AM216 Stochastic Differential Equations

Lecture 04

List of topics in this lecture

- Properties of Wiener process, $dW = O(\sqrt{dt})$
- Discrete version of $W(t)$; arc length of $W(t)$ over finite time is infinity!
- Ito's lemma; $(dW)^2$ can be replaced by dt.
- The gambler's ruin problem; applications of Ito's lemma, law of total probability, law of total expectation; survival probability as a function of (initial cash, time)

Recap

Translation and scaling of normal RVs

If $X \sim N(\mu, \sigma^2)$, then $\frac{X-\mu}{\sigma} \sim N(0,1)$, which is called a standard normal RV.

Theorem:

Sum of independent normal RVs is a normal RV.

If $X \sim N(\mu_1, \sigma_1^2)$ and $Y \sim N(\mu_2, \sigma_2^2)$ are independent, then $(X+Y) \sim N(\mu_1+\mu_2, \sigma_1^2+\sigma_2^2)$.

Stochastic differential equation (SDE)

$$dX = b(X,t)dt + \sqrt{a(X,t)}\,dW$$

Notations: $dW \equiv W(t+dt) - W(t)$, $dX \equiv X(t+dt) - X(t)$

The Wiener process, denoted by $W(t)$, satisfies

1) $W(0) = 0$
2) For $t_2 \ge t_1 \ge 0$, $W(t_2)-W(t_1) \sim N(0, t_2-t_1)$
3) For $t_4 \ge t_3 \ge t_2 \ge t_1 \ge 0$,

   increments $W(t_2)-W(t_1)$ and $W(t_4)-W(t_3)$ are independent.

Note: $W(t)$ is a stochastic process. The full notation is $W(t, \omega)$.

Complication of SDE

Ordinary Difference Equation:

$$\Delta X = b(X,t)\Delta t + o(\Delta t)$$

==> $$\lim_{\Delta t \to 0} \frac{\Delta X}{\Delta t} = b(X,t) + \lim_{\Delta t \to 0} \frac{o(\Delta t)}{\Delta t}$$

==> $$\frac{dX}{dt} = b(X,t)$$

We can work with derivatives, instead of differences.

Stochastic Difference Equation:

$$\Delta X = b(X,t)\Delta t + \sqrt{a(X,t)}\,\Delta W + o(\Delta t)$$

==> $$\lim_{\Delta t \to 0} \frac{\Delta X}{\Delta t} = b(X,t) + \sqrt{a(X,t)} \lim_{\Delta t \to 0} \frac{\Delta W}{\Delta t} + \lim_{\Delta t \to 0} \frac{o(\Delta t)}{\Delta t}$$

Unfortunately $\frac{\Delta W}{\Delta t}$ does not exist as a regular function.

We have to work with differences and finite Δt.

The general approach:

In the discussion of stochastic differential equations, we work with finite dt. Then at the end, we take the limit as $dt \to 0$.

Properties of Wiener process:

1) $dW \sim N(0, dt)$ ==> $dW = \sqrt{dt}\, X$ where $X \sim N(0,1)$
2) $E(dW) = 0$
3) $E((dW)^2) = dt$
4) $dW(t_1)$ and $dW(t_2)$ are independent if the time intervals are disjoint.
5) $dW = O(\sqrt{dt})$ in the statistical sense.

   The RMS (root mean square) of dW is

   $$\text{RMS}(dW) = \sqrt{E\left((dW)^2\right)} = \sqrt{dt}$$

A discrete version of $W(t)$

The discrete version is conceptually easy to understand, and is computationally practical to work with in simulations.

Consider $W(t)$ on a grid over time interval $[0, t_f]$.

Grid points: $\left\{(j\Delta t),\quad j=0,1,\ldots,n\right\},\qquad \Delta t=\dfrac{t_f}{n}$

$W(t)$ on the grid: $\left\{W_j=W(j\Delta t),\quad j=0,1,\ldots,n\right\}$

<u>Question:</u> How to generate a discrete sample path $\{W_j, j = 0, 1, \ldots, n\}$?

<u>Answer:</u> By the definition of $W(t)$, we have

$$\Delta W_j \equiv (W_{j+1}-W_j)=\sqrt{\Delta t}\, X_j\,,\qquad X_j \sim N(0,1)\,,\qquad j=0,1,\ldots,n-1$$

ΔW_j and ΔW_k are independent for $j \neq k$.

<u>Method:</u>

Generate n independent samples of $N(0, 1)$.

$$\left\{X_j,\ j=0,1,\ldots,n-1\right\} \sim \text{(iid)}\, N(0,1)$$

(In Matlab, "randn(1, n)" generates n independent samples of $N(0, 1)$.)

Calculate $\{W_j, j = 0, 1, \ldots, n\}$ as a cumulative sum.

$$W_0=0\,,\qquad W_j=\sqrt{\Delta t}\sum_{k=0}^{j-1}X_k\,,\qquad j=1,2,\ldots,n$$

(In Matlab, "cumsum(X)" calculates the cumulative sum of array X.)

<u>Remarks:</u>

- This method completely specifies the random experiment for generating a discrete sample path $\{W_j, j = 0, 1, \ldots, n\}$.
- On the grid, discrete sample $\{W_j, j = 0, 1, \ldots, n\}$ is exactly the same as the underlying full sample path $W(t)$ (i.e., no approximation error).
- Given a coarse-grid sample $\{W_j, j = 0, 1, \ldots, n\}$, it is desirable to refine it to obtain a fine-grid sample $\{W_k, k = 0, 1, \ldots, 2n\}$ of the same underlying full sample path $W(t)$. The problem of refining a given discrete sample path will be discussed after introducing Bayes' theorem.

If you think $W(t)$ is somewhat unusual and different from the functions we are familiar with, it indeed is. Below we illustrate one peculiar feature of $W(t)$.

<u>A peculiar feature of $W(t)$:</u>

The arc length of $W(t)$ over $[0, t_f]$ is infinity.

<u>Derivation:</u>

We start with the arc length of discrete sample path $\{W_j, j = 0, 1, 2, \ldots\}$.

$$\text{Discrete arc length} = \sum_{j=0}^{n-1} \left| (t_{j+1}, W_{j+1}) - (t_j, W_j) \right| = \sum_{j=0}^{n-1} \left| (\Delta t,\ \Delta W_j) \right|$$

$$= \sum_{j=0}^{n-1} \sqrt{(\Delta t)^2 + (\sqrt{\Delta t}\, X_j)^2}\ , \qquad \{X_j\} \sim \text{iid } N(0,1)$$

$$> \sum_{j=0}^{n-1} \sqrt{\Delta t}\, |X_j| = n\sqrt{\Delta t} \left(\frac{1}{n} \sum_{j=0}^{n-1} |X_j| \right)$$

$$\text{Use } \Delta t = \frac{t_f}{n} \quad \text{and} \quad \frac{1}{n}\sum_{j=0}^{n-1} \left| X_j \right| \approx E(|X|) = \sqrt{\frac{2}{\pi}}$$

$$= \sqrt{n t_f}\sqrt{\frac{2}{\pi}} \ \rightarrow +\infty \qquad \text{as } n \rightarrow +\infty$$

Here we used $E(|X|) = \sqrt{\frac{2}{\pi}}$ for $X \sim N(0, 1)$ (Homework problem)

Therefore, we conclude that the arc length of $W(t)$ over $[0, t_f]$ is infinity!

<u>Ito's lemma</u>:

Suppose $f(t, w)$ is a smooth function of two variables t and w.

Replacing w by $W(t)$ gives us $f(t, W(t))$, a non-smooth random function of single variable t. The randomness comes from the Wiener process $W(t, \omega)$.

We examine the increment of $f(t, W(t))$ corresponding to dt.

$$df(t, W(t)) \equiv f(t+dt, W+dW) - f(t, W).$$

First we expand $f(t, w)$ as a smooth 2-variable function.

$$f(t+dt, w+dw) = f(t, w) + f_t dt + f_w dw$$
$$+\frac{1}{2}\left[f_{tt}(dt)^2 + 2 f_{tw}(dt)(dw) + f_{ww}(dw)^2 \right]$$
$$+O\Big((dt)^3 + (dt)^2(dw) + (dt)(dw)^2 + (dw)^3\Big)$$

We apply the expansion to $f(t+dt, W+dW)$, use $dW = O(\sqrt{dt})$, and neglect $o(dt)$ terms.

$$df(t, W(t)) = f_t dt + f_w dW + \frac{1}{2} f_{ww}(dW)^2 + o(dt) \qquad \text{(E01)}$$

<u>Claim:</u> we can replace $(dW)^2$ with dt and write df as

$$df(t,W(t)) = f_t dt + f_w dW + \frac{1}{2} f_{ww} dt + o(dt) = \left(f_t + \frac{1}{2} f_{ww} \right) dt + f_w dW + o(dt)$$

<u>Theorem (Ito's lemma):</u>

Given $f(0, 0)$, at any $t_f > 0$, the two SDEs below give the same $f(t_f, W(t_f))$.

$$df(t,W(t)) = f_t dt + f_w dW + \frac{1}{2} f_{ww} (dW)^2 + o(dt)$$

$$df(t,W(t)) = \left(f_t + \frac{1}{2} f_{ww} \right) dt + f_w dW + o(dt)$$

<u>Outline of proof:</u>

Let $dt = t_f/n$ and $t_j = jdt$. We calculate $\{f(t_j, W(t_j)), j = 1, 2, ..., n\}$ sequentially.

In one step of dt, the error of replacing $(dW)^2$ with dt is

$$\text{err}_j = \frac{1}{2} f_{ww} \left((dW_j)^2 - dt \right), \qquad dW_j = \sqrt{dt}\, X_j\,, \quad X_j \sim N(0,1)$$

The total error at t_f is

$$\text{err}_{\text{tot}} = \sum_{j=0}^{n-1} \text{err}_j$$

In the simple case of $f_{ww} \equiv 2$, we have

$$E(\text{err}_j) = E\left((dW_j)^2 - dt \right) = 0$$

$$\text{var}(\text{err}_j) = \text{var}\left((dW_j)^2 \right) = 2(dt)^2 \qquad \text{(Homework problem)}$$

Since $\{dW_j, j = 0, 1, 2, ... \}$ are independent, we obtain

$$E(\text{err}_{\text{tot}}) = \sum_{j=0}^{n-1} E(\text{err}_j) = 0$$

$$\text{var}(\text{err}_{\text{tot}}) = \sum_{j=0}^{n-1} \text{var}(\text{err}_j) = 2n(dt)^2 = 2t_f (dt) \to 0 \quad \text{as } dt \to 0$$

We will look at related materials in subsequent lectures/assignments.

<u>The mean-value version of Ito's lemma:</u>

The mean of $df(t, W(t))$ can be calculated exactly using $E(dW)=0$ and $E((dW)^2)=dt$.

$$E_{dW}\left(f(t+dt, W+dW) \right) = f(t,W) + f_t dt + \frac{1}{2} f_{ww} dt + o(dt)$$

We will use this version of Ito's lemma to study the Gambler's ruin problem.

Another version of law of total probability

We start with a unified view of probability and expectation.

Key observation: The probability of an event can be written in terms of the expectation of a random variable.

Given event A, we define random variable X as

$$X(\omega)=\begin{cases}1, & \text{if } \omega\in A\\ 0, & \text{otherwise}\end{cases}$$

==> $\Pr(A) = E(X)$

$E(X) = E(E(X|Y))$ (The law of total expectation)

==> $\boxed{\Pr(A) = E(\Pr(A|Y))}$ (The law of total probability)

Gambler's ruin (applications of Ito's lemma)

Notation and modeling approach:

C: total cash = the sum of your cash and casino's cash (assuming you are the only one playing with the casino).

$X(t)$: your cash at time t. In practice, $C \gg X(0)$.

"Breaking the bank" means "$X(t)$ hits C before hitting 0".

Case 1: we first consider a fair game

$$dX = dW$$

which means $X(t+dt)=X(t)+dW$

It is a fair game because

$$E_{dW}(dX)=E(dW)=0$$

We study the two questions below.

Question #1: How long can you play?

Question #2: What is the chance that you break the bank?

Answer to Question #2 (we address Question #1 after this)

Let $u(x)=\Pr\left(A\middle|\, X(0)=x\right)$, $A\equiv\{X(t) \text{ hits } C \text{ before } 0\}$.

Strategy:

Find a boundary value problem (BVP) governing $u(x)$.

Boundary condition:

$u(C) = 1$ and $u(0) = 0$.

Differential equation:

Start with $X(0) = x \in (0, C)$. After a short time dt, we have

$X(dt) = x + dW$

Recall that $dW = O(\sqrt{dt})$. For a fixed $x \in (0, C)$, when dt is small enough, the probability of $X(t)$ hitting 0 or C in time interval $[0, dt]$ is exponentially small. Here the magnitude of dt depends on how close x is to the two boundaries.

For a fixed $x \in (0, C)$, when dt is small enough (depending on x), we have

$$u(x) = \Pr(A) = E(\underbrace{\Pr(A|X(dt) = x + dW)}_{u(x+dW)}) + o(dt), \qquad A = \{X(t) \text{ hits } C \text{ before } 0\}$$

$$= E_{dW}\left(u(x+dW)\right) + o(dt)$$

Here we used the law of total probability $\boxed{\Pr(A) = E(\Pr(A|Y))}$ (draw a diagram).

Expanding $u(x+dW)$ inside E(), we get

$$u(x) = E_{dW}\left(u(x) + u_x dW + \frac{1}{2}u_{xx}(dW)^2\right) + o(dt)$$

$$= u(x) + \frac{1}{2}u_{xx}dt + o(dt)$$

Divide by dt and then take the limit as $dt \to 0$, we obtain

$$u_{xx} = 0$$

This is the differential equation governing $u(x)$. Thus, function $u(x)$ satisfies the boundary value problem (BVP)

$$\begin{cases} u_{xx}(x) = 0 & \text{differential equation} \\ u(0) = 0, \quad u(C) = 1 & \text{boundary conditions} \end{cases}$$

Solving the differential equation: $u(x) = c_1 + c_2 x$

Enforcing the boundary conditions: $u(x) = \dfrac{x}{C}$

The probability of breaking the bank is proportional to your initial cash and inversely proportional to the total cash.

Answer to Question #1

Let $T(x) = E\left(Z \middle| X(0) = x\right), \quad Z \equiv \left(\text{time from 0 until } X(t) = C \text{ or } X(t) = 0\right)$

Strategy:

Find a boundary value problem (BVP) governing $T(x)$.

Boundary condition:

$T(0) = 0$ and $T(C) = 0$.

Differential equation:

Start with $X(0) = x \in (0, C)$. After a short time dt, we have

$X(dt) = x + dW$

For a fixed $x \in (0, C)$, when dt is small enough (depending on x), we have

$$T(x) = E(Z) = E(E(Z|X(dt) = x + dW)) + o(dt)\,, \quad Z = \begin{pmatrix}\text{time from 0 until} \\ X(t)=C \text{ or } X(t)=0\end{pmatrix}$$

$$= E(\underbrace{E((Z+dt)|X(0) = x + dW)}_{T(x+dW)}) = dt + E_{dW}\left(T(x+dW)\right) + o(dt)$$

Here we used the law of total expectation $\boxed{E(Z) = E(E(Z|Y))}$ (draw a diagram).

Expanding $T(x+dW)$ inside E(), we get

$$T(x) = dt + E_{dW}\left(T(x) + T_x dW + \frac{1}{2}T_{xx}(dW)^2\right) + o(dt)$$

$$= dt + T(x) + \frac{1}{2}T_{xx}dt + o(dt)$$

Divide by dt and then take the limit as $dt \to 0$, we obtain

$$T_{xx} = -2$$

This is the differential equation governing $T(x)$. Thus, function $T(x)$ satisfies the boundary value problem (BVP)

$$\begin{cases} T_{xx}(x) = -2 & \text{differential equation} \\ T(0) = 0, \quad T(C) = 0 & \text{boundary conditions} \end{cases}$$

A particular solution of DE: $T(x) = -x^2$

The general solution of DE: $T(x) = c_1 + c_2 x - x^2$

Enforcing the BCs: $T(x) = x(C - x)$

Remark:

The average does not give us the full picture!

$T(x)$ is the average time until going bankrupt or breaking the bank. However, this average does not give us the full picture of how long we can play with initial cash x.

In particular, when $C = \infty$ (when the casino has infinite amount of cash), we have

$T(x) = x(C-x) = \infty$.

This certainly does not mean we can play forever with initial cash x.

A more detailed answer to Question #1:

We look at the probability of surviving beyond time t.

Assume $C = \infty$. We consider a function of two variables

$$P(x,t) = \Pr\Big(A(t)\,\Big|\,X(0) = x\Big), \qquad A(t) \equiv \Big\{X(\tau) > 0 \text{ for } \tau \in [0,t]\Big\}$$

$P(x, t)$ is the conditional probability of surviving beyond time t given $X(0) = x$.

Strategy:

Find an initial boundary value problem (IBVP) governing $P(x, t)$.

Initial and boundary conditions:

Initial condition: $P(x, 0) = 1$ for $x > 0$

(with $x > 0$, we can certainly survive beyond time 0)

Boundary condition: $P(0, t) = 0$ for $t > 0$

(with $x = 0$, we cannot survive beyond time 0)

Differential equation:

Start with $X(0) = x > 0$. After a short time dt, we have

$X(dt) = x + dW$

For a fixed $x > 0$, when dt is small enough (depending on x), we have

$$P(x,t) = \Pr(A(t)) = E(\Pr(A(t)|X(dt) = x + dW)) + o(dt), \quad A(t) = \Big\{X(\tau) > 0 \text{ for } \tau \in [0,t]\Big\}$$

$$= E(\underbrace{\Pr(A(t-dt)|X(0) = x + dW)}_{P(x+dW,\,t-dt)}) + o(dt) = E_{dW}\Big(P(x+dW, t-dt)\Big) + o(dt)$$

Here we used the law of total probability $\boxed{\Pr(A) = E(\Pr(A|Y))}$ (draw a diagram).

Expanding $P(x+dW, t-dt)$ inside E(), we get

$$P(x,t) = E_{dW}\left(P(x,t) + P_t(-dt) + P_x dW + \frac{1}{2}P_{xx}(dW)^2\right) + o(dt)$$

$$= P(x,t) + P_t(-dt) + \frac{1}{2}P_{xx}dt + o(dt)$$

Divide by dt and then take the limit as $dt \to 0$, we obtain

$$P_t = \tfrac{1}{2}P_{xx}$$

This is the PDE governing $P(x, t)$. Thus, function $P(x, t)$ satisfies the initial boundary value problem (IBVP)

$$\begin{cases} P_t = \frac{1}{2} P_{xx} & \text{partial differential equation} \\ P(0,t) = 0 & \text{boundary condition} \\ P(x,0) = 1 & \text{initial condition} \end{cases}$$

We use the odd extension to convert it to an IVP. The odd extension satisfies the zero-value boundary condition automatically.

Odd extension:

$P(-x, t) = -P(x, t)$

The extended function $P(x, t)$ is governed by the IVP

$$\begin{cases} P_t = \frac{1}{2} P_{xx} \\ P(x,0) = \begin{cases} 1, & x>0 \\ -1, & x<0 \end{cases} \end{cases} \tag{E02}$$

Solution of a general IVP of the heat equation:

$$\begin{cases} u_t = a u_{xx} \\ u(x,0) = f(x) \end{cases} \tag{E03}$$

The solution of IVP (E03) has the expression:

$$u(x,t) = \frac{1}{\sqrt{4\pi a t}} \int_{-\infty}^{+\infty} \exp\left(\frac{-\xi^2}{4at}\right) f(x-\xi)\, d\xi$$

Solution of IVP (E02).

Applying the general formula to (E02), we identify

$$a = \frac{1}{2}, \qquad f(x-\xi) = \begin{cases} 1, & \xi < x \\ -1, & \xi > x \end{cases}$$

We write out $P(x, t)$, the solution of (E02).

$$\begin{aligned} P(x,t) &= \frac{1}{\sqrt{2\pi t}} \int_{-\infty}^{\infty} \exp\left(\frac{-\xi^2}{2t}\right) f(x-\xi)\, d\xi \\ &= \frac{1}{\sqrt{2\pi t}} \int_{-\infty}^{x} \exp\left(\frac{-\xi^2}{2t}\right) d\xi - \frac{1}{\sqrt{2\pi t}} \int_{x}^{\infty} \exp\left(\frac{-\xi^2}{2t}\right) d\xi \\ &= \frac{2}{\sqrt{2\pi t}} \int_{0}^{x} \exp\left(\frac{-\xi^2}{2t}\right) d\xi \end{aligned}$$

Change of variables: $\xi = \sqrt{2t}\, s$

$$= \frac{2}{\sqrt{\pi}} \int_0^{\frac{x}{\sqrt{2t}}} \exp(-s^2) ds = \operatorname{erf}\left(\frac{x}{\sqrt{2t}} \right)$$

Thus, probability $P(x, t)$ has the expression

$$P(x,t) = \operatorname{erf}\left(\frac{x}{\sqrt{2t}} \right)$$

<u>Scaling property</u> of $P(x, t)$:

Start with initial cash x. The survival probability p and the time t are related by

$$p = \operatorname{erf}\left(\frac{x}{\sqrt{2t}} \right)$$

==> $$\frac{x}{\sqrt{2t}} = \operatorname{erfinv}(p)$$

==> $$t = \frac{x^2}{2 \operatorname{erfinv}(p)^2}$$

Given a prescribed threshold p, the maximum time t with surviving probability $\geq p$ is proportional to x^2 with the coefficient depending on p.

<u>A few example values</u> of the coefficient:

$p = 0.1$ ==> $t = 63.33\, x^2$

$p = 0.3$ ==> $t = 6.735\, x^2$

$p = 0.5$ ==> $t = 2.198\, x^2$

$p = 0.7$ ==> $t = 0.931\, x^2$

$p = 0.9$ ==> $t = 0.370\, x^2$