TFR Notes

Veronika Chronholm

August 4, 2023

1 Introduction

2 Theory

2.1 Forward and Backward PDEs

Treatment planning problem (PDE constrained optimisation):

$$\min \frac{1}{2}\|u - d_T\|^2 + \frac{\alpha}{2}\|g\|^2 \tag{1}$$

subject to a constraint

$$\partial_t u + b \cdot \nabla u - \mu \Delta u = g \tag{2}$$

$$u|_{\partial\Omega} = 0 \tag{3}$$

Using the method of Lagrange multipliers, we can solve the above optimisation problem by solving simultaneously the two PDEs

$$-\partial_t z - \mu \Delta z - b \cdot \nabla z = u - d_T \tag{4}$$

$$\partial_t u - \mu \Delta u + b \cdot \nabla u = \frac{1}{\alpha} z \tag{5}$$

2.2 Forward SDE and linear Feynman-Kac formula

Let $b : \mathbb{R}^+ \times \mathbb{R}^d \to \mathbb{R}^d$, $\sigma : \mathbb{R}^+ \times \mathbb{R}^d \to \mathbb{R}^{d\times d}$, and W_t a d-dimensional Brownian motion. Consider the d-dimensional SDE

$$X_t = x + \int_0^t b(s,X_s)ds + \int_0^t \sigma(s,X_s)dW_s \tag{6}$$

for the process X_t.

The infinitesimal generator of X_t is the differential operator

$$\mathcal{L} = \frac{1}{2}\sum_{i,j=1}^{d} [\sigma\sigma^T]_{i,j}(t,x)\partial^2_{x_i,x_j} + \sum_{i=1}^{d} b_i(t,x)\partial_{x_i}. \tag{7}$$

This operator is defined by

$$\mathcal{L}f(x) = \lim_{t\to 0} \frac{\mathbb{E}_x[f(X_t)] - f(x)}{t}. \tag{8}$$

Below, we describe the connection between a PDE involving this differential operator, and the solution to the SDE Equation 6, yielding a probabilistic representation of the solution to the PDE. We consider the terminal value problem

$$\begin{cases} \partial_t u(t,x) + \mathcal{L}u(t,x) - k(t,x)u(t,x) + g(t,x) = 0, \quad t < T, x \in \mathbb{R}^d, \\ u(T,x) = f(x). \end{cases} \tag{9}$$

We note that the probabilistic convention is to formulate this problem as a terminal value problem, as we have done here. However, the PDE convention would be to instead formulate it as an initial value problem. These two formulations are equivalent under a time-reversal $t \mapsto T - t$.

Under some conditions (see for example [2] for the specifics) on the functions f, g, k as well as on b and σ, and provided that the solution u to Equation 9 exists and is in $\mathcal{C}^{1,2}$, and satisfies some further conditions of continuity and boundedness, $u(t,x)$ is given by the Feynman-Kac formula

$$u(t,x) = \mathbb{E}\Big[f\big(X_T^{t,x}\big)e^{-\int_t^T k(r,X_r^{t,x})dr} + \int_t^T g(s,X_s^{t,x})e^{-\int_t^s k(r,X_r^{t,x})dr} ds\Big]. \tag{10}$$

Here $X_T^{t,x}$ denotes the stochastic process X at time T, started at x at time t.

We note that in particular, for $k = 0, g = 0$, the solution to the Kolmogorov backward equation (with terminal condition),

$$\begin{cases} \partial_t u(t,x) + \mathcal{L}u(t,x) = 0, \quad t < T, x \in \mathbb{R}^d, \\ u(T,x) = f(x), \end{cases} \tag{11}$$

is given by

$$u(t,x) = \mathbb{E}\big[f(X_T^{t,x})\big]. \tag{12}$$

Another PDE related to the SDE Equation 6 is the Kolmogorov forward equation

$$\partial_s p(s,y) - \mathcal{L}^* p(s,y) = 0, \tag{13}$$

where the differential operator $\mathcal{L}^*$ is the adjoint of $\mathcal{L}$, given by

$$\mathcal{L}^* = \frac{1}{2}\sum_{i,j=1}^d \partial_{y_i,y_j}[\sigma\sigma^T]_{i,j}(s,y) - \sum_{i=1}^d \partial_{y_i} b_i(s,y). \tag{14}$$

The Kolmogorov forward equation (with initial condition $p(0,y) = \delta(y - x)$) describes the probability distribution of the stochastic process X_t that solves the SDE in Equation 6. We can obtain a Feynman-Kac type formula for the density function $p(s,y)$ (assuming that X_t admits a density), by noting that

$$\mathbb{E}[f(X_T^{t,x})] = \int p_{x,t}(T,z)f(z)dz \tag{15}$$

Hence,

$$p(T,y) = \mathbb{E}[\delta(X_T^{t,x} - y)] = \int p_{t,x}(T,z)\delta(z - y)dz, \tag{16}$$

or more generally

$$p(s,y) = \mathbb{E}[\delta(X_s^{t,x} - y)] = \int p_{t,x}(s,z)\delta(z - y)dz. \tag{17}$$

There are some subtleties that have not been covered here — e.g. what is required for X_t to admit a density function, and the regularity required for the expectation of an indicator function to make sense. We also note that when writing $p(s,y)$, the probability density function of X being at a point y at time s, we are implicitly referring to the probability density conditioned on the initial distribution, i.e. in this case conditioned on X starting at position x at time t, for some $t < s$.

2.3 Forward Backward SDE and nonlinear Feynman-Kac formula

Now, we introduce the forward-backward SDE

$$\begin{cases} X_t &= x + \int_0^t b(s,X_s)ds + \int_0^t \sigma(s,X_s)dW_s \\ \\ Y_t &= f(X_T) + \int_t^T g\big(s,X_s,Y_s,Z_s[\sigma(s,X_s)]^{-1}\big)ds - \int_t^T Z_s dW_s, \end{cases} \tag{18}$$

where the first equation is the forward SDE — identical to Equation 6 — and the second equation is the backward SDE. We note that X_t depends on the values of X prior to time t, whilst Y_t depends on the values of X,Y,Z after time t (and up to time T).

It can be shown that the backward SDE above is related to the PDE (terminal value problem)

$$\partial_t u(t,x) + \mathcal{L}u(t,x) + g(t,x,u(t,x),\nabla u(t,x)) = 0 \tag{19}$$

$$u(T,x) = f(x), \tag{20}$$

where, as before, the operator $\mathcal{L}$ is the infinitesimal generator of the forward SDE. Specifically, if the solution to the terminal value problem exists, then the processes Y_t, Z_t given by

$$Y_t = u(t,X_t) \tag{21}$$

$$Z_t = \sigma(t,X_t)\nabla u(t,X_t) \tag{22}$$

satisfy the FBSDE of Equation 18. We also note that this statement can be extended to a system of k PDEs, and a vector valued stochastic process Y_t.

Similarly to in Equation 10, we can express the solution $u(t,x)$ to the terminal value problem Equation 19 in terms of a (now nonlinear) Feynman-Kac formula as

$$u(t,x) = \mathbb{E}\bigg[f(X_T^{t,x}) + \int_t^T g\big(s,X_s^{t,x},u(s,X_s^{t,x}),\nabla u(s,X_s^{t,x})\big)ds\bigg]. \tag{23}$$

Alternatively, we can write Y_t as

$$Y_t = u(t,X_t) = \mathbb{E}\bigg[f(X_T) + \int_t^T g\big(s,X_s,Y_s,Z_s\big)ds\Big|X_t\bigg]. \tag{24}$$

Comparing Equation 23 to Equation 10, we note that the function g now in general depends on not only t and X_t, but can also depend on u and ∇u. The discount (or attenuation) factor $k(t,x)$ in Equation 9 has been absorbed into the more general source term $g(t,x,u(t,x),\nabla u(t,x))$ in Equation 19. We also note that in Equation 10, only the left hand side depends on u, whilst in Equation 23 the right hand side also depends on u. Hence, to use the latter for numerical simulation of $u(t,x)$, more careful consideration is required.

We have now seen that the solution to the terminal value problem (backward PDE) of Equation 19 is associated with the backward SDE in Equation 18. This is a generalisation of the connection discussed in the previous subsection, and we note that by letting g depend on only

t and x, we can formulate a less general backward SDE and recover a version of the linear Feymnan-Kac formula of Equation 10, glossing over some subtleties related to the discount factor $k(t,x)$.

As before, the PDE directly associated with the forward SDE in Equation 18 is the Kolmogorov forward equation of Equation 13 with initial condition $p(0,y) = \delta(y-x)$. It is not clear whether a Feynman-Kac formula can be obtained for a more general PDE featuring the operator $\mathcal{L}^*$ defined in Equation 14 rather than the operator $\mathcal{L}$ of Equation 7.

2.4 FBSDEs and systems of PDEs

It is also possible to write down a more general version of Equation 18, where the process Y_t which solves the backward SDE is allowed to be K-dimensional. This process can then be used to write down a probabilistic representation of the solution to a system of K coupled PDEs (each with a terminal condition) of the type written down in Equation 19. For simplicity, here we consider the case where the source term g does not depend on the gradient of u. Specifically, we then seek a probabilistic representation of $u = (u^{[1]},...,u^{[K]})$, where

$$\begin{cases} \partial_t u^{[k]}(t,x) + \mathcal{L}^{[k]} u^{[k]}(t,x) + g^{[k]}(t,x,u(t,x)) = 0 \\ u^{[k]}(T,x) = f^{[k]}(x) \end{cases} \tag{25}$$

for $k = 1,2,...,K$, and $g = (g^{[1]},...,g^{[K]})$, $f = (f^{[1]},...,f^{[K]})$. We note that in general each component of g depends on the whole of u, and not just a particular component $u^{[k]}$.

Here, as before, $\mathcal{L}^{[k]}$ denotes the generator of a diffusion process, such as that described in the forward SDE of Equation 18. If the operator $\mathcal{L}^{[k]}$ is the same for each k, it is enough to consider one diffusion process X_t. However, if $\mathcal{L}^{[k]}$ is different for each k, we need K different diffusion processes $(X_t^{[1]},...,X_t^{[K]})$, with corresponding generators $(\mathcal{L}^{[1]},...,\mathcal{L}^{[K]})$. In general, each process $X_t^{[k]}$ is d-dimensional, i.e. of the same dimension as the variable x in the PDE in Equation 25.

For simplicity, here we consider the case where the $X_t^{[k]}$ are one dimensional, and have constant coefficients, but the below argument is easily generalised to the d-dimensional case with nonconstant coefficients. We then have

$$X_t^{[k]} = x + \int_0^t b_k ds + \int_0^t \sigma_k dW_s^{[k]} \tag{26}$$

each with corresponding generator

$$\mathcal{L}^{[k]} = \frac{1}{2}\sigma_k^2 \partial_x^2 + b_k \partial_x. \tag{27}$$

We now claim that the processes $Y_t = (Y_t^{[1]},...,Y_t^{[K]})$ and $Z_t = (Z_t^{[1]},...,Z_t^{[K]})$ defined by

$$Y_t^{[k]} = u^{[k]}(t,X_t^{[k]}) \tag{28}$$

$$Z_t^{[k]} = \sigma_k \partial_x u^{[k]}(t,X_t^{[k]}) \tag{29}$$

obey the backward SDE

$$Y_t^{[k]} = f^{[k]}(X_T^{[k]}) + \int_t^T g^{[k]}\big(s, X_s^{[k]},u(s,X_s)\big) ds - \int_t^T Z_s^{[k]} dW_s^{[k]}, \tag{30}$$

where $u(s,X_s) = \big(u^{[1]}(s,X_s^{[1]}),..., u^{[K]}(s,X_s^{[K]})\big)$.

We show this holds by (componentwise) applying Itô's formula

$$v(s,X_s) = v(s_0,X_{s_0}) + \int_{s_0}^{s} [\partial_t + \mathcal{L}]v(r,X_r)dr + \int_{s_0}^{s} \partial_x v(r,X_r)\sigma(r,X_r)dW_r, \tag{31}$$

where $\mathcal{L}$ is the generator of the diffusion process given by

$$X_s = X_{s_0} + \int_{s_0}^{s} b(r,X_r)ds + \int_{s_0}^{s} \sigma(r,X_r)dW_r, \tag{32}$$

to the process $Y_t = (Y_t^{[1]},...,Y_t^{[K]})$ defined by Equation 28, with $s_0 = t$ and $s = T$. By noting that $u^{[k]}(T,X_T^{[k]}) = f^{[k]}(X_T^{[k]})$ the result follows.

To obtain a Feynman-Kac formula for $u(t,x)$ we take a conditional expectation of Equation 30. What to condition on is slightly subtle. By considering that the equations for $u^{[k]}(t,x)$ for each k must be equations of the same variable x, and are coupled through the function g, we can see that a sensible thing to condition on is $X_t^{[1]} = x,...,X_t^{[K]} = x$. In other words, we start each of the diffusions $X_s^{[k]}$ at the same point x at time t. We then get the Feynman-Kac formula

$$u^{[k]}(t,x) = \mathbb{E}\left[f^{[k]}(X_T^{[k],x,t}) + \int_t^T g^{[k]}\big(s, X_s^{[k],x,t}, u(s,X_s^{x,t})\big)ds\right], \tag{33}$$

where $X_T^{[k],x,t}$ denotes the process $X_s^{[k]}$ started at x at time t, and $X_s^{x,t}$ denotes the (k-dimensional) $(X_s^{[1]},...,X_s^{[K]})$ started at $(x,...,x)$ at time t.

Similarly to before, we can also write down a formula for $Y_t^{[k]}$ as

$$Y_t^{[k]} = \mathbb{E}\left[f^{[k]}(X_T^{[k]}) + \int_t^T g^{[k]}\big(s,X_s^{[k]},Y_s\big)ds \middle| X_t^{[1]},...,X_t^{[K]}\right], \tag{34}$$

by conditioning on the processes $(X_t^{[1]},...,X_t^{[K]})$, at time t, instead of on $(X_t^{[1]} = x,...,X_t^{[K]} = x)$.

2.5 Numerical schemes for FBSDEs

To solve the FBSDE Equation 18 numerically, we need to discretise both the forward and the backward SDE in time. Discretising the forward SDE is straightforward, and can for example be done using the (forward) Euler-Maruyama scheme

$$\begin{cases} X_0^h = x \\ X_{(i+1)h}^{(h)} = X_{ih}^{(h)} + b(ih,X_{ih}^{(h)})h + \sigma(ih,X_{ih}^{(h)})(W_{(i+1)h} - W_{ih}). \end{cases} \tag{35}$$

We will use this discretised version of X_t in the scheme for $Y_t = u(t,X_t)$, or similarly in the scheme for $u(t,x)$. For simplicity, we will here consider the case where g is a function of t,x, and u, but not of ∇u.

By using the tower property of expectation on Equation 24, we can express Y_{t_i} in terms of $Y_{t_{i+1}}$, and in terms of X_t between t_i and t_{i+1} as

$$Y_{t_i} = \mathbb{E}\left[Y_{t_{i+1}} + \int_{t_i}^{t_{i+1}} g(s,X_s,Y_s)ds \middle| X_{t_i}\right]. \tag{36}$$

We recall that we are solving for Y_t backwards in time, so for $t_i < t_{i+1}$, $Y_{t_{i+1}}$ is known.

From the above, we can get the (backward) Euler scheme for Y, namely

$$\begin{cases} Y_T^{(h)} = f(X_T^h) \\ Y_{ih}^{(h)} = \mathbb{E}\left[Y_{(i+1)h}^{(h)} + hg(t_i, X_{ih}^{(h)}, Y_{(i+1)h}^{(h)}) \middle| X_{t_i}^{(h)}\right]. \end{cases} \tag{37}$$

By conditioning on the discretised process $X^{(h)}$ starting at a specific value x at time t_i, we get the same scheme for $u(t,x)$:

$$\begin{cases} u^{(h)}(T,x) = f(x) \\ u^{(h)}(t_i,x) = \mathbb{E}\left[u^{(h)}(t_{i+1}, X_{t_{i+1}}^{(h),t_i,x}) + hg\big(t_i, x, u^{(h)}(t_{i+1}, X_{t_{i+1}}^{(h),t_i,x})\big)\right]. \end{cases} \tag{38}$$

We note here that the accuracy of a numerical simulation of $u(t,x)$ (or Y_t) depends on both the scheme we choose for X_t, and the scheme we choose for Y_t. In [1] a strong stability preserving multistep scheme, which improves the latter, is introduced. This scheme is given by

$$Y_{ih}^{(h)} = \sum_{j=1}^{k} \alpha_j \mathbb{E}\left[Y_{(i+j)h}^{(h)} \middle| X_{t_i}^{(h)}\right] + h \sum_{j=1}^{k} \beta_j \mathbb{E}\left[g\big(t_{i+j}, X_{(i+j)h}^{(h)}, Y_{(i+j)h}^{(h)}\big) \middle| X_{t_i}^{(h)}\right]. \tag{39}$$

The corresponding scheme for $u(t,x)$ becomes

$$u^{(h)}(t_i,x) = \sum_{j=1}^{k} \alpha_j \mathbb{E}\left[u^{(h)}\Big(t_{i+j}, X_{t_{i+j}}^{(h),t_i,x}\Big)\right] + h \sum_{j=1}^{k} \beta_j \mathbb{E}\left[g\Big(t_{i+j}, X_{t_{i+j}}^{(h),t_i,x}, u^{(h)}\big(t_{i+j}, X_{t_{i+j}}^{(h),t_i,x}\big)\Big)\right]. \tag{40}$$

2.6 Connecting the treatment planning problem and the FBSDE

Here, we seek to connect the coupled PDEs that arise from the treatment planning problem to the framework of FBSDEs.

The PDE system arising from the treatment planning problem (with $u^{[1]}$ denoting the primal, and $u^{[2]}$ the dual) is given by

$$\begin{cases} \partial_t u^{[1]} - \mu \Delta u^{[1]} + b \cdot \nabla u^{[1]} & = \frac{1}{\alpha} u^{[2]} \\ -\partial_t u^{[2]} - \mu \Delta u^{[2]} - b \cdot \nabla u^{[2]} & = u^{[1]} - d_T. \end{cases} \tag{41}$$

We also have the boundary condition $u^{[1]} = u^{[2]} = 0$ on the boundary. For simplicity, we consider only one spatial dimension

$$\begin{cases} \partial_t u^{[1]} - \mu \partial_x^2 u^{[1]} + b \partial_x u^{[1]} & = \frac{1}{\alpha} u^{[2]} \\ -\partial_t u^{[2]} - \mu \partial_x^2 u^{[2]} - b \partial_x u^{[2]} & = u^{[1]} - d_T. \end{cases} \tag{42}$$

Rewriting this to be of the form of Equation 25:

$$\begin{cases} \partial_t u^{[1]} - \mu \partial_x^2 u^{[1]} + b \partial_x u^{[1]} - \frac{1}{\alpha} u^{[2]} & = 0 \\ \partial_t u^{[2]} + \mu \partial_x^2 u^{[2]} + b \partial_x u^{[2]} - (u^{[1]} - d_T) & = 0. \end{cases} \tag{43}$$

This is equivalent to

$$\begin{cases} \partial_t u^{[1]} + \mathcal{L}^{[1]} u^{[1]} + g^{[1]} & = 0 \\ \partial_t u^{[2]} + \mathcal{L}^{[2]} u^{[2]} + g^{[2]} & = 0 \end{cases} \tag{44}$$

for

$$\mathcal{L}^{[1]} = -\mu\partial_x^2 + b\partial_x \tag{45}$$

$$\mathcal{L}^{[2]} = \mu\partial_x^2 + b\partial_x \tag{46}$$

and

$$g^{[1]} = -\frac{1}{\alpha}u^{[2]} \tag{47}$$

$$g^{[2]} = -(u^{[1]} - d_T). \tag{48}$$

Questions/uncertainties:

- $\mathcal{L}^{[1]}$ cannot be the generator of a diffusion, as this would require $\frac{1}{2}\sigma^2 = -\mu$, i.e. a diffusion coefficient $\sigma = \sqrt{-2\mu}$, which would be imaginary
- If we could time-reverse the forward equation (change the sign of the ∂_t-term) we'd maybe get the form we want (since the sign of the drift coefficient b doesn't matter), but does this actually make sense to do?
- In the PDEs resulting from the treatment planning problem, is the primal forward in time and the dual backward in time? If so, maybe it would make sense to time-reverse the primal equation, since the FBSDE framework works with multiple equations backward in time (with terminal conditions)?
- If the above works out, what should the terminal condition(s)/initial condition(s) be? I.e. what are the functions $f^{[k]}(x)$?

References

[1] Shuixin Fang, Weidong Zhao, and Tao Zhou. "Strong Stability Preserving Multistep Schemes for Forward Backward Stochastic Differential Equations". In: *Journal of Scientific Computing* 94.3 (2023), p. 53.

[2] Emmanuel Gobet. *Monte-Carlo methods and stochastic processes: from linear to non-linear.* CRC Press, 2016.

[3] Jacopo Werther. *Dose Depth Curves.* 2010. URL: `https://commons.wikimedia.org/wiki/File:Dose_Depth_Curves.svg` (visited on 05/15/2023).