Optimalité des politiques de rabais récurrents pour un modèle bi-niveau avec inertie des clients

Quentin Jacquet[1,2], Wim van Ackooij[1], Clémence Alasseur[1] Stéphane Gaubert[2]

[1] EDF Lab Saclay, Palaiseau, France

{quentin.jacquet, wim.van-ackooij,clemence.alasseur}@edf.fr

[2] INRIA, CMAP, Ecole polytechnique, IP Paris, CNRS, Palaiseau, France

stephane.gaubert@inria.fr

Mots-clés : *Bi-niveau dynamique, contrôle ergodique, tarification, limite champ moyen de processus de décision Markoviens.*

1 Introduction

On considère un modèle de tarification dynamique, dans lequel une population de clients (*followers*) peut changer à tout moment de contrat en fonction des conditions tarifaires et de caractéristiques propres à chaque client, comme l'inertie (propension à rester chez le même fournisseur). Un fournisseur (*leader*) cherche alors à maximiser son revenu moyen par unité de temps, en supposant que la population est de taille infinie. Cela revient à étudier une limite “champ moyen” d'une famille de processus de décision Markoviens [1], ce qui nous ramène ici à un processus de décision déterministe mais dont l'espace d'état est un produit de simplexes.

Nous supposons que, le signal de prix étant fixé, la population évolue selon une dynamique linéaire positive, issue de la maximisation d'une fonction d'utilité (cela inclue notamment les réponses de type *logit*, see [4]). En ce sens, le processus est vu comme un problème bi-niveau itéré dans le temps.

2 Contribution

Résolution via l'équation ergodique. Nous considérons le problème ergodique [2], dans lequel le fournisseur cherche à maximiser son revenu moyen par unité de temps. Ce problème peut être abordé au moyen du problème spectral non-linéaire, dans lequel on cherche une constante g et une fonction h, appelée bias ou potentiel, définie sur un sous-ensemble convenable Ω d'un produit de simplexes, telle que :

$$g + h(\mu) = \max_{a \in A} \Big(R(\mu, a) + h(\mu P(a)) \Big), \qquad \mu \in \Omega$$

où

- $\diamond$ μ représente la répartition de la population sur l'ensemble des contrats,
- $\diamond$ A l'ensemble des actions possibles pour le fournisseur (prix des contrats),
- $\diamond$ $\mu \mapsto \mu P(a)$ décit l'évolution de la population de clients sous l'influence des tarifs,
- $\diamond$ et R le revenu instantané.

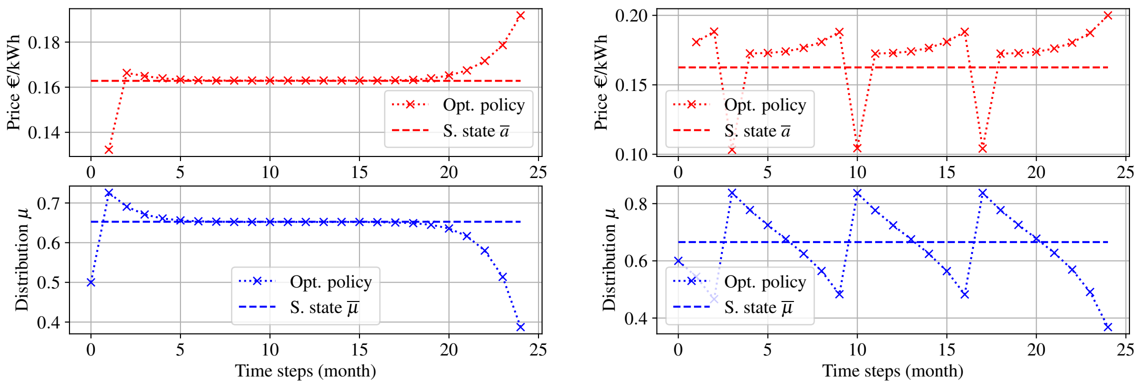
Lorsque ce problème admet une solution (g, h) telle que h soit bornée, la constante g fournit le gain moyen optimal du fournisseur. En outre, pour une distribution μ donnée, une stratégie optimale du fournisseur est obtenue en choisissant une action a atteignant le maximum dans cette équation. Nous montrons, sous une hypothèse de primitivité de l'ensemble des dynamiques, et en exploitant des proprités de contraction pour la métrique de Hilbert, que l'équation ergodique admet effectivement une solution (g, h) telle que la fonction biais h soit convexe et Lipschitzienne.

Calcul des *steady-states*. Nous nous intéressons ensuite au cas des stratégies obtenues en imposant des prix constants au cours du temps. La dynamique converge alors vers un état stationnaire, que nous déterminons analytiquement pour des modèles de type logit avec inertie.

Application à la tarification de l'électricité. Nous présentons enfin une application issue de la tarification de l'énergie, avec 1 ou 2 contrats et 1 ou 2 segments de clients, amenant à résoudre un problème ergodique à espace d'état continu allant jusqu'à la dimension 4, ce que nous faisons en appliquant un algorithme d'itération sur les politiques, après une discrétisation semi-Lagrangienne.

Pour une inertie faible, la politique optimale obtenue correspond à une stratégie *turnpike*, où le fournisseur cherche à atteindre rapidement une stratégie à prix constant. En revanche, lorsque l'inertie de certains clients est suffisante, la stratégie optimale donne lieu à des politiques de tarification cycliques, s'interprétant comme des rabais périodiques, voir Figure 1.

Pour plus de détails, nous renvoyons au travail [3], à paraître dans les actes du IEEE CDC 2022.



(a) Trajectoire optimale pour un horizon de temps fini (action du fournisseur et répartition des consommateurs) pour de *faibles* coûts de changement

(b) Trajectoire optimale pour un horizon de temps fini (action du fournisseur et répartition des consommateurs) pour de *forts* coûts de changement

FIG. 1 – Resultats numérique en horizon de temps fini. *Faible* (resp. *fort*) coûts de changement signifie ici $\gamma = 20$ (resp. 25).

Références

[1] Nicole BÄUERLE. *Mean Field Markov Decision Processes.* 2021. DOI : 10.48550/ARXIV.2106.08755. arXiv : 2106.08755 [math.OC].

[2] Eduardo GARIBALDI et Philippe THIEULLEN. "Minimizing orbits in the discrete Aubry–Mather model". In : *Nonlinearity* 24 (2011), p. 563-611.

[3] Quentin JACQUET, Wim van ACKOOIJ, Clémence ALASSEUR et Stéphane GAUBERT. *Ergodic control of a heterogeneous population and application to electricity pricing.* 2022. DOI : 10.48550/ARXIV.2204.01410.

[4] Polykarpos PAVLIDIS et Paul B. ELLICKSON. "Implications of parent brand inertia for multiproduct pricing". In : *Quantitative Marketing and Economics* 15.4 (juil. 2017), p. 369-407. DOI : 10.1007/s11129-017-9187-8.