Lecture notes by:
Matthias Kredler
Universidad Carlos III de Madrid
This version: May 2018

1 The Kolmogorov forward equation

Suppose we know the distribution of a mass of particles (e.g. economic agents) at $t = 0$, given by a density function $n_0(x)$ over a one-dimensional state variable x. Also suppose that we know the law of motion for the particles (e.g. coming from the optimal policy for a control problem), which is given by a diffusion

$$dX_t = a(X_t)dt + s(X_t)dW_t.$$

The *Kolmogorov forward equation* answers the following question: What is the density function $n(t, x)$ over the space of agents at each point in time $t > 0$ that results from the initial density and the movement of the single particles? Note that we have already answered this question for a deterministic law of motion in the beginning of this course – we found a first-order PDE characterizing the process.

If you have a look back at how we solved the problem there, you will quickly see that the techniques we used then are not helpful anymore with Brownian Motion since a non-zero measure of paths will leave any interval even for a very short interval of time.

Instead, we are going to reason as follows: Two densities $n_1(x)$ and $n_2(x)$ are the same if and only if the following "moments" are the same for any function $f(x)$ and any interval $[a, b]$:

$$\int_a^b f(x)n_1(x)dx = \int_a^b f(x)n_2(x)dx$$

It is easy to see that $n_1 = n_2$ implies that the moments are the same for any f and any interval. To see that also the converse holds, consider the following argument: If n_1 was greater than n_2 on some small set (c, d) (this is without loss of generality – why?), then the above statement would be false for any function that is strictly positive on $[c, d]$ but zero otherwise.

To make our life easier, we will restrict ourselves to functions f that are twice continuously differentiable on some interval $[a, b]$ and whose function value and first derivative tend to zero as we approach the bounds of the support interval. Note that also for this set of functions, the above statement still holds (why?): Two densities are identical if and only if for all f and all associated intervals $[a, b]$ the above moments coincide.

We start the argument by checking how the following moment (or expectation) for a fixed function f with its associated interval $[a, b]$ is changing over a small increment of time:

$$A_f(t) = \int_a^b f(x)n(t, x)dx$$

If our state x is consumption, this could be the average utility over all agents, using some particular utility function $f(x)$. The trick now is that we can think of obtaining the function value $A(t+\Delta t)$ a tiny bit ahead in the future by two means:

- If someone gave us the density, of course we can approximate by linearly extrapolating the density at each point x:

$$A(t+\Delta t) = \int_a^b f(x)\frac{\partial n(t,x)}{\partial t}\Delta t dx$$

- On the other hand, we can also look at the mass of agents at each spot x in the state space and calculate with Ito's Lemma how their utility will change:

$$df(x) = f'(x)\big[a(x)dt + s(x)dW_t\big] + f''(x)\frac{s^2(x)}{2}dt$$

  Now, we integrate over all positions x on the interval, and for each x we then integrate over all possible paths crossing through x at t to get

$$\Delta A(t) \simeq \int_a^b n(t,x)\int_{\Delta W_t}\left[f'(x)a(x)\Delta t + f'(x)s(x)\Delta W_t + f''(x)\frac{s^2(x)}{2}\Delta t\right]dx$$

  Note that the average over the Brownian-Motion term will be zero for each single position x (why?), and we will only be left with the deterministic terms involving Δt:

$$\Delta A(t) \simeq \int_a^b n(t,x)\left[f'(x)a(x) + f''(x)\frac{s^2(x)}{2}\right]\Delta t dx$$

Now, we set the two expressions coming from the different ways of obtaining the moment at $t+\Delta t$ equal to each other, divide by Δt and take the limit as Δt gets small to get

$$\int_a^b n(t,x)\left[f'(x)a(x) + f''(x)\underbrace{\frac{s^2(x)}{2}}_{=\tilde{s}(x)}\right] = \int_a^b \frac{\partial n(t,x)}{\partial t}f(x)dx$$

Then we use the following two tricks:

- For two reasonably-behaved functions f and g, if $g(a)=g(b)=0$ then we have

$$\int_a^b g'(x)h(x)dx = -\int_a^b g(x)h'(x)dx \tag{1}$$

  by integration by parts.

- If we apply the above result to the pairs $g'(x)h(x)$ and $g(x)h'(x)$ (assuming in addition that $g'(a)=g'(b)=0$), we get

$$\int_a^b g''(x)h(x)dx = \int_a^b g(x)h''(x)dx \tag{2}$$

In our problem, we let f play the role of g (recall that we assumed that both f and f' are zero at the boundaries of the interval), and we replace $h(x)$ by $n(t,x)a(x)$ in (1) and $n(t,x)\tilde{s}(x)$ in (2) to get

$$\int_a^b \frac{\partial n(t,x)}{\partial t} f(x)dx = \int_a^b f(x)\Big(- \frac{d}{dx}\Big[n(t,x)a(x)\Big] + \frac{d^2}{dx^2}\Big[n(t,x)\tilde{s}(x)\Big]\Big)dx$$

But now, the same logic as in the very beginning applies: The term in the large parentheses on the right-hand side must be equal to $\partial n(t,x)/\partial t$ – if this was violated for some x, we could find a suitable function f (recall that f was arbitrary!) such that the above inequality didn't hold.

So we get the following second-order PDE for $n(t,x)$, which is called the *Kolmogorov Forward Equation*:

$$\begin{aligned}\frac{\partial n(t,x)}{\partial t} &= -\frac{d}{dx}\Big[n(t,x)a(x)\Big] + \frac{d^2}{dx^2}\Big[n(t,x)\tilde{s}(x)\Big] = \qquad (3)\\ &= -a(x)\frac{\partial n(t,x)}{\partial x} - a'(x)n(t,x) + \frac{\partial^2 n(t,x)}{\partial x^2}\tilde{s}(x) + 2\frac{\partial n(t,x)}{\partial x}\tilde{s}'(x) + \tilde{s}''(x)n(t,x)\end{aligned}$$

2 Forward vs. backward equations

The forward equation is called "forward" equation because we develop the density $n(t,x)$ *forward* in time from $t=0$ on. There is also a *backward* equation – we have already seen it many times but never called it that way. The backward equation answers –for example– the question what an option is worth at $t<T$ if we know the final value $v(T,x)$. Generally, the backward equation characterizes the function $v(t,x)$ that solves the following problem (it is normally stated without discounting, so we will also do this here although we have already stated this equation for a more general case):

$$\begin{aligned} dX_t &= a(X_t)dt + s(X_t)dW_t \\ v(t,x) &= E\big[g(X_T)|X_t = x\big] \\ v(T,x) &= g(x)\end{aligned}$$

It shouldn't be hard for you to verify that v must satisfy the following PDE, which is called the *(Kolmogorov) Backward Equation*:

$$-\frac{\partial v(t,x)}{\partial t} = a(x)\frac{\partial v(t,x)}{\partial x} + \frac{s^2(x)}{2}\frac{\partial^2 v(t,x)}{\partial x^2} \qquad (4)$$

3 Simulating forward equations

In order to numerically solve forward equations, just shift the mass forward on a trinomial grid using the rules that we also used for solving backward in time (remember the homework for the European Option): Take the mass in a cell at t and distribute to the three adjacent cells at $t+\Delta t$ according to the probabilities

p_u, p_m and p_d that we always use. With this algorithm, the total mass at any t will always sum up to one, and we know that it has very good properties when approximating diffusions.