Statistique des structures d'erreur

11 Juin 2003

PLAN DE L'EXPOSE

- Notion intuitive de structure d'erreur
- Extension : Le langage des formes de Dirichlet
- Un exemple en finance
- Lien avec les statistiques (Identification des structures d'erreur)

Notion intuitive de structure d'erreur

On considère une grandeur C (par exemple la concentration d'un polluant dans une rivière) qui peut être connue grâce à des mesures dont le résultat est entaché d'une erreur que l'on notera $\triangle C$. Comme souvent en physique ces grandeurs mal connues seront supposées aléatoires et en général corrélées.

Approche probabiliste :

Il faut connaitre la loi du couple $(C, \triangle C)$. Le phénomène de propagation des erreurs se traduit par le calcul de lois image.

Approche mixte :

$\mathbf{H}_1$: Nous supposons que la quantité $var[\triangle C \mid C]$ est connue et que $\mathbb{E}[\triangle C \mid C] = 0$.

$\mathbf{H}_2$: On considère les erreurs petites ce qui nous permettra d'utiliser les régles de simplification d'usage. Ainsi $\triangle C = \varepsilon Y$ où Y est une variable aléatoire bornée et où ε est un paramètre de calibration des erreurs.

Si f et g sont de classe C^3 avec des dérivées bornées, on a alors les formules de propagation suivantes :

1)

$$var[\triangle f(C) \mid C] = f'^2(C) var[\triangle C \mid C] + \varepsilon^3\ 0(1)$$

$$\mathbb{E}[\triangle f(C) \mid C] = \tfrac{1}{2} f''(C) var[\triangle f(C) \mid C] + \varepsilon^3\ 0(1).$$

2)

$$var[\triangle(h \circ f(C)) \mid C] = h'^2(f(C)) var[\triangle f(C) \mid C] + \varepsilon^3\ 0(1)$$

$$\begin{aligned}\mathbb{E}[\triangle(h \circ f(C)) \mid C] = h'(f(C))\mathbb{E}[\triangle f(C) \mid C] + \\ \frac{1}{2} h''(f(C)) var[\triangle f(C) \mid C] + \varepsilon^3\ 0(1).\end{aligned}$$

Le calcul de variance est un calcul du premier ordre ne faisant intervenir que les variances. Le calcul de biais est un calcul du second ordre faisant intervenir variances et biais.

Sur l'espace $(\mathbb{R}, Bor(\mathbb{R}), loi\ de\ C)$, on peut alors introduire l'opérateur Γ^C, appelé opérateur d'erreur quadratique, défini par

$$\Gamma^C[f](x) = \underset{\varepsilon \to 0}{Lim} \frac{var[\triangle f(C) \mid C = x]}{\varepsilon^2}.$$

Γ^C se polarise en un opérateur symétrique, positif, bilinéaire vérifiant le calcul fonctionnel suivant : si $F \in C^2(\mathbb{R}^2)$ avec des dérivées partielles bornées

$$\begin{aligned}\Gamma^C[F(f,g)] = F_1'^2(f,g)\Gamma^C[f] + F_2'^2(f,g)\Gamma^C[g] \\ +2F_1'(f,g)F_2'(f,g)\Gamma^C[f,g].\end{aligned}$$

Le terme $(\mathbb{R}, Bor(\mathbb{R}), loi\ de\ C, \Gamma^C)$ est appelé une structure d'erreur lorsque Γ^C est symétrique, positif, bilinéaire et vérifie le calcul fonctionnel précedent.

Il s'agit donc d'un espace de probabilité enrichi mais qui garde toute **la souplesse technique :**

- Transport simple via les fonctions régulières

- Définition naturelle du produit de structures d'erreur.

Le langage des formes de Dirichlet

Dorénavant une structure d'erreur sera un terme $(W, \mathcal{W}, m, \mathbb{D}, \Gamma)$ où $(W, \mathcal{W}, m)$ est un espace de probabilité et Γ un opérateur bilinéaire, symétrique, positif de domaine $\mathbb{D}$ dense dans $L^2(m)$, à valeurs dans $L^1(m)$ et vérifiant :

1) Le calcul fonctionnel de classe $C^1 \cap Lip$: soient $U = (U_1, \ldots, U_n) \in \mathbb{D}^n$ et $F \in C^1(\mathbb{R}^n) \cap Lip$ alors $F(U_1, \ldots, U_n) \in \mathbb{D}$ et

$$\Gamma[F(U_1, \ldots, U_n)] = \sum_{i,j=1}^{n} F_i'(U) F_j'(U) \Gamma[U_i, U_j]$$

2) $1 \in \mathbb{D}$, $\Gamma[1] = 0$
3) La forme bilinéaire définie sur $\mathbb{D} \times \mathbb{D}$ par $\mathcal{E}[F, G] = \frac{1}{2} \int \Gamma[F, G] dm$ est fermée au sens où $\mathbb{D}$ muni de la norme du graphe notée

$$\| \,.\, \|_{\mathcal{E}} = (\| \,.\, \|^2_{L^2(m)} + \mathcal{E}[.])^{\frac{1}{2}}$$

est complet.

On notera toujours $\Gamma[F]$ pour $\Gamma[F, F]$ et $\mathcal{E}[F]$ pour $\mathcal{E}[F, F]$.

AVANTAGES

• Extension naturelle de l'approche heuristique (propriété 1)

• Souplesse technique (Image, produits)

• La propriété (3) assure un calcul fonctionnel sur les fonctions Lip avec, lorsque F est contractante,

$$\Gamma[F(U)] \leq \Gamma[U].$$

• Possibilité d'un calcul de biais

Un exemple en finance

On note r le taux des placements sans risques que l'on suppose constant. On considère un actif financier suivant le modèle de Black-Scholes :

$$dS_t = S_t(rdt + \sigma dB_t)$$

de sorte que le cours actualisé soit une martingale. Si le payoff d'une option européenne d'échéance T est donné par $f(S_T)$ on sait que la valeur V_t de l'option à l'instant t est

$$V_t = F(t, S_t, \sigma, r)$$

où

$$F(t, x, \sigma, r) = e^{-r(T-t)} \int f\left(xe^{(r-\frac{\sigma^2}{2})(T-t)+\sigma y\sqrt{T-t}}\right) \frac{e^{\frac{-y^2}{2}}}{\sqrt{2\pi}} dy.$$

Si f est à croissance linéaire, F est très régulière et l'on peut définir les grecs :

$$delta_t = \frac{\partial F}{\partial x}(t, S_t, \sigma, r)$$
$$gamma_t = \frac{\partial^2 F}{\partial x^2}(t, S_t, \sigma, r)$$
$$rho_t = \frac{\partial F}{\partial r}(t, S_t, \sigma, r).$$

On suppose la **volatilité erronée** et l'on modélise cette erreur par la structure d'erreur $(\mathbb{R}_+^*, Bor(R_+^*), m, \mathbb{D}, \Gamma^\sigma)$ vérifiant $Id \in \mathbb{D}^\sigma$ où $\Gamma^\sigma[Id]$ représente la variance conditionnelle de l'erreur sur σ.

On a alors les formules de propagation suivantes :

$$\Gamma^\sigma[S_t] = (B_t - \sigma t)^2 S_t^2 \Gamma^\sigma[Id]$$

$$\Gamma^\sigma[V_t] = (S_t(B_t - \sigma t)delta_t + (T - t)\sigma S_t^2 gamma_t)^2 \Gamma^\sigma[Id]$$

$$\Gamma^\sigma[V_0] = T^2\sigma^2 S_0^4 gamma_0^2 \Gamma^\sigma[Id]$$

Question : Comment déterminer $\Gamma^\sigma[Id]$?

-Faire un choix a priori.

-Mettre en place une determination statistique à partir d'expériences répétées.

Statistique des structures d'erreur

1 Problématique

On considère une variable aléatoire $\Theta : (\Omega, \mathcal{A}, \mathbb{P}) \to \mathbb{R}^d$ qui représente une grandeur deterministe mal connue et nous tentons de trouver une structure d'erreur

$$S^\Theta = (\mathbb{R}^d, Bor(\mathbb{R}^d), \Theta_*\mathbb{P}, \mathbb{D}^\Theta, \Gamma^\Theta)$$

qui soit telle que l'opérateur Γ^Θ exprime la précision de la connaissance que l'on a sur Θ avec les moyens statistiques employés.

Notre approche consiste dans des conditions précisées ci-dessous à identifier Γ^Θ à l'inverse de la matrice d'information de Fisher qui, pour les modèles statistiques paramétriques réguliers, est une mesure de la connaissance disponible.

2 Borne de Cramer-Rao et identification fondamentale

2.1 Hypothèses du modèle régulier

Pour connaitre Θ on peut procéder par estimation pour cela on se donne un modèle paramétrique régulier :

a) $\Theta(\Omega)$ est un ouvert de $\mathbb{R}^d$.

b) Il existe une mesure σ-finie μ telle que, pour tout θ, P_θ soit absolument continue par rapport à μ de densité $f(.,\theta)$ strictement positive.

c) La fonction $\theta \to f(x,\theta)$ est continue en tout point et ce pour μ presque tout x.

d) On note $g(x,\theta) = \sqrt{f(x,\theta)}$. On suppose l'existence d'une fonction ϕ : $\Omega' \times \Theta(\Omega) \to \mathbb{R}^d$ telle que $\forall\, \theta \in \Theta(\Omega)$,

$$\int \parallel \phi(x,\theta) \parallel^2 \, d\mu(x) < \infty$$

et

$$\int |g(x,\theta+h) - g(x,\theta) - (\phi(x,\theta),h)|^2 d\mu(x) = o(\parallel h \parallel^2).$$

On peut alors définir $J(\theta) = 4\int \phi(x,\theta)\phi(x,\theta)^t d\mu(x)$ qui est la quantité d'information de Fisher de notre modèle.

e) $\theta \to \phi(.,\theta)$ est continue dans $L^2(\mu)$.

f) Le modèle est identifiable : $\theta \to P_\theta$ est injective.

2.2 Inégalité de Cramer-Rao (I.C.R)

On se donne une observation X c'est à dire une variable aléatoire telle que la loi de X sachant $\Theta = \theta$ est P_θ.

Théorème : Soit $\psi : \mathbb{R}^d \to \mathbb{R}^m$ une fonction différentiable. Si $(P_\theta)_{\theta \in \Theta(\Omega)}$ vérifie les hypothèses du modèle régulier, si J est inversible et si $T(X)$ est un estimateur sans biais de $\psi(\Theta)$ tel que la fonction $\theta \to \int T^2(x) f(x,\theta) d\mu(x)$ est localement bornée alors $\forall \theta \in \Theta(\Omega)$,

$$\underline{\mathbb{E}[(T(X) - \psi(\Theta))(T(X) - \psi(\Theta))^t \mid \Theta = \theta] \geq \psi'(\theta) J^{-1}(\theta) \psi'(\theta)^t}$$

où $\geq$ est ici la relation d'ordre au sens des matrices symétriques.

2.3 Identification fondamentale (I.F)

Lorsque les composantes de l'identité sont dans $\mathbb{D}^\Theta$, pour $F \in C^1_K(\mathbb{R}^d)$ le calcul fonctionnel donne

$$\Gamma^\Theta[F](\theta) = (\nabla_\theta F)^t \underline{\underline{\Gamma}}^\Theta[Id](\nabla_\theta F) \quad \forall \theta \in \Theta(\Omega)$$

où la matrice $\underline{\underline{\Gamma}}^\Theta[Id](\theta)$ représente la matrice de variance-covariance des erreurs sur Θ sachant $\Theta = \theta$.

On pose alors

$$\boxed{\underline{\underline{\Gamma}}^\Theta[Id] = J^{-1} \quad \textbf{(I.F)}}$$

De même qu'en statistiques on cherche à identifier un espace de probabilité dont la mesure est supposée *a priori* σ-additive, on cherche ici à déterminer à partir de données statistiques une structure d'erreur dans laquelle la forme quadratique a la propriété d'être fermée. Par l'identification fondamentale cette structure s'écrit

$$S^{\Theta} = (\mathbb{R}^d, Bor(\mathbb{R}^d), \Theta_* \mathbb{P}, \mathbb{D}^{\Theta}, \Gamma^{\Theta})$$

où $C^1_K \subset \mathbb{D}^{\Theta}$ et $\forall F \in C^1_K$,

$$\Gamma^{\Theta}[F] = F' J^{-1} (F')^t.$$

Le domaine $\mathbb{D}^{\Theta}$ n'étant alors pas défini de manière unique, nous le prenons minimal au sens de l'inclusion de sorte que
$\underline{C^1_K \text{ est dense dans } \mathbb{D}^{\Theta} \text{ pour } \| \, . \, \|_{\mathcal{E}^{\Theta}}}$.

Retour à Black-Scholes

Comme dans le cadre du modèle de Black-Scholes les variables aléatoires

$$Y_1^N = Log\left(\frac{S_{\frac{T}{N}}}{S_0}\right), \ldots, Y_N^N = Log\left(\frac{S_T}{S_{\frac{(N-1)T}{N}}}\right)$$

sont i.i.d de loi

$$\mathcal{N}((r - \frac{\sigma^2}{2})\frac{T}{N}, \sigma^2\frac{T}{N}),$$

on a un modèle paramétrique naturel d'estimation de la volatilité qui conduit à

$$\Gamma^\sigma[Id](\sigma) = \frac{\sigma^2}{2N + \sigma^2 T}.$$

Ainsi lorsque $N \to +\infty$

$$\frac{\sqrt{\Gamma^\sigma[Id](\sigma)}}{\sigma} = 0.$$

Nous allons tester la robustesse de cette identification vis-à-vis des changements de variables réguliers et des opérations de produit.

2.4 Changements de variables

Hypothèses : On suppose que les composantes de l'identité sont dans $\mathbb{D}^\theta$.

Définition : Soit $\psi : \mathbb{R}^d \to \mathbb{R}^d$ injective de classe $\mathcal{C}^1(\mathbb{R}^d) \cap Lip$. On dira que ce changement de variables est régulier s'il vérifie $\psi'(x)$ inversible pour tout x.

2.4.1 Structure image

Nous sommes capable de prendre sous des conditions peu restrictives l'image d'une structure d'erreur par une variable aléatoire :

Définition : Soit $S = (\Omega, \mathcal{A}, P, D, \Gamma)$ une structure d'erreur et $X : \Omega \to \mathbb{R}^d$ une variable aléatoire dans D^d. On définit

$$\widetilde{D_X} = \{f \in L^2(X_*P) \mid f(X) \in D\}$$

et pour $f \in \widetilde{D_X}$,

$$\widetilde{\Gamma_X}[f](x) = \mathbb{E}[\Gamma[f(X)] \mid X = x].$$

Ainsi

$$\psi_* S = (\mathbb{R}^d, Bor(\mathbb{R}^d), X_*P, D_X, \Gamma_X)$$

où X_*P est la loi de X, D_X est la fermeture de $C^1_K(\mathbb{R}^d)$ dans $(\widetilde{D_X}, \| \,.\, \|_H)$, $\Gamma_X = \widetilde{\Gamma_X}|_{D_X}$ est une structure d'erreur que l'on appelle structure d'erreur image de S par X.

Image de S^Θ par ψ

Pour $F \in \mathcal{C}^1_K(\mathbb{R}^d)$ on a $\forall a \in Im(\psi)$

$$\Gamma^\Theta_\psi[F](a) = \mathbb{E}_{\Theta_* \mathbb{P}}[\Gamma^\Theta[F(\psi)] \mid \psi = a]$$

et donc

$$\Gamma^\Theta_\psi[F](a) = \mathbb{E}_{\Theta_* \mathbb{P}}[\nabla(F(\psi))^t \, J^{-1} \, \nabla(F(\psi)) \mid \psi = a].$$

ψ étant injective, $\Gamma^\Theta_\psi[F](a)$ vaut

$$(\nabla_a F)^t[\psi'(\psi^{-1}(a))] \, [J(\psi^{-1}(a))]^{-1} \, [\psi'(\psi^{-1}(a))]^t(\nabla_a F).$$

2.4.2 Structure d'erreur associée à $\psi(\Theta)$

On estime $\psi(\Theta)$ à partir du modèle dominé $(P_{\psi^{-1}(a)})$.
$\Gamma^{\psi(\Theta)}$ est défini sur $C^1_K(\mathbb{R}^d)$ par

$$\Gamma^{\psi(\Theta)}[F](a) = (\nabla_a F)^t(J^{\psi(\Theta)}(a))^{-1}(\nabla_a F) \quad \forall a \in \psi(\Theta)(\Omega)$$

où $J^{\psi(\Theta)}$ est la matrice d'information de Fisher du modèle régulier $(P_{\psi^{-1}(a),\ a \in \psi(\Theta)(\Omega)})$. Comme

$$J^{\psi(\Theta)}(a) = [\psi'(\psi^{-1}(a))^{-1}]^t \, [J(\psi^{-1}(a))] \, [\psi'(\psi^{-1}(a))^{-1}]$$

pour $F \in C^1_K(\mathbb{R}^d)$ on a le résultat suivant :

proposition :

$$\boxed{\psi_* S^\Theta = S^{\psi(\Theta)}.}$$

2.5 Structure produit

Définition : Soient $S_i = (\Omega_i, \mathcal{A}_i, P_i, D_i, \Gamma_i)$ (i=1,2) deux structures d'erreur, la structure d'erreur produit notée $S_1 \otimes S_2$ est définie comme la structure $(\Omega_1 \times \Omega_2, \mathcal{A}_1 \times \mathcal{A}_2, P_1 \times P_2, D, \Gamma)$ où

$$
\begin{aligned}
D = \{f \in L^2(P_1 \times P_2) |\, & pour\ \ P_2\ \ presque\ \ tout\ \ y\ \ f(.,y) \in D_1 \\
& pour\ \ P_1\ \ presque\ \ tout\ \ x\ \ f(x,.) \in D_2 \\
& \int \Gamma[f](x,y) dP_1(x) dP_2(y) < +\infty\}
\end{aligned}
$$

et

$$\Gamma[f](x,y) = \Gamma_1[f(.,y)](x) + \Gamma_2[f(x,.)](y).$$

On se place alors dans la situation où l'on a un paramètre $\Theta = (\Theta_1, \Theta_2) : (\Omega, \mathcal{A}, \mathbb{P}) \to \mathbb{R}^2$ à évaluer. On suppose que les composantes sont indépendantes et on les évalue à l'aide d'expériences régulières indépendantes. On a alors le résultat suivant :

Proposition : Si ψ_1 et ψ_2 sont injectives de classe $C^1 \cap Lip$ alors

1)

$$\boxed{S^{\Theta} = S^{\Theta_1} \otimes S^{\Theta_2}}$$

2)

$$\boxed{(\psi_1, \psi_2)_* S^{(\Theta_1, \Theta_2)} = \psi_{1\,*} S^{\Theta_1} \otimes \psi_{2\,*} S^{\Theta_2}.}$$

Conclusion

a) Les propriétés analytiques de l'information de Fisher assurent une stabilité algébrique de l'identification fondamentale.

b) L'identification fondamentale est intimement liée à la théorie des statistiques asymptotiques et fournit un cadre de travail souple pour étudier l'estimation d'une grandeur $\psi(\Theta)$ lorsque ψ n'est pas injective.