CAHIERS DE MATHEMATIQUES DE LA DECISION

PRINCIPE D'OPTIMALITÉ ET INFORMATION NÉCESSAIRE POUR LA CONCEPTION DANS LE CAS LE PLUS DÉFAVORABLE

Pierre BERNHARD

N° 7923

CENTRE DE RECHERCHE DE MATHEMATIQUES DE LA DECISION
EQUIPE DE RECHERCHE ASSOCIÉE DU CNRS NUMERO 249
CEREMADE - UNIVERSITE DE PARIS-DAUPHINE

Place du Maréchal De Lattre de Tassigny 75775 PARIS Cedex 16 - Téléphone 505.14.10

La "Conception dans le cas le plus défavorable", encore appelée "Jeu contre la nature" consiste à concevoir la loi de commande d'un système dynamique de façon à minimiser le coût garanti contre toute perturbation (admissible). Cela mène à concevoir le problème comme un jeu entre les perturbations (ou la nature) et nous, rejoignant ainsi la théorie des jeux différentiels à deux joueurs et somme nulle. Ce problème n'a d'intérêt, et bien souvent de solution, que si on est autorisé à moduler sa propre réponse en fonction, de ce que fait l'"adversaire" et qu'on a pu observer. En termes techniques cela veut dire qu'on désire autoriser l'utilisation de stratégies causales et non seulement de lois de commandes en boucle ouverte. Comme en commande optimale stochastique, la présence des perturbations, ou du second joueur, fait que l'information disponible évolue avec le temps, et que la possibilité donnée d'utiliser les observations passées change profondément la nature du problème par rapport à la boucle ouverte.

On sait qu'en commande stochastique, la classe des stratégies à considérer, et donc la difficulté du problème, est radicalement réduite par le principe d'optimalité dont la conséquence est que, si l'état du système est parfaitement connu, les purs feedbacks ou controles markoviens suffisent. L'objet de cet article est d'étudier dans quelle mesure cette situation s'étend à la recherche d'un point selle dans le cadre défini ici.

Nous allons donc définir une classe aussi large que possible de stratégies causales, supposer que le problème a une solution dans cette classe, et étudier si on peut en déduire qu'il existe une solution en pur feedback.

1 - Le système

Le problème concerne un système dynamique d'état $x \in X$, gouverné par les commandes $u \in U$ et $v \in V$ du minimiseur et du maximiseur respectivement. On appelle commandes en boucle ouverte, ou lois de commande des fonctions du temps $t \in \mathbb{R}$: $u(.)$, et $v(.)$, et deux ensembles de lois de commandes admissibles sont donnés : $u(.) \in \Omega_u$, $v(.) \in \Omega_v$. Ces deux ensembles sont supposés clos pour l'opération de concaténation, c'est à dire que si $u_1(.) \in \Omega_u$ et $u_2(.) \in \Omega_u$, alors, quelque soit $\tau \in \mathbb{R}$,

$$(1) \quad u_3(.) : t \to u_3(t) = \begin{cases} u_1(t) & \text{si } t < \tau \\ u_2(t) & \text{si } t \geq \tau \end{cases}$$

appartient encore à Ω_u.

Dans le future, chaque fois qu'aucune confusion n'est possible, nous noterons simplement u et v des lois de commandes.

L'évolution du système est donnée par une fonction de transition

$$\chi : X \times \mathbb{R} \times \Omega_u \times \Omega_v \times \mathbb{R} \to X$$
$$x(t) = \chi(x_o, t_o; u, v; t)$$

satisfaisant les axiomes habituels :

$$(2a) \quad \forall (u,v) \in \Omega_u \times \Omega_v , \quad \chi(x_o, t_o; u, v; t_o) = x_o$$

$$(2b) \quad \forall (u,v) \in \Omega_u \times \Omega_v , \quad \forall \tau, t \quad \chi(\chi(x_o, t_o; u, v; t), t; u, v; \tau) = \chi(x_o, t_o; u, v; \tau)$$

(2c) si $u_1(t) = u_2(t)$ et $v_1(t) = v_2(t)$, pour presque tout $t \leq \tau$,

alors

$$\chi(x_o,t_o;u_1,v_1;\tau) = \chi(x_o,t_o;u_2,v_2;\tau)$$

(On dit que la relation $u,v, \to x(.)$ est causale).

NOTATIONS. On aura constamment besoin par la suite des notations suivantes. La valeur du critère défini ci-dessous ne dépendra souvent que de la restriction de $u(.)$ et $v(.)$ à un intervalle $[t_o,t_1]$. Dans ce cas, $u(.)$ et $v(.)$ n'ont pas besoin d'être définis en dehors de cet intervalle, mais nous gardons les mêmes notations. Par contre, soit $\sigma < \tau$ deux instants distincts, nous noterons

u_σ^τ la restriction de u à (σ,τ), de même pour v_σ^τ

et nous omettrons les indices, inférieur $-\infty$ ou t_o, et supérieur t_1 ou $+\infty$.

De plus l'opération de concaténation définie par (1) sera notée &, de sorte que (1) s'écrira

$$u_3 = u_1^\tau \ \& \ u_{2\tau}$$

De même, pour une application φ de Ω_v dans Ω_u, la notation

$$\varphi = \varphi_1^\tau \ \& \ \varphi_{2\tau}$$

désigne l'application qui à $v \in \Omega_v$ fait correspondre la concaténation de la restriction de $\varphi_1(v)$ à $[t_o,\tau)$ et de la restriction de $\varphi_2(v)$ à $[\tau,t_1)$.

Enfin, par abus de notation, nous écrivons $u^\tau \in \Omega_u$, $v^\tau \in \Omega_v$, et la causalité exprimée par (2b) s'écrira encore

(3) $\quad x(\tau) = \chi(x_o, t_o; u^\tau, v^\tau; \tau).$

Un <u>critère</u> J est donné :

$$J : X \times \mathbb{R} \times \Omega_u \times \Omega_v \rightarrow \mathbb{R} \cup \{-\infty\} \cup \{+\infty\} ,$$

qu'on suppose <u>additif</u> dans le sens suivant :

$$\forall \tau > t_o \ , \ \exists J^\tau : X \times \mathbb{R} \times \Omega_u \times \Omega_v \rightarrow \mathbb{R} \cup \{-\infty\} \cup \{+\infty\}$$

tel que

(4) $\quad \forall x_o, t_o; u, v, \quad J(x_o, t_o; u, v) = J^\tau(x_o, t_o; u^\tau, v^\tau) + J(\chi(x_o, t_o; u^\tau, v^\tau; \tau), \tau; u_\tau, v_\tau)$

Pour le théorème 3, on supposera en outre qu'est donné un sous ensemble C de $X \times \mathbb{R}$, appelé <u>cible</u>, qui intervient dans la définition de J par le fait que t_1 dépend de x_o, t_o, u, v et est le premier instant tel que

$$(x(t_1), t_1) \in C$$

et $\quad J = +\infty \quad$ si $\quad x_o, t_o, u, v$ sont tels qu'un tel t_1 fini n'existe pas.

Ce formalisme recouvre le cas classique d'un système différentiel dans $\mathbb{R}^n$ muni d'un critère intégral et éventuellement d'une cible. Il recouvre aussi d'autre cas, quite, si nécessaire, à limiter la fonction de transition à des évolutions en temps croissant. Cependant ce n'est pas par goût de généraliser que nous choisissons cette forme un peu abstraite, mais pour simplifier l'analyse en ne gardant que les éléments utiles pour notre propos, et en éludant les problèmes techniques liés a un formalisme différentiel par exemple.

Il est important de remarquer que dans le cas où le problème s'arrête quand l'état a atteint une cible, les spécifications $J = +\infty$, ou $J = -\infty$, si $t_1 = \infty$ sont cohérentes avec (4), alors que si on pose dans ce cas J = une valeur finie fixée, (4) n'est plus satisfait.

L'hypothèse d'additivité de J est un peu plus forte que nécessaire, comme l'on montré plusieurs auteurs , mais cela ne nous intéresse pas ici.

2 - Stratégies et point selle.

Notre objectif est de savoir dans quelle mesure des stratégies plus générales que des purs feedbacks sont nécessaires pour exhiber un point-selle. Il nous faut donc choisir a priori des classes très larges de stratégies. Suivant ROXIN [1] et VARAYIA [1], une stratégie φ pour le premier joueur sera une application de $X \times R \times \Omega_v$ dans Ω_u :

$$u(.) = \varphi(x_o, t_o; v(.))$$

ayant la propriété de causalité.

(5) $$u^\tau = \varphi^\tau(x_o, t_o; v^\tau).$$

De même, une stratégie ψ pour le deuxième joueur est une application causale de $X \times R \times \Omega_u$ dans Ω_v.

On suppose que des ensembles Φ et Ψ de stratégies causales pour le premier et le deuxième joueur respectivement, ont été choisis, qui satisfont les conditions suivantes :

(6a) Φ et Ψ contiennent les lois de commandes en boucle ouverte, en ce sens que quelle que soit la fonction $\tilde{u}(x_o,t_o;t)$ telle que $\forall x_o,t_o$, $\tilde{u}(x_o,t_o;.) \in \Omega_u$, la stratégie

$$\tilde{\varphi} : (x_o,t_o;v) \rightarrow \tilde{u}(x_o,t_o;.)$$

est admissible. De même pour ψ

(6b) Tout couple de stratégies admissibles engendre une trajectoire et une seule. C'est a dire que $\forall(\varphi,\psi) \in \Phi \times \Psi$, $\forall x_o,t_o$, le système d'équations

$$u = \varphi(x_o,t_o;v)$$
$$v = \psi(x_o,t_o;u)$$

admet toujours un point fixe et un seul. On dit que la paire $(u,v) \in \Omega_u \times \Omega_v$ est <u>engendrée</u> par la paire $(\varphi,\psi) \in \Phi \times \Psi$.

REMARQUE : Soit $\bar{\varphi},\bar{\psi}$ une paire de stratégies admissible, (x_o,t_o) une phase initiale, $(\bar{u},\bar{v})$ la paire de commande engendrée. Au sens de (6a), $\bar{v}$ est une stratégie admissible (dans Ψ). Alors, $(\bar{u},\bar{v})$ est encore la paire de commandes engendrée par $(\bar{\varphi},\bar{v})$. Elle satisfait en effet les équations de point fixe de $(\bar{\varphi},\bar{v})$:

$$\bar{u} = \bar{\varphi}(\bar{v})$$
$$\bar{v} = \bar{v}$$

dont la solution est par hypothèse unique.

Par abus de notation, nous utiliserons les écritures $J(x_o,t_o;\varphi,\psi)$ ou $J(x_o,t_o;\varphi,v)$, qui sont non ambiguës.

On s'intéresse au cas où le problème présente un point selle. C'est à dire qu'il existe une paire de stratégies $\varphi^*(x_o,t_o;v)$ et $\psi^*(x_o,t_o;u)$ telle que $\forall(\varphi,\psi) \in \Phi \times \Psi$,

$$(7) \quad J(x_o,t_o;\varphi^*,\psi) \leq J(x_o,t_o;\varphi^*,\psi^*) = V(x_o,t_o) \leq J(x_o,t_o;\varphi,\psi^*).$$

(Les quantités φ^*,ψ^* et V ne sont pas nécessairement définies ailleurs qu'en x_o,t_o).

Le résultat ci-dessous est une conséquence de (6a) et est désormais classique.

PROPOSITION. La relation (7) est équivalente à

$$(8) \quad \forall(u,v) \in \Omega_u \times \Omega_v , \quad J(x_o,t_o;\varphi^*,v) \leq J(x_o,t_o;\varphi^*,\psi^*) \leq J(x_o,t_o;u,\psi^*).$$

3 - Résultats

Nous définissons la procédure de mise à jour de la stratégie φ^*. soit (x_o,t_o) une phase initiale pour laquelle un point selle existe. Soit $\hat{\varphi}$ la stratégie ainsi définie :

soit $\tau > t_o$ un instant tel qu'il existe un point selle depuis la phase

$$(x(\tau),\tau) = (\chi(x_o,t_o;\varphi^*(x_o,t_o;v),v;\tau),\tau)$$

alors

$$(8) \qquad \hat{\varphi}(x_o,t_o;v) = \varphi^{*\tau}(x_o,t_o;v^\tau) \ \& \ \varphi^*(x(\tau),\tau;v_\tau)$$

THEOREME 1. La stratégie mise à jour (8) est optimale.

DEMONSTRATION. Montrons qu'on a nécessairement l'inégalité

$$(9)\quad J^{\tau}(x_o,t_o;\varphi^{*\tau}(x_o,t_o;v^{\tau}),v^{\tau}) + V(x(\tau),\tau) \leq V(x_o,t_o)$$

Considérons pour v la stratégie

$$\bar{\psi}(x_o,t_o;u) = v^{\tau} \;\&\; \psi^*(x(\tau),\tau;u_{\tau}),$$

avec le même v^{τ} que précédemment. Quelque soit u_{τ}, on aura

$$J(x(\tau),\tau;u_{\tau},\bar{\psi}_{\tau}) \geq V(x(\tau),\tau)$$

En ajoutant $J^{\tau}(x_o,t_o;\varphi^{*\tau},\psi^{\tau})$ aux deux membres, en utilisant (4) et en choisissant pour u_{τ} la commande engendrée par $\varphi^*(x_o,t_o,.)$ il vient

$$J(x_o,t_o;\varphi^*,\bar{\psi}) \geq V(x(\tau),\tau)$$

d'où (9) en tenant compte des inégalités de point selle (7).

Vérifions alors que $\hat{\varphi}$ est optimale, en utilisant la définition (8).

$$J(x_o,t_o;\hat{\varphi},v) = J^{\tau}(x_o,t_o;\varphi^*(x_o,t_o;v^{\tau}),v^{\tau}) + J(x(\tau),\tau;\varphi^*(x(\tau),\tau;v_{\tau}),v_{\tau})$$

En utilisant l'inégalité de point selle pour le dernier terme ci-dessus on obtient

$$J(x_o,t_o;\hat{\varphi},v) \leq J^{\tau}(x_o,t_o;\varphi^*(x_o,t_o;v^{\tau}),v^{\tau}) + V(x(\tau),\tau)$$

et en utilisant (9) le résultat désiré.

On aurait naturellement le théorème symétrique pour v.

COMMENTAIRE 1. Ce théorème est semblable au Principe d'Optimalité de Bellman, en ce qu'il dit qu'à l'instant τ où il "remarque" que l'autre joueur l'a fait passer par $x(\tau)$ d'où il connait un point selle, le premier joueur peut "oublier le passé" et jouer d'après le point selle de $(x(\tau),\tau)$. Toutes fois, la très grande différence avec ce qui se passe en théorie classique de la commande est qu'on ne peut pas affirmer que φ^*_τ constitue nécessairement une stratégie optimale pour le jeu commençant en $(x(\tau),\tau)$, quelque soit τ.

COMMENTAIRE 2. Un corollaire trivial du théorème 1 est que, par induction, on peut remettre sa stratégie à jour un nombre arbitraire mais fini de fois. Cependant, on ne peut pas en déduire qu'une stratégie remise à jour un nombre infini de fois serait optimale. Nous donnerons un contre exemple.

Le caractère nécessaire de l'optimalité de φ^*_τ quelque soit τ ne peut être établi que dant le cas où les deux joueurs utilisent leur commande optimale, hypothèse particulièrement irréaliste pour la commande dans le cas le plus défavorable.:

THEOREME 2. S'il existe un point selle pour la phase initiale x_o,t_o, alors, quelque soit $\tau > t_o$, les restrictions $\varphi^*_\tau(x_o,t_o;.)$, $\psi^*_\tau(x_o,t_o;.)$ des stratégies optimales constituent un point selle pour le jeu à phase initiale $(x(\tau),\tau) = (\chi(x_o,t_o;\varphi^*,\psi^*;\tau),\tau)$.

DEMONSTRATION. Soit u^*,v^* les commandes engendrées par φ^*,ψ^* depuis x_o,t_o. Considérons les stratégies ci-dessous, pour $x(\tau) = \chi(x_o,t_o;\varphi^*,\psi^*;\tau)$:

$$\hat{\varphi}(x(\tau),\tau;v_\tau) = \varphi^*(x_o,t_o;v^{*\tau} \;\&\; v_\tau),$$
$$\hat{\psi}(x(\tau),\tau;u_\tau) = \psi^*(x_o,t_o;u^{*\tau} \;\&\; u_\tau).$$

D'après (4), on a

$$J(x(\tau),\tau;\hat{\varphi},\hat{\psi}) = V(x_o,t_o) - J^{\tau}(x_o,t_o;u^{*\tau},v^{*\tau}).$$

Puis regardons, pour v_τ quelconque,

$$J(x(\tau),\tau;\hat{\varphi},v_\tau) = J(x_o,t_o;\varphi^*(x_o,t_o;v^{*\tau} \ \& \ v_\tau),v^{*\tau} \ \& \ v_\tau) - J^{\tau}(x_o,t_o;u^{*\tau},v^{*\tau}).$$

Et en utilisant l'inégalité de point selle pour $J(x_o,t_o)$, on a le résultat souhaité. Il en va de même pour $J(x(\tau),\tau;u_\tau,\hat{\psi})$.

COROLLAIRE 1. Si le point selle est unique depuis toute phase initiale, les trajectoires bi-optimales (i.e. engendrées par φ^*,ψ^*) sont aussi engendrées par de purs feedbacks $\bar{\varphi}(x,t),\bar{\psi}(x,t)$:

$$(10) \qquad \begin{aligned} \bar{\varphi}(x,t) &= \varphi^*(x,t;\psi^*)(t) \\ \bar{\psi}(x,t) &= \psi^*(x,t;\varphi^*)(t) \end{aligned}$$

(De plus, si le point selle n'est pas unique en un nombre fini de point sur toute trajectoire bioptimale, tout choix définissant un tel $\bar{\varphi},\bar{\psi}$ est encore optimal, d'après le théorème 1). Toutefois, si $\bar{\varphi},\bar{\psi}$ engendrent bien les trajectoires bioptimales, et permettent donc de calculer V, ce ne sont pas nécessairement des stratégies optimales. L'exemple donné plus loin montre ce phénomène.

REMARQUE. Le corollaire démontre l'existence d'un point fixe aux équations :

$$u(t) = \varphi^*(\chi(x_o,t_o;u,v;t),t;v)$$
$$v(t) = \psi^*(\chi(x_o,t_o;u,v;t),t;u)$$

c'est à dire, dans le cas d'un système différentiel par exemple, l'existence d'une solution à l'équation différentielle initialisée en x_o,t_o, avec les commandes en feedback (10).

Dans l'exemple ci-dessous, les difficultés proviennent du fait que le jeu peut ne pas se terminer. Nous allons montrer que, dans une certaine mesure, pour un jeu a durée limitée à priori, de purs feedbacks suffisent toujours.

Soit $[t_o,t_1]$ un intervalle tel que J ne dépend que de $u_{t_o}^{t_1}$ et $v_{t_o}^{t_1}$, notés u et v.

Soit $\{h_n\}$ une suite décroissante d'entiers positifs, tendant vers zéro.

HYPOTHESE : Les stratégies φ_n et ψ_n "remises à jour" tous les h_n sont définies (depuis x_o,t_o), et $\forall u \in \Omega_u$, $\forall v \in \Omega_v$,

$$(11) \qquad \begin{aligned} J(x_o,t_o;\varphi_n,v) &\to J(x_o,t_o;\bar{\varphi},v) \\ J(x_o,t_o;u,\psi_n) &\to J(x_o,t_o;u,\bar{\psi}) \end{aligned}$$

On a alors, trivialement, le résultat ci-dessous.

COROLLAIRE 2. Pour un jeu à durée finie, et sous l'hypothèse (11), les stratégies en feedback $(\bar{\varphi},\bar{\psi})$ définies en (10) constituent un point selle pour le jeu à phase initiale (x_o,t_o).

DEMONSTRATION. D'après le théorème 1, φ_n et ψ_n étant remises à jour un nombre fini de fois, on a pour tout n

$$J(x_o,t_o;\varphi_n,v) \leq V(x_o,t_o) \leq J(x_o,t_o;u,\psi_n)$$

d'où le résultat par application de (11).

On vérifie que pour un jeu défini par un système différentiel et un critère intégral, l'hypothèse (11) est satisfaite par exemple si

i) $\forall(u,v) \in U \times V, \quad \|\dot{x}\| < k(1+\|x\|)$

(l'ensemble des trajectoires est compact en convergence uniforme sur tout intervalle)

ii) Il existe un ouvert $\mathcal{Z}$ de $X \times [t_o,t_1]$ tel que le point selle est défini depuis toute phase de $\mathcal{Z}$, et en outre, $\forall(x,t) \in \mathcal{Z}$, $\forall(u,v) \in \Omega_u \times \Omega_v$, $\forall\tau \in (t,t_1)$,

$$\chi(x,t;\varphi^*(x,t;v),v,\tau) \in \mathcal{Z}$$

$$\chi(x,t;\ ,\psi^*(x,t;u);\tau) \in \mathcal{Z}$$

iii) Il existe des variétés différentiables de $\mathcal{Z}$, transverses à toutes trajectoire engendrée par un couple (φ^*,v) ou (u,ψ^*), telles qu'en dehors de tout voisinage de ces variétés il existe un $\varepsilon > 0$ (uniforme en x,t) tel que $\varphi^*(x,t;v)$ et $\psi^*(x,t;u)$ soient des fonctions continues du temps sur l'intervalle $(t,t+\varepsilon)$.

En effet, la condition ii) garanti que les trajectoires engendrées par (φ_n,v) et (u,ψ_n) existent et restent dans $\mathcal{Z}$, et i) et iii) garantissent que les trajectoires correspondantes convergent uniformément sur $[t_o,t_1]$ vers celles engendrées par $(\bar{\varphi},v)$ et $(u,\bar{\psi})$.

Montrons enfin que la méthode de la programmation dynamique est justifiée, quand elle réussit. Nous supposons ici l'existence d'une cible C (qui peut être de la forme $X \times \{T\}$, ou au contraire impliquer des conditions sur $(x(t_1),t_1)$).

THEOREME 3. Soit X un espace vectoriel normé dans lequel toutes les trajectoires $(x_o,t_o;u,v,.)$ sont continues et S un sous ensemble fermé de $X \times R$, le séparant en deux parties connexes, l'une contenant (x_o,t_o) et l'autre C. Si, pour toute phase initiale de S le jeu a un point selle, l'existence d'un point selle au jeu de phase initiale (x_o,t_o) est équivalente à l'existence d'un point selle au jeu ayant même phase initiale et même dynamique, et pour critère :

$$\tilde{J}(x_o,t_o;u,v) = J^\tau(x_o,t_o;u^\tau,v^\tau) + V(x(\tau),\tau) \quad \text{si} \quad \tau < \infty\ , \quad \tilde{J} = +\infty \quad \text{sinon,}$$

où τ est le premier instant tel que $(x(\tau),\tau) \in S$, où

$$x(\tau) = \chi(x_o,t_o;u,v;\tau).$$

En outre, les jeux J et $\tilde{J}$ ont la même valeur, et quand les quantités ci-dessous existent,

$$(12) \qquad \begin{aligned} \varphi^*(x_o,t_o;v) &= \tilde{\varphi}(x_o,t_o;v^\tau)\ \&\ \varphi^*(x(\tau),\tau;v_\tau) \\ \psi^*(x_o,t_o;u) &= \tilde{\psi}(x_o,t_o;u^\tau)\ \&\ \psi^*(x(\tau),\tau;u_\tau). \end{aligned}$$

où $\tilde{\varphi}$ et $\tilde{\psi}$ désignent les stratégies optimales du jeu $\tilde{J}$.

DEMONSTRATION. Montrons d'abord que si le jeu $\tilde{J}$ a un point selle, avec des stratégies optimales $\tilde{\varphi}$ et $\tilde{\psi}$ et une valeur $\tilde{V}$, les stratégies (12) forment un point selle du jeu initial.

S'il existe une stratégie $\tilde{\psi}$ telle que l'état ne rencontre jamais S, alors $(\varphi,\tilde{\psi})$ est un point selle dégénéré pour tout φ, avec valeur $+\infty$. Dans ce cas, $\tilde{\psi}$ & v_τ est l'argument d'un point selle dégénéré du jeu global, quelque soit v_τ, puisqu'a fortiori l'état n'atteindra jamais C.

Dans le cas contraire, $\forall v \in \Omega_v$, la paire $(\tilde{\varphi},v)$ engendre un instant τ fini, et

$$J(x_o,t_o;\tilde{\varphi} \,\&\, \varphi^*,v) = J^\tau(x_o,t_o;\tilde{\varphi},v^\tau) + J(x(\tau),\tau;\varphi^*(x(\tau),\tau;v_\tau),v_\tau)$$

soit, en appliquant l'inégalité de point selle définissant $\varphi^*(x(\tau),\tau,.)$

$$J(x_o,t_o;\tilde{\varphi} \,\&\, \varphi^*,v) \leq J\,(x_o,t_o;\tilde{\varphi},v^\tau) + V(x(\tau),\tau) = \tilde{J}(x_o,t_o;\tilde{\varphi}(x_o,t_o;v^\tau),v^\tau)$$

et par définition de $\tilde{\varphi}$, $J(x_o,t_o;\tilde{\varphi} \,\&\, \varphi^*,v) \leq \tilde{V}(x_o,t_o)$.

De même, pour $u \in \Omega_v$, soit la paire $(u,\tilde{\psi})$ donne $\tau = \infty$, soit $\tilde{J} = +\infty$, et a fortiori $J = +\infty > \tilde{V}(x_o,t_o)$ pour la stratégie $\tilde{\psi} \,\&\, \psi^*$ de (12), soit $\tau < \infty$, et on a le raisonnement symétrique du précédent. Donc (12) constitue bien un point selle du jeu d'origine, avec $\tilde{V}(x_o,t_o)$ pour valeur.

Montrons maintenant que si le jeu global a un point selle, le jeu $\tilde{J}$ en a un. Si le point selle global est dégénéré. $J(x_o,t_o;u,\psi^*) = +\infty$, et qu'en outre ψ^* empêche l'état d'atteindre S, alors $\tilde{\varphi} = \psi^{*\tau}$ est un argument d'un point selle dégénéré du jeu $\tilde{J}$, avec la même valeur $+\infty$, Si non, posons

$$\tilde{\varphi}(x_o,t_o;v) = \varphi^{*\tau}(x_o,t_o;v^\tau) \quad \text{et} \quad \tilde{\psi}(x_o,t_o;u) = \psi^{*\tau}(x_o,t_o;u^\tau)$$

(Remarquons que dans ces notations, τ dépend de v dans $\tilde{\varphi}$ et de u dans $\tilde{\psi}$ et peut être $+\infty$ dans le dernier). Le résultat découle alors de l'application du théorème 1. L'inégalité (9) donne directement l'inégalité de point selle pour le jeu J, et une paire $(\tilde{\varphi},v)$ quelconque. De même, pour une paire $(u,\tilde{\psi})$, soit $\tau = +\infty$, et $\tilde{J} = +\infty > V(x_o,t_o)$, soit τ est fini et un calcul en tous points analogue donne a nouveau la deuxième inégalité de point selle.

COMMENTAIRE 3. A nouveau comme pour le théorème de mise à jour, ce théorème ressemble à ce qui se passe en théorie classique de la commande. à cette différence essentielle près que l'existence du point selle au jeu à phase initiale dans S doit être supposée a priori et ne peut être déduite de l'existence du point selle global.

4 - Exemple

Les résultats et certaines difficultés mentionnés ci-dessus sont mis en évidence par l'exemple ci-dessous. Cet exemple est voisin d'un problème proposé par ELAQUIERE et al., dans un autre contexte, et également d'un problème discret de MODEL' [1].

L'état et les commandes sont scalaires. La dynamique est

$$\dot{x} = u + v \qquad -1 \leq u \leq 0, \qquad -1 \leq v \leq 0, \qquad x_o > 0, \quad t_o = 0.$$

La cible est $x \leq 0$, et le critère

$$J = \int_{t_o}^{t_1} (1-x)dt \qquad \text{si} \qquad t_1 \text{ fini}, \qquad J = +\infty \qquad \text{si } t_1 = \infty.$$

L'utilisation des méthodes classiques (BREAKWELL [1]) permettent de trouver des trajectoires candidates optimales et des feedbacks associés :

$$\hat{\phi}(x) = \begin{cases} 0 & \text{si} \quad x > 1 \\ \\ -1 & \text{si} \quad x \leq 1 \end{cases}, \qquad \hat{\psi}(x) = \begin{cases} -1 & \text{si} \quad x > 1 \\ \\ 0 & \text{si} \quad x \leq 1 \end{cases}$$

$$V(x) = x(1 - \frac{x}{2})$$

Toutefois, $\hat{\phi}$ n'est pas une stratégie optimale, car pour $x_0 > 1$, on a

$$J(x_o, t_o; \hat{\phi}, 0) = + \infty > V(x_o).$$

On peut néanmoins exhiber un pont selle en stratégies causales (du à J. LEWIN [1])

La connaissance de $v(.)$ pour les instants passés permet d'évaluer à chaque instant

$$J^1(x_o, t_o; u, v) = \int_{t_o}^{t} (1 - x(s))ds.$$

Le point selle proposé par J. LEWIN est alors

$$\phi^*(x_o, t_o; v)(t) = \begin{cases} 0 & \text{si} \quad x(t) > 1 \quad \text{et} \quad J^1(x_o, t_o; o, v)(t) > - \dfrac{(x_o - 1)^2}{2} \\ \\ -1 & \text{si} \quad x(t) \leq 1 \quad \text{ou} \quad J^1(x_o, t_o; 0, v)(t) \leq - \dfrac{(x_o - 1)^2}{2} \end{cases}$$

$$\psi^{*}(x_o, t_o; u)(t) = \begin{cases} -1 & \text{si } x(t) > 1 \\ \\ 0 & \text{si } x(t) \leqslant 1 \end{cases} = \hat{\psi}(x).$$

PROPOSITION. La paire $(\phi \quad \psi)$, ci-dessus constitue un point selle du jeu.

PREUVE. Nous esquissons la preuve qui est très facile :

i) la condition $J^1 < -1/2(x_o-1)^2$ est satisfaite avant ou à l'instant $t_2 = t_o + x_o - 1$. En effet le problème de maximiser $J^1(t_2)$ en v, avec $u = 0$, a pour solution évidente $v = -1$, $J^1(t_2) = -1/2(x_o - 1)^2$. Donc, contre toute autre politique, cette valeur est atteinte plus tôt.

ii) Donc le jeu se termine en un temps fini, (inférieur à $t_2 + x_o$)

iii) Le coût encouru de l'instant de commutation à la fin du jeu est inférieur ou égal à 1/2. En effet, tant que $x > 1$, l'intégrande est négatif, et le coût encouru de $x = 1$ à $x = 0$ avec $u = -1$ est maximisé par $v = 0$ et vaut 1/2. Pour $u > -1$, $v = 0$, il est supérieur à 1/2.

REMARQUE 3 : Ce point selle n'est pas unique. Le point i) ci-dessus montre qu'on peut remplacer les stratégies ci-dessus par les lois de commande en boucle ouverte

$$u^\circ(x_o,t_o;t) = \begin{cases} 0 & \text{si } x_o > 1 \text{ et } t < t_o + x_o - 1 \\ -1 & \text{si } x_o \leqslant 1 \text{ ou } t > t_o + x_o - 1 \end{cases}$$

$$v^\circ(x_o,t_o;t) = -u^\circ(x_o,t_o;t).$$

(En fait la condition sur x_o est redondante et peut être omise).

De même, on peut reprendre les stratégies ϕ^* et ψ^*, et y remplacer $-1/2(x_o-1)^2$ par toute fonction $W(x_o)$ inférieure à celle-là. A l'opposé, en tenant compte du gain de "U" entre $x > 1$ et $x = 1$ même si $\dot{x} = -2$, on voit qu'on peut aussi prendre comme fonction de commutation

$$J^1 < -\frac{(x_o - 1)^2}{2} + \frac{(x(t) - 1)^2}{4}.$$

REMARQUE 4 : Si "U" et "V" jouent leurs stratégies optimales, l'une quelconque de celles proposées, les commandes engendrées coincident avec celles engendrées par $(\hat{\phi},\hat{\psi})$, qui ont été calculées à l'aide de l'équation d'Hamilton-Jacobi (en fait ce sont u° et v° qui sont calculés. Ils synthétisent $\hat{\phi}$ et $\hat{\psi}$). Donc, le corollaire 1.

et les théorèmes ci-dessus sont bien vérifiés ici. Mais on voit que si "U" décide de "remettre à jour" sa stratégie ϕ^{**} tous les δt fixé, inférieur à $1/2(x_o - 1)$, contre $v = 0$ il obtiendra à nouveau $J = \infty$. Par contre, s'il limite à priori le nombre de remises à jour, il a de nouveau un comportement optimal.

REMARQUE. En introduisant un variable d'état suplémentaire

$$\dot{y} = 1 - x,$$

et en écrivant le coût comme purement final :

$$J = y(t_2),$$

on a un jeu équivalent, dont la valeur est

$$V = y_c + x_o(1 - \frac{x_o}{2}) .$$

En corrigeant ϕ , ci-dessus en ajcutant y_o à la fonction de commutation, cn obtient un point selle où les stratégies optimales ne dépendent que de l'état courant et de l'état initial. (C'est-à-dire un feedback non uniformément optimal dans la terminologie de MODEL'). Mais on ne peut obtenir de stratégie optimale ne dépendant que de l'état courant. (Ce fait est caché si, comme on fait couramment, on pose à priori $y(t_o) = 0$, et on omet de faire varier cette composante de l'état initial.)

CONCLUSION.

On peut justifier la méthode de la programmation dynamique quand elle réussit, c'est à dire quand des points selles existent pour les phases "intermédiaires" entre la phase initiale et la cible. Toutefois nous n'avons pu démontrer que tel était toujours le cas. De même dans une telle situation, on a montré que la valeur pouvait être calculée en reconstruisant les trajectoires bi-optimales à l'aide de purs feedbacks. Mais l'utilisation de ces feedbacks comme stratégies revient à remettre à jour sa stratégie continument, donc un nombre infini de fois, ce qui peut être non optimal comme un exemple l'a montré. Il semble bien que cette pathologie soit liée à la possibilité qu'a le jeu de ne pas se terminer. Cependant, sauf à adopter une définition plus générale de la solution d'une équation différentielle (cf. KRASSOVSKI [1]) les hypothèses sous lesquelles nous savons montrer ce résultat sont impossibles à tester a priori.

BIBLIOGRAPHIE

E. ROXIN [1] "On Varaiya's definition of a differential Game". US.Japan seminar on differential and functional equations. Mineapolis 1967.

P. VARAIYA [1] "The existence of a solution to a differential game". SIAM J^al on Control, 5, 1967.

B.I. MODEL' [1] "Existence of an overall ε-optimal strategy and validity of Bellman's functional equation for an extended class of dynamic processes". Engineering Cybernatics 13, 1975.

J.V. BREAKWELL [1] "Lectures on Differential Games" in Differential Games and Applications, P. Hagedorn et al. ed., Springer lecture notes on Information and Control Sciences, 3, 1977.

J. LEWIN [1] Communication personnelle. Twente University of Technology, 1977.

N.N. KRASSOVSKI et A.J. SOUBBOTINE. Jeux Differentiels. MIR. 1977.

ANNEXE : EQUATIONS DE TCHENTSOV - TCHISTYAKOV

Pour nous placer dans le cadre classique, nous commençons par faire l'extension évidente du concept de valeur du jeu en l'absence de point-selle, cohérente avec notre formalisme. (Cf. VARAIYA[1], ROXIN[1]) :

DEFINITION. On dit que le jeu a une valeur, notée $V(x_o,t_o)$, si

$$\inf_{\varphi \in \Phi} \sup_{v \in \Omega_v} J(x_o,t_o;\varphi(v),v) = \sup_{\psi \in \Psi} \inf_{u \in \Omega_u} J(x_o,t_o;u,\psi(u)) = V(x_o,t_o).$$

De cette définition, on déduit l'existence de stratégies ε-efficaces :

$\forall \varepsilon > 0$, $\exists\, \varphi_\varepsilon \in \Phi$ et $\psi_\varepsilon \in \Psi$:

$$\forall v \in \Omega_v\ , \quad J(x_o,t_o;\varphi_\varepsilon(v),v) \leq V(x_o,t_o) + \varepsilon\ , \tag{A1}$$

$$\forall u \in \Omega_u\ , \quad V(x_o,t_o) - \varepsilon \leq J(x_o,t_o;u,\psi_\varepsilon(u)).$$

Par commodité, nous introduisons la notation suivante :

NOTATION.

$$J^\tau(x_o,t_o;u^\tau,v^\tau) + V(X(x_o,t_o;u^\tau,v^\tau;\tau),\tau) = W^\tau(x_o,t_o;u^\tau,v^\tau). \tag{A2}$$

On établit facilement l'équivalent du lemme 2 :

LEMME A2. Avec les définitions (A1) et (A2), quels que soient $(v,\tau) \in \Omega_v \times [t_o,t_1]$ tels que $V(X(x_o,t_o;\varphi_\varepsilon^\tau,v^\tau;\tau),\tau)$ existe, on a

$$W^\tau(x_o,t_o;\varphi_\varepsilon^\tau(v^\tau),v^\tau) \leq V(x_o,t_o) + \varepsilon \tag{A3}$$

et une inégalité symétrique avec (u,ψ_ε).

DEMONSTRATION. Supposons au contraire qu'il existe $\hat{v}$ et τ tels que

$$W^{\tau}(x_o,t_o;\varphi_\varepsilon^{\tau},\hat{v}^{\tau}) = V(x_o,t_o) + \varepsilon + \delta, \qquad \delta > 0.$$

Alors, il existe $\hat{\psi}_\tau = \psi_{\delta/2}(X(x_o,t_o;\varphi_\varepsilon^{\tau},\hat{v}^{\tau};\tau),\tau;.)$ tel que,

$$\forall u \in \Omega_u, \quad J(x(\tau),\tau;u_\tau,\hat{\psi}_\tau) \geq V(x(\tau),\tau) - \frac{\delta}{2}$$

(où $x(\tau)$ désigne toujours le même état).
Considérons alors pour V la stratégie $\psi(x_o,t_o;.) = \hat{v}^{\tau}$ & $\hat{\psi}_\tau$, et en utilisant la dernière inégalité avec u_τ engendré par $(\varphi_\varepsilon,\psi)$. On obtient alors par sommation des deux dernières relations une contradiction de la première de (A1).
On fait de même pour l'inégalité symétrique.
On en déduit immédiatement le lemme suivant :

LEMME (A1). Si pour τ donné, le jeu a une valeur pour tout état initial x, on a

$$\sup_{v \in \Omega_v} \inf_{u \in \Omega_u} W^{\tau}(x_o,t_o;u^{\tau},v^{\tau}) \leq V(x_o,t_o) \leq \inf_{u \in \Omega_u} \sup_{v \in \Omega_v} W^{\tau}(x_o,t_o;u^{\tau},v^{\tau}). \quad (A4).$$

DEMONSTRATION. L'inégalité (A3) donne immédiatement

$$\forall v \in \Omega_v, \quad \inf_{u \in \Omega_u} W^{\tau}(x_o,t_o;u^{\tau},v^{\tau}) \leq V(x_o,t_o)$$

d'où l'inégalité de gauche, celle de droite s'obtient de la même façon.
On en déduit finalement l'égalité de Tchentsov - Tchistyakov :

THEOREME. Si le jeu a une valeur pour toute phase initiale on a (avec la notation (A2)).

$$V(x_o,t_o) = \sup_{\tau} \sup_{v \in \Omega_v} \inf_{u \in \Omega_u} W^{\tau}(x_o,t_o;u^{\tau},v^{\tau}) = \inf_{\tau} \inf_{u \in \Omega_u} \sup_{v \in \Omega_v} W^{\tau}(x_o,t_o;u^{\tau},v^{\tau}).$$

DEMONSTRATION. On a, par définition, $\forall u, v \in \Omega_u \times \Omega_v$:

$$X(x_o,t_o;u,v;t_o) = x_o \quad , \quad J^{t_o}(x_o,t_o;u,v) = 0 \ ,$$

d'où $W^{t_o} = V(x_o,t_o)$.

REMARQUE 1. Le théorème de Tchentsov - Tchistyakov est établit "classiquement" pour un jeu à coût purement final, i.e. où $J^\tau \equiv 0$, donc où

$$W^\tau(x_o,t_o;u^\tau,v^\tau) = V(X(x_o,t_o;u^\tau,v^\tau,\tau),\tau).$$

REMARQUE 2. Le théorème d'origine est assorti d'une nouvelle définition de valeur inférieure et valeur supérieure, dont on voit immédiatement qu'elles sont encadrées par celles de FLEMMING (ou FRIEDMANN). Donc quand il existe une valeur en ce dernier sens, il en existe aussi une, qui coïncide, au sens de T & T. DANSKIN a montré que cette valeur coïncide avec celle de ROXIN que nous utilisons ici.

REMARQUE 3. Comme l'équation d'Isaacs, il s'agit d'une équation fonctionnelle sur la valeur. Elle a en moins qu'elle n'est pas instantanée (les inf et sup sont pris sur Ω_u et Ω_v, et non pour tout t sur U et V). De ce fait il n'est pas clair qu'elle ait une utilité... En revanche, elle est nécessaire, et ne fait que des hypothèses très faibles sur le jeu. (Aucune sur la régularité de V).