International Journal of Pure and Applied Mathematics
Volume 70 No. 6 2011, 837-841

ESTIMATION OF A PARAMETER OF A KNOWN LAW OF A PROBABILITY BY STOCHASTIC APPROXIMATIONS

Abdelkrim Bennar[1 §], Abdelhalim Bouamaine[2], Youssef Benghabrit[3]

[1]Faculté des Sciences Ben M'sik
Université Hassan II
Sidi Othmane, Casablanca, MOROCCO

[2]Ecole Nationale Supérieure d'éléctricité et de Mécanique
Laboratoire Architecture des Systèmes
Université Hassan II
Ain Chok, Casablanca, MOROCCO

[3]Ecole Nationale Supérieure des arts et Métiers
Université Moulay Ismail
Meknès, MAROCCO

Abstract: We consider a stochastic approximation process in a non-empty closed convex set K of $\mathbb{R}^k$: $X_{n+1} = \Pi\left(X_n - A_n(X_1, X_2, ..., X_n)\Psi_n(Y_n; X_n)\right)$, with for each n, $E[\Psi_n(Y_n; \theta_n)] = 0$, and Π is the projection operator on K.

We denote T_n the sub-σ-algebra generated by the events before time n.

We prove two theorems of almost sure convergence for the process (X_n) and we give two applications for estimation of a parameter of a known law of probability.

AMS Subject Classification: 62L20
Key Words: stochastic approximation, parameter estimation, Likehood maximum

1. Introduction

We define a stochastic approximation process (X_n) in a non-empty closed convex subset K of $\mathbb{R}^k$; we consider:

— observable random variables Y_n in $\mathbb{R}^p$ mutually independent; $\mathbb{R}^k$;

— measurable applications Ψ_n from $\mathbb{R}^p \times \mathbb{R}^k$ into $\mathbb{R}^l$;

Received: March 6, 2011

§Correspondence author

— for $n \geq 1$, a (k, l) random matrix A_n;

— the projection operator Π on K;

— the process (X_n) in K defined recursively by

$$X_{n+1} = \Pi\left(X_n - A_n \Psi_n(Y_n; X_n)\right),$$

— $\mathbb{R}^k$ is named parameter space, $\mathbb{R}^p$ is named space observations, and $\mathbb{R}^l$ is named space functions of estimate.

All random variables are defined on a probability space $(\Omega, \mathcal{A}, P)$. Denote T_n the sub-σ-algebra of $\mathcal{A}$ generated by the events before time n; $X_1, ...X_n$, $A_1, ..., A_n, Y_1, ..., Y_{n-1}$ are measurable with respect to T_n.

We give in Section 2 two almost sure convergence theorems of $\|X_n - \theta_n\|$ to 0. An application of each theorem is given in Section 3, concerning the estimation of a parameter of a known law of probability and the estimation of a parameter of gap under convex constraints.

In the following, $\langle \cdot, \cdot \rangle$ and $\|.\|$ are respectively the usual inner product and norm in $\mathbb{R}^k$; A' denotes the transposed matrix of A, $\lambda_{\min}(B)$ the smallest eigenvalue of B; the abbreviation *a.s.* means almost surely.

2. Lemmas

Let (X_n) be a stochastic process in a subset K of $\mathbb{R}^k$. Suppose:

(H1a) $\forall n \geq 1$, $\|A_n \Psi_n(Y_n; X_n\|^2$ is integrable a.s.

(H1b) $\sum_1^\infty E[\|A_n \Psi_n(Y_n; X_n\|^2 / T_n] < \infty$ a.s.

(H2) $\sum_1^\infty < X_n - \theta_n, E[A_n \Psi_n(Y_n; X_n] >^- < \infty$ a.s.

(H3) $\sum_1^\infty \|\theta_n - \theta_{n+1}\| < \infty$ a.s.

(H4) $E[\|X_1\|^2] < \infty$.

Lemma 1. *Assume H1a, b and H2, 3 and 4 hold; then:*

(1) There exist a random positive variable T *such that* $\|X_n - \theta_n\| \xrightarrow{a.s.} T$.

(2) $\sum_1^\infty < X_n - \theta_n, E[A_n \Psi_n(Y_n; X_n] >^+ < \infty$ *a.s.*

Proof. We have:

$$X_{n+1} - \theta_{n+1} = \Pi(X_n - A_n \Psi_n) - \Pi(\theta_{n+1}),$$

$$\|X_{n+1}-\theta_{n+1}\| \leq \|X_n - A_n\Psi_n - \theta_{n+1}\| \leq \|X_n - \theta_n - A_n\Psi_n + \theta_n - \theta_{n+1}\|,$$

$$\|X_{n+1}-\theta_{n+1}\|^2 \leq \|X_n-\theta_n\|^2 - 2 < X_n - \theta_n, A_n\Psi_n + \theta_{n+1} - \theta_n$$
$$> + \|A_n\Psi_n + \theta_{n+1} - \theta_n\|^2,$$

$$\|X_{n+1}-\theta_{n+1}\|^2 \leq \|X_n-\theta_n\|^2 - 2 < X_n-\theta_n, A_n\Psi_n > -2 < X_n-\theta_n, \theta_{n+1}-\theta_n$$
$$> +4\|A_n\Psi_n\|^2 + 4\|\theta_{n+1}-\theta_n\|^2,$$

$$E[\|X_{n+1}-\theta_{n+1}\|^2/T_n] \leq \|X_n-\theta_n\|^2 - 2 < X_n-\theta_n, E[A_n\Psi_n/T_n]^+$$
$$> +2 < X_n-\theta_n, E[A_n\Psi_n/T_n]^- > -2 < X_n-\theta_n, \theta_{n+1}-\theta_n$$
$$> +4E[\|A_n\Psi_n/T_n\|^2/T_n] + 4\|\theta_{n+1}-\theta_n\|^2.$$

As $| < a, b > | \leq \|a\|^2\|b\| + \|b\|$, we have:

$$E[\|X_{n+1}-\theta_{n+1}\|^2/T_n] \leq \|X_n-\theta_n\|^2[1+2\|\theta_{n+1}-\theta_n\|) - 2 < X_n-\theta_n, E[A_n\Psi_n/T_n]^+$$
$$> +2 < X_n-\theta_n, E[A_n\Psi_n/T_n]^- > +2\|\theta_{n+1}-\theta_n\|$$
$$+4E[\|A_n\Psi_n/T_n\|^2/T_n] + 4\|\theta_{n+1}-\theta_n\|^2.$$

We use in the proof the Robbins-Siegmund Lemma, see [4]:

Lemma 2. *Let $(\Omega, \mathcal{A}, P)$ be a probability space and (T_n) an increasing sequence of sub-σ-algebras of $\mathcal{A}$. For $n \geq 1$, let z_n, β_n, ξ_n and ζ_n be non-negative T_n-measurable random variables such that $E\left[z_{n+1} \mid T_n\right] \leq z_n(1+\beta_n) + \xi_n - \zeta_n$. Suppose $\sum_1^\infty \beta_n < \infty$, $\sum_1^\infty \xi_n < \infty$ a.s. Then $\lim_{n\longrightarrow\infty} z_n$ exists and is finite and $\sum_1^\infty \zeta_n < \infty$ a.s.*

Under hypothesisis H1a, b and H2, 3 and 4, we have:

(1) There exist a random positive variable T such that $\|X_n - \theta_n\| \xrightarrow{a.s} T$.

(2) $\sum_1^\infty < X_n - \theta_n, E[A_n\Psi_n(Y_n; X_n] >^+ < \infty$ a.s.

□

Suppose now:

(H5a) $\forall n \geq 1$, $\exists S_n : K^n \to \mathbb{R}^k$, S_n measurable:

$$E[A_n\Psi_n(Y_n; X_n)/T_n] = S_n(X_1, X_2, ..., X_n),$$

a.s.

(H5b) $\forall 0<\epsilon<1$, $\forall x_1, x_2, ..., x_{n-1}, ... \in K$

$$\sum_1^\infty \inf_{\left\{x:\ x\in K,\, \epsilon\|x-\theta_n\|<\frac{1}{\epsilon}\right\}} < x-\theta_n, S_n(x_1, x_2, ..., x_{n-1}, x) >^+ = +\infty.$$

(H6) $\|A_n\Psi_n(Y_n; X_n) - E[A_n\Psi_n(Y_n; X_n)/T_n]\| \xrightarrow{a.s} 0$

Lemma 3. *Assume H1a, b, H4, H6, then:* $\|X_{n+1} - X_n\| \xrightarrow{a.s} T$

Proof. $\omega \in \Omega$ is fixed throughout the proof, belonging to the intersection of the fowlling defined $a.s.$ convergence sets:

$C = \{\omega \in \Omega : \|X_n(\omega) - \theta_n\| \longrightarrow T(\omega)\}$

$C_1 = \{\omega \in \Omega : \sum_1^\infty < X_n(\omega) - \theta_n, S_n(X_1(\omega), X_2\omega), ..., X_n(\omega)) >^+ < \infty\}$.

Suppose $T(\omega) \neq 0$. Below ω is omitted.

Then: By H1a, there exist $0 < \epsilon_1 < 1$ and an integer $N(\omega)$ such that for $n \geq N(\omega), \epsilon_1 < \|X_n(\omega) - \theta_n\| < \frac{1}{\epsilon_1}$.

Hence $\sum_{N(\omega)}^\infty < X_n(\omega) - \theta_n, S_n(X_1(\omega), X_2(\omega), ..., X_{n-1}(\omega), X_n(\omega)) >^+ = +\infty$, a contradiction with $\omega \in C_1$, then $T(\omega) = 0$. □

3. Theorems of Almost Sure Convergence

Theorem 4. *Assume H1a, b and H2, 3 and 4, H5a, b hold; then:* $\|X_n - \theta_n\| \longrightarrow 0$ *a.s.*

Proof. It is a consequence of the lemmas 1 and 3 □

Suppose now:

(H7a) $\forall\ n \geq 1,\ \exists\ S_n : K \longrightarrow \mathbb{R}^k$,

$S_n measurable$;

$$E[A_n\Psi_n(Y_n; X_n)/T_n] = S_n(X_n) \text{ a.s.}$$

(H7b) $\forall 0<\epsilon<1,\ S(\epsilon) = sup_n sup_{\{x:\ x\in K, \epsilon<\|x-\theta\|<\frac{1}{\epsilon}\}} \|S_n(x)\| < \infty$;

(H7c) For all $\epsilon > 0$, there exists $\eta > 0$ such that $((x_1, x_2 \in K\ : \|x_1 - x_2\| < \eta) \Rightarrow (\sup_n \|S_n(x_1) - S_n(x_2)\| < \epsilon))$.

(H7d) There exist a positive integer q, a sequence of integers (n_l), for all $0 < \epsilon < 1$ an integer $L(\epsilon)$ such that $n_{l+1} \geq n_l + q$ and

$$b(\epsilon) = \inf_{l > L(\epsilon)} \inf_{\left\{x \in K,\, \epsilon < \|x-\theta\| < \frac{1}{\epsilon}\right\}} \sum_{j \in I_l} < x - \theta_j, S_j(x) >^+,$$

with $I_l = \{n_l, n_l + 1, ..., n_{l+1} - 1\}$; $\sum_1^\infty b_l(\epsilon) = +\infty$.

Theorem 5. *Assume H1*a, b, *H2, 3, 4, 6, H7*a, b, c, d *hold, then:* $\|X_n - \theta\| \longrightarrow 0$ *a.s.*

References

[1] A.E. Albert, L.A. Gardner, *Stochastic Approximation and Nonlinear Regression*, Research Monograph, **42**, The MIT Press, Cambridge, Massachusetts (1967).

[2] A. Bennar, *Approximation Stochastique, Convergence dans le Cas de Plusieurs Solutions et Étude de Modèles de Corrélation*, Thèse de Doctorat de Troisième Cycle, Université de Nancy 1 (1985).

[3] A. Bennar, *Sur L'analyse D'une Classe de Processus Stochastiques*, Thèse de Doctorat détat, Université Hassan II Mohammedia (2009).

[4] Y. Benghabrit, *Contribution à L'etude de Léstimation Séquentielle par Approximation Stochastique dún Paramètre Statistique*, Thèse de Doctorat de Troisième cycle, Université de Nancy 1 (1984).

[5] A. Bouamaine, *Méthodes D'approximation Stochastique en Analyse des Données*, Thèse de Doctorat détat, EMI, Université Mohammed V, Rabat (1996).

842