MA8109 fall 2011

The Kolmogorov and Fokker-Planck Equations

Harald E. Krogstad, IMF/rev. 2011

This short note summarizes the material about the Kolmogorov and Fokker-Planck Equations which is covered in the lectures, but not in Øksendal.

The topic is still diffusion, and we have limited ourselves to time homogeneous (autonome) *Itô diffusions* of the form

$$dX_t = b(X_t)\,dt + \sigma(X_t)\,dB_t. \tag{1}$$

The conditions on b and σ are in general the same as in Theorem 5.2.1, thus ensuring well-behaved solutions. In particular, X_t is continuous *a.s.*

Itô diffusions are Markov processes, and the main result in Chapter 7.1 was Theorem 7.1.2. Changing the notation slightly, we shall from now on write

$$u(t,x) = \mathsf{E}\left(f\left(X_t^x\right)\right) = \mathsf{E}^x\left(f\left(X_t\right)\right). \tag{2}$$

Observe that $u(t,x)$ is the same as $G_t(x)$ in the note about diffusion. Here f is a bounded Borel function as in Theorem 7.1.2, or a smoother function in $C_c^2(\mathbb{R}^n)$ if we need to compute $A(f)$ for the infinitesimal generator of the diffusion. Theorem 7.1.2 may be restated as

$$\mathsf{E}^x\left(f\left(X_{r+t}\right)|\mathcal{F}_r\right)(\omega) = \mathsf{E}^{X_r(\omega)} f\left(X_t\right) = G_t\left(X_r(\omega)\right) = u\left(t, X_r(\omega)\right). \tag{3}$$

We also recall that the generator A is the second order differential operator defined by

$$Af = \sum_i b_i(x)\frac{\partial f}{\partial x}(x) + \frac{1}{2}\sum_{i,j}\left(\sigma(x)\sigma(x)'\right)_{ij}\frac{\partial^2 f}{\partial x_i \partial x_j}(x). \tag{4}$$

It is convenient to assume that b and σ are bounded, continuous functions, and that f and its first and second derivatives in Eqn. 4 are continuous functions vanishing at ∞, that is, $f \in C_0^2(\mathbb{R}^n)$. Then $Af(x)$ will be a bounded, continuous function in x. This in turn implies that in Dynkin's Formula, the function within the integral,

$$t \to Af\left(X_t^x(\omega)\right), \tag{5}$$

is bounded and continuous a.s. Also the expectation,

$$t \to \mathsf{E}^x\left[Af\left(X_t\right)\right] = \int_\Omega Af\left(X_t^x(\omega)\right) dP(\omega), \tag{6}$$

will be a continuous function of t by dominated convergence.

The main result in Chapter 8.1 is the following theorem:

Let $f \in C_0^2(\mathbb{R}^n)$ and $u(t,x) = \mathsf{E}^x(f(X_t))$. Then

$$\begin{array}{lc} (i) & \frac{\partial u}{\partial t} = Au,\ t>0,\ x \in \mathbb{R}^n, \\ (ii) & \lim_{t\to 0} u(t,x) = f(x),\ x \in \mathbb{R}^n \\ (iii) & u \text{ is unique.} \end{array} \tag{7}$$

The theorem is called Kolmogorov Backward Equation in B.Ø., but this does not seem to be quite in accordance with other books on the same topic. We will return to this later. For readers familiar to PDEs, the theorem states that $u(t,x)$ is the unique solution to the equation $\frac{\partial u}{\partial t} = Au$ for $t > 0,\ x \in \mathbb{R}^n$, satisfying the initial condition $u(0,x) = f(x)$. For standard Brownian motion, the equation is the familiar parabolic heat (or diffusion) equation,

$$\frac{\partial u}{\partial t} = \frac{1}{2}\nabla^2 u. \tag{8}$$

There are three key points in the proof of this theorem. The first is to establish that $\partial u/\partial t$ really exists, and this follows from Dynkin's Formula applied with the constant stopping time $\tau = t$:

$$u(t,x) = \mathsf{E}^x(f(X_t)) = f(x) + \mathsf{E}^x\left[\int_0^t Af(X_s)\,ds\right]. \tag{9}$$

Since $Af(X_s)$ is bounded, the expectation may be moved inside the integral by Fubini's Theorem, and since the integrand of the s-integral is then a continuous function in s, we have immediately that

$$\frac{\partial u}{\partial t} = \mathsf{E}^x Af(X_t) \tag{10}$$

by the Fundamental Theorem of Calculus. It is tempting to write

$$\mathsf{E}^x Af(X_t) = A\mathsf{E}^x(f(X_t)), \tag{11}$$

but this seems to be rather unlikely, since $Af(X_t)$ contains terms like

$$\cdots + b_i(X_t)\frac{\partial f}{\partial x_i}(X_t) + \cdots. \tag{12}$$

However, it is nevertheless true, and the idea is to show that $Au(t,x)$ indeed exists by considering the Generator Theorem,

$$Au(t,x) = \lim_{r\to 0}\frac{\mathsf{E}^x[u(t,X_r)] - u(t,x)}{r}. \tag{13}$$

Note that here, t is just a parameter. We do not yet know whether this limit exists, so we consider the first term in the numerator on the RHS of Eqn. 13 separately:

$$\begin{aligned}\mathsf{E}^x[u(t,X_r)] &= \mathsf{E}^x\left[\mathsf{E}^{X_r}(f(X_t))\right]\\ &= \mathsf{E}^x[\mathsf{E}^x(f(X_{r+t})|\mathcal{F}_r)]\\ &= \mathsf{E}^x(f(X_{r+t}))\\ &= u(t+r,x).\end{aligned} \tag{14}$$

The first equality is the definition of u, the second is Theorem 7.1.2 (here Eqn. 3), and the third is the "double-expectation rule" for conditional expectations, *viz.*,

$$\mathsf{E}(\mathsf{E}(X|\mathcal{H})) = \int_\Omega \mathsf{E}(X|\mathcal{H})\,dP = \int_\Omega X\,dP = \mathsf{E}(X). \tag{15}$$

We now introduce Eqn. 14 into 13, and note that we already know (from Eqn. 10) that this limit exists and is equal to $\partial u/\partial t$. Dynkin's Formula (Eqn. 9) shows immediately that

$$\lim_{t\to 0} u(t,x) = f(x) = u(0,x). \tag{16}$$

The third key element in the proof is uniqueness of the solution, which is shown by an elegant argument in the book. Because the equation is linear, it is enough to show that if there is a function $w \in C^{1,2}(\mathbb{R}\times\mathbb{R}^n)$ so that

$$-\frac{\partial w}{\partial t} + Aw = 0,\ t \geq 0,\ x\in\mathbb{R}^n, \tag{17}$$

$$w(0,x) = 0, \tag{18}$$

then $w(t,x) = 0$ for all $t > 0$.

Let Y_t be the Itô diffusion on $\mathbb{R}\times\mathbb{R}^n$ defined by

$$dY_t = \begin{bmatrix} -dt \\ dX_t \end{bmatrix} = \begin{bmatrix} -1 \\ b \end{bmatrix} dt + \begin{bmatrix} 0 \\ \sigma \end{bmatrix} dB_t \tag{19}$$

The generator of Y_t, $\tilde{A}$, as a differential operator acting on functions on $\mathbb{R}\times\mathbb{R}^n$, is just the operator on the left hand side of Eqn. 17 (check it!). Note that for the start position $Y_0 = (s,x)$,

$$Y_t = (s-t, X_t^x),\ t \geq 0. \tag{20}$$

We apply Dynkin's Formula for Y_t, the function w, the start position (s,x), and, following BØ, the stopping time $\tau = t\wedge\tau_R$, $\tau_R = \inf\{t>0, |X_t| \geq R\}$,

$$\mathsf{E}^{s,x}(w(Y_\tau)) = w(Y_0) + \mathsf{E}^{s,x}\int_0^\tau \tilde{A}w(Y_u)\,du. \tag{21}$$

Clearly, $\mathsf{E}\tau < \infty$ and the last term vanishes since $\tilde{A}w = -\frac{\partial w}{\partial t} + Aw = 0$ according to the assumptions. Thus,

$$w(s,x) = \mathsf{E}^{(s,x)}(w(Y_\tau)) \underset{R\to\infty}{\longrightarrow} \mathsf{E}^{(s,x)}(w(Y_t)). \tag{22}$$

The final trick is to let $t = s$:

$$w(s,x) = \mathsf{E}^{(s,x)}(w(Y_s)) = \mathsf{E}^{(s,x)}(w(0,X_s)) = w(0,X_s) = 0, \tag{23}$$

again according to the assumption. This proves the theorem (7).

We recall that Markov processes may be characterized in terms of *transition functions,* and in some fortunate cases these transition functions may even be expressed by means of *transition densities*, which we here shall write $p(t,x,y)$ (a slight change compared to the previous note about diffusion). The transition density is a probability density in y for all x and t-s, and we obtain the probability for $X_t\in A$ when $X_0 = x$ simply as

$$\mathsf{P}(X_t^x \in A) = \int_A p(t,x,y)\,dy. \tag{24}$$

Recall that the transition densities for the standard Brownian motion have the simple form

$$p(t,x,y) = \frac{1}{(2\pi t)^{n/2}} \exp\left(-\frac{|x-y|^2}{2t}\right). \tag{25}$$

Moreover, the transition densities satisfy the *Chapman-Kolmogorov Equations*,

$$p(t+s,x,y) = \int_{\mathbb{R}^n} p(s,x,z)\, p(t,z,y)\, dz. \tag{26}$$

If the diffusion we are considering has transition densities, we may write the solution to the problem above as

$$u(t,x) = \mathsf{E}^x(f(X_t)) = \int_{\mathbb{R}^n} f(y)\, p(t,x,y)\, dy. \tag{27}$$

By applying the uniqueness in the theorem above to the two problems

$$\frac{\partial u}{\partial t} = Au,\ \ u(0,x) = f(x), \tag{28}$$

$$\frac{\partial \tilde{u}}{\partial t} = A\tilde{u},\ \ \tilde{u}(0,x) = u(t_0,x), \tag{29}$$

we observe that

$$u(t_0+t_1,x) = \tilde{u}(t_1,x). \tag{30}$$

We leave to the reader also to verify Eqn. 30 by applying the Chapman-Kolmogorov equations.

What about the equation for p? We shall assume that $p(t,x,y)$ for $t>0$ are smooth and nice functions so that we may move derivations in and out of integrals, etc. First of all, we recall Eqn. 10,

$$\frac{\partial u}{\partial t} = \mathsf{E}^x(Af(X_t)), \tag{31}$$

and hence, introducing Eqn. 27 and moving $\partial/\partial t$ inside the integral,

$$\int_{\mathbb{R}^n} \left[f(y) \frac{\partial p(t,x,y)}{\partial t} - Af(y)\, p(t,x,y) \right] dy = 0. \tag{32}$$

We now apply a well-known technique from Hilbert space theory, namely the concept of the *adjoint operator*. The adjoint operator is defined by the identity

$$\int A\phi(y) \cdot \psi(y)\, dy = \int \phi(y) \cdot A^*\psi(y)\, dy, \tag{33}$$

supposed to hold for all functions ϕ and ψ so that $A\phi$, $A^*\psi$, and the integrals exist. In the present case, partial integrations (applying functions in $C_c^2(\mathbb{R}^n)$) show that the adjoint operator has the form

$$A^*\psi(y) = -\sum_i \frac{\partial}{\partial y_i}\left[b_i(y)\,\psi(y)\right] + \frac{1}{2}\sum_{i,j} \frac{\partial^2}{\partial y_i \partial y_j}\left[\left(\sigma(y)\,\sigma(y)'\right)_{ij} \psi(y)\right]. \tag{34}$$

(If necessary, read more about this in any book on PDE's or Hilbert space theory). If we apply the definition of the adjoint operator in Eqn. 32, we obtain to separate $f(y)$ from the rest,

$$\int_{\mathbb{R}^n} f(y)\left[\frac{\partial p(t,x,y)}{\partial t} - A^*_{(y)} p(t,x,y)\right] dy = 0. \tag{35}$$

The subscript (y) simply indicates that y is the variable to use in A. The identity will hold for all functions $f \in C_c^2(\mathbb{R}^n)$. If the bracket is a continuous function, as will be the case if p is smooth $(t > 0)$, the only possibility is that

$$\frac{\partial p(t,x,y)}{\partial t} - A^*_{(y)} p(t,x,y) = 0. \tag{36}$$

(This is a simple, but *very* useful argument: *If g is continuous and*

$$\int_{\mathbb{R}^n} f(y)\, g(y)\, dy = 0 \tag{37}$$

for all f-s in a dense set in L^1 so that the integral exists, then $g(y) \equiv 0$. Here $C_c^2(\mathbb{R}^n)$ is dense in $L^1(\mathbb{R}^n)$. What is the related result in $L^2(\mathbb{R}^n)$?).

Starting from a point x, Eq. 36 defines the development of $p(t,x,y)$ forward in time. For Brownian motion,

$$A = A^* = \frac{1}{2}\nabla^2, \tag{38}$$

and the equation for p is identical to the equation for u, but this is not true in general. If you know about distribution theory, you will notice that the solution converges towards a δ-function at x when $t \to 0$,

$$\lim_{t\to 0} p(t,x,y) = \delta(x-y). \tag{39}$$

(Do not worry about this if you no not know about distribution theory!).

Equation 36 is called the *Fokker-Planck Equation* or *Kolmogorov's Forward Equation*.

It is now also easy to derive what is commonly denoted *Kolmogorov's Backward Equation*, starting with the expression for u in Eq. 27 and $\partial u/\partial t - Au = 0$. Again we move $\partial/\partial t$ and A inside the integral and obtain

$$\begin{aligned} &\int_{\mathbb{R}^n} \left[\frac{\partial}{\partial t} f(y)\, p(t,x,y) - A_{(x)} f(y)\, p(t,x,y)\right] dy \\ &= \int_{\mathbb{R}^n} f(y)\left[\frac{\partial p(t,x,y)}{\partial t} - A_{(x)} p(t,x,y)\right] dy. \end{aligned} \tag{40}$$

The same argument as above leads to

$$\frac{\partial p(t,x,y)}{\partial t} = A_{(x)} p(t,x,y). \tag{41}$$

Note that here the generator works on x, whereas y is fixed. The equation enables us to determine p *backward* in time, say

$$\tilde{p}(s,x) = p(t_0 - s, x, y_0),\ 0 < s \le t_0. \tag{42}$$

The equation for $\tilde{p}$ becomes

$$\frac{\partial \tilde{p}}{\partial s} = -A\tilde{p}, \tag{43}$$

and *this* is the equation that is usually called *Kolmogorov's Backward Equation*.