Journées SMAI MODE 2018

Un algorithme stochastique de type min-plus pour un problème de contrôle optimal déterministe

Benoît Tran
École Nationale des Ponts et Chaussées - CERMICS

Marianne Akian
École Polytechnique - CMAP et INRIA

Jean-Philippe Chancelier
École Nationale des Ponts et Chaussées - CERMICS

Résumé. Nous étudions un problème de contrôle optimal discret en horizon fini, de dynamique linéaire, de coûts instantanés et final quadratiques et de contrôles à la fois continus et discrets. Nous développons un algorithme et prouvons sa convergence en s'inspirant du travail effectué dans [1] et [2]. Les opérateurs de Bellman associés à l'équation de la programmation dynamique de notre problème sont min-plus linéaires sur les formes quadratiques, permettant ainsi de déterminer explicitement les fonctions valeurs comme des infimums finis de formes quadratiques. Une telle approche n'est pas soumise à la fameuse malédiction de la dimension, toutefois le nombre de formes quadratiques nécessaires au calcul de l'infimum croît exponentiellement avec le temps.
L'algorithme proposé ici sélectionne certaines formes quadratiques impliquées dans le calcul de l'infimum en testant si une forme quadratique améliore l'approximation courante en un point tiré aléatoirement uniformément sur la sphère. Dans notre cadre, le calcul explicite des opérateurs de Bellman à contrôle discret fixé est donné par une équation de Riccati et nous prouvons la convergence de notre algorithme en nous basant sur l'existence d'un ensemble borné, pour l'ordre de Loewner sur les matrices symétriques, stable par tous les opérateurs de Bellman.

Mots-clefs : Contrôle optimal, opérateurs de Bellman, méthode min-plus, équation de Riccati.

1 Présentation du contexte et remarques générales

Soit $\mathcal{M}$ un ensemble de contrôles de cardinal fini. On s'intéresse dans ce qui suit a un problème d'optimisation linéaire-quadratique en temps discret d'horizon fini mélangeant contrôles continus et discrets :

$$\begin{aligned} &\min_{\substack{x\in(\mathbb{R}^n)^T \\ (u,\mu)\in(\mathbb{R}^m\times\mathcal{M})^T}} \sum_{t=0}^{T-1} c_t^{\mu_t}(x_t,u_t) + \phi(x_T) \\ &\text{t.q.} \quad \begin{cases} x_{t+1} = f_t^{\mu_t}(x_t,u_t) \\ x_0 \in \mathbb{R}^n \text{ donné,} \end{cases} \end{aligned} \tag{1}$$

où on effectue les hypothèses suivantes.

Hypothèses 1. *A chaque temps $t \in \{0, \dots, T-1\}$ et pour tout contrôle discret $\mu \in \mathcal{M}$:*

1. *La dynamique $f_t^\mu : (\mathbb{R}^n, \mathbb{R}^m) \longrightarrow \mathbb{R}^n$ est linéaire :*
$$f_t^\mu(x, u) = A_t^\mu x + B_t^\mu u,$$
pour des matrices A_t^μ et B_t^μ de dimensions adéquates.
2. *La fonction de coût $c_t^\mu : (\mathbb{R}^n, \mathbb{R}^m) \longrightarrow \mathbb{R}$ est une forme quadratique pure,* i.e.
$$c_t^\mu(x, u) = x^T C_t^\mu x + u^T D_t^\mu u,$$
où la matrice C_t^μ est symétrique semi-définie positive et la matrice D_t^μ est symétrique définie positive.
3. *La fonction de coût final $\phi : \mathbb{R}^n \longrightarrow \mathbb{R}$ est une forme quadratique pure et convexe.*
4. *La plus grande valeur propre de la matrice symétrique réelle $\left(A_t^\mu\right)^T A_t^\mu$ est inférieure à 1.*

La solution du problème (1) peut-être obtenue par programmation dynamique en calculant de façon récursive en temps rétrograde la suite de fonction valeurs
$$\begin{cases} V_T = \phi \\ \forall t \in \{0, \dots, T-1\}, V_t : x \mapsto \mathcal{B}_t(V_{t+1})(x)\,, \end{cases}$$
où les *opérateurs de Bellman* $\mathcal{B}_t$, pour $t \in \{0, \dots, T-1\}$, sont définis de la façon suivante :
$$\begin{cases} \mathcal{B}_t^\mu(\phi) := \inf_{u \in \mathbb{R}^m} c_t^\mu(\cdot, u) + \phi\left(f_t^\mu(\cdot, u)\right) \\ \mathcal{B}_t := \inf_{\mu \in \mathcal{M}} \mathcal{B}_t^\mu. \end{cases}$$
Sous les hypothèses 1, pour un contrôle discret $\mu \in \mathcal{M}$ fixé et un temps discret t fixé, l'opérateur $\mathcal{B}_t^\mu$ transforme les formes quadratiques pures en forme quadratiques pures avec une forme explicite du résultat (équation de Riccati) pour les formes quadratiques semi définies positives. On peut ainsi considérer $\mathcal{B}_t^\mu$ comme un opérateur défini sur les matrices symétriques réelles $\mathcal{S}_n$ a à valeurs dans $\mathcal{S}_n$ qui est aussi *min-plus linéaire* et *monotone*.
Notons que de la min-plus linéarité des opérateurs $\mathcal{B}_t^\mu$ on déduit l'existence pour tout temps t d'un ensemble non-vide Q_t de formes quadratiques pures telles que :
$$V_t = \inf_{q \in Q_t} q$$
Toutefois le cardinal de Q_t croît exponentiellement lorsque t décroît vers 0, ainsi un calcul naïf des fonctions valeurs est trop coûteux. Nous allons construire un algorithme qui approxime la fonction valeur V_t au temps $t \in \{0, \dots, T-1\}$ par une fonction $V_t^{(k)}$ définie comme un infimum sur un ensemble bien choisi $Q_t^{(k)}$ inclus dans Q_t :
$$V_t^{(k)} := \inf_{q \in Q_t^{(k)}} q \approx \inf_{q \in Q_t} q =: V_t.$$
L'originalité de notre démarche (déjà mise en avant dans [1] et [2]) réside en la caractérisation d'un ensemble borné pour l'ordre de Loewner sur $\mathcal{S}_n$, qui est invariant sous l'action de n'importe quel opérateur de Bellman $\mathcal{B}_t^\mu$: c'est l'existence de cet ensemble qui nous assurera la convergence de notre algorithme.

2 Un algorithme stochastique de type min-plus

Algorithme 2. *Nous construisons itérativement pour $k \in \mathbb{N}$ une suite d'approximations $\left\{V_t^{(k)}\right\}_{t\in\{0,\dots,T-1\}}$ des fonctions valeurs. Pour t et k fixé la fonction $V_t^{(k)}$ est un infimum sur un ensemble fini de formes quadratiques pures $Q_t^{(k)} \subset Q_t$.*

1. *La fonction de coût final ϕ étant une fonction quadratique pure, pour tout entier $k \geq 0$, on définit $V_T^{(k)}$ par $V_T^{(k)} := \phi$.*
2. *On initialise l'algorithme pour $k = 0$, en posant pour tout temps $t \in \{0, \dots, T-1\}$, $Q_t^{(0)} := \emptyset$ et $V_t^{(0)} := +\infty$.*
3. *Supposons, pour k fixé, la suite $\left\{V_t^{(k)}\right\}_{t\in\{0,\dots,T-1\}}$ construite et donnée sous la forme*

$$V_t^{(k)} = \min_{q\in Q_t^{(k)}} q \quad \textit{avec} \quad Q_t^{(k)} \subset Q_t.$$

   On construit la suite $\left\{V_t^{(k+1)}\right\}_{t\in\{0,\dots,T-1\}}$ de façon retrograde en temps en commençant par $t := T-1$ de la façon suivante :

   (a) On tire uniformément sur la sphère unité de $\mathbb{R}^n$ un point $X_t^{(k)}$.

   (b) On détermine un couple d'indices $(i^*, \mu^*) \in \{1, \dots, |Q_{t+1}^{(k)}|\} \times \mathcal{M}$ tels que

$$(i^*, \mu^*) \in \operatorname*{arg\,min}_{\{1,\dots,|Q_t^{(k)}|\}\times\mathcal{M}} \mathcal{B}_t^{\mu}\left(V_{t+1}^{(k,i)}\right)\left(X_t^{(k)}\right),$$

   où $V_{t+1}^{(k,i)}$ est la i-ième forme quadratique pure composant $Q_{t+1}^{(k)}$.

   (c) Si $\mathcal{B}_t^{\mu^*}\left(V_{t+1}^{(k,i^*)}\right)\left(X_t^{(k)}\right) + \delta < V_t^{(k)}\left(X_t^{(k)}\right)$, où $\delta > 0$ est un paramètre de précision fixé, alors on pose :

$$\begin{cases} Q_t^{(k+1)} := Q_t^{(k)} \cup \left\{\mathcal{B}_t^{\mu^*}\left(V_{t+1}^{(k,i^*)}\right)\right\} \quad \subset Q_t \\ V_t^{(k+1)} = \min_{q\in Q_t^{(k+1)}} q. \end{cases}$$

   Sinon, poser $Q_t^{(k+1)} = Q_t^{(k)}$ et $V_t^{(k+1)} = V_t^{(k)}$.

   (d) Retourner à l'étape a) en utilisant $t' = t-1$ tant que $t' < 0$.

3 Convergence théorique de notre algorithme

Théorème 3. *Soit $\delta > 0$ un paramètre de précision et notons $\left(V_T^{(k)}, \dots, V_0^{(k)}\right)_{k\geq 0}$ une suite de surapproximations des fonctions valeurs $(V_T, \dots, V_0)$, générées par l'algorithme 2. Notons pour chaque temps $t \in \{0, \dots, T-1\}$, $\mathbb{P}_t$ la probabilité sur les tirages des points effectués au temps t par l'algorithme 2 et notons $\|\cdot\|_\infty$ la norme infinie sur la sphère unité de $\mathbb{R}^n$. Alors la suite $\left(V_t^{(k)}\right)_{k\geq 0}$ converge uniformément sur la sphère unité et*

$$\lim_{k\to+\infty} \mathbb{P}_t\left(\|V_t^{(k)} - V_t\|_\infty \leq \delta\right) = 1.$$

Références

[1] Z. Qu. *Nonlinear Perron-Frobenius theory and max-plus numerical methods for Hamilton-Jacobi equations*. 2013 Thèse, Ecole Polytechnique.

[2] Z. Qu. *A max-plus based randomized algorithm for solving a class of HJB PDEs*. 2014 IEEE 53rd IEEE Conference on Decision and Control.