SEQUENTIAL METHODS IN STATISTICS
BANACH CENTER PUBLICATIONS, VOLUME 16
PWN–POLISH SCIENTIFIC PUBLISHERS
WARSAW 1985

PARAMETER ESTIMATION IN STOCHASTIC SYSTEMS: SOME RECENT RESULTS AND APPLICATIONS

VIVEK S. BORKAR

Department of Applied Mathematics,
Twente University of Technology,
Enschede, The Netherlands

Some recent work on the characterization of almost sure limit sets for maximum likelihood estimates for stochastic systems is reviewed. Applications to allied topics such as input selections for identification, model selection, self-tuning etc. are briefly discussed.

I. Introduction

This paper reviews some recent work by the author and others ([1]–[4]) concerning the characterization of limit sets for maximum likelihood estimates for stochastic systems. A general framework for these problems is developed and specific results presented for Ito processes and controlled Markov chains. For proof etc., see [1]–[4]. Applications to related topics are discussed.

II. The general framework

Consider a stochastic process $\{X_t\}_{t\in J}$ where J is either $R^+ = [0, \infty)$ or N (the set of natural numbers). For some compact metric space D, let P_θ, $\theta \in D$, be a family of probability measures on the trajectory ("canonical") space of $\{X_t\}_{t\in J}$, equipped with the appropriate topology and the corresponding Borel σ-field. Assume that the "true" probability measure is P_{θ_0} for some $\theta_0 \in D$. $P_{\theta T}$, $T\in J$, will denote the restriction of P_θ to the σ-field generated by $\{X_t, t \leqslant T\}$. For each $T\in J$, we assume that $P_{\theta T}$, $\theta \in D$, is a family of mutually absolutely continuous probability measures. Then for $t\in J$, $\theta\in D$, the Radon–Nikodym derivative

$$\Lambda_t(\theta) = \frac{dP_{\theta T}}{dP_{\theta_0 t}}$$

is a martingale with respect to the progressive σ-fields $\mathscr{F}_t$ generated by $\{X_s, s \leqslant t\}$. By Jensen's inequality, it follows that

$$L_t(\theta) = \ln \Lambda_t(\theta)$$

is an $\mathscr{F}_t$-supermartingale. Under suitable conditions, we can apply the Doob–Meyer decomposition theorem to obtain the unique decomposition

$$L_t(\theta) = M_t(\theta) - A_t(\theta) \tag{2.1}$$

for each $\theta \in D$, where $M_t(\theta)$ is a zero-mean $\mathscr{F}_t$-martingale and $A_t(\theta)$ is an $\mathscr{F}_t$-predictable nonnegative increasing process.

Under appropriate moment conditions, the following "strong law of large numbers" holds for each $\theta \in D$:

$$\lim_{t \to \infty} \frac{M_t(\theta)}{t} = 0 \quad \text{a.s.} \tag{2.2}$$

Suppose we can prove that both $\dfrac{M_t(\theta)}{t}$ and $\dfrac{A_t(\theta)}{t}$ are continuous in θ on D, uniformly in t belonging to some subset of $\boldsymbol{J}$ of the form $\{s \in \boldsymbol{J} \mid s \geqslant a\}$, $a \in \boldsymbol{J}$. Then (2.2) yields

$$\lim_{t \to \infty} \sup_{\theta \in D} \frac{M_t(\theta)}{t} = 0 \quad \text{a.s.} \tag{2.3}$$

Let $\hat{\theta}_t$ be the maximum likelihood estimate of θ_0 at t, i.e., at each $t \in \boldsymbol{J}$

$$L_t(\hat{\theta}_t) \geqslant L_t(\theta) \qquad \forall \theta \in D.$$

In particular,

$$L_t(\hat{\theta}_t) \geqslant L_t(\theta_0) = 0.$$

Hence

$$\frac{M_t(\hat{\theta}_t)}{t} \geqslant \frac{A_t(\hat{\theta}_t)}{t} \geqslant 0.$$

By (2.3)

$$\lim_{t \to \infty} \frac{A_t(\hat{\theta}_t)}{t} = 0 \quad \text{a.s.}$$

Thus

$$\hat{\theta}_t \to \{\theta \in D \mid \liminf_{t \to \infty} \frac{A_t(\theta)}{t} = 0\} \quad \text{a.s.} \tag{2.4}$$

This gives a very general characterization of limit sets for maximum likelihood estimates from which stronger results can be derived for more specific cases. The arguments leading from (2.2) to (2.3) depend very much on the problem at hand, as seen by comparing [2] and [3].

We next see examples of how the relation (2.1) may arise.

EXAMPLE 1. Let $J = N$. Assuming that the appropriate probability densities and conditional probability densities exist,

$$\Lambda_t(\theta) = \frac{P_\theta(X_1, X_2, \ldots, X_t)}{P_{\theta_0}(X_1, X_2, \ldots, X_t)} = \prod_{i=1}^{t} \frac{P_\theta(X_i | \mathscr{F}_{i-1})}{P_{\theta_0}(X_i | \mathscr{F}_{i-1})}$$

with $\mathscr{F}_i = \sigma(X_1, X_2, \ldots, X_i)$ for $i \geqslant 1$, $\mathscr{F}_0$ the trivial σ-field. Then we have

$$L_t(\theta) = \sum_{j=1}^{t} \ln \frac{P_\theta(X_j | \mathscr{F}_{j-1})}{P_{\theta_0}(X_j | \mathscr{F}_{j-1})} = \sum_{j=1}^{t} \left(\ln \frac{P_\theta(X_j | \mathscr{F}_{j-1})}{P_{\theta_0}(X_j | \mathscr{F}_{j-1})} - \right.$$
$$\left. - E\left(\ln \frac{P_\theta(X_j | \mathscr{F}_{j-1})}{P_{\theta_0}(X_j | \mathscr{F}_{j-1})} \middle| \mathscr{F}_{j-1} \right) \right) - \sum_{j=1}^{t} E\left(-\ln \frac{P_\theta(X_j | \mathscr{F}_{j-1})}{P_{\theta_0}(X_j | \mathscr{F}_{j-1})} \middle| \mathscr{F}_{j-1} \right).$$

Applying Jensen's inequality to the summands in the second sum, we can verify that the above expression is in the form (2.1).

EXAMPLE 2. Let $J = \boldsymbol{R}^+$ and suppose that $\{X_t\}_{t \in \boldsymbol{R}}$ has continuous sample paths. In many cases, the Ito differential rule can be applied to write the following version of (2.1):

$$L_t(\theta) = \int_0^t \frac{1}{\Lambda_s} d\Lambda_s - \frac{1}{2} \int_0^t \frac{1}{\Lambda_s^2} d\langle \Lambda, \Lambda \rangle_s$$

where $\langle \Lambda, \Lambda \rangle_t$ is the compensator of Λ_t^2, $t \in \boldsymbol{R}$.

EXAMPLE 3. Let $J = \boldsymbol{R}^+$ and suppose that $\{X_t\}_{t \in \boldsymbol{R}}$ is a point process with an intensity $\{\lambda_t(\theta)\}_{t \in \boldsymbol{R}}$, which, for each $\theta \in A$, is a positive predictable process. Then

$$\ln \Lambda_t = \int_0^t \ln \frac{\lambda_s(\theta)}{\lambda_s(\theta_0)} dX_s - \int_0^t (\lambda_s(\theta) - \lambda_s(\theta_0)) ds$$
$$= \int_0^t \ln \left(\frac{\lambda_s(\theta)}{\lambda_s(\theta_0)} \right) (dX_s - \lambda_s(\theta_0) ds) - \int_0^t \left(\frac{\lambda_s(\theta)}{\lambda_s(\theta_0)} - 1 - \ln \left(\frac{\lambda_s(\theta)}{\lambda_s(\theta_0)} \right) \right) \lambda_s(\theta_0) ds.$$

From the inequality $x - 1 \geqslant \ln x$, one can verify that this has the form (2.1).

III. Parameter estimation for certain processes

Let D be a compact subset of $\boldsymbol{R}^n$. Suppose $\{X_t\}_{t\in\boldsymbol{R}^+}$ satisfies the stochastic differential equation

$$dX_t = m_t(\theta, X_{[0,t]})\,dt + dW_t, \qquad X_0 = 0 \quad \text{a.s.} \tag{3.1}$$

where $m_t: D \times C[0, t] \to \boldsymbol{R}$ is a measurable map, and $X_{[0,t]}$ denotes the trajectory $\{X_s, s \in [0, t]\}$. Under suitable conditions on m_t, the above equation has a unique strong solution $\{X_t^\theta\}_{t\in\boldsymbol{R}^+}$, which induces on $C[0, \infty)$ a probability measure P_θ such that for each $T \in [0, \infty)$,

$$\frac{dP_{\theta T}}{dP_{\theta_0 T}} = \exp\left\{ \int_0^T \left(m_t(\theta, X^{\theta_0}_{[0,t]}) - m_t(\theta_0, X^{\theta_0}_{[0,t]})\right) d\tilde{W}_t - \frac{1}{2}\int_0^T \left(m_t(\theta, X^{\theta_0}_{[0,t]}) - m_t(\theta_0, X^{\theta_0}_{[0,t]})\right)^2 dt \right\}$$

where $\tilde{W}_t$, $t \in \boldsymbol{R}^+$, is a Wiener process under P_{θ_0}. We have the following correspondence with (2.1):

$$M_t(\theta) = \int_0^t \left(m_s(\theta, X^{\theta_0}_{[0,s]}) - m_s(\theta_0, X^{\theta_0}_{[0,s]})\right) d\tilde{W}_s,$$

$$A_t(\theta) = \tfrac{1}{2}\int_0^t \left(m_s(\theta, X^{\theta_0}_{[0,s]}) - m_s(\theta_0, X^{\theta_0}_{[0,s]})\right)^2 ds.$$

Under appropriate continuity and moment conditions on m_t, one can prove the uniform continuity of $\dfrac{M_t(\theta)}{t}$ and $\dfrac{A_t(\theta)}{t}$ in θ with respect to $t \in [1, \infty)$. Then (2.4) yields

$$\hat{\theta}_t \to \left\{\theta \in D \,\middle|\, \liminf_{t\to\infty} \int_0^t \left(m_s(\theta, X^{\theta_0}_{[0,s]}) - m_s(\theta_0, X^{\theta_0}_{[0,s]})\right)^2 ds = 0\right\} \quad \text{a.s.} \tag{3.2}$$

If the drift term $m_t(\theta, X^{\theta_0}_{[0,t]})$ has the form $\tilde{m}(\theta, X_t^{\theta_0})$ for some $\tilde{m}$: $D \times \boldsymbol{R} \to \boldsymbol{R}$, then the process $X_t^{\theta_0}$, $t \in \boldsymbol{R}^+$ is Markov. Suppose, in addition, that it has a unique invariant probability measure ν, then

$$\hat{\theta}_t \to \{\theta \in D \mid E_\nu[(\tilde{m}(\theta, X_t^{\theta_0}) - \tilde{m}(\theta_0, X_t^{\theta_0}))^2] = 0\} \quad \text{a.s.} \tag{3.3}$$

where $E_\nu(\cdot)$ denotes the expectation with respect to ν.

The basic methodology used above can also be applied to some closely related problems. An important problem is the situation when $\theta_0 \notin D$. (3.2), (3.3) then have to be replaced by (3.4), (3.5) respectively:

$$\liminf_{t\to\infty} \frac{1}{t} \int_0^t (m_s(\theta, X^{\theta_0}_{[0,s]}) - m_s(\theta_0, X^{\theta_0}_{[0,s]}))^2 \, ds$$

$$\geqslant \limsup_{t\to\infty} \frac{1}{t} \int_0^t (m_s(\hat{\theta}_t, X^{\theta_0}_{[0,s]}) - m_s(\theta_0, X^{\theta_0}_{[0,s]}))^2 \, ds \quad \text{a.s.} \qquad \forall \theta \in D \tag{3.4}$$

and

$$\hat{\theta}_t \to \{\bar{\theta} \in D \mid E_\nu[(\tilde{m}(\bar{\theta}, X_t^{\theta_0}) - \tilde{m}(\theta_0, X_t^{\theta_0}))^2] = \min_{\theta \in D} E_\nu[(\tilde{m}(\theta, X_t^{\theta_0}) - \tilde{m}(\theta_0, X_t^{\theta_0}))^2]\} \quad \text{a.s.} \tag{3.5}$$

Also interesting is the case when θ_0 is an $\boldsymbol{R}^n$-valued random variable independent of W_t, $t \in \boldsymbol{R}^+$. It can be shown that (3.2)/(3.3) hold a.s. on the set $\{\theta_0 \in D\}$ and (3.4)/(3.5) hold a.s. on the set $\{\theta_0 \notin D\}$ [3].

Let X_n, $n = 1, 2, \ldots$, be a controlled Markov chain on a state space $S = \{1, 2, \ldots\}$ with transition probabilities $p(i, j; z_i, \theta)$, θ as before and z_i a control parameter in some compact metric space $H(i)$. Let Z_n, $n = 1, 2, \ldots$, be a control sequence of $H(X_n)$-valued random variables adapted to the progressive σ-fields generated by $X_1, X_2, \ldots$ Under suitable assumptions

$$\frac{dP_{\theta n}}{dP_{\theta_0 n}} = \prod_{m=1}^{n-1} \frac{p(X_m, X_{m+1}; Z_m, \theta)}{p(X_m, X_{m+1}; Z_m, \theta_0)}.$$

The log-likelihood function can be decomposed as in Example 1. Under suitable stability conditions,

$$\hat{\theta}_n \to \left\{\theta \in D \mid \liminf_{n\to\infty} \frac{1}{n} \sum_{m=1}^{n} \sum_{i \in S} p(X_n, i; Z_n, \theta_0) \ln \frac{p(X_n, i; Z_n, \theta)}{p(X_n, i; Z_n, \theta_0)} = 0\right\} \quad \text{a.s.}$$

For a Markov chain with stationary transition probabilities which is positive recurrent, this implies that the transition probabilities under θ_0 agree with those under any limit point of $\hat{\theta}_n$. For $\theta_0 \notin D$, an analog of (3.5) can be derived ([2]).

IV. Applications

(i) Identifiability conditions: The above results lead to easily characterized sufficient conditions for local or global strong consistency. Compare with the similar results for prediction error estimates [10].

(ii) Input selection for identification: The input for an identification experiment has a threefold role – to ensure model discrimination, to improve its rate and to ensure stability. The last is less apparent, but the proofs of [1]–[4] show that the stability assumption plays a key role in the asymptotic analysis. Suppose X_t, $t \in J$, is stationary under the input schemes being considered and

$$\lim_{t\to\infty} \frac{1}{t} \ln\left(\frac{dP_{\theta t}}{dP_{\theta_0 t}}\right) = -G(\theta, \theta_0) \quad \text{a.s. } P_{\theta_0}$$

for some function G: $D \times D \to [0,\infty)$. (Nonnegativity is obvious from (2.1)–(2.2).) For $\theta \neq \theta_0$, the discrimination between θ, θ_0 is ensured by making $G(\theta, \theta_0) > 0$ and its rate improved by increasing its value. Thus $\min_{\theta_2} \max_{\theta_1} G(\theta_1, \theta_2)$ can be used as a performance criterion for input schemes. (This is closely related to the various "information distances" and criteria based on those.) Randomized inputs will often do strictly better than nonrandomized ones ([1], [2]). Also, some nonstationary inputs may do better than the stationary ones. Heuristically, the trend of the estimates may suggest ruling out part of D or enlarging D after a finite time, causing a change in the input scheme. A nice way to achieve such adaptation would be to have a random input sequence generated according to some fixed parametrized distribution, these parameters being updated according to the trend in $\hat{\theta}_t$.

(iii) Model selection: As a paradigm, consider (3.5) with $\hat{m}_t = m(\theta_0, X_t^{\theta_0})$ being the "true" drift. We may choose

$$m(\theta, x) = \sum_{i=1}^{n} \theta(i) b_i(x)$$

for $\theta = [\theta(1), \theta(2), \ldots, \theta(n)] \in D \subset \mathbf{R}^n$ and b_i are suitably chosen functions based on prior intuition. Then (3.5) becomes

$$\hat{\theta}_t \to \{\theta \in D \mid \theta \text{ minimizes } E_v([\sum_{i=1}^{n} \theta(i) b_i(X_t) - \hat{m}_t]^2)\} \quad \text{a.s.}$$

Thus we can select the parametrization and the model set as per an appropriate trade-off between computational convenience and accuracy, an estimate of the latter being given by formulas such as (3.5).

(iv) Self-tuning control: The estimation scheme can be coupled with a feedback law dependent upon the current estimate to produce a self-tuning controller. The problem then is to ensure that the stability conditions assumed at the outset for the analysis above hold for this closed-loop situation. This program was carried out for controlled Markov chains in [1]–[2], where ε-optimal controls are derived using suitable randomizations.

It is apparent from [1]–[2] that the exact optimal cost may be attainable, in principle, by an adaptive randomization scheme such as the one proposed in (ii) above. Performance can also be improved by "holding" an estimate for an extended length of time instead of updating it at each instant [8]. Heuristically speaking, adding some "noise" to the self-tuning input helps. This may not have to be explicitly done, since the roundoff error etc. in numerical schemes may provide a natural "jitter". This might explain why the simulation studies for self-tuners give promising results even when the exact theoretical results seem hard to establish ([12]). It is also worth noting that stronger results are possible when D is finite ([5]) or some special parametrization is used ([11]). These are false in general, as pointed out in [7].

V. Related problems

(i) Consider a situation when a control system is given an "identifying" input for a finite length of time and the estimate at the end of that interval taken on faith as the true parameter, to select the optimal input thereafter. The choice of length of the estimation interval then becomes an optimal stopping problem.

(ii) In many cases, the penalty for a mistake in identification may be known. For example: In a self-tuning controller, let $C(\theta, \theta')$ be the average reward when the optimal strategy under θ is used and the true parameter is θ'. Then the penalty for choosing θ is simply $C(\theta, \theta_0)-C(\theta_0, \theta_0)$. Since θ_0 is unknown, we may use $\max_{\tilde{\theta}}|C(\theta, \tilde{\theta})-C(\tilde{\theta}, \tilde{\theta})|$ as the penalty function. It seems reasonable to expect better performance if such penalty functions are explicitly incorporated in estimation schemes, e.g., by "weighting" the estimation criterion in favour of θ with low penalty.

(iii) For applications, it is important to have recursive schemes. There is an increasing literature about these for linear systems, but the only effort in Markov chains seems to be [9].

(iv) A problem that has drawn some attention recently is the situation when many agents try to estimate the same system based on different observations. In [6], it was shown that in certain simple cases, their estimates converge to the same value if they incorporate in their own estimation schemes the estimates of others. Much more needs to be done in this direction.

References

[1] V. S. Borkar, *Identification and Adaptive Control of Markov Chains*, Ph. D. Thesis, Dept. of E.E.C.S., University of California, Berkeley 1980.

[2] V. S. Borkar and P. Varaiya, *Identification and Adaptive Control of Markov Chains*, SIAM J. Control Optim., to appear.
[3] V. S. Borkar and A. Bagchi, *Parameter Estimation in Continuous-Time Stochastic Processes*, Memorandum no. 331, Department of Applied Mathematics, Twente University of Technology, Enschede, The Netherlands, 1981.
[4] A. Bagchi and V. S. Borkar, *Parameter Identification in Infinite Dimensional Linear Systems*, in preparation.
[5] V. S. Borkar and P. Varaiya, *Adaptive Control of Markov Chains: Finite Parameter Set*, IEEE Trans. Automat. Control, **24** (1979).
[6] V. S. Borkar and P. Varaiya, *Asymptotic Agreement in Distributed Estimation*, IEEE Trans. Automat. Control, to appear.
[7] P. Kumar, *Adaptive Control with a Compact Parameter Set*, Math. Res. Report, No. 80–16, Dept. of Mathematics, University of Maryland, Baltimore County 1980.
[8] P. Kumar and A. Becker, *A New Family of Optimal Adaptive Controllers*, Math. Res. Report, No. 80–18, Dept. of Mathematics, University of Maryland, Baltimore County 1980.
[9] Y. M. El-Fattah, *Recursive Algorithms for Adaptive Control of Finite Markov Chains*, IEEE Trans. Systems Man Cybernet. **11** (1981).
[10] L. Liung, *On the Consistency of Prediction Error Identification Methods*, in: R. K. Mehra and D. G. Lainiotis (eds.), *System Identification: Advances and Case Studies*, Academic Press, New York 1976.
[11] B. Sagalovsky, *Adaptive Control and Parameter Estimation in Markov Chains: A Linear Case*, preprint.
[12] Sunil Shah, Talk given in the Dept. of E.E.C.S., University of California, Berkeley 1980.

Presented to the semester
Sequential Methods in Statistics
September 7–December 11, 1981