MTH 9831, FALL 2015. LECTURE 8

Abstract. Connections with PDEs.

1. SDE: definition, Markov property of solutions, associated differential operators.
2. Feynman-Kac formula.
3. Kolmogorov's backward and forward equations.

1 Stochastic Differential Equations

Definition 1.1. *Let $b(t,x)$ be a deterministic d-dimensional vector valued function on $[0,\infty)\times\mathbb{R}^d$ and $\sigma(t,x)$ be a deterministic $d\times r$ matrix for all $(t,x)\in[0,\infty)\times\mathbb{R}^d$. Denote by $B(t)$ the r-dimensional standard Brownian motion. A stochastic differential equation with drift $b(t,x)$ and volatility $\sigma(t,x)$ is an equation of the form*

$$dX(t) = b(t,X(t))dt + \sigma(t,X(t))dB(t). \tag{1.1}$$

A strong solution to 1.1 with initial condition $x\in\mathbb{R}^d$ is a stochastic process $X(t)_{t\geq 0}$ with continuous sample paths and with the following properties

1. $X(0)=x$, a.s.
2. $X(t)_{t\geq 0}$ is adapted to filtration $\mathcal{F}(t)_{t\geq 0}$ generated by $B(t)_{t\geq 0}$.
3. $\int_0^t \left[|B_i(s,X(s))| + \sigma_{ij}^2(s,X(s))\right] ds < \infty$, a.s. for $1\leq i\leq d$, $1\leq j\leq r$, $0\leq t<\infty$.
4. For $t\in[0,T]$, $X(t) = X(0) + \int_0^t B(u,X(u))du + \int_0^t \sigma(u,X(u))dB(u)$, a.s.

Remark. Often in applications we need to start at time t from a given point $x\in\mathbb{R}^d$ and find X: $X(t)=x$ and

$$X(u) = X(t) + \int_t^u b(s,X(s))ds + \int_t^u \sigma(s,X(s))dB(s), \quad \text{a.s. for } u\geq t.$$

Some conditions on $b(t,x)$ and $\sigma(t,x)$ are needed to ensure the existence and uniqueness of solutions of 1.1.

Example 1.1. $dX(t) = X^3(t)dt - X^2(t)dB(t)$, $X(0)=1$. Let us try to find $X(t)$ in the form $f(t,B(t))$. Then

$$dX(t) = df(t,B(t)) = f_t dt + f_x dB(t) + \frac{1}{2}f_{xx}dt = \left(f_t + \frac{1}{2}f_{xx}\right)dt + f_x dB(t),$$

where f has to satisfy: $f_x = -f^2$ and $f_t + \frac{1}{2}f_{xx} = f^3$.

$$f_x = -f^2 \Rightarrow f(t,x) = -\frac{1}{x+C(t)}.$$

Substituting in the second equation gives

$$\begin{aligned}
&-\frac{C'(t)}{(x+C(t))^2} + \frac{1}{2}\frac{2}{(x+C(t))^3} = \frac{1}{(x+C(t))^3}\\
&\Rightarrow C'(t)=0 \Leftrightarrow C(t)\equiv \text{const}\\
&\Rightarrow f(t,x) = \frac{1}{x+C}.
\end{aligned}$$

Initial condition $X(0) = f(0,0) = \frac{1}{C} \Rightarrow C=1$. Thus,

$$X(t) = f(t,B(t)) = \frac{1}{B(t)+1}.$$

The problem is that with probability 1 $B(t)$ hits -1, i.e. $\mathbb{P}(\tau_{-1}<\infty)=1$. Thus, the solution exists only up to the "explosion time" τ_{-1}.

Typical conditions which prevent explosions and guarantee existence and uniqueness of 1.1 are:

$$\begin{aligned}
&|b(t,x)-b(t,y)| + \| \sigma(t,x)-\sigma(t,y) \| \leq K|x-y|,\\
&|b(t,x)|, \| \sigma(t,x) \| \leq K(1+|x|),
\end{aligned}$$

for all $t \geq 0$, $x, y \in \mathbb{R}$. Note, these are Lipshitz conditions. In fact, for $d = 1$ the conditions can be weakened (for example, CIR is included). Here $\| A = \| \sqrt{\sum_{i,j=1}^{d} a_{ij}^2}$.

Markov property. Intuition at $d = 1$. How can one try to simulate $X(t)_{0 \leq t \leq T}$? Suppose $X(0) = x$ is given. Consider a small increment of time δ. Then according to 1.1,

$$\begin{aligned} X(\delta) &\approx x + b(0,x)\delta + \sigma(0,x) \underbrace{B(\delta)}_{\sim N(0,\delta)} \\ &= x + b(0,x)\delta + \sigma(0,x)\sqrt{\delta} Z_1, \quad Z_1 \sim N(0,1). \end{aligned}$$

Then

$$\begin{aligned} X((k+1)\delta) &= X(k\delta) + b(k\delta, X(k\delta))\delta + \sigma(k\delta, X(k\delta))(B((k+1)\delta) - B(k\delta)) \\ &= X(k\delta) + b(k\delta, X(k\delta))\delta + \sigma(k\delta, X(k\delta))\sqrt{\delta} Z_{k+1}, \end{aligned}$$

and $Z_1, Z_2, \cdots, Z_k, \cdots$ are i.i.d. $N(0,1)$. Then $X(n\delta) = X(T)$ gives us a simulated value at time T. If we could justify passing to the limit (in some sense) as $\delta \to 0$, then the limiting process would give us a solution to 1.1. Moreover, our procedure suggests that the limiting process would be memoriless ($\equiv$ Markov process), since given $X(k\delta)$ we can simulate $X((k+1)\delta), \cdots, X(n\delta)$ without knowing $X(0), X(\delta), \cdots, X((k-1)\delta)$. More rigorously, the following theorem holds:

Theorem 1.1. *Let $0 \leq t \leq T$ and h be a Borel measurable function. Set*

$$g(t,x) = E^{t,x} h\left(X(T)\right) = E\left[h(X(T)) | X(t) = x\right].$$

In other words, take $X(t) = x$, solve 1.1 for $u \in [t,T]$, get $X(T)$, and compute the expectation of $h(X(T))$, repeat for all $t \in [0,T]$, $x \in \mathbb{R}$. Get a deterministic function $g(t,x)$. Then for $t \in [0,T]$,

$$\mathbb{E}\left[h(X(T)) | \mathcal{F}(t)\right] = g(t, X(t)),$$

i.e. solutions to SDEs are Markov process.

Remark. There exists a Borel-measurable function $g(t,x)$ such that $\mathbb{E}\left[h(X(T)) | \mathcal{F}(t)\right] = g(t, X(t))$. Moreover, $g(t,x) = E^{t,x} h\left(X(T)\right) = E\left[h(X(T)) | X(t) = x\right]$.

Proof. Proof is omitted. Heuristics: meanings of $b(t,x)$ and $\sigma^2(t,x)$ at $d = 1$.

$$\begin{aligned} &E\left[X(t+\delta) - X(t) | X(t) = x\right] \approx E\left[b(t, X(t))\delta + \sigma(t, X(t))\sqrt{\delta} Z | X(t) = x\right] = b(t,x)\delta. \\ &\Rightarrow b(t,x) = \lim_{\delta \to 0} \frac{E\left[X(t+\delta) - X(t) | X(t) = x\right]}{\delta}. \\ &E\left[(X(t+\delta) - X(t))^2 | X(t) = x\right] \approx E\left[\left(b(t, X(t))\delta + \sigma(t, X(t))\sqrt{\delta} Z\right)^2 \Big| X(t) = x\right] \\ &= b^2(t,x)\delta^2 + \delta\sigma^2(t,x) \\ &\Rightarrow a(t,x) \triangleq \sigma^2(t,x) = \lim_{\delta \to 0} \frac{E\left[(X(t+\delta) - X(t))^2 | X(t) = x\right]}{\delta}. \end{aligned}$$

Here $b(t,x)$ is called the drift coefficient, $a(t,x)$ is called the diffusion coefficient. These heuristics associate the coefficients in SDE with the PDE. $\square$

If $u(t,x)$ is a smooth function then

$$\begin{aligned} &E\left[u(t+\delta, X(t+\delta)) - u(t, X(t)) | X(t) = x\right] \\ &= E\left[u_t(t, X(t))\delta + u_x(t, X(t))(X(t+\delta) - X(t)) + \frac{1}{2} u_{xx}(t, X(t))(X(t+\delta) - X(t))^2 + \cdots \Big| X(t) = x\right] \\ &= \delta\left(u_t(t, X(t)) + b(t,x) u_x(t, X(t)) + \frac{1}{2} u_{xx}(t,x) a(t,x)\right) + o(\delta). \end{aligned}$$

Thus,

$$\begin{aligned} &\lim_{\delta \to 0} \delta^{-1}\left[u(t+\delta, X(t+\delta)) - u(t, X(t)) | X(t) = x\right] \\ &= u_t(t, X(t)) + b(t,x) u_x(t, X(t)) + \frac{1}{2} u_{xx}(t,x) a(t,x) \\ &\triangleq u_t(t, X(t)) + (\mathcal{A}_t u)(t,x), \end{aligned}$$

where

$$\mathcal{A}_t \triangleq b(t,x)\frac{\partial}{\partial x} + \frac{1}{2}a(t,x)\frac{\partial^2}{\partial x^2}, \quad t \geq 0. \tag{1.2}$$

Similarly, for $d \geq 1$, we set

$$\mathcal{A}_t \triangleq \sum_{i=1}^{d} b_i(t,x)\frac{\partial}{\partial x_i} + \frac{1}{2}\sum_{i,j=1}^{d} a_{ij}(t,x)\frac{\partial^2}{\partial x_i \partial x_j}, \quad t \geq 0, \tag{1.3}$$

where

$$\begin{aligned} b^T(t,x) &= \lim_{\delta \to 0} \frac{E\left[(X(t+\delta) - X(t))^T | X(t) = x\right]}{\delta}, \\ a(t,x) &= \lim_{\delta \to 0} \frac{E\left[(X(t+\delta) - X(t))(X(t+\delta) - X(t))^T | X(t) = x\right]}{\delta} = \left(\sigma\sigma^T\right)(t,x). \end{aligned}$$

Remark. Given $a(t,x)$, σ is not unique: $(\sigma O)(\sigma O)^T = \sigma O O^T \sigma^T = \sigma\sigma^T$, $\forall$ orthogonal matrix O. When b and σ do not depend on t, i.e. $b = b(x)$, $\sigma = \sigma(x)$, X is called a (time homogeneous, or Itô) diffusion process, and $\mathcal{A}$ (now it does not depend on t) is called the *generator* of X.

2 Feynman-Kac Formula

This formula gives a stochastic representation of a solution to a PDE. More precisely,

Theorem 2.1 (Feynman-Kac formula). *Let $g(t,x)$ be the solution of*

$$\begin{aligned} & g_t + b(t,x)g_x + \frac{1}{2}\sigma^2(t,x)g_{xx} = r(t,x)g, \\ & g(T,x) = h(x). \end{aligned}$$

Let $X(u)_{t \leq u \leq T}$ solve 1.1 with $X(t) = x$. Then

$$g(t,x) = E^{t,x}\left[e^{-\int_t^T r(s,X(s))ds} h(X(T))\right].$$

We assumed that some mild integrability conditions are satisfied so that all integrals make sense.

Proof. The idea of the proof. For $0 \leq t \leq u \leq T$, we compute (t is fixed and u is changing)

$$\begin{aligned} & d\left(e^{-\int_t^u r(s,X(s))ds} g(u,X(u))\right) \\ & = -r(u,X(u))e^{-\int_t^u r(s,X(s))ds} g(u,X(u))du + e^{-\int_t^u r(s,X(s))ds} dg(u,X(u)) \\ & = e^{-\int_t^u r(s,X(s))ds} \Bigg[g_x(u,X(u))\sigma(u,X(u))dB(u) + \\ & \left(-r(u,X(u))g(u,X(u)) + g_t(u,X(u)) + b(u,X(u))g_x(u,X(u)) + \frac{1}{2}\sigma^2(u,X(u))g_{xx}(u,X(u))\right) du \Bigg]. \end{aligned}$$

Since g solves the PDE, the drift term is 0. Then

$$e^{-\int_t^T r(s,X(s))ds} g(T,X(T)) - g(t,X(t)) = \int_t^T e^{-\int_t^u r(s,X(s))ds} g_x(u,X(u))\sigma(u,X(u))dB(u).$$

Taking conditional expectation (given $X(t) = x$) we get

$$E^{t,x}\left[e^{-\int_t^T r(s,X(s))ds} h(X(T))\right] = g(t,x).$$

□

Theorem 2.2 (Multidimensional Feynman-Kac formula). *Let $a(t,x) = \sigma\sigma^T(t,x)$ and assume that $g(t,x)$ satisfies on $[0,T) \times \mathbb{R}^d$*

$$g_t + \sum_{j=1}^{d} b_j(t,x)\frac{\partial g}{\partial x_j} + \frac{1}{2}\sum_{i,j=1}^{d} a_{ij}(t,x)\frac{\partial^2 g}{\partial x_i \partial x_j} = r(t,x)g$$

with the terminal condition $g(T,x) = h(x)$, $x \in \mathbb{R}^d$. Then (under integrability conditions)

$$g(t,x) = \mathbb{E}^{t,x}\left[e^{-\int_t^T r(s,X(s))ds} h(X(T))\right],$$

where $X(u)_{t \le u \le T}$ solves 1.1 with $X(t) = x$, i.e. $\forall j = 1,2,\cdots,d$,

$$dX_j(u) = b_j(u,X(u))du + \sum_{k=1}^{d} \sigma_{jk}(u,X(u))dB_k(u),$$

and $X_j(t) = x$, $j = 1,2,\cdots,d$, $t \le u \le T$.

Example 2.1 (Heston Stochastic Volatility Model $d=2$)**.** Suppose that under $\tilde{P}$ the stock price follows

$$dS(t) = rS(t)dt + \sqrt{V(t)}S(t)d\tilde{B}_1(t),$$

where $V(t)$ is a stochastic process and

$$dV(t) = (a - bV(t))dt + \sigma\sqrt{V(t)}d\tilde{B}_2(t).$$

Here $a, b, \sigma > 0$ and $d[\tilde{B}_1, \tilde{B}_2](t) = \rho dt$, $-1 < \rho < 1$.

- Find the corresponding generator $\mathcal{A}$.
- Waht does the F.K. formula say about the solution of $\frac{\partial g}{\partial t} + \mathcal{A}g = rg$, $g(T,x,v) = h(x,v)$.

It is easy to find the drift $b(t,x,v) = \begin{pmatrix} rx \\ a - bv \end{pmatrix}$. There are two ways to find $\| a_{ij}(t,x,v) \|$.

I. Find $\sigma(t,x,v)$ and compute $\sigma\sigma^T(t,x,v)$. For this we have to "decorrelate" our Brownian motion

$$\begin{aligned} d\tilde{B}_1(t) &= d\tilde{W}(t), \\ d\tilde{B}_2(t) &= \rho d\tilde{W}_1(t) + \sqrt{1-\rho^2}d\tilde{W}_2(t), \end{aligned}$$

where $(\tilde{W}_1(t), \tilde{W}_2(t))$ is a standard 2-dimensional Brownian motion, and

$$\begin{aligned} dS(t) &= rS(t)dt + \sqrt{V(t)}S(t)d\tilde{W}_1(t), \\ dV(t) &= (a - bV(t))dt + \sigma\sqrt{V(t)}\left(\rho d\tilde{W}_1(t) + \sqrt{1-\rho^2}d\tilde{W}_2(t)\right). \end{aligned}$$

Thus,

$$\sigma(t,x,v) = \begin{Vmatrix} \sqrt{v}x & 0 \\ \rho\sigma\sqrt{v} & \sqrt{1-\rho^2}\sigma\sqrt{v} \end{Vmatrix} \Rightarrow \sigma\sigma^T(t,x,v) = \begin{Vmatrix} vx^2 & \sigma xv\rho \\ \sigma xv\rho & \sigma^2 v \end{Vmatrix}.$$

II. Compute directly $\| a_{ij}(t,x,v) \|$.

$$\begin{aligned} d[S,S](t) &= V(t)S^2(t)dt \Rightarrow a_{11}(t,x,v) = vx^2, \\ d[S,V](t) &= \sigma V(t)S(t)\rho dt \Rightarrow a_{12}(t,x,v) = a_{21}(t,x,v) = \sigma\rho vx, \\ d[V,V](t) &= \sigma^2 V(t)dt \Rightarrow a_{22}(t,x,v) = \sigma^2 v. \end{aligned}$$

Why does this work? Recall our heuristics:

$$E\left[(S(t+\delta) - S(t))^2 \,|S(t) = x, V(t) = v\right] \approx E\left[\left(rS(t)\delta + \sqrt{V(t)}S(t)\sqrt{\delta}Z\right)^2 |S(t) = x, V(t) = v\right]$$

$$\approx vx^2\delta + o(\delta) \Rightarrow a_{11}(t,x,v) = vx^2,$$

$$E\left[(S(t+\delta) - S(t))\,(V(t+\delta) - V(t))\,|S(t) = x, V(t) = v\right]$$

$$\approx E\left[\left(rS(t)\delta + \sqrt{V(t)}S(t)\sqrt{\delta}Z_1\right)\left((a - bV(t))\delta + \sigma\sqrt{V(t)}\sqrt{\delta}Z_2\right)\Big|S(t) = x, V(t) = v\right]$$

$$\approx \sigma vx\delta \underbrace{E\left[Z_1 Z_2\right]}_{=\rho} + o(\delta) \Rightarrow a_{12}(t,x,v) = \sigma vx\rho.$$

Finally, $a_{22}(t,x,v)$ is determined similarly. Thus,

$$\mathcal{A} \triangleq rx\frac{\partial}{\partial x} + (a - bv)\frac{\partial}{\partial v} + \frac{1}{2}vx^2\frac{\partial^2}{\partial x^2} + \sigma\rho vx\frac{\partial^2}{\partial x \partial v} + \frac{1}{2}\sigma^2 v\frac{\partial^2}{\partial v^2}.$$

By the F.K. formula $g(t,x,v) = \mathbb{E}^{t,x,v}\left[h(S(T), V(T))\right]$ represents the solution to

$$\frac{\partial g}{\partial t} + \mathcal{A}g = 0$$

with the terminal condition $g(T,x,v) = h(x,v)$.

Main point: we have an SDE 1.1 and its solution X such that $X(t) = x$. For a given T, $X(T)$ is a random variable whose distribution depends on parameters of 1.1 and also on t, x.

- $X(T)$ has a density denoted by $p(t, x; T, y)$. Here t, x, T are fixed parameters, and to get the density of $X(T)$ you have to consider it as a function of y. Question: how to find this density?
- Variables (t, x) are called *backward variables* as they correspond to the starting point of the process X: $X(t) = x$.
- Variables (T, y) are called *forward variables* as y is the *location* of the process at a future time T.
- $X(u)_{t \le u \le T}$ is a Markov process, and $p(t, x; T, y)$ is its transition probability density, that is

$$P\left(X(T) \in B | X(t) = x\right) = \int_B p(t, x; T, y) dy, \quad \forall B \in \mathcal{B}.$$

Remark. When the coefficients b and σ do not depend on t, that is when $b = b(x)$ and $\sigma = \sigma(x)$, then p is a function of $(T - t), x, y$ (as for Brownian motion).

- All in all $p(t, x; T, y)$ has 2 pairs of variables (t, x) - backward and (T, y) - forward.
- Kolmogorov *backward* equation is the PDE to which $p(t, x; T, y)$ is a solution (with some special terminal data) when T, y are treated as fixed parameteres and t, x as variables.
- Kolmogorov *forward* equation is the PDE to which $p(t, x; T, y)$ is a solution (with some special initial data) when x, t are treated as fixed parameters and T, y as variables. In particular, the density of $X(T)$ satisifies the forward equation in variables T, y.

Example 3.1 (Heat equation)**.**

$$\frac{\partial g}{\partial t} + \frac{1}{2}\frac{\partial^2 g}{\partial x^2} = 0, \quad g(x, T) = h(x) \tag{3.1}$$

Recognize $b \equiv 0, a \equiv 1 \Rightarrow$ the corresponding SDE is

$$dX(u) = dB(u), \quad X(t) = x.$$

Thus, $X(u) = x + B(u) - B(t)$, $t \le u \le T$. Generateor $\mathcal{A} = \frac{\partial^2}{\partial x^2}$, and

$$g(t, x) = \mathbb{E}^{t,x} h\left(X(T)\right) = \mathbb{E}^{t,x} h\left(x + B(T) - B(t)\right).$$

is a solution to 3.1. Since $B(T) - B(t) \sim N(0, T - t)$, we can compute $B(T) - B(t) \overset{\mathrm{d}}{=} \sqrt{T - t}Z$:

$$g(t, x) = \int_{-\infty}^{+\infty} h\left(x + \sqrt{T - t}z\right) \frac{1}{\sqrt{2\pi}} e^{-\frac{1}{2}z^2} dz. \tag{3.2}$$

Look at it now from the perspective of $p(t, x; T, y)$,

$$g(t, x) = \mathbb{E}^{t,x} h(X(T)) = \int_{-\infty}^{+\infty} h(y) p(t, x; T, y) dy. \tag{3.3}$$

- If $h(x) = \mathbf{1}_B(x)$ for some set B, then

$$g(t, x) = \int_B p(t, x; T, y) dy = P\left(X(T) \in B | X(t) = x\right) = \mathbb{E}^{t,x}\left[\mathbf{1}_B(X(T))\right] \quad \sim \text{F.K. formula}$$

- Comparing 3.2 and 3.3 (setting $y = x + \sqrt{T - t}z$ in 3.2), we get that

$$p(t, x; T, y) = \frac{1}{\sqrt{2\pi(T - t)}} e^{-\frac{(x-y)^2}{2(T-t)}},$$

the transition probability density of Brownian motion, where $p(t, x; T, y)dy$ is the probability that the Brownian motion that started at x at time t will be in $(y, y + dy)$ at time T.

Theorem 3.1 (Kolmogorov Backward Equation)**.** *Let X solve 1.1 on $[t, T]$ and $x(t) = x$. Let*

$$p(t, x; T, y) = \lim_{|B| \to 0} \frac{1}{|B|} P\left(X(T) \in B | X(t) = x\right)$$

where $|B|$ is the Lebesgue measure of set $B \in \mathcal{B}$. Then for $x \in \mathbb{R}^d$, $0 \le t < T$,

$$\frac{\partial p}{\partial t} + \mathcal{A}_t p = 0, \quad \textit{in variables } t, x \tag{3.4}$$

with the terminal condition $p(T, x; T, y) = \delta_y(x)$. The latter means that $\forall f \in C_B(\mathbb{R})$,

$$\lim_{t \nearrow T} \int f(x) p(t, x; T, y) dx = f(y),$$

or equivalently, that for every $B \in \mathcal{B}$,

$$\lim_{t \nearrow T} \int_B p(t, x; T, y) dx = \mathbf{1}_B(y).$$

Equation 3.4 is Kolmogorov backward equation, $\mathcal{A}$ is given by 1.2.

Theorem 3.2 (Kolmogorov Forward Equation)**.** *Define the adjoint operator*

$$\mathcal{A}_T^* f(T, y) = -\frac{\partial}{\partial y}\left(b(T, y) f(T, y)\right) + \frac{1}{2}\frac{\partial^2}{\partial y^2}\left(\sigma^2(T, y) f(T, y)\right).$$

Then $p(t, x; T, y)$ satisfies for $T > t$, $x \in \mathbb{R}^d$,

$$\frac{\partial p}{\partial T} = \mathcal{A}^* p, \quad \textit{in variables } T, y$$

with the initial condition $p(t, x; t, y) = \delta_x(y)$, meaning that $\forall f \in \mathcal{C}_B(\mathbb{R})$,

$$\lim_{T \searrow t} \int_{\mathbb{R}} f(y) p(t, x; T, y) dy = f(x),$$

or equivalently, that $\forall B \in \mathcal{B}$,

$$\lim_{T \searrow t} \int_B p(t, x; T, y) dy = \mathbf{1}_B(x).$$

Example 3.2. An example to have in mind: $X(t)$ is a standard Brownian motion (d-dimensional), $\mathcal{A}_t = \frac{1}{2}\Delta_x = \frac{1}{2}\sum_{i=1}^d \frac{\partial^2}{\partial x_i^2}$,

$$p(t, x; T, y) = \frac{1}{(2\pi(T-t))^{d/2}} \exp\left(-\frac{1}{2(T-t)}\sum_{i=1}^d (y_i - x_i)^2\right).$$

Then the backward equation: $\frac{\partial p}{\partial t} + \frac{1}{2}\Delta_x p = 0$, $0 \le t < T$, $x \in \mathbb{R}^d$, set $z_i = \frac{x_i - y_i}{\sqrt{T-t}}$,

$$\lim_{t \nearrow T} \int_{\mathbb{R}^d} f(x) p(t, x; T, y) dx = \lim_{t \nearrow T} \int_{\mathbb{R}^d} f(y + \sqrt{2(T-t)} z) \frac{1}{(2\pi)^{d/2}} e^{-\frac{\|z\|^2}{2}} dz = f(y),$$

for all continuous bounded functions f. Computing the adjoint operator we get that, in this case,

$$\mathcal{A}_T^* = \frac{1}{2}\Delta_y = \frac{1}{2}\sum_{i=1}^d \frac{\partial^2}{\partial y_i^2},$$

and the forward equation is

$$\frac{\partial p}{\partial T} = \frac{1}{2}\Delta_y p, \quad t < T, x \in \mathbb{R}^d,$$

and set $z_i = \frac{y_i - x_i}{\sqrt{T-t}}$,

$$\lim_{T \searrow t} \int_{\mathbb{R}^d} f(y) p(t, x; T, y) = \lim_{T \searrow t} \int_{\mathbb{R}^d} f(x + \sqrt{T-t} z) \frac{1}{(2\pi)^{d/2}} e^{-\frac{\|z\|^2}{2}} dz = f(x)$$

for all $f \in \mathcal{C}_B(\mathbb{R}^d)$.