Available online at www.sciencedirect.com
ScienceDirect
Stochastic Processes and their Applications 119 (2009) 562–587
stochastic processes and their applications
www.elsevier.com/locate/spa

Splitting for rare event simulation: A large deviation approach to design and analysis

Thomas Dean, Paul Dupuis*

Lefschetz Center for Dynamical Systems, Division of Applied Mathematics, Brown University, United States

Received 24 April 2007; received in revised form 3 February 2008; accepted 4 February 2008
Available online 8 April 2008

Abstract

Particle splitting methods are considered for the estimation of rare events. The probability of interest is that a Markov process first enters a set B before another set A, and it is assumed that this probability satisfies a large deviation scaling. A notion of subsolution is defined for the related calculus of variations problem, and two main results are proved under mild conditions. The first is that the number of particles generated by the algorithm grows subexponentially if and only if a certain scalar multiple of the importance function is a subsolution. The second is that, under the same condition, the variance of the algorithm is characterized (asymptotically) in terms of the subsolution. The design of asymptotically optimal schemes is discussed, and numerical examples are presented.

MSC: 60F10; 65C05; 60J85

Keywords: Rare event; Monte Carlo; Branching process; Large deviations; Subsolutions; Hamilton–Jacobi–Bellman equation; Simulation; Variance reduction

1. Introduction

Numerical estimation of probabilities of rare events is a difficult problem. There are many potential applications in operations research and engineering, insurance, finance, chemistry, biology, and elsewhere, and many papers (and by now even a few books) have proposed numerical schemes for particular settings and applications. Because the quantity of interest is very small, standard Monte Carlo simulation requires an enormous number of samples for

* Corresponding author.
E-mail address: Paul_Dupuis@Brown.edu (P. Dupuis).

0304-4149/$ - see front matter
doi:10.1016/j.spa.2008.02.017

the variance of the resulting estimate to be comparable to the unknown probability. It quickly becomes unusable, and more efficient alternatives are sought.

The two most widely considered alternatives are those based on change-of-measure techniques and those based on branching processes. The former is usually called *importance sampling*, and the latter is often referred to as *multi-level splitting*. While good results on a variety of problem formulations have been reported for both methods, it is also true that both methods can produce inaccurate and misleading results. The design issue is critical, and one can argue that proper theoretical tools for the design of importance sampling and splitting algorithms were simply not available for complicated models and problem formulations. [An alternative approach based on interacting particles has also been suggested as in [3]. However, we are unaware of any analysis of the performance of these schemes as the probability of interest becomes small.]

Suppose that the probability of interest takes the form $p = P\{Z \in G\} = \mu(G)$, where G is a subset of some reasonably regular space (e.g., a Polish space S) and μ a probability measure. In ordinary Monte Carlo one generates a number of independent and identically distributed (iid) samples $\{Z_i\}$ from μ, and then estimates p using the sample mean of $1_{\{Z_i \in G\}}$. In the case of importance sampling, one uses an alternative sampling distribution ν, generates iid samples $\left\{\bar{Z}_i\right\}$ from ν, and then estimates via the sample mean of $[\mathrm{d}\mu/\mathrm{d}\nu](\bar{Z}_i)1_{\{\bar{Z}_i \in G\}}$. The Radon–Nikodim derivative $[\mathrm{d}\mu/\mathrm{d}\nu](\bar{Z}_i)$ guarantees that the estimate is unbiased. The goal is to choose ν so that individual samples $[\mathrm{d}\mu/\mathrm{d}\nu](\bar{Z}_i)1_{\{\bar{Z}_i \in G\}}$ cluster tightly around p, thereby reducing the variance. However, for complicated process models or events G the selection of a good measure ν may not be simple. The papers [10,11] show how certain standard heuristic methods based on ideas from large deviations could produce very poor results. The difficulty is due to points in S with low probability under ν for which $\mathrm{d}\mu/\mathrm{d}\nu$ is very large. The aforementioned large deviation heuristic does not properly account for the contribution of these points to the variance of the estimate, and it is not hard to find examples where the corresponding importance sampling estimator is much worse than even ordinary Monte Carlo. These estimates exhibit very inaccurate and/or unstable behavior, though the instability may not be evident from numerical data until massive amounts have been generated.

The most discussed application of splitting type schemes is to first entrance probabilities, and to continue the discussion we specialize to that case. Thus Z is the sample path of a stationary stochastic process $\{X_i\}$ (which for simplicity is taken to be Markovian), and G is the set of trajectories that first enter a set B prior to entering a set A. More precisely, for disjoint B and A and $x \notin A \cup B$,

$$p = p(x) = P\left\{X_j \in B, X_i \notin A, i \in \{0, \ldots, j\}, j < \infty | X_0 = x\right\}.$$

In the most simple version of splitting, the state space is partitioned according to certain sets $B \subset C_0 \subset C_1 \subset \cdots \subset C_K$, with $x \notin C_K$ and $A \cap C_K = \emptyset$. These sets are often defined as level sets of a particular function V, which is commonly called an *importance function*. Particles are generated and killed off according to the following rules. A single particle is started at x. Generation of particles (splitting) occurs whenever an existing particle reaches a threshold or level C_i for the first time. At that time, a (possibly random) number of new particles are placed at the location of entrance into C_i. Future evolutions of these particles are independent of each other (and all other particles), and follow the law of $\{X_i\}$. Particles are killed if they enter A before B. Attached to each particle is a weight. Whenever a particle splits, the weight of each descendent equals that of the parent times a discount factor. A random tree is thereby produced, with each leaf corresponding to a particle that has either reached B or been killed. A random variable (roughly