4 SINGULAR PERTURBATIONS OF MARKOV CHAINS AND DECISION PROCESSES

Konstantin E. Avrachenkov

Jerzy Filar*

Moshe Haviv

Abstract: In this survey we present a unified treatment of both singular and regular perturbations in finite Markov chains and decision processes. The treatment is based on the analysis of series expansions of various important entities such as the perturbed stationary distribution matrix, the deviation matrix, the mean-passage times matrix and others.

4.1 BACKGROUND AND MOTIVATION

Finite state Markov Chains (MC's) are among the most widely used probabilistic models of discrete event stochastic phenomena. Named after A.A. Markov, a famous Russian mathematician, they capture the essence of the existentialist "here and now" philosophy in the so-called "Markov property" which, roughly speaking, states that probability transitions to a subsequent state depend only on the current state and time. This property is less restrictive than might appear at first because there is a great deal of flexibility in the choice of what constitutes the "current state". Because of their ubiquitous nature Markov Chains are, nowadays, taught in many undergraduate and graduate courses ranging from mathematics, through engineering to business administration and finance.

*This research was supported in part by the ARC grant #A49906132.

Whereas a MC often forms a good description of some discrete event stochastic process, it is not automatically equipped with a capability to model such a process in the situation where there may be a "controller" or a "decision-maker" who—by a judicious choice of actions—can influence the trajectory of the process. This innovation was not introduced until the seminal works of Howard [44] and Blackwell [16] that are generally regarded as the starting point of the modern theory of Markov Decision Processes (or MDP's for short). Since then, MDP's have evolved rapidly to the point that there is now a fairly complete existence theory, and a number of good algorithms for computing optimal policies with respect to criteria such as maximization of limiting average expected reward, or the discounted expected reward.

The bulk of the, now vast, literature on both MCs and MDPs deals with the "perfect information" situations where all the model parameters—in particular probability transitions—are assumed to be known precisely. However, in most applications this assumption will be violated. For instance, a typical parameter, ρ, would normally be replaced by an estimate

$$\hat{\rho} = \rho + \varepsilon(n)$$

where the error term, $\varepsilon(n)$, comes from a statistical procedure used to estimate ρ and n is the number of observations used in that estimation. In most of the valid statistical procedures $|\varepsilon(n)| \downarrow 0$ as $n \uparrow \infty$, in an appropriate sense. Thus, from a perturbation analysis point of view, it is reasonable to suppress the argument n and simply concern ourselves with the effects of $\varepsilon \to 0$.

Roughly speaking, the subject of perturbation analysis of MC's and MDP's divides naturally into the study of "regular" and "singular" perturbations. Intuitively, regular perturbations are "good" in the sense that the effect of the perturbation dissipates harmlessly as $\varepsilon \to 0$, whereas singular perturbations are "bad" in the sense that small changes of ε (in a neighborhood of 0) can induce "large" effects. Mathematically, it can be shown that singular perturbations are associated with a change of the rank of a suitably selected matrix. This can be easily seen from the now classical example due to Schweitzer [66] where the perturbed probability transition matrix

$$P(\varepsilon) = \begin{pmatrix} 1-\frac{\varepsilon}{2} & \frac{\varepsilon}{2} \\ \frac{\varepsilon}{2} & 1-\frac{\varepsilon}{2} \end{pmatrix} \xrightarrow{\varepsilon \downarrow 0} P(0) = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}.$$

but the limit matrix for $\varepsilon > 0$

$$P^*(\varepsilon) \equiv \lim_{t\to\infty} P^t(\varepsilon) = \begin{pmatrix} \frac{1}{2} & \frac{1}{2} \\ \frac{1}{2} & \frac{1}{2} \end{pmatrix} \not\to P^*(0) \equiv \lim_{t\to\infty} P^t(0) = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}.$$

Indeed, the rank of $P^*(\varepsilon)$ is 1 for all $\varepsilon > 0$ (regardless of how close it is to zero), but it jumps to 2 at $\varepsilon = 0$; despite the fact that $P(\varepsilon) \to P(0)$. Thus, we see that singular perturbations can occur in MC's in a very natural and essential way. The latter point can be underscored by observing the behavior of $(I - \lambda P)^{-1}$ as $\lambda \to 1$, where P is a Markov (probability transition) matrix. It is well-known (e.g., see [16] or [57]) that this inverse can be expanded as a

Laurent series (with a pole of order 1) in the powers of $\varepsilon := 1-\lambda$. Indeed, much of the theory devoted to the connections between the discounted and limiting average MDP's exploits the asymptotic properties of this expansion, as $\varepsilon \downarrow 0$. Of course, the rank of $(I - \lambda P)$ changes when $\lambda = 1(\varepsilon = 0)$.

It is not surprising, therefore, that the literature devoted to singularly perturbed MC's and MDP's has been growing steadily in recent years. In fact there have been quite a few developments since the 1995 survey by Abbad and Filar [2].

The purpose of this survey paper is to present an up to date outline of a unified treatment of both singular and regular perturbations in MC's and MDP's that is based on series expansions of various important entities such as the perturbed stationary distribution matrix, the deviation matrix, the mean-passage times matrix and the resolvent-like matrix $(I - \lambda P)^{-1}$. From this series expansion perspective, the regular perturbations are simply the cases where Laurent series reduce to power series. Consequently, the capability to characterize and/or compute the coefficients of these expansions and the order of the pole (if any, at $\varepsilon = 0$) becomes of paramount importance.

This survey covers only the results on *discrete time* MC's and MDP's. For a parallel development of the *continuous time* models we refer an interested reader to the comprehensive book of Yin and Zhang [76]. For related emerging results in game theory see Altman et al. [6].

The logical structure of the survey is as follows: in Section 2, perturbations of (uncontrolled) Markov chains are discussed from the series expansions perspective; in Section 3 the consequences of these results are discussed in the context of optimization problems arising naturally in the (controlled) MDP case; and, finally, in Section 4 applications of perturbed MDP's to the Hamiltonian Cycle Problem are outlined. This last section demonstrates that the theory of perturbed MDP's has applications outside of its own domain.

4.2 UNCONTROLLED PERTURBED MARKOV CHAINS

4.2.1 Introduction and preliminaries

Let $P \in R^{n\times n}$ be a transition stochastic matrix representing transition probabilities in a Markov chain. Suppose that the structure of the underlying Markov chain is aperiodic. Let $P^* = \lim_{t\to\infty} P^t$ which is well-known to exist for aperiodic processes. In the case when the process is also ergodic, P^* has identical rows, each of which is the stationary distribution of P, denoted by π. Let Y be the *deviation matrix* of P which is defined by $Y = (I - P + P^*)^{-1} - P^*$. It is well known (e.g., see [51]) that Y exists and it is the unique matrix satisfying

$$Y(I-P) = I - P^* = (I-P)Y \text{ and } P^*Y = 0 = Y\underline{1} \tag{4.1}$$

(where $\underline{1}$ is a matrix full of 1's) making it the group inverse of $I - P$. Finally, $Y = \lim_{T\to\infty} \Sigma_{t=0}^{T}(P^t - P^*)$. Let M_{ij} be the mean passage time from state-i into state-j. It is also known that when the corresponding random variable is proper, then M_{ij} is finite. Of course, the matrix M is well-defined if and only