ELSEVIER

Statistics & Probability Letters 56 (2002) 143–146

STATISTICS & PROBABILITY LETTERS

www.elsevier.com/locate/stapro

Hoeffding's inequality for uniformly ergodic Markov chains

Peter W. Glynn[a], Dirk Ormoneit[b,*]

[a] *Department of Management Science and Engineering, Stanford University, Stanford, CA 94305-9010, USA*
[b] *Department of Computer Science, Stanford University, Stanford, CA 94305-9010, USA*

Received March 2001; received in revised form June 2001

Abstract

We provide a generalization of Hoeffding's inequality to partial sums that are derived from a uniformly ergodic Markov chain. Our exponential inequality on the deviation of these sums from their expectation is particularly useful in situations where we require uniform control on the constants appearing in the bound. © 2002 Elsevier Science B.V. All rights reserved

Keywords: Hoeffding's inequality; Markov chains; Large deviations

1. Introduction

Hoeffding's inequality is a key tool in the analysis of many problems arising in both probability and statistics. Given a sequence $\mathbf{Y} \equiv (Y_i : i \geqslant 0)$ of independent and bounded random variables, Hoeffding's inequality provides an exponential bound on partial sums of the form $S_n = Y_0 + \cdots + Y_{n-1}$.

Theorem 1 (Hoeffding's inequality)**.** *Suppose that for each $i \geqslant 0$ there exist real numbers a_i and b_i such that $P(Y_i \in [a_i, b_i]) = 1$. Then for any $\varepsilon > 0$, we have*

$$\mathbf{P}(S_n - \mathbf{E}[S_n] \geqslant n\varepsilon) \leqslant \exp\left(-2n^2\varepsilon^2 \Big/ \sum_{i=0}^{n-1} (b_i - a_i)^2 \right).$$

For a proof, see Hoeffding (1963).

As indicated above, this result has found broad applicability in many different settings. See, for example, Serfling (1980) for various statistical contexts within which the inequality plays a central role. Devroye et al. (1996) illustrate the importance of this inequality in the classification setting. The explicit nature of the constants in the bound makes it especially attractive in contexts within which one needs to establish that the probability in question decays exponentially in n in some uniform fashion.

* Corresponding author. Fax: 49-89-211-13311
E-mail address: ormoneit@cs.stanford.edu (D. Ormoneit).

PII: S0167-7152(01)00158-4

Our interest in the inequality arose from its rule in the analysis of a reinforcement learning algorithm; see Ormoneit and Glynn (2001). Our specific application requires an extension of Hoeffding's inequality to the setting of Markov chains. It is important to recognize that some appropriate mixing assumption must be enforced, in addition to requiring boundedness of the summands, if the qualitative form of the inequality is to be retained. To see this, consider the degenerate setting in which $Y_i = Y_0$ for $i \geqslant 1$, in which case any upper bound on $\mathbf{P}(S_n - \mathbf{E}[S_n] \geqslant n\varepsilon)$ cannot decay to zero as $n \to \infty$. Consequently, the extension of Hoeffding's inequality to the Markov context will require imposing an appropriate recurrence condition on the chain (that will induce an appropriate level of mixing).

Specifically, let $\mathbf{X} \equiv (X_n\colon n \geqslant 0)$ be a Markov chain taking values in a state space S. We require the following condition on X:

(A.1) There exists a probability measure φ on S, $\lambda > 0$, and an integer $m \geqslant 1$ such that

$$\mathbf{P}_x(X_m \in \cdot) \geqslant \lambda\varphi(\cdot)$$

for each $x \in S$.

Here $\mathbf{P}_x(\cdot)$ denotes the conditional probability $\mathbf{P}(\cdot|X_0 = x)$. Condition (A.1) is closely related to the assumptions of uniform ergodicity (see Meyn and Tweedie, 1993) and Doeblin recurrence (see Doob, 1953). It is often easy to verify for chains taking values in a compact state space. It should be noted that chains satisfying (A.1) automatically possess a unique stationary distribution π (Meyn and Tweedie, 1993).

To present our main result, let $f : S \to \mathbb{R}$, set $Y_i \equiv f(X_i)$, and let $S_n \equiv \sum_{i=0}^{n-1} Y_i$. For example, Y_i may be the reward for visiting the state X_i, as in the case of a Markov reward process (Puterman, 1994). Furthermore, let the norm and the long-term expected value of f be defined according to $\|f\| \equiv \sup\{|f(x)|\colon x \in S\}$ and $\alpha \equiv \int_S f(x)\pi(\mathrm{d}x)$, respectively. Our main result is the following Markovian extension of Hoeffding's inequality.

Theorem 2. *Assume* (A.1) *and suppose that* $\|f\| < \infty$. *Then we have*

$$\mathbf{P}_x(S_n - \mathbf{E}[S_n] \geqslant n\varepsilon) \leqslant \exp\left(-\frac{\lambda^2(n\varepsilon - 2\|f\|m/\lambda)^2}{2n\|f\|^2 m^2}\right)$$

for $n > 2\|f\|m/(\lambda\varepsilon)$.

Note that this bound has the same degree of explicitness in terms of the underlying "problem data" as does the classical Hoeffding's inequality, and displays the same qualitative behavior. One interesting feature of our proof, supplied in Section 2, is that it depends critically on the "additive form" of Poisson's equation for the function f. This is perhaps surprising in view of the recent work of Balaji and Meyn (2000) which emphasizes the role of the *multiplicative* Poisson's equation in the large deviations theory for S_n. However, it should be noted that our analysis assumes the recurrence hypothesis (A.1), whereas Meyn's work covers more general chains.

2. Proof of Theorem 2

Set $f_c(x) \equiv f(x) - \alpha$ and note that $S_n - n\alpha = \sum_{i=0}^{n-1} f_c(X_i)$. Under condition (A.1), it is known that

$$|\mathbf{E}_x f_c(X_n)| \leqslant \|f\| \cdot (1-\lambda)^{\lfloor n/m \rfloor};$$

see Asmussen et al. (1992) or Rosenthal (1992) for a proof.[1] (Here, $\mathbf{E}_x(\cdot) \equiv \mathbf{E}[\,\cdot\,|X_0 = x]$.) Hence,

$$g(x) = \sum_{n=0}^{\infty} \mathbf{E}_x f_c(X_n)$$

[1] The bound is proven there with $\|f_c\|$ instead of $\|f\|$ on the right-hand side, from which the inequality above follows.

converges absolutely and

$$\|g\| \leqslant \|f\| \cdot m/\lambda. \tag{1}$$

Furthermore, g solves Poisson's equation:

$$g(x) - \mathbf{E}_x g(X_1) = f_c(x) \tag{2}$$

for $x \in S$. Observe that

$$D_i \equiv g(X_i) - \mathbf{E}_x[g(X_i)|X_0, \ldots, X_{i-1}]$$

is a martingale difference for $i \geqslant 0$. Furthermore, (2) implies that

$$S_n - n\alpha = \sum_{i=1}^{n} D_i + g(X_0) - g(X_n).$$

It follows that for $\theta \geqslant 0$,

$$\mathbf{E}_x \exp(\theta(S_n - n\alpha)) \leqslant \exp(2\theta\|g\|) \cdot \mathbf{E}_x \exp\left(\theta \sum_{i=1}^{n} D_i\right). \tag{3}$$

But

$$\mathbf{E}_x \exp\left(\theta \sum_{i=1}^{n} D_i\right) = \mathbf{E}_x \exp\left(\theta \sum_{i=1}^{n-1} D_i\right) \mathbf{E}_x[\exp(\theta D_n)|X_0, \ldots, X_{n-1}]. \tag{4}$$

However, we may now take advantage of a key step in the proof of the classical Hoeffding's inequality (conditional on $X_0, \ldots, X_{i-1}$) to conclude that, because D_i lies a.s. in an interval of length $2\|g\|$,

$$\mathbf{E}_x[\exp(\theta D_n)|X_0, \ldots, X_{n-1}] \leqslant \exp(\theta^2 \|g\|^2/2); \tag{5}$$

see Lemma 8.1 of Devroye et al. (1996) for a proof of inequality (5). Inequalities (3) and (5), together with (4) recursively applied, yield the inequality

$$\mathbf{E}_x \exp(\theta(S_n - n\alpha)) \leqslant \exp(2\theta\|g\| + n\theta^2 \|g\|^2/2).$$

Markov's inequality then establishes the tail bound

$$\mathbf{P}_x(S_n - n\alpha \geqslant n\varepsilon) \leqslant \exp(-\theta n\varepsilon + 2\theta\|g\| + n\theta^2 \|g\|^2/2). \tag{6}$$

The value of θ that minimizes this bound is

$$\theta_n = \frac{n\varepsilon - 2\|g\|}{n\|g\|^2}.$$

Substituting θ_n into (6) and exploiting the bound (1) establishes the theorem. □

Note that we also could have directly applied Azuma's inequality (see Azuma, 1967) to our martingale representation for S_n, in order to obtain an exponential inequality for the sum S_n. However, Azuma's inequality involves an a.s. bound on $|D_i|$, which in our setting could be as large as $2\|g\|$. This introduces an extra factor of 4 into the exponent of (5), and thereby reduces the effectiveness of our probability bound by a similar factor.

References

Asmussen, S., Glynn, P.W., Thorisson, H., 1992. Stationarity detection in the initial transient problem. ACM Trans. Modeling and Comput. Simulation 2, 130–157.
Azuma, K., 1967. Weighted sums of certain dependent random variables. Tôhoku Math. J. 19 (3), 357–367.
Balaji, S., Meyn, S.P., 2000. Multiplicative ergodicity for irreducible Markov chains. Stochastic Process. Appl., 90, 123–144.
Devroye, L., Györfi, L., Lugosi, G., 1996. A Probabilistic Theory of Pattern Recognition. Springer, New York.
Doob, J.L., 1953. Stochastic Processes. Wiley, New York.
Hoeffding, W., 1963. Probability inequalities for sums of bounded random variables. J. Amer. Statist. Assoc. 58, 13–30.
Meyn, S.P., Tweedie, R.L., 1993. Markov Chains and Stochastic Stability. Springer, New York.
Ormoneit, D., Glynn, P.W., 2001. Kernel-based reinforcement learning in average-cost problems. IEEE Transactions on optimal control, to be published.
Puterman, M.L., 1994. Markov Decision Processes: Discrete Stochastic Dynamic Programming. Wiley, New York.
Rosenthal, J.S., 1992. Rates of Convergence for Gibbs Sampler and other Markov Chains. Doctoral Dissertation, Harvard University.
Serfling, R.J., 1980. Approximation Theorems of Mathematical Statistics. Wiley, New York.