Chaotic Modeling and Simulation (CMSIM) 1: 57–67, 2017

In short about some stochastic techniques useful in the systems analysis

Gabriel V. Orman[1], Irinel Radomir[1], and Sorina-Mihaela Stoian[2]

[1] Department of Mathematics and Computer Science "Transilvania" University of Braşov, 500091 Braşov, Romania
(E-mail: ogabriel@unitbv.ro)
[2] "Excelsior" Excellency Centre, Braşov, Romania

Abstract. We shall refer shortly to some problems involving stochastic calculus in connection with the Brownian motion, a very useful model in the study of the systems and implicitly in the study of *the chaotic and complex systems*. Finally, an application based on the Brownian motion model is presented.
Keywords: stochastic differential equations, stochastic approximation, Markov processes, Brownian motion, convergence.

1 Introduction

Many practical problems conduct us to the following notion: *the equation obtained by allowing randomness in the coefficients of a differential equation is called a "stochastic differential equation"*.

Therefore, it is clear that any solution of a stochastic differential equation must involve some randomness. In other words one can hope to be able to say something about the probability distribution of the solutions.

At the same time, results on almost sure convergence of stochastic approximation processes are often proved by a separation of deterministic (pathwise) and stochastic considerations. The basic idea is to show that a "distance" between estimate and solution itself has the tendency to become smaller. The so-called *first Lyapunov method of investigation* does not use knowledge of a solution. Thus, in deterministic numerical analysis gradient of Newton procedures for minimizing or maximizing F by a recursive sequence (X_n) are investigated by a Taylor expansion of $F(X_{n+1})$ around X_n - a device which has been used in stochastic approximation for the first time by Blum, Kushner, Z. Schuss, M.T. Wasan, Nevel'son and Has'minskii. The second Lyapunov method in its deterministic and stochastic version uses knowledge of a solution.

On the other hand, the Brownian motion was firstly investigated by L. Bachelier and A. Einstein, and then N. Wiener had the possiblity to put it on a firm mathematical foundation. Then, many of the scientific works has been done on their applicatios in physics, chemistry, communications, population

Received: 19 March 2016 / Accepted: 15 December 2016

ISSN 2241-0503

genetics, and other fields. In this way the Markov processes and many related topics become of a special interest.

In fact, the construction of the Brownian motion as a limit of a rescaled random walk can be generalized to a class of Markov chain. In such a context one can refer to Markov processes from a pespective of K Itô.

Some aspects will be discussed below.

2 A problem of stochastic approximation

The basic Stochastic Approximation Algorithms introduced by H. Robbins & S. Monro and by J. Kiefer & J. Wolfowitz have been the subject of an enormous literature.

This is due to the large number of applications and the interesting theoretical issues in the analysis of *dynamically defined stochastic processes*.

In recent years, algorithms of the stochastic approximation type have found applications in new and diverse areas, and new techniques have been developed for proofs of convergence and rate of convergence. The actual and potential applications in signal processing have exploded. Indeed, whether or not they are called stochastic approximations, such algorithms occur frequently in practical systems for the purposes of noise or interface cancellation, the optimization of *post processing* or *equalization* filters in time varying communication channels, adaptive antenna systems, and many related applications.

In such applications, the underlying processes are often nonstationary, the optimal value of the parameter of the system (say, for example θ) changes with time, and we keeps the *step size* (say for example ε_n) strictly away from zero in order to allow *tracking*. Such tracking applications lead to new problems in the asymptotic analysis (e.g. when ε_n are adjusted adaptively); one wishes to estimate the tracking errors and their dependence on the structure of the algorithm.

We shall refer to a problem discussed by J.H. Venter and J.L. Gastwirth and also studied by M.T. Wasan.

For an unknown parameter $\lambda > 0$ let us consider a distribution function defined as follows

$$\begin{aligned} F(x) &= 1 - e^{-\lambda x} \quad \text{for} \quad x \geq 0 \\ F(x) &= 0 \qquad\qquad\ \ \text{for} \quad x < 0. \end{aligned} \tag{1}$$

$F(x)$ can be the distribution function of a *system* or *item*, with a *life time* for which inspections are made at time $t_1, t_2, t_3, \cdots$. If the conclusion of the inspections is that the system is inoperative, then it will be repared or replaced. In any other case nothing is done. Thus, the problem is to choose the inspection plan, that is to choose the sequence $t_1, t_2, t_3, \cdots$ in an optimal way in a suitable sense. In this way a problem of stochastic approximation, discussed by J.H. Venter & J.L. Gastwirth and M.T. Wasan, is obtained.

Let $0 < a < b$ two constants and it is supposed that $a < \lambda < b$. The inspection times are defined as follows

$$T_1 = t_1, \quad T_i = t_i - t_{i-1}, \quad i = 2, 3, 4, \cdots. \tag{2}$$

Now let us consider an arbitrary sequence of random variables $\{X_n\}$ for which the joint distribution of any finite number does not depend on λ. Denote $T_1 = \max\{0, X_1\}$ and one defines $\{T_n\}$ as follows

$$T_{n+1} = \max\{0, f_n(Y_1, \cdots, Y_n) + X_n\} \quad \text{for} \quad n = 1, 2, 3, \cdots \tag{3}$$

where f_n is a real-valued measurable function of $(Y_1, \cdots, Y_n)$, functionally independent of λ. And Y_i, $i = 1, 2, 3, \cdots$, are random variables with conditional distribution

$$Y_i : \begin{pmatrix} 0 & 1 \\ 1 - e^{-\lambda T_i} & e^{-\lambda T_i} \end{pmatrix} \tag{4}$$

given $\{Y_1, \cdots, Y_{i-1}, T_1, \cdots, T_i\}$. Hence one has $Y_i = 1$ if the ith inspection conducts to the conclusion that the system is operative and $Y_i = 0$ if it is inoperative.

We can understand that after n inspections, the next inspection time T_{n+1} depends on the past observations $(Y_1, \cdots, Y_n)$ through f_n while X_n allows for additional randomization.

Let us denote by $\mathcal{G}$ the class of all inspection plans and let I be a generic element of $\mathcal{G}$.

Now, a criterion of optimality can be obtained (following the plan of M.T. Wasan).

Let $J_n(I, \lambda)$ be defined by the equality

$$J_n(I, \lambda) = \frac{E\left(\dfrac{d}{d\lambda} \log L_n(\lambda)\right)^2}{n} \tag{5}$$

the average information obtainable from a plan I after n inspections, where $L_n(\lambda)$ is the probability function of λ based on $(Y_1, Y_2, \cdots, Y_n, T_1, T_2, \cdots, T_n)$. At the same time, the limiting average information $J(I, \lambda)$, obtainable from plan I, is defined by the following equality

$$J(I, \lambda) = \lim_{n \to \infty} \inf J_n(I, \lambda). \tag{6}$$

Let $J_n(I, \lambda)$ be defined by the equality

$$J_n(I, \lambda) = \frac{E\left(\dfrac{d}{d\lambda} \log L_n(\lambda)\right)^2}{n} \tag{7}$$

the average information obtainable from a plan I after n inspections, where $L_n(\lambda)$ is the probability function of λ based on $(Y_1, Y_2, \cdots, Y_n, T_1, T_2, \cdots, T_n)$. At the same time, the limiting average information $J(I, \lambda)$, obtainable from plan I, is defined by the following equality

$$J(I, \lambda) = \lim_{n \to \infty} \inf J_n(I, \lambda). \tag{8}$$

Now, the following question arises: *how one can maximizing* $J_n(I, \lambda)$ *and* $J(I, \lambda)$ *by a judicious choice of* I ? Hence we attain to a well known method of efficient estimation of λ, which conduct us to the theorem below

Theorem 1.

$$J_n(I,\lambda) \leq \frac{T_\lambda(2-\lambda T_\lambda)}{\lambda} \tag{9}$$

for each n and for all λ and I, where T_λ is the solution of the equation

$$\lambda T = 2(1-e^{-\lambda T}). \tag{10}$$

[$T_\lambda = -\frac{\log p}{\lambda}$ and T_λ is the $100(1-p)th$ percentile of exponential distribution, where $p \approx 0.203$].

Let us observe that the equality in (9) follows if and only if $T_i = T_\lambda$ with probability one for each i. That is to say if λ were known, the optional inspection plan in the sense of maximizing $J_n(I,\lambda)$ for each n and λ would call for periodic inspection with interinspection times T_λ. But within the class $\mathcal{G}$ (what means that λ is unknown) there is no optimal plan for which the equality is obtained in (9).

Now an optimality criterion can be defined by using the concept of *adaptive inspection plan*. But this will be discussed on another occasion.

Other problems, proofs and related topica can be found in [1], [2], [12], [7], [23], [19].

3 Markov processes in the generalized sense and Brownian motion

The Brownian motion can be represented as a random sum of integrals of orthogonal functions. Such a representation satisfies the theoretician's need to prove the existence of a process with the four defining properties of Brownian motion, but it also serves more concrete demands, one of the most important being the "chaotic and complex systems analysis".

Especially, the series representation can be used to derive almost all of the most important analytical properties of Brownian motion.

It can also give a powerful numerical method for generating the Brownian motion paths that are required in computer simulation.

Starting from the observation that many a time we refer to *chaos* and *chaotic and complex systems* to describe the comportment of some natural phenomena, it is very useful, from a mathematical point of view, to talk about a passing *from chaotic and complex systems to Brownian motion*. In this way we can refer to the *Brownian motion* which is a more realistic model of such phenomena.

Brownian motion was frequently explained as due to the fact that particles were alive. An origin of mathematical Brownian motion is a game theoretic model for fluctuations of stock prices due to L. Bachelier from 1900. In his doctoral thesis, *Théorie de la spéculation*, Ann. Sci. École Norm. Sup., 17, 1900, 21-86, he hinted that it could apply to physical Brownian motion. Therein, and in his subsequent works, he used the heat equation and, proceeding by analogy with *heat propagation* he found, albeit formally, distributions of various functionals of mathematical Brownian motion. Heat equations and related parabolic type equations were used rigorously by Kolmogorov, Petrovsky, Khintchine.

It was frequently explained as due to the fact that particles were alive. It is only in 1905 that kinetic molecular theory led Einstein to the first mathematical model of Brownian motion. He began by deriving its possible existence and then only learned that it had been observed.

Now we shall refer shortly to some results due to K. Itô relating to the *3-dimensional Brownian motion* and the *k-dimensional Brownian motion.*

Let now consider $C = C(S)$ to be the space of all continuous functions (it is a separable Banach space with the supremum norm). The *transition operators* can be defined in a similar manner:

Definition 1. The operators p_t, defined by

$$(p_t f)(x) = \int_S p_t(x, dy) f(y), \quad f \in C$$

are called "transition operators".

And the conditions for the transition probabilities can be adapted to the transition operators.

On the other hand, a *Markov process* is usually defined as a system of stochastic processes

$$\{X_t(\omega), t \in T, \omega \in (\Omega, K, P_a)\}_{a \in S},$$

that is for each $a \in S$, $\{X_t\}_{t \in S}$ is a stochastic process defined on the probability space (Ω, K, P_a).

The transition probabilities of a Markov process will be denoted by $\{p(t, a, B)\}$; also let us denote by $\{H_t\}$ the transition semigroup and by R_α the resolvent operator of $\{H_t\}$.

The next results shows that $p(t, a, B)$, H_t and R_α can be expressed in terms of the process as follows:

Theorem 2. *Let f be a function in $C(S)$. Then*

1. $p(t, a, B) = P_a(X_t \in B)$.
2. *For* $E_a(\cdot) = \int_\Omega \cdot P_a(d\omega)$ *one has* $H_t f(a) = E_a(f(X_t))$.
3. $R_\alpha f(a) = E_a\left(\int_0^\infty e^{-\alpha t} f(X_t) dt\right)$.

Now regarding to the *Markov process* one can observe that a definition as it is given above not correspond to many processes that are of a real interest. For this reason it is useful to obtain an extension of this notion (as it was proposed by K. Itô).

Let E be a separable Banach space with real coefficients and norm $||\cdot||$ and let also $L(E, E)$ be the space of all bounded linear operators $E \longrightarrow E$. It can be observed that $L(E, E)$ is a linear space.

Definition 2. The collection of stochastic processes

$$X = \{X_t(\omega) \equiv \omega(t) \in S,\, t \in T,\, \omega \in (\Omega, K, P_a)\}_{a \in S}$$

is called a "Markov process" if the following conditions are satisfied:

1) the "state space" S is a complete separable metric space and $K(S)$ is a topological σ-algebra on S;
2) the "time internal" $T = [0, \infty)$;
3) the "space of paths" Ω is the space of all right continuous functions $T \longrightarrow S$ and K is the σ-algebra $K[X_t : t \in T]$ on Ω;
4) the probability law of the path starting at a, $P_a(H)$, is a probability measure on (Ω, K) for every $a \in S$ which satisfy the following conditions:
4a) $P_a(H)$ is $K(S)$-measurable in a for every $H \in K$;
4b) $P_a(X_0 = a) = 1$;
4c) $P_a(X_{t_1} \in E_1, \cdots, X_{t_n} \in E_n) =$
$$\int \dots \int_{a_i \in E_i} P_a(X_{t_1} \in da_1) P_{a_1}(X_{t_2 - t_1} \in da_2) \dots$$
$$\dots P_{a_{n-1}}(X_{t_n - t_{n-1}} \in da_n) \qquad \text{for} \quad 0 < t_1 < t_2 < \dots < t_n.$$

According to Definition 2, X will be referred as a *Markov process in the generalized sense.*

Now let X be a Markov process in a generalized sense and let us consider a function $f \in B(S)$ where $B(S)$ denote the space of all bounded real $K(S)$-measurable functions.

Now let us suppose that

$$E_a\left(\int_0^\infty |f(X_t)|dt\right) \tag{11}$$

is bounded in a such that

$$Uf(a) = E_a\left(\int_0^\infty f(X_t)dt\right) \tag{12}$$

is well-defined and is a bounded $K(S)$-measurable function of $a \in S$.

The Uf is called *the potential* of f with respect to X and having in view that $Uf = \lim_{\alpha \downarrow 0} R_\alpha f$, it is reasonable to write R_0 instead of U. Based on this fact, $R_\alpha f$ will be called *the potential of order* α of f.

Now the name *potential* is justified by the following theorem on *the 3-dimensional Brownian motion*

Theorem 3. *Let X be the 3-dimensional Brownian motion. If $f \in \mathbf{B}(S)$ has compact support, then f satisfies (11) and*

$$Uf(a) = \frac{1}{2\pi} \int_{R^3} \frac{f(b)db}{|b - a|} = \frac{1}{2\pi} \times Newtonian\ potential\ of\ t. \tag{13}$$

Now let X be a k-dimensional Brownian motion defined as follows

Definition 3. The *k-dimensional Brownian motion* is defined on $S = R^k$ by the equality

$$p_t(a, db) = (2\pi t)^{-\frac{k}{2}} e^{-\frac{|b-a|^2}{2t}} db = N_t(b - a)db,$$

where $|b - a|$ is the norm of $b - a$ in R^k.

Now let us define

$$\mathcal{N}_{\mathbf{B}} = R_\alpha \mathbf{B}(S)$$

and

$$\mathcal{N} = \{f \in \mathbf{B}(\mathbf{S}) : R_\alpha f = 0\}.$$

The $\mathcal{N}_{\mathbf{B}}$ and $\mathcal{N}$ are both independent of α. Therefore we can regard R_α as a one-to-one map from the quotient space $\mathbf{B}(\mathbf{S})/\mathcal{N}$ onto $\mathcal{N}_{\mathbf{B}}$.

By using the *k-dimensional Brownian motion* the following result is obtained

Theorem 4. *Let X be a k-dimensional Brownian motion. Then*

$$\mathcal{N} = \{f \in \mathbf{B}(R^k) : f = 0 \ a.e.\},$$

$$\mathcal{D}(A) = \{u \in \mathbf{B}(R^k) : \Delta u \ is \ a \ measurable \ function \ bounded \ a.e.\},$$

and

$$Au = \frac{1}{2}\Delta u,$$

where the Laplacian Δ is to be understood in the Schwarts distibution sense.

For more details and some related topics see ([8], [9], [10]).

4 Application

It is known that the construction of the Brownian motion as a limit of a rescaled random walk can be generalized to a class of Markov chains. In such a context we shall consider an application to genetics studied by W. Feller ([4]). Such a problem was firstly discussed in detail by S. Wright and R. A. Fisher; and its Markovian nature was pointed out by G. Malécot[1]. The topic has been also considered by H.J. Kushner ([11]); and then has been developed by Z. Schuss as an example in connection with a study concerning the problem of diffusion approximation to Markov chains ([21]). It was also presented at lenght in [18].

It is known that the heritable characters depend on special carries, called *genes,* which appear in pairs. Each gene of a particular pair can suggest two forms A and a which determine a genetic type in a population. Therefore, three different pairs can be formed AA, Aa, aa, such that the organism belongs to one of these three *genotypes*. On the other hand, the reproductive celles, called *gametes*, are assumed to have one gene. In this way, the gametes of an organism of genotype AA or aa have the gene A or respective the gene a, whereas the gametes of an organism of genotype Aa may have the gene A or the gene a with equal probability. We can view such a problem in the context of the *binomial distribution*

$$P_n(k) = C_n^k p^k q^{n-k}.$$

We can imagine an urn with $2N$ elements (that is the genes of types A and a). Then, the genotype structure of N offsprings will be the result of

[1]G. Malécot: *Sur un problèm de probabilités en chaine que pose la génétique*, Comptes Rendus de l'Académie des Sciences, vol.219, 1944, pp.379-381.

$2N$ independent drawings from the urn. Furthermore, it is considered that a population consists of N individuals in each generation. Now, if i of the genes are of the type A $(0 \leq i \leq 2N)$ in some generation, then it is said that the generation is in the state E_i. In this way a Markov chain is connected to such a genetic process. It has $2N+1$ states $E_0, E_1, E_2, \cdots, E_{2N}$. Calculating the transition probability from the state E_i to E_j, in one generation, one gets

$$P_{ij}^N = C_{2N}^j \left(\frac{i}{2N}\right)^j \left(1 - \frac{i}{2N}\right)^{2N-j}. \tag{14}$$

Now, one considers a population of N individuals consisting of $X_N(n) = i$ individuals of type A in the nth generation. Then, the next generation consists of N individuals randomly selected from a practically infinite offspring of the previous generation. Obviously the selection process is binomial with probability $x = \frac{i}{N}$ for type A, the proportion $\frac{i}{N}$ of A types being equal to the probability in a large offspring population. Therefore, the transition probability is given by the following equality

$$P_{ij}^N = P(X_N(n+1) = j \,|\, X_N(n) = i) = C_N^j \, x^j (1-x)^{N-j}.$$

Let now denote by s the $fitness$ of A relative to a when selection forces act on the population

$$s = \frac{\overline{x} - x}{x(1 - \overline{x})}$$

where

$$\overline{x} = \frac{x(1+s)}{1+sx}.$$

Therefore, one obtains

$$P_{ij}^N = C_N^j \, \overline{x}^j (1 - \overline{x})^{N-j}.$$

Now, if $s = s_N(n)$ is a random variable then, the *probability of extinction* of a genotype, or the *time until extinction*, or the *total A population*, or other characteristics of interest, become very hard to calculate. For this reason the Markov chains $\{X_N(n)\}$ can be approximated by a diffusion process, or more exactly, by a solution of a stochastic differential equation.

To this end the following form is considered for the process

$$X_N([Nt]) = N \, x_N(t)$$

where $[Nt]$ is the greatest integer not exceeding Nt and t is any positive number. Thus, $x_N(t)$ represents the *proportion* of A types in population. Now it is supposed that

$$N \, E s_N(n) \to \sigma(t)$$

as $N \to \infty$, $n = [Nt]$, and $NEs_N^2(n) \to \nu(t)$ and $NEs_N^k(n) \to 0$ for all $k > 2$.

Let now be $\Delta t = \frac{1}{N}$. Then, one gets

$$a_N \equiv \frac{1}{\Delta t} E_{x,t} \left[x_N(t + \Delta t) - x_N(t) \right] = \\ = E[X_N([Nt] + 1) - X_N([Nt]) \,|\, X_N([Nt]) = x_N].$$

But $X_N(n)$ is a binomial variable $B(N, \overline{x})$ such that it results

$$a_N = NE(\overline{x} - x) = NE\left(\frac{(1 + s_N(n))x}{1 + s_N(n)x} - x \right) \to \\ [\sigma(t) - \nu(t)x]x(1 - x) \equiv a(x, t) \tag{15}$$

and respective

$$\tfrac{1}{\Delta t} E_{x,t}[x_N(t + \Delta t) - x_N(t)]^2 \to x(1 - x)[1 + \nu(t)x(1 - x)] \equiv b(x, t). \tag{16}$$

The moments of higher order tend to zero as $N \to \infty$. Now it can be shown that the convergence is sufficiently rapid as to satisfy the imposed conditions. Let now consider again the specified model. Thus, we have

$$E[x_N(t + \Delta t) - x_N(t) \,|\, x_N(t)] = 0$$

and

$$E[|x_N(t + \Delta t) - x_N(t)|^2 \,|\, x_N(t)] = \frac{x_N(t)[1 - x_N(t)]}{N}$$
$$E[|x_N(t + \Delta t) - x_N(t)\,|^4 \, x_N(t)] \le \frac{K}{N^2}$$

for a constant K.

Now, it can be seen that the conditions for convergence hold, provided that the stochastic differential equation

$$dx(t) = a(x, t)dt + \sqrt{b(x, t)}\, dw(t)$$
$$x(0) = x \tag{17}$$

has a unique solution (with absorption at $x = 0$ and $x = 1$). But the conditions of the existence and uniqueness theorem are not satisfied by the coefficient $\sqrt{b(x, t)}$ in (17). To show the existence and uniqueness it is necessary to consider (17) in the interval $I_\varepsilon[\varepsilon, 1 - \varepsilon]$ with absorption at the boundary δI of I_ε. The conditions of the existence and uniqueness theorem are satisfied in I_ε. Thus, it is a unique solution $x_\varepsilon(t)$ in I_ε up to the time

$$\tau_\varepsilon = \inf\{t \,|\, x_\varepsilon(t) \in \delta I_\varepsilon\}.$$

If $\varepsilon_1 < \varepsilon_2$ then, $\tau_{\varepsilon_1} \ge \tau_{\varepsilon_2}$ and

$$x_{\varepsilon_1}(t) = x_{\varepsilon_2}(t) \quad \text{where} \quad 0 \le t \le \tau_{\varepsilon_2},$$

by the so-called the *localization principle*.[2] Now, as $\varepsilon \to 0$, one gets that $x_{\varepsilon(t)}$ converges to a limit $x(t)$ and $\tau_\varepsilon \to \tau$ (here τ is the absorption of $x(t)$).

In this way it results that $x_N(t) \to x(t)$, with $x(t)$ the solution of the stochastic differential equation

$$dx(t) = a(x,t)dt + \sqrt{b(x,t)}\, dw(t)$$

with absorbing boundaries at $x = 0$ and $x = 1$.

Thus, once a genotype is extinct, it will stay extinct for all future generations unless mutation occurs. Therefore, the pobability of extinction is the probability of exit of $x(t)$ from the interval (0, 1).

4.1 Conclusion

Obviously, verious situation may exist when the survival of a particular genotype can be very dynamic.

In general, the interaction of a population can have a great complexity, which lead to the enhancement of the interdisciplinary coordination in these studies.

But, as we have already emphasized, for a random variable $f = f_N(n)$ the "probability of extinction" of a genotype, or the "time until extinction", or the "total A population", or other characteristics of interest, become very hard to calculate. And this is the reason for which a Markov chain $\{X_N(n)\}$ is useful to be approximated by a solution of a stochastic differential equation.

References

1. Comer, J.P.Jr., *Application of stochastic approximation to process control.* J. Roy. Stat. Soc., Series B, 27, 1965, pp.321-331.
2. Fabian, V., On asymptotic normality in stochastic approximation. Michigan State University, Statistical Laboratory Publications, 1967.
3. Feller, W., An Introduction to Probability Theory and its Applications, vol. I. John Wiley & Sons, Inc., New York, London, 1960.
4. Feller, W., An Introduction to Probability Theory and its Applications, vol. II. John Wiley & Sons, Inc., New York, London, 1966.
5. Gihman, I.I. and Skorohod, A.V., Stochastic Differential Equations. Springer-Verlag, Berlin, 1972.
6. Gnedenko, B.V., The Theory of Probability. Mir Publisher, Moscow, 1976.

[2]The *Localization principle*. Let us consider $a_1(x,t) = a_2(x,t)$ and $b_1(x,t) = b_2(x,t)$ for $c \le x \le d$ and $t \ge 0$. Suppose that $a_i(x,t)$ and $b_i(x,t)$, with $i = 1, 2$, satisfy the conditions of the existence and uniqueness theorem in R for the stochastic differential equations

$$dx_i(t) = a_i(x,t)dt + b_i(x,t)dw(t), \qquad i = 1,2$$
$$x(0) = x_0 \in [c,d].$$

Let us denote $\tau_i = \inf\{t \,|\, x_i(t) \notin (c,d)\}$ with $i = 1, 2$. Then, $\tau_1 = \tau_2\, a.s.$ and $x_1(t) = x_2(t)$ for all $t \le \tau_1\, a.s.$ [For more details see [5]].

7. Gray, K.B., *The application of stochastic approximation to the optimization of random circuits*. American Mathematical Society, Proceedings of Simposium in Applied Mathematics, 16, 1962, pp. 178-192.9++
8. Itô, K. Selected Papers. Springer-Verlag, Berlin Heidelberg, 1987.
9. Itô, K. and McKean, H.P. Jr., Diffusion Processes and their Sample Paths. Springer-Verlag, Berlin Heidelberg, 1996.
10. Itô, K., Stochastic Processes. Eds. Ole E. Barndorff-Nielsen and Ken-iti Sato. Springer, 2004.
11. Kushner, H.J. *On the weak convergence of interpolated Markov chains to a diffusion*. Ann. Probab., 2, 1974, 40-50.
12. Kushner, H.J. and Yin, G.G. Stochastic Approximation Algorithms and Applications. Springer-Verlag New York, Inc., 1997.
13. Loève, M., Probability theory I. Springer-Verlag, New York, Heidelberg Berlin, 1977.
14. Loève, M., Probability theory II. Springer-Verlag, New York, Heidelberg Berlin, 1978.
15. Øksendal, B., Stochastic Differential Equations: An Introduction with Applications. Sixth Edition. Springer-Verlag, 2003.
16. Orman, G.V., Lectures on *Stochastic Approximation Methods and Related Topics*. Preprint. "Gerhard Mercator" University, Duisburg, Germany, 2001.
17. Orman, G.V., Handbook of Limit Theorems and Stochastic Approximation. "Transilvania" University Press, Brasov, 2003.
18. Orman, G.V., Capitole de matematici aplicate. Ed. Albastră, MicroInformatica, Cluj-Napoca, 1999.
19. Orman, G.V., Aspects of convergence and approximation in random systems analysis. Lambert Academic Publishing, Germany, 2012.
20. Orman, G.V., Basic Probability Theory, Covergence, Stochastic Procesees And Applications. *Concepts and Stochastic Models useful in the Study of the Systems*. Transilvania University Press, Braşov, 2016.
21. Schuss Z., Theory and Application of Stochastic Differential Equations. John Wiley & Sons, New York, 1980.
22. Stroock D. W., Markov Processes from K. Itô Perspective. Princeton Univ. Press, Princeton, 2003.
23. Wasan, M.T., Stochastic Approximation. Cambridge University Press, 1969.