SUR LES FEEDBACKS LOCAUX OPTIMAUX EN COMMANDE STOCHASTIQUE

Jean-Pierre QUADRAT
IRIA LABORIA
B.P 105
78150 - LE CHESNAY.

INTRODUCTION

La recherche du feedback optimal S^* dans le cas de l'observation complète de l'état X conduit à la recherche d'une fonction dépendant de l'état et du temps. Dans le cas général où on n'a pas une expression explicite de cette fonction, il faudra la tabuler. Si $X \in \mathbb{R}^n$, S^* à valeurs dans $\mathbb{R}^m$ il faut tabuler une fonction de $[0,T] \times \mathbb{R}^n \to \mathbb{R}^m$. Discrètisons chaque composante de l'état et le temps en k points, on est donc conduit à calculer mk^{n+1} valeurs. Dans les applications n est plus souvent de l'ordre de 20 que de 2, k doit souvent être souvent au moins de l'ordre de 10, on est alors conduit à calculer 10^{21} paramètres. Si l'on considère que les gros ordinateurs ont des mémoires de l'ordre de 10^6 mots, on voit que l'on est très loin du compte. Il nous semble donc que la recherche de la commande optimale est un problème mal posé. On peut restreindre ses ambitions et chercher la commande en boucle ouverte ou en boucle sur un sous-ensemble des variables d'état (feedback partiel), on montrera que ce problème est encore plus difficile, au moins dans le cas général, que le problème précédent. Par contre dans certains cas particuliers dont le prototype est le cas de système à dynamiques découplées pouvant avoir un critère couplant, on montrera qu'on peut calculer l'optimum dans la classe des feedbacks locaux.

Pour montrer ces résultats on se placera dans le cadre de la commande de diffusions stochastiques. Pour celà nous introduirons les diffusions, l'équation de Fokker Planck, puis les équations de Kolmogorov et de Hamilton Jacobi. Nous montrerons que la recherche d'un feedback partiel optimal se ramène au problème de la commande de l'équation de Fokker Planck. Nous remarquerons que ce problème est plus compliqué que la résolution de l'équation de Hamilton Jacobi donnant le feedback global optimal. Nous étudierons enfin le cas particulier des systèmes à dynamiques découplées pour lesquels nous montrerons que les conditions d'optimalité dans la classe de feedbacks locaux, (optimalité au sens joueur par joueur), se ramène à la résolution d'un système d'équation du type Hamilton Jacobi couplé. Chaque équation exprime les conditions d'optimalité d'un contrôleur local, la stratégie des autres contrôleurs étant fixée. Le couplage entre ces équations provient du fait que chaque contrôleur fait la moyenne sur des informations auxquels il n'a pas accès. Un algorithme pour résoudre ce système d'équation sera donné.

1 - RAPPELS SUR LES DIFFUSIONS, FORMULE D'ITO, EQUATION DE FOKKER PLANCK, DE KOLMOGOROV.

Dans ce paragraphe nous donnons le formalisme mathématique conduisant aux résultats d'éxistence et d'unicité les plus généraux, pour la notion intuitive de processus régi par une équation différentielle ordinaire perturbée par un bruit. Ce formalisme bien qu'un peu abstrait conduit à la présentation la plus économe (nous semble-t-il!) des résultats sur les équations différentielles stochastiques, (une fois admise l'existence et l'unicité d'une solution au "problème de martingale"). D'autre part les feedbacks "bang-bang", important en pratique ne pose grâce à cette présentation aucun problème.

On se donne donc l'espace probabilisable (Ω, F_t, F) avec :
$\Omega = C([0,T];\ R^n)$ est l'ensemble des fonctions continues de $(0,T) \to \mathbb{R}^n$.
Pour $\omega \in \Omega$, on note $X_t(\omega) = \omega_t$ l'application coordonnée.
F_t désigne plus petite tribu engendrée par $(X_s,\ s \leq t)$. $F = F_T$.
On se donne alors deux fonctions b et a

(1.1) $b : [0,T] \times \mathbb{R}^n \to \mathbb{R}^n$ borelienne bornée

(1.2) $a : [0,T] \times \mathbb{R}^n \to \mathcal{M}n,n$ continue bornée où $\mathcal{M}n,n$ désigne l'ensemble des matrices symétriques définies strictement positives de type nxn.

Définition :

P mesure de probabilité sur (Ω, F_t, F) sera dite solution du problème de martingale (x,b,a) si et seulement si pour $\forall \varphi \in C_b^{1,2}([0,T] \times \mathbb{R}^n)$*.

(1.3) $\varphi(t,X_t) - \int_o^t L_{a,b}\, \varphi(s,Xs)$ dans est une (P,F_t) martingale

(1.4) $P\ (X(o) = x) = 1$

où $L_{a,b}$ désigne l'opérateur parabolique suivant :

$$L_{a,b} = \frac{\partial}{\partial t} + \sum_i b_i \frac{\partial}{\partial x_i} + \frac{1}{2} \sum_{i,j} a_{ij} \frac{\partial^2}{\partial x_i \partial x_j}$$

rappelons qu'une (P,F_t) Martingale est un processus $M_t)$ vérifiant la propriété $E^{F_S}(M_t) = M_s$ pour $t \geq s$.

L'espérance est prise au sens de la loi P.

---

* $C_b^{1,2}$ fonctions deux fois différentiable en x, une fois en temps, les dérivées étant continues et bornées.

Nous appelons le processus canonique X_t , diffusion stochastique un théorème de Stroock-Varadhan [8] nous indique qu'il existe une seule mesure de probabilité P, solution du problème de martingale (x,b,a) pour b et a vérifiant les hypothèses (1.1) et (1.2). Le processus canonique X_t défini sur (Ω,F_t,F,P) sera qualifié de diffusion faible de dérive b, de terme de diffusion a et de condition initiale x. Ce processus est markovien STROOCK-VARADHAN [8]. Si on se donne la loi d'entrée (loi sur la condition initiale) on appellera encore P la loi $P_x\,(d\omega)\;\mu\,(dx)$.

Formule d'Ito et de Dynkin.

La formule de définition du problème de martingale implique la formule de différentiation suivante, connue sous le nom de formule d'Ito.

pour $\forall \varphi \in C_b^{1,2}\,([0,T]\times \mathbb{R}^n)$

$$(1.5)\quad d\varphi(t,X_t) = \Big(\frac{\partial\varphi}{\partial t}(t,X_t) + \sum_{i=1}^{n} b_i\,(t,X_t)\,\frac{\partial\varphi}{\partial x_i}(t,X_t) + \sum_{i=1}^{n}\sum_{j=1}^{n}\frac{1}{2}a_{ij}\,(t,X_t)\,\frac{\partial^2\varphi}{\partial x_i\,\partial x_j}(t,X_t)\Big)\,dt + dM_t$$

où M_t est une martingale

Puisque M_t est une martingale $E^{F_t}\,dM_t = 0$ et donc on a la formule de Dynkin.

$$(1.6)\quad E^{Ft}\,d\varphi(t,X_t) = \Big[\frac{\partial\varphi}{\partial t}(t,X_t) + \sum_{i=1}^{n} b_i\,(t,X_t)\,\frac{\partial\varphi}{\partial x_i}(t,X_t) + \sum_{i=1}^{n}\sum_{j=1}^{n}\frac{1}{2}a_{ij}\,(t,X_t)\,\frac{\partial^2\varphi}{\partial x_i\,\partial x_j}(t,X_t)\Big]\,dt$$

La formule (1.5) indique d'autre part en l'appliquant à la fonction $\varphi(x) = x$ que

$$dX_t = b\,(t,X_t)\,dt + dM_t \qquad X(0) = x$$

et donc que X_t peut être considéré comme un système différentiel $\dot{X}_t = b\,(t,X_t)$ $X(0) = x$, perturbé par un bruit "$\frac{dM_t}{dt}$" en fait (1.5) caractérise complètement la loi du bruit. La formule d'Ito ou de Dynkin indique que pour de tels processus on n'a pas la formule de différentiation habituelle, mais qu'il faut faire un développement au second ordre.

La grande importance pratique de ces processus provient du fait, qu'en gros pour tout processus $X(t)$ admettant une fonction d'autocorrélation $\in L^1$, le processus $\frac{1}{\varepsilon} X \left(\frac{t}{\varepsilon^2}\right)$ tend lorsque $\varepsilon \to 0$ vers un bruit blanc cf. par ex. Bensoussan - Lions-Papanicolaoü [1]. De ce résultat on déduit des théorèmes pouvant être considérés comme des théorèmes du type centrale limite pour les processus, les diffusions jouant alors le rôle de la loi gaussienne, pour le théorème centrale limite habituel.

Equation de Kolmogorov.

Sous les hypothèses (1.1) et (1.2), et si l'équation (1.7) :

$$(1.7) \quad \frac{\partial V}{\partial t} + \sum_{i=1}^{n} b_i \frac{\partial V}{\partial x_i} + \sum_i \sum_j \frac{1}{2} a_{ij} \frac{\partial^2 V}{\partial x_i \partial x_j} + f = 0 \quad (t,x) \in [0,T[\times \mathbb{R}^n$$

$$V(T,x) = \emptyset\,(x) \, ,$$

admet une solution $C_b^{1,2}$ alors $V(t,x)$ s'interprète :

$$(1.8) \quad V(t,x) = E \left(\int_t^T f\,(t,X_s)\, ds + \emptyset\,(X_T) \;\Big|\; X_t = x \right)$$

La difficulté mathématique essentielle étant ici de répondre à la question : quelles sont les hypothèses minimales sur b, a, f, $\emptyset$ pour que $V \in C_b^{1,2}$. Ou plus généralement quelle sont les hypothèses minimales sur b, a, f, $\emptyset$ assurant la régularité minimum sur V permettant l'interprétation (1.8). Bensoussan - Lions [2] répondent à ce genre de questions. Disons que sous les hypothèses b et a vérifiant (1.1) f borelienne bornée, $\emptyset$ continue bornée, V est $W^{1,2,p}$*. Ce qui est suffisant comme régularité pour avoir la formule de Dynkin donc l'interprétation (1.8). Montrons la validité de cette interprétation sous l'hypothèse : V solution (1.7) est $C_b^{1,2}$.

S'il en est ainsi grâce à (1.3) on a :

$$E^{F_t}\, V\,(T,X_T) - V\,(t,X_t) = E^{F_t} \int_t^T L_{ab}\; V\,(s,Xs)\, ds$$

or grâce à (1.7), on a :

$$V\,(T,X_T) = \emptyset\;(X_T) \, ,$$

---

* $W^{1,2,p}$ ensemble des fonctions appartenant à L^p ainsi que sa dérivée en temps et ses deux premières dérivées en espace.

et,

$$L_{a,b}\ V\ (s,X_s) = -\ f\ (s,X_s)\ ,$$

et donc

$V\ (t,X_t) = E^{Ft}\ (\int_t^T f\ (s,X_s)ds + \emptyset\ (X_T))$ ce qui est le résultat désiré.

Equation Fokker Planck.

prenons $\varphi \in C_K^{1,2}$**

$$E\,\varphi(T,X_t) - E\,\varphi(0,X_o) = E \int_o^T L_{a,b}\ \varphi\,(s,X_s)\ ds$$

Si l'on note $p(t,x)$ la loi marginale de X_t à l'instant t on a :

$$(1.9)\quad (p_T,\ \varphi_T) - (p_o,\ \varphi_o) = \int_o^T (L_{a,b}\ \varphi_t,\ p_t)$$

où $\qquad (p_t,\ \varphi_t) = \int_{\mathbb{R}^n} p(t,x)\ \varphi(t,x)\ dx$

Mais alors (1.9) se réécrit :

$$(1.10)\quad \int_o^T (L^*_{a,b}\ p,\varphi) = 0 \qquad \forall\ \varphi \in C_K^{1,2}$$

où $L^*_{a,b}p = -\ \frac{\partial p}{\partial t} - \sum_i \frac{\partial}{\partial x_i}(b_i p) + \sum_{i,j} \frac{1}{2} \frac{\partial^2}{\partial x_i \partial x_j}(a_{ij}p)$

et si $L^*_{a,b}p$ est borné dans L^2 (1.10) entraine $L^*_{a,b}p = 0$

puisque $\varphi \in C_K^{1,2}$ est dense dans L^2 et donc le résultat peut s'énoncer ainsi Si p_t densité de probabilité de la loi de X_t est suffisam-

---

** ensemble des fonctions étant continues et à support compact ainsi que ses dérivées une fois en temps et deux fois en espace.

ment régulière $W^{1,2,2}$ alors, elle est solution de l'équation de Fokker Planck

$$\left\{ \begin{array}{l} -\dfrac{\partial}{\partial t}\, p - \sum_i \dfrac{\partial}{\partial x_i}\, (b_i p) + \sum_{i,j} \dfrac{1}{2} \dfrac{\partial^2}{\partial x_i \partial x_j}\, (a_{ij} p) = 0 \\ p_o = \mu \end{array} \right.$$

où μ désigne la loi initiale.

2 - PROBLEME DE COMMANDE OPTIMALE - EQUATION D'HAMILTON JACOBI

On définit le problème de commande stochastique, étant données :

- une commande en boucle fermée $S : [0,T] \times \mathbb{R}^n \times U \to \mathbb{R}^m$, $(t, x) \to S(t,x)$, U compact de $\mathbb{R}^m$;

- une dérive dépendant d'un paramètre à valeur dans U :
  $b : [0,T] \times \mathbb{R}^n \times U \to \mathbb{R}^n$, $(t, x, u) \mapsto b(t,x,u)$ borélienne fornée, lipschitz en u ;

- un terme de diffusion défini par (1.2) ;

On appellera P_S la solution au problème de martingale $(x, b\circ S, a)$ où
$b\circ S : [0,T] \times \mathbb{R}^n \to \mathbb{R}^n$, $(t, x) \mapsto b(t,x,S(t,x))$

On définit alors deux fonctions coûts :

$C : [0,T] \times \mathbb{R}^n \times U \to \mathbb{R}^+$ borélienne bornée lipschitz en u ;
$\phi : \mathbb{R}^n \to \mathbb{R}^+$ continue.

Le critère s'écrit alors :

$$\int_o^T C\circ S\ (t, X_t)\ dt + \phi(X_T) \quad \text{avec } C\circ S(t,x) = C(t,X,S(t,X)).$$

Le problème de commande stochastique s'énonce trouver la commande feedback réalisant :

$$\underset{S}{\mathrm{Min}}\ E_{P_S} \int_o^T C\circ S(t,X_t)dt$$

avec P_S solution du problème de martingale $(X, b\circ S, a)$.

La solution de ce problème est obtenue en résolvant l'équation de Hamilton Jacobi.

$$(2.1)\quad \begin{cases} \dfrac{\partial V}{\partial t} + \underset{u\in U}{\mathrm{Min}}\{ \sum_{i=1}^n b_i(t,x,u)\dfrac{\partial V}{\partial x_i} + C(t,x,u)\} + \dfrac{1}{2}\sum_i \sum_j a_{ij}(t,x)\dfrac{\partial^2 V}{\partial x_i \partial x_j} = 0 \\ V(T,x) = \phi(x) \qquad x \in \mathbb{R}^n \end{cases}$$

$S^*(t,x) \in \operatorname{Arg}\operatorname{Min}_u \left\{ \sum_{i=1}^n b_i(t,x,u)\frac{\partial V}{\partial Xi} + C(t,x,u)\right\}$ défini un feedback S^* optimal.

$V(t,x)$ solution de (2.1) s'interprète comme

$$\operatorname{Min}_S E_{P_S} \left(\int_t^T C \circ S\,(s,Xs)\,ds + \phi\,(X_T) \,\middle|\, X_t = x\right)$$

Pour montrer cela supposons que la solution de (2.1) soit $C_b^{1,2}$ grâce à la définition du problème de martingale on a :

$$V(T,X_T) - V(0,X_o) - \int_o^T L_{a,b\circ S}\, V(t,X_t)\, dt \quad \text{est une} \quad (P_S, F_t)$$

martingale partant de 0, pour toute stratégie S et donc de (2.1) on obtient :

$$V(0,x) = E_{P_S}\left\{V(T,X_T) - \int_o^T L_{a,b\circ S}\, V(t,X_t)\, dt\right\}$$

$$\leq E_{P_S}\left\{\phi(X_T) + \int_o^T C\circ S(t,X_t)\, dt\right\}$$

On obtient l'égalité lorsqu'on fait le même raisonnement avec le feedback optimal S^*.

On a supposé ici que la solution de (2.1) est $C_b^{1,2}$, en fait cette hypothèse est trop forte. Dans Bensoussan - Lions [2] on montre la régularité de la solution de (2.1) suffisante pour avoir l'interprétation. La difficulté mathématique réside bien sur, dans cette partie.

Remarque 1.

Si $u \to (b(t,x,u),\ C(t,x,u))$ est linéaire et U un pavé on voit que la commande optimale est bang bang et donc que des hypothèses du type Lipschitz sur le terme de dérive, qui assurerait l'existence d'une diffusion, seraient insuffisantes. Ceci explique le choix fait de parler du problème de martingale, plutôt que de définir les diffusions au sens fort Gikhman-Skorohod [7].

Remarque 2.

On peut montrer par un raisonnement du type de celui employé pour montrer l'optimalité de l'équation de la programmation dynamique, que l'on ne perd rien à optimiser dans la classe des Feedbacks $S : [0,T] \times \mathbb{R}^n \to \mathbb{R}^m$, plutôt que dans la classe des commande adaptées $S : [0,T]\times\Omega \to \mathbb{R}^m$ S_t étant F_t mesurable. Autrement dit on ne gagne rien à retenir tout le passé (dans le cas markovien bien sur), il suffit de mémoriser le présent et d'optimiser dans la classe des commandes dépendant seulement du présent.

En conclusion, mis a part quelques difficultés d'ordre théorique du type régularité de l'équation de Hamilton Jacobi, la caractérisation des commandes optimales ne pose pas de difficulté. Par contre une difficulté du type numérique, presque insurmontable, apparaît : Comment résoudre numériquement (2.1) lorsque $x \in R^n$ avec n grand, disons supérieur 3 ou 4.

. En déterministe on pouvait intégrer (2.1) le long des caractéristiques (principe de Pontriaguine) et conserver en mémoire la commande seulement le long de la trajectoire optimale, ici cette méthode est exclue puisqu'en quelque sorte les caractéristiques sont aléatoires et donc qu'il faudrait calculer $S(t,x)$ $\forall x,t$. En stochastique on a donc à faire face au dilemme suivant - soit on résout l'équation d'H.J. soit on renonce à l'optimalité. Si l'on renonce à l'optimalité comment poser le problème de commande stochastique ? La première idée venant à l'esprit consiste à rechercher la commande dans une sous classe de Feedbacks. C'est la problème que nous nous posons dans le paragraphe suivant.

3 - OPTIMISATION DANS LA CLASSE DES FEEDBACKS LOCAUX.

Si $I = \{1,2,\dots k\}$ et $n_1, n_2, \dots, n_k$ désignent k nombres entiers

On se donne :

- $S_i : [0,T] \times \mathbb{R}^{n_i} \to U_i \subset \mathbb{R}^{m_i}$, U_i compact $\mathbb{R}^{m_i}$, $\forall i \in I$;

$$t, \quad x_i \quad S_i(t,x_i)$$

- une dérive dépendant d'un paramètre à valeur dans U :

$$b : [0,T] \times \mathbb{R}^n \times \prod_{i=1}^{k} U_i \to \mathbb{R}^n, \quad \text{avec} \quad n = \Sigma n_i \ ;$$
$$t \quad x \quad (u_1,u_2,\dots u_k) \to b(t,x,u)$$

- un terme de diffusion défini par (1.2).

On appellera P_S la solution du problème de martingale $(x, b\circ S, a)$

où $b \circ S : [0,T] \times \mathbb{R}^n \to \mathbb{R}^n$
$$t \quad x \quad b\,(t,x,S_1(t,x_1),\dots S_k(t,x_k))$$

Autrement dit, par rapport au paragraphe (2), la commande dépend ici que de l'état local, ou bien, chaque composante de la diffusion à un contrôleur observant seulement cette composante là du système.

Etant donnée la fonction coût

$$C : [0,T] \times \mathbb{R}^n \times \prod_{i=1}^{k} U_i \to \mathbb{R}^+$$
$$t, \quad x \quad (u_1,u_2 \dots u_k) \to C(t,x,u_1,u_2,\dots u_k),$$

et $\emptyset$ défini au paragraphe 2 ; le problème de commande stochastique s'énonce trouver la commande feedback local optimal réalisant

$$\operatorname*{Min}_{S} E_{P_S} \int_0^T C(t,X_t,S_1(t,X_1(t)), S_2(t,X_2(t)), \dots, S_k(t,X_t))\, dt + \phi(X_T)$$

P_S étant solution du problème de martingale $(x, b\circ S, a)$.

On a donc ici un problème d'une équipe ayant une information partielle contribuant à optimiser le même critère.

Une façon simple d'obtenir des conditions nécessaires d'optimalité et de voir ce problème comme un problème de contrôle le l'E.D.P. de Fokker Planck.

$$
(3.1)\quad \begin{cases}
\underset{\substack{S \\ \text{feedback} \\ \text{locaux}}}{\text{Min}} \{J(S) = \int_o^T \int_{\mathcal{O}} C(t, x, S_1(t,x), \dots S_k(t,x)) p(t,x_1,x_2\dots x_n)\, d\mathcal{O}\} \\
\qquad + \int \phi(\ ,x) p(T,x)\, d\mathcal{O} \\
\mathcal{O} = \mathbb{R}^n \\
d\mathcal{O} = dx_1 \dots dx_n \\
p \text{ satisfaisant} \\
-\frac{\partial p}{\partial t} - \sum_{i=1}^n \frac{\partial}{\partial x_i} (b_i(t,x,S(t,x))p) + \frac{1}{2} \sum_i \sum_j \frac{\partial^2}{\partial x_i \partial x_j} (a_{ij} p) = 0 \\
p(o) = \mu
\end{cases}
$$

Théorème 1 :

Une C.N.S pour que $J(R) \geq J(S)$ est que $p(R)$ désignant la solution de (3.5) pour le feedback R, $V(S)$ désignant la solution de (3.6) pour le feedback S, l'on ait :

$$
(3.2)\quad \int_{\mathcal{O}} \left(C o R + \sum_{i=1}^n b_i o R \frac{V(S)}{\partial x_i}\right) p(R)\, d\mathcal{O} \geq \int \left(C o S + \sum_{i=1}^n b_i o S \frac{\partial V}{\partial x_i}(S)\right) p(R)\, d\mathcal{O}
$$

pp en temps

Démonstration :

Montrons le résultat dans le cas $\phi = o$ (uniquement pour simplifier les intégrations par parties)

On note :

$$\langle f,g\rangle = \int_o^T \int_{\mathcal{O}} f(x,t)g(x,t)\,d\mathcal{O}\,dt \qquad (f,g)_t = \int_{\mathcal{O}} f(x,t)g(x,t)\,d\mathcal{O}$$

$$A^*(S) = -\frac{\partial\cdot}{\partial t} - \sum_{i=1}^{n} \frac{\partial}{\partial x_i} b_i \circ S\,\cdot + \frac{1}{2}\sum_{i=1}^{n}\sum_{j=1}^{n} \frac{\partial^2}{\partial x_i \partial x_j} a_{ij}\cdot$$

alors si $*$ désigne la transposition on a :

$$A(S) = \frac{\partial}{\partial t} + \sum_{i=1}^{n} b_i \circ S \frac{\partial}{\partial x_i} + \frac{1}{2}\sum_{i=1}^{n}\sum_{j=1}^{n} a_{ij} \frac{\partial^2}{\partial x_i \partial x_j}$$

or :

$$J(R) = \langle C\circ R, p(R)\rangle$$

L'inégalité (3.2) se réécrit :

$$(C\circ R + A(R)V(S), p(R)) \geq (C\circ S + A(S)V(S), p(R)) \quad \text{p.p}$$

démontrons son équivalence avec $J(R) \geq J(S)$

$J(R) \geq J(S)$

$\iff \langle C\circ R, p(R)\rangle \geq \langle C\circ S, p(S)\rangle$

$\iff \langle C\circ R - C\circ S, p(R)\rangle + \langle C\circ S, p(R) - p(S)\rangle \geq o$

par définition de $V(S)$ on a l'équivalence

$\iff \langle C\circ R - C\circ S, p(R)\rangle - \langle A(S)V(S), p(R) - p(S)\rangle \geq o$

$\iff \langle C\circ R - C\circ S, p(R)\rangle - \langle V(S), A^*(S)\,(p(R) - p(S))\rangle \geq o$

Mais par définition de $p(S)$ et de $p(R)$

$A^*(S)\,p(S) - A^*(R)\,p(R) = o$

et donc

$A^*(S)\,(p(S) - p(R)) + (A^*(S) - A^*(R)\,p(R)) = o$

et donc on a encore

$\iff \langle C\circ R - C\circ S, p(R)\rangle + \langle V(S), (A^*(R) - A^*(S))\,p(R)\rangle \geq o$

$\iff \langle C\circ R + A(R)V(S) - (C\circ S + A(S)V(S)), p(R)\rangle \geq o$

d'où le résultat.

Remarque

On peut montrer également ce résultat directement grâce à l'interprétation probabiliste.

Théorème 2 :

Une C.N pour que S^* soit optimal dans la classe des feedbacks locaux supposé exister est que S^* satisfasse le système d'équations suivant :

$$(3.5) \quad \begin{cases} -\dfrac{\partial p_{S^*}}{\partial t} - \displaystyle\sum_{i=1}^{n} \dfrac{\partial}{\partial x_i} (b_i o S^* \, p_{S^*}) + \dfrac{1}{2} \sum_{i=1}^{n} \sum_{j=1}^{n} \dfrac{\partial}{\partial x_i \partial x_j} (a_{ij} p_{S^*}) = o \\ p_{S^*}(o,x) = \mu \end{cases}$$

$$(3.6) \quad \begin{cases} \dfrac{\partial V_{S^*}}{\partial t} \displaystyle\sum_{i=1}^{n} b_i o S^* \dfrac{\partial V_{S^*}}{\partial x_i} + \sum_{i,j} \dfrac{1}{2} a_{ij} \dfrac{\partial^2}{\partial x_i \partial x_j} V_{S^*} + CoS^* = o \\ V_{S^*}(T,x) = \phi(x) \end{cases}$$

$$(3.7) \quad S^*(t,.) \in \operatorname{Arg\,Min}_{S} \int_{\mathcal{O}} (CoS + \sum_{i=1}^{n} b_i o S \, \dfrac{\partial V_{S^*}}{\partial x_i}) \, p_{S^*} \, d\mathcal{O}$$

l'intégrale étant évaluée au point t.

Idée de la démonstration.

Ce résultat découle du th. précédent, lorsqu'on l'applique à des $S = \begin{cases} S^* \text{ sur } C_A \\ \text{qcq sur } A \end{cases}$ où S^* est la commande optimale et A est un intervalle en temps suffisamment petit pour perturber P_S sur les temps postérieurs aux éléments de A de façon aussi faible que l'on veut.

Pour avoir des résultats précis de ce type on pourra se reporter à LIONS [5].

Remarque 1 :

Le théorème 1, donne un algorithme pour obtenir la solution. S^*.

Etape 1 : S_o donné

Etape 2 : Calculer P_{S_o}

Etape 3 : Etape 3.1. se donner t

étape 3.2. résoudre (3.7) (minimisation complète)

étape 3.3. intégrer d'un pas (3.6) (pb discrétisé en temps)

étape 3.4. $t \to t-h$

étape 3.5. retourner à l'étape 3.1 jusqu'à $t = o$

Etape 4 : S_o devient S_1 obtenu au cours de l'étape 3.2.

Etape 5 : retourner à l'étape 3.2 jusqu'à convergence.

On obtient ainsi une suite décroissante de coûts, un point stationnaire de l'algorithme vérifiera les conditions nécessaires d'optimalité du th.2. En gros ce sera un minimum local.

Remarque 2 :

La description de l'algorithme de la remarque 1, montre qu'on a à résoudre une suite de problème de difficulté voisine de la résolution de l'équation de Hamilton Jacobi (étape 3). Et donc, en conclusion, se restreindre à rechercher l'optimum de la classe des feedbacks locaux conduit en général à la résolution d'un problème plus difficile encore que la résolution de l'équation de Hamilton Jacobi, et donc que d'obtenir un feedback sur l'état complet. Cette approche nous semble donc (au moins pour le cas général) stérile du point de vue de la commande des grands systèmes stochastiques. Nous allons voir que dans des cas particuliers elle peut être très utile.

4 - UN EXEMPLE NON TRIVIAL OU LA RECHERCHE DANS LA CLASSE DES FEEDBACKS LOCAUX PERMET D'ELARGIR LA CLASSE DES PROBLEMES RESOLUBLES NUMERIQUEMENT.

Nous étudions un système composé de n sous-systèmes à dynamiques découplées, le critère étant couplant. Pour simplifier la présentation, nous supposerons les sous-systèmes scalaires bien que le résultat énoncé reste vrai dans le cas vectoriel.

Plus précisemment étant donnés :

- le terme de dérive :

$$b : I\times[0,T]\times\mathbb{R}\times U \to \mathbb{R} \text{ borelienne, lipschitz en } u,$$
$$i,t,x,u_i \to b_i(t,x_i,u_i)$$

- le terme de diffusion vérifiant :

$$a_{ij} = o \quad i \neq j, \quad a_{ii} : [0,T] \times \mathbb{R} \to [\alpha,\beta] \subset \mathbb{R}^+ \quad \alpha > o, \text{ continues ;}$$

- des feedbacks locaux :

$$S : I \times [0,T] \times \mathbb{R} \to U \subset \mathbb{R}$$
$$i \quad t \quad x_i \quad S_i(t,x_i)$$

On note : $b\circ S : (i,t,x) \to b_i(t,x_i, Si(t,x_i))$.

Soit P_S la solution du problème de martingale $(x,b\circ S,a)$.

Soit C la fonction coût

$$C : [0,T]\times \mathbb{R}^n\times U^n \to \mathbb{R}^t \text{ borélienne, lipschitz en } u.$$
$$t, x, u \quad C(t,x,u)$$

On note de même $C\circ S : t,x \to C(t,x,S_1(t,x_1), S_2(t,x_2)\dots, S_n(t,x_n))$

Le problème consiste alors à :

$$\underset{S}{\mathrm{Min}}\ \mathbb{E}_{P_S} \int_o^T C\circ S(t,X_t)dt$$

La différence essentielle avec le paragraphe précédent est la dépendance en x de la dérive, ici b_i est fonction de x_i, et donc, se restreindre à la classe des feedbacks locaux maintient le caractère découplé à la dynamique. Il est à remarquer que en général on perd quelque chose à se restreindre à la classe des feedbacks locaux puisque le critère étant couplant, il n'y a aucune raison pour que le feedback global optimal ait une structure découplée. Là encore, on voit une différence avec le cas déterministe , dans ce dernier cas les feedbacks locaux contenant les commandes boucles ouvertes, l'optimum dans la classe des feedbacks locaux est égal à l'optimum dans la classe de feedbacks globaux.

Montrons que pour des problèmes du type dynamique découplé, se restreindre à la classe des feedbacks locaux permet de résoudre numériquement des problèmes inaccessibles par l'équation de Hamilton Jacobi du problème global.

Théorème.

Une conditon suffisante pour que les Si soit optimaux, contrôleur par contrôleur est que le système d'équations (supposer exister et avoir une solution suffisamment régulière) suivant soit satisfait :

$$(4.1)\quad \begin{cases} -\dfrac{\partial p_S^i}{\partial t} - \dfrac{\partial}{\partial x_i}\left[b_i \circ S_i \, p_S^i \right] + \dfrac{1}{2}\dfrac{\partial^2}{\partial x_i^2}\left[a_{ii} \, p_S^i \right] = o & i = 1\dots,n \\ p_S^i(o,x^i) = \delta_{x_o^i}(x^i) \ , \ \delta \text{ mesure de Direc ;} & \end{cases}$$

$$(4.2)\quad \begin{cases} \dfrac{\partial V_S^i}{\partial t} + \operatorname{Min}_{ui} \left\{ b_i(u) \dfrac{\partial V_S^i}{\partial x_i} + \psi_S^i(u_i) \right\} + \dfrac{1}{2} a_{ii} \dfrac{\partial^2 V_S^i}{\partial x_i^2} = o & \\ V_S^i(T,x) = o & i = 1\dots,n \end{cases}$$

avec :

$$\psi_S^i(t,x_i,u_i) = \int C\left[t, x_1,\dots x_n, S_1(t,x_1), S_{i-1}(t,x_{i-1}), \ , \ S_{i+1}(t,x_{i+1}),\dots,S_n(t,x_n) \right] \prod_{j \neq i} p_S^j(t,x_j)\, dx_j$$

$$S(t,x_i) \in \operatorname{Arg} \min_{u_i} \left\{ b_i(u) \frac{\partial V^i}{\partial x_i} + \psi_S^i(u_i) \right\}$$

Démonstration.

D'après la définition du problème de martingale on a pour V_S^i, solution de (4.1), (4.2), R_i étant un feedback local et P_{R_i} la solution du Pb de

$(X, (b_1 o S_1 \ldots, b_i o R_i, b_{i+1} o S_{i+1}, \ldots b_n o S_n), a)$

$$E_{P_{R_i}} \{V_S^i (T, X_T^i) - V_S^i (o, x_o^i) - \int_o^T \left(\frac{\partial V_S}{\partial t} + b_i o R_i \frac{\partial V_S^i}{\partial x_i} + \frac{1}{2} a_{ii} \frac{\partial^2 V_S^i}{\partial x_i^2} (t, X_t^i)\right) dt\} = o$$

et donc :

$$V_S^i (o, x_o^i) \leq E_{P_{R_i}} \int_o^T \psi_S^i \; R_i (t, x_t) dt$$

$$\leq E_{P_{R_i}} \int_o^T C[t, X_t, S_1(t, X_t^1), \ldots, R_i(t, X_t^i), S_{i+1}(t, X_t^{i+1})) \ldots] dt$$

On a l'égalité pour $R_i = S_i$, et donc on a :

$$V_S^1 (o, x_o^1) = V_S^2 (o, x_o^2) \ldots = V_S^n (o, x_o^n) = E_{P_S} \int_o^T C o S(t, X_t) dt$$

De plus avec la notation :

$$J(S_1, \ldots S_{i-1}, R_i, S_{i+1}, \ldots S_n) = E_{P_{R_i}} \int_o^T C(t, X_t, S_1(t, X_t^1), \ldots, R_i(t, X_t^i), \ldots S_n(t, X_t^n)) dt$$

on a :

$$J(S) \leq J(S_1, \ldots S_{i-1}, R_i, S_{i+1}, \ldots S_n) \; \forall R_i.$$

On a donc le résultat.

Pour résoudre (4.1) (4.2) on peut suivant Pronovost [6] utiliser l'algorithme de relaxation suivant :

Etape 1 : S donné, calculer p_S en résolvant (4.1) ;

Etape 2.i : Calculer $\psi_S^i(u_i)$

Etape 3.i : Calculer V_S^i et R_i en résolvant (4.2)

Etape 4.i : S devient $S_1,\dots S_{i-1}, R_i, S_{i+1},\dots S_n$

Etape 5..i : Calculer p_S^i en résolvant (4.1)i ;

Etape 6.i : i devient i+1 modulo^n, retourner en (2.i) jusqu'à convergence.

On obtient une suite de V_S^i ayant la propriété que $V_S^i(o, x_o^i)$ décroit.

On voit donc que pour résoudre (4.1) (4.2), on a été amené à résoudre une suite d'équations aux dérivées partielles sur $[0,T]\times \mathbb{R}$, au lieu d'avoir à résoudre l'équation de Hamilton Jacobi dans $[0,T]\times \mathbb{R}^n$ donnant le feedback global optimal. Il y a là un gain énorme.

Remarque 1 :

Bien que cela ne ressorte pas très clairement de cette présentation rapide, le point important n'est pas que le système soit à dynamiques découplées, mais que la solution de l'équation de Fokker Planck ait la forme produit $p(x_1,\dots x_n) = \prod_{i=1}^{n} p_i(x_i)$, il faut dans ce cas se restreindre à des feedbacks conservant cette structure produit, il est alors possible d'obtenir le même genre de résultat. Nous ne détaillerons pas ce dernier cas qui fera l'objet d'une prochaine publication COHEN, QUADRAT, VIOT [3].

Remarque 2 :

Une difficulté numérique subsiste, c'est le calcul de $\psi_S^i(u_i)$ qui demande le calcul numérique d'une intégrale multiple. La méthode générale sera une méthode de Montecarlo, il suffit de générer des variables aléatoires Scalaires de loi indépendants pour donner une approximation de cette intégrale multiple. On peut parfois se ramener à des produits de convolution cas où le critère s'écrit $C(\sum_i g_i(X_i, u_i))$ ou bien dans ce dernier cas approximer la loi de $\sum_i g_i$ en utilisant le théorème centrale limite. Pour un exemple d'application, on pourra consulter DELEBECQUE - QUADRAT [3] ou PRONOVOST [6]. Sur un exemple dans la première référence on verra qu'il faut peu d'itérations pour atteindre l'optimum.

Conclusion :

La recherche de Feedback étant essentielle dans le cas de la commande stochastique, on a montré que la recherche du Feedback optimal est un idéal qu'il sera difficile de réaliser, qui est en tout cas inaccessible pour l'instant, dès que la taille du système devient grande (supérieure à 4 ou 5). La recherche de l'optimum dans une classe de feedback moins grande est en général encore plus difficile. Cependant dans certains cas particuliers par exemple système à dynamiques découplées, mais pouvant avoir un critère couplant, il est possible de calculer effectivement le feedback local optimal joueur par joueur pour des systèmes de taille beaucoup plus grande par exemple un système ayant une trentaine de composantes semble parfaitement accessible. Il nous semble important d'explorer ces cas particuliers.

[1] A.BENSOUSSAN - J.L.LIONS - PAPANICOLAOU
Asymptotic Analysis for periodic structure
North Holland 1978

[2] A.BENSOUSSAN - J.L.LIONS
Application des Inéquations variationnelles en contrôle stochastique
Dunod 1978

[3] G.COHEN - J.P.QUADRAT - M.VIOT
Sur le contrôle optimal de système dont la loi de probabilité à la forme produit à paraître

[4] F.DELEBECQUE - J.P.QUADRAT
Contribution of stochastic control, singular perturbation averaging and team theories to an example of large scale systems : Management of Hydropower production IEEE Automatic control AC-28 n° 2 p.209-222 Av.78

[5] J.L.LIONS - Contrôle optimal des systèmes gouvernés par des équations aux dérivées partielles
Dunod 1968

[6] R.PRONOVOST - Planification de la production énergétique au moyen de modèles à réservoir multiple col. Th. des Syst. et ap. à la gestion des services publics. Presses Univ. de Montréal 1975

[7] J.I.GIKHMAN - A.V.SKOROHOD
Stochastic differentiel equations
Spinger Verlag, New York 1972

[8] D.W.STOOCK - S.R.S.VARADHAN
Diffusion with continuous coefficients I and II Com. on pure and applied math. 1967