

6. **E. Fernández-Gaucherand**, “A Note on the Ross-Taylor Theorem,” *Applied Mathematics & Computation* **64** (1994) 207-212.

A Note on the Ross-Taylor Theorem

Emmanuel Fernández-Gaucherand

*Systems and Industrial Engineering Department
The University of Arizona
Tucson, Arizona 85721*

Transmitted by L. Duckstein

ABSTRACT

In this note the conditions used in proving a result due to S. M. Ross and H. M. Taylor are examined. These results pertain to the existence of bounded solutions to the average cost optimality equation for controlled Markov processes with an average cost criterion. In particular, we show how the use of commonly found convexity (concavity) properties of value functions can be employed to verify a seemingly rather restrictive equicontinuity condition. In addition, we remark that several results in the literature can be viewed as special cases of the result by Ross and Taylor, contrary to claims otherwise.

1. INTRODUCTION

Consider a controlled Markov process (CMP) described by the quadruplet $(\mathbf{X}, \mathbf{U}, Q, c)$, where \mathbf{X} is the state space, a (Borel) subset of a complete and separable metric space, \mathbf{U} is a finite set of actions (or decisions), Q is a stochastic kernel describing the distribution of the next state X_{t+1} , given the current state-action pair (X_t, U_t) , $c: \mathbf{X} \times \mathbf{U} \rightarrow \mathbb{R}$ is the *bounded* one-stage cost function. A (stationary) policy is a rule $\pi: \mathbf{X} \rightarrow \mathbf{U}$ for making decisions, such that $U_t = \pi(X_t)$. We refer to [1-4] for more details on the description of the model.

Given a policy π and an initial state x_0 , two criteria commonly used to measure the performance of the system are the *discounted cost* (DC),

$$J_{\beta}^{\pi}(x_0) := \lim_{N \rightarrow \infty} E_{x_0}^{\pi} \left\{ \sum_{t=0}^N \beta^t c(X_t, U_t) \right\}, \quad (1)$$

where $0 < \beta < 1$ is the discount factor, and the *average cost* (AC),

$$J^\pi(x_0) := \limsup_{N \rightarrow \infty} \frac{1}{N+1} E_{x_0}^\pi \left\{ \sum_{t=0}^N c(X_t, U_t) \right\}, \quad (2)$$

where $E_{x_0}^\pi$ denotes the expectation operator with respect to the (canonical) probability measure induced by π and x_0 [1–4]. The optimal values for (1) and (2) above, i.e., their infimums over all policies, are denoted by $J_{\beta^*}^*(x_0)$ and $J^*(x_0)$, respectively. An optimal policy for a particular criterion is a policy which attains the corresponding infimum value. The stochastic control problem is that of characterizing optimal values and policies.

For the DC case many results are available, and the problem can be considered as well understood [2, 3, 5, 6]. On the other hand, for the AC case there are still many open issues [1]. In this respect, of particular interest to us are conditions under which there exist *bounded* solutions (ρ^*, h) to the average cost optimality equation (ACOE):

$$\rho^* + h(x) = \min_{u \in U} \left\{ c(x, u) + \int_{\mathbf{X}} h(y) Q(dy | x, u) \right\}, \quad (3)$$

where $\rho^* \in \mathbb{R}$, and $h: \mathbf{X} \rightarrow \mathbb{R}$ is a bounded function. Building upon results by Taylor [7], Ross showed in [8] that if a bounded solution exists for the ACOE, then $\rho^* = J^*(x_0)$ for all x_0 , and minimizing actions in the ACOE determine an optimal policy. Hence, conditions under which such solutions can be obtained are of much interest; see [9] for some *necessary* conditions.

2. THE ROSS-TAYLOR THEOREM

Ross and Taylor developed an approach based on the idea of studying the AC problem as a limit of the DC problem, as $\beta \uparrow 1$. The idea is to obtain, along some sequence $\beta_n \uparrow 1$,

$$(1 - \beta_n) J_{\beta_n}^*(\bar{x}) \xrightarrow{\beta_n \uparrow 1} \rho^*, \quad (4)$$

$$h_{\beta_n}(x) \xrightarrow{\beta_n \uparrow 1} h(x), \quad \forall x \in \mathbf{X}, \quad (5)$$

for some arbitrarily selected *reference state* $\bar{x} \in \mathbf{X}$, and where the *differential* discounted cost function $h_{\beta}(\cdot)$ is given by

$$h_{\beta}(x) := J_{\beta}^*(x) - J_{\beta}^*(\bar{x}).$$

Using the Arzela–Ascoli theorem [10] to obtain a sequence $\beta_n \uparrow 1$ as required above, the following was proved, among other things, in [8].

THEOREM 1 [Ross–Taylor]. *If*

(C1) $\{h_{\beta}\}$ *is uniformly bounded, and*

(C2) $\{h_{\beta}\}$ *is an equicontinuous family of functions, then there exists a bounded solution* (ρ^*, h) *to the ACOE, with* $h(\cdot)$ *a continuous function.*

3. SATISFYING THE CONDITIONS

We now turn our attention to conditions C1, C2 in Theorem 1. The first condition has been well studied in the literature, and there are many interesting applications in which it is known to be satisfied [1, 4–6, 11–13]. When \mathbf{X} is a countable set and C1 holds, then C2 is immediately satisfied. However, many applications in e.g., inventory control [5], equipment replacement [5, 7], and systems with incomplete state information [5, 11, 13] require the consideration of a state space \mathbf{X} which is not a countable set. For the latter situation, very few and rather specific cases have been reported in which C2 is satisfied, c.f., [13]. We observe that the assumption that \mathbf{X} is a (Borel) subset of a complete and separable metric space is needed for the underlying probabilistic structure to be well defined [1–4]. Furthermore, separability is crucial for the proof of Theorem 1, in order for the Arzela–Ascoli theorem to be applicable [10].

The paucity of cases in which C2 has been shown to be satisfied can be traced, to a large extent, to the selection of inadequate metrics for \mathbf{X} since equicontinuity is a topological property determined by the metric used. We make the following observation.

REMARK 1. To verify C2, any metric for \mathbf{X} inducing completeness, separability, and equicontinuity of $\{h_{\beta}\}$ can be employed.

As an illustration of the usefulness of the above remark, we mention that the results in [12] and [14] can be easily shown to be subsumed by those in [8], contrary to claims otherwise. This is done using a metric introduced by Platzman [12], easily verified to be complete and separable in finite dimensional simplices. Also subsumed are cases in which C2 was shown to hold in [13, 15]. In the above references $\{h_{\beta}\}$ is a family of concave functions. In these and other similar cases, by using convex analytic results, C2 may be easily verified as follows.

THEOREM 2. *Suppose that* $(\mathbf{X}, \|\cdot\|)$ *is a complete, separable normed vector space. If* C1 *holds and* $\{h_{\beta}\}$ *is a family of convex (or concave) functions, then* C2 *is satisfied.*

PROOF. The results follow via the *local Lipschitz property* of convex functions [16, 17] as follows. Let $x_0 \in \mathbf{X}$ and $\epsilon > 0$ be given, and define

$$\bar{B}(x_0, 2\epsilon) := \{y' \in \mathbf{X} \mid \|x_0 - y'\| \leq 2\epsilon\}.$$

Let $x, y \in B(x_0, \epsilon) := \{y' \in \mathbf{X} \mid \|x_0 - y'\| < \epsilon\}$, with $x \neq y$, and define

$$z := x + \frac{\epsilon}{\|x - y\|}(x - y).$$

Thus $\|x_0 - z\| \leq \|x_0 - x\| + \epsilon < 2\epsilon$, and then

$$\begin{aligned} \|x - y\|z &= \|x - y\|x + \epsilon(x - y) \\ \implies x &= \frac{\|x - y\|}{\|x - y\| + \epsilon}z + \frac{\epsilon}{\|x - y\| + \epsilon}y; \end{aligned}$$

hence x is a convex combination of $z, y \in B(x_0, 2\epsilon)$. By C1, there exists a constant M such that

$$|h_\beta(y')| \leq M, \quad \forall y' \in \mathbf{X}.$$

Now, by the convexity of $h_\beta(\cdot)$,

$$h_\beta(x) \leq \frac{\|x - y\|}{\|x - y\| + \epsilon}h_\beta(z) + \frac{\epsilon}{\|x - y\| + \epsilon}h_\beta(y),$$

and therefore

$$\begin{aligned} h_\beta(x) - h_\beta(y) &\leq \frac{\|x - y\|}{\|x - y\| + \epsilon} [h_\beta(z) - h_\beta(y)] \\ &\leq \frac{2M\|x - y\|}{\|x - y\| + \epsilon} \\ &\leq \left\lceil \frac{2M}{\epsilon} \right\rceil \|x - y\|. \end{aligned}$$

Interchanging x and y , and since x_0 was arbitrary, we then conclude that $\{h_\beta(\cdot)\}$ is locally equi-Lipschitzian, and in particular (locally) equicontinuous. ■

REMARK 2. The proof above would also apply if \mathbf{X} is a convex open subset of a complete and separable normed vector space. In fact, \mathbf{X} could

be a countable union of such open subsets, and then a diagonalization argument can be employed to extract one subsequence along which (4), (5) can be taken. In addition, only a local analysis is needed in the proof of Theorem 2, thus C1 could be relaxed to some sort of local boundedness condition; see [18].

4. CONCLUSIONS

The observation made in Remark 1 and the result in Theorem 2 can be very useful in verifying condition C2 in Theorem 1. For example, in [12, 14, 18] the results in [8] where deemed as not applicable, and instead the convexity of $\{h_\beta\}$ was used to invoke [17, Theorem 10.9]. However, the latter result is simply a version of the Arzela-Ascoli theorem, and equicontinuity is obtained in a similar manner as in Theorem 2; see [17, Theorems 10.4, 10.6]. Furthermore, the result in Theorem 2 finds applications in other contexts also, e.g., see [19]. General conditions for CMP to exhibit the needed convexity property of value functions can be found in [20]. Examples of CMP exhibiting this condition are problems with imperfect state information [5], some gambling models [6], linear systems with quadratic cost [5], some inventory control models [5] and problems in Bayesian sequential analysis [6].

ACKNOWLEDGMENT

The author thanks his colleague Professor F. Szidarovsky for very useful comments on an earlier version of this paper. This research was partially supported by The Engineering Foundation under Grant RI-A-93-10, in part by the National Science Foundation under Grant NSF-INT 9201430, and in part by a grant from the AT&T Foundation.

REFERENCES

- 1 A. Arapostathis, V. S. Borkar, E. Fernández-Gaucherand, M. K. Ghosh and S. I. Marcus, Discrete-time controlled Markov processes with an average cost criterion: A survey, *SIAM J. Control Optim.* 31:282-344 (1993).
- 2 D. P. Bertsekas and S. E. Shreve, *Stochastic Optimal Control: The Discrete Time Case*, Academic Press, New York, 1978.
- 3 E. B. Dynkin and A. A. Yushkevich, *Controlled Markov Processes*, Springer-Verlag, New York, 1979.
- 4 O. Hernández-Lerma, *Adaptive Markov Control Processes*, Springer-Verlag, New York, 1989.

- 5 D. P. Bertsekas, *Dynamic Programming: Deterministic and Stochastic Models*, Prentice-Hall, Englewood Cliffs, NJ, 1987.
- 6 S. M. Ross, *Introduction to Stochastic Dynamic Programming*, Academic Press, New York, 1983.
- 7 H. M. Taylor, Markovian sequential replacement processes, *The Annals of Math. Statist.* 38:1677-1694 (1965).
- 8 S. M. Ross, Arbitrary state Markovian decision processes, *The Annals of Math. Statist.* 39:2118-2122 (1968).
- 9 E. Fernández-Gaucherand, A. Arapostathis and S. I. Marcus, Remarks on the existence of solutions to the average cost optimality equation in Markov decision processes, *Systems and Control Letters* 15:425-432 (1990).
- 10 H. L. Royden, *Real Analysis*, 2nd ed., Macmillan, New York, 1968.
- 11 E. Fernández-Gaucherand, A. Arapostathis, and S. I. Marcus, On the average cost optimality equation and the structure of optimal policies for partially observable Markov decision processes, *Ann. Oper. Res.* 29:439-470 (1991).
- 12 L. K. Platzman, Optimal infinite-horizon undiscounted control of finite probabilistic systems, *SIAM J. Control Optim.* 18:362-380 (1980).
- 13 S. M. Ross, Quality control under Markovian deterioration, *Management Sci.* 17:587-596 (1971).
- 14 M. Ohnishi, H. Mine and H. Kawai, An optimal inspection and replacement policy under incomplete state information: Average cost criterion, in *Stochastic Models in Reliability Theory* (S. Osaki and Y. Hatoyama, Eds.), Lecture Notes in Economics and Mathematical Systems 235, Springer-Verlag, Berlin, 1984, pp. 187-197.
- 15 C. C. White, A Markov quality control process subject to partial observation, *Management Sci.* 23:843-852 (1977).
- 16 R. R. Phelps, *Convex Functions, Monotone Operators, and Differentiability*, Lecture Notes in Mathematics 1364, Springer-Verlag, New York, 1989.
- 17 R. T. Rockafellar, *Convex Analysis*, Princeton University Press, 1972.
- 18 R. Hartley, Dynamic Programming and an undiscounted, infinite horizon, convex stochastic control problem, in *Recent Developments in Markov Decision Processes* (R. Hartley, L. C. Thomas, and D. J. White, Eds), Academic Press, London, 1980, pp. 277-300.
- 19 E. Fernández-Gaucherand, *Controlled Markov Processes on the Infinite Planning Horizon: Optimal and Adaptive Control*, Ph.D. Thesis, The University of Texas at Austin, August 1991.
- 20 K. F. Hinderer, On the structure of solutions of stochastic dynamic programs, in *Proc. of the 7th Conference on Probability Theory*, Brasov, Romania, 1984, pp. 173-182.