Chaotic Modeling and Simulation (CMSIM) 4: 711–718, 2012

New aspects in approximation of a Markov chain by a solution of a stochastic differential equation

Gabriel V. Orman and Irinel Radomir

Transilvania University of Brasov, 500091 Braşov, Romania.
E-mail: ogabriel@unitbv.ro

Abstract. The usual class of Markov processes which we involve many times has some restrictions that it does not cover many interesting processes. We shall refer, in this paper, to some problems involving stochastic calculus, diffusion approximation and Markov processes. In this context the problem of absorbing and reflecting barriers is also discussed.
Keywords: stochastic differential equations, Markov chains, transition probabilities, Brownian motion.

1 Introduction

When a stochastic differential equation is considered if it is allowed for some randomness in some of its coefficients, it will be often obtained a so-called *stochastic differential equation* which is a more realistic mathematical model of the considered situation.

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.

In the sequel we shall refer to some aspects relating to the approximation in the study of Markov processes and Brownian motion. Such problems were developed particularly by Schuss[13], Kushner and Yin[5], Itô and McKean Jr.[3], Wasan[14].

Results on almost sure convergence of stochastic approximation processes are often proved by a separation of deterministic (pathwise) and stochastic considerations. A key problem in effective applications concerns the *amount of noise* in the observations, and this leads to variations that incorporate variance reduction methods. With the use of these methods, the algorithm becomes

Received: 2 April 2012 / Accepted: 16 October 2012

ISSN 2241-0503

more effective, but also more complex. Hence, it is desirable to have robust algorithms, which are not overly sensitive to unusually large noise values.

More details and related topics can be found in Schuss[13], Kushner and Yin[5], Itô and McKean Jr.[3], Itô[4], Øksendal[6], Øksendal and Sulem[7], Stroock[12], Orman[9], [10], Wasan[14] which are also the basis in our development.

2 Preliminaries

Definition 1. The *sample space* Ω of a random experiment is the collection of all possible outcomes. An *event* A is a subset of the sample space, that is, a set of outcomes.

Definition 2. A *probability measure* on a sample space Ω of a random experiment is a function $\mathrm{P}[\cdot]$ that maps events in Ω to real numbers such that: (i) $\mathrm{P}[\Omega] = 1$,

(ii) $\mathrm{P}[\mathrm{A}] \geq 0$ for all events A,

(iii) $P\left[\bigcup_{i\in I} A_i\right] = \sum_{i\in I} P[A_i]$ where I is a finite or countable infinite set of integers and any pair of the events $A_1, A_2, A_3, \cdots$ is disjoint.

Let us consider the triplet $(\Omega, \mathcal{K}, P)$ where

- Ω is the *sample space*. Its elements are referred to as *sample points*;
- $\mathcal{K}$ is a σ-field of subsets of Ω containing Ω itself. Its elements are *events*;
- P is a *probability measure* on the measurable space $(\Omega, \mathcal{K})$.

If an event A is of the type $A = \{\omega \in \Omega \,|\, R(\omega)\}$ for some property $R(\cdot)$, (of the probability) we may write $P(R)$ for $P(A)$. An event is called a *sure event* if $P(A) = 1$ and a *null event* if $P(A) = 0$. Alternatively, $R(\cdot)$ is said to hold *a.s.* if $P(R) = 1$.

The triplet $(\Omega, \mathcal{K}, P)$ is referred to as a *probability space*.

Let now consider an experiment that is repeated n times and suppose that m $(m \leq n)$ times the event B occurred. Also suppose that k times $(k \leq m)$ the event A occurred, provided that B occurred. Then, the event $A \cap B$ occurred k $(k \leq n)$ times, such that we have $P(A \cap B) = \frac{k}{n}$. Now $\frac{k}{n} = \frac{m}{n} \cdot \frac{k}{m} = P(B) \cdot P(A \,|\, B)$. In this way the following relation is obtained $P(A \cap B) = P(B) \cdot P(A \,|\, B)$ or

$$P(A \,|\, B) = \frac{P(A \cap B)}{P(B)}, \quad P(B) > 0. \tag{1}$$

$P(A \,|\, B)$ in (1) is called a *conditional probability* whenever $P(B) > 0$. We retain that the function

$$P_B(A) = P(A \,|\, B) \tag{2}$$

is a probability measure in B, where B is now considered as a smaller sample space.

Hence, the measure $P(A \mid B)$ is the probability of the event A provided that B occurs.

A fundamental concept of probability is the notion of random variable. A *random variable* is a function that gives a numerical value to each outcome of a random experiment. The distinction between the random variable and the experimental outcome itself can become blurry in this case, because if ω denotes an outcome, then $X(\omega) = \omega$. But we retain that the random variable is a function and the outcome is its input. The domain of the random variable is the sample space and that its range is a set of numbers.

Definition 3. Let $(\Omega, \mathcal{K}, P)$ be a probability space and let us denote by E a subset of R^n. A *random variable* X is a function from Ω into E. We refer to E as being the *state space* of the random variable.

So, a random variable encodes an experimental outcome as a number, or a vector of real numbers in the multidimensional case. When a random variable has a multidimensional state space, we emphasize that fact by calling it a *random space*.

Let (E, ξ) be a measurable space and $X : (\Omega, \mathcal{K}, P) \to (E, \xi)$ a random variable (i.e. a measurable map). The image μ of P under X is a probability measure on (E, ξ), called the *law of* X and denoted by $\mathcal{L}(X)$. The events $\{\omega \mid X(\omega) \in A\}$ for $A \in \xi$ form a sub-σ-field of $\mathcal{K}$ called the σ-field generated by X and denoted by $\sigma(X)$. More gneral, given a family X_α, $\alpha \in I$, of random variables on $(\Omega, \mathcal{K}, P)$ taking values in measurable spaces (E_α, ξ_α), $\alpha \in I$, respectively, the σ-field generated by X_α, $\alpha \in I$, denoted by $\sigma(X_\alpha, \alpha \in I)$, is the smallest sub-$\sigma$-field with respect to which they are all measurable. They may be situations where it is preferable to view $\{X_\alpha, \alpha \in I\}$ as a single random variable taking values in the product space $\prod E_\alpha$ endowed with the product σ-field $\prod \xi_\alpha$. If so, this definition reduces to the preceding one.

Two (or more) random variables are said *to agree in law* if their laws coincides. They could be defined on different probability spaces. A random variable $X(\omega)$ generates a field (σ-field) $\mathcal{K}_X$ of events generated by events of the form $\{\omega \mid X(\omega) = a\}$ where a is any number. The field consists of events which are unions of events of the form $\{\omega \mid X(\omega) = a\}$. The probability function P on the events of this field $\mathcal{K}_X$ generated by $X(\omega)$ is called the *probability distribution of* $X(\omega)$.

Suppose we have n random variable $X_1(\omega), \cdots, X_n(\omega)$ defined on a probability space. The random variables $X_1, \cdots, X_n$ *are said to be independent* if the fields (σ-fields) $\mathcal{K}_{X_1}, \cdots, \mathcal{K}_{X_n}$ generated by them are independent.

Definition 4. A *stochastic process* is a parametrized collecion of random variables

$$\{X_t\}_{t \in T}$$

defined on a probability space $(\Omega, \mathcal{K}, P)$ and assuming values in R^n.

The parameter space T may be the halfline $[0, +\infty)$, or it may also be an interval $[a, b]$, or the non-negative integers and even subsets of R^n, for $n \geq 1$.

Now, for each $t \in T$ fixed, we have a random variable $\omega \to X_t(\omega),\ \omega \in \Omega$. On the other hand, fixing $\omega \in \Omega$, we can consider the function

$$t \to X_t(\omega), \quad t \in T \tag{3}$$

which is called a *path* of the random variable X_t. It is useful to think of t as *time* and each ω as an individual *particle* or *experiment*. Thus, $X_t(\omega)$ would represent the position (or the result) at time t of the particle (experiment) ω. In some cases it is convenient to write $X(t,\omega)$ instead of $X_t(\omega)$, such that the process can be regarded as a function of two variables $(t,\omega) \to \psi(t,\omega)$ from $T \times \Omega$ into R^n. In stochastic analysis this is often a natural point of view, because there it is crucial to have $X(t,\omega)$ jointly measurable in (t,ω).

In this paper we shall denote a stochastic process by $X(t)$.

3 Markov process and diffusion process

Definition 5. A stochastic process $X(t)$ on $[0,T]$ is called a *Markov process* if for $n = 1,2,3,\cdots$ and any sequences $0 \le t_0 < t_1 < \cdots < t_n \le T$ and $x_0, x_1, \cdots, x_n$, the following equality is satisfied:

$$\begin{aligned} P(X(t_n) < x_n \,|\, X(t_{n-1}) = x_{n-1}, X(t_{n-2}) = x_{n-2}, \cdots, X(t_0) = x_0) = \\ = P(X(t_n) < x_n \,|\, X(t_{n-1}) = x_{n-1}). \end{aligned} \tag{4}$$

The equation (4) means the fact that the process $forget$ the past, provided that t_{n-1} is regarded as the present.

Let Ω_X be the state space of the random variables X_t. Take $\mathcal{K}_X$ as the σ-field of measurable subsets of Ω_X. For convenience, assume that there is a first point to the set T. The probability structure is specified in terms of an initial probability measure and a transition probability function describing how transitions take place from one time to another.

We denote by $P(t_0, A)$ a probability measure on the sets A of $\mathcal{K}_X$. This is the probability distribution at the initial time t_0. Further let the transition probability function $p(t,x;\tau,A),\ t_0 \le \tau < t,\ x \in \Omega_X,\ A \in \mathcal{K}_X$ be a function with the following properties:

i $p(t,x;\tau,A)$ is a probability measure in $A \in \mathcal{K}_X$ for fixed t,x,τ;
ii $p(t,x;\tau,A)$ is measurable in x with respect to $\mathcal{K}_X$ for fixed t,τ,A;
iii $p(t,x;\tau,A)$ satisfies the integral equation (commonly called the *Chapman-Kolmogorov equation*)

$$p(t,x;\tau,A) = \int_{\Omega_X} p(s,y;\tau,A)p(t,x;s,dy) \tag{5}$$

for any s with $t < s < \tau$.

As it is shown in the theory of stochastic processes, the transition probability function $p(t,x;\tau,A)$ is the conditional probability

$$p(t,x;\tau,A) = P[X_\tau(\omega) \in A \,|\, X_t(\omega) = x]. \tag{6}$$

Now the *transition distribution function*

$$F(t,x;\tau,y) = P(X_\tau(\omega) < y \,|\, X_t(\omega) = x) \tag{7}$$

can be obtained, corresponding to the case when in $p(t,x;\tau,A)$ we take A of the form $(-\infty,y)$. It verifies the following relation

$$F(t,x;\tau,y) = \int_R F(s,z;\tau,y) d_z F(t,x;s,z).$$

Then, the *transition density function* with respect to y is as follows

$$f(t,x;\tau,y) = \frac{\partial}{\partial y} F(t,x;\tau,y), \tag{8}$$

and verifies the equalities

$$F(t,x;\tau,y) = \int_{-\infty}^{y} f(t,x;\tau,z)\, dz, \qquad \int_R f(t,x;\tau,y)\, dy = 1. \tag{9}$$

Furthermore, the Markov property (4) implies that

$$f(t,x;\tau,y) = \int_R f(s,z;\tau,y) f(t,x;s,z) dz, \qquad t < s < \tau \tag{10}$$

that is, the probability that $X(t)$ goes from x to y in the time interval $[t,T]$ is that probability that $X(\cdot)$ goes to any point z at any time s and then, *independently* of the way it reached z, it goes to y. The equality (10) is also referred to as the *Chapman-Kolmogorov equation* for Markov processes.

In certain conditions of existence, the transition density function satisfies the following two equations which are referred to as the *backward Kolmogorov equation* and respective the *forward Kolmogorov equation*

$$\frac{\partial f(t,x;\tau,y)}{\partial t} = -a(t,x)\, \frac{\partial f(t,x;\tau,y)}{\partial x} - \frac{1}{2}\, b(t,x)\, \frac{\partial^2 f(t,x;\tau,y)}{\partial x^2} \tag{11}$$

and

$$\frac{\partial f(t,x;\tau,y)}{\partial \tau} = -\frac{\partial}{\partial y}\, [a(\tau,y) f(t,x;\tau,y)] + \frac{1}{2}\, \frac{\partial^2}{\partial y^2}\, [b(\tau,y) f(t,x;\tau,y)] \tag{12}$$

where $a(t,x),\ b(t,x),\ a(\tau,y),\ b(\tau,y)$ are functions satisfying some conditions to assure the existence and the uniqueness of the solution of the equations. The forward Kolmogorov equation is also referred to as the *Fokker-Planck equation*.

Definition 6. A Markov process $X(t)$ is called a *diffusion process* if the following conditions are satisfied:

i For every $\varepsilon > 0$, t and x,

$$\lim_{\Delta t \to 0}\ \frac{1}{\Delta t} \int_{|y-x|>\varepsilon} F(t,x,t+\Delta t,y) dy = 0. \tag{13}$$

ii There exist the functions $a(t,x)$ and $b(t,x)$ such that for all $\varepsilon > 0$, t and x,

$$\lim_{\Delta t \to 0} \frac{1}{\Delta t} \int_{|y-x| \leq \varepsilon} (y-x)\, f(t,x,t+\Delta t,y)dy = a(t,x), \tag{14}$$

$$\lim_{\Delta t \to 0} \frac{1}{\Delta t} \int_{|y-x| \leq \varepsilon} (y-x)^2\, f(t,x,t+\Delta t,y)dy = b(t,x). \tag{15}$$

The function $a(t,x)$ is called the *(infinitesimal) drift coefficient* of $X(t)$ and $b(t,x)$ is called the *(infinitesimal) diffusion coefficient*. The intuitive meaning of conditions (13) - (15) and of the coefficients $a(t,x)$ and $b(t,x)$ is the following. In a short time interval h, the displacement of $X(\cdot)$ from a point x at time t is given by $a(t,x)\Delta t + \delta x + 0(\Delta t)$, where $a(t,x)$ is the velocity of the medium in which a particle (whose motion is described by $X(\cdot)$) drifts, δx is the random fluctuation of the particle due to random collision or thermal fluctuation, a.s.o. Furthermore, $E\,\delta x = 0, Var\,\delta x = b(t,x)\Delta t$. That is to say $b(t,x)$ is proportional to the average energy of the fluid molecules in the neighborhood of the particle. One can observe that the following conditions imply the conditions *i* and *ii* above:

(a) For any positive number δ, as $\Delta t \to 0$

$$\frac{1}{\Delta t} E_{x,t} \left| X(t+\Delta t) - X(t) \right|^{2+\delta} \to 0$$

(b) and

$$\frac{1}{\Delta t} E_{x,t} \left[X(t+\Delta t) - X(t) \right] \to a(t,x),$$
$$\frac{1}{\Delta t} E_{x,t} \left[X(t+\Delta t) - X(t) \right]^2 \to b(t,x).$$

4 Absorbing and reflecing barriers

Let us consider that a particle located on a straight line moves along the line via random impacts occurring at times $t_1, t_2, t_3, \cdots$. The particle can be at points with integral coordinates $a, a+1, a+2, \cdots, b$. At points a and b there are absorbing barriers. Each impact displaces the particle to the right with probability p and to the left with probability $q = 1-p$ so long as the particle is not located at a barrier. If the particle is at a barrier then, it remains in the states A_1 and A_{n-1} with probability 1.

A similar example can be considered for a particle being in a random walk, when at points a and b there are reflecting barriers. The conditions remain the same as in the former case, the only difference being that if the particle is at a barrier, any impact will transfer it one unit inside the gap between the barriers.

1. Let now be the case of a Brownian motion with an absorbing barrier. The forward Kolmogorov equation (12) for a Brownian motion on $x > 0$ with an absorbing boundary at $x = 0$ is given by

$$\begin{aligned}
&\frac{\partial p}{\partial t} = \frac{1}{2}\frac{\partial^2 p}{\partial y^2} \quad \text{in} \quad y > 0\\
&p(0,t,y) = 0, \quad t > 0, \;\; y > 0\\
&p(x,t,y) \to \delta(x-y) \quad \text{as} \quad t \downarrow 0, \;\; x > 0, \;\; y > 0.
\end{aligned}$$

The solution of such an initial boundary value problem is as follows

$$p(x,t,y) = \frac{1}{t\sqrt{2\pi}}\left[e^{-\frac{(x-y)^2}{2t^2}} - e^{-\frac{(x+y)^2}{2t^2}}\right].$$

It can be seen that by symmetry, $p(x,t,0) = 0$. Then, it can be shown that

$$\frac{1}{t\sqrt{2\pi}}\int_{-\infty}^{+\infty} e^{-\frac{(x+y)^2}{2t^2}}\varphi(x)dx \to \varphi(-y) = 0$$

as $t \downarrow 0$ if $y > 0$. Therefore,

$$p(x,t,y) \to \delta(x-y) \;\text{ as }\; t \downarrow 0 \;\text{ for all }\; x > 0,\, y > 0.$$

2. Now let us consider the Brownian motion on $x > 0$ but with a reflection barrier at the origin.

The forward Kolmogorov equation for a Brownian motion on $x > 0$ with an absorbing boundary at $x = 0$ is given by

$$\begin{aligned}
&\frac{\partial p}{\partial t} = \frac{1}{2}\frac{\partial^2 p}{\partial y^2}, \quad y > 0\\
&\left.\frac{\partial p(x,t,y)}{\partial y}\right|_{y=0} = 0\\
&p(x,t,y) \to \delta(x-y) \quad \text{as} \quad t \downarrow 0, \;\; x > 0, \;\; y > 0.
\end{aligned}$$

In this case the following solution is found

$$p(x,t,y) = \frac{1}{t\sqrt{2\pi}}\left[e^{-\frac{(x-y)^2}{2t^2}} + e^{-\frac{(x+y)^2}{2t^2}}\right]$$

and the condition

$$\left.\frac{\partial p(x,t,y)}{\partial x}\right|_{x=0} = 0$$

holds too.

References

1.Bharucha-Reid, A.T. Elements Of The Theory Of Markov Processes And Their Applications. Mineola, New York, 1997. Dover Publications, Inc.

2.Gnedenko, B.V. The Theory of Probability. Moskow, 1976. Mir Publisher.
3.Itô K. and McKean H.P. Jr. Diffusion Processes and their Sample Paths. Berlin Heidelberg, 1996. Springer-Verlag.
4.K. Itô. Stochastic Processes. Ole E. Barndorff-Nielsen and Ken-iti Sato, editors, 2004. Springer.
5.Kushner, H.J. and Yin, G.G. Stochastic Approximation Algorithms and Applications. New York, 1997. Springer-Verlag.
6.Øksendal, B. Stochastic Differential Equations: An Introduction with Applications, Sixth Edition. 2003. Springer-Verlag.
7.Øksendal B. and Sulem A. Applied Stochastic Control of Jump Diffusions. 2007. Springer-Verlag.
8.Orman, G.V. Lectures on *Stochastic Approximation Methods and Related Topics.* Preprint, "Gerhard Mercator" University, 2001. Duisburg, Germany.
9.Orman, G.V. Handbook of Limit Theorems and Stochastic Approximation. Brasov, 2003. "Transilvania" University Press.
10.Orman, G.V. On a Problem of Approximation of Markov Chains by a Solution of a Stochastic Differential Equation. In Christos H. Skiadas, Ioannis Dimotikalis and Charilaos Skiadas, editors, *Chaos Theory: Modeling, Simulation and Applications*, pages 30-40, 2011. World Scientific Publishing Co Pte Ltd.
11.Steele, J. M. Stochastic Calculus and Financial Applications. 2001. Springer-Verlag New York, Inc.
12.Stroock D. W. Markov Processes from K. Itô Perspective. Princeton, 2003. Princeton Univ. Press.
13.Schuss Z. Theory and Application of Stochastic Differential Equations. New York, 1980. John Wiley & Sons.
14.Wasan, M.T. Stochastic Approximation. Cambridge, 1969. Cambridge University Press.