Diffusion Models と SDE と平衡完成!

Stochastic Diff. Eq.

Goal. DM について知り、
SDE, 平衡完成との関連を知ること

1. 正規化
2. Variational Free Energy
3. DM
4. SDE との関係

1. 正規分布

1.1 2次関数

Prop.（平均の特徴づけ）

$x_i \in \mathbb{R}$

$$\bar{x} = \frac{1}{N} \sum x_i \quad : \text{平均}$$

このとき

$$\bar{x} = \boxed{\underset{t}{\operatorname{argmin}}} \sum (t - x_i)^2$$

右側にtの関数と思って最小にさせるtの値

prf

$$\text{"RHS"} = Nt^2 - 2\sum x_i t + ▨$$

$$= N\left(t - \frac{\sum x_i}{N}\right)^2 + ▨$$

Prop

$x_i \in \mathbb{R}$

$w_i \in \mathbb{R}_{>0}$

$$\overline{x}^w := \frac{1}{\sum w_i} \sum w_i x_i$$

then

$$\overline{x}^w = \operatorname*{argmin}_t \sum w_i (t - x_i)^2$$

prf

$$\text{"RHS"} = \left(\sum w_i\right) t^2 - 2\left(\sum w_i x_i\right) t +$$

$$= \sum w_i \left(t - \frac{\sum w_i x_i}{\sum w_i} \right)^2 +$$

1.2 正規分布

$x \in \mathbb{R}$ or $x \in \mathbb{R}^n$ or $x \in \Omega \subset \mathbb{R}^n$

$[0,1]$とか、$\mathbb{R}_{>0}$とか、

$\mathbb{R}$とか[illegible]

Def

$p(x)$: func.

is prob. density func.

$\overset{\text{def.}}{\Longleftrightarrow}$ $p(x) \geq 0$, $\int p(x)\,dx = 1$

Def

平均　　分散

$\mu \in \mathbb{R}$, $\sigma^2 \in \mathbb{R}_{>0}$

$$p_{\mu,\sigma^2}(x) = \frac{1}{\sqrt{2\pi\sigma^2}} \exp\left(-\frac{(x-\mu)^2}{2\sigma^2}\right)$$

で定める prob. dist. を正規分布という

$$p_{\mu,\lambda}(x) = \sqrt{\frac{\lambda}{2\pi}} \exp\left(-\frac{1}{2}\lambda(x-\mu)^2\right)$$

$\lambda = \frac{1}{\sigma^2}$: 精度

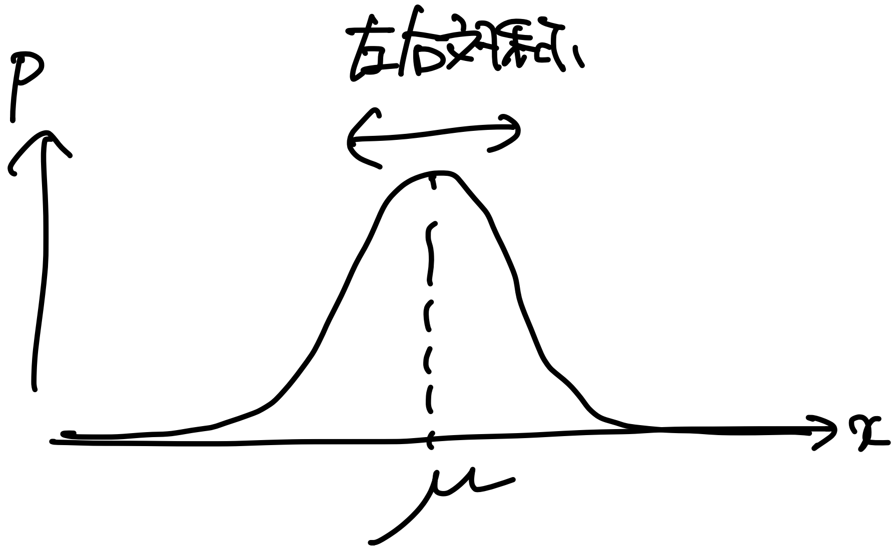
Rem

▷ $\int_{-\infty}^{+\infty} p_{\mu,\sigma^2}(x)\,dx = 1$, $p \geq 0$ なので prob. meas.

確率変数

▷ $X \sim p_{\mu,\sigma^2}$ に対し

$$E[X] = \mu,\quad V[X] = \sigma^2$$



(Ω, μ) : meas. sp. $\mu(\Omega) = 1$

$X : \Omega \to \mathbb{R}^n$: meas. を確率変数という

Def

Σ : symm. pos. definite $n \times n$ mat.

$\mu \in \mathbb{R}^n$ に対して,

$$p_{\mu,\Sigma}(x) := \frac{1}{\sqrt{(2\pi)^n \det \Sigma}} \times \exp\left(-\frac{1}{2} {}^t(x-\mu)\Sigma^{-1}(x-\mu)\right)$$

の定める prob. dist. を多変量正規分布という

Rem

$p_{\mu,\Sigma}$: prob. density func.

$X \sim p_{\mu,\Sigma}$ ならば

今日は diagonal のみ

$$E[X] = \mu, \quad Cov[X] = \Sigma$$

1.3 平方完成

Thm

$a, b, c \in \mathbb{R}$, $a \neq 0$

$$ax^2 + bx + c = a\left(x + \frac{b}{2a}\right)^2 - \frac{b^2}{4a} + c$$

Thm

A : symm. $\det A \neq 0$

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

$c \in \mathbb{R}$

$${}^t x A x + {}^t b x + c$$

$$= {}^t\left(x + \frac{1}{2}A^{-1}b\right) A \left(x + \frac{1}{2}A^{-1}b\right)$$

$$- \frac{1}{4}\,{}^t b A^{-1} b + c$$

prf

$$\text{RHS} = {}^t x A x + \underbrace{\frac{1}{2}\, {}^t x \cancel{A} \cancel{A^{-1}} b}$$

$$+ \underbrace{\frac{1}{2}\, {}^t b\, \cancel{{}^t A^{-1}} \cancel{A} x}_{A^{-1}} + \cancel{\frac{1}{4}\, {}^t b\, {}^t A^{-1} A A^{-1} b}$$

$$\cancel{- \frac{1}{4}\, {}^t b A^{-1} b} + c$$

$$= {}^t x A x + {}^t b x + c$$

$$= \text{LHS}$$

1.4 積とたたみ込み

X, Y, Z : prob. var.

Z z

<u>case1</u>

$$p(Z=z \mid X=x) = \underline{P_{x,\sigma_x^2}(z)}$$

$\lambda_x = \frac{1}{\sigma_x^2}$

Z は平均 x 分散 σ_x^2 の正規分布に従う

$z = x + \underline{\varepsilon_x} \sim P_{0,\sigma_x^2}$

$$p(Z=z \mid Y=y) = P_{y,\sigma_y^2}(z)$$

$\lambda_y = \frac{1}{\sigma_y^2}$

<u>例1)</u> Z : 物体の重さ

X, Y : 測定結果

x, y が分かったとき Z の値をいくつと思うべきか?

Ans. Zの値は $\dfrac{\lambda_x x+\lambda_y y}{\lambda_x+\lambda_y}$ と選ぶべし

$\because\ p(Z=z \mid X=x, Y=y)$

$$= \frac{p(z,x,y)}{\int p(z,x,y)\,dz}$$

$$= \frac{\sqrt{\frac{\lambda_x}{2\pi}}\exp\left(-\frac{1}{2}\lambda_x(z-x)^2\right)\sqrt{\frac{\lambda_y}{2\pi}}\exp(\omega)}{\int \rule{3cm}{0.4pt}\, dz}$$

平均は重み付き平均。

$$= C\exp\left(-\frac{1}{2}\lambda_x(z-x)^2\right)\exp\left(-\frac{1}{2}\lambda_y(z-y)^2\right)$$

$$= C\exp\left(-\frac{1}{2}\left(\lambda_x(z-x)^2+\lambda_y(z-y)^2\right)\right)$$

$$= C\exp\left(-\frac{1}{2}\left((\lambda_x+\lambda_y)z^2-2(\lambda_x x+\lambda_y y)z + (\text{定})\right)\right)$$

精度は和

$$= D\exp\left(-\frac{1}{2}(\lambda_x+\lambda_y)\left(z-\frac{\lambda_x x+\lambda_y y}{\lambda_x+\lambda_y}\right)^2\right)$$

case2

$$p(Y=y \mid X=x) = p_{x,\sigma_x^2}(y)$$

$$p(Z=z \mid Y=y) = p_{y,\sigma_y^2}(z)$$

このとき、$p(Z=z \mid X=x)$ は?

Ans. $p(Z=z \mid X=x) = p_{x,\sigma_x^2+\sigma_y^2}(z)$

∵ $y = x + \varepsilon_x$

$z = y + \varepsilon_y$

独立

$\Rightarrow z = x + \varepsilon_x + \varepsilon_y$

$E[z] = x + E[\varepsilon_x] + E[\varepsilon_y] = x$

$V[z] = V[\varepsilon_x] + V[\varepsilon_y]$

$= \sigma_x^2 + \sigma_y^2$

1.5 KL-divergence

Kullback-Leibler

Def

p, q : prob. density func.

$$KL[p \| q] = \int p(x) \log \frac{p(x)}{q(x)} dx$$

: KL-div.

Rem

$$KL[p \| q] = E_{p(x)}\left[\log \frac{p(x)}{q(x)}\right]$$

$\log p(x) - \log q(x)$

log確率は本質っぽい量なことが多い。

Fact

- $KL[p \| q] \geq 0$
- $KL[p \| q] = 0 \Longleftrightarrow p = q$
- $KL[p \| q] \neq KL[q \| p]$
- $KL[p \| q]$ は p と q の"キョリの2乗"的に扱うことが多い.

(KLの2階微分が計量を定める
Fisher計量)

Prop

$$KL[P_{\mu,\sigma^2} \| P_{\eta,\tau^2}] = \frac{1}{2}\left(-\log\left(\frac{\sigma^2}{\tau^2}\right) - 1 + \left(\frac{\sigma^2}{\tau^2}\right)\right) + \frac{(\eta-\mu)^2}{2\tau^2}$$

$$H[p] = \int -\log p(x) \quad p(x)\,dx$$