MICHEL ÉMERY

En cherchant une caractérisation variationnelle des martingales

Séminaire de probabilités (Strasbourg), tome 22 (1988), p. 147-154
<http://www.numdam.org/item?id=SPS_1988__22__147_0>

Université de Strasbourg
Séminaire de Probabilités

EN CHERCHANT UNE CARACTÉRISATION VARIATIONNELLE DES MARTINGALES

par M. Emery[1]

> Nous n'aimons pas le désordre, l'aléatoire, le hasard. Les comportements superstitieux, les martingales, la magie témoignent aussi de notre besoin d'imaginer qu'il existe *un ordre caché* derrière le hasard et le désordre.
>
> J.N. KAPFERER, *Rumeurs.*

Nous ne nous occuperons ici que de processus réels, bien que cet exposé trouve son origine dans une question de géométrie différentielle stochastique : Dans une variété, la même structure (une connexion) permet de définir une classe de courbes, les géodésiques, et une classe de processus, les martingales, qui obéissent à des énoncés souvent analogues. Il est tentant de prolonger cette analogie et de se demander, dans le cas riemannien, si les propriétés variationnelles des géodésiques (elles minimisent l'énergie et la longueur) auraient pour pendant des caractérisations variationnelles des martingales. C'est faute de savoir traiter cette question dans une variété riemannienne générale que nous nous placerons dans le cadre usuel des processus à valeurs réelles.

On se fixe un espace filtré $(\Omega, \mathcal{F}, \mathbb{P}, (\mathcal{F}_t)_{t\geq 0})$ vérifiant les conditions habituelles.

Problème 1. — Soit μ une probabilité sur la demi-droite $[0,\infty[$; *parmi les processus X mesurables adaptés et tels que l'intégrale $\int_0^\infty X_s\,\mu(ds)$ soit une variable aléatoire donnée, quel est celui qui minimise la fonctionnelle* $\mathbb{E}\left[\int_0^\infty X_s^2\,\mu(ds)\right]$?

Nous poserons $\alpha = \sup(\text{support } \mu) = \inf\{t : \mu(]t,\infty[) = 0\}$, de sorte que le plus petit intervalle qui contienne 0 et porte μ est $I = [0,\alpha[$ si $\mu(\{\alpha\}) = 0$ et $I = [0,\alpha]$ si $\mu(\{\alpha\}) > 0$ (dans ce dernier cas, α est fini). Nous appellerons F la fonction de répartition de μ, donnée par $F(t) = \mu([0,t])$, et f la fonction

[1] Je remercie S.Q. Song et C. Stricker qui ont considérablement simplifié cet exposé.

$f(t) = \mu([t,\infty[) = 1 - F(t-)$. Enfin nous désignerons par $\mathcal{H}$ l'espace de Hilbert constitué des (classes de) processus X mesurables, adaptés et tels que $\|X\|^2_{\mathcal{H}} = \mathbb{E}\left[\int_0^\infty X_s^2\,\mu(ds)\right]$ soit fini; c'est simplement le sous-espace fermé de $L^2(\Omega\times\mathbb{R}_+, \mathcal{F}\otimes\mathcal{B}(\mathbb{R}_+), \mathbb{P}\otimes\mu)$ formé des processus ayant un représentant adapté.

Pour $X \in \mathcal{H}$, on peut écrire

$$\mathbb{E}\Big[\Big(\int_0^\infty X_s\,\mu(ds)\Big)^2\Big] \leq \mathbb{E}\Big[\int_0^\infty X_s^2\,\mu(ds)\Big] = \|X\|^2_{\mathcal{H}}\,;$$

ainsi l'opérateur $\Phi : X \longmapsto \int_0^\infty X_s\,\mu(ds)$ est continu de $\mathcal{H}$ dans L^2. Son noyau $\mathcal{H}_0 = \operatorname{Ker}\Phi = \left\{X \in \mathcal{H} : \int_0^\infty X_s\,\mu(ds) = 0\right\}$ est donc un sous-espace fermé.

Les solutions du problème 1 sont faciles à décrire à l'aide de $\mathcal{H}_0$: si l'on se donne $Z \in L^2$, l'ensemble $\Phi^{-1}(Z)$ est un sous-espace affine fermé; s'il n'est pas vide, il contient un unique élément de norme minimale, la projection orthogonale de l'origine, que l'on peut caractériser comme le seul élément de $\Phi^{-1}(Z)$ orthogonal à $\mathcal{H}_0$. Ainsi, *un processus $X \in \mathcal{H}$ résout le problème 1 si et seulement si il est orthogonal à $\mathcal{H}_0$*. Nous allons voir que ces processus sont exactement les martingales.

Nous noterons $\mathcal{M}$ l'ensemble des martingales $(M_t)_{t\in I}$ qui sont dans $\mathcal{H}$ (puisque μ néglige I^c, la donnée de M sur I suffit à définir M comme élément de $\mathcal{H}$). C'est un sous-espace vectoriel fermé de $\mathcal{H}$: Si une suite $(M^n)_{n\in\mathbb{N}}$ de $\mathcal{M}$ converge dans $\mathcal{H}$ vers une limite X, $\mathbb{E}\left[\int_0^\infty (M_s^n - X_s)^2\,\mu(ds)\right]$ tend vers zéro, donc, après extraction d'une sous-suite, M_s^n tend dans L^2 vers X_s pour tout $s \in J$, où $\mu(J) = 1$. Donc $(X_s)_{s\in J}$ est une martingale, et, comme tout élément de $I \setminus J$ est majoré par un élément de J, X est dans $\mathcal{M}$.

LEMME. — *Dans $\mathcal{H}$, le supplémentaire orthogonal de $\mathcal{H}_0$ est $\mathcal{M}$.*

Démonstration. — Puisque $\mathcal{M}$ est fermé, il suffit d'établir que $\mathcal{M}^\perp = \mathcal{H}_0$. Nous allons pour cela vérifier que l'espace $\mathcal{M}^2$ des martingales de $\mathcal{M}$ qui sont bornées dans L^2 est dense dans $\mathcal{M}$, puis que ${\mathcal{M}^2}^\perp = \mathcal{H}_0$.

D'abord la densité de $\mathcal{M}^2$. Soit $M \in \mathcal{M}$. Dans le cas où I est un compact $[0,\alpha]$, M s'écrit $M_t = \mathbb{E}[M_\alpha|\mathcal{F}_t]$ avec $M_\alpha \in L^2$, donc $\mathcal{M} = \mathcal{M}^2$. Dans le cas contraire, pour tout $t \in I$, la martingale arrêtée[1] $M^{|t}$ est bornée dans L^2, et il ne reste qu'à voir que, quand t croît vers α, $M^{|t}$ tend vers M dans $\mathcal{H}$; or ceci résulte de ce que

$$\begin{aligned}\|M^{|t} - M\|^2_{\mathcal{H}} &= \mathbb{E}\Big[\int_{]t,\infty[}(M_s - M_t)^2\,\mu(ds)\Big] = \int_{]t,\alpha[}\|M_s - M_t\|_2^2\,\mu(ds)\\ &\leq \int_{]t,\alpha[}\left(\|M_s\|_2 + \|M_t\|_2\right)^2\mu(ds) \leq \int_{]t,\alpha[}\left(2\,\|M_s\|_2\right)^2\mu(ds)\end{aligned}$$

tend vers zéro par convergence dominée quand t croît vers α.

---

[1] Nous empruntons à G. Letta l'excellente notation $X^{|T}$ pour désigner le processus X arrêté au temps T.

Montrons maintenant que ${\mathcal{M}^2}^\perp = \mathcal{H}_0$. Soient X dans $\mathcal{H}$ et M dans $\mathcal{M}^2$; il y a une v.a. V dans L^2 telle que $M_t = \mathbb{E}[V|\mathcal{F}_t]$ pour tout $t \in I$ (et *a fortiori* pour μ-presque tout $t \geq 0$). Le produit scalaire de M et X vaut

$$\begin{aligned}
\langle M, X\rangle_{\mathcal{H}} &= \mathbb{E}\Big[\int_0^\infty M_s X_s\, \mu(ds)\Big] \\
&= \int_0^\infty \mathbb{E}[M_s X_s]\, \mu(ds) \quad (\text{car } \textstyle\int \mathbb{E}[|M_s||X_s|]\mu(ds) \leq \|M\|\|X\| < \infty) \\
&= \int_0^\infty \mathbb{E}[V X_s]\, \mu(ds) \quad (\text{car } V \text{ et } \mu\text{-presque tout } X_s \text{ sont dans } L^2) \\
&= \mathbb{E}\Big[\int_0^\infty V X_s\, \mu(ds)\Big] \quad (\text{car } \mathbb{E} \textstyle\int |V||X_s|\mu(ds) \leq \|V\|_2 \|\Phi(X)\|_2 < \infty) \\
&= \mathbb{E}\Big[V \int_0^\infty X_s\, \mu(ds)\Big] = \mathbb{E}[V\, \Phi(X)]\ .
\end{aligned}$$

Ceci montre que si X est dans $\mathcal{H}_0$, il est orthogonal à $\mathcal{M}^2$. Réciproquement, un X de $\mathcal{H}$ qui est orthogonal à toute $M \in \mathcal{M}^2$ doit, par le même calcul, vérifier $\mathbb{E}[V \int_0^\infty X_s\, \mu(ds)] = 0$ pour tout $V \in L^2$, donc $\Phi(X) = 0$ et X est dans $\mathcal{H}_0$. ∎

Nous sommes maintenant en mesure de donner la solution du problème 1.

PROPOSITION. — *Pour $Z \in L^2(\Omega, \bigvee_{t\in I} \mathcal{F}_t, \mathbb{P})$, notons Z_t la martingale $\mathbb{E}[Z|\mathcal{F}_t]$, ϕ la fonction $\phi(t) = \mathbb{E}[Z_t^2]$ et $(X_t)_{t\in I}$ la martingale $\displaystyle\int_0^t \frac{dZ_s}{f(s)}$ (y compris le saut en 0 : $X_0 = Z_0$). Les conditions suivantes sont équivalentes :*

(i) *la v.a. Z est dans l'image $\Phi(\mathcal{H})$;*

(ii) *la martingale X est dans $\mathcal{H}$ (donc dans $\mathcal{M}$) ;*

(iii) *l'intégrale $\displaystyle\int_I \frac{d\phi(s)}{f(s)}$ est finie.*

Lorsqu'elles sont réalisées, $\Phi(X) = Z$ et X est l'élément de $\Phi^{-1}(Z)$ qui a la plus petite norme.

Réciproquement, toute martingale $(M_t)_{t\in I}$ de $\mathcal{M}$ est obtenue de cette manière à partir de la v.a. $Z = \Phi(M)$.

<u>Démonstration</u>. — Remarquons avant tout que l'intégrale qui définit X a un sens puisque f est minorée sur tout compact de I.

L'équivalence entre (ii) et (iii) résulte de

$$\mathbb{E}[X_t^2] = \mathbb{E}[[X,X]_t] = \mathbb{E}\Big[\int_0^t \frac{d[Z,Z]_s}{f^2(s)}\Big] = \int_0^t \frac{d\phi(s)}{f^2(s)}\ ;$$

$$\mathbb{E}\Big[\int_0^\infty X_s^2\, \mu(ds)\Big] = \int_I \mu(dt) \int_0^t \frac{d\phi(s)}{f^2(s)} = \int_I \frac{d\phi(s)}{f(s)}\ .$$

Supposons maintenant (i) : Z s'écrit $\Phi(Y)$ pour un Y de $\mathcal{H}$. En notant M la projection orthogonale de Y sur $\mathcal{M}$, on a aussi $Z = \Phi(M) = \int_0^\infty M_s\, \mu(ds)$; pour tout t dans I, on peut écrire (l'intégrale commute avec l'espérance conditionnelle car $\mathbb{E}\left[\int_0^\infty |M_s|\,\mu(ds)\right]$ est fini)

$$\begin{aligned} Z_t = \mathbb{E}[Z|\mathcal{F}_t] &= \mathbb{E}\Big[\int_0^\infty M_s\,\mu(ds)\;\Big|\;\mathcal{F}_t\Big] = \int_0^\infty \mathbb{E}[M_s|\mathcal{F}_t]\,dF(s) \\ &= \int_{[0,t]} M_s\,dF(s) + \big(1-F(t)\big)M_t = \int_{[0,t]} \big(1-F(s-)\big)\,dM_s\;, \end{aligned}$$

ou encore $Z_t = \int_0^t f(s)\,dM_s$. Comme f est minorée sur $[0,t]$, il en résulte que $M_t = \int_0^t f(s)^{-1} dZ_s = X_t$, et $X = M$ est dans $\mathcal{H}$. Nous avons établi du même coup que $Z = \Phi(X)$; et, par le lemme, X est la projection de l'origine sur $\Phi^{-1}(Z)$, c'est-à-dire l'élément de norme minimale.

Montrons que (ii) entraîne (i). Supposant la martingale X dans $\mathcal{H}$, et écrivant comme ci-dessus

$$Z_t = \int_0^t f(s)\,dX_s = \int_{[0,t]} X_s\,\mu(ds) + \big(1-F(t)\big)X_t\;,$$

nous devons établir que $Z = \int_0^\infty X_s\,\mu(ds)$. Si α est dans I, $F(\alpha) = 1$ d'où $Z_\alpha = \int_{[0,\alpha]} X_s\,\mu(ds)$ et c'est terminé. Si, au contraire, $I = [0,\alpha[$, puisque, quand t croît vers α, l'intégrale $\int_{[0,t]} X_s\,\mu(ds)$ a une limite p.s. et dans L^2, nous devons vérifier que $\big(1-F(t)\big)X_t$ tend vers zéro dans L^2 (car alors la limite sera $\lim_{t\uparrow\alpha} Z_t$, c'est-à-dire Z). En posant, pour t dans I, $\psi(t) = \mathbb{E}[X_t^2]$, ψ est croissante sur l'intervalle $I = [0,\alpha[$ car X est une martingale, et vérifie $\int_{[0,\alpha[} \psi(s)\,\mu(ds) < \infty$ car X est dans $\mathcal{H}$. Il en résulte que

$$\big(1-F(t)\big)\psi(t) = \psi(t)\,\mu(]t,\alpha[) \le \int_{]t,\alpha[} \psi(s)\,\mu(ds)$$

tend vers zéro quand t croît vers α et, *a fortiori*,

$$\big\|\big(1-F(t)\big)X_t\big\|_2^2 = \big(1-F(t)\big)^2\psi(t)$$

tend aussi vers zéro. Donc $\big(1-F(t)\big)X_t$ tend vers zéro dans L^2.

Réciproquement, si une martingale $(M_t)_{t\in I}$ est dans $\mathcal{H}$, en posant $Z = \Phi(M)$, M est l'unique martingale de $\Phi^{-1}(Z)$ (unique car par le lemme l'espace des martingales est orthogonal à $\Phi^{-1}(Z)$) ; elle est donc nécessairement égale à $X = \int f^{-1}\,dZ$. ∎

REMARQUES. — a) Il peut paraître curieux que lorsque μ varie, la solution du problème reste toujours la même (les martingales). J.B. Walsh nous a fait observer que ça n'est, au fond, pas si surprenant : Le lemme dit que X est une martingale si et seulement si l'on a $\mathbb{E}\left[\int_0^\infty X_s Y_s\,\mu(ds)\right] = 0$ pour tout Y tel que $\int_0^\infty Y_s\,\mu(ds) = 0$; en prenant $dB_s = Y_s\mu(ds)$, la mesure μ n'intervient plus, et on a en réalité caractérisé les martingales comme les processus adaptés X tels que l'on ait $\mathbb{E}\left[\int_0^\infty X_s\, dB_s\right] = 0$ pour une classe suffisamment riche de processus B à variation bornée et vérifiant $B_\infty = 0$.

b) Par changement de temps, tout ceci s'étend au cas où la mesure déterministe μ est remplacée par une probabilité aléatoire, c'est-à-dire un processus croissant A tel que $A_\infty = 1$; mais bien entendu, les processus obtenus ne sont alors plus exactement les martingales, mais celles des martingales locales qui sont réduites par les temps d'arrêt correspondant aux temps constants par le changement de temps associé à A. En revanche, si l'on ne suppose pas $A_\infty = 1$ p.s., mais seulement $\mathbb{E}[A_\infty] = 1$, le résultat ne subsiste plus : si par exemple T est une variable aléatoire intégrable de loi équivalente à la mesure de Lebesgue sur $\mathbb{R}_+$, en substituant à μ le processus croissant $A_t = t \wedge T$, l'opérateur Φ devient injectif et tous les éléments de $\mathcal{H}$ sont donc des solutions du problème (bien que dans $\mathcal{H}$ les seules martingales soient les constantes).

c) Cette caractérisation des martingales s'étend au cas où la probabilité μ est remplacée par une mesure positive σ-finie. Dans cette situation, on peut encore définir le sous-espace $\mathcal{H}_0$ de $\mathcal{H}$ comme l'ensemble des X tels que presque sûrement on ait $\int_0^\infty |X_s|\, \mu(ds) < \infty$ et $\int_0^\infty X_s\, \mu(ds) = 0$, mais il n'est plus nécessairement fermé; il reste vrai que les solutions du problème 1 sont exactement les processus orthogonaux à $\mathcal{H}_0$; nous allons voir que ce sont les martingales. Désignons par $r > 0$ une fonction bornée sur $\mathbb{R}_+$ telle que la mesure $r.\mu$ soit une probabilité, et par $\mathcal{K}$, $\mathcal{K}_0$ les espaces $\mathcal{H}$ et $\mathcal{H}_0$ formés à l'aide de cette probabilité. Comme r est bornée, tout élément de $\mathcal{H}$ est aussi dans $\mathcal{K}$. Soit X dans $\mathcal{H}_0^\perp$; pour tout Y dans $\mathcal{K}_0$, le processus $r(t)Y_t$ est dans $\mathcal{H}_0$, et l'on a $\langle X, Y\rangle_{\mathcal{K}} = \langle X, rY\rangle_{\mathcal{H}} = 0$; donc X est orthogonal dans $\mathcal{K}$ à $\mathcal{K}_0$, et par le lemme $(X_t)_{t\in I}$ est une martingale. Réciproquement, si $(X_t)_{t\in I}$ est une martingale, la même démonstration que dans le lemme montre que X est orthogonal à $\mathcal{H}_0$ (si $\alpha \in I$, écrire $X_t = \mathbb{E}[X_\alpha|\mathcal{F}_t]$; si $\alpha \notin I$, procéder par approximation). Il faut toutefois observer que l'espace de processus ainsi obtenu (les martingales $(X_t)_{t\in I}$ qui sont dans $\mathcal{H}$) peut être très pauvre : puisque la fonction $\psi(t) = \mathbb{E}[X_t^2]$ est croissante sur I et intégrable pour μ, elle est nulle pour tout t tel que $\mu([t,\infty[) = \infty$; en particulier si μ n'est pas bornée on a toujours $X_0 = 0$, et si μ donne une masse infinie à tout voisinage de l'infini X doit être identiquement nul. Ceci se voit aussi sur la formule explicitant X en fonction de son intégrale $Z = \Phi(X)$, formule qui s'écrit comme précédemment

$$X_t = \int_0^t \frac{1}{\mu([s,\infty[)}\, dZ_s \;;$$

avec la convention évidente $1/\infty = 0$, on retrouve le fait que la martingale X reste nulle tant que $\mu([t,\infty[)$ est infini.

Problème 1^{bis}. — Soient ν et π deux mesures σ-finies sur la demi-droite $[0,\infty[$ équivalentes (et finies sur les mêmes ensembles) ; *parmi les processus X mesurables adaptés et tels que l'intégrale $\int_0^\infty X_s\, \nu(ds)$ soit une variable aléatoire donnée, quel est celui qui minimise la fonctionnelle* $\mathbb{E}\left[\int_0^\infty X_s^2\, \pi(ds)\right]$?

Choisissons une densité $\rho > 0$ de ν par rapport à π, et posons $Y_t = X_t/\rho(t)$. On a alors $\int_0^\infty X_s\, \nu(ds) = \int_0^\infty Y_s \rho^2(s)\, \pi(ds)$ et $\int_0^\infty X_s^2\, \pi(ds) = \int_0^\infty Y_s^2 \rho^2(s)\, \pi(ds)$; en posant $\mu(ds) = \rho(s)\, \nu(ds) = \rho^2(s)\, \pi(ds)$, on est ramené au problème 1. Les solutions du problème 1^{bis} sont donc les processus de $L^2(\mathbb{P}\otimes\pi)$ qui sont de la forme $\rho(t)M_t$, où M est une martingale; si l'on s'est donné la v.a. $Z = \int_0^\infty X_s\, \nu(ds)$, la solution s'écrit

$$X_t = \rho(t) \int_0^t \frac{1}{\mu([s,\infty[)}\, dZ_s\,.$$

Quel est l'analogue, dans une variété riemannienne, des propriétés ci-dessus? Elles font intervenir les intégrales $\int X_s\, \mu(ds)$ et $\int X_s^2\, \mu(ds)$; lorsque X prend ses valeurs dans une variété, $X_{t_1} + X_{t_2}$ n'a déjà aucun sens, et en attribuer un à $\int X_s\, \mu(ds)$ semble désespéré — à moins d'utiliser une connection pour recoller, par transport parallèle, les quantités $X_s\, \mu(ds)$, considérées comme vecteurs tangents infinitésimaux; mais puisqu'on sait (Bismut, Meyer) que la propriété d'être une martingale est stable par relèvement dans l'espace tangent au point de départ, ceci ne peut rien apporter de nouveau. Nous ne voyons donc pas comment généraliser la proposition précédente aux variétés.

En revanche, dans une variété riemannienne, on sait, pour les bons processus, définir une variation quadratique qui généralise le crochet des processus réels. Il est donc tentant de poser la question suivante, qui sera, elle, généralisable aux variétés : Parmi les semimartingales $(X_t)_{0\le t\le 1}$ (pour la filtration $(\mathcal{F}_t)_{0\le t\le 1}$) telles que X_1 soit une variable aléatoire donnée, quelle est celle qui minimise la variation quadratique moyenne $\mathbb{E}\big[[X,X]_1 - [X,X]_0\big]$?

Il est clair que s'il existe un processus continu et à variation finie A satisfaisant la contrainte, il sera solution, puisqu'il annule la quantité à minimiser! Plus généralement, si X est une solution à ce problème et si A est n'importe quel processus continu à variation finie nul en 0 et 1 (par exemple, un processus déterministe), $X+A$ est aussi une solution. Quand elle existe, la solution est donc loin d'être unique, et n'a aucune raison de ressembler à une martingale. Pour avoir une chance de restaurer l'unicité, travaillons à temps discret : remplaçons $(\mathcal{F}_t)_{t\ge 0}$ par une filtration discrète $(\mathcal{F}_k)_{k=0,\dots,n}$ sur $(\Omega, \mathcal{F}, \mathbb{P})$.

Problème 2. *Parmi les processus $(X_k)_{k=0,\dots,n}$ adaptés à la filtration $(\mathcal{F}_k)$ et tels que X_n soit une v.a. donnée dans L^2, chercher celui qui minimise la variation quadratique moyenne $\mathbb{E}\big[\sum_{k=1}^n (X_k - X_{k-1})^2\big]$ (ou plus généralement la variation quadratique pondérée $\mathbb{E}\big[\sum_{k=1}^n a_k(X_k - X_{k-1})^2\big]$, où les coefficients $a_k > 0$ sont des réels donnés).*

Étant donné un processus adapté $X = (X_k)_{k=0,\dots,n}$ de carré intégrable, posons $Y_k = X_k + \mathbb{E}[X_n - X_k|\mathcal{F}_0]$, de sorte que $Y_n = X_n$ et $Y_0 = \mathbb{E}[X_n|\mathcal{F}_0]$. On a

$$Y_k - Y_{k-1} = X_k - X_{k-1} - \mathbb{E}\big[X_k - X_{k-1}\big|\mathcal{F}_0\big]\ ,$$

donc $\|Y_k - Y_{k-1}\|_2 \leq \|X_k - X_{k-1}\|_2$ avec égalité ssi $\mathbb{E}[X_k - X_{k-1}|\mathcal{F}_0] = 0$. Il en résulte qu'en ce qui concerne le critère à minimiser, Y fait au moins aussi bien que X, et même strictement mieux sauf si $\mathbb{E}[X_k|\mathcal{F}_0]$ est indépendant de k. Dans l'étude du problème 2, on peut donc se restreindre à ne considérer que des processus X tels que $X_0 = \mathbb{E}[X_n|\mathcal{F}_0]$.

Effectuons le changement de variable $D_k = X_k - X_{k-1}$. Le problème peut être ainsi reformulé : donnés X_n et $X_0 = \mathbb{E}[X_n|\mathcal{F}_0]$, chercher le processus adapté $(D_1, \ldots, D_n)$ vérifiant $\sum_{k=1}^n D_k = X_n - X_0$ et qui minimise $\mathbb{E}\left[\sum_{k=1}^n a_k D_k^2\right]$. À la formulation discrète près, nous sommes ramenés au problème 1^{bis}, et même au problème 1 si les a_k sont tous égaux.

Ceci permet de résoudre le problème : le processus $a_k D_k$ doit être une martingale, et, en posant $f_k = \dfrac{1}{a_k} + \ldots + \dfrac{1}{a_n}$ (avec $f_{n+1} = 0$) et $M_k = \mathbb{E}[X_n|\mathcal{F}_k]$, on obtient pour solution

$$a_k D_k = \sum_{i=1}^{k} \frac{1}{f_i}(M_i - M_{i-1}) ,$$

$$X_k = M_k - f_{k+1} \sum_{i=1}^{k} \frac{1}{f_i}(M_i - M_{i-1}) .$$

Dans le cas le plus simple où tous les a_k valent 1, la solution est donc caractérisée par le fait que les accroissements D_k du processus X forment une martingale. Ceci laisse espérer que, bien que tous les processus continus et à variation finie soient solution du problème 2 non discret, ceux dont la dérivée est une martingale soient, en un sens, encore plus optimaux que les autres; mais je ne vois pas comment rendre ceci rigoureux, et ce serait de toute façon impossible à transférer aux martingales dans les variétés.

Revenant au problème discret, avec des a_k quelconques, on peut se demander quand la solution X est une martingale. Cela se produit lorsque, pour $k = 1, \ldots, n$, $\mathbb{E}[D_k|\mathcal{F}_{k-1}] = 0$, c'est-à-dire $D_{k-1} = 0$; la condition est donc que $\mathbb{E}[X_n|\mathcal{F}_{n-1}]$ soit mesurable pour $\mathcal{F}_0$ (curieusement, cette condition est indépendante des a_k), et, lorsque cela a lieu, les processus X, M et D sont constants sur l'intervalle $[0, n-1]$.

Bien qu'en général la solution X ne soit pas une martingale, il est possible, en jouant sur les coefficients a_k, d'obtenir qu'elle le soit approximativement : choisissons pour les a_k une progression géométrique a^k, dont nous faisons tendre la raison a vers $+\infty$. Les coefficients f_{k+1}/f_i qui figurent dans l'expression de X_k tendent vers zéro, et X tend donc vers la martingale M. Ceci a une interprétation très simple : chaque coefficient a_k étant négligeable devant le suivant a_{k+1}, pour construire le X optimal, il faut d'abord minimiser le dernier terme $a_n \mathbb{E}\left[(X_n - X_{n-1})^2\right]$, ce qui donne $X_{n-1} = \mathbb{E}[X_n|\mathcal{F}_{n-1}]$, puis l'avant-dernier, etc...; de proche en proche, on obtient ainsi la martingale.

En définitive, dans chacun de ces deux cas, on voit que ça ne marche réellement bien que si $n = 1$, et ce qui se laisse aisément caractériser, ce ne sont pas les martingales, mais beaucoup plus prosaïquement les espérances conditionnelles. Encore une fois, c'est là quelque chose qui refuse de se laisser transporter aux variétés : on sait bien que dans une variété la notion de martingale n'est pertinente que pour des processus continus et n'a aucun sens en temps discret, puisqu'il n'existe aucune définition de l'espérance conditionnelle qui satisfasse la propriété d'associativité $\mathbb{E}[\mathbb{E}[X|\mathcal{G}]|\mathcal{H}] = \mathbb{E}[X|\mathcal{H}]$ pour $\mathcal{H} \subset \mathcal{G}$.

NOTE SUR LES ÉPREUVES

Tout ceci est bien connu! Stricker a attiré mon attention sur la référence

S.R. Pliska : Duality theory for some stochastic control models. *Stochastic Differential Systems,* Lecture Notes in Control and Information Theory N° 43, Springer 1983.

qui traite des mêmes questions dans un contexte plus général et contient en outre une abondante bibliographie.