Random <u>walks</u>

Last time

- → Wiener process
- SDEs
  stochastic calculus
- Diffusion models

Generative model :

<u>Output</u>

$x_{m+1}, x_{m+2}, \ldots \sim \pi$

- (target probability distribution)
- unknown

<u>Input</u>:

$x_1, x_2, x_3, \ldots, x_m \sim \pi$

Score-generative models (SGMs) based on SDEs

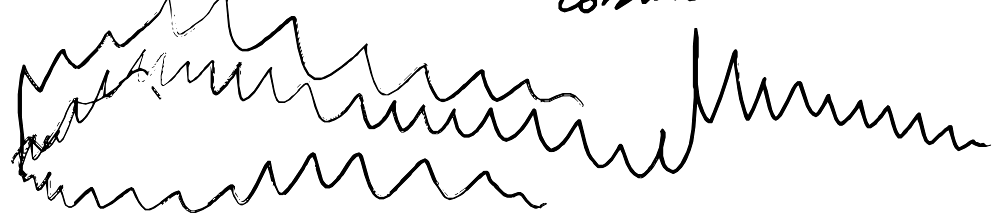
Wiener process

Scaling limit of random walks

- $W : [0, T] \times \Omega \longrightarrow \mathbb{R}^n$

  Stochastic process with properties

  - $W_0 = 0$ almost surely
  - $W_t - W_s \sim \mathcal{N}(0, (t-s) Id)$ for all $s \leq t$
  - $W_{t_0} - W_{t_1}$, $W_{t_1} - W_{t_2}, \dots$ $t_0 > t_1 > t_2 > \dots$ are independent RVs.
  - W_t is almost surely continuous.



- Wiener measure : Gaussian measure

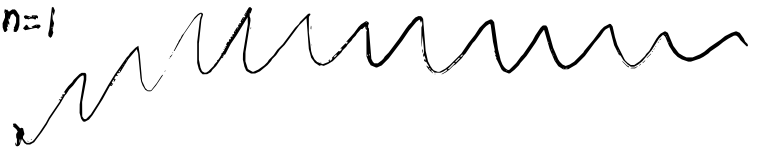
Infinite-dimensional Gaussian distribution

$\{X_1, \dots X_n\} \sim \mathbb{R}^n$ is multivariate Gaussian if $\sum_{i=1}^{n} a_i X_i$ is 1D Gaussian $\{a_i\} \in \mathbb{R}^n$ $a_i \neq 0$

$$P(X) = \frac{1}{(\sqrt{2\pi})^n |\Sigma|^{1/2}} e^{-(X-\mu)^T \Sigma^{-1} (X-\mu)}$$

$C([0,T], \mathbb{R}^n)$: space of continuous functions from $[0,T]$ to $\mathbb{R}^n$

n=1



Any Linear functional (function from $C([0,T], \mathbb{R}^n)$ to $\mathbb{R}$) on γ is a 1D Gaussian

$W_t - W_s$ $\qquad$ $W_t - W_0$

Fokker-Planck equation:
Deterministic evolution of probability densities

$$\underset{\uparrow}{dX_t} = dW_t \quad \text{(Brownian motion)}$$

Variation in stochastic process

$$dX_t = \lim_{\delta t \to 0} \left(X_{t+\delta t} - X_t \right)$$

$n=1$

$$\frac{\partial \rho(x,t)}{\partial t} = \underset{\uparrow}{D} \frac{\partial^2 \rho}{\partial x^2}(x,t)$$

Diffusion Coefficient

Particle in fluid moving "randomly"

- ε with $\frac{1}{2}$, $-\varepsilon$ with $\frac{1}{2}$

$$\rho(\underset{\mathbb{R}}{x}, t + \delta t) = \rho(x,t) + \delta t \frac{\partial \rho}{\partial t}(x, \delta t) + O(\delta t^2)$$

$$= \rho(x-\varepsilon, t) \times \frac{1}{2} + \rho(x+\varepsilon, t) \times \frac{1}{2}$$

$$= \frac{1}{2}\left(\rho(x,t) - \varepsilon \frac{\partial \rho}{\partial x}(x,t) + \rho(x,t) + \varepsilon \frac{\partial \rho}{\partial x}(x,t) + \frac{2\varepsilon^2}{2} \frac{\partial^2 \rho}{\partial x^2}(x,t) + O(\varepsilon^4) \right)$$

$$\frac{\partial \rho}{\partial t} = \boxed{\frac{\varepsilon^2}{2\delta t}} \frac{\partial^2 \rho}{\partial x^2} + O\left(\frac{\varepsilon^4}{(\delta t)^2} \right)$$

(Fokker-Planck equation)

$$\frac{\partial \rho}{\partial t} = D \frac{\partial^2 \rho}{\partial x^2} \qquad D = \frac{\varepsilon^2}{2\delta t}$$

$$\rho(x,t) = \frac{1}{\sqrt{2\pi}\sqrt{2Dt}} e^{-x^2/4Dt}$$

General SDE

$$dX_t = \underbrace{f(t, X_t)\,dt}_{\text{drift}} + \underbrace{\sigma(t, X_t)\,dW_t}_{\text{diffusion}}$$

→ Ito process

→ Chain rule Ito's lemma

$$df(t, X_t) = \left(\frac{\partial f}{\partial t} dt + \frac{\partial f}{\partial X} dX_t \right)$$

$$dX_t = f(t, X_t)\,dt + \sigma(t, X_t) \circ dW_t$$

Stratanovich

ODE

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

SDE

$$dX_t = f(t, X_t)dt + \sigma(t, X_t)dW_t$$

Forward Euler

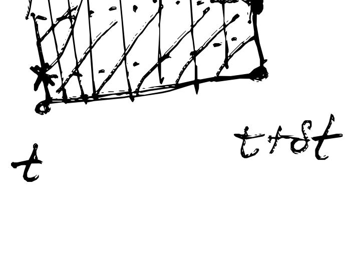
$$\frac{X(t+\delta t) - X(t)}{\delta t} = f(t, X(t))$$

$$X(t+\delta t) = f(X, X(t))\delta t + X(t)$$

Euler-Maruyama

$$X_{t+\delta t} - X_t = f(t, X_t)\delta t + \sigma(t, X_t)\eta_t$$

Runge-Kutta



t $\quad$ $t+\delta t$

midpoint value of stochastic process to define time-integration

: Stratanovich

Fokker-Planck equation of General SDE

$\rightarrow dX_t = f(t, X_t)\,dt + \underbrace{\sigma(t, X_t)}_{\in \mathbb{R}^{n\times n}} \underset{\in \mathbb{R}^n}{dW_t}$

$X_t \in \mathbb{R}^n$

FPE:

$$\frac{\partial \rho(x,t)}{\partial t} = \nabla\cdot(\rho f)(t,x) + \sum_{i,j} \frac{\partial^2 (D_{ij}\,\rho(x,t))}{\partial x_i \partial x_j}$$

$$D_{ij} = \frac{1}{2}(\sigma\sigma^T)_{ij}$$

1D:

$$\rightarrow \frac{\partial \rho}{\partial t} = \frac{\partial}{\partial x}(\rho f) + \frac{\sigma^2}{2}\frac{\partial^2 \rho}{\partial x^2}$$

PDE

Asymptotic solution of FPE:

steady: $\frac{\partial \rho}{\partial t} = 0$

Invariant measure of stochastic process

$dX_t = \ldots\ldots$

$\varphi_t(X_0) = X_t$

$\varphi_{t\#}\rho = \rho$

$\frac{\partial \rho}{\partial t} = 0$

O-U process

Ornstein- Uhlenbeck

$$dX_t = -a X_t \, dt + \sigma \, dW_t$$

FPE:

$$\frac{\partial \rho}{\partial t} = a \frac{\partial}{\partial x}(x \rho) + \frac{\sigma^2}{2} \frac{\partial^2 \rho}{\partial x^2}$$

Others :
→ Langevin diffusion
→ Overdamped Langevin diffusion

$$-a \frac{\partial}{\partial x}(x \rho) = \frac{\sigma^2}{2} \frac{\partial^2 \rho}{\partial x^2}$$

$$x \rho = \frac{\sigma^2}{2a} \frac{\partial \rho}{\partial x} + c$$

$$x \, dx = c_1 \frac{\partial \rho}{\rho}$$

$$\rho(x) = c_2 e^{-c_1 x^2}$$

MCMC methods :

π : target distribution

- Construct stochastic process with invariant dist π.
- Sample stochastic process for a long time

Input : $\nabla \log \pi$: Score of π

$$\nabla \log \pi : \mathbb{R}^n \to \mathbb{R}^n$$

$$\pi(x) = \frac{f(x)}{Z}$$

$$\nabla \log \pi(x) = \nabla \log f(x)$$

Know f ⇒ know score

- Score- Generative models :
  finite - time SDE evolution
  but can still provably sample π.