Operations Research Letters 44 (2016) 297–301

Contents lists available at ScienceDirect

Operations Research Letters

journal homepage: www.elsevier.com/locate/orl

Singularly perturbed linear programs and Markov decision processes

Konstantin Avrachenkov [a,*], Jerzy A. Filar [b], Vladimir Gaitsgory [c], Andrew Stillman [b]

[a] *Inria Sophia Antipolis, 2004 Route des Lucioles, 06902 Sophia Antipolis, France*

[b] *School of Computer Science, Engineering & Mathematics, Flinders University, Australia*

[c] *Department of Mathematics, Macquarie University, Australia*

ARTICLE INFO

Article history:
Received 26 November 2015
Accepted 4 February 2016
Available online 16 February 2016

Keywords:
Markov Decision Processes (MDPs)
Discounted MDPs
Long-run average MDPs
Singularly perturbed linear programs
Limiting linear program

ABSTRACT

Linear programming formulations for the discounted and long-run average MDPs have evolved along separate trajectories. In 2006, E. Altman conjectured that the two linear programming formulations of discounted and long-run average MDPs are, most likely, a manifestation of general properties of singularly perturbed linear programs. In this note we demonstrate that this is, indeed, the case.

1. Introduction

The connection between linear programming and Markov Decision Processes (MDPs) was launched in the 1960s, with the papers by D'Epenoux [8], De Ghellinck [7] and Manne [16]. While the linear programming formulation for the discounted MDP was relatively straightforward, extension to the long-run average, multi-chain, MDP proved challenging and required nearly two decades to arrive at a single linear program supplied, by Hordijk and Kallenberg [11,12], that completely solves such a multi-chain MDP. We refer the reader to Kallenberg [14,15], Puterman [19] and Altman [1] for excellent, comprehensive, treatments of linear programming methods for discrete time Markov decision processes. Even though the approaches to discounted and long-run average MDPs evolved along separate trajectories, Tauberian theorems provided a theoretical connection between the two cases with the discount parameter approaching unity from below; e.g. see Blackwell [6] and Veinott [20,21].

Parametric linear programming has a long history that is well documented in many excellent textbooks (e.g., see Murty [17]). However, majority of the so-called sensitivity analyses presented in operations research books focus on perturbations of the objective function coefficients or of the right hand side vector; sometimes extending also to changes in non-basic columns. To the best of our knowledge, Jeroslow [13] was, perhaps, the first to consider perturbations of the entire coefficient matrix of a linear program. In the context of MDP, the results of [13] have been applied to Blackwell optimality [10] and to perturbed MDPs [2,4]. In Pervozvanskii and Gaitsgori [18] the authors focus on the singularly perturbed case where a discontinuity can arise as the perturbation parameter approaches a critical value. In the latter and in the more recent book by Avrachenkov et al. [4] the main cause of that discontinuity has been the change in the rank of the coefficient matrix at the critical value of the perturbation parameter. Hence, it was perhaps surprising that such discontinuities can also arise when the rank does not change, as shown very recently in [3].

In 2006, Eitan Altman conjectured that the two linear programming formulations of discounted and long-run average MDPs must be a manifestation of some general properties of singularly perturbed linear programs. In this note we demonstrate that this is, indeed, the case by first extending the results in [3] and then formally applying new singular perturbation results to the MDP problem.

2. General perturbed linear programming problem

Consider the family of linear programming problems parameterized by $\varepsilon > 0$:

$$\begin{aligned} &\max\langle c^{(0)} + \varepsilon c^{(1)}, x\rangle \\ \text{s. t. } &(A^{(0)} + \varepsilon A^{(1)})x = b^{(0)} + \varepsilon b^{(1)}, \\ &x \geq 0, \end{aligned} \tag{1}$$

where $c^{(0)}, c^{(1)} \in \mathbb{R}^n$, $b^{(0)}, b^{(1)} \in \mathbb{R}^m$ and $A^{(0)}, A^{(1)} \in \mathbb{R}^{m\times n}$. The optimal value, the solution set and the feasible set of Problem (1) are denoted as $F^*(\varepsilon)$, $\theta^*(\varepsilon)$ and $\theta(\varepsilon)$, respectively.

* Corresponding author.
E-mail address: K.Avrachenkov@inria.fr (K. Avrachenkov).

http://dx.doi.org/10.1016/j.orl.2016.02.005
0167-6377/

The goal of the perturbed linear programming approach is to construct, if possible, a linear programming problem that does not depend on ε and such that its optimal solutions are *feasible limiting optimal* for (1) in the sense prescribed below by Definition 1. The linear program with this property will be called a *limiting LP*.

Definition 1. A vector $x \in \mathbb{R}^n$ is called *feasible limiting optimal* for the perturbed linear program (1) if $x \in \liminf_{\varepsilon\downarrow 0}\theta(\varepsilon)$ and $\lim_{\varepsilon\downarrow 0}F^*(\varepsilon) = \langle c^{(0)}, x\rangle$.

Let us introduce and discuss a set of assumptions:

Assumption (H_0): *There exists a positive γ_0 and a bounded set $B \subset \mathbb{R}^n$ such that $\theta(\varepsilon) \subset B$ for every $\varepsilon \in (0, \gamma_0]$.*

Assumption (H_0^*): *There exists a positive γ_0 and a bounded set $B \subset \mathbb{R}^n$ such that $\theta^*(\varepsilon) \subset B$ for every $\varepsilon \in (0, \gamma_0]$.*

Assumption (H_1): *The matrix $A^{(0)}$ has rank m.*

Assumption (H_2): *For all ε sufficiently small and positive, the rank of $A^{(0)} + \varepsilon A^{(1)}$ is equal to m.*

Note that Assumption (H_1) implies Assumption (H_2). Also, Assumption (H_0) implies Assumption (H_0^*).

The unperturbed problem is said to satisfy Slater condition if

$$\theta(0) \cap \mathbb{R}^n_{++} \neq \emptyset, \quad \text{where } \mathbb{R}^n_{++} \stackrel{\text{def}}{=} \{x \in \mathbb{R}^n \ : \ x > 0\}. \tag{2}$$

In [18], it has been shown that if Assumptions (H_0) and (H_1) are valid and if the Slater condition (2) is satisfied, then the unperturbed LP is the limiting problem for the perturbed program (1). That is, every optimal solution of the former is limiting optimal for the latter. In [18] it has also been shown that if Assumption (H_1) is not satisfied, the discontinuity of $\theta(\varepsilon)$ at $\varepsilon = 0$ may occur. This is a case of so-called *singular perturbation*. The authors of [18] proposed a limiting LP to deal with the case of singular perturbation. Then, in [3] it has been demonstrated that if the Slater condition is not satisfied for the unperturbed LP, the discontinuity of $\theta(\varepsilon)$ at $\varepsilon = 0$ may occur with Assumptions (H_0) and (H_1) being satisfied. The authors of [3] have constructed a limiting LP for the case when the Slater condition is not satisfied for the unperturbed problem. Below we show that a result similar to that obtained in [3] can be established with the replacement of (H_0) by (H_0^*).

Assume that (H_1) is satisfied and define the set

$$J_0 := \{i \in \{1, \dots, n\} : \exists\, x \in \theta(0) \text{ such that } x_i > 0\}. \tag{3}$$

According to this definition, if $j \notin J_0$, then $x_j = 0$ for every $x \in \theta(0)$. Moreover, if $J_0 \neq \emptyset$, convexity of $\theta(0)$ implies that there exists $\hat{x} \in \theta(0)$ such that $\hat{x}_j > 0$ for every $j \in J_0$. Note that J_0 can be determined by solving n independent linear programming problems $\max_{x\in\theta(0)} x_j$, with $j = 1, \dots, n$.

Consider the following linear program

$$\max\{\langle c^{(0)}, x^0\rangle \ : \ x^0 \in \theta_1\} \stackrel{\text{def}}{=} F_1^*, \tag{4}$$

where

$$\theta_1 \stackrel{\text{def}}{=} \{x^0 \ : \ \exists\, (x^0, x^1) \in \Theta_1\}, \tag{5}$$

and

$$\begin{aligned}\Theta_1 = \{&(x^0, x^1) \in \mathbb{R}^n \times \mathbb{R}^n \ : \ x^0 \in \theta(0),\\ &A^{(0)}x^1 + A^{(1)}x^0 = b^{(1)},\ x_j^1 \geq 0\ \forall j \notin J_0\}.\end{aligned} \tag{6}$$

Note that,

$$\theta_1 \subset \theta(0) \quad \text{and therefore } F_1^* \leq F^*(0).$$

Slater condition (2) is equivalent to having $J_0 = \{1, 2, \dots, n\}$. If this is the case, then $\theta_1 = \theta(0)$ (provided that Assumption (H_1) is satisfied), and the problem (4) is equivalent to the unperturbed problem. If the Slater condition is not satisfied, these two problems are not equivalent.

Following [3], let us introduce the following extended version of the Slater condition.

Definition 2. We shall say that the *extended Slater condition of order 1* (or, for brevity, ES-1) is satisfied if there exists $(\hat{x}^0, \hat{x}^1) \in \Theta_1$ such that $\hat{x}_j^1 > 0$ for every $j \notin J_0$ and $\hat{x}_j^0 > 0$ for every $j \in J_0$.

Theorem 1. *Let Assumptions (H_0^*) and (H_2) be satisfied. Then*

$$\limsup_{\varepsilon\downarrow 0}\theta^*(\varepsilon) \subset \theta_1 \tag{7}$$

and

$$\limsup_{\varepsilon\downarrow 0}F^*(\varepsilon) \leq F_1^*. \tag{8}$$

If, in addition, Assumption (H_1) and the ES-1 condition are satisfied, then

$$\limsup_{\varepsilon\downarrow 0}\theta^*(\varepsilon) \subset \theta_1^*, \tag{9}$$

where θ_1^* is the set of optimal solutions of problem (4), and

$$\lim_{\varepsilon\downarrow 0}F^*(\varepsilon) = F_1^*. \tag{10}$$

Also, any optimal solution x^0 of the problem (4) is limiting optimal for the perturbed problem (1).

Proof. Most steps of the proof are similar to the corresponding steps of the proof of Theorem 2.1 in [3], and we will only indicate the steps that differ from those used in the aforementioned proof.

Let us introduce the following notations. Given a finite set S, denote by $|S|$ the number of elements of S. Let $S_m := \{J \subset \{1, 2, \dots, n\} : |J| = m\}$, so $|S_m| = \binom{n}{m}$. Given a matrix $D \in \mathbb{R}^{m\times n}$ and an index set $J \in S_m$, the matrix $D_J \in \mathbb{R}^{m\times m}$ is constructed by extracting from D the set of m columns indexed by the elements of J. In a similar way, given a vector $x \in \mathbb{R}^n$ and $J \in S_m$, we denote by x_J the vector of $\mathbb{R}^m$ constructed by extracting from x the coordinates $x_j,\ j \in J$ (that is, $x_J \stackrel{\text{def}}{=} \{x_j\},\ j \in J$).

In Lemmas 3.1 and 3.2 of [3] (see also [5,9]) it was established that

$$S_m = \Omega_1 \cup \Omega_2 \quad \text{with } \Omega_1 \cap \Omega_2 = \emptyset,$$

where Ω_1 and Ω_2 are defined by the equations

$$\begin{aligned}\Omega_1 &:= \{J \in S_m : (A^{(0)} + \varepsilon A^{(1)})_J \text{ is nonsingular for } \varepsilon \in (0, \gamma)\}\\ &\neq \emptyset,\end{aligned}$$

$$\Omega_2 := \{J \in S_m : (A^{(0)} + \varepsilon A^{(1)})_J \text{ is singular for all } \varepsilon \in [0, \gamma)\}$$

(here and in what follows, γ stands for a positive number small enough).

Also, it was established that, if

$$x_J(\varepsilon) := [(A^{(0)} + \varepsilon A^{(1)})_J]^{-1}(b_0 + \varepsilon b_1) \tag{11}$$

$(J \in \Omega_1)$ and if

$$\limsup_{\varepsilon\downarrow 0}\|x_J(\varepsilon)\| < \infty, \tag{12}$$

then $x_J(\varepsilon)$ allows the power series expansion

$$x_J(\varepsilon) = \sum_{l=0}^{\infty}\varepsilon^l x_J^l, \quad \forall \varepsilon \in (0, \gamma). \tag{13}$$

Let $\Omega_1^* \subset \Omega_1$ be such that $J \in \Omega_1^*$ if and only if there exists a subsequence $\varepsilon' \to 0$ such that the vector $x(\varepsilon) = \{x_j(\varepsilon)\},\ j = 1, \dots, n$, the non-zero elements of which are equal to the corresponding non-zero elements of $x_J(\varepsilon) = \{x_j(\varepsilon)\},\ j \in J$ (with $x_J(\varepsilon)$ being as in (11)) satisfies the inclusion

$$x(\varepsilon') \in \theta^*(\varepsilon').$$