

Các mô hình ngẫu nhiên và ứng dụng:

CHUYỂN ĐỘNG BROWN, MÔ HÌNH HỌC SÂU VÀ NHÚNG DỮ LIỆU

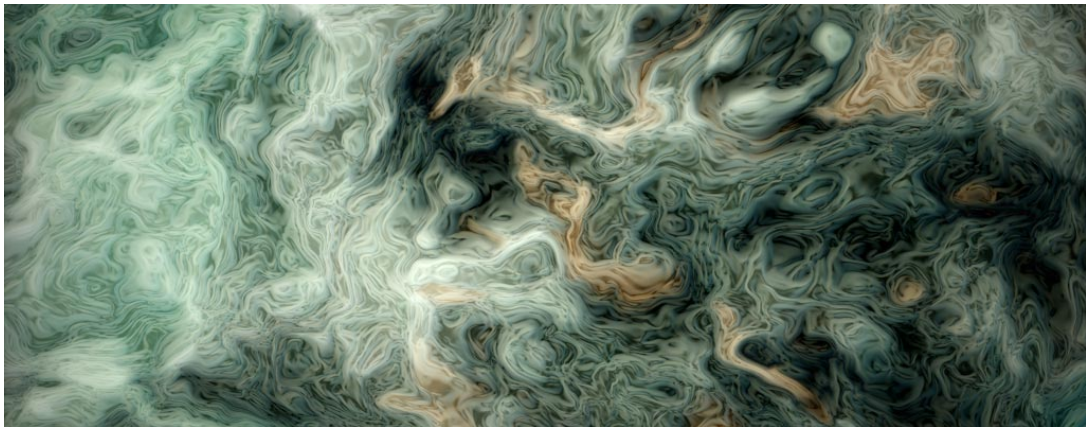
Giảng viên hướng dẫn: TS. Nguyễn Thị Ngọc Anh

Học viên: Nguyễn Đức Hùng ~ 20212498M

Hà Nội, Tháng 03 năm 2023

1. Giới thiệu
2. Chuyển động Brown
3. Chuyển động Brown và Transformer
4. Chuyển động Brown và Δ VAE
5. Kết luận

- ◇ Vật lý
- ◇ Tài chính
- ◇ ???



Hình 1. Inigo Quilez— $f(p) = \text{fBM}(p + \text{fBM}(p + \text{fBM}(p)))^{(1)}$

⁽¹⁾<https://iquilezles.org/articles/warp/>

1. Giới thiệu

2. Chuyển động Brown

3. Chuyển động Brown và Transformer

4. Chuyển động Brown và Δ VAE

5. Kết luận

Mệnh đề 1

Chuyển động Brown hầu chắc chắn không khả vi tại $t > 0$ [2].

[2] Kempthorne Peter và cộng sự. "Lecture 17: Stochastic Processes II". Trong: *Topics in Mathematics with Applications in Finance-18.S096*. MIT OpenCourseWare. Cambridge MA, 2013

$$f(x + \Delta x) - f(x) = \sum_{k=1}^{\infty} \frac{(\Delta x)^k}{k!} f^{(k)}(x) \quad (1)$$

$$\mathbb{E}[(\Delta B_t)^2] = \mathbb{E}[(B_{t+\Delta t} - B_t)^2] \quad (2)$$

$$= \text{Var}[\Delta B_t] + \mathbb{E}[\Delta B_t] = \Delta t. \quad (3)$$

Giải tích

$$df(t, x) = \frac{\partial f}{\partial x} dx + \frac{\partial f}{\partial t} dt, \quad (4)$$

Giải tích Itô

$$\begin{aligned} df(t, B_t) &= \frac{\partial f}{\partial x} dB_t + \frac{\partial f}{\partial t} dt + \frac{1}{2} \cdot \frac{\partial^2 f}{\partial x^2} (dB_t)^2 \\ &= \left(\frac{\partial f}{\partial t} + \frac{1}{2} \cdot \frac{\partial^2 f}{\partial x^2} \right) dt + \frac{\partial f}{\partial x} dB_t. \end{aligned} \quad (5)$$

1. Giới thiệu
2. Chuyển động Brown
3. Chuyển động Brown và Transformer
4. Chuyển động Brown và Δ VAE
5. Kết luận

Định nghĩa 1 (Layer Normalization)

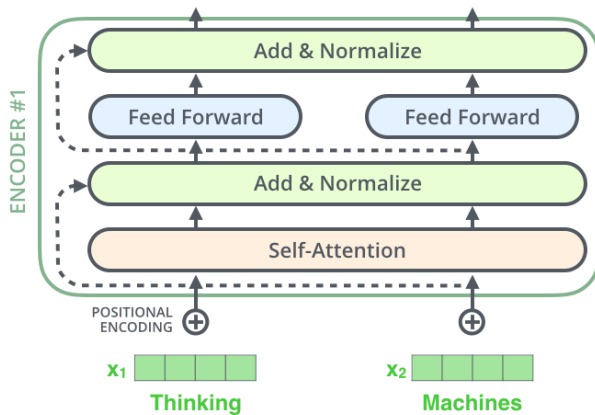
$$\text{LayerNorm}(x) = \frac{x - \mathbb{E}[x]}{\sqrt{\text{Var}[x]}} \cdot \gamma + \beta, \quad (6)$$

Định nghĩa 2 (Softmax)

$$\text{Softmax}(x) = \left[\frac{\exp(x_j)}{\sum_{i=1}^d \exp(x_i)} \right]_{j=\overline{1,d}} \quad (7)$$

Định nghĩa 3 (Scale-dot Attention)

$$\text{Attention}(Q, K, V) = \text{Softmax}\left(\frac{QK^{\top}}{\sqrt{d}}\right) \cdot V \quad (8)$$



Hình 2. Một lớp transformer

Quan sát 1

- ◇ Tính chất chuẩn hoá[4]

$$\|\text{LayerNorm}(x)\| = \sqrt{d}; \quad (9)$$

- ◇ Ma trận chuyển[5]

$$p_{i,j} = \exp\left(\frac{-\|v_i - v_j\|^2}{2\sqrt{d}}\right); \quad (10)$$

- ◇ Thuật toán tối ưu bậc 2[5].

[4]Ruibin Xiong và cộng sự. *On Layer Normalization in the Transformer Architecture*. 2020

[5]Yingshi Chen. *The Brownian motion in the transformer model*. 2021

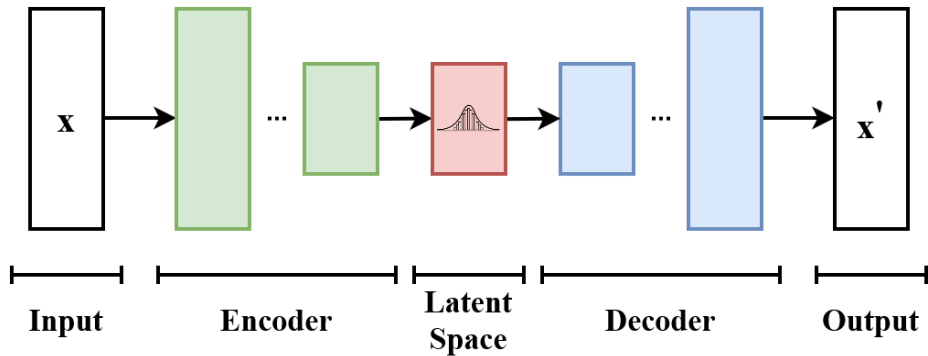
Quan sát 2

- ◇ Multi-head?
- ◇ Kết nối tắt
- ◇ Lớp chiếu

$$y = f(x) + x; \quad (11)$$

$$y = x \cdot A^T + b. \quad (12)$$

1. Giới thiệu
2. Chuyển động Brown
3. Chuyển động Brown và Transformer
4. Chuyển động Brown và Δ VAE
5. Kết luận



Hình 3. Ý tưởng của VAE

- ◇ Không gian biến ẩn Z ;
- ◇ Phân phối biến ẩn \mathbb{P}_Z ;
- ◇ Họ phân phối mã hoá \mathbb{Q}_Z ;
- ◇ Họ phân phối giải mã \mathbb{P}_X ;
- ◇ Hàm đổi biến g .

Định nghĩa 4 (Hàm mục tiêu)

$$L(x) = -\mathbb{E}_{z \sim \mathbb{Q}_Z^{\alpha(x)}} \left[\log p_X^{\beta(z)}(x) \right] + D_{\text{KL}} \left(\mathbb{Q}_Z^{\text{enc}(x)} \parallel \mathbb{P}_Z \right). \quad (13)$$

Sự phù hợp với dữ liệu

- ◇ Phân phối chuẩn/Gauss?
- ◇ Không gian \mathbb{R}^d ?
- ◇ Không gian con bất kỳ?

Định nghĩa 5 (Xây dựng đổi biến[3])

$$g: \Gamma \times (0, \infty) \times Z \rightarrow Z \quad (14)$$

$$g(\gamma, t, z) \sim \mathbb{Q}_Z^{t,z} \quad (15)$$

[3]Luis A. Perez Rey và cộng sự. “Diffusion Variational Autoencoders”. Trong: *Proceedings of the Twenty-Ninth International Joint Conference on Artificial Intelligence, IJCAI-20*. Soạn bởi Christian Bessiere. Main track. International Joint Conferences on Artificial Intelligence Organization, tháng 7 2020, tr. 2704–2710

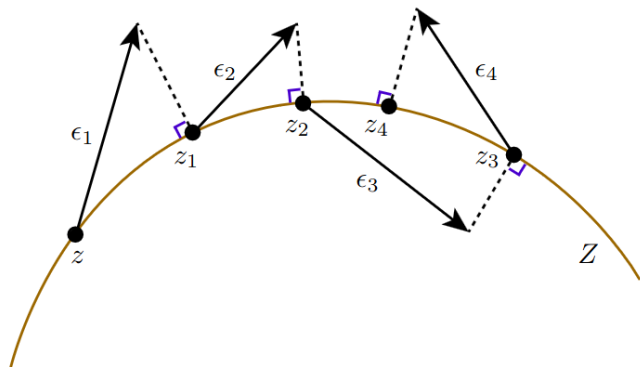
Định nghĩa 6 (Xây dựng du động ngẫu nhiên[3])

$$z_1 = P(z + \sqrt{\tau}\varepsilon_1), \quad (16)$$

$$z_2 = P(P(z + \sqrt{\tau}\varepsilon_1) + \sqrt{\tau}\varepsilon_2) \quad (17)$$

$$\vdots$$

[3] Luis A. Perez Rey và cộng sự. “Diffusion Variational Autoencoders”. Trong: *Proceedings of the Twenty-Ninth International Joint Conference on Artificial Intelligence, IJCAI-20*. Soạn bởi Christian Bessiere. Main track. International Joint Conferences on Artificial Intelligence Organization, tháng 7 2020, tr. 2704–2710



Hình 4. Du động ngẫu nhiên trên một mặt tròn Z . [3]

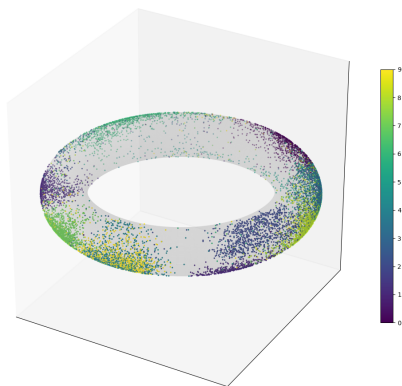
[3] Luis A. Perez Rey và cộng sự. "Diffusion Variational Autoencoders". Trong: *Proceedings of the Twenty-Ninth International Joint Conference on Artificial Intelligence, IJCAI-20*. Soạn bởi

Định nghĩa 7 (Phép đổi biến với chuyển động Brown.[3])

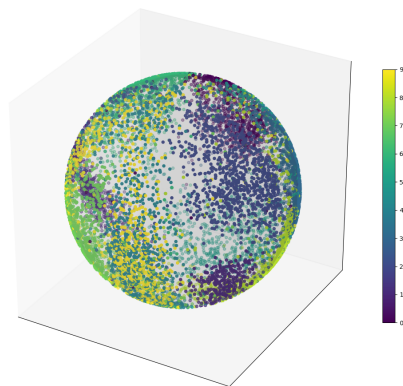
$$g: \mathcal{E}^N \times (0, \infty) \times Z \rightarrow Z, \quad (18)$$

$$g(\varepsilon_1, \dots, \varepsilon_N, t, z) = P\left(\dots P\left(P\left(z + \sqrt{\frac{t}{N}}\varepsilon_1\right) + \frac{t}{N}\varepsilon_2\right)\dots + \sqrt{\frac{t}{N}}\varepsilon_N\right). \quad (19)$$

[3]Luis A. Perez Rey và cộng sự. “Diffusion Variational Autoencoders”. Trong: *Proceedings of the Twenty-Ninth International Joint Conference on Artificial Intelligence, IJCAI-20*. Soạn bởi Christian Bessiere. Main track. International Joint Conferences on Artificial Intelligence Organization, tháng 7 2020, tr. 2704–2710



(a) Mặt hình xuyên



(b) Cầu đơn vị

Hình 5. Dữ liệu MNIST nhúng lên các mặt khác nhau

1. Giới thiệu
2. Chuyển động Brown
3. Chuyển động Brown và Transformer
4. Chuyển động Brown và Δ VAE
5. Kết luận

- ◇ Dùng chuyển động Brown kết hợp VAE để nhúng dữ liệu;
- ◇ Transformer ~ kết hợp, biến đổi chuyển động Brown;
- ◇ Giải thích lỗ hổng trong [5] với ý tưởng trong [3];
- ◇ Nhiều ứng dụng khác.

[5]Yingshi Chen. *The Brownian motion in the transformer model*. 2021

[3]Luis A. Perez Rey và cộng sự. “Diffusion Variational Autoencoders”. Trong: *Proceedings of the Twenty-Ninth International Joint Conference on Artificial Intelligence, IJCAI-20*. Soạn bởi Christian Bessiere. Main track. International Joint Conferences on Artificial Intelligence Organization, tháng 7 2020, tr. 2704–2710

Cảm ơn mọi người đã chú ý lắng nghe!

- [1] Elton P Hsu. “A brief introduction to Brownian motion on a Riemannian manifold”. Trong: *một bài giảng của Viện Toán, Đại học Northwestern* (2008). Mục 1.4 chứa lý thuyết cho phương pháp dùng phép chiếu trong nghiên cứu của Δ VAE.
- [2] Kempthorne Peter và cộng sự. “Lecture 17: Stochastic Processes II”. Trong: *Topics in Mathematics with Applications in Finance–18.S096*. MIT OpenCourseWare. Cambridge MA, 2013.
- [3] Luis A. Perez Rey và cộng sự. “Diffusion Variational Autoencoders”. Trong: *Proceedings of the Twenty-Ninth International Joint Conference on Artificial Intelligence, IJCAI-20*. Soạn bởi Christian Bessiere. Main track. International Joint Conferences on Artificial Intelligence Organization, tháng 7 2020, tr. 2704–2710.

- [4] Ruibin Xiong và cộng sự. *On Layer Normalization in the Transformer Architecture*. 2020.
- [5] Yingshi Chen. *The Brownian motion in the transformer model*. 2021.