

ANNALES SCIENTIFIQUES
DE L'UNIVERSITÉ DE CLERMONT-FERRAND 2
Série Mathématiques

J. P. BERTRAN

Méthodes stochastiques de dualité

Annales scientifiques de l'Université de Clermont-Ferrand 2, tome 58, série *Mathématiques*, n° 12 (1976), p. 35-63

http://www.numdam.org/item?id=ASCFM_1976__58_12_35_0

© Université de Clermont-Ferrand 2, 1976, tous droits réservés.

L'accès aux archives de la revue « *Annales scientifiques de l'Université de Clermont-Ferrand 2* » implique l'accord avec les conditions générales d'utilisation (<http://www.numdam.org/conditions>). 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 STOCHASTIQUES DE DUALITE

J. P. BERTRAN, Université de Nancy

V.1 - Enoncé du problème, données

V.1.1 - Données, hypothèses

Nous nous situons dans un contexte théorique envisagé par J.L. Lions, R. Tremolières, R. Glowinski [1] (dans le cas quadratique) ou par A. Bensoussan, J.L. Lions, R. Temam [1] .

Soit \mathcal{U} un espace de Hilbert séparable sur \mathbb{R} et $\mathcal{U}_{ad} \subset \mathcal{U}$, \mathcal{U}_{ad} convexe fermé.

Soit $J : \begin{matrix} v & \xrightarrow{\quad} & J(v) \\ \mathcal{U} & \xrightarrow{\quad} & \mathbb{R} \end{matrix}$ une fonctionnelle vérifiant

$\underline{J}_1 : J \text{ est } C_2$

$\underline{J}_2 : (\nabla J(u) - \nabla J(v), u-v)_{\mathcal{U}} > m \|u-v\|_{\mathcal{U}}^2, \quad m > 0$

$\underline{J}_3 : \|\nabla J(u) - \nabla J(v)\|_{\mathcal{U}} \leq M \|u-v\|_{\mathcal{U}}$

Soient \mathcal{U}_{ad} et \mathcal{U}'_{ad} deux convexes fermés de \mathcal{U} .

On peut repérer \mathcal{U}'_{ad} par la donnée

- de Λ , cône convexe de sommet 0 dans un espace vectoriel normé Ψ à préciser

- et d'une fonction $(v,q) \longrightarrow \Phi(v,q)$

$$\mathcal{U} \times \Lambda \longrightarrow \mathbb{R}$$

convexe, s.c.i. en v et telle que $\Phi(v,\rho q) = \rho \Phi(v,q) \quad \forall \rho \geq 0$

concave, s.c.s. en p .

Alors $v \in \mathcal{U}'_{ad} \iff \Phi(v,q) \leq 0 \quad \forall q \in \Lambda$.

Pour simplifier et parce que ce sera suffisant les applications projetées, nous supposons que Ψ est un espace de Hilbert séparable identifié à son dual et que $\Phi(v,q)$ est de la forme $\Phi(v,q) = (q, Bv)_{\Psi}$

où $B : v \longrightarrow B(v)$ vérifie $\|B(v) - B(v')\|_{\Psi} \leq C \|v - v'\|_{\mathcal{U}} \quad \forall v, v' \in \mathcal{U}_{ad}$.
 $\mathcal{U}'_{ad} \longrightarrow \Psi$

Pour inclure dans la théorie les problèmes de minimisation de fonctionnelles non différentiables, nous supposons seulement que Λ est convexe fermé. Enfin, nous ferons l'hypothèse de qualification :

$$\underline{Q} : \exists v_0 \in \mathcal{U}'_{ad} \text{ tel que } (q, B(v_0))_{\Psi} < 0 \quad \forall q \in \Lambda.$$

V.1.2 - Problème

On considère le problème $\min_{v \in \mathcal{U}_{ad} \cap \mathcal{U}'_{ad}} J(v)$.

L'introduction du repérage de \mathcal{U}'_{ad} permet de reformuler ce problème :

problème primal :

$$(5.1.1) \quad \inf_{v \in \mathcal{U}_{ad}} \sup_{q \in \Lambda} \{J(v) + (q, B(v))_{\Psi}\}$$

auquel on associe le

problème dual

$$(5.1.2) \quad \sup_{q \in \Lambda} \inf_{v \in \mathcal{U}_{ad}} \{J(v) + (q, B(v))_{\Psi}\}$$

V.1.3 - Théorèmes d'existence et d'unicité, reformulation du problème dual

Posons $\mathcal{L}(v, q) = J(v) + (q, B(v))_{\Psi}$.

Remarques sur les hypothèses :

1° $m > 0$ n'est pas nécessaire pour un théorème d'existence. Il ne sera requis que pour l'unicité.

$$2° \underline{J_2} \implies \|u\| \longrightarrow +\infty \implies J(u) \longrightarrow +\infty$$

3° \underline{J}_1 sera nécessaire seulement pour utiliser le théorème des fonctions implicites. Pour l'instant, il suffit que J soit C^1 .

4° $\underline{Q} \implies \mathcal{U}'_{ad}$ admet au moins un point intérieur.

5° On peut remplacer \underline{Q} par une autre hypothèse de qualification comme on le fait couramment en programmation mathématique.

Énonçons maintenant sans démonstration quelques résultats classiques. On pourra consulter J Céa [1] ou J.Céa, R. Glowinski [1].

Théorème 5.1.3.1

(5.1.3)

Sous les hypothèses \underline{Q} , \underline{J}_1 , \underline{J}_2 , \underline{J}_3 , il existe
 $(u, p) \in \mathcal{U}_{ad} \times \Lambda$ solution de $\mathcal{L}(u, q) \leq \mathcal{L}(u, p) < \mathcal{L}(v, p)$
 $\forall (v, q) \in \mathcal{U}_{ad} \times \Lambda$.

Corollaire 5.1.2.2

(5.1.4)

Un couple (u, p) solution de (5.1.3) est aussi solution de (5.1.4) et réciproquement

$$\left[\begin{array}{l} u \in \mathcal{U}_{ad}, p \in \Lambda \\ (\nabla J(u), v-u)_{\mathcal{U}} + (p, B(v)-B(u))_{\Psi} \geq 0 \quad \forall v \in \mathcal{U}_{ad} \\ (q-p, B(u))_{\Psi} \leq 0 \quad \forall q \in \Lambda \end{array} \right.$$

Ce corollaire montre que la théorie qui va suivre s'applique aux problèmes posés en termes d'inéquations : remplaçons dans (5.1.4) $\nabla J(u)$ par $A(u)$ où A vérifie les hypothèses :

\underline{A}_1 : A est C^1

\underline{A}_2 : $(A(u)-A(v), u-v)_{\mathcal{U}} \geq m \|u-v\|_{\mathcal{U}}^2 \quad (m > 0)$

\underline{A}_3 : $\|A(u) - A(v)\|_{\mathcal{U}} \leq M \|u-v\|_{\mathcal{U}}$

Les paragraphes suivants donneront des moyens pour résoudre le problème :

$$(5.1.5) \quad \begin{cases} u \in \mathcal{U}_{ad}, p \in \Lambda \\ (A(u), v-u)_{\mathcal{U}} + (p, B(v) - B(u))_{\Psi} \geq 0 \quad \forall v \in \mathcal{U}_{ad} \\ (q-p, B(u))_{\Psi} \leq 0 \quad \forall q \in \Lambda. \end{cases}$$

Remarque : on a implicitement identifié \mathcal{U} à son dual. Ce n'est pas essentiel pour la suite. On pourrait écrire $\langle \nabla J(u), v-u \rangle_{\mathcal{U}^*, \mathcal{U}}$ ou $\langle A(u), v-u \rangle_{\mathcal{U}^*, \mathcal{U}}$.

Corollaire 5.1.3.3

Si (u_1, p_1) et (u_2, p_2) sont 2 solutions de (5.1.3),
alors $u_1 = u_2$.

(On utilise ici $m > 0$)

Donc :

Corollaire 5.1.3.4

Si (u, p) est solution de (5.1.3) u est la solution
unique du problème primal.

Corollaire 5.1.3.5

Si p est une solution du problème dual et si u mini-
mise $J(v) + (p, B(v))_{\Psi}$ sur \mathcal{U}_{ad} , alors (u, p) vérifie
(5.1.3) et u est la solution du problème primal.

On peut maintenant reformuler le problème dual

$$\begin{cases} \sup_{q \in \Lambda} [J(u_q) + (q, B(u_q))_{\Psi}] \\ u_q \text{ défini par :} \\ q \in \Lambda, u_q \in \mathcal{U}_{ad} \\ (\nabla J(u_q), v-u_q)_{\mathcal{U}} + (q, B(v)-B(u_q))_{\Psi} \geq 0 \quad \forall v \in \mathcal{U}_{ad} \end{cases}$$

V.2 - Exemples

On pourra trouver dans A. Bensoussan, J.L. Lions, R. Temam [1] des exemples économiques du problème envisagé.

Donnons ici des exemples en théorie du contrôle.

V.2.1 - Remarques générales (cf. J.L. Lions [1])

Soit le problème

$$(5.2.1) \quad \inf_{v \in \mathcal{U}_{ad}} J(v)$$

où $J(v) = \frac{1}{2} \pi(v,v) - L(v) + J(0)$ est une fonctionnelle quadratique vérifiant :

- . $\pi(u,v)$ est une forme bilinéaire symétrique, continue sur \mathcal{U} telle que $\pi(v,v) \geq m \|v\|^2$, $m > 0$ $v \in \mathcal{U}$.
- . $L(v)$ est une forme linéaire continue sur \mathcal{U} .

Nous voulons remplacer le problème (5.2.1) par un problème de point-selle équivalent.

Introduisons la fonction $v \longrightarrow F(v)$ définie par

$$F(v) = \begin{cases} -L(v) & v \in \mathcal{U}_{ad} \\ +\infty & v \notin \mathcal{U}_{ad} \end{cases}$$

$$(5.2.1) \iff (5.2.2) \quad \inf_{v \in \mathcal{U}} \left[\frac{1}{2} \pi(v,v) + F(v) \right]$$

F est convexe, s.c.i., à valeurs dans $]-\infty, +\infty]$, $F \neq +\infty$.

Soit $F^*(w)$ la fonction duale de F par rapport à π :

$$F^*(w) = \sup_{v \in \mathcal{U}} [\pi(w,v) - F(v)].$$

Dans ces conditions, on sait (J.L. Lions [1]) que

$$(5.2.3) \quad \left[\inf_{v \in \mathcal{U}} \left[\frac{1}{2} \pi(v,v) + F(v) \right] + \inf_{v \in \mathcal{U}} \left[\frac{1}{2} \pi(v,v) + F^*(v) \right] = 0, \right.$$

Si u est l'élément qui réalise la 1ère borne inférieure, alors la 2ème borne inférieure est atteinte en $-u$.

En posant $\mathcal{L}(v,w) = \frac{1}{2} \pi(v,v) + \pi(v,w) - F(w)$

(5.2.3) revient à résoudre le problème de point-selle :

$$\mathcal{L}(u,w) \leq \mathcal{L}(u,v) \leq \mathcal{L}(t,v) \quad \forall w \in \mathcal{U}, \quad \forall t \in \mathcal{U}.$$

Remarque : On peut associer au problème (5.2.1) d'autres problèmes de point-selle. Par exemple, remarquons que $F(v) = +\infty$ si $v \notin \mathcal{U}_{ad}$.

$$\text{Donc } F^*(w) = \sup_{v \in \mathcal{U}_{ad}} (\pi(w,v) - F(v)) = \sup_{v \in \mathcal{U}_{ad}} (\pi(w,v) + L(v)).$$

En posant $\mathcal{L}'(v,w) = \frac{1}{2} \pi(v,v) + \pi(v,w) + L(w)$

(5.2.3) équivaut donc à résoudre le problème de point-selle

$$\mathcal{L}'(u,w) \leq \mathcal{L}'(u,v) \leq \mathcal{L}'(t,v) \quad \forall w \in \mathcal{U}_{ad}, \quad \forall t \in \mathcal{U}.$$

V.2.2 - Un problème de contrôle elliptique.

On reprend le problème posé au § I.1.1 que nous avons déjà étudié selon un autre point de vue au § II.5.1.

Le problème est donc : $\inf_{v \in \mathcal{U}_{ad}} J(v)$

où $J(v) = \|\mathcal{C}y(v) - z_d\|_{\mathcal{H}}^2 + (Nv, v)_{\mathcal{U}}$

$$J(v) \text{ s'écrit : } \frac{1}{2} J(v) = \frac{1}{2} \pi(v,v) - L(v) + \frac{1}{2} \|z_d - \mathcal{C}y(0)\|_{\mathcal{H}}^2.$$

avec $\pi(u,v) = (\mathcal{C}(y(u) - y(0)), \mathcal{C}(y(v) - y(0)))_{\mathcal{H}} + (Nu, v)_{\mathcal{U}}$

$$L(v) = (z_d - \mathcal{C}y(0), \mathcal{C}(y(v) - y(0)))_{\mathcal{H}}$$

$$\pi(v,v) \geq \nu \|v\|^2 \quad (\text{propriété de } N)$$

$v \longrightarrow L(v)$ est linéaire car $v \longrightarrow y(v)$ est affine.

On est donc bien dans le contexte du § V.2.1 (cf. J.L. Lions [1]).

Soit $S \in \mathcal{L}(\mathcal{U}, \mathcal{V})$ défini par $Sv = y(v) - y(0)$.

Étudions sur cet exemple comment se transmet, dans le problème de point-selle associé, un bruit sur l'observation.

On peut écrire $\Pi(u,v) = (Mu,v)_{\mathcal{U}}$

$$L(v) = (L,v)_{\mathcal{U}}$$

où $\Pi = S^* \mathcal{C}^* \mathcal{C} S + N$

et $L = S^* \mathcal{C}^* (z_d - \mathcal{C}y(0))$.

L'observation fournit la connaissance de $\mathcal{C}y(v) + \eta(v) = \mathcal{C}Sv + \eta(v)$.

On en déduit donc $\Pi v - L + Z(v)$

où $Z(v) = S^* \mathcal{C}^* \eta(v)$ est un bruit dont les propriétés stochastiques se déduisent aisément de celles de $\eta(v)$.

V.2.3 - Un problème de contrôle parabolique

Soit un problème parabolique avec (par exemple) contrôle et observation frontière. (cf. J.L. Lions [1]).

On reprend une partie du contexte du § I.1.2

On considère le système :

$$\left[\begin{array}{l} \frac{\partial y(v)}{\partial t} + Ay(v) = f \quad \text{dans } Q = \Omega \times [0,T] \\ \frac{\partial y(v)}{\partial \nu_A} = v \quad \text{sur } \Sigma = \partial \Omega \\ y(x, 0 ; v) = y_0(x), \quad x \in \Omega \end{array} \right.$$

et le problème de contrôle

$$\inf_{v \in \mathcal{U}_{ad}} J(v) = \inf_{v \in \mathcal{U}_{ad}} \int_{\Sigma} (y(v) - z_d)^2 d\Sigma + (Nv,v)_{L^2(\Sigma)},$$

L'opérateur \mathcal{C} d'observation est donc ici l'opérateur trace sur Σ .

Pour se placer dans le contexte du § V.2.1, on peut poser :

$$\pi(u,v) = \int_{\Sigma} (y(u) - y(0)) (y(v) - y(0)) d\Sigma + (Nu,v)_{L^2(\Sigma)}$$

$$L(v) = \int_{\Sigma} (z_d - y(0)) (y(v) - y(0)) d\Sigma.$$

Ici encore, l'observation bruitée $\mathcal{C}y(v) + \eta(v) = y(v)|_{\Sigma} + \eta(v)$

permet de déduire $\Pi v - L + Z(v)$.

Ces deux exemples montrent l'intérêt d'étudier des algorithmes de recherche de point-selle utilisant une observation entachée d'un bruit.

V.3 - Algorithmes stochastiques de recherche de points-selle, généralités.

Nous allons étudier dans la suite de ce chapitre deux types de méthodes : une méthode "de type Uzawa" et une méthode "de type Arrow, Hurwicz".

(cf. K.J. Arrow, L. Hurwicz, H. Uzawa [1]).

V.3.1 - Algorithme d'Uzawa

Dans les algorithmes de type Uzawa, l'idée de base est la suivante : on résout le problème dual par une méthode de gradient et on récupère comme "sous-produit" une approximation de la solution optimale du problème primal.

L'algorithme déterministe est défini par :

$$\left[\begin{array}{l} p_0 \text{ quelconque } \in \Lambda \\ p_n \text{ étant connue, soit } u_n \text{ l'élément de } \mathcal{U}_{ad} \text{ qui minimise} \\ \quad J(v) + (p_n, B(v))_{\Psi} \text{ sur } \mathcal{U}_{ad} \\ \text{puis } p_{n+1} = p_{\Lambda} (p_n + \rho_n B(u_n)) \\ p_{\Lambda} : \text{projecteur sur } \Lambda, \rho_n \text{ à choisir.} \end{array} \right.$$

On suppose que l'observation permet de calculer à q fixé le minimum u_q de $g_q(v) = J(v) + (q, B(v))_{\Psi}$ sur \mathcal{U}_{ad} par exemple par un algorithme d'optimisation déterministe. Comme on n'obtiendra pas en général la solution exacte, on peut considérer qu'on trouve seulement $u_q + Z(q)$ où $Z(q)$ est une erreur que nous traiterons d'un point de vue stochastique. B. Martinet [1], étudie ce type d'erreur d'un point de vue déterministe. Voir aussi R.T. Rockafellar [1].

En outre l'algorithme de calcul de u_q peut être lui-même un algorithme d'optimisation stochastique si l'observation fournit seulement $\nabla g_q(v)$ entaché d'un bruit.

Nous supposons donc que l'observation permet de trouver (une réalisation de) $u_q + Z(q)$

où $Z(q)$ est une variable aléatoire vérifiant les hypothèses

$$\begin{aligned} \underline{S_1} \quad & E [B(u_q + Z(q))] = B(u_q) \\ \underline{S_2} \quad & E [||Z(q)||^2] \leq B^2. \end{aligned}$$

L'algorithme stochastique sera donc

$$(5.3.1) \quad \left[\begin{array}{l} P_0 \text{ quelconque } \in \Lambda \\ U_0 = u_{P_0} + Z(P_0) \\ P_n \text{ étant connu, } U_n = u_{P_n} + Z_n \\ P_{n+1} = P_\Lambda (P_n + \rho_n B(U_n)) \\ \text{où } u_{P_n} \text{ est l'élément de } \mathcal{U}_{ad} \text{ qui minimise } J(v) + (P_n, B(v))_\Psi \\ \text{sur } \mathcal{U}_{ad} \text{ et où } Z_n \text{ est une variable aléatoire qui à } P_n = q \\ \text{fixé a même loi que } Z(q). \end{array} \right.$$

Explicitons cet algorithme sur l'exemple V.2.2

Repérons \mathcal{U}_{ad} à l'aide de $\Lambda \subset \Psi$:

$$v \in \mathcal{U}_{ad} \iff (q, B(v))_\Psi \leq 0 \quad \forall q \in \Lambda \subset \Psi .$$

On pose $\mathcal{L}(v, q) = ||\mathcal{G}_y(v) - z_d||^2 + (Nv, v) + (q, B(v))_\Psi$.

Soit $P_n \in \Lambda$.

On résout le problème :

$\inf_{v \in \mathcal{U}} \mathcal{L}(v, P_n)$ à l'aide de l'observation entachée d'un bruit $\mathcal{G}_y(v) + r(v)$.

Soit u_{P_n} la résolution de ce problème.

En raison de la présence de η , on pourra seulement trouver $U_n = u_{P_n} + Z_n$.

On pose ensuite

$$P_{n+1} = P_{\Lambda} (P_n + \rho_n B(U_n)).$$

On retrouve sur cet exemple la situation habituelle dans la méthode d'Uzawa : l'algorithme n'est pas complètement déterminé. Le mode de calcul de U_n reste à choisir. La méthode d'Arrow Hurwicz est une méthode pour préciser ce point : on remplace le calcul de U_{ρ_n} par le calcul d'une valeur approchée à l'aide d'une seule itération de gradient.

V.3.2 - Algorithme d'Arrow Hurwicz

On suppose à présent que $\mathcal{U}_{ad} \cap \mathcal{U}'_{ad} = \mathcal{U}'_{ad}$ i.e. $\mathcal{U}_{ad} = \mathcal{U}$ et que

$$B \in \mathcal{L}(\mathcal{U}, \Psi)$$

$g_q(v)$ admet donc le gradient :

$$\nabla g_q(v) = J(v) + B^* q.$$

L'observation nous fournit pour tout v (une réalisation de) $\nabla J(v) + Z(v)$,

L'algorithme stochastique sera donc :

$$(5.3.2) \quad \left[\begin{array}{l} u_0 \in \mathcal{U} \text{ quelconque, } p_0 \in \Lambda \text{ quelconque} \\ U_{n+1} = U_n - \rho_n^{(1)} (\nabla J(U_n) + B^* P_n + Z_n) \\ P_{n+1} = P_{\Lambda} (P_n + \rho_n^{(2)} B^* U_{n+1}) \\ \text{où } Z_n \text{ est une variable aléatoire } \in L^2(\underline{\omega}, \mathcal{U}') \text{ dont la loi à} \\ U_n = v \text{ fixé est identique à celle de } Z(v), \rho_n^{(1)} \text{ et } \rho_n^{(2)} \\ \text{sont à définir plus loin.} \end{array} \right.$$

Explicitons cet algorithme sur l'exemple V.2.2

$$\nabla J(v) = 2 \Pi v - 2L$$

$$\Pi = S^* \mathcal{C}^* \mathcal{C} S + N$$

$$L = S^* \mathcal{C}^* (z_d - \mathcal{C}y(0)).$$

L'observation nous fournit $\mathcal{C}y(v) + \eta(v) = \mathcal{C}[Sv + y(0)] + \eta(v)$.

$$\begin{aligned} \text{On en déduit } 2S^* \mathcal{C}^* [\mathcal{C}Sv + \eta(v)] + 2Nv - 2S^* \mathcal{C}^* [z_d - \mathcal{C}y(0)] \\ = \nabla J(v) + 2S^* \mathcal{C}^* \eta(v) = \nabla J(v) + Z(v). \end{aligned}$$

Remarquons que le calcul de $\nabla J(v) + Z(v)$ nécessite d'appliquer l'opérateur S^* , donc de résoudre un système aux dérivées partielles adjoint,

$$\text{En effet } \nabla J(v) + Z(v) = 2S^* \mathcal{C}^* [\mathcal{C}y(v) - z_d + \eta(v)]$$

et on a vu au § II.5.1 que si $p_1(v)$ est solution de

$$A^* p_1(v) = \mathcal{C}^* \Lambda_{\mathcal{H}} (\mathcal{C}y(v) - z_d + \eta(v))$$

$$\text{alors } \nabla J(v) + Z(v) = 2(B^* p_1(v) + \Lambda u Nv).$$

Pour mettre en oeuvre l'algorithme (5.3.2), connaissant U_n et P_n , on calcule $\nabla J(U_n) + Z_n$ comme il vient d'être dit, on en déduit U_{n+1} puis P_{n+1} par (5.3.2).

V.3.3 - Introduction de nouveaux espaces

L'étude de la convergence de la variable primale pourrait être conduite à l'aide des seules notions définies jusqu'ici. Cependant, l'introduction de quelques notations supplémentaires permettra l'étude de la convergence de la variable duale.

Notons $\tilde{\mathcal{U}} = L^2(\underline{\omega}, \mathcal{U})$ espace des variables aléatoires à valeurs dans \mathcal{U} dont le carré de la norme est intégrable.

De même $\tilde{\Psi} = L^2(\underline{\omega}, \Psi)$

$$\text{et } \tilde{\Lambda} = \{q \in \tilde{\Psi} \mid \tilde{q}(\omega) \in \Lambda \text{ presque sûrement}\}$$

$$\mathfrak{F}(\tilde{v}, \tilde{q}) = E [(\tilde{q}, B(\tilde{v}))_{\Psi}] = (\tilde{q}, B(\tilde{v}))_{\tilde{\Psi}}$$

$$\tilde{\mathcal{U}}'_{ad} = \{\tilde{v} \in \tilde{\mathcal{U}} \mid \tilde{v}(\omega) \in \mathcal{U}'_{ad} \text{ p.s.}\}$$

$$\tilde{\mathcal{U}}_{ad} = \{\tilde{v} \in \tilde{\mathcal{U}} \mid \tilde{v}(\omega) \in \mathcal{U}_{ad} \text{ p.s.}\}.$$

En identifiant \mathcal{U}_{ad} à un sous-ensemble de \mathcal{U} formé de variables aléatoires presque sûrement constantes, $\mathcal{U}_{ad} \subset \tilde{\mathcal{U}}_{ad}$. De même $\Lambda \subset \tilde{\Lambda}$.

Posons $L(v, q) = I(v) + (q, B(v))_{\Psi}$

$$E \left[\mathcal{L}(\tilde{v}, \tilde{q}) \right] = E \left[J(\tilde{v}) \right] + E \left[(\tilde{q}, B(\tilde{v}))_{\Psi} \right].$$

Nous allons comparer les points-cols de $\mathcal{L}(v, q)$ et ceux de $L(\tilde{v}, \tilde{q})$ sur $\tilde{\mathcal{U}}_{ad} \times \tilde{\Lambda}$.

Proposition 5.3.3.1

Sous l'hypothèse J₃, $I(\tilde{v})$ est dérivable au sens de Frechet dans $\tilde{\mathcal{U}}$ et $\nabla I(\tilde{v})(\omega) = \nabla J(\tilde{v})(\omega)$.

Démonstration :

$$J : v \longrightarrow [J(v)] \qquad I : \tilde{v} \longrightarrow E [J(\tilde{v})]$$

$$\mathcal{U} \longrightarrow \mathcal{R} \qquad \mathcal{U} \longrightarrow \mathcal{R}$$

$$J(\tilde{v} + \delta\tilde{v}) - J(\tilde{v}) = \delta J = (\nabla J(\tilde{v}), \delta\tilde{v})_{\mathcal{U}} + \int_0^1 d\theta (\nabla J(\tilde{v} + \theta\delta\tilde{v}) - \nabla J(\tilde{v}), \delta\tilde{v})_{\mathcal{U}}$$

$$\delta I = \int_{\underline{\omega}} \delta J \, dp$$

$$\delta I = (\nabla J(\tilde{v}), \delta\tilde{v})_{\mathcal{U}} + \int_{\underline{\omega}} dp(\omega) \int_0^1 d\theta (\nabla J(\tilde{v} + \theta\delta\tilde{v}) - \nabla J(\tilde{v}), \delta\tilde{v})_{\mathcal{U}}$$

$$\left| \int_{\underline{\omega}} dp \int_0^1 d\theta (\dots) \right| \leq \int_{\underline{\omega}} dp \frac{M}{2} \|\delta\tilde{v}(\omega)\|^2 = \frac{M}{2} \|\delta\tilde{v}\|_{\mathcal{U}}^2 \text{ d'après } \underline{J_3}.$$

On peut donc écrire $\delta I = (\nabla J(\tilde{v}), \delta\tilde{v})_{\mathcal{U}} + R(\tilde{v}, \delta\tilde{v}) \times \|\delta\tilde{v}\|_{\mathcal{U}}$

où $R(\tilde{v}, \delta\tilde{v}) \longrightarrow 0$.

$$\|\delta\tilde{v}\|_{\mathcal{U}} \longrightarrow 0.$$

Donc $\nabla J(\tilde{v})(\omega) = \nabla I(\tilde{v})(\omega)$.

Considérons à présent le problème de point col

$$L(\tilde{u}, \tilde{q}) \leq L(\tilde{u}, \tilde{p}) \leq L(\tilde{v}, \tilde{p}) \qquad \forall (\tilde{v}, \tilde{q}) \in \tilde{\mathcal{U}}_{ad} \times \tilde{\Lambda}.$$

Proposition 5.3.3.2

Sous les hypothèses \underline{J}_2 et \underline{J}_3 , si $(\tilde{u}_1, \tilde{p}_1)$ et $(\tilde{u}_2, \tilde{p}_2)$ sont points col de $L(\tilde{v}, \tilde{q})$, alors $\tilde{u}_1 = \tilde{u}_2$.

Démonstration :

Il suffit de déduire des hypothèses \underline{J}_2 et \underline{J}_3 (faites sur J) des hypothèses analogues relatives à I.

$$\underline{J}_3 \implies \|\nabla J(\tilde{u}(\omega)) - \nabla J(\tilde{v}(\omega))\|_{\mathcal{U}} \leq M \|\tilde{u}(\omega) - \tilde{v}(\omega)\|_{\mathcal{U}}.$$

En élevant au carré et en intégrant :

$$\|\nabla I(\tilde{u}) - \nabla I(\tilde{v})\|_{\tilde{\mathcal{U}}} \leq M \|u - v\|_{\tilde{\mathcal{U}}}.$$

On opère de même pour \underline{J}_2 .

Proposition 5.3.3.3

Soit (u, p) un point col de $\mathcal{L}(v, q)$. (u, p) est point col de $L(\tilde{v}, \tilde{q})$.

Démonstration :

On sait que $\mathcal{L}(u, q) \leq \mathcal{L}(u, p) \leq \mathcal{L}(v, p) \quad \forall v \in \mathcal{U}_{ad}, \forall q \in \Lambda$,

En choisissant $v = \tilde{v}(\omega), q = \tilde{q}(\omega), \tilde{v} \in \tilde{\mathcal{U}}_{ad}, \tilde{q} \in \tilde{\Lambda}$

et en prenant l'espérance mathématique dans la double inégalité,

$$L(u, \tilde{q}) \leq L(u, p) \leq L(\tilde{v}, p) \quad \forall (\tilde{v}, \tilde{q}) \in \tilde{\mathcal{U}}_{ad} \times \tilde{\Lambda}.$$

V.4 - Convergence de l'algorithme d'Uzawa stochastique

On complète la définition (5.3.1) par le choix de ρ_n

$$\sum_n \rho_n = +\infty, \quad \sum_n \rho_n^2 < +\infty.$$

Nous allons établir un théorème de convergence de la variable primale, puis un théorème de convergence de la variable duale.

V.4.1 - Convergence de la variable primale

Théorème 5.4.1

Soit u la solution du problème primal. Sous les hypothèses $\underline{J}_1, \underline{J}_2, \underline{J}_3, \underline{Q}, \underline{S}_1, \underline{S}_2$ il existe une sous-suite $\{U_{n_\ell}\}$ telle que $U_{n_\ell} - Z_n \xrightarrow{\ell \rightarrow \infty} u$ en moyenne quadratique. Par suite

$$\lim_{\ell} \sup ||U_{n_\ell} - u||_{\mathcal{U}}^2 \leq B^2.$$

Démonstration :

Soit u la solution (unique) du problème primal et p une solution du problème dual. $p = p_{\Lambda}(p + \rho_n B(u))$.

Posons $R_n = P_n - p$

$$(5.4.1) \quad ||R_{n+1}||_{\Psi}^2 \leq ||R_n + \rho_n (B(U_n) - B(u))||_{\Psi}^2 \quad (\text{proposition 1.5.1.1})$$

$$U_n - Z_n \text{ minimise } J(v) + (P_n, B(v))_{\Psi} \text{ sur } \mathcal{U}_{ad} \iff$$

$$(5.4.2) \quad (\nabla J(U_n - Z_n), v - (U_n - Z_n))_{\mathcal{U}} + (P_n, B(v) - B(U_n - Z_n))_{\Psi} \geq 0 \quad \forall v \in \mathcal{U}_{ad}$$

$$\implies \text{ avec } v = u$$

$$(5.4.3) \quad (\nabla J(U_n - Z_n), u - (U_n - Z_n))_{\mathcal{U}} + (P_n, B(u) - B(U_n - Z_n))_{\Psi} \geq 0.$$

La deuxième inégalité de (5.1.3) s'écrit :

$$(5.4.4) \quad (\nabla J(u), v - u)_{\mathcal{U}} + (p, B(v) - B(u))_{\Psi} \geq 0 \quad \forall v \in \mathcal{U}_{ad}$$

$$\implies \text{ avec } v = U_n - Z_n$$

$$(5.4.5) \quad (\nabla J(u), U_n - Z_n - u)_{\mathcal{U}} + (p, B(U_n - Z_n) - B(u))_{\Psi} \geq 0$$

$$\implies (5.4.5) - (5.4.3) :$$

$$(5.4.6) \quad (\nabla J(U_n - Z_n) - \nabla J(u), U_n - Z_n - u)_{\mathcal{U}} + (P_n - p, B(U_n - Z_n) - B(u))_{\Psi} \leq 0$$

$$(5.4.6) \text{ et } \underline{J}_2 \implies$$

$$(5.4.7) \quad -(P_n - p, B(U_n - Z_n) - B(u))_{\Psi} \geq m ||U_n - Z_n - u||_{\mathcal{U}}^2$$

$$(5.4.1) \implies$$

$$(5.4.8) \quad ||R_{n+1}||_{\Psi}^2 \leq ||R_n||_{\Psi}^2 + 2\rho_n (R_n, B(U_n) - B(u))_{\Psi} + \rho_n^2 ||B(U_n) - B(u)||_{\Psi}^2.$$

D'après les hypothèses sur le bruit Z_n ,

$$E [B(U_n) | P_n] = E [B(U_n - Z_n + Z_n) | P_n] = B(U_n - Z_n)$$

$$\begin{aligned} E [(R_n, B(U_n) - B(u))_{\Psi}] &= (R_n, B(U_n) - B(u))_{\Psi} \\ &= (R_n, B(U_n) - B(U_n - Z_n))_{\Psi} + (R_n, B(U_n - Z_n) - B(u))_{\Psi} \end{aligned}$$

$$\begin{aligned} (R_n, B(U_n) - B(U_n - Z_n))_{\Psi} &= E \{ E [(R_n, B(U_n) - B(U_n - Z_n))_{\Psi} | P_n] \} \\ &= E \{ (R_n, E [B(U_n) | P_n] - B(U_n - Z_n))_{\Psi} \} \\ &= 0. \end{aligned}$$

$$\text{Donc } (R_n, B(U_n) - B(u))_{\Psi} = (R_n, B(U_n - Z_n) - B(u))_{\Psi}$$

On montre de même que

$$||B(U_n) - B(u)||_{\Psi}^2 = ||B(U_n) - B(U_n - Z_n)||_{\Psi}^2 + ||B(U_n - Z_n) - B(u)||_{\Psi}^2,$$

Donc (5.4.8) devient

$$\begin{aligned} ||R_{n+1}||_{\Psi}^2 \leq & ||R_n||_{\Psi}^2 + 2\rho_n (R_n, B(U_n - Z_n) - B(u))_{\Psi} + \rho_n^2 [||B(U_n) - B(U_n - Z_n)||_{\Psi}^2 \\ & + ||B(U_n - Z_n) - B(u)||_{\Psi}^2] \end{aligned}$$

(5.4.7) \implies

$$||R_{n+1}||_{\Psi}^2 \leq ||R_n||_{\Psi}^2 - 2\rho_n - m ||U_n - Z_n - u||_{\mathcal{U}}^2 + \rho_n^2 c^2 [||Z_n||_{\mathcal{U}}^2 + ||U_n - Z_n - u||_{\mathcal{U}}^2]$$

ou :

$$(5.4.9) \quad ||R_{n+1}||_{\Psi}^2 \leq ||R_n||_{\Psi}^2 + (\rho_n^2 c^2 - 2m\rho_n) ||U_n - Z_n - u||_{\mathcal{U}}^2 + c^2 \rho_n^2 ||Z_n||_{\mathcal{U}}^2$$

$$\rho_n \rightarrow 0 \implies 2m\rho_n - \rho_n^2 c^2 > 0 \quad \text{pour } n \text{ assez grand}$$

$$||Z(q)||_{\mathcal{U}}^2 \leq B^2 \implies ||Z_n||_{\mathcal{U}}^2 \leq B^2.$$

$$\text{D'où } ||R_n||_{\Psi}^2 - ||R_{n+1}||_{\Psi}^2 \geq (2m\rho_n - \rho_n^2 c^2) ||U_n - Z_n - u||_{\mathcal{U}}^2 - \rho_n^2 c^2 B^2.$$

Comme on a supposé que $\sum_n \rho_n^2 < +\infty$, $\|R_n\|_{\tilde{U}}^2$ a une limite d'après la proposition (1.5.2.2) et par suite (prop. (1.5.2.3))

$$-\sum_n (2m\rho_n - \rho_n^2 c^2) \cdot \|U_n - Z_n - u\|_{\tilde{U}}^2 < +\infty$$

et il existe une sous-suite telle que $\|U_{n_\ell} - Z_{n_\ell} - u\|_{\tilde{U}} \xrightarrow{\ell \rightarrow \infty} 0$

$$\|U_{n_\ell} - u\|_{\tilde{U}}^2 = \|U_{n_\ell} - Z_{n_\ell} - u\|_{\tilde{U}}^2 + \|Z_{n_\ell}\|_{\tilde{U}}^2 \quad \text{car } U_{n_\ell} - Z_{n_\ell} \text{ est mesurable}$$

$$\text{par rapport à } P_{n_\ell} \implies \limsup_\ell \|U_{n_\ell} - u\|_{\tilde{U}}^2 \leq B^2.$$

V.4.2 - Convergence de la variable duale

Théorème 5.4.2

Toute suite extraite de $\{P_{n_\ell}\}$ ($\{n_\ell\}$ définie au théorème (5.4.1) converge faiblement dans $\tilde{\Psi}$ vers P' tel que (u, P') soit point-col de $L(\tilde{v}, \tilde{q})$.

Démonstration :

Prouvons d'abord que $(B(U_n - Z_n), P_{n+1} - P_n)_{\tilde{\Psi}}$ est le terme général > 0 d'une série convergente.

Si u_q est défini par $(\nabla J(u_q), v - u_q)_{\tilde{U}} + (q, B(v) - B(u_q))_{\tilde{\Psi}} > 0 \quad \forall v \in \mathcal{U}_{ad}$, en posant $g(q) = \mathcal{L}(u_q, q)$, la 2ème inégalité de (5.1.4) prouve que $\nabla g(q) = B(u_q)$

$$g(P_{n+1}) - g(P_n) = (B(U_{P_n}), P_{n+1} - P_n)_{\tilde{\Psi}} + \int_0^1 \frac{d\theta}{\theta} (B(U_{P_n + \theta(P_{n+1} - P_n)}), P_n + \theta(P_{n+1} - P_n) - P_n)_{\tilde{\Psi}}.$$

En utilisant la définition de u_q d'une part avec $q = P_n + \theta(P_{n+1} - P_n)$

$$v = U_{P_n} \quad \text{et d'autre part } q = P_n, \quad v = U_{P_n + \theta(P_{n+1} - P_n)}$$

$$-c \|U_{P_n + \theta(P_{n+1} - P_n)} - U_{P_n}\|_{\tilde{U}} \|P_n + \theta(P_{n+1} - P_n) - P_n\|_{\tilde{\Psi}} \leq$$

$$(B(U_{P_n + \theta(P_{n+1} - P_n)}), P_n + \theta(P_{n+1} - P_n) - P_n)_{\tilde{\Psi}} \leq$$

$$\leq -(\nabla J(U_{P_n + \theta(P_{n+1} - P_n)} - \nabla J(U_{P_n}), U_{P_n + \theta(P_{n+1} - P_n)} - U_{P_n}) \mathcal{U} \leq$$

$$\leq -m \|U_{P_n + \theta(P_{n+1} - P_n)} - U_{P_n}\|^2 \leq 0$$

$$\implies \|U_{P_n + \theta(P_{n+1} - P_n)} - U_{P_n}\| \leq \frac{c\theta}{m} \|P_{n+1} - P_n\|_{\Psi}$$

$$\text{Donc } 0 \geq \int_0^1 \geq - \int_0^1 d\theta c \|U_{P_n + \theta(P_{n+1} - P_n)} - U_{P_n}\| \|P_{n+1} - P_n\|_{\Psi}$$

$$\geq - \frac{c^2}{2m} \|P_{n+1} - P_n\|_{\Psi}^2$$

$$\text{Donc (5.4.10) } g(P_{n+1}) - g(P_n) = (B(U_n - Z_n), P_{n+1} - P_n)_{\Psi}$$

$$- \theta_n^2 \frac{c^2}{2m} \|P_{n+1} - P_n\|_{\Psi}^2$$

θ_n^2 aléatoire tel que $0 \leq \theta_n^2 \leq 1$.

Nous allons maintenant prendre l'espérance mathématique après avoir relié

$$(B(U_n), P_{n+1} - P_n)_{\Psi} \text{ et } (B(U_n - Z_n), P_{n+1} - P_n)_{\Psi}$$

$P_{n+1} - P_n + \rho_n B(U_n - Z_n)$ est mesurable par rapport à $P_n \implies$

$$0 \leq (P_{n+1} - P_n, B(U_n) - B(U_n - Z_n))_{\Psi} = (P_{n+1} - P_n + \rho_n B(U_n - Z_n), B(U_n) - B(U_n - Z_n))_{\Psi}$$

$$\leq \rho_n \|B(U_n) - B(U_n - Z_n)\|_{\Psi}^2 \leq \rho_n c^2 \|Z_n\|^2 \leq \rho_n c^2 B^2$$

$$\|P_{n+1} - P_n\|_{\Psi}^2 \leq (P_{n+1} - P_n, \rho_n B(U_n))_{\Psi} = (P_{n+1} - P_n, \rho_n B(U_n - Z_n))_{\Psi} +$$

$$(P_{n+1} - P_n, \rho_n (B(U_n) - B(U_n - Z_n)))_{\Psi}$$

$$(5.4.12) \quad \|P_{n+1} - P_n\|_{\Psi}^2 \leq \rho_n (P_{n+1} - P_n, B(U_n - Z_n))_{\Psi} + \rho_n^2 B^2 c^2$$

$$\text{ou } \|P_{n+1} - P_n\|_{\Psi}^2 = \theta_n \rho_n (P_{n+1} - P_n, B(U_n - Z_n))_{\Psi} + \theta_n \rho_n^2 B^2 c^2$$

$$0 \leq \theta_n \leq 1.$$

En prenant l'espérance mathématique dans (5.4.10)

$$(5.4.13) \quad E[g(P_{n+1})] - E[g(P_n)] = \rho_n (P_{n+1} - P_n, B(U_n - Z_n))_{\Psi} \left[\frac{1}{\rho_n} - \rho_n \frac{C^2}{2m} \right] - \theta'_n \rho_n^2 B^2 C^2$$

$$0 \leq \theta_n \leq 1$$

$$\begin{aligned} (P_{n+1} - P_n, B(U_n - Z_n))_{\Psi} &= E \{ E [(P_{n+1} - P_n, B(U_n - Z_n))_{\Psi} \mid P_n] \} \\ &= E \{ (E(P_{n+1} \mid P_n) - P_n, B(U_n - Z_n))_{\Psi} \} \\ &= (P_n + \rho_n B(U_n - Z_n) - P_n, B(U_n - Z_n))_{\Psi} > 0. \end{aligned}$$

Donc (5.4.13) s'écrit $E[g(P_{n+1})] - E[g(P_n)] = v_n + w_n$

où $\sum_n |w_n| < +\infty$ $v_n \geq 0$ car $\left[\frac{1}{\rho_n} - \theta_n \frac{C^2}{2m} \right] > 0$ pour n assez grand.

D'après la proposition 1.5.2.2 $\sum v_n < +\infty$

et donc aussi (5.4.14) $\sum_n (P_{n+1} - P_n, B(U_n - Z_n))_{\Psi} < +\infty$.

Nous allons à présent montrer que $(R_{n_l}, B(U_{n_l} - Z_{n_l}))_{\Psi} \xrightarrow[n \rightarrow \infty]{} 0$.

Considérons la suite $\{n\}$ extraite au théorème (5.4.1) $\|R_n\|_{\Psi}^2$ a une limite.

Donc $\{R_n\}$ est bornée dans Ψ .

On peut extraire de $\{n_l\}$ une suite, soit $\{n'\}$ telle que $R_{n'} \xrightarrow{\sim} R$ dans Ψ faible.

D'après le théorème (5.4.1) $U_{n'} - Z_{n'} \xrightarrow{\sim} u$ dans \tilde{U} fort

$$\text{et } \|B(U_{n'} - Z_{n'}) - B(u)\|_{\Psi}^2 \xrightarrow{\sim} 0$$

Il résulte alors de (5.4.7) que

$$(5.4.15) \quad (R_{n'}, B(U_{n'} - Z_{n'}))_{\Psi} \xrightarrow{\sim} (R, B(u))_{\Psi}.$$

$P_n \in \tilde{\Lambda}$ convexe fermé, donc faiblement fermé.

Donc $R + p \in \tilde{\Lambda} \implies (5.4.16) \quad (R, B(u))_{\Psi} \leq 0$ (analogue à 5.1.4)

D'après (5.4.14), $(R_{n+1} - R_n, B(U_n - Z_n))_{\Psi} \xrightarrow[n \rightarrow \infty]{} 0$ et avec (5.4.15).

$$(5.4.17) \quad (R_{n'+1}, B(U_n, -Z_n))_{\Psi} \longrightarrow (R, B(u))_{\Psi}$$

Comme conséquence des propriétés de l'opérateur de projection P :

$$(P_n + \rho_n)_{\Lambda} - P_{n'+1}, z - P_{n'+1})_{\Psi} \leq 0 \quad \forall z \in \Lambda$$

avec $z = P$:

$$(5.4.18) \quad -(R_n, -R_{n'+1} + \rho_n, B(U_n), R_{n'+1})_{\Psi} \leq 0.$$

$$\text{Soit } \rho_n, (B(U_n), R_{n'+1})_{\Psi} \geq (R_{n'+1}, R_{n'+1} - R_n)_{\Psi} \geq$$

$$\geq \frac{1}{2} \|R_{n'+1}\|_{\Psi}^2 - \frac{1}{2} \|R_n\|_{\Psi}^2.$$

D'où

$$(5.4.19) \quad (B(U_n), R_{n'+1})_{\Psi} \geq \frac{1}{2\rho_n} \left[\|R_{n'+1}\|_{\Psi}^2 - \|R_n\|_{\Psi}^2 \right]$$

$$(5.4.12) \implies \frac{1}{\rho_n} \|R_{n'+1} - R_n\|_{\Psi}^2 \leq (B(U_n - Z_n), P_{n'+1} - P_n)_{\Psi} + \rho_n B^2 C^2 \xrightarrow{n \rightarrow \infty} 0$$

$$\|R_{n'+1}\|^2 \leq 2 \left[\|R_{n'+1} - R_n\|_{\Psi}^2 + \|R_n\|_{\Psi}^2 \right]$$

$$\implies \frac{1}{\rho_n} \|R_{n'+1} - R_n\|_{\Psi}^2 \geq \frac{1}{2\rho_n} \left[\|R_{n'+1}\|_{\Psi}^2 - \|R_n\|_{\Psi}^2 \right]$$

$$\text{c'est-à-dire (5.4.20) } \frac{1}{2\rho_n} \left[\|R_{n'+1}\|_{\Psi}^2 - \|R_n\|_{\Psi}^2 \right] \xrightarrow{n \rightarrow \infty} 0$$

$$(5.4.21) \quad (B(U_n - Z_n), R_{n'+1})_{\Psi} = (B(U_n), R_{n'+1})_{\Psi} - (B(U_n) - B(U_n - Z_n), R_{n'+1})_{\Psi}$$

(5.4.11) s'écrit

$$0 \leq (R_{n'+1} - R_n, B(U_n) - B(U_n - Z_n))_{\Psi} \leq \rho_n B^2 C^2.$$

$$\text{Mais } (R_n, B(U_n) - B(U_n - Z_n))_{\Psi} = 0.$$

Donc :

$$(5.4.22) \quad 0 \leq (R_{n'+1}, B(U_n) - B(U_n - Z_n))_{\Psi} \leq \rho_n B^2 C^2.$$

En portant (5.4.19) et (5.4.22) dans (5.4.21)

$$(B(U_n - Z_n), R_{n'+1})_{\Psi} \geq \frac{1}{2\rho_n} \left[\|R_{n'+1}\|_{\Psi}^2 - \|R_n\|_{\Psi}^2 \right] - \rho_n B^2 C^2.$$

Il résulte alors de (5.4.17) et (5.4.20) que

$$(5.4.23) \quad (R, B(u))_{\Psi} \geq 0$$

ce qui comparé à (5.4.15) donne

$$(5.4.24) \quad (R, B(u))_{\Psi} = 0.$$

Donc 0 est le seul point d'accumulation de la suite $(R_{n_l}, B(U_{n_l} - Z_{n_l}))_{\Psi}$

D'où

$$(5.4.25) \quad (R_{n_l}, B(U_{n_l} - Z_{n_l}))_{\Psi} \xrightarrow{l \rightarrow \infty} 0$$

Il en résulte que l'on a

$$(5.4.26) \quad (P_{n_l}, B(U_{n_l} - Z_{n_l}))_{\Psi} \xrightarrow{l \rightarrow \infty} (P, B(u))_{\Psi}.$$

Par définition de p, $\mathcal{L}(u, q) \leq \mathcal{L}(u, p) \quad \forall q \in \Lambda$

$$\iff (q, B(u))_{\Psi} \leq (p, B(u))_{\Psi} \quad \forall q \in \Lambda$$

$$\text{Donc} \quad (\tilde{q}, B(u))_{\Psi} \leq (p, B(u))_{\Psi} \quad \forall \tilde{q} \in \tilde{\Lambda}$$

et (5.4.26) s'écrit

$$(5.4.27) \quad (P_{n_l}, B(U_{n_l} - Z_{n_l}))_{\Psi} \xrightarrow{l \rightarrow \infty} \sup_{\tilde{q} \in \tilde{\Lambda}} (\tilde{q}, B(u))_{\Psi}.$$

Soit une sous-suite $\{P_{n_l}\}$ extraite de $\{P_{n_l}\}$ telle que

$$\begin{aligned} P_{n_l} &\longrightarrow P' \in \tilde{\Lambda} && \text{dans } \tilde{\Psi} \text{ faible} \\ U_{n_l} - Z_{n_l} &\longrightarrow u && \text{dans } \tilde{\mathcal{U}} \text{ fort.} \end{aligned}$$

D'où

$$(P_{n_l}, B(U_{n_l} - Z_{n_l}))_{\Psi} \longrightarrow (P', B(u))_{\Psi} = \sup_{\tilde{q} \in \tilde{\Lambda}} (\tilde{q}, B(u))_{\Psi}$$

c'est-à-dire

$$(5.4.28) \quad (\tilde{q} - P', B(u))_{\Psi} \leq 0 \quad \forall \tilde{q} \in \tilde{\Lambda}.$$

Soit

$$(5.4.29) \quad L(u, \tilde{q}) \leq L(u, P') \quad \forall \tilde{q} \in \tilde{\Lambda}.$$

Les hypothèses faites sur J impliquent :

$$\nabla J(U_{n_l} - Z_{n_l}) \longrightarrow \nabla J(u) \text{ dans } \tilde{\mathcal{U}} \text{ faible,}$$

(5.4.2) \implies

$$(\nabla J(U_{n_l} - Z_{n_l}), \tilde{v} - (U_{n_l} - Z_{n_l}))_{\tilde{u}} + (P_{n_l}, B(\tilde{v}) - B(U_{n_l} - Z_{n_l}))_{\Psi} \geq 0 \quad \forall \tilde{v} \in \tilde{u}_{ad}.$$

En passant à la limite :

$$(\nabla J(u), \tilde{v} - u)_{\tilde{u}} + (P', B(\tilde{v}) - B(u))_{\Psi} \geq 0 \quad \forall \tilde{v} \in \tilde{u}_{ad}$$

c'est-à-dire : $L(u, P') \leq L(\tilde{v}, P') \quad \forall \tilde{v} \in \tilde{u}_{ad}.$

V.4.3 - Comparaison avec les résultats antérieurs

L'algorithme d'Uzawa envisagé sur un plan stochastique n'a été l'objet à notre connaissance, d'aucune publication jusqu'à ce jour.

Il n'en est pas de même pour les algorithmes déterministes qui en dérivent.

Avec un choix de ρ_n différent ($\rho_n \geq \delta > 0$).

A. Bensoussan, J.L. Lions, R. Temam [1] démontrent des théorèmes déterministes comparables aux théorèmes (5.4.1) et (5.4.2).

Avec l'hypothèse $\rho_n \longrightarrow 0 \quad \sum_n \rho_n = +\infty$ (algorithme à petits pas),

A. Auslender [2] démontre la convergence de la variable primale dans le cas déterministe.

En revanche, nous n'avons pas connaissance de résultat déterministe analogue au théorème (5.4.2) pour l'algorithme à petit pas.

Un tel résultat est fourni par le théorème (5.4.2) évidemment encore valable dans le cas déterministe.

V.5 - Convergence de l'algorithme d'Arrow-Hurwicz stochastique

On considère l'algorithme (5.3.2) où on va choisir $\rho_n^{(1)} = \rho_n$

$\rho_n^{(2)} = k \rho_n$ vérifiant :

$$\rho_n > 0 \quad \sum_n \rho_n = +\infty, \quad \sum_n \rho_n^2 < +\infty, \quad \rho_n \leq \frac{m}{M^2}$$

V.5.1 - Convergence de la variable primale

Théorème 5.5.1

Soit u la solution du problème primal. Sous les hypothèses $\underline{J}_1, \underline{J}_2, \underline{J}_3, \underline{Q}, \underline{S}_1, \underline{S}_2$, il existe une sous-suite $\{U_{n_\ell}\}$ telle que $U_{n_\ell} \xrightarrow{\ell \rightarrow \infty} u$ en moyenne quadratique.

Démonstration :

D'après \underline{J}_2 et \underline{J}_3 , on a

$$(5.5.1) \quad ||v-u-\rho (\nabla J(v)-\nabla J(u))||_{\mathcal{U}}^2 \leq [1-2\rho m + \rho^2 M^2] ||v-u||_{\mathcal{U}}^2.$$

Comme $1-\rho m \geq 1-2\rho m + \rho^2 M^2$ dès que $\rho \leq \frac{m}{M^2}$

$$(5.5.2) \quad ||v-u-\rho (\nabla J(v)-\nabla J(u))||_{\mathcal{U}}^2 \leq (1-\rho m) ||v-u||_{\mathcal{U}}^2.$$

Soit (u,p) point-col de $\mathcal{L}(v,q)$

$$(5.5.3) \quad \begin{cases} u = u - \rho_n (\nabla J(u) + B^* p) \\ p = P_{\Lambda}(p - k\rho_n B u). \end{cases}$$

Posons $W_n = U_n - u, R_n = P_n - p$

$$W_{n+1} = W_n - \rho_n (\nabla J(U_n) - \nabla J(u) + B^* R_n + Z_n)$$

$$||R_{n+1}||_{\Psi}^2 = ||P_{\Lambda}(P_n + k\rho_n B U_{n+1}) - P_{\Lambda}(P - k\rho_n B u)||_{\Psi}^2$$

$$(5.5.4) \quad ||R_{n+1}||_{\Psi}^2 \leq ||R_n + k\rho_n B W_{n+1}||_{\Psi}^2 = ||R_n||_{\Psi}^2 + 2k\rho_n (R_n, B W_{n+1})_{\Psi} + \rho_n^2 k^2 ||B W_{n+1}||_{\Psi}^2$$

$$(5.5.5) \quad ||W_{n+1}||_{\mathcal{U}}^2 = (W_n - \rho_n (\nabla J(U_n) - \nabla J(u)), W_{n+1})_{\mathcal{U}} - \rho_n (B^* R_n, W_{n+1})_{\mathcal{U}} - \rho_n (Z_n, W_{n+1})_{\mathcal{U}}$$

$$E \left[-\rho_n (Z_n, W_{n+1})_{\mathcal{U}} \mid U_n \right] = E \left[-\rho_n (Z_n, W_n - \rho_n (\nabla J(U_n) - \nabla J(u) + B^* R_n + Z_n))_{\mathcal{U}} \mid U_n \right]$$

$$= \underbrace{\rho_n^2 E \left[\|z_n\|_{\mathcal{U}}^2 \mid U_n \right]}_{\leq B^2} + \underbrace{E \left[-\rho_n (z_n, w_n - \rho_n (\nabla J(U_n) - \nabla J(u)) + B^* R_n) \mid U_n \right]}_{= 0 \text{ (mesurabilité du 2ème terme par rapport à } U_n)}$$

Donc :

$$(5.5.6) \quad -\rho_n (z_n, w_{n+1})_{\mathcal{U}} \leq \rho_n^2 B^2$$

$$(5.5.5) \text{ et } (5.5.6) \implies \|w_{n+1}\|_{\mathcal{U}}^2 \leq (w_n - \rho_n (\nabla J(U_n) - \nabla J(u)), w_{n+1})_{\mathcal{U}} - \rho_n (B^* R_n, w_{n+1})_{\mathcal{U}} + \rho_n^2 B^2.$$

Puisque $\rho_n \leq \frac{m}{M^2}$, (5.5.2) peut être utilisé, soit en posant

$$\beta_n^2 = 1 - \rho_n^m, \quad \beta_n > 0$$

$$(5.5.7) \quad \|w_{n+1}\|_{\mathcal{U}}^2 \leq \beta_n \|w_n\|_{\mathcal{U}} \|w_{n+1}\|_{\mathcal{U}} - \rho_n (B^* R_n, w_{n+1})_{\mathcal{U}} + \rho_n^2 B^2$$

$$(5.5.4) \implies$$

$$(5.5.8) \quad \|R_{n+1}\|_{\Psi}^2 \leq \|R_n\|_{\Psi}^2 + 2k\rho_n (B^* R_n, w_{n+1})_{\mathcal{U}} + k^2 \rho_n^2 \|B^* w_{n+1}\|_{\Psi}^2.$$

$$\text{D'où } \|R_{n+1}\|_{\Psi}^2 \leq \|R_n\|_{\Psi}^2 + 2k \left[\beta_n \|w_n\|_{\mathcal{U}} \|w_{n+1}\|_{\mathcal{U}} - \|w_{n+1}\|_{\mathcal{U}}^2 + \rho_n^2 B^2 \right] + \rho_n^2 k^2 c^2 \|w_{n+1}\|_{\mathcal{U}}^2$$

$$(5.5.9) \quad \|R_{n+1}\|_{\Psi}^2 \leq \|R_n\|_{\Psi}^2 + k \left[\beta_n \|w_n\|_{\mathcal{U}}^2 - \beta_{n+1} \|w_{n+1}\|_{\mathcal{U}}^2 \right] + 2k\rho_n^2 B^2 + \left[\rho_n^2 k^2 c^2 - 2k + k(\beta_n + \beta_{n+1}) \right] \|w_{n+1}\|_{\mathcal{U}}^2.$$

$$\text{En posant } \tau_n = \|R_n\|_{\Psi}^2 + k\beta_n \|w_n\|_{\mathcal{U}}^2, \quad (5.5.9) \iff$$

$$(5.5.10) \quad \tau_n - \tau_{n+1} \geq \left[k(2 - (\beta_n + \beta_{n+1})) - \rho_n^2 k^2 c^2 \right] \|w_{n+1}\|_{\mathcal{U}}^2 - 2k\rho_n^2 B^2.$$

A fortiori, en posant $\gamma_n = k(1 - \beta_n) - \rho_n^2 k^2 c^2$

$$(5.5.11) \quad \tau_n - \tau_{n+1} \geq \gamma_n \|w_{n+1}\|_{\mathcal{U}}^2 - 2k\rho_n^2 B^2.$$

Puisque $\rho_n \longrightarrow 0 > \gamma_n \sim \frac{m}{2} \rho_n$ lorsque $n \longrightarrow +\infty$.

Donc (5.5.12) $\sum_n \gamma_n = +\infty$ et $\gamma_n > 0$ pour n assez grand.

D'après la proposition (1.5.2.2), τ_n a une limite

et $\sum_n \gamma_n \|w_{n+1}\|_{\mathcal{U}}^2 < +\infty$.

D'après la proposition (1.5.2.), 0 est point d'adhérence de $\{w_n\}$ i.e.

Il existe une sous-suite U_{n_ℓ} telle que

$U_{n_\ell} \xrightarrow[n \rightarrow +\infty]{} u$ en moyenne quadratique.

V.5.2 - Comparaison avec les résultats antérieurs. Problèmes ouverts concernant la convergence de la variable duale.

Comme pour la méthode d'Uzawa, nous n'avons pas connaissance de publications concernant la méthode d'Arrow-Hurwicz stochastique. Il est cependant utile de comparer les résultats obtenus aux résultats connus pour les algorithmes déterministes :

Algorithme à petit pas : $\rho_n \longrightarrow 0 \quad \sum_n \rho_n = +\infty$.

La convergence de l'algorithme déterministe est donnée par A. Auslender [2], mais nous ne connaissons pas de résultat déterministe concernant la convergence de la variable duale.

Algorithme à grand pas : $\rho_n \geq \delta > 0$.

La convergence de l'algorithme déterministe est donnée par J.L. Lions, R. Glowinski, R. Trémolières [1] dans le cas quadratique et (5.5.1) en permet la généralisation pour une fonctionnelle strictement convexe. On a alors convergence forte dans \mathcal{U} de toute la suite U_n .

Nous ne connaissons pas de résultat déterministe concernant la convergence de la variable duale.

Reprenant une idée de A. Bensoussan, J.L. Lions, R. Tremam [1] développée dans le cadre de l'algorithme d'Uzawa, démontrons le :

Théorème 5.5.2

La suite $\{P_n\}$ est bornée et toute suite extraite converge faiblement vers p' tel que (u, P') soit point-col de $\mathcal{L}(v, q)$.

Démonstration :

$u_n \longrightarrow u$ dans \mathcal{U} fort. Donc $Bu_n \longrightarrow B_u$ dans fort. La suite $\{r_n\}$ est bornée, il existe donc une sous-suite $\{r_{n'}\}$ telle que $r_{n'} \xrightarrow[n' \rightarrow +\infty]{} r$ dans Ψ faible.

$P_{n'} = r_{n'} + p \longrightarrow r+p$ dans Ψ faible.

Λ est convexe fermé, donc faiblement fermé.

Donc $r + p \in \Lambda$.

$$(q-p, B(u))_{\Psi} \leq 0 \quad \forall q \in \Lambda \implies$$

$$(5.5.13) \quad (r, B(u))_{\Psi} \leq 0.$$

Grâce à la convergence forte de $B(u_n)$ et la convergence faible de $r_{n'}$,

$$(5.5.14) \quad (B(u_{n'}), r_{n'})_{\Psi} \xrightarrow[n' \rightarrow +\infty]{} (B(u), r)_{\Psi}$$

Comme conséquence des propriétés de l'opérateur P_{Λ}

$$(5.5.15) \quad (P_{n'-1} + \rho_{n'-1} B(u_{n'}) - P_{n'}) - P_{n'}, z - P_{n'})_{\Psi} \leq 0 \quad \forall z \in \Lambda$$

Avec $z = p$:

$$(5.5.16) \quad \rho_{n'-1} (Bu_{n'}, r_{n'})_{\Psi} \geq (r_{n'}, r_{n'} - r_{n'-1})_{\Psi} \geq \frac{1}{2} [\|r_{n'}\|_{\Psi}^2 - \|r_{n'-1}\|_{\Psi}^2]$$

$$\{ \|r_{n'}\|_{\Psi}^2 \} \text{ converge } \rho_{n'-1} \geq \delta > 0.$$

Donc par passage à la limite

$$(5.5.17) \quad (r, B(u))_{\Psi} \geq 0.$$

Par comparaison avec (5.5.13)

$$(5.5.18) \quad (r, B(u))_{\Psi} = 0.$$

On peut maintenant achever la démonstration comme A. Bensoussan, J.L. Lions, R. Temam [1] p. 151-152.

Le problème pour trouver dans le cas stochastique un résultat analogue au théorème (5.4.2) est donc le passage de $\rho_n \geq \delta > 0$ à $\rho_n \longrightarrow 0$, $\sum \rho_n = +\infty$ et non pas le passage du cas déterministe au cas stochastique. Nous avons résolu cette question pour l'algorithme d'Uzawa, mais le problème reste ouvert pour l'algorithme d'Arrow-Hurwicz.

BIBLIOGRAPHIE

K.J. ARROW, L. HURWICZ, H. UZAWA

- [1] Studies in linear and non linear programming, Stanford University Press (1958)

A. AUSLENDER

- [1] Méthodes et théorèmes de dualité, RIRO (4ème année, n° R1, 1970, pp. 9-45)
- [2] Problèmes de minimax via l'analyse convexe et les inégalités variationnelles : théorie et algorithmes, Springer (1972)

A. BENSOUSSAN

- [1] Filtrage optimal des systèmes linéaires, Dunod (1971)
- [2] L'identification et le filtrage, Cahier IRIA n° 1 (février 1969)

A. BENSOUSSAN, J.L. LIONS, R. TEMAM

- [1] Sur les méthodes de décomposition, de décentralisation et de coordination et applications, Cahier IRIA n° 11 (Juin 1972)

J.P. BERTRAN

- [1] Optimisation stochastique dans un espace de Hilbert, méthode de gradient, C.R. Ac. Sc. Paris t. 276, série A, pp. 613-616 (1973)
- [2] Optimisation stochastique dans un espace de Hilbert, Colloque d'analyse numérique, Epinal (1972)
- [3] Optimisation stochastique dans un espace de Hilbert, méthode de séries divergentes, Colloque d'analyse numérique, La Colle Sur Loup (1973)
- [4] Méthodes stochastiques de dualité, Colloque d'analyse numérique, Gourette (1974)

[1] Intégration - livre VI - chapitre 1.4, Hermann A.S.I., 1175

[2] Espaces vectoriels topologiques,
Livre V - chapitre 3.5, Hermann A.S.I., 1220

J. CEA

[1] Optimisation - Théorie et Algorithmes, Dunod (1971)

J. CEA, R. GLOWINSKI

[1] Minimisation de fonctionnelles non différentiables,
IRIA INF/7105 (Mars 1971)

G. CHAVENT

[1] Analyse fonctionnelle et identification de coefficients répartis
dans les équations aux dérivées partielles,
Thèse Paris (1971)

J.P. COMER

[1] Some stochastic approximation procedures for use in Process control,
Ann. Math. Stat. vol. 35, n° 3, pp. 1136-1146 (1964)

J.L. DOOB

[1] Stochastic processes,
Wiley New-York (1953)

L.E. DUBBINS, D.A. FRIEDMAN

[1] A sharper form of the Borel-Cantelli lemma and the strong law,
Ann. Math. Stat. vol. 36, pp. 800-807 (1965)

M. ENGELHARDT

[1] On upper bounds for variances in stochastic approximation,
SIAM J. Appl. Math. vol. 24, N° 2, pp. 145-151 (1973)

B. FICHET

[1] Sur l'approximation et l'optimisation stochastique,
Thèse d'ingénieur docteur, Toulouse (1970)

E. GLADYSHEV

[1] On stochastic approximation. Theory of probability and its appli-
cations,
Vol. X, n° 2, pp. 275-278 (1965)

A.A. GOLDSTEIN

[1] Minimizing functionals on Hilbert spaces, computing methods
in optimization problems
Academic Press (1964)

O. HANS, M. DRIML

[1] Conditionnal expectations for generalized random variables
Trans. second Prague conf. on Info. Theory, Prague (1960)

P. KENNETH, M. SIBONY, J.P. YVON

[1] La méthode de pénalisation et ses applications aux problèmes de
contrôle optimal
Cahier IRIA, n°2, Algorithmes numériques d'optimisation (1970)

J. KIEFER, J. WOLFOWITZ

[1] Stochastic estimation of the maximum of a regression function,
Ann. Math. Stat. 23, pp. 462-466 (1962)

H.J. KUHNER

- [1] Stochastic approximation algorithms for the local optimisation of functions with nonunique stationary points, IEEE. Trans. Autom. Contr. vol n°17, pp. 646-654 (1972)

H.J. KUSHNER

- [2] Stochastic approximation algorithms for constrained optimisation problems, Brown University, Providence R.I, CDS. Rep. 72.1 et à paraître Ann. Statist.

H.J. KUSHNER, T. GAVIN

- [1] Stochastic approximation type methods for constrained systems : Algorithms and numerical results, IEEE Trans. Autom. Contr. vol. AC 19, pp. 349-357 (1974)

J.L. LIONS

- [1] Contrôle optimal de systèmes gouvernés par des équations aux dérivées partielles, Dunod (1968)

J.L. LIONS, R. TREMOLIERES, R. GLOWINSKI

- [1] Applications des méthodes d'optimisation (à paraître)

B. MARTINET

- [1] Maximisation d'une fonction concave connue de façon approchée. Application à la méthode duale de Rockafellar, Colloque d'analyse numérique, La Colle Sur Loup (1973)

M. METIVIER

- [1] Martingales à valeurs vectorielles, applications à la dérivation des mesures vectorielles, Ann. Inst. Fourier, Grenoble 17.2 (1967) pp. 175-208

E. MOURIER

- [1] Eléments aléatoires dans un espace de Banach, Thèse Paris (1952)

J. NEVEU

- [1] Calcul des probabilités, Masson - Paris (1970)

B.T. POLJAK

- [1] A general method of solving extremum problems, Soviet Math. Dokl. vol. 8, N° 3, pp. 593-597 (1967)

H. ROBBINS, S. MONRO

- [1] A stochastic approximation method, Ann. Math. Stat. vol n° 22, pp. 400-407 (1951)

R.T. ROCKAFELLAR

- [1] A dual approach to solving non linear programming problems by unconstrained optimization, Mathematical programmin, pp. 354-473 (1973)

JOCELYNE ROUYER

- [1] Propriétés stochastiques du gradient de la fonctionnelle coût dans la méthode de l'état adjoint lorsque l'observation est stochastique,
Rapport de DEA, Nancy (1972)

JOSEPH ROUYER

- [1] Méthodes de Frank-Wolfe appliquée à l'optimisation stochastique
Rapport de DEA, Nancy (1973)

J. SACKS

- [1] Asymptotic distributions of stochastic approximation procedures,
Ann. Math. Stat. vol. 29, pp. 373-405 (1958)

L. SCHEMETTERER

- [1] Sur l'itération stochastique,
Le calcul des probabilités et ses applications 87, pp. 55-63 (1958)
- [2] Multidimensional stochastic approximation
Multivariate analysis II - Proc. 2nd Int. Symp. Dayton Ohio -
Ac Press. New-York, pp. 443-460 (1969)

J.H. VENTER

- [1] On Dvoretzky stochastic approximation theorems,
Ann. Math. Stat. vol. 37, pp. 1534-1544 (1966)
- [2] On convergence of the Kiefer-Wolfowitz approximation procedures,
Ann. Math. Stat. vol. 38, pp. 1031-1036 (1967)