?

생성 : p_{data} 분포를 따르는 새로운 샘플 z를 뽑는 것.

- 조건부 생성 $p(\cdot|y)$
- 비조건부 〃 $p(\cdot)$

$x \sim p_{init} = N(0, I)$

⇓

$\boxed{\text{Generative Model}}$

⇓

$z \sim p_{data}$

ODE. $\frac{d}{dt}X_t = u_t(X_t)$ or $dX_t = u_t(X_t)\,dt$.

입자의 순간속도

입자가 위치한 곳의 물살 .

Flow $X_t = \psi_t(X_0)$ "t초 흐른 ODE를 나타냈을 때 최종위치"

Picard - Lindelof Thm.

예시 $u_t = -\theta x$

$\psi_t(x_0) = \exp(-\theta t)\cdot x_0$

물살 u_t가 충분히 부드러우면 ⇒ 모든 X_0에 대해 ODE는 유일한 해 ψ_t를 가진다.

① 연속으로 미분가능

② 도함수가 유계

복잡한 u_t에 대한 ψ_t를 구하기 위한 수치해석이 필요하다.

① 오일러 방법

$$\frac{X_{t+h} - X_t}{h} = u(X_t)$$

② 호인 방법

- $X'_{t+h} = X_t + h\,u_t(X_t)$
- $X_{t+h} = X_t + h\left(\frac{u_t(X_t) + u_{t+h}(X'_{t+h})}{2}\right)$

Flow Model (목표 $p_{init} \longrightarrow p_{data}$)

① $X_0 \sim p_{init} = \mathcal{N}(0, I)$

② define ODE $\Rightarrow$ $\frac{d}{dt} X_t = u_t(X_t)$ (Neural Net)

③ solve ODE $\Rightarrow$ X_1

④ train NN $\Rightarrow$ $X_1 \sim p_{data}$.

브라운 운동 $\Rightarrow$ 시뮬레이션

SDE의 핵심이다.

① $W_0 = 0$

· $W_{t+h} = W_t + \sqrt{h} \cdot \varepsilon_t$

$\sim \mathcal{N}(0, I)$

② $W_t - W_s \sim \mathcal{N}(0, (t-s) I)$

③ 겹치지 않는 시간 구간의

변화량은 서로 독립적

ODE : $X_0 \longrightarrow X_1$

deterministic

$$X_{t+h} = X_t + h\, u_t(X_t)$$

$$dX_t = u_t(X_t)\, dt.$$

SDE : $X_0 \longrightarrow X_1$

stocastic

$$X_{t+h} = X_t + h\, u_t(X_t) + \sigma_t \cdot (W_{t+h} - W_t)$$

$$\approx \sqrt{h} \cdot \sigma_t \cdot \varepsilon_t$$

$$dX_t = u_t(X_t)\, dt + \sigma_t \cdot dW_t$$

SDE에는 유일해가 존재한다. if $u_t(x)$ ① 연속으로 미·가 ② 도함수 유계 ; σ_t ③ 연속

- $X_0 = x_0$ (init cond)

- $dX_t = \underbrace{u(X_t)\,dt}_{\text{drift}} + \underbrace{\sigma_t\,dW_t}_{\text{diffusion}}$.

$$\begin{cases} \text{ODE} : & dX_t = f(X_t, t)\,dt \\ \text{SDE} : & dX_t = f(X_t, t)\,dt + g(t)\,dW_t \end{cases}$$

solution path $\{X_t\}_{t=0}^{T} \Longrightarrow$ prob. distribution $\{p_t\}_{t=0}^{T}$

- càdlàg $\begin{cases} \text{right-conti} \\ \exists \text{ left limit.} \end{cases}$

- $X_t = X_0 + \int_0^t f(X_s, s)\,ds + \int_0^t g(X_s, s)\,dW_s$

- $\int_0^t g(X_s, s)\,dW_s = \sum_{k=0}^{t/\Delta t} g(X_{k\Delta t}, \varepsilon_t)\underbrace{\sqrt{\Delta t}\cdot Z_k}_{=\Delta W_k} \quad (\Delta t \to 0)$

  ε_t: $[t_k, t_{k+1}]$, Z_k: $N(0, I)$

(example) OU process

$$dX_t = -\theta \cdot X_t\,dt + \sigma\,dW_t$$

$$X_\infty \sim N\left(0, \frac{\sigma^2}{2\theta}\right)$$

How to simulate SDE

< Euler-Maruyama Method >

$$X_{t+h} = X_t + h\,u_t(X_t) + \sqrt{h}\cdot\sigma_t\,\varepsilon_t$$

$\sim N(0, I)$

Diffusion 모델이 풀어야 하는 문제

$$X_0 \sim p_{init}$$

$$dX_t = u_t(X_t)\,dt + \sigma_t\,dW_t$$

$$\Rightarrow X_1 \sim p_{data}.$$

Fokker-Planck Equation

SDE

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

1차원

$$\partial_t p_t(x) = -\partial_x\left(f(x,t)\,p_t(x)\right) + \frac{g^2(t)}{2}\,\partial_x^2\left(p_t(x)\right) \quad \text{and} \quad X_t \sim p_t$$

---

① 극한 형태의 시간 미분

$$\partial_t \mathbb{E}[\phi(X_t)] = \lim_{\Delta t \to 0} \frac{\mathbb{E}[\phi(X_{t+\Delta t})] - \mathbb{E}[\phi(X_t)]}{\Delta t}$$

② 오일러-마루야마 근사

$$X_{t+\Delta t} \approx X_t + \Delta X_t$$

$$\approx X_t + f\,\Delta t + g\sqrt{\Delta t}\,z$$

③ 테일러 전개

$$\phi(X_{t+\Delta t}) \approx \phi(X_t) + \phi'(X_t)\cdot(f\Delta t + g\sqrt{\Delta t}\,z) + \frac{1}{2}\phi''(X_t)\cdot(g^2\Delta t\cdot z^2 + \cdots)$$

Δt에 대한 고차항

④ 기댓값

$$\mathbb{E}[\phi(X_{t+\Delta t})] - \mathbb{E}[\phi(X_t)]$$

$$\approx \mathbb{E}\left[\phi'(X_t)\cdot(f\Delta t + g\sqrt{\Delta t}\cdot \cancel{z}) + \frac{1}{2}\phi''(X_t)(g^2\Delta t\cdot \cancel{z^2})\right]$$

$E[z]=0$

$E[z^2]=1$

$$\approx \Delta t\cdot \mathbb{E}\left[\phi'(X_t)f + \frac{1}{2}\phi''(X_t)g^2\right]$$

⑤ 극한 계산 완료

$$\partial_t \mathbb{E}[\phi(X_t)] \approx \mathbb{E}\left[\phi'(X_t)\cdot f(X_t,t) + \frac{1}{2}\phi''(X_t)\, g^2(t)\right]$$

LHS RHS

⑥ 기댓값의 정의

$$\partial_t \int \phi(x)\, p_t(x)\, dx \qquad \int \phi'(x)(f p_t)\, dx$$

$$= \int \phi(x)\left(\partial_t p_t\right) dx \qquad + \int \frac{1}{2}\phi''(x)\,(g^2 p_t)\, dx$$

$$\int \phi(x)\cdot\left[-\partial_x (f p_t)\right] dt \qquad \int \phi(x)\cdot\left[\frac{g^2}{2}\partial_x^2 (p_t)\right] dx$$

---

⑦ 부분적분.

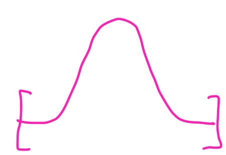
assume ϕ, f sufficiently smooth

& decay sufficiently quickly at $|x| \to \infty$

$\phi \in C_c^\infty(\mathbb{R})$

$$\int_{\mathbb{R}} \phi(x) f'(x)\, dx = \cancel{\left[\phi(x) f(x)\right]_{-\infty}^{\infty}}^{\,0} - \int_{\mathbb{R}} \phi'(x) f(x)\, dx$$

C_c^∞



$$\forall \phi \in C_c^\infty \qquad \int \phi(x)\underbrace{\left(\partial_t p_t + \partial_x (f p_t) - \frac{g^2}{2}\partial_x^2 (p_t)\right)}_{=0} dx = 0$$

다차원에서는...

$$\partial_t p_t = -\nabla_x \cdot (f p_t) + \frac{1}{2}\mathrm{tr}\left(g^T (\nabla_x^2 p_t)\, g\right)$$

OU process

- $dX_t = -\beta X_t dt + \sigma dW_t$

- $X_t = e^{-\beta t} X_0 + \sigma e^{-\beta t} \int_0^t e^{\beta s} dW_s$

- $X_t | X_0 \sim N\left(e^{-\beta t} X_0,\ \frac{\sigma^2}{2\beta} \cdot (1 - e^{-2\beta t})\right)$

Stationary distribution $N(0, \frac{\sigma^2}{2\beta})$

- $\mathbb{E}[X_t] = \mathbb{E}[\mathbb{E}[X_t | X_0]]$

$$= \mathbb{E}[e^{-\beta t} \cancel{X_0}^{\,0}] = 0$$

- $Var[X_t] = \mathbb{E}[Var(X_t | X_0)] + Var(\mathbb{E}[X_t | X_0])$

$$= \left[\frac{\sigma^2}{2\beta} \cdot (1 - e^{-2\beta t})\right] + (e^{-\beta t})^2 \cdot \frac{\sigma^2}{2\beta}$$

$$= \cdots = \frac{\sigma^2}{2\beta}$$

- $p_t(X_t) = \frac{1}{\sqrt{\pi\sigma^2/\beta}} \exp\left(-\frac{\beta}{\sigma^2} \cdot (X_t)^2\right)$

image corruption with OU

$$X_T \mid X_0 \sim \mathcal{N}\left(e^{-\beta t} x_0 , \frac{\sigma^2}{2\beta}(1 - e^{-2\beta t}) \cdot I\right)$$

output input

$$\sim \mathcal{N}\left(0 , \frac{\sigma^2}{2\beta}\right) \quad \text{if} \quad T \approx \infty$$

---

in ODE, $\begin{cases} X(0) \\ \frac{dX}{dt}(t) = f(X(t), t) \end{cases}$

i.e $X_{k+1} = X_k + \Delta t \cdot f(X_k, k\Delta t)$

hence $X_k = X_{k+1} - \Delta t \cdot f(X_{k+1}, k\Delta t)$

what about SDE?

$X_t = A_t + B_t$ and

① 결정론적 $A_t = e^{-\beta t} X_0$

② 확률론적 $B_t = \underbrace{\sigma \cdot e^{-\beta t}}_{=: Y_t} \cdot \underbrace{\int_0^t e^{\beta s} dW_s}_{=: Z_t}$

step ① 일단 미분.

- $dA_t = \dfrac{dA_t}{dt} \cdot dt = (-\beta \cdot e^{-\beta t} X_0)\, dt$
- $dY_t = \dfrac{dY_t}{dt} \cdot dt = (-\beta \cdot \sigma \cdot e^{-\beta t})\, dt$
- $dZ_t = d\left(\int_0^t e^{\beta s} dW_s\right) = e^{\beta t} \cdot dW_t$

---

step ② 이토 미분들의 세계

(i) $d(Y_t Z_t) = Y_t\, dZ_t + Z_t\, dY_t + dY_t\, dZ_t$.

(ii) $dt \cdot dt = 0$

$dt \cdot dW_t = 0$

$dW_t \cdot dW_t = dt$

hence

$dY_t \cdot dZ_t = (-\beta \cdot \sigma\, e^{-\beta t} dt) \cdot (e^{\beta t} \cdot dW_t) = 0$

---

step ③ 정리.

$$dX_t = \begin{pmatrix} (-\beta \cdot e^{-\beta t} X_0)\, dt \\ + (-\beta \cdot \sigma \cdot e^{-\beta t})\, dt \cdot \int_0^t e^{\beta s} dW_s \end{pmatrix} \Rightarrow -\beta X\, dt.$$

$$\left(+ \sigma \cdot e^{\cancel{-\beta t}} \cdot e^{\cancel{\beta t}} \cdot dW_t \right) \Rightarrow \sigma\, dW_t.$$

$$\Delta(YZ) = (Y+\Delta Y)(Z+\Delta Z) - YZ$$

$$= Y\Delta Z + Z\Delta Y + \Delta Y \Delta Z \, .$$

일반

$$\Delta Y \approx Y'(t)\cdot \Delta t$$

$$\Delta Z \approx Z'(t)\cdot \Delta t \, .$$

$$\Delta Y \Delta Z \approx Y'(t) Z'(t) \cdot (\cancel{\Delta t})^2$$

작음.

이토

- $dW_t \approx \sqrt{dt} \quad (dW_t \sim N(0, dt))$

- $\begin{cases} dY = a\cdot dt + b\cdot dW_t \\ dZ = c\cdot dt + d\cdot dW_t \, . \end{cases}$

- $$dYdZ = a\cdot c\,(\underbrace{\cancel{dt\cdot dt}}_{0}) + ad\,(\underbrace{\cancel{dt\cdot dW_t}}_{0}) + bc\,(\underbrace{\cancel{dW_t\cdot dt}}_{0}) + bd\,(\underbrace{\cancel{dW_t\cdot dW_t}}_{dt})$$

생존!

$$X_t = e^{-\beta t}\cdot X_0 + \sigma \cdot e^{-\beta t} \cdot \int_0^t e^{\beta s} dW_s$$

① $$E[X_t] = E[e^{-\beta t}\cdot X_0] + E[\sigma \cdot e^{-\beta t} \cdot \underbrace{\int_0^t e^{\beta s} dW_s}_{=0}]$$

$$= e^{-\beta t}\cdot E[X_0]$$

② $$Var[X_t] = e^{-2\beta t}\cdot Var(X_0) + \sigma^2 \cdot e^{-2\beta t} \cdot \underbrace{Var\left(\int_0^t e^{\beta s} \cdot dW_s\right)}_{= E\left[\int_0^t e^{2\beta s} ds\right]}$$

이토 등측정리

$$= \underbrace{e^{-2\beta t}\cdot Var(X_0) + \sigma^2 \cdot e^{-2\beta t} \cdot \underbrace{\int_0^t e^{2\beta s} ds}_{= \frac{1}{2\beta}\cdot(e^{2\beta s} - 1)}}_{= \frac{\sigma^2}{2\beta}(1 - e^{-2\beta t})}$$

이토 등측정리

$$\mathbb{E}\left[\left(\int_0^t f(s)\,dW_s\right)^2\right] = \mathbb{E}\left[\int_0^t (f(s))^2\,ds\right]$$

(pf) $\mathbb{E}\left[\left(\sum_i f(t_i)\cdot \Delta B_i\right)^2\right]$

$$= \mathbb{E}\left[\sum_i (f(t_i))^2 \cdot \underbrace{(\Delta B_i)^2}_{=\Delta t_i} + \sum_{i \neq j} f(t_i) f(t_j) \underbrace{\Delta B_i \Delta B_j}_{\mathbb{E}[\Delta B_i]\cdot\mathbb{E}[\Delta B_j] = 0}\right]$$