Morfismos, Vol. 3, No. 2, 1999, pp. 69–79

A SADDLE-POINT THEOREM FOR CONSTRAINED MARKOV CONTROL PROCESSES *

Raquiel R. López-Martínez [1]

Abstract

This paper considers constrained Markov control processes in Borel spaces, with unbounded costs. The criterion to be minimized is a long-run expected average cost, and the constraints are imposed on similar average costs. We first give conditions under which the constrained problem is equivalent to a convex programming problem, and then we present a saddle-point theorem for the Lagrange function associated with the convex program. This theorem gives the existence of an optimal solution to the constrained problem.

1991 Mathematics Subject Classification: 90C40, 93E20, 90C25.
Keywords and phrases: Constrained Markov control processes, convex problem, saddle-point.

1 Introduction

Constrained Markov control processes (MCPs) form an important class of stochastic control problems with applications in many areas; see, for instance, [2, 4-6, 11-14, 17-19], as well as the books [1] and [16] and their extensive bibliographies.

In this paper we study constrained MCPs in *Borel spaces*, with *unbounded costs*. The criterion to be minimized is a long-run expected average cost, and the constraints are imposed on similar average cost functionals. This *constrained problem* (CP) was studied in [9] using

*Research partially support by a CONACyT scholarship.
[1] Graduate Student, Departamento de Matemáticas, CINVESTAV-IPN.

the so-called "direct method" and also infinite-dimensional linear programming. In particular, that paper gives conditions under which CP is solvable and equivalent to an equality-constrained linear program. Here, on the other hand, we shall consider CP as a *convex programming problem* . This approach has been previously used assuming that the state space is compact [13] or countable [18].

In this paper, we begin in §2 by introducing some basic terminology and notation. In §3 we define the concept of stable policy, and state Lemma 3.4, which ensures that we can consider CP as a convex programming problem. Finally, in §4 we study the convex problem and we give a saddle-point theorem for the associated Lagrange function, which gives an optimal solution for CP.

2 Constrained MCPs

The *constrained* Markov control model is of the form

$$(\mathrm{X},\ \mathrm{A},\ \{\mathrm{A}(x)\,|\,x\in\mathrm{X}\},Q,c,\mathbf{d},\mathbf{k}), \tag{2.1}$$

where X and A are the *state space* and the *control space*, respectively. We shall assume that X and A are Borel spaces, endowed with the corresponding Borel σ-algebras $\mathcal{B}(\mathrm{X})$, $\mathcal{B}(\mathrm{A})$. For each $x\in\mathrm{X}$, $\mathrm{A}(x)$ in $\mathcal{B}(\mathrm{A})$ denotes the nonempty set of *feasible controls* or *actions* when the system is in state $x\in X$. We suppose that the set

$$\mathbb{K}:=\{(x,a)\,|\,x\in\mathrm{X},\ a\in\mathrm{A}(x)\} \tag{2.2}$$

of feasible state-action pairs is a *closed* (hence Borel measurable) subset of $\mathrm{X}\times\mathrm{A}$. Moreover, Q stands for the *transition law*, and $c\colon\mathbb{K}\to\mathbb{R}$ is a measurable function that denotes the *cost-per-stage* . Finally, $\mathbf{d}=(d_1,\ldots,d_q):\mathbb{K}\to\mathbb{R}^q$ is a given function and $\mathbf{k}=(k_1,\ldots,k_q)$ is a given vector in $\mathbb{R}^q$, which are used to define the constrained problem (CP) in (2.3) and (2.4), below.

Let Π be set of all (randomized, history-dependent) admisible control policies. If necessary, see [1, 3, 7, 8, 10] for further information on policies. To guarantee that Π is nonempty, we shall assume that the set Φ of all stochastic kernels φ such that $\varphi(\mathrm{A}(x)|\,x)=1$ for all $x\in\mathrm{X}$ is nonempty.

Let c and $\mathbf{d}=(d_1,\ldots,d_q)$ be as in (2.1), and let $\mathbb{P}(\mathrm{X})$ be the set of probabilty measures (p.m.'s) on X. For each control policy $\pi\in\Pi$ and initial distribution $\nu \in \mathbb{P}(X)$, consider the long-run expected average costs

$$J_0(\pi,\nu) := \limsup_{n\to\infty} \frac{1}{n} E_\nu^\pi \left[\sum_{t=0}^{n-1} c(x_t, a_t)\right]$$

and

$$J_i(\pi,\nu) := \limsup_{n\to\infty} \frac{1}{n} E_\nu^\pi \left[\sum_{t=0}^{n-1} d_i(x_t, a_t)\right] \quad \text{for } i = 1,\ldots,q.$$

Furthermore, letting $\mathbf{k} = (k_1,\ldots,k_q)$ be the q-vector in (2.1), define a subset Δ of $\Pi \times \mathbb{P}(X)$ as

$$\Delta := \{(\pi,\nu) | J_0(\pi,\nu) < \infty \ \text{ and } J_i(\pi,\nu) \le k_i \ (i = 1,\ldots,q)\}. \tag{2.3}$$

With this notation, we may then define the **constrained problem** (CP) we are concerned with as follows:

$$\begin{aligned} \mathbf{CP}: \ & \text{Minimize } J_0(\pi,\nu) \\ & \text{subject to: } (\pi,\nu) \in \Delta. \end{aligned} \tag{2.4}$$

If there exists a pair (π^*,ν^*) in Δ such that

$$J_0(\pi^*,\nu^*) = \inf\{J_0(\pi,\nu) \,|\, (\pi,\nu) \in \Delta\} =: \rho^*, \tag{2.5}$$

then (π^*,ν^*) is called a *constrained optimal pair*, and ρ^* is called the *optimal value* of CP.

3 Reduction of CP to stable policies

The following conditions are used to reduce CP to a problem on a set of stable policies.

Assumption 3.1

(a) CP is *consistent*; that is, the set Δ in (2.3) is nonempty.

(b) $c(x,a)$ is nonnegative and *inf-compact*, which means that for each $r \in \mathbb{R}$ the set $\{(x,a) \in \mathbb{K} \,|\, c(x,a) \le r\}$ is compact.

(c) $d_i(x,a)$ is nonnegative and lower semicontinuous (l.s.c), for $i = 1,\ldots,q$.

(d) The transition law Q is *weakly continuous*, that is (denoting by $C_b(\mathrm{S})$ the space of continuous bounded functions on a topological spaces S), Q is such that $\int_{\mathrm{X}} u(y)Q(dy|\cdot\,)$ belongs to $C_b(\mathbb{K})$ for each function u in $C_b(\mathrm{X})$.

Remark 3.2 (a) (See, for instance, pp. 88-89 in [3], or pp. 89 in [10].) If μ is a p.m. on X×A concentrated on $\mathbb{K}$, then there exists $\varphi \in \Phi$ such that μ can be "disintegrated" as

$$\mu(\mathrm{B} \times \mathrm{C}) = \int_B \varphi(\mathrm{C}|x)\widehat{\mu}(dx) \quad \forall\ \mathrm{B} \in \mathcal{B}(\mathrm{X}), \mathrm{C} \in \mathcal{B}(\mathrm{A}), \tag{3.1}$$

where $\widehat{\mu}(\mathrm{B}) := \mu(\mathrm{B} \times \mathrm{A})$ for all B in $\mathcal{B}(\mathrm{X})$ is the *marginal* (or projection) of μ on X. Conversely, for each $\phi \in \Phi$ and $\nu \in \mathbb{P}(\mathrm{X})$, the p.m. μ on X× A defined by

$$\mu(\mathrm{B} \times \mathrm{C}) := \int_B \varphi(\mathrm{C}|x)\nu(dx) \qquad \forall\ \mathrm{B} \in \mathcal{B}(\mathrm{X}), \mathrm{C} \in \mathcal{B}(\mathrm{A}) \tag{3.2}$$

is concentrated on $\mathbb{K}$ and its marginal on X is $\widehat{\mu} = \nu$. The p.m. μ in (3.1) and (3.2) will be written as $\mu = \widehat{\mu}\cdot\varphi$ and $\mu = \nu\cdot\varphi$, respectively.

(b) For each $\varphi \in \Phi$ and $x \in \mathrm{X}$ we write

$$c(x,\varphi) := \int_A c(x,a)\varphi(da|x) \quad \text{and} \quad Q(x,\varphi) := \int_A Q(x,a)\varphi(da|x),$$

and similarly for $d_i(x,\varphi)$.

Definition 3.3 (Stable policies) Let $\mu = \widehat{\mu}\cdot\varphi$ be as in (3.1). Then the p.m. μ (or the randomized stationary policy $\varphi \in \Phi$) is said to be *stable* if

(a) $\langle \mu, c\rangle := \int c(x,a)\mu(d(x,a)) = \int c(x,\varphi)\widehat{\mu}(dx) < \infty$, and

(b) the marginal $\widehat{\mu}$ is an *invariant probablity measure* (i.p.m.) for the transition kernel $Q(\cdot\,|\cdot\,,\varphi)$, that is,

$$\widehat{\mu}(\mathrm{B}) = \int_{\mathrm{X}} Q(\mathrm{B}|x,\varphi)\widehat{\mu}(dx) \qquad \forall\ \mathrm{B} \in \mathcal{B}(\mathrm{X}).$$

We shall denote by $\mathbb{P}(\mathbb{K})$ the family of p.m.'s on X×A concentrated on $\mathbb{K}$, and by $\mathbb{P}_{\mathrm{s}}(\mathbb{K}) \subset \mathbb{P}(\mathbb{K})$ the subset of stable p.m.'s

By the Individual Ergodic Theorem (see, for instance, p. 388 in [20] or Theorem E.11 in [7]), if $\mu = \widehat{\mu}\cdot\varphi$ is stable, then the long-run expected average cost $J_0(\varphi, \widehat{\mu})$ when using the policy $\varphi \in \Phi$ and the initial distribution is $\widehat{\mu}$ is given by

$$J_0(\varphi, \widehat{\mu}) = \lim_{n\to\infty} \frac{1}{n} E_{\widehat{\mu}}^{\varphi} \left[\sum_{t=0}^{n-1} c(x_t, a_t) \right] = \langle \mu, c \rangle.$$

Thus, for $\mu = \widehat{\mu} \cdot \varphi$ in $\mathbb{P}_{\mathrm{s}}(\mathbb{K})$ we have

$$J_0(\varphi, \widehat{\mu}) = \langle \mu, c \rangle = \int_{\mathrm{X}} c(x, \varphi) \widehat{\mu}(dx),$$

and, similarly,

$$J_i(\varphi, \widehat{\mu}) = \langle \mu, d_i \rangle = \int_{\mathrm{X}} d_i(x, \varphi) \widehat{\mu}(dx) \quad \text{for } i = 1, \ldots, q.$$

With this notation we can now state the following key fact.

Lemma 3.4 (Reduction of CP to stable policies) Under Assumption 3.1, for each feasible pair $(\pi, \nu) \in \Delta$ for CP there exists a stable p.m. $\mu = \widehat{\mu} \cdot \varphi$ such that
(a) $(\varphi, \widehat{\mu})$ is in Δ , and
(b) $J_0(\pi, \nu) \geq J_0(\varphi, \widehat{\mu}) = \langle \mu, c \rangle$.
Hence, we can write ρ^* in (2.5) as

$$\rho^* = \inf\{\langle \mu, c \rangle | \mu \in \Delta_{\mathrm{s}}\}, \tag{3.3}$$

where

$$\begin{aligned} \Delta_{\mathrm{s}} &:= \{\mu \in \mathbb{P}_{\mathrm{s}}(\mathbb{K}) | \text{ if } \mu = \widehat{\mu} \cdot \varphi, \text{then } (\varphi, \widehat{\mu}) \in \Delta\} \\ &= \{\mu \in \mathbb{P}_{\mathrm{s}}(\mathbb{K}) | \ \langle \mu, d_i \rangle \leq k_i \ , i = 1, \ldots, q\}. \end{aligned}$$

Proof. See the proof of Lemma 3.5 in [9].

4 The problem CP as a convex programming problem

In this section we see that CP is equivalent to a convex programming problem, which is shown to have an optimal solution. This solution is then used to obtain a constrained optimal pair for CP.

Consider the functions

$$f : \mathbb{P}_{\mathrm{s}}(\mathbb{K}) \to \mathbb{R} \ \text{ and } \ G : \mathbb{P}_{\mathrm{s}}(\mathbb{K}) \to \mathbb{R}^{\mathrm{q}}$$

defined as $f(\mu) := \langle \mu, c \rangle$, and $G(\mu) := (G_1(\mu), \ldots, G_q(\mu))$ with $G_i(\mu) := \langle \mu, d_i \rangle - k_i$ for $i = 1, \ldots, q$. Obviously, f and G are convex functions and $\mathbb{P}_{\mathrm{s}}(\mathbb{K})$ is a convex set. Thus, by Lemma 3.4 we can to represent CP as the convex problem

$$\begin{aligned} &\text{Minimize} \quad f(\mu) \\ &\text{subject to: } \mu \in \mathbb{P}_{\mathrm{s}}(\mathbb{K}) \text{ and } G(\mu) \leq \theta, \end{aligned} \tag{4.1}$$

where θ is the vector zero in $\mathbb{R}^{\mathrm{q}}$, and $G(\mu) \leq \theta$ means that $G_i(\mu) \leq 0$, for all $i = 1, \ldots, q$.

The *Lagrangian* $\mathrm{L} : \mathbb{P}_{\mathrm{s}}(\mathbb{K}) \times \mathbb{R}^{\mathrm{q}}_{+} \to \mathbb{R}$ associated with problem (4.1) is given by

$$\mathrm{L}(\mu, \boldsymbol{\alpha}) := f(\mu) + (G(\mu), \boldsymbol{\alpha}), \tag{4.2}$$

where $\boldsymbol{\alpha} = (\alpha_1, \ldots, \alpha_q)$ is in $\mathbb{R}^{\mathrm{q}}_{+}$, and $(\cdot, \cdot)$ denote the inner product in $\mathbb{R}^{\mathrm{q}}$.

The following saddle-point result gives conditions for problem (4.1) to have a solution.

Theorem 4.1 Suppose that there exists $(\mu^*, \boldsymbol{\alpha}^*) \in \mathbb{P}_{\mathrm{s}}(\mathbb{K}) \times \mathbb{R}^{\mathrm{q}}_{+}$ such that the Lagrangian L has a saddle point at $(\mu^*, \boldsymbol{\alpha}^*)$, i.e.,

$$\mathrm{L}(\mu^*, \boldsymbol{\alpha}) \leq \mathrm{L}(\mu^*, \boldsymbol{\alpha}^*) \leq \mathrm{L}(\mu, \boldsymbol{\alpha}^*), \tag{4.3}$$

for all $(\mu, \boldsymbol{\alpha})$ in $\mathbb{P}_{\mathrm{s}}(\mathbb{K}) \times \mathbb{R}^{\mathrm{q}}_{+}$. Then (a) μ^* solves problem (4.1), and (b) the disintegration $\mu^* = \widehat{\mu^*} \cdot \varphi^*$ of μ^* satisfies that (φ^*, μ^*) is a constrained optimal pair for CP.

Proof. The proof of the part (a) is similar to that of Theorem 2 in [15], p. 221 and, therefore, is omitted. Part (b) follows from (a) and the equivalence of CP and Problem (4.1) ■

In view of Theorem 4.1, to prove that problem (4.1) is solvable it suffices to show the existence of a saddle point for L. To do this, we shall suppose the following.

Assumption 4.2 (Slater condition) There exists $\mu_1 \in \mathbb{P}_{\mathrm{s}}(\mathbb{K})$ such that $G(\mu_1) < \theta$, that is, $G_i(\mu_1) < 0$ for $i = 1, \ldots, q$.

Let us now consider the functions

$$\mathrm{L}_1(\boldsymbol{\alpha}) := \inf_{\mu\in\mathbb{P}_{\mathrm{s}}(\mathbb{K})} \mathrm{L}(\mu,\boldsymbol{\alpha}), \tag{4.4}$$

$$\mathrm{L}_2(\mu) := \sup_{\boldsymbol{\alpha}\geq\theta} \mathrm{L}(\mu,\boldsymbol{\alpha}), \tag{4.5}$$

Remark 4.3 Note that for all $\boldsymbol{\alpha} \in \mathbb{R}^{\mathrm{q}}_{+}$

$$\mathrm{L}_1(\boldsymbol{\alpha}) \leq \inf_{\mu\in\Delta_{\mathrm{s}}} \mathrm{L}(\mu,\boldsymbol{\alpha}) \leq \inf_{\mu\in\Delta_{\mathrm{s}}} \langle\mu, c\rangle = \rho^*,$$

that is, $\mathrm{L}_1(\boldsymbol{\alpha}) \leq \rho^*$ for all $\boldsymbol{\alpha} \in \mathbb{R}^{\mathrm{q}}_{+}$, and, similarly, $\rho^* \leq \mathrm{L}_2(\mu)$ for all $\mu \in \Delta_{\mathrm{s}}$. It is also clear that

$$\sup_{\boldsymbol{\alpha}\geq\theta} \mathrm{L}_1(\boldsymbol{\alpha}) \leq \inf_{\mu\in\mathbb{P}_{\mathrm{s}}(\mathbb{K})} \mathrm{L}_2(\mu).$$

Theorem 4.4 Under Assumptions 3.1 and 4.2, there exists a saddle point $(\mu^*, \boldsymbol{\alpha}^*)$ for the Lagrangian L.

Before proving Theorem 4.4, let us first prove the following.

Lemma 4.5 Under Assumption 4.2 , there exists $\boldsymbol{\alpha}^*$ in $\mathbb{R}^{\mathrm{q}}_{+}$ such that

$$\mathrm{L}_1(\boldsymbol{\alpha}^*) = \sup_{\boldsymbol{\alpha}\geq\theta} \mathrm{L}_1(\boldsymbol{\alpha}) = \rho^*,$$

with ρ^* as in (2.5) or (3.3).

Proof. In the space $\mathbb{R} \times \mathbb{R}^{\mathrm{q}}$ define the sets

$$\begin{aligned}
\mathrm{B}_1 &:= \{(x,\boldsymbol{\alpha}) |\ x \geq f(\mu), \boldsymbol{\alpha} \geq G(\mu) \text{ for some } \mu \in \mathbb{P}_{\mathrm{s}}(\mathbb{K})\ \} \\
\mathrm{B}_2 &:= \{(x,\boldsymbol{\alpha}) |\ x \leq \rho^*, \boldsymbol{\alpha} \leq \theta\ \}.
\end{aligned}$$

The set B_2 is obviously convex, and so is B_1 because f and G are convex. By (3.3), B_1 contains no interior points of B_2. On other hand, since $G(\mu_1)$ is an interior point of $\mathbb{R}^{\mathrm{q}}_{-}$ (Slater condition), the set B_2 contains an interior point. Thus, by the Separating Hyperplane Theorem (see, for example [15], p. 133, Theorem 3), there is a vector $(x^*, \boldsymbol{\alpha}^*) \in \mathbb{R}\times\mathbb{R}^{\mathrm{q}}$ such that

$$x^* x_1 + (\boldsymbol{\alpha}_1, \boldsymbol{\alpha}^*) \geq x^* x_2 + (\boldsymbol{\alpha}_2, \boldsymbol{\alpha}^*)$$

for all $(x_1, \boldsymbol{\alpha}_1) \in \mathrm{B}_1$ and all $(x_2, \boldsymbol{\alpha}_2) \in \mathrm{B}_2$. From the nature of B_2 it follows that $x^* \geq 0$ and $\boldsymbol{\alpha}^* \geq \theta$. We next show that in fact $x^* > 0$. Indeed, as the vector $(\rho*, \theta)$ is in B_2, we have

$$x^* x + (\boldsymbol{\alpha}, \boldsymbol{\alpha}^*) \geq x^* \rho^* \tag{4.6}$$

for all $(x, \boldsymbol{\alpha}) \in \mathrm{B}_1$. Thus, if $x^* = 0$, then taking $\boldsymbol{\alpha} = G(\mu_1)$ we obtain $(G(\mu_1), \boldsymbol{\alpha}^*) \geq 0$. Therefore, $G_i(\mu_1) \geq 0$ for some $i = 1, \ldots, q$, which contradicts Assumption 4.2. It follows that $x^* > 0$ and, without loss of generality, we may assume $x^* = 1$.

Now, since the point (ρ^*, θ) is in the closure of both B_1 and B_2, we have (with $x^* = 1$ in (4.6))

$$\begin{aligned} \rho^* &= \inf_{(x,\boldsymbol{\alpha}) \in \mathrm{B}_1} [x + (\boldsymbol{\alpha}, \boldsymbol{\alpha}^*)] \leq \inf_{\mu \in \mathbb{P}_{\mathrm{s}}(\mathbb{K})} [f(\mu) + (G(\mu), \boldsymbol{\alpha}^*)] \\ &= \mathrm{L}_1(\boldsymbol{\alpha}^*) \leq \inf_{\mu \in \Delta_{\mathrm{s}}} f(u) = \rho^*; \end{aligned}$$

see Remark 4.3. Hence the lemma is proved. ■

The following Lemma 4.6 is a minimax result.

Lemma 4.6 Under Assumptions 3.1 and 4.2, we have

$$\sup_{\boldsymbol{\alpha} \geq \theta} \mathrm{L}_1(\boldsymbol{\alpha}) = \inf_{\mu \in \mathbb{P}_{\mathrm{s}}(\mathbb{K})} \mathrm{L}_2(\mu) = \rho^*. \tag{4.7}$$

Proof. Since $(G(\mu), \boldsymbol{\alpha}) \leq \theta$ for all $\mu \in \Delta_{\mathrm{s}}$, we see that

$$\mathrm{L}_2(\mu) = \sup_{\boldsymbol{\alpha} \geq \theta} \mathrm{L}(\mu, \boldsymbol{\alpha}) = \langle \mu, c \rangle \quad \text{for all } \mu \in \Delta_{\mathrm{s}}.$$

Hence, by (3.3),

$$\inf_{\mu \in \Delta_{\mathrm{s}}} \mathrm{L}_2(\mu) = \rho^*.$$

Therefore,

$$\inf_{\mu \in \mathbb{P}_{\mathrm{s}}(\mathbb{K})} \mathrm{L}_2(\mu) \leq \rho^*,$$

so that, by Remark 4.3 and Lemma 4.5, the equality (4.7) holds. ■

Lemma 4.7 Under Assumptions 3.1 and 4.2, there exists a p.m. μ^* in $\mathbb{P}_{\mathrm{s}}(\mathbb{K})$ such that

$$\mathrm{L}_2(\mu^*) = \inf_{\mu \in \mathbb{P}_{\mathrm{s}}(\mathbb{K})} \mathrm{L}_2(\mu) = \rho^*.$$

Proof. If μ is in $\mathbb{P}_{\mathrm{s}}(\mathbb{K})$ but not in Δ_{s}, then there exists i_0 in $\{1, \ldots, q\}$ such that $G_{i_0}(\mu) > 0$, which implies that $\mathrm{L}_2(\mu) = +\infty$. Therefore,

$$\inf_{\mu \in \mathbb{P}_{\mathrm{s}}(\mathbb{K})} \mathrm{L}_2(\mu) = \inf_{\mu \in \Delta_{\mathrm{s}}} \mathrm{L}_2(\mu) = \inf_{\mu \in \Delta_{\mathrm{s}}} \langle \mu, c \rangle.$$

On the other hand, for all $\mu \in \Delta_{\mathrm{s}}$ and $\boldsymbol{\alpha} \geq \theta$, we have $(G(\mu), \boldsymbol{\alpha}) \leq 0$, and so it follows that

$$\mathrm{L}_2(\mu) = \sup_{\boldsymbol{\alpha} \geq \theta} \mathrm{L}(\mu, \boldsymbol{\alpha}) = \langle \mu, c \rangle \ \forall \ \mu \in \Delta_{\mathrm{s}}.$$

Therefore, from the latter equality and Corollary 3.6 in [9], the desired conclusion follows. ■

Proof of Theorem 4.4. From lemma 4.6 we have that

$$\mathrm{L}(\mu^*, \boldsymbol{\alpha}^*) = \rho^*,$$

where ρ^* is the optimal value of CP —see (2.5) or (3.3). Now, by the latter equality, together with Lemma 4.5, (4.7) and the definition of L_1 and L_2, it follows that

$$\mathrm{L}(\mu^*, \boldsymbol{\alpha}^*) = \mathrm{L}_1(\boldsymbol{\alpha}^*) \leq \mathrm{L}(\mu, \boldsymbol{\alpha}^*) \text{ for all } \mu \in \mathbb{P}_{\mathrm{s}}(\mathbb{K}),$$

and, similarly,

$$\mathrm{L}(\mu^*, \boldsymbol{\alpha}^*) = \mathrm{L}_2(\mu^*) \geq \mathrm{L}(\mu^*, \boldsymbol{\alpha}) \text{ for all } \boldsymbol{\alpha} \geq \theta.$$

Therefore, the pair $(\mu^*, \boldsymbol{\alpha}^*)$ is a saddle point. ■

To summarize, Theorem 4.4 gives the existence of a saddle point $(\mu^*, \boldsymbol{\alpha}^*)$ for L, which, by Theorem 4.1 yields a constrained optimal pair $(\varphi^*, \widehat{\mu^*})$ for CP. It turns out the converse is also true, as we next show.

Proposition 4.8 If $\mu^* = \widehat{\mu^*} \cdot \varphi^* \in \Delta_{\mathrm{s}}$ is such that $(\varphi^*, \widehat{\mu^*})$ is a constrained optimal pair for CP and Assumption 4.2 holds, then the Lagrangian L has a saddle point.

Proof. Let $\boldsymbol{\alpha}^*$ be as in Lemma 4.5. By the definition of L_1, we have that

$$\mathrm{L}(\mu^*, \boldsymbol{\alpha}^*) \leq \mathrm{L}(\mu, \boldsymbol{\alpha}^*) \text{ for all } \mu \in \mathbb{P}_{\mathrm{s}}(\mathbb{K}),$$

which gives the second inequality in (4.3). On other hand, $(G(\mu^*), \boldsymbol{\alpha}^*) = 0$ because, as $G(\mu^*) \leq \theta$, we have

$$\rho^* \leq f(\mu^*) + (G(\mu^*), \boldsymbol{\alpha}^*) \leq f(\mu^*) = \rho^*$$

and hence $(G(\mu^*), \boldsymbol{\alpha}^*) = 0$. Therefore,

$$\mathrm{L}(\mu^*, \boldsymbol{\alpha}) - \mathrm{L}(\mu^*, \boldsymbol{\alpha}^*) = (G(\mu^*), \boldsymbol{\alpha}) - (G(\mu^*), \boldsymbol{\alpha}^*) = (G(\mu^*), \boldsymbol{\alpha}) \leq 0,$$

and the first inequality in (4.3) follows. ■

Acknowledgement

The author wishes to thank Dr. Onésimo Hernández-Lerma for his very valuable comments and suggestions.

Raquiel R. López-Martínez
Departmento de Matemáticas,
CINVESTAV-IPN,
A.P. 14-740,
México D.F. 07000,
México,
raquiel@math.cinvestav.mx

References

[1] E.Altman, *Constrained Markov Decision Processes*, Chapman & Hall/CRC, Boca Raton, FL, 1999.

[2] V.S. Borkar, *Ergodic control of Markov chains with constraints—the general case,* SIAM J. Control Optim., 32(1994), pp. 176-186.

[3] E.B. Dynkin and A.A. Yushkevich, *Controlled Markov Processes,* Springer-Verlag, Berlin, 1979.

[4] E. Feinberg and A. Shwartz, *Constrained discounted dynamic programming,* Math. Oper. Res., 21(1996), pp. 922-945.

[5] E. Feinberg and A. Shwartz,*Constrained dynamic programming with two discount factors: applications and an algorithm,* IEEE Trans. Autom. Control, 44(1999), pp. 628-631.

[6] K. Golabi, R.B. Kulkarni, and G.B. Way, *A statewide pavement management system,* Interfaces, 12(1982), pp. 5-21.

[7] O. Hernández-Lerma and J.B. Lasserre, *Discrete-Time Markov Control Processes: Basic Optimality Criteria,* Springer-Verlag, New York, 1996.

[8] O. Hernández-Lerma and J.B. Lasserre, *Further Topics on Discrete-Time Markov Control Processes,* Springer-Verlag, New York, 1999.

[9] O. Hernández-Lerma, J. González-Hernández and R.R. López-Martínez, *Constrained average cost Markov control processes in Borel spaces*, Reporte interno # 262, CINVESTAV-IPN, México, 1999. (Submitted)

[10] K. Hinderer, *Foundations of Non-stationary Dynamic Programming with Discrete-Time Parameter*, Lecture Notes Oper. Res. Math. Syst. 33, Springer-Verlag, Berlin, 1970.

[11] A. Hordijk and F. Spieksma, *Constrained admission control to a queueing system,* Adv. Appl. Prob., 21(1989), pp. 409-431.

[12] Y. Huang and M. Kurano, *The LP approach in average rewards MDPs with multiple cost constraints: The countable state case,* J. Inform. Optim. Sci., 18(1997), pp. 33-47.

[13] M. Kurano, J.-I. Nakagami, and Y. Huang, *Constrained Markov decision processes with compact state and action spaces: the average case,* preprint, Faculty of Education, Chiba University, Japan, 1999.

[14] A. Lazar, *Optimal flow control of a class of queueing networks in equilibrium,* IEEE Trans. Autom. Control, 28(1983), pp. 1001-1007.

[15] D. G. Luenberger, *Optimization by Vector Space Methods,* Wiley, New York, 1969.

[16] A.B. Piunovskiy, *Optimal Control of Random Sequences in Problems with Constraints,* Kluwer, Boston, 1997.

[17] K. Ross and R. Varadarajan, *Multichain Markov decision processes with a sample path constraint: a decomposition approach,* Math. Oper. Res., 16(1991), pp. 195-207.

[18] L.I. Sennott, *Constrained average cost Markov decision chains,* Prob. Eng. Inform. Sci., 7(1993), pp. 69-83.

[19] F. Vakil and A.A. Lazar, *Flow control protocols for integrated networks with partially observed voice traffic,* IEEE Trans. Autom. Control, 32(1987), pp. 2-14.

[20] K. Yosida, *Functional Analysis,* 5th edition, Springer-Verlag, Berlin, 1978.