

Backward and Forward equations for Diffusion processes.

Arka P. Ghosh
Department of Statistics
Iowa State University
Ames, IA 50011-1210
apghosh@iastate.edu
(515) 294-7851.

February 1, 2010

Abstract

This section is devoted to the discussion of two fundamental (partial) differential equations, that arise in the context of Markov diffusion processes. After giving a brief introduction of continuous-time continuous state Markov processes, we introduce the forward and backward equation, and provide a heuristic derivation of these equations for diffusion processes. We also discuss some examples and features of these two equations.

In this section we discuss two partial differential equations (PDE) that arise in the theory of continuous-time continuous-state Markov processes, which was introduced by Kolmogorov in 1931. Here, we focus only on Markov diffusion processes (see Section 2.1.6.1) and describe the forward and backward equation for such processes. The forward equation is also known as Fokker-Planck equation (and was already known in the physics literature before Kolmogorov formulated these). We begin by a brief introduction to continuous-time continuous-state Markov processes which are continuous analogs of Discrete Time Markov Chains (DTMC) and Continuous Time Markov Chains (CTMC) discussed earlier in Section 2.1.1 and 2.1.2 followed by some basic properties of Markov processes. Then we state the two equations and provide sketches of the proofs. Finally, we conclude the section with some specific examples and features of these equations.

Preliminaries.

Diffusion processes have been discussed in Section 2.1.6.1. For simplicity of the exposition, we consider the following time-homogeneous version of the diffusion process for this section: A (time-homogeneous) Itô diffusion is a stochastic process $\{X(t)\}$ satisfying a stochastic differential equation of the form

$$dX(t) = b(X(t))dt + \sigma(X(t))dW(t), t > 0; \quad X(0) = x, \quad (1)$$

where $\{W(t)\}$ is a (standard) Brownian motion and b, σ are functions that satisfy :

$$|\sigma(x) - \sigma(y)| < D|x - y|; \quad x, y \in \mathbb{R}.$$

It can be shown that for $\{\mathcal{F}_t^W\}, \{\mathcal{F}_t^X\}$ representing the filtrations generated by W and X ,

$$\mathcal{F}_t^X \subseteq \mathcal{F}_t^W. \quad (2)$$

Markov property: The diffusion satisfies the *Markov property*: If f is a bounded measurable function, then $E^x [f(X(t+h)) | \mathcal{F}_t^W] = E^{X(t)} [f(X(h))]$, where the superscript in the expectation represents that these are conditional expectation given $X(0) = t$. Thus, from (2), we have that: $E^x [f(X(t+h)) | \mathcal{F}_t^X] = E^{X(t)} [f(X(h))]$ This intuitively means that the (future) evolution of the diffusion process is completely specified by the current value of the process (and the knowledge of the history of the process is not necessary). Compare this property to the Markov property introduced in the context of Markov chains (discrete time versions in Sections 2.1.1 and 2.1.2) above. In fact, the above property holds in a stronger sense, in which the time t can be replaced by a random (stopping) time τ (strong Markov property).

Infinitesimal generator: For Markov processes, the infinitesimal generator is defined as

$$Af(x) = \lim_{t \downarrow 0} \frac{E^x(f(X(t)) - f(x))}{t}, \quad x \in \mathbb{R}. \quad (3)$$

The set of functions f for which the limit on the right side exists (for all x) is called the domain of the operator, and denoted by \mathcal{D}_A . For an Itô diffusion that we are concerned with, this operator is a second order partial differential operator, and can be written down explicitly as follows:

$$Af(x) = b(x) \frac{df}{dx} + \frac{1}{2} \sigma^2(x) \frac{d^2 f}{dx^2}, \quad (4)$$

and \mathcal{D}_A contains all twice-differentiable functions with compact support. This operator plays a fundamental role in the study of diffusion processes, as well as relevant for understanding the forward and backward Kolmogorov equations for diffusions.

Examples: Here are generators of some basic examples of diffusion processes discussed in Section 2.1.6.1. For a (general) Brownian motion (with drift b and diffusion coefficient $\sigma > 0$) the generator is:

$$Af(x) = bf'(x) + \frac{1}{2} \sigma^2 f''(x). \quad (5)$$

For geometric Brownian motion process (with parameters b and σ), the generator is

$$Af(x) = bx f'(x) + \frac{1}{2} x^2 \sigma^2 f''(x). \quad (6)$$

For Ornstein Uhlenbeck process (with parameters b and σ), it is

$$Af(x) = -bx f'(x) + \frac{1}{2} \sigma^2 f''(x). \quad (7)$$

Forward and Backward Equations for Diffusion processes.

These are two (partial) differential equation that characterize the dynamics of the distribution of the diffusion process. Kolmogorov's forward equation addresses the following: If at time t the state of the system is x , what can we say about the distribution of the state at a future time $s > t$

(hence the term “forward”). The backward equation, on the other hand, is useful when address the question that given that the system at a future time s has a particular behavior, what can we say about the distribution at time $t < s$. This imposes a terminal condition condition on the PDE, which is integrated backward in time, from s to t (hence the term “backward” is associated with this). Historically, the forward equation is was discovered (as the Fokker-Plank equation) before the backward equation. However, the backward equation is somewhat more general and we will describe that first.

The forward and backward equations are expressed in various equivalent forms. Here, we describe them in the following form, which illustrates the use of the terms forward and backward more clearly: Fix an interval $[0, T]$, and we will deal with $t \in [0, T]$.

Backward Equation: Let $g(x)$ be a bounded smooth (twice continuously differentiable having compact support) function, and let

$$u(t, x) = E^{x,t}(g(X(T))) \equiv E(g(X(T))|X(t) = x). \quad (8)$$

Then u satisfies:

$$\frac{\partial u}{\partial t} + Au = 0, \text{ with the “terminal” condition } u(T, x) = g(x), \quad (9)$$

where the right hand side is to be interpreted as A (as in (4)) applied to the $u(t, x)$ as a function of x .

Forward Equation: In addition, if $X(t)$ has a density $p(t, x)$, then for a probability density function $\mu(\cdot)$, the probability densities satisfy the following:

$$\frac{\partial}{\partial t} p(t, x) = (A^*p)(t, x) \text{ with the initial condition } p(0, x) = \mu(x). \quad (10)$$

Here A^* is the adjoint operator of A , defined as:

$$A^*v(t, y) = -\frac{\partial}{\partial y} (b(y)v(t, y)) + \frac{1}{2} \frac{\partial^2}{\partial y^2} (\sigma^2(y)v(t, y)). \quad (11)$$

Proof: We now provide a sketch of the proof of the backward equation (9). First note that for the Markov (diffusion) process $\{X(t)\}$ with some transition probability function

$$P(t, x, A) = P(X(t) \in A | X(0) = x),$$

the following Chapman Kolmogorov equation holds:

$$P(t + s, x, A) = \int P(t, x, du)P(s, u, A) \quad (12)$$

Recall the Chapman-Kolmogorov Equation introduced earlier in Section 2.1.1 in the context of Markov chains. Also note that for the diffusion in (1), the following properties are satisfied for the transition probabilities:

$$\begin{aligned} \lim_{t \rightarrow \infty} \frac{1}{t} \int_{\{\|x-y\| \geq \delta\}} P(t, x, dy) &= 0, \\ \lim_{t \rightarrow \infty} \frac{1}{t} \int_{\{\|x-y\| < \delta\}} (y - x)P(t, x, dy) &= b(x), \\ \lim_{t \rightarrow \infty} \frac{1}{t} \int_{\{\|x-y\| \geq \delta\}} (y - x)^2 P(t, x, dy) &= \sigma(x), \end{aligned} \quad (13)$$

for all $\delta > 0$. Informally, it means the following: in a small time interval, it has negligible probability of being away from x ; also the mean and variance of the “displacements” of the diffusion process is approximately the drift and the diffusion coefficients, respectively. We will use these properties in the proofs of the forward and backward equations here. First, observe that from (12) and (8) it follows that

$$u(t+h, x) = \int P(h, x, dy)u(t, y).$$

Hence,

$$\begin{aligned} \frac{u(t+h) - u(t)}{h} &= \frac{\int P(h, x, dy)[u(t, y) - u(t, x)]}{h} \\ &\approx \frac{1}{h} \int P(h, x, dy) \left[(y-x) \frac{\partial u(t, x)}{\partial x} + \frac{1}{2} (y-x)^2 \frac{\partial^2 u(t, x)}{\partial x^2} \right], \end{aligned} \quad (14)$$

where the last expression follows from Taylor’s expansion, and the approximate equality is a consequence of (13), for h small. Hence, by taking limit as $h \rightarrow 0$, using (13) and the form of the generator A in (4), the backward equation follows.

Now we sketch the proof of the forward equation (10) using the backward equation as follows: Assume that the random variable $X(t)$ has a density $p(t, x)$ and hence, from (8) and properties of conditional expectations, we have

$$E(g(X(T))) = E\left(E^{x,t}(g(X(T)))\right) = \int u(t, x)p(t, x)dx.$$

Since the left side is free of t , we get by taking derivatives with respect to t ,

$$0 = \int p(t, x) \frac{\partial u(t, x)}{\partial t} dx + \int \frac{\partial p(t, x)}{\partial t} u(t, x) dx. \quad (15)$$

Since u satisfies the backward equation (9), using the form of A in (4), we get from (15) that

$$0 = \int p(t, x) \left(-b(x) \frac{\partial u(t, x)}{\partial x} - \frac{1}{2} \sigma^2(x) \frac{\partial^2 u(t, x)}{\partial x^2} \right) dx + \int \frac{\partial p(t, x)}{\partial t} u(t, x) dx.$$

Finally, using integration by parts (which replaces x -derivatives of u by those of p), assuming that the integrals decay fast enough (as $|x| \rightarrow \infty$) and using the form of A^* in (11), we get that

$$\int \left(-(A^*p)(x, t) + \frac{\partial p(t, x)}{\partial t} \right) u(t, x) dx = 0.$$

Since this equation should be true for all functions u , we get that $(A^*p)(x, t) = \frac{\partial p(t, x)}{\partial t}$ which proves the forward equation.

Examples.

Here we discuss some special cases of the forward and backward diffusions.

Example 1 [Martingale Property]: When $b \equiv 0$, then the forward equation reduces to

$$\frac{1}{2} \frac{\partial^2}{\partial x^2} (\sigma^2(x)p(t, x)) = \frac{\partial}{\partial t} p(t, x). \quad (16)$$

Hence, using integration by parts, one gets

$$\begin{aligned}\frac{d}{dt}E[X(t)] &= \frac{d}{dt} \int_{-\infty}^{\infty} xp(x,t)dx = \int_{-\infty}^{\infty} x \frac{\partial}{\partial t} p(t,x)dx = \int_{-\infty}^{\infty} x \frac{1}{2} \frac{\partial^2}{\partial x^2} (\sigma^2(x,t)p(x,t))dx \\ &= - \int_{-\infty}^{\infty} \frac{1}{2} \frac{\partial}{\partial x} (\sigma^2(x,t)p(x,t))dx = 0.\end{aligned}$$

In other words, $E[X(t)]$ does not change with t . In fact, one can prove a stronger statement: then $\{X(t)\}$ is a martingale, i.e. $E[X(t) | \mathcal{F}_s^W] = X(s)$ for all $s < t$.

Example 2: Consider the Ornstein Uhlenbeck process with parameters b and σ (see Section 2.1.6.1), whose generator was discussed earlier (for simplicity, we assume $\sigma = 1$ here). The backward equation for this process follows from the general form of the equation (9) and the generator in (7):

$$\frac{\partial}{\partial t} u(t,x) = -bx \frac{\partial}{\partial x} u(t,x) + \frac{1}{2} \frac{\partial^2}{\partial x^2} u(t,x), \quad (17)$$

which can be solved explicitly. It turns out, with a change of variable, it reduces to the standard diffusion equation (16) (with $\sigma(x) \equiv 1$). From that, it can be deduced that the transition probability $P(t,x,\cdot)$ is the normal (Gaussian) probability with mean xe^{-bt} and variance parameter $\frac{(1-e^{-2bt})}{2b}$.

Example 3: Similarly, for the geometric Brownian motion process, one can get the forward equation by substituting $b(x) = bx, \sigma(x) = \sigma x$ in (9) and (6). Solving these explicitly yields that the transition probability in this case is a log-normal probability:

$$P(t,x,A) = \int_A \frac{1}{\sigma y \sqrt{2\pi t}} \exp\left(-\frac{\log(y/x) - (b - \frac{1}{2}\sigma^2)t}{2\sigma^2 t}\right) dy.$$

Remark: The backward equation holds whenever b is continuous and $\sigma > 0$. But for the forward equation to hold, one clearly needs the derivatives of these two functions. In that sense, backward equation is more general than the forward equation. However, when such derivatives exist, one can show that there exists a “minimal” solution for the transition function $P(t,x,A)$ such that $u(\cdot,\cdot)$ in (8) solves the backward equation. But this transition function in such cases need not be a “proper” probability. To guarantee a proper transition probability can be obtained from these equation, one usually assumes additional boundary conditions. Different choices of such boundary conditions (e.g. absorbing boundary condition, reflecting boundary conditions etc.) uniquely characterizes the associated process (see Chapter X of [3] for more discussion on this).

Further Reading.

More detail about these equations can be found in classical textbooks on probability, such as [3]. Other references for these equations for diffusion processes are [5], [6], [16], [17] etc. For students and researchers that are trying to learn this material for the first time, more accessible reference could be [4] and [15]. A reference for general Markov processes can be found in [2]. For partial differential equations in general, one is encouraged to consult [11].

As mentioned earlier, these equations are very important in physics and there is a vast literature ([1], [9], [8] etc.) on this topic from the physics community as well. For applications to financial mathematics, see [12], [10]. A general reference for students in operations research to learn stochastic systems in general is [7]. For more applications in queueing theory models involving diffusions processes, see [14] and [13].

References

- [1] Berry, S. Functional Integration and Quantum Physics. 1979. Academic Press.
- [2] Ethier, SN, Kurtz, TG. Markov processes. Characterization and convergence. Wiley Series in Probability and Mathematical Statistics: Probability and Mathematical Statistics. 1986. John Wiley & Sons, Inc., New York.
- [3] Feller, W. An Introduction to Probability Theory and Its Applications, (Vol. 2) 1991. Wiley, 2nd Ed.
- [4] Karatzas, I, Shreve, S. Brownian Motion and Stochastic Calculus. 1991. Springer, 2nd edition.
- [5] Ikeda, N. Ito's Stochastic Calculus And Probability Theory. 1996. Springer.
- [6] Itô, K, McKean, HP. Diffusion Processes and their Sample Paths. 1996. Springer.
- [7] Kulkarni, VG. Modeling and Analysis of Stochastic Systems. 1996. Chapman & Hall; 1 edition.
- [8] Risken, H, Frank, T. The Fokker-Planck Equation: Methods of Solutions and Applications. 1996. Springer, 2nd Ed.
- [9] Kadanoff, LP. Statistical Physics: statics, dynamics and renormalization. 2000. World Scientific.
- [10] Steele, JM. Stochastic Calculus and Financial Applications. 2000. Springer.
- [11] Gilbarg, D, Trudinger, NS. Elliptic Partial Differential Equations of Second Order. 2001. Springer.
- [12] Karatzas, I, Shreve, S. Methods of Mathematical Finance. 2001. Springer.
- [13] Kushner, HJ. Heavy Traffic Analysis of Controlled Queueing and Communications Networks. 2001. Springer, 1st Ed.
- [14] Whitt, W. Stochastic-Process Limits. 2002. Springer.
- [15] Oksendal, BK. Stochastic Differential Equations: An Introduction with Applications. 2003. Springer, 6th Ed.
- [16] Protter, P. Stochastic Integration and Differential Equations. 2005. Springer, 2nd Ed.
- [17] Stroock, DW, Varadhan, SRS. Multidimensional Diffusion Processes. 2005. Springer.