Lecture 19. Path integral

Tiejun Li[1,2]

[1]School of Mathematical Sciences (SMS),
&
[2]Center for Machine Learning Research (CMLR),
Peking University,
Beijing 100871,
P.R. China
tieli@pku.edu.cn

Office: No. 1 Science Building, Room 1376E

Path integral: brief introduction

- The path integral can be dated to R. Feynman to construct a new formulation to understand quantum mechanics.

Path integral: brief introduction

- The path integral can be dated to R. Feynman to construct a new formulation to understand quantum mechanics.
- The path integral gives very powerful formal approach to deal with the probability measures on path space and compute the expectation for some functionals of Wiener paths.

Path integral: brief introduction

- The path integral can be dated to R. Feynman to construct a new formulation to understand quantum mechanics.
- The path integral gives very powerful formal approach to deal with the probability measures on path space and compute the expectation for some functionals of Wiener paths.
- Briefly speaking, path integral is a formal infinite dimensional limit of the considered stochastic process under finite dimensional approximations.

Wiener process

- Let us start with the formal representation of the Wiener measure P_* defined on the canonical space $(C[0,1], \mathcal{B}(C[0,1]))$ for the standard Wiener process.

Wiener process

- Let us start with the formal representation of the Wiener measure P_* defined on the canonical space $(C[0,1], \mathcal{B}(C[0,1]))$ for the standard Wiener process.
- From the definition of Wiener process, we have the joint pdf for $(W_{t_1}, W_{t_2}, \ldots, W_{t_n})$

$$p_n(w_1, w_2, \ldots, w_n) = \frac{1}{Z_n} \exp(-I_n(w)),$$

where $0 < t_1 < t_2 < \cdots < t_n \leq 1$ and

$$Z_n = (2\pi)^{\frac{n}{2}} \left[t_1(t_2 - t_1) \cdots (t_n - t_{n-1})\right]^{\frac{1}{2}},$$

$$I_n(w) = \frac{1}{2} \sum_{j=1}^{n} \left(\frac{w_j - w_{j-1}}{t_j - t_{j-1}}\right)^2 (t_j - t_{j-1}), \quad t_0 := 0, w_0 := 0.$$

Formal Wiener Measure

- Now we take the formal limit as $n \to \infty$, we obtain

$$p_n dw_1 dw_2 \cdots dw_n \to \frac{1}{Z} \exp(-I[w])\delta(w_0)\mathcal{D}w,$$

where the δ-function $\delta(w_0)$ is to fix $w_0 = 0$, $I[w]$ is called the action functional of the Wiener process defined as

$$I[w] = \frac{1}{2} \int_0^1 \dot{w_t}^2 dt.$$

$\mathcal{D}w$ is a shortcut for $\prod_{0 \le t \le 1} dw_t$, which is the formal volume element in the path space $C[0, 1]$. Z is the normalization factor.

Formal Wiener Measure

- Now we take the formal limit as $n \to \infty$, we obtain

$$p_n dw_1 dw_2 \cdots dw_n \to \frac{1}{Z} \exp(-I[w]) \delta(w_0) \mathcal{D}w,$$

where the δ-function $\delta(w_0)$ is to fix $w_0 = 0$, $I[w]$ is called the action functional of the Wiener process defined as

$$I[w] = \frac{1}{2} \int_0^1 \dot{w_t}^2 dt.$$

$\mathcal{D}w$ is a shortcut for $\prod_{0 \le t \le 1} dw_t$, which is the formal volume element in the path space $C[0,1]$. Z is the normalization factor.

- For notations, we use the lowercase w_t for dumb variables, but the uppercase W_t for the stochastic process.

Remark on the Formal Wiener Measure

- We should emphasize that this interpretation is purely formal and all of the results induced by the path integral need to be reproved in rigorous mathematical language before we want to use it as an theorem.

Remark on the Formal Wiener Measure

- We should emphasize that this interpretation is purely formal and all of the results induced by the path integral need to be reproved in rigorous mathematical language before we want to use it as an theorem.
- One reason to understand it is only formal is that we have no infinite dimensional Lebesgue measure.

Issue on the Infinite Dimensional Lebesgue Measure

- Let us consider an infinite dimensional Hilbert space H with orthonormal basis $\{e_1, e_2, \ldots\}$.

Issue on the Infinite Dimensional Lebesgue Measure

- Let us consider an infinite dimensional Hilbert space H with orthonormal basis $\{e_1, e_2, \ldots\}$.
- Define the balls

$$B_n = B_{\frac{1}{2}}(e_n) = \{x | \|x - e_n\| \leq 1/2\}, \quad B = B_2(0) = \{x | \|x\| \leq 2\}.$$

Issue on the Infinite Dimensional Lebesgue Measure

- Let us consider an infinite dimensional Hilbert space H with orthonormal basis $\{e_1, e_2, \ldots\}$.
- Define the balls

$$B_n = B_{\frac{1}{2}}(e_n) = \{x | \|x - e_n\| \leq 1/2\}, \quad B = B_2(0) = \{x | \|x\| \leq 2\}.$$

- As a Lebesgue measure, it should be translation invariant and finite for bounded sets.

Issue on the Infinite Dimensional Lebesgue Measure

- Let us consider an infinite dimensional Hilbert space H with orthonormal basis $\{e_1, e_2, \ldots\}$.
- Define the balls

$$B_n = B_{\frac{1}{2}}(e_n) = \{x | \|x - e_n\| \leq 1/2\}, \quad B = B_2(0) = \{x | \|x\| \leq 2\}.$$

- As a Lebesgue measure, it should be translation invariant and finite for bounded sets.
- If the Lebesgue measure on H exists as $\mu(\cdot)$, then we have

$$0 < \mu(B_1) = \mu(B_2) = \cdots = \mu(B_n) = \cdots < \infty, \quad 0 < \mu(B) < \infty.$$

Issue on the Infinite Dimensional Lebesgue Measure

- However from the disjointness of $\{B_n\}$ and $B_n \subset B$ for any n, we obtain
$$\mu(B) \geq \sum_n \mu(B_n) = \infty,$$
which is a contradiction!

Issue on the Infinite Dimensional Lebesgue Measure

- However from the disjointness of $\{B_n\}$ and $B_n \subset B$ for any n, we obtain
$$\mu(B) \geq \sum_n \mu(B_n) = \infty,$$
which is a contradiction!
- Thus the notation $\mathcal{D}w$ is totally meaningless!

Issue on the Infinite Dimensional Lebesgue Measure

- However from the disjointness of $\{B_n\}$ and $B_n \subset B$ for any n, we obtain
$$\mu(B) \geq \sum_n \mu(B_n) = \infty,$$
which is a contradiction!
- Thus the notation $\mathcal{D}w$ is totally meaningless!
- But the glamor of path integral is that it can give some extremely insightful results in a very efficient way. That is why it is also useful for applied mathematicians.

Expectation of a Wiener functional

Example

Compute the expectation

$$\mathbb{E} \exp\Big(-\frac{1}{2}\int_0^1 W_t^2 dt \Big).$$

Expectation of a Wiener functional

Example

Compute the expectation

$$\mathbb{E}\exp\Big(-\frac{1}{2}\int_0^1 W_t^2 dt\Big).$$

- Note that it is not straightforward to compute this expectation since the integrand involves the whole Wiener path, i.e. a Wiener functional.

Expectation of a Wiener functional

Solution.

- From the Karhunen-Loeve expansion,

$$\begin{aligned}\int_0^1 W_t^2 dt &= \int_0^1 \sum_{k,l} \sqrt{\lambda_k \lambda_l} \alpha_k \alpha_l \phi_k(t) \phi_l(t) dt \\ &= \sum_k \int_0^1 \lambda_k \alpha_k^2 \phi_k^2(t) dt = \sum_k \lambda_k \alpha_k^2.\end{aligned}$$

Expectation of a Wiener functional

Solution.

- From the Karhunen-Loeve expansion,

$$\int_0^1 W_t^2 dt = \int_0^1 \sum_{k,l} \sqrt{\lambda_k \lambda_l} \alpha_k \alpha_l \phi_k(t) \phi_l(t) dt$$
$$= \sum_k \int_0^1 \lambda_k \alpha_k^2 \phi_k^2(t) dt = \sum_k \lambda_k \alpha_k^2.$$

- Then

$$\mathbb{E} \exp\Big(- \frac{1}{2} \int_0^1 W_t^2 dt \Big) = \mathbb{E}\Big(\prod_k \exp(-\frac{1}{2} \lambda_k \alpha_k^2) \Big)$$
$$= \prod_k \mathbb{E} \exp(-\frac{1}{2} \lambda_k \alpha_k^2).$$

Expectation of a Wiener functional

- From the identity

$$\mathbb{E}\exp(-\frac{1}{2}\lambda_k\alpha_k^2) = \int_{-\infty}^{+\infty}\frac{1}{\sqrt{2\pi}}e^{-\frac{x^2}{2}}\cdot e^{-\frac{1}{2}\lambda_k x^2}dx = \sqrt{\frac{1}{1+\lambda_k}}$$

we obtain

$$\mathbb{E}\exp\Big(-\frac{1}{2}\int_0^1 W_t^2 dt\Big) \quad = \quad \prod_k\sqrt{\frac{1}{1+\lambda_k}} := M,$$

where

$$M^{-2} = \prod_{k=1}^{\infty}\Big(1+\frac{4}{(2k-1)^2\pi^2}\Big).$$

Expectation of a Wiener functional

- From the identities for infinite product series we have

$$\cosh(x) = \prod_{n=1}^{\infty} \left(1 + \frac{4x^2}{(2n-1)^2\pi^2}\right),$$

where $\cosh(x) = (e^x + e^{-x})/2$.

Expectation of a Wiener functional

- From the identities for infinite product series we have

$$\cosh(x) = \prod_{n=1}^{\infty} \left(1 + \frac{4x^2}{(2n-1)^2\pi^2}\right),$$

where $\cosh(x) = (e^x + e^{-x})/2$.

- Thus

$$M = (\cosh(1))^{-\frac{1}{2}} = \sqrt{\frac{2e}{1+e^2}}.$$

Path integral approach

Here we show how to apply the path integral approach to compute the expectation of this Wiener functional. The path integral approach to compute the expectation is composed of the following two steps.

- Step 1. Discretize the problem into finite dimensions.

Path integral approach

Here we show how to apply the path integral approach to compute the expectation of this Wiener functional. The path integral approach to compute the expectation is composed of the following two steps.

- Step 1. Discretize the problem into finite dimensions.
- Step 2. Take the formal limit as $n \to \infty$.

Step 1. Discretize the problem into finite dimensions

At first let us take finite dimensional approximation to the functional

$$\begin{aligned} &\exp\left(-\frac{1}{2}\int_0^1 W_t^2 dt\right) \\ \approx &\exp\Big(-\frac{1}{2}\sum_{j=1}^{n} W_{t_j}^2 \Delta t\Big) \\ = &\exp\left(-\frac{1}{2}\Delta t X^T A X\right), \end{aligned}$$

where $\Delta t = t_j - t_{j-1}$ for $j = 1, 2, \ldots, n$, $A = I$, and $X = (W_{t_1}, W_{t_2}, \cdots, W_{t_n})^T$.

Step 1. Discretize the problem into finite dimensions

Thus

$$\mathbb{E}\exp\Big(-\frac{1}{2}\int_0^1 W_t^2 dt\Big)$$
$$\approx \int_{\mathbb{R}^n} \exp\left(-\frac{1}{2}\Delta t \boldsymbol{x}^T A \boldsymbol{x}\right) \cdot \frac{1}{Z_n} \exp\left(-\frac{1}{2}\Delta t \boldsymbol{x}^T B \boldsymbol{x}\right) d\boldsymbol{x},$$

where $\boldsymbol{x} = (x_1, x_2, \ldots, x_n)$, $Z_n = (2\pi)^{\frac{n}{2}}(\det(\Delta t B)^{-1})^{\frac{1}{2}}$, and

$$B = \frac{1}{\Delta t^2}\begin{pmatrix} 2 & -1 & & & \\ -1 & 2 & -1 & & \\ & \ddots & \ddots & \ddots & \\ & & -1 & 2 & -1 \\ & & & -1 & 1 \end{pmatrix}.$$

Step 1. Discretize the problem into finite dimensions

From equation

$$\mathbb{E}\exp\Big(-\frac{1}{2}\int_0^1 W_t^2 dt\Big)$$
$$\approx \int_{\mathbb{R}^n} \exp\left(-\frac{1}{2}\Delta t \boldsymbol{x}^T A \boldsymbol{x}\right) \cdot \frac{1}{Z_n} \exp\left(-\frac{1}{2}\Delta t \boldsymbol{x}^T B \boldsymbol{x}\right) d\boldsymbol{x},$$

we have

$$\mathbb{E}\exp\Big(-\frac{1}{2}\int_0^1 W_t^2 dt\Big) \approx \frac{(2\pi)^{\frac{n}{2}}(\det(\Delta t(A+B))^{-1})^{\frac{1}{2}}}{(2\pi)^{\frac{n}{2}}(\det(\Delta t B)^{-1})^{\frac{1}{2}}}$$
$$= \left(\frac{\det(B)}{\det(A+B)}\right)^{\frac{1}{2}} = \left(\frac{\prod_i \lambda_i^B}{\prod_i \lambda_i^{A+B}}\right)^{\frac{1}{2}},$$

where $\lambda_i^B, \lambda_i^{A+B}$ are eigenvalues of B and $A+B$, respectively.

Step 2. Take the formal limit as $n \to \infty$

- If we take the formal limit as $n \to +\infty$, the matrix B will converge to the differential operator $\mathcal{B} = -d^2/dt^2$ with zero Dirichlet boundary condition at $t = 0$ and free Neumann boundary condition at $t = 1$.

Step 2. Take the formal limit as $n \to \infty$

- If we take the formal limit as $n \to +\infty$, the matrix B will converge to the differential operator $\mathcal{B} = -d^2/dt^2$ with zero Dirichlet boundary condition at $t = 0$ and free Neumann boundary condition at $t = 1$.
- Thus the eigenvalues of $\mathcal{B}$ corresponds to the following Sturm-Liouville boundary value problem

$$-\frac{d^2 u}{dt^2} = \lambda u(t), \quad u(0) = 0, u'(1) = 0.$$

Step 2. Take the formal limit as $n \to \infty$

- Note that the quadratic form

$$\int_0^1 W_t^2 dt = (\mathcal{A} W_t, W_t),$$

where $\mathcal{A} = I$ and $(f, g) := \int_0^1 fg dt$.

Step 2. Take the formal limit as $n \to \infty$

- Note that the quadratic form

$$\int_0^1 W_t^2 dt = (\mathcal{A}W_t, W_t),$$

where $\mathcal{A} = I$ and $(f, g) := \int_0^1 fg dt$.

- We have the formal path integral limit

$$\begin{aligned}
&\mathbb{E} \exp\Big(-\frac{1}{2}\int_0^1 W_t^2 dt\Big) \\
&= \int \exp\Big(-\frac{1}{2}(\mathcal{A}w_t, w_t)\Big) \cdot \frac{1}{Z} \exp\Big(-\frac{1}{2}(\mathcal{B}w_t, w_t)\Big)\delta(w_0)\mathcal{D}w
\end{aligned}$$

where the operator $\mathcal{B}u(t) := d^2u/dt^2$ and

$$Z = \int \exp\Big(-\frac{1}{2}(\mathcal{B}w_t, w_t)\Big)\delta(w_0)\mathcal{D}w.$$

Step 2. Take the formal limit as $n \to \infty$

Now we formally apply the Gaussian integrals in infinite dimensions to obtain

$$\mathbb{E} \exp\Big(- \frac{1}{2} \int_0^1 W_t^2 dt \Big) = \Big(\frac{\det \mathcal{B}}{\det(\mathcal{A} + \mathcal{B})} \Big)^{\frac{1}{2}},$$

where $\det \mathcal{B},\ \det(\mathcal{A} + \mathcal{B})$ mean the products of all eigenvalues for the following boundary value problems:

$$\begin{cases} \mathcal{B}u = \lambda u, \quad \text{or} \quad (\mathcal{A} + \mathcal{B})u = \lambda u, \\ u(0) = 0, \quad u'(1) = 0. \end{cases}$$

This yield the same result as before.

Girsanov transformation

- We have seen that the Wiener measure over $[0,1]$ can be formally expressed as

$$d\mu_W = Z^{-1} \exp\left(-\frac{1}{2}\int_0^1 \dot{w}_t^2 dt\right) \delta(w_0)\mathcal{D}w.$$

Girsanov transformation

- We have seen that the Wiener measure over $[0, 1]$ can be formally expressed as

$$d\mu_W = Z^{-1} \exp\left(-\frac{1}{2}\int_0^1 \dot{w}_t^2 dt\right) \delta(w_0)\mathcal{D}w.$$

- The solution of the SDE

$$dX_t = b(X_t, t) + \sigma(X_t, t)dW_t, \quad X_0 = 0.$$

can be viewed as a map between the Wiener path $\{W_t\}$ and $\{X_t\}$:

$$\{W_t\} \xrightarrow{\Phi} \{X_t\}.$$

Girsanov transformation

- We have seen that the Wiener measure over $[0,1]$ can be formally expressed as

$$d\mu_W = Z^{-1} \exp\left(-\frac{1}{2}\int_0^1 \dot{w}_t^2 dt\right) \delta(w_0)\mathcal{D}w.$$

- The solution of the SDE

$$dX_t = b(X_t,t) + \sigma(X_t,t)dW_t, \quad X_0 = 0.$$

can be viewed as a map between the Wiener path $\{W_t\}$ and $\{X_t\}$:

$$\{W_t\} \xrightarrow{\Phi} \{X_t\}.$$

- Consequently, the mapping Φ induces another measures on $C[0,1]$, which is nothing but the distribution of $\{X_t\}$.

Girsanov transformation

- We now ask the question how the measure $d\mu_W$ changes under the mapping Φ?

Girsanov transformation

- We now ask the question how the measure $d\mu_W$ changes under the mapping Φ?
- Let us first consider the case when $\sigma = 1$ in one dimension. The more general conditions can be derived in a similar way.

Girsanov transformation

- We now ask the question how the measure $d\mu_W$ changes under the mapping Φ?
- Let us first consider the case when $\sigma = 1$ in one dimension. The more general conditions can be derived in a similar way.
- We will perform the path integral through two steps as in the previous section: that is, making discretization first and then taking the formal continuum limit.

Step 1. Discretize the problem into finite dimensions

- With the Euler-Maruyama discretization, we obtain

$$X_{t_{j+1}} = X_{t_j} + b(X_{t_j}, t_j)(t_{j+1} - t_j) + (W_{t_{j+1}} - W_{t_j}).$$

Step 1. Discretize the problem into finite dimensions

- With the Euler-Maruyama discretization, we obtain

$$X_{t_{j+1}} = X_{t_j} + b(X_{t_j}, t_j)(t_{j+1} - t_j) + (W_{t_{j+1}} - W_{t_j}).$$

- In matrix form we have

$$B\cdot \begin{pmatrix} X_{t_1} \\ X_{t_2} \\ \vdots \\ X_{t_n} \end{pmatrix} - \begin{pmatrix} b(X_{t_0}, t_0)(t_1 - t_0) \\ b(X_{t_1}, t_1)(t_2 - t_1) \\ \vdots \\ b(X_{t_{n-1}}, t_{n-1})(t_n - t_{n-1}) \end{pmatrix} = B\cdot \begin{pmatrix} W_{t_1} \\ W_{t_2} \\ \vdots \\ W_{t_n} \end{pmatrix},$$

where $t_0 = 0$, $X_{t_0} = 0$, and the matrix B has the form

$$B = \begin{pmatrix} 1 & & & & \\ -1 & 1 & & & \\ & \ddots & \ddots & & \\ & & \ddots & \ddots & \\ & & & -1 & 1 \end{pmatrix}_{n\times n}.$$

Step 1. Discretize the problem into finite dimensions

- The equation

$$X_{t_{j+1}} = X_{t_j} + b(X_{t_j}, t_j)(t_{j+1} - t_j) + (W_{t_{j+1}} - W_{t_j}).$$

indeed introduces a finite dimensional transformation Φ_n as

$$\{W_{t_1}, W_{t_2}, \cdots, W_{t_n}\} \xrightarrow{\Phi_n} \{X_{t_1}, X_{t_2}, \cdots, X_{t_n}\}.$$

Step 1. Discretize the problem into finite dimensions

- The equation

$$X_{t_{j+1}} = X_{t_j} + b(X_{t_j}, t_j)(t_{j+1} - t_j) + (W_{t_{j+1}} - W_{t_j}).$$

  indeed introduces a finite dimensional transformation Φ_n as

$$\{W_{t_1}, W_{t_2}, \cdots, W_{t_n}\} \xrightarrow{\Phi_n} \{X_{t_1}, X_{t_2}, \cdots, X_{t_n}\}.$$

- With dumb variables representation for the equation, we have

$$x_{j+1} = x_j + b(x_j, t_j)(t_{j+1} - t_j) + (w_{j+1} - w_j), \quad j = 0, \dots, n-1$$

  where $w_0 = 0$ and x_0 is fixed.

Step 1. Discretize the problem into finite dimensions

- The equation

$$X_{t_{j+1}} = X_{t_j} + b(X_{t_j}, t_j)(t_{j+1} - t_j) + (W_{t_{j+1}} - W_{t_j}).$$

  indeed introduces a finite dimensional transformation Φ_n as

$$\{W_{t_1}, W_{t_2}, \cdots, W_{t_n}\} \xrightarrow{\Phi_n} \{X_{t_1}, X_{t_2}, \cdots, X_{t_n}\}.$$

- With dumb variables representation for the equation, we have

$$x_{j+1} = x_j + b(x_j, t_j)(t_{j+1} - t_j) + (w_{j+1} - w_j), \quad j = 0, \ldots, n-1$$

  where $w_0 = 0$ and x_0 is fixed.
- It is not difficult to find that the Jacobian of the transformation

$$\frac{\partial(w_1, \ldots, w_n)}{\partial(x_1, \ldots, x_n)} = 1.$$

Step 1. Discretize the problem into finite dimensions

- Suppose we want to compute the average $\langle F[X_t] \rangle$, then

  $$\langle F[X_t] \rangle \approx \langle F(X_{t_1}, X_{t_2}, \cdots, X_{t_n}) \rangle = \langle G(W_{t_1}, W_{t_2}, \cdots, W_{t_n}) \rangle,$$

  where $G = F \circ \Phi_n$.

Step 1. Discretize the problem into finite dimensions

- Suppose we want to compute the average $\langle F[X_t]\rangle$, then

  $$\langle F[X_t]\rangle \approx \langle F(X_{t_1}, X_{t_2}, \cdots, X_{t_n})\rangle = \langle G(W_{t_1}, W_{t_2}, \cdots, W_{t_n})\rangle,$$

  where $G = F \circ \Phi_n$.
- Furthermore with transformation of variables

  $$\begin{aligned}\langle F[X_t]\rangle &\approx \int G(w_1, w_2, \cdots, w_n)\frac{1}{Z_n}\exp(-I_n(\boldsymbol{w}))dw_1dw_2\cdots dw_n \\ &= \int F(x_1, x_2, \cdots, x_n)\frac{1}{Z_n}\exp(-\tilde{I}_n(\boldsymbol{x}))dx_1dx_2\cdots dx_n,\end{aligned}$$

  where the transformation holds because of the performance of the Jacobian, and $\tilde{I}_n(\boldsymbol{x}) = I_n \circ \Phi_n^{-1}(\boldsymbol{x})$ by definition

  $$\begin{aligned}\tilde{I}_n(\boldsymbol{x}) =& \frac{1}{2}\sum_{j=1}^{n}\left(\frac{x_j - x_{j-1}}{t_j - t_{j-1}}\right)^2(t_j - t_{j-1}) + \frac{1}{2}\sum_{j=1}^{n} b^2(x_{j-1}, t_{j-1})(t_j - t_{j-1}) \\ &- \sum_{j=1}^{n}(x_j - x_{j-1}) \cdot b(x_{j-1}, t_{j-1}).\end{aligned}$$

Step 1. Discretize the problem into finite dimensions

Changing the dumb variables x_i to w_i, we obtain

$$
\begin{aligned}
\langle F[X_t] \rangle &\approx \int F(w_1, w_2, \cdots, w_n) \frac{1}{Z_n} \exp(-I_n(\boldsymbol{w})) \\
&\quad \cdot \exp\Big(-\frac{1}{2} \sum_{j=1}^{n} b^2(w_{j-1}, t_{j-1})(t_j - t_{j-1}) \Big) \\
&\quad \cdot \exp\Big(\sum_{j=1}^{n} b(w_{j-1}, t_{j-1}) \cdot (w_j - w_{j-1}) \Big) dw_1 dw_2 \cdots dw_n \\
&= \Big\langle F(W_{t_1}, W_{t_2}, \cdots, W_{t_n}) \\
&\quad \cdot \exp\Big(-\frac{1}{2} \sum_{j=1}^{n} b^2(W_{t_{j-1}}, t_{j-1})(t_j - t_{j-1}) \Big) \\
&\quad \cdot \exp\Big(\sum_{j=1}^{n} b(W_{t_{j-1}}, t_{j-1}) \cdot (W_{t_j} - W_{t_{j-1}}) \Big) \Big\rangle .
\end{aligned}
$$

Step 2. Take the formal limit as $n \to \infty$

- Now with the finite dimensional discretization, we can take formal continuum limit

$$\langle F[X_t] \rangle = \left\langle F[W_t] \exp\left(-\frac{1}{2}\int_0^1 b^2(W_t,t)dt + \int_0^1 b(W_t,t)dW_t\right)\right\rangle.$$

Step 2. Take the formal limit as $n \to \infty$

- Now with the finite dimensional discretization, we can take formal continuum limit

$$\langle F[X_t] \rangle = \Big\langle F[W_t] \exp\Big(-\frac{1}{2}\int_0^1 b^2(W_t,t)dt + \int_0^1 b(W_t,t)dW_t\Big)\Big\rangle.$$

- Since the transformation is valid for arbitrary F, in mathematical language, this asserts that the distribution μ_X is absolutely continuous with respect to μ_W, and

$$\frac{d\mu_X}{d\mu_W} = \exp\Big(-\frac{1}{2}\int_0^1 b^2(W_t,t)dt + \int_0^1 b(W_t,t)dW_t\Big).$$

Step 2. Take the formal limit as $n \to \infty$

The above derivations can be done directly with continuum version if one gets familiar enough

$$
\begin{aligned}
&\langle F[X_t] \rangle \\
=&\langle G[W_t] \rangle \quad \text{(where } G = F \circ \Phi) \\
=&\int G[w_t] \cdot \frac{1}{Z} \exp\Big(-\frac{1}{2}\int_0^1 \dot{w}_t^2 dt \Big) \delta(w_0) \mathcal{D}w \\
=&\int F[x_t] \cdot \frac{1}{Z} \\
&\cdot \exp\Big(-\frac{1}{2}\int_0^1 \dot{x}_t^2 dt - \frac{1}{2}\int_0^1 b^2(x_t, t) dt + \int_0^1 b(x_t, t)\dot{x}_t dt \Big) \delta(x_0) \mathcal{D}x \\
=&\int F[w_t] \cdot \frac{1}{Z} \\
&\cdot \exp\Big(-\frac{1}{2}\int_0^1 \dot{w}_t^2 dt - \frac{1}{2}\int_0^1 b^2(w_t, t) dt + \int_0^1 b(w_t, t)\dot{w}_t dt \Big) \delta(w_0) \mathcal{D}w \\
=&\Big\langle F[W_t] \exp\Big(-\frac{1}{2}\int_0^1 b^2(W_t, t) dt + \int_0^1 b(W_t, t) dW_t \Big) \Big\rangle.
\end{aligned}
$$

Cameron-Martin formula

- A special case of this representation is the Cameron-Martin formula, for the transformation

$$\boldsymbol{X}_t = \boldsymbol{W}_t + \boldsymbol{\phi}(t)$$

where ϕ is a smooth function.

Cameron-Martin formula

- A special case of this representation is the Cameron-Martin formula, for the transformation

$$\boldsymbol{X}_t = \boldsymbol{W}_t + \boldsymbol{\phi}(t)$$

  where ϕ is a smooth function.
- This can be obtained from SDE with $\boldsymbol{b}(X_t, t) = \dot{\boldsymbol{\phi}}(t)$. In this case, we get

$$\frac{d\mu_X}{d\mu_W} = \exp\Big(-\frac{1}{2}\int_0^1 \dot{\boldsymbol{\phi}}^2(t)dt + \int_0^1 \dot{\boldsymbol{\phi}}(t)d\boldsymbol{W}_t\Big).$$

Girsanov formula

- A slight generalization is the Girsanov formula. Consider two SDE's:

$$\begin{aligned} d\boldsymbol{X}_t &= \boldsymbol{b}(\boldsymbol{X}_t, t)dt + \boldsymbol{\sigma}(\boldsymbol{X}_t, t)d\boldsymbol{W}_t, \\ d\boldsymbol{Y}_t &= (\boldsymbol{b}(\boldsymbol{Y}_t, t) + \boldsymbol{\gamma}(t, \omega))dt + \boldsymbol{\sigma}(\boldsymbol{Y}_t, t)d\boldsymbol{W}_t, \end{aligned}$$

  where $\boldsymbol{X}, \boldsymbol{Y}, \boldsymbol{b}, \boldsymbol{\gamma} \in \mathbb{R}^n$, $\boldsymbol{W} \in \mathbb{R}^m$ and $\boldsymbol{\sigma} \in \mathbb{R}^{n\times m}$. Assume that $\boldsymbol{X}_0 = \boldsymbol{Y}_0 = \boldsymbol{x}$.

Girsanov formula

- A slight generalization is the Girsanov formula. Consider two SDE's:

$$dX_t = b(X_t, t)dt + \sigma(X_t, t)dW_t,$$
$$dY_t = (b(Y_t, t) + \gamma(t, \omega))dt + \sigma(Y_t, t)dW_t,$$

  where $X, Y, b, \gamma \in \mathbb{R}^n$, $W \in \mathbb{R}^m$ and $\sigma \in \mathbb{R}^{n\times m}$. Assume that $X_0 = Y_0 = x$.

- Then the distributions of $\{X_t\}$ and $\{Y_t\}$ over $[0, 1]$ are absolutely continuous with respect to each other. Moreover the Radon-Nikodym derivative is given by

$$\frac{d\mu_Y}{d\mu_X}[X.] = \exp\left(-\frac{1}{2}\int_0^1 |\phi(t,\omega)|^2 dt + \int_0^1 \phi(t,\omega)dW_t\right),$$

  where ϕ is the solution of

$$\sigma(X_t, t)\phi(t, \omega) = \gamma(t, \omega).$$

Girsanov theorem

- Mathematically, the above two results have another formulation whose idea can be explained as follows.

Girsanov theorem

- Mathematically, the above two results have another formulation whose idea can be explained as follows.
- Suppose we have n independent standard Gaussian random variables $Z_1, Z_2, \ldots, Z_n \sim N(0,1)$ on probability space $(\Omega, \mathcal{F}, P)$. Given a vector $(\mu_1, \mu_2, \ldots, \mu_n) \in \mathbb{R}^n$, the new random variables with translation

$$\tilde{Z}_k = Z_k + \mu_k, \quad k = 1, 2 \ldots, n$$

are no longer $N(0,1)$ distributed.

Girsanov theorem

- Mathematically, the above two results have another formulation whose idea can be explained as follows.
- Suppose we have n independent standard Gaussian random variables $Z_1, Z_2, \dots, Z_n \sim N(0,1)$ on probability space $(\Omega, \mathcal{F}, P)$. Given a vector $(\mu_1, \mu_2, \dots, \mu_n) \in \mathbb{R}^n$, the new random variables with translation

$$\tilde{Z}_k = Z_k + \mu_k, \quad k = 1, 2 \dots, n$$

  are no longer $N(0,1)$ distributed.
- But we can define another probability measure

$$\tilde{P}(d\omega) = \exp\Big(- \sum_{k=1}^{n} \mu_k Z_k(\omega) - \frac{1}{2} \sum_{k=1}^{n} \mu_k^2 \Big) P(d\omega).$$

Girsanov theorem

Then we have

$$\begin{aligned}
&\tilde{P}\left(\tilde{Z}_1 \in [\tilde{z}_1, \tilde{z}_1 + d\tilde{z}_1), \dots, \tilde{Z}_n \in [\tilde{z}_n, \tilde{z}_n + d\tilde{z}_n)\right) \\
=& \exp\Big(-\sum_{k=1}^{n} \mu_k(\tilde{z}_k - \mu_k) - \frac{1}{2}\sum_{k=1}^{n} \mu_k^2 \Big) \\
&\cdot P\left(\tilde{Z}_1 \in [\tilde{z}_1, \tilde{z}_1 + d\tilde{z}_1), \dots, \tilde{Z}_n \in [\tilde{z}_n, \tilde{z}_n + d\tilde{z}_n)\right) \\
=& \exp\Big(-\sum_{k=1}^{n} \mu_k(\tilde{z}_k - \mu_k) - \frac{1}{2}\sum_{k=1}^{n} \mu_k^2 \Big) \\
&\cdot (2\pi)^{-\frac{n}{2}} \exp\Big(-\frac{1}{2}\sum_{k=1}^{n} (\tilde{z}_k - \mu_k)^2 \Big) d\tilde{z}_1 \cdots d\tilde{z}_n \\
=& (2\pi)^{-\frac{n}{2}} \exp\Big(-\frac{1}{2}\sum_{k=1}^{n} \tilde{z}_k^2 \Big) d\tilde{z}_1 \cdots d\tilde{z}_n.
\end{aligned}$$

Girsanov theorem

- This reveals that the variables $\{\tilde{Z}_k\}_{k=1,\ldots,n}$ are again independent $N(0,1)$ random variables on space $(\Omega, \mathcal{F}, \tilde{P})$.

Girsanov theorem

- This reveals that the variables $\{\tilde{Z}_k\}_{k=1,\ldots,n}$ are again independent $N(0,1)$ random variables on space $(\Omega, \mathcal{F}, \tilde{P})$.
- If we take

$$Z_k = \frac{\Delta W_k}{\sqrt{\Delta t_k}}, \quad \tilde{Z}_k = \frac{\Delta \tilde{W}_k}{\sqrt{\Delta t_k}}, \quad \mu_k = \phi_k \sqrt{\Delta t_k}$$

and take the formal limit as $n \to \infty$, where $\Delta W_k = W_{t_{k+1}} - W_{t_k}$ and W_t is the standard Wiener process on $(\Omega, \mathcal{F}, P)$, we may claim that

$$\tilde{W}_t = W_t + \int_0^t \phi(s) ds$$

is again a standard Wiener process on $(\Omega, \mathcal{F}, \tilde{P})$ with

$$\tilde{P}(d\omega) = \exp\Big(- \int_0^t \phi(s) dW_s - \frac{1}{2} \int_0^t \phi^2(s) ds\Big) P(d\omega).$$

This claim is indeed true even for multidimensional case and the translation $\phi(t)$ can be ω-dependent.

Girsanov theorem

Theorem (Girsanov theorem I)

For Itô process

$$d\tilde{\boldsymbol{W}}_t = \boldsymbol{\phi}(t,\omega)dt + d\boldsymbol{W}_t, \quad \tilde{\boldsymbol{W}}_0 = 0,$$

where $\boldsymbol{W} \in \mathbb{R}^d$ *is a* d*-dimensional standard Wiener process on* $(\Omega, \mathcal{F}, \mathbb{P})$*. Define*

$$Z_t(\omega) = \exp\Big(- \int_0^t \boldsymbol{\phi}(s,\omega)d\boldsymbol{W}_s - \frac{1}{2}\int_0^t \boldsymbol{\phi}^2(s,\omega)ds\Big).$$

Assume $\boldsymbol{\phi}(t,\omega)$ *satisfies* $\mathbb{E}\exp\left(\frac{1}{2}\int_0^T |\boldsymbol{\phi}|^2(s,\omega)ds\right) < \infty$ *(Novikov's condition), where* $T \leq \infty$ *is a fixed constant. Define* $\tilde{\mathbb{P}}$ *as*

$$\tilde{\mathbb{P}}(d\omega) = Z_T(\omega)\mathbb{P}(d\omega),$$

then we have $\tilde{\boldsymbol{W}}$ *is a* d*-dimensional Wiener process with respect to* $(\Omega, \mathcal{F}_T, \tilde{\mathbb{P}})$ *for* $t \leq T$*.*

Girsanov theorem

- The Novikov's condition is to ensure the process Z_t in

$$Z_t(\omega) = \exp\Big(- \int_0^t \boldsymbol{\phi}(s,\omega) d\boldsymbol{W}_s - \frac{1}{2}\int_0^t \boldsymbol{\phi}^2(s,\omega) ds \Big).$$

is an *exponential martingale*.

Girsanov theorem

- The Novikov's condition is to ensure the process Z_t in

$$Z_t(\omega) = \exp\Big(- \int_0^t \boldsymbol{\phi}(s,\omega) d\boldsymbol{W}_s - \frac{1}{2}\int_0^t \boldsymbol{\phi}^2(s,\omega) ds \Big).$$

is an *exponential martingale*.

- The rigorous proof of Girsanov Theorem may be referred to:
  I. Karatzas and S. E. Shreve, *Brownian motion and stochastic calculus*,
  B. Øksendal, *Stochastic differential equations: An introduction with applications*.

Girsanov theorem

- The definition of Z_t does not contradict

$$\frac{d\mu_X}{d\mu_W} = \exp\Big(-\frac{1}{2}\int_0^1 \dot{\boldsymbol{\phi}}^2(t)dt + \int_0^1 \dot{\boldsymbol{\phi}}(t)d\boldsymbol{W}_t \Big).$$

Indeed, they are consequences of each other.

Girsanov theorem

- The definition of Z_t does not contradict

$$\frac{d\mu_X}{d\mu_W} = \exp\Big(-\frac{1}{2}\int_0^1 \dot{\boldsymbol{\phi}}^2(t)dt + \int_0^1 \dot{\boldsymbol{\phi}}(t)d\boldsymbol{W}_t\Big).$$

  Indeed, they are consequences of each other.

- To see this, we note that for any functional F

$$\begin{aligned}
&\Big\langle F[\tilde{\boldsymbol{W}}_t]\Big\rangle_{\tilde{\mathbb{P}}} = \Big\langle F[\tilde{\boldsymbol{W}}_t]Z_T\Big\rangle_{\mathbb{P}} \\
=&\Big\langle F[\tilde{\boldsymbol{W}}_t]\exp\Big(-\int_0^T \boldsymbol{\phi}(s,\omega)d\tilde{\boldsymbol{W}}_s + \frac{1}{2}\int_0^T \boldsymbol{\phi}^2(s,\omega)ds\Big)\Big\rangle_{\mathbb{P}} \\
=&\Big\langle F[\boldsymbol{W}_t]\exp\Big(-\int_0^T \boldsymbol{\phi}(s,\omega)d\boldsymbol{W}_s + \frac{1}{2}\int_0^T \boldsymbol{\phi}^2(s,\omega)ds\Big)\frac{d\mu_{\tilde{\boldsymbol{W}}}}{d\mu_{\boldsymbol{W}}}\Big\rangle_{\mathbb{P}} \\
=&\Big\langle F[\boldsymbol{W}_t]\Big\rangle_{\mathbb{P}}.
\end{aligned}$$

Girsanov theorem

It can also be verified by path integrals as follows

$$
\begin{aligned}
& \langle F[\boldsymbol{W}_t]\rangle_{\mathbb{P}} \\
= \; & \int F[\boldsymbol{w}_t]\cdot\frac{1}{Z}\exp\Big(-\frac{1}{2}\int_0^T \dot{\boldsymbol{w}}_t^2 dt\Big)\delta(\boldsymbol{w}_0)D\boldsymbol{w} \\
= \; & \int F[\tilde{\boldsymbol{w}}_t]\cdot\frac{1}{Z}\exp\Big(-\frac{1}{2}\int_0^T \dot{\tilde{\boldsymbol{w}}}_t^2 dt\Big)\delta(\tilde{\boldsymbol{w}}_0)D\tilde{\boldsymbol{w}} \\
= \; & \int F\circ\Phi[\boldsymbol{w}_t] \\
& \cdot\frac{1}{Z}\exp\Big(-\frac{1}{2}\int_0^T \dot{\boldsymbol{w}}_t^2 dt-\frac{1}{2}\int_0^T \boldsymbol{\phi}^2 dt-\int_0^T \boldsymbol{\phi}(t)\dot{\boldsymbol{w}}_t dt\Big)\delta(\boldsymbol{w}_0)D\boldsymbol{w} \\
= \; & \Big\langle G[\boldsymbol{W}_t]\exp\Big(-\frac{1}{2}\int_0^T \boldsymbol{\phi}^2(t)dt-\int_0^T \boldsymbol{\phi}(t)d\boldsymbol{W}_t\Big)\Big\rangle_{\mathbb{P}} \\
= \; & \Big\langle F[\tilde{\boldsymbol{W}}_t]\exp\Big(-\frac{1}{2}\int_0^T \boldsymbol{\phi}^2(t)dt-\int_0^T \boldsymbol{\phi}(t)d\boldsymbol{W}_t\Big)\Big\rangle_{\mathbb{P}} \\
= \; & \Big\langle F[\tilde{\boldsymbol{W}}_t]Z_T\Big\rangle_{\mathbb{P}}=\Big\langle F[\tilde{\boldsymbol{W}}_t]\Big\rangle_{\tilde{\mathbb{P}}}.
\end{aligned}
$$

Girsanov theorem

Corresponding to

$$\frac{d\mu_Y}{d\mu_X}[X.] = \exp\left(-\frac{1}{2}\int_0^1 |\boldsymbol{\phi}(t,\omega)|^2 dt + \int_0^1 \boldsymbol{\phi}(t,\omega) d\boldsymbol{W}_t\right),$$

we have another form of Girsanov theorem.

Girsanov theorem

Theorem (Girsanov theorem II)

For Itô processes $\boldsymbol{X}, \boldsymbol{Y}$ satisfy

$$\begin{cases} d\boldsymbol{X}_t = \boldsymbol{b}(\boldsymbol{X}_t, t)dt + \boldsymbol{\sigma}(\boldsymbol{X}_t, t)d\boldsymbol{W}_t, \quad \boldsymbol{X}_0 = \boldsymbol{x}, \\ d\boldsymbol{Y}_t = (\boldsymbol{b}(\boldsymbol{Y}_t, t) + \boldsymbol{\gamma}(t,\omega))dt + \boldsymbol{\sigma}(\boldsymbol{Y}_t, t)d\boldsymbol{W}_t, \quad \boldsymbol{Y}_0 = \boldsymbol{x}, \end{cases}$$

where $\boldsymbol{X}, \boldsymbol{Y}, \boldsymbol{b}, \boldsymbol{\gamma} \in \mathbb{R}^n$, $\boldsymbol{W} \in \mathbb{R}^m$ and $\boldsymbol{\sigma} \in \mathbb{R}^{n\times m}$, and assume $\boldsymbol{b}$ and $\boldsymbol{\sigma}$ satisfy the global Lipschitz and linear growth conditions. Suppose there exists unique $\phi(t,\omega)$ such that $\boldsymbol{\sigma}(\boldsymbol{X}_t, t)\boldsymbol{\phi}(t,\omega) = \boldsymbol{\gamma}(t,\omega)$ and the Novikov's condition $\mathbb{E}\exp\left(\frac{1}{2}\int_0^T |\boldsymbol{\phi}|^2(s,\omega)ds\right) < \infty$ holds. Define $\tilde{\boldsymbol{W}}_t$, Z_t and $\tilde{\mathbb{P}}$ as in Girsanov theorem I, then $\tilde{\boldsymbol{W}}$ is a standard Wiener process under $(\Omega, \mathcal{F}_T, \tilde{\mathbb{P}})$ and $\boldsymbol{Y}$ satisfies

$$d\boldsymbol{Y}_t = \boldsymbol{b}(\boldsymbol{Y}_t, t)dt + \boldsymbol{\sigma}(\boldsymbol{Y}_t, t)d\tilde{\boldsymbol{W}}_t, \quad \boldsymbol{Y}_0 = \boldsymbol{x}, \quad t \le T.$$

Girsanov theorem

- Thus the law of $\boldsymbol{Y}_t$ under $\tilde{\mathbb{P}}$ is the same that of $\boldsymbol{X}_t$ under $\mathbb{P}$ for $t \leq T$.

Girsanov theorem

- Thus the law of $\boldsymbol{Y}_t$ under $\tilde{\mathbb{P}}$ is the same that of $\boldsymbol{X}_t$ under $\mathbb{P}$ for $t \leq T$.
- The readers may be referred to
  I. Karatzas and S. E. Shreve, *Brownian motion and stochastic calculus*,
  B. Øksendal, *Stochastic differential equations: An introduction with applications*.
  for proof details.

Feynman-Kac formula: revisited

- Earlier we have known that the solution of PDE

$$\partial_t v = \frac{1}{2}\Delta v + q(x)v, \quad v|_{t=0} = f(x)$$

can be represented as

$$v(x,t) = \mathbb{E}^x\Big(\exp\Big(\int_0^t q(W_s)ds\Big)f(W_t)\Big).$$

Feynman-Kac formula: revisited

- Earlier we have known that the solution of PDE

$$\partial_t v = \frac{1}{2}\Delta v + q(x)v, \quad v|_{t=0} = f(x)$$

can be represented as

$$v(x,t) = \mathbb{E}^x\Big(\exp\Big(\int_0^t q(W_s)ds\Big)f(W_t)\Big).$$

- In path integral form

$$v(x,t) = \int \delta(w_0 - x)\frac{1}{Z}\exp\Big(-\int_0^t \big(\frac{1}{2}\dot{w}_s^2 - q(w_s)\big)ds\Big)f(w_t)\mathcal{D}w,$$

where the delta-function $\delta(w_0 - x)$ is to shift the starting point of the Wiener process to x.

Feynman-Kac formula: revisited

Feynmann-Kac formula originates from Feynmann's interpretation of quantum mechanics, namely that solution of linear Schrödinger equation

$$i\hbar\partial_t\psi = -\frac{\hbar^2}{2m}\Delta\psi + V(x)\psi, \quad \psi|_{t=0} = \psi_0(x)$$

can be expressed formally as

$$\psi(x,t) = \int \delta(w_0 - x)\frac{1}{Z}\exp\Big(\frac{i}{\hbar}I[w]\Big)\psi_0(w_t)\mathcal{D}w,$$

where $I[\cdot]$ is the Lagrangian defined as

$$I[w] = \int_0^t \Big(\frac{m}{2}\dot{w}_s^2 - V(w_s)\Big)ds.$$

Feynman-Kac formula: revisited

- Formally if we take

$$m = 1, \quad \hbar = -i$$

in the Schrödinger equation and path integral form of the Schrödinger equation, we exactly obtain the above formulation for Feynman-Kac problem! Indeed, that is the real story on how Feynman-Kac formula is created.

Feynman-Kac formula: revisited

- Formally if we take

$$m = 1, \quad \hbar = -i$$

  in the Schrödinger equation and path integral form of the Schrödinger equation, we exactly obtain the above formulation for Feynman-Kac problem! Indeed, that is the real story on how Feynman-Kac formula is created.
- Feynman's formally expression is yet to be made rigorous. However, Kac's reinterpretation for the heat equation instead of Schrödinger's equation can be readily proved. The Feynman-Kac formula can also be generalized to the case when Δ is replaced by more general second order differential operator as we did in previous lecture.