

J. P. QUADRAT

M. VIOT

**Méthodes de simulation en programmation
dynamique stochastique**

Revue française d'automatique, informatique, recherche opérationnelle. Mathématique, tome 7, n° 1 (1973), p. 3-22.

http://www.numdam.org/item?id=M2AN_1973__7_1_3_0

© AFCET, 1973, tous droits réservés.

L'accès aux archives de la revue « Revue française d'automatique, informatique, recherche opérationnelle. Mathématique » implique l'accord avec les conditions générales d'utilisation (<http://www.numdam.org/legal.php>). Toute utilisation commerciale ou impression systématique est constitutive d'une infraction pénale. Toute copie ou impression de ce fichier doit contenir la présente mention de copyright.

NUMDAM

Article numérisé dans le cadre du programme
Numérisation de documents anciens mathématiques
<http://www.numdam.org/>

METHODES DE SIMULATION EN PROGRAMMATION DYNAMIQUE STOCHASTIQUE

par J. P. QUADRAT et M. VIOT ⁽¹⁾

Résumé. — On démontre une propriété analogue, à la loi forte des grands nombres, pour les problèmes de programmation dynamique stochastique (avec contraintes) :

à toute suite de simulations indépendantes des paramètres aléatoires du système et à tout entier n , on associe un problème de programmation dynamique (déterministe), construit sur les n premières réalisations de la suite de simulations. On montre que les coûts minima et les stratégies optimales déterminées par ces problèmes (simulés), vont converger (presque sûrement) vers le coût minimum et la stratégie optimale du problème initial.

INTRODUCTION

On étudie des problèmes de programmation dynamique stochastique avec contraintes presque sûre, en probabilité ou en moyenne. La valeur Φ ou coût minimum, de tels problèmes, dépend de la mesure de probabilité μ , associée aux paramètres aléatoires (ou bruit) du système, ainsi que des bornes α des contraintes, à savoir $\Phi(\mu, \alpha)$.

On montre que, pour un choix convenable de l'ensemble des stratégies feed-back utilisées dans le contrôle du système, cette fonction Φ est semi-continue inférieurement (s.c.i.), pour la topologie de la convergence étroite sur les mesures.

De plus, on montre qu'elle possède une propriété de continuité, le long de certaines suites (μ_r, α_r) correctement choisies. C'est en particulier le cas pour les mesures discrètes (notées $\mu_r(\omega)$) issues de simulations indépendantes sur le bruit du système. Nous avons appelé cette propriété : loi forte des grands nombres en programmation dynamique stochastique. Elle conduit, entre autres, à des approximations numériques intéressantes dans le cas où la dimension du système ne permet pas l'utilisation des procédés classiques, ou encore la loi de probabilité du bruit est imparfaitement connue (cf. [5] et § 4).

(1) I.R.I.A. Département d'Informatique Numérique, Rocquencourt.

§ 1. SYSTEMES ET STRATEGIES

Pour tout $i = 0, 1, \dots, N$ (respectivement $i = 0, 1, \dots, N - 1$) soient n_i (respectivement m_i et p_i), des entiers ≥ 1 , et f_i , des applications continues de $R^{m_i} \times R^{m_i} \times R^{p_i}$ dans $R^{n_{i+1}}$. On considère un système dont l'état est donné par l'équation de récurrence :

$$(1) \quad x_{i+1} = f_i(x_i, u_i, e_i); \quad i = 0, 1, \dots, N - 1.$$

Les paramètres aléatoires du système seront les vecteurs

$$e = (x_0, e_0, e_1, \dots, e_{N-1})$$

de l'espace produit :

$$E = R^{n_0} \times \prod_{i=0}^{N-1} R^{p_i}.$$

Par la suite on désigne par μ une mesure de probabilité de Radon sur E .

La partie observable du système est définie de la manière suivante : on suppose que $n_i = n_i^{\text{ob}} + n_i^{\text{inob}}$ avec $n_i^{\text{ob}} \geq 0$; soit alors (x_0, x_1, \dots, x_N) une solution de l'équation (1) et x_i^{ob} , les n_i^{ob} premières composantes de x_i : le vecteur x_i^{ob} correspond à la partie observable du système à l'instant i . Lorsque $n_i^{\text{ob}} = 0$, le système est dit inobservable à l'instant i .

On note Σ_i , l'espace des applications continues de $R^{n_i^{\text{ob}}}$ dans R^{m_i} , muni de la topologie de la convergence compacte. Lorsque $n_i^{\text{ob}} = 0$, Σ_i s'identifie à R^{m_i} . Soit par ailleurs Σ l'espace produit

$$\prod_{i=1}^{N-1} \Sigma_i$$

(muni de la topologie produit) : un élément $S = (S_0, S_1 \dots S_{N-1})$ de Σ sera appelé stratégie.

Remarquons que Σ est métrisable séparable et que d'après le théorème d'Ascoli, une partie \mathcal{S} de Σ est relativement compacte si et seulement si — $\mathcal{S}_i = \text{proj}_{\Sigma_i} \mathcal{S}$ est faiblement borné dans Σ_i ; — \mathcal{S}_i est une partie équicontinue de Σ_i .

A une stratégie $S \in \Sigma$ et à un vecteur $e = (x_0, e_0, \dots, e_{N-1})$ de E , est associée une solution de l'équation (1), notée $X(S, e) = (X_0(S, e), X_1(S, e), \dots, X_N(S, e))$ par les relations :

$$(2) \quad \begin{cases} X_0(S, e) = x_0 \\ X_{i+1}(S, e) = f_i(X_i(S, e), S_i(X_i^{\text{ob}}(S, e)), e_i). \end{cases}$$

Une solution $X(S, e)$ est donc un élément de l'espace produit

$$F = \prod_{i=0}^N R^{n_i}.$$

REMARQUES

1. Lorsque le système est inobservable à tout instant

$$(n_i^{\text{ob}} = 0, i = 0, 1, \dots, N - 1),$$

les stratégies, obtenues ne sont autres que les stratégies open-loop.

2. En apparence, on se limite à des stratégies markoviennes (ne dépendant que de l'observation à l'instant considéré). En fait, il suffirait d'augmenter l'espace des états de l'équation (1), en gardant en mémoire tout ou une partie des états précédents, pour obtenir des stratégies non markoviennes.

Nous allons étudier maintenant certaines propriétés de l'application $X : (S, e) \rightarrow X(S, e)$ de $\Sigma \times E$ dans F qui a été définie en (2).

Proposition 1. *L'application $X : (S, e) \rightarrow X(S, e)$ est continue de $\Sigma \times E$ dans F .*

Démonstration.

L'espace Σ étant métrisable (et E de dimension finie), il suffit de montrer que pour toute suite (S_r, e^r) de $\Sigma \times E$ convergeant vers (S, e) , la suite $X(S_r, e^r)$ converge vers $X(S, e)$. En fait on va montrer par récurrence sur $i = 0, 1, \dots, N$ que l'on a :

$$\lim_r X_i(S_r, e^r) = X_i(S, e).$$

Or cela est vrai pour $i = 0$, car d'après (2) :

$$X_0(S_r, e^r) = x_0^r.$$

Supposons donc la propriété vérifiée pour i , et désignons par L_i^{ob} un compact de $R^{n_i^{\text{ob}}}$ contenant la suite convergente :

$$x_i^r = X_i^{\text{ob}}(S_r, e^r).$$

Comme la suite $S_{r,i}$ converge vers S_i uniformément sur tout compact, on a :

(3) $\forall \varepsilon > 0, \exists R(\varepsilon)$ tel que

$$r \geq R(\varepsilon) \Rightarrow \sup_{x \in L_i^{\text{ob}}} |S_{r,i}(x) - S_i(x)| \leq \varepsilon.$$

De la majoration :

$$|S_{r,i}(x_i^r) - S_i(x_i)| \leq |S_{r,i}(x_i^r) - S_i(x_i^r)| + |S_i(x_i^r) - S_i(x_i)|$$

et de la condition (3), on déduit que la suite $S_{r,i}(x_i^r)$ est convergente et de limite

$$S_i(x_i) = S_i(X_i^{\text{ob}}(S, e)).$$

La continuité des applications f_i entraînant finalement que :

$$\lim_r X_{i+1}(S_r, e^r) = X_{i+1}(S, e).$$

Corollaire. Soient \mathcal{S} un compact de Σ et K un compact de E , il existe alors un compact L de F tel que :

$$\forall (S, e) \in \mathcal{S} \times K; \quad X(S, e) \in L.$$

On en déduit en particulier que si μ est une mesure à support compact de E et si \mathcal{S} est une partie compacte de Σ , l'ensemble des vecteurs aléatoires :

$$\{ X(S, \cdot); S \in \mathcal{S} \} \text{ est une partie bornée de } L^\infty(E, \mu; F).$$

De manière plus générale, désignons par K_n une suite croissante de compacts de E telle que $\mu(\bigcup K_n) \leq \frac{1}{2^n}$ et posons :

$$\lambda_n = \max_{S \in \mathcal{S}, e \in K_n} |X(S, e)|.$$

Alors l'ensemble $\{ X(S, \cdot) \mid S \in \mathcal{S} \}$ formera une partie bornée de $L^p(E, \mu; F)$ à condition que :

$$\sum_n \frac{\lambda_n^p}{2^n} < +\infty.$$

Notons $\mathcal{M}_+^1(E)$ l'ensemble des mesures de probabilité de Radon sur E , muni de la topologie de la convergence étroite : c'est un espace métrisable de type dénombrable (cf. [2], chap. IX, § 5, n° 4, p. 62).

Dans la suite un rôle essentiel sera joué par les intégrales de la forme :

$$(4) \quad G(S, \mu) = \int_E g(X(S, e), S) d\mu(e)$$

où

- $X(S, e)$ est la solution de (2) associée à $S \in \Sigma$ et $e \in E$,
- g une fonction s.c.i. et bornée inférieurement sur $F \times \Sigma$,
- μ une mesure de $\mathcal{M}_+^1(E)$.

Proposition 2. Soit g une fonction continue sur $F \times \Sigma$ et \mathcal{S} une partie de Σ . On suppose qu'il existe une fonction ≥ 0 , μ -intégrable, ρ telle que :

$$(5) \quad \forall S \in \mathcal{S} : |g(X(S, e), S)| \leq \rho(e), \mu\text{-p.p.}$$

La fonction $S \rightarrow G(S, \mu)$ est alors continue sur \mathcal{S} .

Démonstration. Soit S_r une suite de \mathcal{S} convergeant vers un élément $S \in \mathcal{S}$. D'après la Proposition 1, on a :

$$\forall e \in E, \lim_{r \rightarrow \infty} g(X(S_r, e), S_r) = g(X(S, e), S).$$

De plus, la majoration (5) donne pour tout r :

$$|g(X(S_r, e), S_r)| \leq \rho(e), \mu\text{-p.p.}$$

Le résultat découle donc du théorème de convergence dominée de Lebesgue.

Proposition 3. Soit g une fonction s.c.i. et bornée inférieurement sur $F \times \Sigma$, et soit G la fonction définie sur $\Sigma \times \mathcal{M}_+^1(E)$ par (4).

Alors G est s.c.i. et bornée inférieurement. De plus G est continue bornée si g est continue bornée.

Démonstration. Supposons d'abord g continue bornée et posons

$$\tilde{g}(e, S) = g(X(S, e), S).$$

La fonction \tilde{g} définie sur $E \times \Sigma$ est également continue bornée :

$$\forall e \in E, \quad \forall S \in \Sigma \quad ; \quad |\tilde{g}(e, S)| \leq M.$$

Soit alors (S_r, μ_r) une suite de $\Sigma \times \mathcal{M}_+^1(E)$ convergeant vers (S, μ) ; on a :

$$\begin{aligned} & |G(S_r, \mu_r) - G(S, \mu)| \\ & \leq \left| \int \tilde{g}(e, S) d\mu_r(e) - d\mu(e) \right| + \left| \int \tilde{g}(e, S_r) - \tilde{g}(e, S) d\mu_r(e) \right|. \end{aligned}$$

Le premier terme du second membre de l'inégalité tend vers 0 par définition même de la convergence étroite.

Considérons maintenant un compact K de E tel que $\mu(E - K) \leq \varepsilon$ et soit L un ouvert relativement compact contenant K . Il existe $R_1(\varepsilon)$ tel que :

$$r \geq R_1(\varepsilon) \Rightarrow \mu_r(E - L) \leq \varepsilon.$$

Par ailleurs, comme \tilde{g} est continue, il existe $R_2(\varepsilon)$ tel que :

$$r \geq R_2(\varepsilon) \Rightarrow \sup_{e \in L} |\tilde{g}(e, S_r) - \tilde{g}(e, S)| \leq \varepsilon.$$

Donc pour $r \geq \max(R_1(\varepsilon), R_2(\varepsilon))$ on a

$$\left| \int \tilde{g}(e, S_r) - \tilde{g}(e, S) d\mu_r(e) \right| \leq 2M\mu_r(E - L) + \varepsilon\mu_r(L) \\ \leq (2M + 1)\varepsilon.$$

D'où la continuité de G lorsque g est continue bornée.

Supposons maintenant g s.c.i. et bornée inférieurement. Les espaces F et Σ étant métrisables, la fonction g est l'enveloppe supérieure d'une suite croissante (g_r) de fonctions continues bornées. Et le résultat découle de la relation :

$$G(S, \mu) = \sup_r \int_E g_r(X(S, e), S) d\mu(e).$$

EXEMPLES

1° Soit \mathcal{O} un ouvert de F , alors la fonction :

$$(S, \mu) \rightarrow \mu \{ e : X(S, e) \in \mathcal{O} \}$$

est s.c.i. sur $\Sigma \times \mathcal{N}_+^1(E)$.

Il suffit d'appliquer la proposition 3 à $g =$ fonction caractéristique de l'ouvert \mathcal{O} .

2° Soit B un fermé de F et α un scalaire, $0 \leq \alpha \leq 1$. La fonction :

$$(S, \mu) \rightarrow \alpha - \mu \{ e : X(S, e) \in B \}$$

est s.c.i. sur $\Sigma \times \mathcal{N}_+^1(E)$.

(Appliquer la proposition 3 à $g = \alpha - f$ caractéristique (B)).

3° Soient p_i ($i = 1, \dots, N$) des fonctions s.c.i. et bornées inférieurement sur $R^{n_i} \times R^{m_i-1}$; la fonction :

$$(S, \mu) \rightarrow \sum_{i=1}^N \int_E p_i[X_i(S, e), S_{i-1}(X_{i-1}^{ob}(S, e))] d\mu(e)$$

est également s.c.i. sur $\Sigma \times \mathcal{N}_+^1(E)$.

En conséquence les intégrales du type (4) englobent les principales formes de contraintes presque-sûre en probabilité ou en moyenne, de même que les principales formes de critères que l'on rencontre en programmation dynamique stochastique.

Soient donc g_0, g_1, \dots, g_K , $K + 1$ fonctions s.c.i. et bornées inférieurement sur $F \times \Sigma$, $\alpha = (\alpha^1 \dots \alpha^K)$ un point de R^K , μ une mesure de $\mathcal{N}_+^1(E)$ et S une

partie de Σ ; un problème de programmation dynamique stochastique s'écrit sous la forme :

$$(6) \quad P(\mu, \alpha) \begin{cases} \Phi(\mu, \alpha) = \inf_{S \in \mathcal{S}} G_0(S, \mu) \\ X_{i+1}(S, e) = f_i(X_i(S, e), S_i(X_i^{ob}(S, e), e_i)) \\ G_k(S, \mu) \leq \alpha_k \quad k = 1, \dots, K. \end{cases}$$

En particulier lorsque \mathcal{S} est une partie compact de Σ et que l'ensemble des stratégies admissibles est non vide, on est assuré de l'existence d'une stratégie optimale pour $P(\mu, \alpha)$ (cf. proposition 3).

Orientation générale.

Lorsque la mesure μ sur E est discrète : $\mu = \sum_{j=1}^r p_j \delta_{e^j}$, le programme (6) s'écrit simplement :

$$(6') \quad P(\mu, \alpha) \begin{cases} \Phi(\mu, \alpha) = \inf_{S \in \mathcal{S}} \sum_{j=1}^r p_j g_0(X(S, e^j), S) \\ X_{i+1}(S, e^j) = f_i(X_i(S, e^j), S_i(X_i^{ob}(S, e^j), e_i^j)) \\ \sum_{j=1}^r p_j g_k(X(S, e^j), S) \leq \alpha_k \quad , \quad k = 1, \dots, K. \end{cases}$$

On peut donc espérer le résoudre (au moins dans certains cas), par des méthodes purement déterministes.

Par ailleurs, on sait que l'ensemble des mesures discrètes est dense dans $\mathcal{M}_+^1(E)$. Si donc la valeur Φ était une fonction continue sur $\mathcal{M}_+^1(E)$ on aurait ainsi une méthode d'approximation simple des problèmes (6) par discrétisation de la mesure de probabilité des bruits du système.

En fait on ne peut espérer mieux qu'une semi-continuité inférieure globale pour la fonction Φ (théorème 1). Le travail principal consistera donc à trouver des conditions suffisantes de continuité de Φ le long de certaines suites (proposition 7) et à montrer que ces conditions suffisantes sont réalisées dans le cas particulièrement intéressant de la simulation indépendante des paramètres aléatoires du système (§ 3).

§ 2. QUELQUES PROPRIETES DE LA VALEUR Φ

Dans ce paragraphe \mathcal{S} désignera une partie compact de Σ .

Soit $\Gamma(\mu, \alpha)$, l'ensemble des stratégies admissibles du programme $P(\mu, \alpha)$:

$$(7) \quad \Gamma(\mu, \alpha) = \{ S \in \mathcal{S} : G_k(S, \mu) \leq \alpha_k, k = 1, \dots, K \}.$$

Proposition 4. *La multi-application Γ est semi-continue supérieurement de $\mathcal{N}_+^1(E) \times \mathbf{R}^k$ dans \mathcal{S} .*

Démonstration. Les fonctions G_k étant s.c.i. et \mathcal{S} compact, les $\Gamma(\mu, \alpha)$ sont des compacts.

Soit (S_r, μ_r, α_r) une suite de points du graphe de Γ , convergeant vers (S, μ, α) ; on a donc pour tout $k = 1 \dots K$:

$$G_k(S, \mu) \leq \liminf_r G_k(S_r, \mu_r) \leq \liminf_r \alpha_r = \alpha$$

D'où $S \in \Gamma(\mu, \alpha)$. En conséquence Γ est de graphe fermé, ce qui entraîne, \mathcal{S} étant compact, la s.c.s. de Γ ([1], chap. VI, cor. théor. 7).

Corollaire. *Soit (μ_r, α_r) une suite de $\mathcal{N}_+^1(E) \times \mathbf{R}^K$ convergeant vers (μ, α) . Alors :*

$$(8) \quad \overline{\liminf_r \Gamma(\mu_r, \alpha_r)} \subset \overline{\limsup_r \Gamma(\mu_r, \alpha_r)} \subset \Gamma(\mu, \alpha).$$

Démonstration. Il suffit de montrer que :

$$\limsup_r \Gamma(\mu_r, \alpha_r) \subset \Gamma(\mu, \alpha).$$

Dire que $S \in \limsup_r \Gamma(\mu_r, \alpha_r)$ entraîne qu'il existe une sous-suite (r_j) telle que $S \in \Gamma(\mu_{r_j}, \alpha_{r_j})$ pour tout r_j . D'où pour tout $k = 1, \dots, K$:

$$G_k(S, \mu) \leq \liminf_{r_j} G_k(S, \mu_{r_j}) \leq \liminf_{r_j} \alpha_{r_j} = \alpha.$$

Et on a bien $S \in \Gamma(\mu, \alpha)$.

Théorème 1. *La valeur $\Phi(\mu, \alpha)$ de $P(\mu, \alpha)$ est une fonction semi-continue inférieurement et bornée inférieurement sur $\mathcal{N}_+^1(E) \times \mathbf{R}^K$.*

Démonstration.

1° Posons $\Lambda = \{ (\mu, \alpha) \in \mathcal{N}_+^1(E) \times \mathbf{R}^K \mid \Gamma(\mu, \alpha) \neq \emptyset \}$ et montrons que Λ est fermé.

Soit $(\mu_r, \alpha_r) \in \Lambda$ convergeant vers (μ, α) et pour tout r , soit S_r un élément de $\Gamma(\mu_r, \alpha_r)$.

Comme \mathcal{S} est compact on peut toujours supposer la suite (S_r) convergente vers un $S \in \mathcal{S}$. La s.c.s. de Γ entraîne alors que $S \in \Gamma(\mu, \alpha)$. D'où $(\mu, \alpha) \in \Lambda$.

2° La restriction de Γ à Λ étant s.c.s. et non vide, la restriction de Φ à Λ est s.c.i. d'après [1] (chap. VI, § 3, Th. 2).

Et comme en dehors du fermé Λ , Φ est identique à $+\infty$, la fonction Φ est donc partout s.c.i.

Proposition 5. Soit (μ_r, α_r) une suite de $\mathcal{M}_+^1(E) \times \mathbf{R}^K$ convergeant vers (μ, α) et soit \tilde{S}_r une solution optimale du problème $P(\mu_r, \alpha_r)$. Supposons que

$$(9) \quad \liminf_r \Phi(\mu_r, \alpha_r) = \Phi(\mu, \alpha)$$

il existe alors une sous-suite de (\tilde{S}_r) convergeant vers une solution optimale de $P(\mu, \alpha)$ (1).

Démonstration.

Posons :

$$\beta_r = \Phi(\mu_r, \alpha_r) = G_0(\tilde{S}_r, \mu_r)$$

et

$$\beta = \liminf_{r_j} \Phi(\mu_r, \alpha_r) = \Phi(\mu, \alpha).$$

Soit $(\beta_{r_j'})$ une sous-suite de (β_r) telle que

$$\lim_{r_j'} \beta_{r_j'} = \beta.$$

De la suite $(\tilde{S}_{r_j'})$ on peut extraire une sous-suite (\tilde{S}_{r_j}) convergeant vers un certain $\tilde{S} \in \Gamma(\mu, \alpha)$, d'après la s.c.s. de Γ . Enfin la s.c.i. du critère G_0 entraîne :

$$G_0(\tilde{S}, \mu) \leq \liminf_r G_0(\tilde{S}_{r_j}, \mu_{r_j}) = \liminf_{r_j} \Phi(\mu_{r_j}, \alpha_{r_j}) = \Phi(\mu, \alpha).$$

Donc \tilde{S} est une solution optimale de $P(\mu, \alpha)$.

Corollaire. Si on remplace l'hypothèse (9) par :

$$(10) \quad \lim_r \Phi(\mu_r, \alpha_r) = \Phi(\mu, \alpha)$$

alors toute sous-suite convergente de (\tilde{S}_r) converge vers une solution optimale de $P(\mu, \alpha)$.

Démonstration.

Si (\tilde{S}_{r_j}) converge vers \tilde{S} , on a $\tilde{S} \in \Gamma(\mu, \alpha)$ et de plus :

$$G_0(\tilde{S}, \mu) \leq \liminf_{r_j} \Phi(\mu_{r_j}, \alpha_{r_j}) = \Phi(\mu, \alpha).$$

Proposition 6. Soit (μ_r, α_r) une suite convergeant vers (μ, α) et supposons que :

a) l'application $S \rightarrow G_0(S, \mu)$ est continue sur \mathcal{S} ;

(1) On rappelle que μ_r converge vers μ pour la topologie de la convergence étroite et que la sous-suite de (\tilde{S}_r) convergera au sens de la topologie de la convergence compacte sur Σ .

b) *il existe une partie dénombrable* : $\Gamma^d \subset \limsup_r \Gamma(\mu_r, \alpha_r)$ *dense dans* $\Gamma(\mu, \alpha)$ *et telle que :*

$$\forall S \in \Gamma^d, \quad \lim_r G_0(S, \mu_r) = G_0(S, \mu).$$

Alors on a : $\liminf_r \Phi(\mu_r, \alpha_r) = \Phi(\mu, \alpha)$ *et si* (\tilde{S}_r) *est optimale pour* $P(\mu_r, \alpha_r)$ *il existe une sous-suite de* (\tilde{S}_r) *convergeant vers une solution optimale de* $P(\mu, \alpha)$.

Démonstration.

Pour $S \in \Gamma^d$, il existe (r_j) telle que $S \in \Gamma(\mu_{r_j}, \alpha_{r_j})$ et donc :

$$\liminf_r \Phi(\mu_r, \alpha_r) \leq \lim_{r_j} G_0(S, \mu_{r_j}) = G_0(S, \mu).$$

La continuité de $G_0(\cdot, \mu)$ et la condition b) entraînent alors que :

$$\liminf_r \Phi(\mu_r, \alpha_r) \leq G_0(S, \mu), \quad \forall S \in \Gamma(\mu, \alpha).$$

D'où l'inégalité :

$$\liminf_r \Phi(\mu_r, \alpha_r) \leq \Phi(\mu, \alpha).$$

L'inégalité inverse découle de la s.c.i. de la fonction Φ (théor. 1) et le reste de la proposition 5.

Proposition 7. *Soit* (μ_r, α_r) *une suite convergeant vers* (μ, α) *et supposons que :*

a) *l'application* $S \rightarrow G_0(S, \mu)$ *est continue sur* S ;

b) *il existe une partie dénombrable* $\Gamma^d \subset \liminf_r \Gamma(\mu_r, \alpha_r)$ *et dense dans* $\Gamma(\mu, \alpha)$, *telle que*

$$\forall S \in \Gamma^d, \quad \lim_r G_0(S, \mu_r) = G_0(S, \mu).$$

Alors on a : $\lim_r \Phi(\mu_r, \alpha_r) = \Phi(\mu, \alpha)$ *et toute sous-suite convergente de* (\tilde{S}_r) , *converge vers une solution optimale de* $P(\mu, \alpha)$.

Démonstration.

Soit (β_{r_j}) une sous-suite convergente de la suite $\beta_r = \Phi(\mu_r, \alpha_r)$ et soit $S \in \Gamma^d$. Comme $S \in \liminf_r \Gamma(\mu_r, \alpha_r)$, à partir d'un certain rang $r_j \geq r_0$, on aura $S \in \Gamma(\mu_{r_j}, \alpha_{r_j})$. D'où :

$$\beta = \lim_{r_j} \Phi(\mu_{r_j}, \alpha_{r_j}) \leq \lim_{r_j} G_0(S, \mu_{r_j}) = G_0(S, \mu).$$

Par continuité de $G_0(\cdot, \mu)$ on obtient encore :

$$\beta \leq G_0(S, \mu), \quad \forall S \in \Gamma(\mu, \alpha).$$

Donc tout point d'accumulation de la suite (β_r) est inférieur ou égal à $\Phi(\mu, \alpha)$; mais comme on a toujours :

$$\liminf_r \beta_r \geq \Phi(\mu, \alpha) \quad (\text{théorème 1}),$$

cela n'est possible que si

$$\Phi(\mu, \alpha) = \lim_r \Phi(\mu_r, \alpha_r).$$

Le reste découle alors du corollaire, proposition 5.

§ 3. LOI FORTE DES GRANDS NOMBRES EN PROGRAMMATION DYNAMIQUE STOCHASTIQUE

3.1. Rappels sur la loi forte des grands nombres

Soit (Ω, \mathcal{F}, P) un espace probabilisé et (η_j) $j \geq 1$, une suite de variables aléatoires indépendantes équidistribuées, définies sur cet espace. On suppose que la loi commune des η_j admet un moment du 1^{er} ordre a , et un écart type σ fini.

Posons :

$$\xi_r = \sum_{j=1}^r (\eta_j - a).$$

La martingale ξ_r possède alors les propriétés suivantes :

Proposition 8 : Pour tout réel $\varepsilon > 0$, on a :

$$(11) \quad \lim_r \frac{1}{r^{1/2+\varepsilon}} \sum_{j=1}^r (\eta_j - a) = 0 \text{ (p.s.)}$$

Démonstration.

Soit $u_r = r^{1/2+\varepsilon}$. La série $\sum_r u_r^{-2} \sigma$, étant convergente, le résultat découle de [3], proposition IV-6-1.

Remarquons que la condition (11) est équivalente à :

$$(11') \quad \forall \lambda > 0, \quad P\left(\liminf_r \left\{ \omega : \left| \frac{1}{r} \xi_r \right| \leq \lambda r^{-1/2+\varepsilon} \right\}\right) = 1.$$

Proposition 9. *Supposons que la loi commune des η_j est à support compact. Alors :*

$$(12) \quad P(\limsup_r \xi_r = +\infty) = P(\liminf_r \xi_r = -\infty) = 1$$

(pour la démonstration voir [3], prop. IV-6-3).

Une conséquence de cette dernière proposition est que, si on désigne par $\bar{\eta}_r = \frac{1}{r} \sum_{j=1}^r \eta_j$, la moyenne arithmétique de la suite (η_j) , la suite $\bar{\eta}_r$ converge presque sûrement vers a en oscillant une infinité dénombrable de fois autour de cette valeur :

$$P(\limsup_r \{ \bar{\eta}_r \leq a \}) \geq P(\liminf_r \xi_r < 0) = 1$$

$$P(\limsup_r \{ \bar{\eta}_r \geq a \}) \geq P(\limsup_r \xi_r > 0) = 1.$$

3.2. Application à la simulation indépendante

Dans toute la suite de cette partie, on désignera par $P(\mu, \alpha)$ un problème de programmation dynamique stochastique vérifiant les hypothèses supplémentaires :

H_1 , le critère $G_0(S, \mu)$ est continu sur \mathcal{S} (μ fixé).

Une condition suffisante pour que H_1 soit vérifiée est donnée par la proposition 2.

H_2 , pour tout $k = 1, \dots, K$ et tout $S \in \mathcal{S}$:

$$\int_E g_k^2(X(S, e), S) d\mu(e) < +\infty.$$

Par ailleurs on notera (Ω, \mathcal{F}, P) l'espace produit $(E^N, \otimes_{n \in N} \mathcal{B}, \otimes_{n \in N} \mu)$: c'est donc l'espace des simulations (ou tirages) indépendantes des paramètres aléatoires e du système.

A tout $\omega = (e_r)_{r \geq 1} \in \Omega$, on associe une suite de mesures discrètes $(\mu_r(\omega))$ sur E , en posant :

$$(13) \quad \mu_r(\omega) = \frac{1}{r} \sum_{j=1}^r \delta_{e_j}.$$

D'après le théorème de Glivenko-Cantelli ([4], chap. 10, § 10), la suite $\mu_r(\omega)$ converge étroitement vers μ , pour presque tout $\omega \in \Omega$.

Par ailleurs remarquons que $G_k(S, \mu_r(\omega))$ s'écrit :

$$(14) \quad G_k(S, \mu_r(\omega)) = \frac{1}{r} \sum_{j=1}^r g_k(S, e^j); \quad k = 0, 1, \dots, K.$$

C'est donc la moyenne arithmétique de la suite de variables aléatoires indépendantes équadistribuées $\gamma_k^j(\omega) = g_k(S, e^j)$.

On étudie maintenant l'approximation de $P(\mu, \alpha)$ par des problèmes $P(\mu_r(\omega), \alpha_r)$.

Théorème 2. *Supposons les hypothèses H_1, H_2 , vérifiées et soit α_r une suite de R^K convergeant vers α de telle façon qu'il existe $\varepsilon > 0$ et $\lambda > 0$ tels que :*

$$\alpha_k^r - \alpha_k \geq \lambda r^{-\frac{1}{2} + \varepsilon}; \quad k = 1, \dots, K.$$

Soit $\tilde{S}_r(\omega)$ une solution optimale du problème $P(\mu_r(\omega), \alpha^r)$, lorsque $\Gamma(\mu_r(\omega), \alpha^r)$ est non vide, ou dans le cas contraire un élément quelconque de S .

Alors sauf sur un P négligeable de Ω , on a :

a)
$$\Phi(\mu, \alpha) = \lim_r \Phi(\mu_r(\omega), \alpha^r);$$

b) *toute sous-suite convergente de $\tilde{S}_r(\omega)$, converge vers une solution optimale de $P(\mu, \alpha)$.*

Démonstration.

Il s'agit de montrer que pour presque tout $\omega \in \Omega$, les conditions a) et b) de la proposition 7 sont vérifiées.

La condition a) découle de l'hypothèse H_1 .

Posons maintenant :

$$\varepsilon_k^r = \alpha_k^r - \alpha_k$$

$$A_k^r(S) = \{ \omega \in \Omega : |G_k(S, \mu_r(\omega)) - G_k(S, \mu)| \leq \varepsilon_k^r \}.$$

D'après (14) et (11'), la moyenne arithmétique $G_k(S, \mu_r(\omega))$ va converger p.s. vers $G_k(S, \mu)$ avec une vitesse supérieure à $r^{-1/2 + \varepsilon} \leq \varepsilon_k^r$.

Donc pour tout S , et tout k :

(15)
$$P(\liminf_r A_k^r(S)) = 1.$$

Soit par ailleurs Γ^d une partie dénombrable dense dans $\Gamma(\mu, \alpha)$. On a les inclusions :

$$\begin{aligned} \Omega_1 &= \{ \omega : \Gamma^d \subset \liminf_r \Gamma(\mu_r(\omega), \alpha^r) \} \\ &= \bigcap_{S \in \Gamma^d} \liminf_r \{ \omega : S \in \Gamma(\mu_r(\omega), \alpha^r) \} \\ &\supset \bigcap_{S \in \Gamma^d} \liminf_r \left(\bigcap_{k=1}^K A_k^r(S) \right) \\ &= \bigcap_{S \in \Gamma^d} \bigcap_{k=1}^K \liminf_r A_k^r(S). \end{aligned}$$

En conséquence, d'après (15) :

$$P(\Omega_1) = 1.$$

Désignons maintenant par Ω_2 l'ensemble des ω tels que $\mu_r(\omega)$ converge étroitement vers μ , et par Ω_3 :

$$\Omega_3 = \{ \omega : G_0(S, \mu) = \lim_r G_0(S, \mu_r(\omega)), \forall S \in \Gamma^d \}.$$

On sait que $P(\Omega_2) = 1$ (théorème de Glivenko-Cantelli) et de même $P(\Omega_3) = 1$ d'après la loi forte des plus grands nombres et le fait que Γ^d est dénombrable.

Posons alors $\Omega_0 = \Omega_1 \cap \Omega_2 \cap \Omega_3$. Dans ces conditions, on a montré que $P(\Omega_0) = 1$ et que pour tout $\omega \in \Omega_0$, la suite $(\mu_r(\omega), \alpha^r)$ converge vers (μ, α) en vérifiant :

$$\left\{ \begin{array}{l} - \Gamma^d \subset \liminf_r \Gamma(\mu_r(\omega), \alpha^r) \\ - \Gamma^d \text{ dense dans } \Gamma(\mu, \alpha) \\ - \forall S \in \Gamma^d, \lim_r G_0(S, \mu_r(\omega)) = G_0(S, \mu) \end{array} \right.$$

c.q.f.d.

REMARQUE. Dans la pratique, ce résultat a un inconvénient : celui de perturber les bornes des contraintes. En particulier, si on part d'une contrainte presque-sûre pour $P(\mu, \alpha)$:

$$\mu \{ e : X(S, e) \in B \} = 1$$

cette contrainte est transformée en contrainte en probabilité pour $P(\mu_r(\omega), \alpha^r)$:

$$\frac{1}{r} \text{card} \{ 1 \leq j \leq r : X(S, e^j) \in B \} \geq 1 - \varepsilon_r$$

avec

$$\varepsilon_r > r^{-\frac{1}{2} + \varepsilon}.$$

3.3. Cas de contraintes presque-sûres

On considère ici un problème $P(\mu)$ vérifiant l'hypothèse H_1 et dont toutes les contraintes sont du type presque-sûre. Dans ce cas, on peut toujours se ramener à la donnée d'un sous-ensemble B de F tel que l'on ait l'unique contrainte :

$$\mu \{ e : X(S, e) \in B \} = 1.$$

Dans ces conditions, le programme s'écrit :

$$P(\mu) \begin{cases} \Phi(\mu) = \min_{S \in \mathcal{S}} G_0(S, \mu) \\ X_{i+1}(S, e) = f_i(X_i(S, e), S_i(X_i^{ob}(S, e)), e) \\ \mu \{ e : X(S, e) \in B \} = 1. \end{cases}$$

Et les stratégies admissibles sont données par :

$$\Gamma(\mu) = \{ S \in \mathcal{S} : \mu(X(S, e) \in B) = 1 \}.$$

Remarquons que les problèmes simulés : $P(\mu_r(\omega))$, $\omega = (e_j)_{j \leq r}$, se mettent sous la forme simple :

$$P(\mu_r(\omega)) \begin{cases} \Phi(\mu_r(\omega)) = \min_{S \in \mathcal{S}} \frac{1}{r} \sum_{j=1}^r g_0(S, e^j) \\ X_{i+1}(S, e^j) = f_i(X_i(S, e^j), S_i(X_i^{ob}(S, e^j)), e^j) \\ X(S, e^j) \in B \quad 1 \leq j \leq r. \end{cases}$$

Théorème 3. *Supposons que H_1 est vérifié et que B soit fermé dans F .*

Soit $\tilde{S}_r(\omega)$ une solution optimale de $P(\mu_r(\omega))$, lorsque $\Gamma(\mu_r(\omega))$ est non vide ou, dans le cas contraire, un élément quelconque de \mathcal{S} .

Alors, sauf sur un P -négligeable de Ω :

a)
$$\Phi(\mu) = \lim_r \Phi(\mu_r(\omega))$$

b) *toute sous-suite convergente de $\tilde{S}_r(\omega)$ converge vers une solution optimale de $P(\mu)$.*

Démonstration.

On applique encore la proposition 7. La condition a) est obtenue par l'hypothèse H_1 .

Posons maintenant pour $S \in \mathcal{S}$:

$$B_S = \{ e : X(S, e) \in B \}$$

$$A_r(S) = \{ \omega : \mu_r(\omega)(B_S) \geq \mu(B_S) \}.$$

Remarquons que :

$$\mu_r(\omega)(B_S) = \frac{1}{r} \sum_{j=1}^r \chi_S(e_j)$$

où χ_S est la fonction caractéristique de B_S .

L'application de la proposition 9, nous dit que la fréquence empirique $\mu_r(\omega)(B_S)$ va converger vers $\mu(B_S)$ en oscillant une infinité dénombrable de fois autour de cette valeur. Donc pour tout S :

$$(16) \quad P(\limsup_r A_r(S)) = 1.$$

Soit maintenant Γ^d une partie dénombrable dense de $\Gamma(\mu)$. En utilisant la décroissance des ensembles $\Gamma(\mu_r(\omega))$, on vérifie les inclusions :

$$\begin{aligned} \Omega_1 &= \{ \omega : \Gamma^d \subset \liminf_r \Gamma(\mu_r(\omega)) \} \\ &= \bigcap_{S \in \Gamma^d} \liminf_r \{ \omega : S \in \Gamma(\mu_r(\omega)) \} \\ &= \bigcap_{S \in \Gamma^d} \limsup_r \{ \omega : S \in \Gamma(\mu_r(\omega)) \} \\ &\supset \bigcap_{S \in \Gamma^d} \limsup_r A_r(S). \end{aligned}$$

Donc, d'après (16), on a :

$$P(\Omega_1) = 1.$$

Et comme dans le théorème 2, on pose :

$$\begin{aligned} \Omega_2 &= \{ \omega \in \Omega : \mu_r(\omega) \text{ converge étroitement vers } \mu \} \\ \Omega_3 &= \{ \omega \in \Omega : G_0(S, \mu) = \lim_r G_0(S, \mu_r(\omega)), \quad \forall S \in \Gamma^d \} \\ \Omega_0 &= \Omega_1 \cap \Omega_2 \cap \Omega_3. \end{aligned}$$

On a encore $P(\Omega_0) = 1$ et la condition $b)$ de la proposition 7 se trouve ainsi vérifiée.

REMARQUE. Soit T une transformation sur E laissant la mesure μ invariante. Le problème $P(\mu)$ est alors équivalent à minimiser sur \mathcal{S} , le nouveau critère :

$$\tilde{G}(S, \mu) = \frac{1}{2} \left\{ \int_E g_0(X(S, e), S) d\mu(e) + \int_E g_0(X(S, Te), S) d\mu(e) \right\}$$

sous la nouvelle contrainte :

$$\mu \{ e : X(S, e) \in B \} = \mu \{ e : X(S, Te) \in B \} = 1.$$

Pratiquement, cela revient à ajouter l'état supplémentaire $X(S, Te)$ au système initial. Le théorème 3 s'applique donc encore à ce nouveau problème et les essais numériques semblent montrer que les vitesses de convergence de l'algorithme se trouvent ainsi améliorées (cf. § 4, exemple 1).

§ 4. EXEMPLES NUMERIQUES

On présente dans ce paragraphe deux essais numériques sur des problèmes $P(\mu)$, complètement observable.

Le premier est l'exemple « test » d'un système de Kalman, le second un problème de gestion-production à plusieurs niveaux. Dans les deux cas, les contrôles feed-back ont été pris dans un ensemble d'applications k -lipschitziennes :

$$\mathcal{S}_k = \{ S : |S_i(x) - S_i(x')| \leq k |x - x'| \quad ; \quad \forall x, x' \quad ; \quad i = 0, 1 \dots N-1 \}.$$

Les problèmes $P(\mu_r(\omega))$, $\omega = (e^j)_{j \geq 1}$, sont alors de la forme :

$$(17) \quad \begin{cases} x_i^j = f_i(x_i^j, u_i^j, e_i^j) \quad ; \quad x_0 \text{ fixe} \quad ; \quad i = 0, 1 \dots N-1 \quad ; \quad j = 1 \dots r. \\ x_i^j \in B_i \quad , \quad u_i^j \in C_i \\ |u_i^j - u_i^{j'}| \leq k |x_i^j - x_i^{j'}| \quad , \quad \forall (j, j') \\ \min_u \frac{1}{r} \sum_{j=1}^r \left\{ \sum_{i=0}^{N-1} L_i(x_i^j, u_i^j) + M(x_N^j) \right\}. \end{cases}$$

Les méthodes de résolution utilisées ont été des méthodes de gradient sur le problème pénalisé (la principale difficulté provenant de la contrainte (17) non convexe en $u_i = (u_i^j)_{1 \leq j \leq r}$).

EXEMPLE 1. *Système linéaire, coût quadratique.*

$$(18) \quad \begin{cases} x_{i+1} = x_i + 2u_i + e_i \quad ; \quad i = 0, 1, 2. \\ x_0 = 10 \\ e_i \quad : \quad \text{loi normale centrée d'écart-type 10} \\ \min_{S \in \mathcal{S}_k} E \left\{ x_3^2 + \sum_{i=0}^2 u_i^2 \right\} \quad , \quad k = 1. \end{cases}$$

La solution théorique est obtenue en résolvant les équations de Riccati donnant le feed-back déterministe.

$$\tilde{u}_0 = -0,154 x_0$$

$$\tilde{u}_1 = -0,222 \tilde{x}_1$$

$$\tilde{u}_2 = -0,4 \tilde{x}_2$$

Le coût optimal est de : 138,8.

Pour les problèmes $P(\mu_r(\omega))$ issus de (18), on utilise la remarque du théorème 3 : la loi normale étant symétrique, les e_i^j seront générés par couples

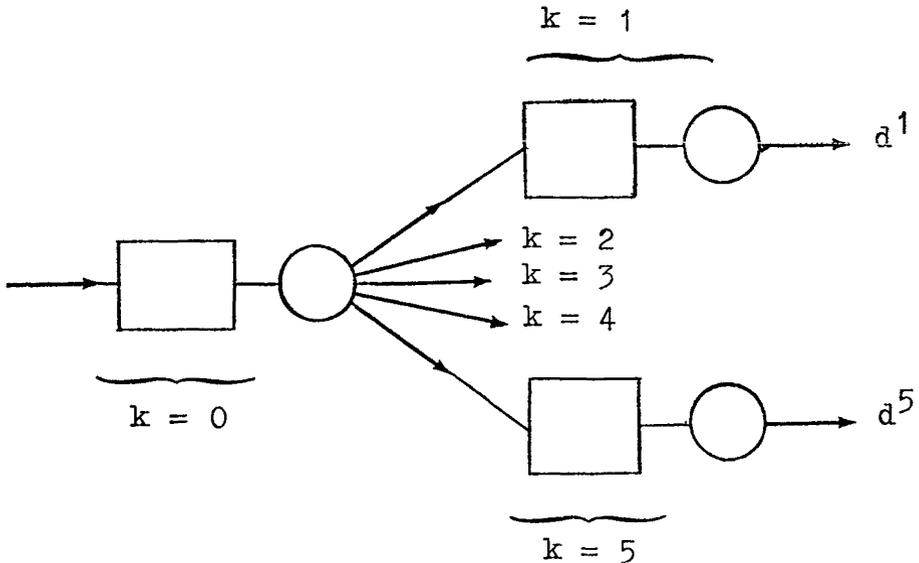
symétriques. Dans le tableau suivant, on présente les valeurs de la commande initiale u_0 et les valeurs de $\Phi(\mu_r(\omega))$ pour différents r .

	THÉORIQUE	$r = 10$	$r = 20$	$r = 50$
u_0	- 1,54	- 1,54	- 1,54	- 1,54
Coût.	138,8	59	70	110

Remarquons la bonne approximation de la commande optimale initiale : elle s'explique certainement par le fait que les e_i ont été générés par couples symétriques. Quant à l'évolution du coût optimal, on ne peut bien entendu espérer mieux qu'une vitesse de convergence comparable à celle de la loi forte des grands nombres (déjà supérieure à $r^{-1/2}$!).

EXEMPLE 2. *Gestion. Production.*

Le système est représenté par le diagramme :



Chaque carré étant un système de production régi par l'équation :

$$(19) \quad p_{i+1}^k = p_i^k + u_i^k \quad k = 0, 1 \dots 5 \ ; \ i = 0, \dots 8.$$

- où p_i^k est le niveau de production du système k à l'instant i ;
- où u_i^k désigne la variation de production décidée à l'instant i .

Chaque rond représente un stock régi par l'équation :

$$(20) \quad s_{i+1}^k = s_i^k + p_i^k - d_i^k \quad ; \quad k = 0, \dots, 5 \quad i = 0, \dots, 8.$$

— où s_i^k désigne le niveau de stock à l'instant i pour l'unité k ,

— où d_i^k désigne la sortie du stock à l'instant i .

Pour $k = 0$:

$$d_i^0 = \sum_{k=1}^5 p_i^k.$$

Pour $k = 1 \dots 5$, les d_i^k sont aléatoires (indépendantes en i et k) et de loi de probabilité égale au produit de convolution de trois lois uniformes sur $[0 ; 0, 1]$.

Les contraintes sont de la forme :

— pour les niveaux de production,

$$0 \leq p_i^k \leq 2 \quad ; \quad \forall k \text{ et } i.$$

— pour les stocks,

$$s_i^k \geq 0 \quad ; \quad \forall k \text{ et } i.$$

Le critère à minimiser étant :

$$E \left\{ \sum_{i=0}^8 [p_i + 0,2s_i + u_i^2] \right\}.$$

Pour 10 réalisations et les états initiaux :

$$p_0^0 = 1,5 \quad ; \quad p_0^k = 0,3 \quad ; \quad s_0^k = 0 \quad , \quad k = 1 \dots 5,$$

on a obtenu :

coût optimal : 17,2

commande initiale optimale :

$$u_0^1 = -0,0621$$

$$u_0^2 = -0,0623$$

$$u_0^3 = -0,0619$$

$$u_0^4 = -0,0617$$

$$u_0^5 = -0,0625$$

$$u_0^0 = -0,257$$

temps machine : 3 mn sur CII 10.070.

n° avril 1973, R-1.

BIBLIOGRAPHIE

- [1] BERGE C., *Espaces topologiques, fonctions multivoques*, Dunod, Paris, 1959.
- [2] BOUBAKI N., *Intégration*, chap. IX, Hermann, Paris, 1969.
- [3] NEVEU J., *Bases mathématiques du Calcul des Probabilités*, Masson, Paris, 1964.
- [4] FISZ M., *Probability Theory and Mathematical Statistics*, J. Wiley, 3^e édition, 1969.
- [5] QUADRAT J. et VIOT M., *Approximation numérique des problèmes de programmation dynamique stochastique*, IRIA, Cahier n° 9, mars 1972 : Méthodes numériques d'analyse de systèmes, tome 1.