HAL
open science

Analyse de sensibilité dans une classe de processus de décision markoviens

Samuel Vercraene, Jean-Philippe Gayon, Fikri Karaesmen

To cite this version:

Samuel Vercraene, Jean-Philippe Gayon, Fikri Karaesmen. Analyse de sensibilité dans une classe de processus de décision markoviens. ROADEF 2013, Feb 2013, Troyes, France. hal-00789175

HAL Id: hal-00789175

https://hal.science/hal-00789175v1

Submitted on 16 Feb 2013

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

Analyse de sensibilité dans une classe de processus de décision markoviens

Samuel Vercraene[1], Jean-Philippe Gayon[1], Fikri Karaesmen[2]

[1] G-Scop ; Grenoble INP
46 avenue Félix Viallet, 38 031 Grenoble Cedex 1, France
{samuel.vercraene, jean-philippe.gayon}@grenoble-inp.fr
[2] Koç University
158, Rumeli Feneri Yolu, 34450 Sariyer, Istanbul, Turquie
fkaraesmen@ku.edu.tr

Mots-clés : *Processus de décision markovien, analyse de sensibilité, contrôle optimal.*

1 Introduction

En recherche opérationnelle stochastique certains problèmes de contrôle optimal peuvent être résolus grâce à une formulation en processus de décision markovien (MDP) [5]. Ainsi, pour certains d'entre eux, nous pouvons montrer que la politique optimale de gestion est caractérisée par un ou plusieurs seuils (points, courbes, surfaces de commutation) [2, 7].

Dans ce qui suit, nous présentons un système de file d'attente modélisant un problème de production et de stockage à deux étages [7, 8], afin d'illustrer l'effet des variations des paramètres de ce système sur la politique optimale de celui-ci. Ce système est décrit figure 1.

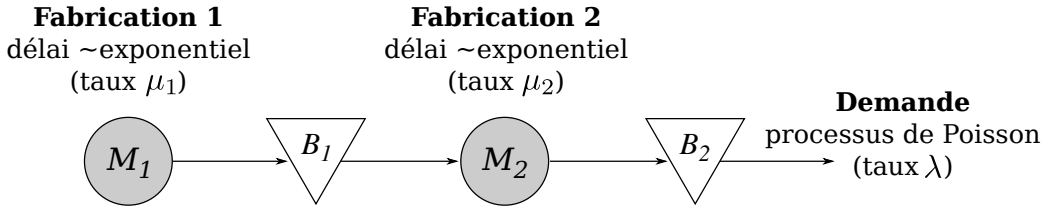
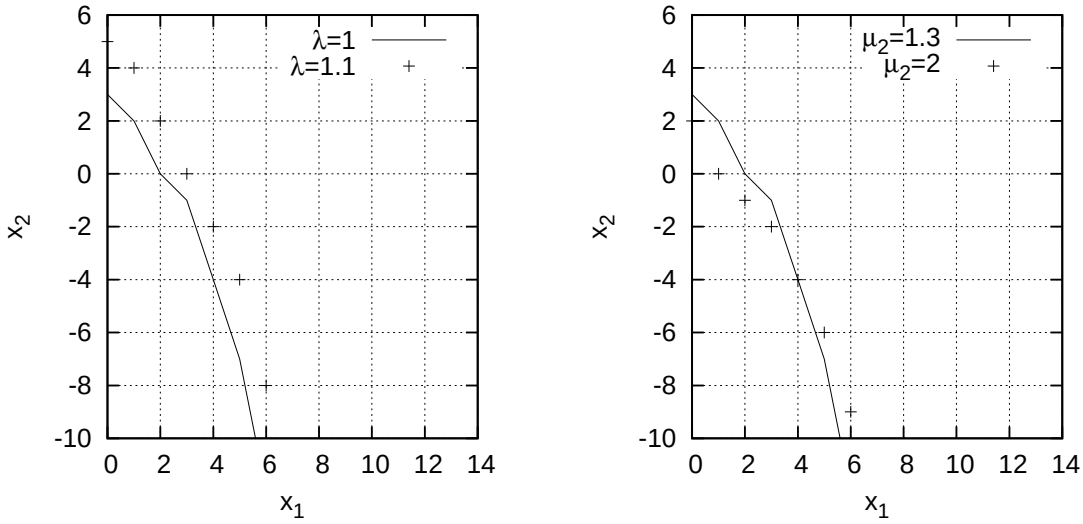


FIG. 1 – Système de production et de stockage à deux étages, avec production pour le stock.

Les serveurs M_1 et M_2 produisent les articles un par un avec des délais de traitement distribués selon des lois exponentielles, de taux respectivement μ_1 et μ_2. Un article produit au serveur i est stocké dans le stock B_i. Le stock final B_2 voit les demandes des clients arrivées selon un processus de Poisson de taux λ. Nous supposons que les commandes insatisfaites sont différées et que la préemption est autorisée. L'état du système est décrit par (x_1, x_2) avec x_1 le nombre de produits en cours de production dans le stock B_1 et x_2 le nombre de produits finis dans un stock B_2 moins le nombre de demande en attente. Le système paye un coût h_i par unité de temps et unité de produit dans le stock B_i et un coût b par unité de temps et unité de demande en attente. L'objectif est de minimiser l'espérance de coût actualisé sur un horizon infini. D'après [7] nous savons que la politique optimale existe et qu'elle est caractérisée par deux courbes de commutation $S_1(x_2)$ et $S_2(x_1)$.

La figure 2a (resp. 2b) montre l'influence du taux de demande λ (resp. taux de service μ_2) sur la courbe de commutation $S_1(x_2)$. Nous observons que la courbe $S_1(x_2)$ a un comportement homogène (augmentation) quand λ augmente, ce qui n'est pas le cas pour une augmentation de μ_2. Dans ce travail nous adressons la question suivante : pouvons-nous prédire l'effet de la variation d'un paramètre d'un système modélisé en MDP sur la politique optimale de celui-ci ?



(a) Effet de λ.

(b) Effet de μ_2.

FIG. 2 – Effet des paramètres du système sur la courbe de commutation $S_1(x_2)$ pour l'instance de base $\{\lambda = 1,\ \mu_1 = 2,\ \mu_2 = 1.3,\ h_1 = 1,\ h_2 = 3,\ b = 2\}$.

A notre connaissance, dans la littérature, seuls les travaux de [1] s'intéressent à la même problématique. Cependant leur étude se limite aux MDP dont l'espace d'état n'est décrit que par une seule variable (MDP mono-dimensionnels). Dans ce travail nous étendons leur étude à une classe de MDP multidimensionnels.

Dans la section 2 nous présentons la classe des MDP dans laquelle nous faisons notre étude. Dans la section 3 nous présentons le lien entre l'analyse de sensibilité et certaines propriétés, notamment la supermodularité. Enfin dans la section 4 nous concluons sur ces travaux et présentons quelques perspectives de recherche.

2 Formulation générique

Nous nous plaçons dans le cadre de la programmation dynamique basée sur les évènements [3] permettant de modéliser de façon standard une grande majorité des systèmes de file d'attente. Dans ce cadre nous nous intéressons aux MDP où l'équation d'optimalité peut s'écrire :

$$v^\star = \mathcal{T} v^\star, \quad \text{tel que} \quad \mathcal{T} v = C + \sum_{i=1}^{l} p_i A_i v, \tag{1}$$

avec $\mathbf{x}$ un vecteur quelconque de dimension n, v et C des fonctions de coût sur $\mathbf{x}$ à valeurs réelles, p_i la probabilité d'occurrence de l'événement i, et A_i l'opérateur sur v représentant l'évènement i. Un opérateur A_i pouvant être de deux formes (avec $\mathbf{a}$ un vecteur de dimension n) :

– un opérateur de translation

$$T^t_{\mathbf{a}} v(\mathbf{x}) = \begin{cases} v(\mathbf{x}+\mathbf{a}) & \text{si } \mathbf{x}+\mathbf{a} \in \mathbb{X}, \\ v(\mathbf{x}) + r & \text{sinon.} \end{cases} \tag{2}$$

– un opérateur de choix

$$T^c_{\mathbf{a}} v(\mathbf{x}) = \begin{cases} \min\{v(\mathbf{x}), v(\mathbf{x}+\mathbf{a}) + c\} & \text{si } \mathbf{x}+\mathbf{a} \in \mathbb{X}. \\ v(\mathbf{x}) + r & \text{sinon.} \end{cases} \tag{3}$$

L'opérateur de translation permet de modéliser des actions systématiques dans l'espace d'état, comme par exemple des arrivées de clients non contrôlées ou encore des départs de produits. Notons que le coût r sert à modéliser des ruptures de stock ou des capacités maximales atteintes. De la même façon l'opérateur de choix permet de modéliser des actions contrôlées comme la décision de démarrer/arrêter un serveur ou la décision d'accepter un client. Plus généralement ces deux opérateurs généralisent la majorité des opérateurs présentés par [4].

3 Analyse de sensibilité

Supposons que l'opérateur $\mathcal{T}$ est composé d'un opérateur de choix $T^c_{\mathbf{d}}$, alors la politique optimale sera fonction de la quantité $\Delta_{\mathbf{d}} v^\star(\mathbf{x}) = v^\star(\mathbf{x}+\mathbf{d}) - v^\star(\mathbf{x})$; car pour un système dans l'état $\mathbf{x}$, si $\Delta_{\mathbf{d}} v^\star(\mathbf{x}) \leq c$, alors il est optimal de faire évoluer le système vers l'état $\mathbf{x}+\mathbf{a}$ et si $\Delta_{\mathbf{d}} v^\star(\mathbf{x}) > c$, alors il est optimal de ne pas changer d'état.

Donc s'il existe un seuil séparant l'espace d'état en deux entre les états où le système doit évoluer et les états où le système ne doit pas évoluer, alors l'évolution de ce seuil sera fonction de l'évolution de la quantité $\Delta_{\mathbf{d}} v(\mathbf{x})$ quand le système est soumis à une perturbation. Ainsi, une augmentation (resp. diminution) de $\Delta_{\mathbf{d}} v(\mathbf{x})$ pour tout $\mathbf{x}$ quand le système est perturbé induit une diminution (resp. augmentation) homogène du seuil dans le direction $\mathbf{d}$.

En redéfinissant l'espace d'état et en incorporant les paramètres du système dans celui-ci, nous montrons que le comportement homogène de la courbe de commutation peut être démontré analytiquement en montrant que la fonction de valeur optimale $v^\star$ est supermodulaire dans les directions $\mathbf{d}$ et $\boldsymbol{\epsilon}$ [6]. La propriété de supermodularité sur v est notée $\mathtt{S}_{\mathbf{d},\boldsymbol{\epsilon}}$ et est vraie si et seulement si $\Delta_{\mathbf{d}}\Delta_{\boldsymbol{\epsilon}} v \geq 0$.

Pour montrer que $v^\star$ est supermodulaire dans les directions $\mathbf{d}$ et $\boldsymbol{\epsilon}$, nous utilisons la démarche suivante. Premièrement nous montrons que $\mathcal{T}$ est contractant [4], donc que toute suite $v^{k+1} = Tv^k$ converge vers $v^\star$, l'unique solution de l'équation $v^\star = \mathcal{T} v^\star$. Ainsi, si $\mathcal{T}$ propage une propriété de supermodularité alors $v^\star$ sera supermodulaire. Ensuite, nous trouvons des conditions sur les opérateurs T^t et T^c pour que $\mathcal{T}$ propage la propriété de supermodularité dans les directions $\mathbf{d}$ et $\boldsymbol{\epsilon}$. Plus précisément, sachant que $\Delta_{\mathbf{d}}\Delta_{\boldsymbol{\epsilon}} v \geq 0$, nous déterminons des conditions suffisantes pour avoir

$$\Delta_{\mathbf{d}}\Delta_{\boldsymbol{\epsilon}} T^t_{\mathbf{a}} v \geq 0, \tag{4}$$

$$\Delta_{\mathbf{d}}[T^t_{\mathbf{a}} v - v] \geq 0, \tag{5}$$

$$\Delta_{\mathbf{d}}\Delta_{\boldsymbol{\epsilon}} T^c_{\mathbf{a}} v \geq 0, \tag{6}$$

$$\Delta_{\mathbf{d}}[T^c_{\mathbf{a}} v - v] \geq 0, \tag{7}$$

quel que soit l'espace d'état, quel que soit la perturbation $\boldsymbol{\epsilon}$ et quel que soient les vecteurs $\mathbf{a}$ et $\mathbf{d}$ (de dimension n). L'expression de ces conditions constitue notre principal théorème.

Les conditions que nous obtenons sont écrites sous forme booléenne pour faciliter leur lecture. La limitation de page ne nous permet pas de présenter l'intégralité de ces résultats, cependant dans le but d'illustrer leur structure, nous présentons la condition obtenue pour que l'équation (4) soit vérifiée :

$$(\mathtt{S}_{\mathbf{d}-\mathbf{a},\boldsymbol{\epsilon}} . |\epsilon_r \geq 0| + \mathtt{R}_{\mathbf{a},\mathbf{d}}).(\mathtt{S}_{\mathbf{d}+\mathbf{a},\boldsymbol{\epsilon}} . |\epsilon_r \leq 0| + \mathtt{R}_{\mathbf{a},-\mathbf{d}}) \tag{8}$$

Dans cette équation, "." et "+" représentent respectivement les opérateurs booléens "et" et "ou", $|\epsilon_r \geq 0|$ est une condition sur la perturbation agissant sur le paramètre r, $\mathtt{R}_{\mathbf{a},\mathbf{d}}$ est une condition sur la forme de l'espace d'état et $\mathtt{S}_{\mathbf{d}-\mathbf{a},\boldsymbol{\epsilon}}$ est une condition de supermodularité sur v.

Grâce à ce théorème, nous montrons pour l'exemple présenté en introduction, l'influence des variations des paramètres sur les courbes de commutation optimales $S_1(x_2)$ et $S_2(x_1)$. Ces résultats sont donnés dans le tableau 1. Notons qu'ils justifient le comportement observé figure 2a.

Bien que nouveaux, les résultats présentés dans le tableau 1 ne sont qu'un exemple d'application de nos travaux et nous insistons sur le fait que les résultats obtenus dans nos théorèmes

Paramètre (p)	Augmente ($\epsilon_p > 0$)	Diminue ($\epsilon_p < 0$)
λ	Augmente	Diminue
h_1	Diminue si $\epsilon_{h_1} \leq \min\{\epsilon_{h_2}, -\epsilon_b\}$	Augmente si $\epsilon_{h_1} \geq \max\{\epsilon_{h_2}, -\epsilon_b\}$
h_2	Diminue	Augmente
b	Augmente	Diminue

TAB. 1 – Évolution des courbes de commutation $S_1(x_2)$ et $S_2(x_1)$ en fonction des variations des paramètres p

s'appliquent pour tous les MDP se formulant comme une combinaison convexe d'opérateur de translation et d'opérateur de choix (voir section 2).

4 Conclusion

Nous observons que des outils de la programmation dynamique stochastique permettent la caractérisation de la politique optimale dans de nombreux problèmes. Les travaux présentés dans ce document visent à développer ces outils. Nous présentons un cadre général pour étudier l'effet des paramètres d'un système formulé en MDP sur la politique optimale de celui-ci. Nos contributions dans ce travail sont diverses, tant d'un point de vue de la généricité de la formulation adoptée que des résultats obtenus. Les pistes de recherches sont nombreuses. Parmi elles, l'étude d'un opérateur générique permettant de modéliser des fréquences d'occurrence d'événements fonction de l'espace d'état du système nous parait particulièrement prometteuse. Ceci nous permettrait notamment de modéliser des problèmes comportant plusieurs serveurs en parallèle ou encore de faire de la tarification variable de produit. Pour d'autres pistes de recherche envisagées, comme par exemple l'étude d'un opérateur de choix avec trois choix possibles, les démonstrations nécessaires à notre étude semble très longues et répétitives. C'est pourquoi l'automatisation de ces preuves semble une approche intéressante pour les traiter.

Références

[1] Eren Çil, E. Örmeci, and Fikri Karaesmen. Effects of system parameters on the optimal policy structure in a class of queueing control problems. *Queueing Systems*, 61 :273–304, 2009.

[2] Albert Y. Ha. Inventory Rationing in a Make-to-Stock Production System with Several Demand Classes and Lost Sales. *MANAGEMENT SCIENCE*, 43(8) :1093–1103, 1997.

[3] Ger Koole. Structural results for the control of queueing systems using event-based dynamic programming. *Queueing Systems*, 30(3) :323–339, December 1998.

[4] Ger Koole. Monotonicity in markov reward and decision chains : Theory and applications. *Foundations and Trends in Stochastic Systems*, 1 :1–76, January 2006.

[5] Martin L. Puterman. *Markov Decision Processes, Discrete stochastic, Dynamic programming.* ed. Jhon Wiley & Sons, Inc., 1994.

[6] Donald M. Topkis. *Supermodularity & Complementarity.* 1998.

[7] Michael H. Veatch and Lawrence M. Wein. Monotone control of queueing networks. *Queueing Systems*, 12(3) :391–408, 1992.

[8] Michael H. Veatch and Lawrence M. Wein. Optimal control of a Two-Station tandem Production/Inventory system. *Operations Research*, 42(2) :337–350, April 1994.