \\ \\ = \frac{\partial^2 F}{\partial x^2} (dx)^2 + 2 \frac{\partial^2 F}{\partial y \partial x} dx dy + \frac{\partial^2 F}{\partial y^2} (dy)^2 + \frac{\partial F}{\partial x} d^2x + \frac{\partial F}{\partial y} d^2y \end{array} \right.$$

ce qui peut se retenir en écrivant

$$\begin{aligned} d^2Z = d(dF) &= d\left(\frac{\partial F}{\partial x} dx + \frac{\partial F}{\partial y} dy\right) \\ &= d\left(\frac{\partial F}{\partial x}\right) dx + \frac{\partial F}{\partial x} d^2x + d\left(\frac{\partial F}{\partial y}\right) dy + \frac{\partial F}{\partial y} d^2y \end{aligned}$$

b) Randomisation. Les formules (6) sont identiquement valides si les accroissements du, dv, dw sont supposés aléatoires.

Simplement les symboles de Landau $o(\cdot)$ sont aléatoires et il convient de faire attention à l'intégrabilité des dérivées des fonctions dont nous prenons des espérances.

Nous pouvons donc réécrire avec des grandes lettres comme il est d'usage en calcul des probabilités : si $Z = F(X, Y)$

$$(7) \quad \left\{ \begin{array}{l} dZ = \frac{\partial F}{\partial x} dX + \frac{\partial F}{\partial y} dY \\ d^2Z = \frac{\partial^2 F}{\partial x^2} (dX)^2 + 2 \frac{\partial^2 F}{\partial y \partial x} dX dY + \frac{\partial^2 F}{\partial y^2} (dY)^2 + \frac{\partial F}{\partial x} d^2X + \frac{\partial F}{\partial y} d^2Y. \end{array} \right.$$

Si les fonctions f, g et F sont à dérivées bornées et si la v.a. dU, dV, dW sont de carré intégrable, on voit que dX, dY et dZ sont de carré intégrable et d^2X, d^2Y et d^2Z sont dans L^1 .

Nous introduisons alors les opérateurs de biais et de variance A et Γ associés à la variable aléatoire erronée (X, Y) ainsi définis :

$$(8) \quad \left\{ \begin{array}{l} A[F](x, y) = \mathbb{E}[dZ + \frac{1}{2}d^2Z \mid X=x, Y=y] \\ \Gamma[F](x, y) = \mathbb{E}[(dZ)^2 \mid X=x, Y=y] \end{array} \right.$$

indiquons tout de suite que dans la formule donnant $A[F]$ on peut mettre ou ne pas mettre dZ , nous allons y revenir, pour l'instant nous le laissons.

- Notons que si on applique A à la fonction F^2 , on a $d^2(F^2) = d(2FdF) = 2(dF)^2 + 2Fd^2F$ d'où

$$(9) \quad \begin{aligned} A[F^2] &= \mathbb{E}^{x,y}[2FdF + (dF)^2 + Fd^2F] = 2FA[F] + \Gamma[F] \\ \Gamma[F] &= A[F^2] - 2FA[F]. \end{aligned}$$

- Si nous considérons

$$H(X, Y) = \Phi(F_1(X, Y), F_2(X, Y), \dots, F_d(X, Y))$$

nous voyons par les formules de changement de variables (7) que

$$(10) \quad \begin{cases} A[H] = \sum_{i=1}^d \frac{\partial \Phi}{\partial F_i} A[F_i] + \frac{1}{2} \sum_{i,j=1}^d \frac{\partial^2 \Phi}{\partial F_i \partial F_j} \Gamma[F_i, F_j] \\ \Gamma[H] = \sum_{i,j=1}^d \frac{\partial \Phi}{\partial F_j} \frac{\partial \Phi}{\partial F_i} \Gamma[F_i, F_j] \end{cases}$$

la seconde relation généralise celle de Gauss au cas d'erreur éventuellement corrélées.

c) L'analogie des formules (10) avec la formule d'Ito pour les semi-martingales continues et leur crochet est frappante. On peut la préciser ainsi :

Considérons le champ de vecteurs

$$\begin{aligned} A[X](x, y) &= \mathbb{E}[dX + \frac{1}{2}d^2X | X = x, Y = y] \\ A[Y](x, y) &= \mathbb{E}[dY + \frac{1}{2}d^2Y | X = x, Y = y] \end{aligned}$$

et le champ de matrices

$$\underline{\underline{\Gamma}}(x, y) = \begin{pmatrix} \Gamma[X](x, y) & \Gamma[X, Y](x, y) \\ \Gamma[X, Y](x, y) & \Gamma[Y](x, y) \end{pmatrix}$$

et soit $\Sigma(x, y)$ une racine carrée régulière de $\underline{\underline{\Gamma}}$, considérons la diffusion S solution de l'eds

$$\begin{cases} dS_t^1 = \Sigma_{11}(S_t)dB_t^1 + \Sigma_{12}(S_t)dB_t^2 + A[X](S_t)dt \\ dS_t^2 = \Sigma_{21}(S_t)dB_t^1 + \Sigma_{22}(S_t)dB_t^2 + A[Y](S_t)dt \end{cases}$$

le calcul d'Ito donne

$$\begin{aligned} dF_i(S)_t &= \frac{\partial F_i}{\partial X}(S_t)dS_t^1 + \frac{\partial F_i}{\partial Y}(S_t)dS_t^2 \\ &+ \frac{1}{2}[\frac{\partial^2 F_i}{\partial X^2}(S_t)\Gamma[X](S_t) + 2\frac{\partial^2 F_i}{\partial X \partial Y}(S_t)\Gamma[X, Y](S_t) + \frac{\partial^2 F_i}{\partial Y^2}(S_t)\Gamma[Y](S_t)]dt \end{aligned}$$

et aussi, en notant $(.)^*$ la partie à variation finie continue

$$(11) \left\{ \begin{array}{l} (dF_i(S)_t)^* = \frac{\partial F_i}{\partial X}(S_t)A[X](S_t)dt + \frac{\partial F_i}{\partial Y}(S_t)A[Y](S_t)dt \\ \quad + \frac{1}{2}[\frac{\partial^2 F_i}{\partial X^2}(S_t)\Gamma[X](S_t) + 2\frac{\partial^2 F_i}{\partial X\partial Y}(S_t)\Gamma[X, Y](S_t) + \frac{\partial^2 F_i}{\partial Y^2}(S_t)\Gamma[Y](S_t)]dt \\ \text{et} \\ d \langle F_i(S), F_j(S) \rangle_t = \frac{\partial F_i}{\partial X}(S_t)\frac{\partial F_j}{\partial X}(S_t)\Gamma[X](S_t)dt + (\frac{\partial F_i}{\partial X}\frac{\partial F_j}{\partial Y} \\ \quad + \frac{\partial F_i}{\partial Y}\frac{\partial F_j}{\partial X})\Gamma[X, Y]dt + \frac{\partial F_i}{\partial Y}(S_t)\frac{\partial F_j}{\partial Y}(S_t)\Gamma[Y](S_t)dt \end{array} \right.$$

de sorte que les formules de changement de variables pour les biais et les variances des erreurs (10) sont vérifiées si nous définissons $A[F_i](x, y)$ et $\Gamma[F_i, F_j](x, y)$ par les relations

$$A[F_i](S_t) = \frac{(dF_i(S)_t)^*}{dt}$$

$$\Gamma[F_i, F_j](S_t) = \frac{d \langle F_i(S), F_j(S) \rangle_t}{dt}$$

de même pour H et tout autre fonction régulière de X et Y .

En résumé

Nous voyons que la donnée d'une variable aléatoire erronée (X, Y) à valeurs \mathbb{R}^2 peut se représenter par un opérateur différentiel d'ordre 1 aléatoire

$$F \mapsto b_{(X,Y)}[F] = \frac{\partial F}{\partial X}(X, Y)dX + \frac{\partial F}{\partial Y}(X, Y)dY$$

et un opérateur différentiel d'ordre 2 aléatoire

$$F \mapsto a_{(X,Y)}[F] = \frac{\partial F}{\partial x}d^2X + \frac{\partial F}{\partial y}d^2Y + \frac{\partial^2 F}{\partial x^2}(dX)^2 + 2\frac{\partial^2 F}{\partial y\partial x}dXdY + \frac{\partial^2 F}{\partial y^2}(dY)^2$$

où dX, dY, d^2X, d^2Y sont des variables aléatoires a priori quelconques.

Ensuite, pour un calcul d'erreur qui ne retient que les deux premiers moments de la loi conditionnelle de l'erreur sachant (X, Y) , on résume par deux opérateurs différentiels (déterministes)

l'opérateur de biais

$$A[F](x, y) = \mathbb{E}[b_{(X,Y)} + \frac{1}{2}a_{(X,Y)}[F] | (X, Y) = (x, y)]$$

et l'opérateur de variance

$$\Gamma[F](x, y) = \mathbb{E}[(b_{(X,Y)}[F])^2 | (X, Y) = (x, y)].$$

et nous avons les formules de changement de variables

$$(12) \quad \begin{cases} A[H] = \sum_{i=1}^d \frac{\partial \Phi}{\partial F_i} A[F_i] + \frac{1}{2} \sum_{i,j=1}^d \frac{\partial^2 \Phi}{\partial F_i \partial F_j} \Gamma[F_i, F_j] \\ \Gamma[H] = \sum_{i,j=1}^d \frac{\partial \Phi}{\partial F_j} \frac{\partial \Phi}{\partial F_i} \Gamma[F_i, F_j]. \end{cases}$$

a) Notons que nous avons identiquement

$$(b_{(X,Y)}[F])^2 = \frac{1}{2} a_{(X,Y)}[F^2] - F a_{(X,Y)}[F]$$

d'où il résulte que

$$(13) \quad \Gamma[F] = A[F^2] - 2FA[F]$$

l'opérateur Γ se déduit de l'opérateur $a_{(x,y)}$.

b) Si au lieu de $a_{(x,y)}$ nous avons pris

$$\hat{a}_{(X,Y)} = c_{(X,Y)} + a_{(X,Y)}$$

où c est un champ aléatoire d'opérateurs d'ordre 1, et donc

$$\hat{A}[F] = \mathbb{E}[b_{(X,Y)}[F] + \frac{1}{2} \hat{a}_{(X,Y)}[F] | (X, Y) = (x, y)]$$

les formules (12) et (13) seraient encore vérifiées avec \hat{A} , Γ restant inchangé.

Les remarques a) et b) ci-dessus font que si, pour la propagation des erreurs, on ne retient que les opérateurs A et Γ et les formules (12), l'opérateur A présente une certaine ambiguïté quant à ses termes d'ordre 1. On peut oublier l'opérateur du 1er ordre b . Cela se comprend par la nature même de la notion de biais qui nécessite une origine de référence. Cette référence change si l'on ajoute au départ une erreur déterministe du type parallaxe qui se propage selon un calcul du premier ordre.

d) Dans les variétés.

Rappelons qu'étant donnée une variété M , un *vecteur tangent d'ordre 2* en a est un opérateur différentiel au point a sans terme constant d'ordre ≤ 2 , on note $\tau_a(M)$ leur ensemble et $\tau(M)$ l'espace des champs de vecteurs d'ordre 2.

Une forme différentielle d'ordre 2 est une fonction \mathcal{C}^∞ sur $\tau(M)$ linéaire sur chaque $\tau_a(M)$. On note $\tau^*(M)$ l'espace des formes différentielles d'ordre 2.

Exemple : Si f et g sont des fonctions réelles sur M on définit les formes différentielles d'ordre deux $d^2 f = \lambda \in \tau(M) \mapsto \lambda(f)$ puis

$$df \cdot dg = \frac{1}{2}(d^2(fg) - fd^2g - gd^2f)$$

On dit que $\lambda \in \tau(M)$ est de type elliptique, si $\langle \lambda, df \cdot df \rangle \geq 0 \quad \forall f$. On note $\tau^e(M)$ l'espace des champs de vecteurs d'ordre 2 de type elliptique.

On se donne une variable aléatoire X à valeurs M et, en chaque point $x \in M$ une probabilité sur $\tau_x^e(M)$ régulière en x .

Autrement dit, on se donne un champ aléatoire de vecteurs tangents d'ordre 2 de type elliptique (dont on utilisera que les lois marginales d'ordre 1). On note ce champ ΔX : c'est l'"erreur" sur X .

Puis on définit

$$(14) \quad \begin{cases} A[f](x) = \frac{1}{2}\mathbb{E}[\langle d^2 f, \Delta X \rangle | X = x] \\ \Gamma[f](x) = \mathbb{E}[\langle df \cdot df, \Delta X \rangle | X = x] \end{cases}$$

Alors A est un champ de vecteurs tangents d'ordre 2 et Γ un champ d'opérateurs différentiels bilinéaires d'ordre 1 de type positif donné par

$$\Gamma[f] = A[f^2] - 2fA[f].$$

Si on considère $h = \varphi(f_1, f_2, \dots, f_k)$, on a en notant x^i un système de coordonnées

$$\begin{aligned} A[\varphi(f_1, f_2, \dots, f_k)](x) &= \frac{1}{2}\mathbb{E}_x[\langle \Delta X, d^2 h \rangle] \\ &= \frac{1}{2}\mathbb{E}_x[\langle \Delta X, \sum_i D_i h d^2 x^i + \sum_{ij} D_{ij} h dx^i \cdot dx^j \rangle] \end{aligned}$$

et

$$\begin{aligned} \Gamma[\varphi(f_1, f_2, \dots, f_k)](x) &= \mathbb{E}_x[\langle \Delta X, dh \cdot dh \rangle] \\ &= \mathbb{E}_x[\langle \Delta X, \sum_{ij} D_i h D_j h dx \cdot dx^j \rangle] \end{aligned}$$

avec

$$\begin{aligned} D_i h &= \sum_p \varphi'_p D_i f_p \\ D_{ij} h &= \sum_{p,q} \varphi''_{pq} D_i f_p D_j f_q + \sum_p \varphi'_p D_{ij} f_p \end{aligned}$$

d'où en remplaçant :

$$(15) \quad \begin{cases} A[\varphi(f_1, \dots, f_k)] = \sum_p \varphi'_p A[f_p] + \frac{1}{2} \sum_{p,q} \varphi''_{pq} \Gamma[f_p, f_q] \\ \Gamma[\varphi((f_1, \dots, f_k))] = \sum_{p,q} \varphi'_p \varphi'_q \Gamma[f_p, f_q] \end{cases}$$

L'interprétation en termes de diffusion est analogue au cas plat. On sait que si la différentielle d'Ito est $dF(S)_t$, sa partie à variation finie $(dF(S)_t)^*$ est intrinsèque, avec l'interprétation que nous avons donnée dans le cas plat, $F \mapsto A[F]$ pris en S_t est le vecteur tangent d'ordre 2 des caractéristiques locales de S .

Remarquons que si nous ajoutons à l'erreur ΔX un vecteur tangent d'ordre 1 aléatoire $\Delta X + b$, la formule (14) devient

$$A_b[f](x) = \frac{1}{2} \mathbb{E}_x[\langle df, b \rangle + \langle d^2 f, \Delta X \rangle]$$

car $d^2 f|_{TM} = df$ (cf. P.-A. Meyer *Sém. Prob* 15, 1981, page 49) ce qui ne change pas Γ et A_b vérifie encore les formules (15).

On voit donc que l'interprétation en termes d'erreur s'accorde avec le fait que sur une variété, la partie d'ordre 1 d'un vecteur tangent d'ordre 2 n'est pas définie, (sauf à disposer d'une connexion linéaire permettant de séparer la partie déterministe et la partie stochastique (martingale locale continue) de l'erreur). Déjà dans le cas plat, la partie d'ordre 1 du biais nécessite une convention au départ.

Remarque. Pourquoi nous sommes nous limités à ne retenir de la loi conditionnelle de l'erreur que les moments d'ordre 1 et 2 alors que le formalisme des différentielles d'ordre n et des vecteurs tangents d'ordre n nous permettait de traiter aussi bien la propagation des autres moments ? Disons d'abord par souci de simplicité, mais une réponse plus précise pourra être donnée plus loin (Remarque page 23).

Variété produits. (pour mémoire)

L'idée que les erreurs d'un couple de variables aléatoires indépendantes sont non-corrélées s'écrit parfaitement avec le langage que nous venons d'introduire.

Images. (pour mémoire)

Une application régulière (déterministe mais pas nécessairement injective) d'une variété M dans une variété N transporte les erreurs. Les opérateurs de biais et de variance s'écrivent explicitement.

III. Calcul d'erreur symétrique et complet.

a) La symétrie comme invariant

Si on part d'une situation où l'erreur est centrée, par image non-linéaire l'erreur n'est plus centrée.

Qu'est-ce qui se conserve par image ?

Si X_n est une approximation de X , et si la loi du couple (X, X_n) est symétrique, alors la loi de $(\psi(X), \psi(X_n))$ est symétrique évidemment. Concrètement cela représente des situations où l'on ne sait pas de deux variables aléatoires proches X_0 et X_1 quelle est la bonne quelle est l'erronée, et dans cette hésitation on travaille avec (X_B, X_{1-B}) où B est une variable de Bernoulli indépendante.

Reprenons les notations du cas plat. Si $(X, X + \Delta X)$ est un couple symétrique, nous avons

$$\begin{aligned} F(X + \Delta X) - F(X) &= dF + \frac{1}{2}d^2F + \text{reste} \\ G(X + \Delta X) - G(X) &= dG + \frac{1}{2}d^2G + \text{reste} \end{aligned}$$

d'où

$$\begin{aligned} \mathbb{E}[F(X + \Delta X)G(X) - F(X)G(X + \Delta X)] \\ &= \mathbb{E}[G(X)(dF + \frac{1}{2}d^2F) - F(X)(dG + \frac{1}{2}d^2G) + \text{reste}] \\ &= \mathbb{E}[G(X)A[F](X)] - \mathbb{E}[F(X)A[G](X)] + \text{reste} \\ &= \langle G, AF \rangle_\nu - \langle F, AG \rangle_\nu + \text{reste} \end{aligned}$$

en notant ν la loi de X .

Nous voyons que si le couple $(X, X + \Delta X)$ est symétrique et l'erreur petite, l'opérateur A est symétrique par rapport à la loi de X . Cette propriété de symétrie (qui résoud l'ambiguïté des termes d'ordre 1 de A) est un invariant par image : l'opérateur de biais de l'erreur de l'image de X par une application φ est symétrique par rapport la loi de $\varphi(X)$.

Le cas symétrique permet de construire un cadre beaucoup plus puissant grâce à la théorie des formes de Dirichlet, cadre qui s'étend sans difficultés à la dimension infinie (cf. N. Bouleau, *Error calculus for finance and physics*, De Gruyter 2003). Nous l'esquisons maintenant.

b) Structures d'erreur

Une structure d'erreur est un terme

$$(\Omega, \mathcal{A}, \mathbb{P}, \mathbb{D}, \Gamma)$$

où $(\Omega, \mathcal{A}, \mathbb{P})$ est un espace de probabilité, satisfaisant les propriétés:

- 1.) \mathbb{D} est un sous-espace dense de $L^2(\Omega, \mathcal{A}, \mathbb{P})$
- 2.) Γ est une application bilinéaire symétrique de $\mathbb{D} \times \mathbb{D}$ dans $L^1(\mathbb{P})$ vérifiant le calcul de Gauss de classe $\mathcal{C}^1 \cap \text{Lip}$, c'est à dire que si $u \in \mathbb{D}^m$ et $v \in \mathbb{D}^n$, pour F et G de classe \mathcal{C}^1 et lipschitziennes de \mathbb{R}^m [resp. \mathbb{R}^n] dans \mathbb{R} , on a $F \circ u \in \mathbb{D}$ et $G \circ v \in \mathbb{D}$ et

$$\Gamma[F \circ u, G \circ v] = \sum_{i,j} F'_i(u) G'_j(v) \Gamma[u_i, v_j] \quad \mathbb{P}\text{-p.s.}$$

- 3.) La forme bilinéaire $\mathcal{E}[f, g] = \mathbb{E}\Gamma[f, g]$ est fermée, i.e. \mathbb{D} est complet pour la norme $\|\cdot\|_{\mathbb{D}} = (\|\cdot\|_{L^2(\mathbb{P})}^2 + \mathcal{E}[\cdot, \cdot])^{\frac{1}{2}}$.
- 4.) $1 \in \mathbb{D}$ (donc $\Gamma[1, 1] = 0$ markovianité).

On écrit $\mathcal{E}[f]$ pour $\mathcal{E}[f, f]$ et $\Gamma[f]$ pour $\Gamma[f, f]$.

Avec cette définition, la forme \mathcal{E} est une *forme de Dirichlet*.

Il lui correspond un opérateur de Dirichlet A (générateur du semi-groupe associé à la forme de Dirichlet \mathcal{E}) qui vérifie (sous certaines hypothèses sur F) :

$$A[F \circ u] = \sum_i F'_i \circ u A[u_i] + \frac{1}{2} \sum_{i,j} F''_{ij} \circ u \Gamma[u_i, u_j] \quad \mathbb{P}\text{-p.s.}$$

Exemple 1. (Structure d'Ornstein-Uhlenbeck en dimension 1)

$$\begin{aligned}
 \Omega &= \mathbb{R} \\
 \mathcal{A} &= \text{tribu borélienne } \mathcal{B}(\mathbb{R}) \\
 \mathbb{P} &= \mathcal{N}(0, 1) \text{ loi normale réduite} \\
 \mathbb{D} = H^1(\mathcal{N}(0, 1)) &= \{u \in L^2(\mathbb{P}), u' \text{ au sens des distributions} \\
 &\quad \text{appartient à } L^2(\mathbb{P})\} \\
 \Gamma[u] &= u'^2
 \end{aligned}$$

alors, $(\mathbb{R}, \mathcal{B}(\mathbb{R}), \mathcal{N}(0, 1), H^1(\mathcal{N}(0, 1)), \Gamma)$ est une structure d'erreur. Nous pouvons aussi obtenir l'opérateur de biais (le générateur associé):

$$\mathcal{D}A = \{f \in L^2(\mathbb{P}): f'' - xf' \text{ au sens des distributions} \in L^2(\mathbb{P})\}$$

et $Af = \frac{1}{2}f'' - \frac{1}{2}I \cdot f'$ où I est l'application identité de \mathbb{R} .

Exemple 2. (Structure de Monte-Carlo en dimension 1)

$$\begin{aligned}
 \Omega &= [0, 1] \\
 \mathcal{A} &= \text{tribu borélienne} \\
 \mathbb{P} &= \text{mesure de Lebesgue} \\
 \mathbb{D} = \{u \in L^2([0, 1], dx): \text{la dérivée } u' \text{ au sens des distributions} \\
 &\quad \text{sur }]0, 1[\text{ appartient à } L^2([0, 1], dx)\} \\
 \Gamma[u] &= u'^2.
 \end{aligned}$$

L'espace \mathbb{D} ainsi défini est noté $H^1([0, 1])$.

Exemple 3. Soient U un domaine (ouvert connexe) de \mathbb{R}^d de volume unité, $\mathcal{B}(U)$ la tribu borélienne et $dx = dx_1, \dots, dx_d$ la mesure de Lebesgue,

$$\begin{aligned}
 \mathbb{D} &= \{u \in L^2(U, dx): \text{le gradient } \nabla u \text{ au sens des distributions} \\
 &\quad \text{appartient à } L^2(U, dx; \mathbb{R}^d)\} \\
 \Gamma[u] &= |\nabla u|^2 = \left(\frac{\partial u}{\partial x_1}\right)^2 + \dots + \left(\frac{\partial u}{\partial x_d}\right)^2.
 \end{aligned}$$

Alors $(U, \mathcal{B}(U), dx, \mathbb{D}, \Gamma)$ est une structure d'erreur. De la relation $\mathcal{E}[f, g] = \langle -Af, g \rangle$ on tire que le domaine du générateur contient les fonctions de classe \mathcal{C}^2 à support compact dans U , $\mathcal{DA} \supset \mathcal{C}_K^2(U)$ et que pour de telles fonctions

$$Af = \frac{1}{2}\Delta f = \frac{1}{2} \sum_{i=1}^d \frac{\partial^2 f}{\partial x_i^2}.$$

Exemple 4.

Soit D un ouvert connexe de \mathbb{R}^d de volume unité. Soit $\mathbb{P} = dx$ la mesure de Lebesgue sur D . Soit Γ défini sur $\mathcal{C}_K^\infty(D)$ par

$$\Gamma[u, v] = \sum_{ij} \frac{\partial u}{\partial x_i} \frac{\partial v}{\partial x_j} a_{ij}, \quad u, v \in \mathcal{C}_K^\infty(D)$$

où les fonctions a_{ij} vérifient les hypothèses suivantes

- $a_{ij} \in L_{\text{loc}}^2(D) \quad \frac{\partial a_{ij}}{\partial x_k} \in L_{\text{loc}}^2(D) \quad i, j, k = 1, \dots, d$
- $\sum_{i,j} a_{ij}(x) \xi_i \xi_j \geq 0 \quad \forall \xi \in \mathbb{R}^d \quad \forall x \in D$
- $a_{ij}(x) = a_{ji}(x) \quad \forall x \in D.$

Alors la pré-structure $(D, \mathcal{B}(D), \mathbb{P}, \mathcal{C}_K^\infty(D), \Gamma)$ est fermable.

Commentaire sur le caractère complet

Le fait que nous exigeons que la forme \mathcal{E} soit *fermée* est une restriction, mais une restriction extrêmement féconde. La situation est tout à fait analogue à la question de la σ -additivité en calcul des probabilités : sans cette propriété, on ne peut rien dire de la transmission des erreurs par des objets qui sont définis par des limites. Or beaucoup d'objets des mathématiques contemporaines sont définis par des limites (intégrales, solutions d'edo, solutions d'edp, intégrales stochastiques, solutions d'eds, etc.)¹

¹Le philosophe Karl Popper est tombé dans ce piège en soulignant que sa théorie des probabilités (additives) contenait celle de Kolmogorov (σ -additives) Cf. N.Bouleau "Some thoughts upon axiomatized languages, a focus on probability theory and error calculus with Dirichlet forms" Butlleti de la Societat Catalana de Matemàtiques Vol. 18 n2 p25-36, (2004) cf.<http://www.enpc.fr/HomePages/bouleau/nb1.html>

Cet outil d'extension permet

- d'étendre le calcul d'erreur relatif aux variances des fonctions \mathcal{C}^1 aux fonctions lipschitziennes,
- d'établir une propriété d'existence de densité des lois qui généralise la méthode de Malliavin : la propriété de densité de l'énergie image,
- la propriété de fermeture est conservée par image et par produit, même par produit infini. Ceci permet de construire naturellement des structures d'erreur sur les espaces de processus aléatoires,
- en particulier sur l'espace de Wiener, où l'on obtient pour A l'opérateur d'Ornstein-Uhlenbeck. Le calcul de Malliavin s'interprète alors comme un calcul d'erreur.

Théorème des produits

Soient $S_n = (\Omega_n, \mathcal{A}_n, \mathbb{P}_n, \mathbb{D}_n \Gamma_n)$, $n \geq 1$, des structures d'erreur. La structure produit

$$S = (\Omega, \mathcal{A}, \mathbb{P}, \mathbb{D}, \Gamma) = \prod_{n=1}^{\infty} S_n$$

est définie par

$$(\Omega, \mathcal{A}, \mathbb{P}) = \left(\prod_{n=1}^{\infty} \Omega_n, \otimes_{n=1}^{\infty} \mathcal{A}_n, \prod_{n=1}^{\infty} \mathbb{P}_n \right)$$

$$\mathbb{D} = \left\{ f \in L^2(\mathbb{P}) : \forall n, \text{ pour presque tout } w_1, w_2, \dots, w_{n-1}, w_{n+1}, \dots \right. \\ \left. \text{pour la mesure produit} \right. \\ \left. x \rightarrow f(w_1, \dots, w_{n-1}, x, w_{n+1}, \dots) \in \mathbb{D}_n \text{ et} \right. \\ \left. \int \sum_n \Gamma_n[f] d\mathbb{P} < +\infty \right\}$$

et pour $f \in \mathbb{D}$

$$\Gamma[f] = \sum_{n=1}^{\infty} \Gamma_n[f].$$

S est une structure d'erreur, markovienne si chaque S_n l'est.

Lorsque nous écrivons $\Gamma_n[f]$, Γ_n opère sur le n -ième argument de f seulement.

A partir des briques de base que sont les structures d'erreur en dimension 1 nous pouvons ainsi, par produit, obtenir des structures d'erreur sur des espaces fonctionnels. Ceci fournit aisément des structures sur

- l'espace de Wiener, cf N. B. & F. Hirsch, *Dirichlet Forms and Analysis on Wiener Space*, De Gruyter 1991,

- l'espace de Poisson général, l'espace de Monte Carlo, cf N. B. *Error Calculus for Finance and Physics*, De Gruyter, 2003.

Images de structures d'erreur

L'opération est aussi simple et presque'aussi générale que l'image d'un espace de probabilité par une application.

Si $(\Omega, \mathcal{A}, \mathbb{P}, \mathbb{D}, \Gamma)$ est une structure d'erreur et X une variable aléatoire à valeurs \mathbb{R}^d dont les composantes sont dans \mathbb{D} , le terme $(\mathbb{R}^d, \mathcal{B}(\mathbb{R}^d), \mathbb{P}_X, \mathbb{D}_X, \Gamma_X)$ est une structure d'erreur où

$$\begin{aligned} \mathbb{P}_X & \text{ est la loi de } X, \\ \mathbb{D}_X & = \{f \in L^2(\mathbb{P}_X) : f \circ X \in \mathbb{D}\} \\ \Gamma_X[f](x) & = \mathbb{E}\{\Gamma[f \circ X] | X = x\}, \quad f \in \mathbb{D}. \end{aligned}$$

En fait on peut définir l'image par des variables aléatoires plus générales.

La structure image par X peut être appelée la "Dirichlet-loi" de X . C'est une structure sur \mathbb{R}^d telle que l'identité I a ses composantes dans le domaine de Γ_X et on a les formules:

$$\begin{aligned} \mathbb{E}[\Gamma[X] | X] & = \Gamma_X[I] \circ X \\ \mathbb{E}[\Gamma[\varphi(X)] | X] & = \Gamma_X[\varphi] \circ X. \end{aligned}$$

Plusieurs théorèmes des probabilités ont des analogues en théorie des structures d'erreur sous certaines conditions: Gateaux-Lévy, Strassen, etc. (cf. Bouleau-Hirsch [de Gruyter 1991].)

Cas de l'espace de Wiener

- Soit $(\chi_n)_{n \in \mathbb{N}}$ une base orthonormale de $L^2(\mathbb{R}_+, \mathcal{B}(\mathbb{R}_+), dx)$, et $(g_n)_{n \in \mathbb{N}}$ une suite de variables i.i.d. gaussiennes réduites.

A une fonction $f \in L^2(\mathbb{R}_+, \mathcal{B}(\mathbb{R}_+), dx)$ on associe l'intégrale de Wiener

$$I(f) = \sum_n \langle f, \chi_n \rangle g_n,$$

homomorphisme de $L^2(\mathbb{R}_+, dx)$ dans $L^2(\Omega, \mathcal{A}, \mathbb{P})$. Si nous posons

$$B(t) = \sum_n \langle 1_{[0,t]}, \chi_n \rangle g_n = \sum_n \int_0^t \chi_n(y) dy \cdot g_n$$

alors $B(t)$ est un mouvement brownien standard.

A cause du cas où f est en escalier, $I(f)$ est noté $I(f) = \int_0^\infty f(s) dB_s$

- La construction précédente utilise l'espace produit

$$(\Omega, \mathcal{A}, \mathbb{P}) = (\mathbb{R}, \mathcal{B}(\mathbb{R}), \mathcal{N}(0, 1))^{\mathbb{N}},$$

les g_n 's étant les coordonnées. Si sur chaque facteur on place une structure d'erreur

$$(\mathbb{R}, \mathcal{B}(\mathbb{R}), \mathcal{N}(0, 1), \mathbf{d}_n, \gamma_n),$$

on obtient une structure d'erreur sur $(\Omega, \mathcal{A}, \mathbb{P})$:

$$(\Omega, \mathcal{A}, \mathbb{P}, \mathbb{D}, \Gamma) = \prod_{n=0}^{\infty} (\mathbb{R}, \mathcal{B}(\mathbb{R}), \mathcal{N}(0, 1), \mathbf{d}_n, \gamma_n)$$

telle qu'une variable aléatoire $F(g_0, g_1, \dots, g_n, \dots)$ appartient à \mathbb{D} ssi

$$\begin{cases} \forall n \quad x \rightarrow F(g_0, \dots, g_{n-1}, x, g_n, \dots) \text{ appartient à } \mathbf{d}_n \quad \mathbb{P}\text{-p.s.} \\ \text{et } \Gamma[F] = \sum_n \gamma_n[F], \quad \text{appartient à } L^1(\mathbb{P}) (\gamma_n \text{ opérant sur la } n\text{-ième variable de } F). \end{cases}$$

- Si sur chaque facteur on prend la structure d'Ornstein-Uhlenbeck de dimension 1, on obtient

$$\begin{aligned}\Gamma[g_n] &= 1 \\ \Gamma[g_m, g_n] &= 0 \quad \text{if } m \neq n.\end{aligned}$$

Pour $f \in L^2(\mathbb{R}_+)$, de $\int_0^\infty f(s) dB_s = \sum_n \langle f, \chi_n \rangle g_n$ on tire

$$\Gamma \left[\int_0^\infty f(s) dB_s \right] = \sum_n \langle f, \chi_n \rangle^2 = \|f\|_{L^2(\mathbb{R}_+)}^2,$$

D'où l'on déduit par le calcul fonctionnel $\forall F \in \mathcal{C}^1 \cap \text{Lip}(\mathbb{R}^m)$

$$\Gamma \left[F \left(\int f_1(s) dB_s, \dots, \int f_n(s) dB_s \right) \right] = \sum_{i,j} \frac{\partial F}{\partial x_i} \frac{\partial F}{\partial x_j} \int f_i(s) f_j(s) ds.$$

C'est la structure d'Ornstein-Uhlenbeck sur l'espace de Wiener.

- Restreignons-nous à $t \in [0, 1]$ pour simplifier et prenons pour χ_n la base trigonométrique. Si sur chaque facteur on prend la structure d'Ornstein-Uhlenbeck de dimension 1 homothétiquement modifiée par un coefficient constant dépendant de n :

$$\prod_{n=0}^{\infty} (\mathbb{R}, \mathcal{B}(\mathbb{R}), \mathcal{N}(0, 1), H^1(\mathcal{N}(0, 1)), u \rightarrow (2\pi n)^{2q} u'^2)$$

on obtient sur l'espace de Wiener une structure d'erreur vérifiant sur le premier chaos :

$$\Gamma \left[\int_0^1 f(s) dB_s \right] = \int_0^1 f^{(q)2}(s) ds$$

où $f^{(q)}$ est la dérivée q -ième de f . C'est une structure où l'erreur perturbe longitudinalement la trajectoire brownienne, qui entre dans la famille des structures dites de type "Mehler généralisé".

IV. Les quatre opérateurs de biais.

Nous abordons maintenant la question suivante : comment une erreur au sens habituel en mathématiques, erreur d'approximation, engendre-t-elle une structure d'erreur.

Considérons une variable aléatoire Y définie sur l'espace de probabilité $(\Omega, \mathcal{A}, \mathbb{P})$ à valeurs dans l'espace mesurable (E, \mathcal{F}) et des approximations Y_n , $n \in \mathbb{N}$, définies également sur $(\Omega, \mathcal{A}, \mathbb{P})$ à valeurs dans (E, \mathcal{F}) .

On suppose qu'il existe une algèbre \mathcal{D} de fonctions bornées de E dans \mathbb{R} ou \mathbb{C} dense dans $L^2(E, \mathcal{F}, \mathbb{P}_Y)$ contenant les constantes et une suite $(\alpha_n)_{n \in \mathbb{N}}$ de nombres positifs, avec lesquels on considère les hypothèses suivantes :

$$(H1) \quad \begin{cases} \forall \varphi \in \mathcal{D}, \text{ il existe } \bar{A}[\varphi] \in L^2(E, \mathcal{F}, \mathbb{P}_Y) \quad \text{t.q.} \quad \forall \chi \in \mathcal{D} \\ \lim_{n \rightarrow \infty} \alpha_n \mathbb{E}[(\varphi(Y_n) - \varphi(Y))\chi(Y)] = \mathbb{E}_Y[\bar{A}[\varphi]\chi]. \end{cases}$$

l'espérance \mathbb{E}_Y étant relative à la loi \mathbb{P}_Y .

$$(H2) \quad \begin{cases} \forall \varphi \in \mathcal{D}, \text{ il existe } \underline{A}[\varphi] \in L^2(E, \mathcal{F}, \mathbb{P}_Y) \quad \text{t.q.} \quad \forall \chi \in \mathcal{D} \\ \lim_{n \rightarrow \infty} \alpha_n \mathbb{E}[(\varphi(Y) - \varphi(Y_n))\chi(Y_n)] = \mathbb{E}_Y[\underline{A}[\varphi]\chi]. \end{cases}$$

$$(H3) \quad \begin{cases} \forall \varphi \in \mathcal{D}, \text{ il existe } \tilde{A}[\varphi] \in L^2(E, \mathcal{F}, \mathbb{P}_Y) \quad \text{t.q.} \quad \forall \chi \in \mathcal{D} \\ \lim_{n \rightarrow \infty} \alpha_n \mathbb{E}[(\varphi(Y_n) - \varphi(Y))(\chi(Y_n) - \chi(Y))] = -2\mathbb{E}_Y[\tilde{A}[\varphi]\chi]. \end{cases}$$

Dès que deux des hypothèses (H1) (H2) (H3) sont vérifiées (avec la même algèbre \mathcal{D} et la même suite α_n), la troisième l'est aussi grâce à la relation

$$\tilde{A} = \frac{\bar{A} + \underline{A}}{2}.$$

- Lorsqu'il est défini l'opérateur \bar{A} qui considère l'erreur asymptotique du point de vue du modèle limite, sera appelé *l'opérateur de biais théorique*.

- l'opérateur \underline{A} qui considère l'erreur asymptotique du point de vue du modèle approchant, sera appelé *l'opérateur de biais pratique*.

- A cause de la propriété

$$\langle \tilde{A}[\varphi], \chi \rangle_{L^2(\mathbb{P}_Y)} = \langle \varphi, \tilde{A}[\chi] \rangle_{L^2(\mathbb{P}_Y)}$$

l'opérateur \tilde{A} sera appelé *l'opérateur de biais symétrique*.

Le résultat suivant montre qu'une forme de Dirichlet (parfois non-locale) se cache souvent derrière une approximation :

Théorème. *Sous l'hypothèse (H3)*

a) *la limite*

$$\tilde{\mathcal{E}}[\varphi, \chi] = \lim_n \frac{\alpha_n}{2} \mathbb{E}[(\varphi(Y_n) - \varphi(Y))(\chi(Y_n) - \chi(Y))] \quad \varphi, \chi \in \mathcal{D}$$

définit une forme bilinéaire positive fermable dont la plus petite extension fermée est notée $(\mathcal{E}, \mathbb{D})$.

b) $(\mathcal{E}, \mathbb{D})$ est une forme de Dirichlet

c) $(\mathcal{E}, \mathbb{D})$ admet un opérateur carré du champ Γ vérifiant $\forall \varphi, \chi \in \mathcal{D}$

$$\Gamma[\varphi] = \tilde{A}[\varphi^2] - 2\varphi\tilde{A}[\varphi]$$

$$\mathbb{E}_Y[\Gamma[\varphi]\chi] = \lim_n \alpha_n \mathbb{E}[(\varphi(Y_n) - \varphi(Y))^2(\chi(Y_n) + \chi(Y))/2]$$

d) $(\mathcal{E}, \mathbb{D})$ est locale si et seulement si $\forall \varphi \in \mathcal{D}$

$$\lim_n \alpha_n \mathbb{E}[(\varphi(Y_n) - \varphi(Y))^4] = 0.$$

J'introduis maintenant le quatrième opérateur de biais \mathfrak{A} défini sous (H1) et (H2) sur \mathcal{D} par

$$\mathfrak{A} = \frac{1}{2}(\overline{A} - \underline{A}).$$

Comme $\mathbb{E}_Y[\mathfrak{A}[\varphi]\chi] = \lim_n \mathbb{E}[(\varphi(Y_n) - \varphi(Y))(\chi(Y) + \chi(Y_n))/2]$ nous voyons que \mathfrak{A} représente l'erreur asymptotique du point de vue d'un observateur extérieur accordant le même poids au modèle théorique et au modèle approché et mesurant l'erreur algébriquement sur le même axe. A cause des propriétés que nous verrons plus bas l'opérateur \mathfrak{A} sera appelé *l'opérateur de biais singulier*.

On dira qu'un opérateur B de \mathcal{D} dans $L^2(\mathbb{P}_Y)$ est du *premier ordre* s'il satisfait

$$B[\varphi\chi] = B[\varphi]\chi + \varphi B[\chi] \quad \forall \varphi, \chi \in \mathcal{D}$$

Proposition. *Sous (H1) à (H3)*

a) *la variance théorique $\lim_n \alpha_n \mathbb{E}[(\varphi(Y_n) - \varphi(Y))^2 \psi(Y)]$ et la variance pratique $\lim_n \alpha_n \mathbb{E}[(\varphi(Y) - \varphi(Y_n))^2 \psi(Y_n)]$ existent et nous avons $\forall \varphi, \chi, \psi \in \mathcal{D}$*

$$\begin{aligned} \lim_n \alpha_n \mathbb{E}[(\varphi(Y_n) - \varphi(Y))(\chi(Y_n) - \chi(Y))\psi(Y)] &= \mathbb{E}_Y[-\underline{A}[\varphi\psi]\chi + \underline{A}[\psi]\varphi\chi - \overline{A}[\varphi]\chi\psi] \\ \lim_n \alpha_n \mathbb{E}[(\varphi(Y) - \varphi(Y_n))(\chi(Y_n) - \chi(Y))\psi(Y_n)] &= \mathbb{E}_Y[-\overline{A}[\varphi\psi]\chi + \overline{A}[\psi]\varphi\chi - \underline{A}[\varphi]\chi\psi] \end{aligned}$$

b) *Ces deux variances coïncident si et seulement si \mathbb{A} est du premier ordre, et alors elles sont égales à $\mathbb{E}_Y[\Gamma[\varphi]\psi]$.*

c) *Si la forme de Dirichlet est locale, alors \mathbb{A} est du premier ordre.*

Remarque. Sous (H3) la condition d) du théorème caractérisant le cas où la forme \mathcal{E} est locale est équivalente à l'une ou l'autre des conditions suivantes :

$$\begin{aligned} \text{(j)} \quad \exists \lambda > 2 \quad \lim_n \alpha_n \mathbb{E}[|\varphi(Y_n) - \varphi(Y)|^\lambda] &= 0 \quad \forall \varphi \in \mathcal{D}. \\ \text{(jj)} \quad \forall \lambda > 2 \quad \lim_n \alpha_n \mathbb{E}[|\varphi(Y_n) - \varphi(Y)|^\lambda] &= 0 \quad \forall \varphi \in \mathcal{D}. \end{aligned}$$

Ceci apporte une réponse à la question posée page 12 concernant le calcul d'erreur avec des moments d'ordre supérieurs à 2 en utilisant des vecteurs tangents d'ordre supérieur à 2 : Dans les cas où la forme de Dirichlet est locale (et où nous pouvons propager les erreurs par un calcul différentiel) les moments d'ordre > 2 sont négligeables devant les moments d'ordre 1 ou 2.

Revenons à la situation où seule est supposée l'hypothèse (H3). Le résultat suivant montre que, pour les variances, le calcul d'erreur sur le modèle limite effectué avec l'erreur asymptotique coïncide avec l'erreur d'approximation obtenue asymptotiquement sur les fonctions \mathcal{C}^1 .

Proposition. *Sous (H3). Si la forme $(\mathcal{E}, \mathbb{D})$ est locale, alors le principe du calcul d'erreur asymptotique est valide sur*

$$\tilde{\mathcal{D}} = \{F(f_1, \dots, f_p) : f_i \in \mathcal{D}, F \in \mathcal{C}^1(\mathbb{R}^p, \mathbb{R})\}$$

$$\begin{aligned} i.e. \quad \lim_n \alpha_n \mathbb{E}[(F(f_1(Y_n), \dots, f_p(Y_n)) - F(f_1(Y), \dots, f_p(Y)))^2] \\ = \mathbb{E}_Y[\sum_{i,j=1}^p F'_i(f_1, \dots, f_p) F'_j(f_1, \dots, f_p) \Gamma[f_i, f_j]]. \end{aligned}$$

Exemples

- Prenons pour Y un mouvement brownien B indexé par $[0, 1]$ en tant que variable aléatoire à valeurs $\mathcal{C}([0, 1])$ et prenons pour Y_ε l'approximation $Y_\varepsilon = B + \sqrt{\varepsilon}W$ où W est

un mouvement brownien indépendant, nous pouvons appliquer le théorème avec pour \mathcal{D} les combinaisons linéaires de fonctions $\varphi(B) = e^{i \int_0^1 f dB}$ avec f régulière disons \mathcal{C}_b^1 .

L'hypothèse (H3) est vérifiée. Le théorème fournit la structure d'Ornstein-Uhlenbeck sur l'espace de Wiener.

- Séries à accroissements indépendants. Soit

$$S = \sum_{n=1}^{\infty} \frac{X_n}{n^2} + \frac{Z_n}{n}$$

où $X_n, Z_n \in L^{2+\varepsilon}$, Z_n centrée, (X_n, Z_n) i.i.d., on approche S par sa somme partielle $S_n = \sum_{k=1}^n \frac{X_k}{k^2} + \frac{Z_k}{k}$.

Par l'inégalité de Burholder, on a $n\mathbb{E}[|S - S_n|^{2+\varepsilon}] \rightarrow 0$ as $n \rightarrow \infty$. De sorte que, prenant $\mathcal{D} = \mathcal{C}_K^\infty$, on a pour $\varphi, \chi \in \mathcal{D}$

$$\begin{aligned} \lim_n n\mathbb{E}[(\varphi(S) - \varphi(S_n))\chi(S_n)] &= \lim_n n\mathbb{E}[(S - S_n)\varphi'(S_n)\chi(S_n) + \frac{1}{2}(S - S_n)^2\varphi''(S_n)\chi(S_n)] \\ &= \frac{1}{2}\mathbb{E}[Z_1^2]\mathbb{E}[\varphi''(S)\chi(S)] + \mathbb{E}[X_1]\mathbb{E}[\varphi'(S)\chi(S)]. \end{aligned}$$

Donc l'hypothèse (H2) est satisfaite et

$$\underline{A}[\varphi] = \frac{\mathbb{E}[Z_1^2]}{2}\varphi'' + \mathbb{E}[X_1]\varphi'.$$

De même

$$\lim_n n\mathbb{E}[(\varphi(S) - \varphi(S_n))^2] = \lim_n n\mathbb{E}[(S - S_n)^2\varphi'^2(S_n)] = \mathbb{E}[Z_1^2]\mathbb{E}[\varphi'^2(S)]$$

(H3) est vérifiée dès que la loi de S satisfait la condition de Hamza (cf. M. Fukushima et al. [1994]) et alors la forme de Dirichlet est locale, et $\Gamma[\varphi] = \underline{A}[\varphi^2] - 2\varphi\underline{A}[\varphi] = \mathbb{E}[Z_1^2]\varphi'^2$.

- Intégrale stochastique. Considérons l'intégrale

$$Y = \int_0^1 H_s dB_s$$

approchée par la somme

$$Y_n = \sum_{k=0}^{n-1} H_{\frac{k}{n}}(B_{\frac{k+1}{n}} - B_{\frac{k}{n}})$$

(B_t) est un brownien standard défini comme le processus des coordonnées sur l'espace $\mathcal{C}([0, 1])$ muni de la mesure de Wiener, $H_s = H_0 + \int_0^s \xi_u dB_u + \int_0^s \eta_u du$ est un processus d'Ito régulier au sens de Malliavin. Sous de bonnes hypothèses on obtient :

- l'hypothèse (H3) est vérifiée sous des conditions de régularité assez simples et

$$\langle \tilde{A}[\varphi], \chi \rangle = -\frac{1}{4} \mathbb{E} \left[\int_0^1 \xi_s^2 ds \varphi'(Y) \chi'(Y) \right].$$

- l'hypothèse (H1) suppose des conditions de régularité plus fines et

$$\begin{aligned} \langle \bar{A}[\varphi], \chi \rangle &= \frac{1}{2} \mathbb{E} \left[\int_0^1 \xi_s D_s D_s [\varphi'(Y) \chi(Y)] ds \right] + \frac{1}{2} \mathbb{E} \left[\int_0^1 \eta_s D_s [\varphi'(Y) \chi(Y)] ds \right] \\ &\quad - \frac{1}{4} \mathbb{E} \left[\int_0^1 \xi_s^2 (\varphi' \chi)'(Y) ds \right] - \frac{1}{4} \mathbb{E} \left[\int_0^1 \xi_s^2 \varphi'(Y) \chi'(Y) ds \right] \end{aligned}$$

où D désigne le gradient de Malliavin sur l'espace de Wiener. Alors \mathbb{A} est un opérateur du premier ordre.

• EDS et schéma d'Euler. L'erreur d'approximation de la solution d'une EDS par le schéma d'Euler a fait l'objet de nombreux travaux dont un des aboutissements est une forme de théorème central limite fonctionnel. Voir notamment Jean Jacod et Philip Protter "Asymptotic error distributions for the Euler method for stochastic differential equations" *Ann. Probab.* 26, 267-307, (1998) et les références indiquées. Ce théorème de limite centrale permet de simplifier l'étude de la limite

$$\lim_{n \rightarrow \infty} \alpha_n \mathbb{E} [(\varphi(Y_n) - \varphi(Y))^2]$$

mais l'existence de l'opérateur \tilde{A} , et donc la fermabilité de la forme, n'est établie à ce jour qu'en dimension 1. En revanche l'opérateur \bar{A} a été déterminé par Paul Malliavin et Anton Thalmaier "Numerical error for SDE: asymptotic expansion and hyperdistributions" *Note C. R. A. S.* sI, vol 336, n10, p851, 2003, il a une allure similaire à celui donné plus haut dans le cas de l'intégrale stochastique.

V. Statistique et erreurs.

Dans le calcul d'erreur par le langage des formes de Dirichlet *toutes les grandeurs erronées sont aléatoires*. On peut penser la loi a priori de la grandeur comme "le champ et l'usage" de l'appareil de mesure.

Comment déterminer expérimentalement l'opérateur Γ d'une structure d'erreur ?

Supposons que la structure d'erreur à identifier soit sur \mathbb{R}^d :

$$(\mathbb{R}^d, \mathcal{B}(\mathbb{R}^d), \mathbb{P}, \mathbb{D}, \Gamma)$$

Concrètement $(\mathbb{R}^d, \mathcal{B}(\mathbb{R}^d), \mathbb{P})$ est l'espace image d'une grandeur x que l'on mesure avec une certaine précision. \mathbb{P} est sa loi a priori.

Nous allons considérer que procéder à une mesure de la grandeur x , c'est estimer x statistiquement comme paramètre d'une famille de probabilités \mathbb{Q}_x

On sait alors que si on a un estimateur T disons sans biais de x , ($\mathbb{Q}_x[T] = x$) la précision sur x est limitée par l'inégalité de Cramer-Darmois-Fisher-Rao

$$\mathbb{Q}_x[(T - x)(T - x)^t] \geq [J(x)]^{-1}$$

avec égalité si T est efficace. Rappelons la définition de l'information de Fisher et l'inégalité de Cramer et al.

Soit $x \in \mathbb{R}^d$ et \mathbb{Q}_x une famille de probabilités sur un certain espace dominées par la probabilité \mathbb{Q}

$$\mathbb{Q}_x = L(x, \cdot) \mathbb{Q} \quad \text{avec } L(x, \cdot) \text{ régulière en } x$$

Alors pour toute variable aléatoire $Y \in L^2(\mathbb{Q})$ on a

$$\mathbb{E}_x[Y - \mathbb{E}_x(Y)]^2 \geq (\text{grad} \mathbb{E}_x(Y))^t [J(x)]^{-1} \text{grad} \mathbb{E}_x(Y)$$

où $J(x)$ est la matrice d'information de Fisher du modèle

$$J(x) = \left(\mathbb{E}_x \left[\frac{\partial \log L(x)}{\partial x_i} \frac{\partial \log L(x)}{\partial x_j} \right] \right)_{ij}$$

Soit I l'application identique de \mathbb{R}^d dans lui-même, $J(x)$ se comporte comme une information, $\Gamma[I](x)$ est une précision.

Nous posons l'identification

$$\boxed{\Gamma[I](x) = J^{-1}(x)}$$

Si nous choisissons ainsi $\Gamma[I](x) = J^{-1}(x)$, comme Γ vérifie le calcul fonctionnel, si $f : \mathbb{R}^d \rightarrow \mathbb{R}^p$ est de classe $\mathcal{C}^1 \cap \text{Lip}$, nous avons

$$\Gamma[f](x) = (\text{grad} f)^t \cdot \Gamma[I](x) \cdot \text{grad} f$$

Autrement dit l'opérateur Γ est déterminé, ainsi que pour toutes les structures images de notre structure d'erreur $(\mathbb{R}^d, \mathcal{B}(\mathbb{R}^d), \mathbb{P}, \mathbb{D}, \Gamma)$.

Donc la question se pose de savoir si nous obtenons une précision compatible avec ce calcul lorsque f est injective et que nous mesurons $f(x)$. La réponse est positive.

Sous les hypothèses d'un modèle statistique "régulier", si nous considérons $y = f(x)$ où $f \in \mathcal{C}^1 \cap \text{Lip}$ et injective, la structure d'erreur obtenue pour y est l'image de la structure obtenue pour x par l'application f . L'identification est aussi stable par produits en un sens naturel. Cf. N. B. et Chr. Chorro, "Structures d'erreur et estimation paramétrique", *Note C.R.A.S.*, Ser.I 338 (2004), 305-310.

Cela signifie que l'erreur sur x ainsi obtenue *ne dépend pas du paramétrage et a donc un sens physique*.

Quelques ouvrages liés à ces questions

AZEMA, J.; YOR, M. *Séminaire de Probabilités XVI, 1980/81 Supplément : Géométrie Différentielle Stochastique*, LNM 921, Springer, 1982.

BOULEAU N. *Error Calculus for Finance and Physics, the Language of Dirichlet Forms*, De Gruyter, 2003.

BOULEAU N., HIRSCH F. *Dirichlet Forms and Analysis on Wiener Space*, De Gruyter, (1991).

DELLACHERIE, CL., MEYER, P.-A., *Probabilités et Potentiel*, Hermann 1987.

EMERY, M. *Stochastic Calculus in Manifolds*, Springer, 1989.

FUKUSHIMA, M.; OSHIMA, Y.; TAKEDA, M. *Dirichlet Forms and Symmetric Markov Processes*, De Gruyter 1994.

HALL, P., HEYDE, C. C., *Martingale Limit Theory and its Applications*, Academic Press (1980).

JACOD, J., SHIRYAEV, A.N., *Limit Theorems for Stochastic Processes*, Springer 1987.

MA, Z.-M., RÖCKNER, M. *Introduction to the Theory of (Non-symmetric) Dirichlet Forms*, Springer 1992.

MALLIAVIN, P., THALMAIER, A. *Stochastic Calculus of Variations in Mathematical Finance*, Springer 2005.

NUALART, N. : *The Malliavin Calculus and Related Topics*, Springer, 1995.