European Journal of Operational Research 241 (2015) 391–401

Contents lists available at ScienceDirect

European Journal of Operational Research

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

Stochastics and Statistics

Partially Observable Markov Decision Processes incorporating epistemic uncertainties

R. Faddoul[a], W. Raphael[a,*], A.-H. Soubra[b], A. Chateauneuf[c]

[a] Civil and Environmental Engineering Department ESIB, Saint-Joseph University, Mar Roukos, PoB 11-154, Riad El Solh, Beirut, Lebanon
[b] Institut de Recherche en Génie Civil et Mécanique, BP 152, 44603 St-Nazaire Cedex, France
[c] LaMI, Polytech'Clermont-Ferrand, BP 206, 63174 Aubière Cedex, France

ARTICLE INFO

Article history:
Received 10 April 2013
Accepted 19 August 2014
Available online 22 September 2014

Keywords:
Uncertainty modeling
Markov process
Dynamic programming
Epistemic uncertainty
Decision analysis
Maintenance

ABSTRACT

The use of Markov Decision Processes for Inspection Maintenance and Rehabilitation of civil engineering structures relies on the use of several transition matrices related to the stochastic degradation process, maintenance actions and imperfect inspections. Point estimators for these matrices are usually used and they are evaluated using statistical inference methods and/or expert evaluation methods. Thus, considerable epistemic uncertainty often veils the true values of these matrices. Our contribution through this paper is threefold. First, we present a methodology for incorporating epistemic uncertainties in dynamic programming algorithms used to solve finite horizon Markov Decision Processes (which may be partially observable). Second, we propose a methodology based on the use of Dirichlet distributions which answers, in our sense, much of the controversy found in the literature about estimating Markov transition matrices. Third, we show how the complexity resulting from the use of Monte-Carlo simulations for the transition matrices can be greatly overcome in the framework of dynamic programming. The proposed model is applied to concrete bridge under degradation, in order to provide the optimal strategy for inspection and maintenance. The influence of epistemic uncertainties on the optimal solution is underlined through sensitivity analysis regarding the input data.

1. Introduction

Most countries worldwide face the problem of ageing civil engineering infrastructures. This is especially the case in the developed countries where investment focus is increasingly shifting from extending the infrastructure to the inspection, maintenance and rehabilitation (IM&R) of the existing assets. For example, government statistics in the United States show that the proportion of public non-capital spending for infrastructure increased from 39 percent in 1960 to 57 percent in 1994 (CBO, 1999). Although mathematical optimization and operational research methods started to be gradually applied for maintenance optimization since the early 1940s, it is until the last two decades that maintenance optimization became a topic of highest priority for civil engineering infrastructure managers. Several researches emphasized on the usefulness of the characterization of the uncertainties (pertaining to reliability analysis and to risk-based decision problems in engineering) as epistemic or aleatory (Der Kiureghian & Ditlevsen, 2009; Hofer, 1996). The decision process regarding the optimal maintenance policy is often very sensitive to the input statistical parameters (e.g. mean values, standard deviations, Markov transition matrices) of the probabilistic models (Avrachenkov, Filar, & Haviv, 2001; Cranahan, 1988). However, these parameters are estimated by classical inference methods where the available data is often very limited (DeStefano & Grivas, 1998; Madanat, Mishalani, & Wan Ibrahim, 1995; Mishalani & Madanat, 2002). This lack of information leads to second-order uncertainties due to the inevitable difference between the observed sample from which inference is made and the real population. This second-order statistical uncertainty, shrouding the real values of the parameters, could be usually quantified by several means, such as confidence intervals, Bayesian probability distributions and fuzzy sets (Corotis, 2009). Keeping in mind that exact and tractable modeling of reality is a far-fetch if not an impossible goal to achieve, it remains that a desired attribute of a mathematical model is to be able to use, as best as possible, the available information. In fact, a mathematical optimization model using only a portion of the relevant information will produce biased optimal solutions. Parametric uncertainty is one type of information that is seldom taken into account by most of the existing models for IM&R optimization. Therefore, point estimators for the transition matrices in Markov Decision Processes (MDP) are simply used instead of exploiting the additional information that can be available. If, for example, a probability

* Corresponding author. Tel.: +961 3 731450.
E-mail addresses: faddoul_rafic@yahoo.com (R. Faddoul), wassim.raphael@fi.usj.edu.lb, wassim012@hotmail.com (W. Raphael), Abed.Soubra@univ-nantes.fr (A.-H. Soubra), alaa.chateauneuf@cust.univ-bpclermont.fr (A. Chateauneuf).

http://dx.doi.org/10.1016/j.ejor.2014.08.032
0377-2217/

density function is available for the elements of MDP transition matrices, then the use of the whole information contained in such a distribution will lead to more realistic life cycle cost estimation and therefore to better IM&R decisions. It will also be shown that the expected costs usually increase with full uncertainty consideration. Hence, neglecting the epistemic uncertainty leads in general to misleading low life cycle costs.

In the next two sections, a literature review on the epistemic uncertainties in the transition matrices and on the dynamic programming of finite horizon MDPs or POMDPs (Partially Observable Markov Decision Processes) are presented. Next, we describe the parameters uncertainties in MDPs. Then, we propose a methodology for incorporating the epistemic uncertainty in MDPs. The paper ends with the presentation of a numerical example that considers IM&R optimization management of a highway concrete bridge deck.

2. Epistemic uncertainties in the transition matrices

Several studies have been suggested in the literature since the early 1950s in order to take into account the effect of uncertainties in the transition matrices. Bellman (1961) proposed an "adaptive control" of Markov chains by recognizing that the problem can be transformed into an equivalent dynamic programming problem with a completely known transition law. The state space of the equivalent problem is the Cartesian product of the original state space by the set of all probability distributions on the parameter set. This approach has the advantage of being adaptive but it suffers, as pointed out by Bellman himself, from the "dimensionality curse"; i.e. the problem becomes quickly intractable as the state space grows exponentially with higher dimensions. A customized version of this approach in the field of IM&R of civil engineering structures was proposed by Durango and Madanat (2008). One can say that most of the research tackling the issue of parametric uncertainty in sequential decision optimization problems focuses on investigating robust optimization techniques (Givan, Leach, & Dean, 2000; Iyengar, 2005; Nilim & El Ghaoui, 2005). In the same framework, ambiguity is investigated in the context of infinite horizon MDP with finite state and action spaces where it is modeled by constraining the transition probability matrix to lie in a pre-specified polytope (Bagnell, Ng, & Schneider, 2001; Satia & Lave, 1973; White & Eldeib, 1994). Such an approach (robust optimization) considers the optimization problem as a game between the decision maker and the nature which is supposed to be malevolent. In other terms, the worst case scenario is assumed, leading to policies which are deemed too conservative by several authors. For this reason, several studies attempted to avoid this drawback. Kuhn and Madanat (2006) used the Hurwicz criterion (Revelle, Whitlatch, & Wright, 1997) which allows the decision maker to set his or her own 'optimism level', as a number between 0 and 1. This number is used to balance the decisions by interpolating between the best case and the worst case scenarios. Delage and Mannor (2009) proposed a chance constrained MDP with a set of percentile criteria that represent the trade-off between the optimistic and the pessimistic views. However, as for the models of Kuhn and Madanat, a so-called risk of the policy level must be chosen at the discretion of the decision maker. Moreover, these approaches do not use all of available information about the epistemic uncertainty.

The authors of the present paper believe that for problems which do not include hard constraints (i.e. constraints which must be met compulsory otherwise the solution found is deemed not feasible), the whole information available about the uncertainty can be taken into account by the optimization model. A particular example of optimization problems, where the constraints are soft, is the management optimization of infrastructure facilities which is subject to limited budget constraints. In this case, constraining the expected direct cost that is to be paid by the manager during each time period, to be less than or equal to the imposed budgets limits is deemed to be sufficient. In this paper, we present an extension of dynamic programming algorithms used to solve finite horizon MDPs or Partially Observable MDPs (POMDPs) in order to take into account second-order epistemic uncertainties. The whole information available about the uncertainty is taken into account by the optimization model. In addition, we propose a methodology based on the use of Dirichlet distributions which avoid, in our sense, much of the controversy dominating the literature about the different methods used to estimate the Markov transition matrices (maximum likelihood estimation, regression using state expectation, regression using state distribution, etc.). We illustrate the methodology by applying it to a Generalized Partially Observable Markov Decision Process (GPOMDP) (Faddoul, Raphael, & Chateauneuf, 2009). A literature review on POMDPs and GPOMDPs is provided in the following section. Although the proposed methodology for including epistemic uncertainties in the decision process is motivated by IM&R optimization, the formulation is quite general and can be relevant to any MDP.

3. Dynamic programming for finite horizon POMDPs

The methodology presented in this paper uses dynamic programming to solve finite horizon MDPs. This technique has been used extensively for maintenance optimization (Wang, 2002). However, a legitimate question arises about the suitability of the Markov property to describe the state evolution of civil engineering structures. Conflicting arguments are presented in literature concerning this subject. Although Neves, Frangopol, and Cruz (2006) have criticized the Markov property assumption for civil engineering structures, Orcesi and Cremona (2010) have demonstrated that the homogeneous Markov assumption was justified in the case of the French national bridge stock. Whatever the case may be, a technique proposed by Robelin and Madanat (2007) allows the dynamic programming to take into account the possible non-Markovian effects of actions and/or deterioration processes.

A dynamic programming model is said to be deterministic if the application of an action or a degradation process at the beginning of stage n will result in a precisely known state of the system at the beginning of stage $n + 1$. On the other hand, in probabilistic dynamic programming, the application of an action or a degradation process at the beginning of stage n will lead to a probabilistic distribution of the state of the system at the beginning of the next stage (Pham & Wang, 1996).

In a maintenance model, large uncertainties are associated with the results of any maintenance action and/or degradation process, due to two main factors:

1 imperfectness of maintenance actions;
2 stochastic nature of the degradation process between the application time of the maintenance action and the beginning of the next stage.

Due to the abovementioned uncertainties, probabilistic dynamic programming models were used in the majority of maintenance optimization problems.

However, classical probabilistic dynamic programming models in maintenance assume perfect inspections at the beginning of each time period. This type of models suffers from two main drawbacks; namely:

1 The inspections are assumed to be perfect and this is rarely the case;
2 The optimization of inspection planning (which is vital because inspections have usually significant costs) is not possible.

The use and development of POMDPs was essentially driven by the goal of dealing with the first one of these two shortcomings of classical probabilistic MDPs.