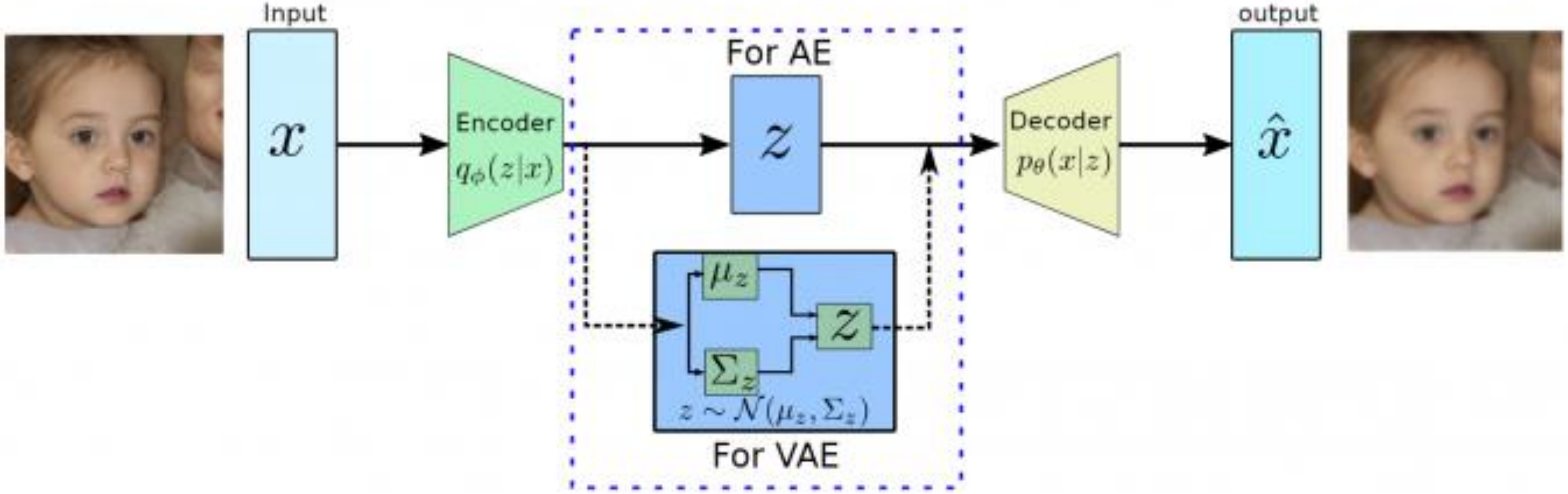
Variational Autoencoder

DEEP LEARNING
IN SCIENCE AND ENGINEERING

Variational Autoencoder (VAE)

自編碼器



變分自編碼器

與傳統自動編碼器不同，變分自動編碼器(VAE)的潛在空間有更加嚴格的結構。VAE假設潛在空間中的表示z遵循某種概率分佈(如高斯分佈)，並且潛在空間是連續的，這使得VAE可以生成新數據。

Variational Autoencoder (VAE)

變分自動編碼器 (Variational Autoencoder, VAE) 是一種生成模型，通過最大化邊際對數似然來學習資料的潛在分佈。其核心是將資料表示為一組潛在變量 z，並使用這些變量生成新的資料樣本。

VAE 通常由兩部分組成：

1. **編碼器 (Encoder)**：將輸入資料 x 映射到潛在空間的隱變量 z。
2. **解碼器 (Decoder)**：從隱變量 z 重構輸入資料 x。

解碼器　　潛在空間

似然(證據效力)　先驗(無證據)

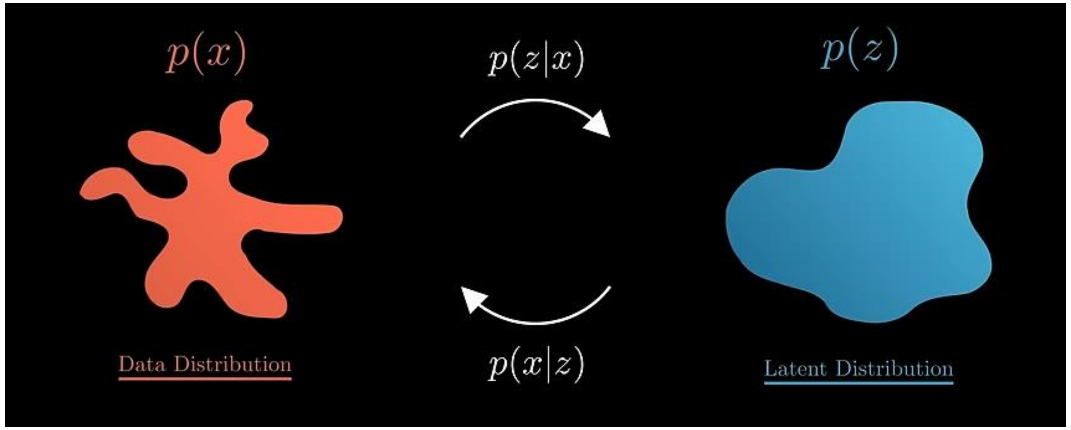
$$p(z \mid x) = \frac{p(x \mid z)\, p(z)}{p(x)}$$

後驗(有證據)　　邊際似然(證據)

編碼器

後驗　　先驗

資料分佈　　編碼　　潛在分佈



Q：證據是什麼？　　邊際似然　　解碼　似然

Variational Autoencoder (VAE)

1. 對數邊際似然

VAE 的目標是最大化輸入資料 x 的對數邊際似然： 使模型生成與輸入資料分佈相似的樣本

邊際似然：所有先驗的似然(所有潛在的解碼) 似然(解碼) 先驗(潛在)

$$\log p_\theta(x) = \log \int p_\theta(x|z)p(z)dz$$

聯合機率

$$p(x) = \int p(x,z)dz$$

p(x)：模型生成輸入x的機率(越大越好)

其中：

$$p(x|z) = \frac{p(x,z)}{p(z)}$$

條件機率

- $p_\theta(x|z)$ 是解碼器的生成分佈，給定隱變量 z 生成資料 x 的條件概率。
- $p(z)$ 是隱變量的先驗分佈，通常假設為標準正態分佈 $\mathcal{N}(0, I)$。

由於對這個積分直接求解非常困難，我們使用變分推論來近似最大化這個目標。

微分：函數求極值 變分：泛函求極值 變分推論：機率分佈的最佳化

Variational Autoencoder (VAE)

2. 變分下界 (ELBO)

後驗分佈(編碼)未知

VAE 使用一個近似的後驗分佈 $q_\phi(z|x)$ 來近似真實的後驗分佈 $p_\theta(z|x)$。我們希望最大化對數邊際似然的變分下界 (ELBO, Evidence Lower Bound)：

$$\log p_\theta(x) \geq \mathcal{L}(\theta, \phi; x) = \mathbb{E}_{q_\phi(z|x)}[\log p_\theta(x|z)] - \mathrm{KL}(q_\phi(z|x)||p(z))$$

稍後說明

重構誤差×(−1)　推論誤差

其中：

- 第一項 $\mathbb{E}_{q_\phi(z|x)}[\log p_\theta(x|z)]$ 是重構誤差，衡量模型能多好地重構輸入資料 x。
- 第二項 $\mathrm{KL}(q_\phi(z|x)||p(z))$ 是編碼器生成的近似後驗分佈 $q_\phi(z|x)$ 與先驗分佈 $p(z)$ 之間的 KL 散度。

近似後驗分佈也是未知？　先驗已知？

最大化ELBO：最小化(重構誤差+推論誤差)

Variational Autoencoder (VAE)

3. 重構誤差

這部分衡量解碼器在給定隱變量 z 時，能否很好地重構輸入資料 x：

$$\mathbb{E}_{q_\phi(z|x)}[\log p_\theta(x|z)]$$ 稍後說明

如果 $p_\theta(x|z)$ 是高斯分佈，則這項可以解釋為資料 x 與生成樣本的對數似然。

4. KL 散度

KL 散度用來測量編碼器的近似後驗分佈 $q_\phi(z|x)$ 與先驗分佈 $p(z)$ 的差異：

$$\mathrm{KL}(q_\phi(z|x)||p(z)) = \frac{1}{2}\sum_{i=1}^{d}\left(\sigma_\phi^2(x)_i + \mu_\phi(x)_i^2 - 1 - \log(\sigma_\phi^2(x)_i)\right)$$ 稍後證明

其中 $\mu_\phi(x)$ 和 $\sigma_\phi^2(x)$ 是編碼器對隱變量的均值和方差。

Variational Autoencoder (VAE)

5. 總結

VAE 的最終目標是最大化 ELBO：

對數似然的期望值　　近似後驗　先驗

$$\mathcal{L}(\theta, \phi; x) = \mathbb{E}_{q_\phi(z|x)}[\log p_\theta(x|z)] - \mathrm{KL}(q_\phi(z|x)||p(z))$$

稍後說明

這同時考慮了生成樣本的質量（重構誤差）和隱變量分佈的正則化（KL 散度）。

強制模型的後驗分佈接近先驗分佈

目標：最大化對數邊際似然
手段：最大化證據下界
意義：最小化重構+推論誤差

重構誤差：輸入x和輸出f(z)均方誤差的期望值
推論誤差：近似後驗q(z|x)與先驗p(z)的分佈距離(KL散度)

Q：為何後驗機率要用近似函數？

VAE變分推論的主要想法：最佳化近似後驗(編碼器)來最大化邊際似然(所有潛在空間的解碼器)

Evidence Lower Bound (ELBO)

ELBO（證據下界）的推導基於變分推論，用來近似複雜的對數邊際似然 $\log p_\theta(x)$。我們希望最大化這個下界來逼近對數邊際似然。這裡，我將詳細推導 ELBO 的方程式。

1. 對數邊際似然

我們的目標是最大化對數邊際似然 $\log p_\theta(x)$：

$$\log p_\theta(x) = \log \int p_\theta(x|z)p(z)dz$$

其中 $p_\theta(x|z)$ 是解碼器的生成分佈，給定潛在變量 z 生成 x 的條件概率，而 $p(z)$ 是先驗分佈，通常假設為標準正態分佈 $\mathcal{N}(0, I)$。

由於對這個積分直接求解很困難，我們引入一個近似後驗分佈 $q_\phi(z|x)$，用來近似真實後驗分佈 $p_\theta(z|x)$。

Evidence Lower Bound (ELBO)

2. 引入變分近似

我們在對數邊際似然中引入後驗分佈 $q_\phi(z|x)$ 來進行變分推導：

Q：which is PDF?

$$\log p_\theta(x) = \log \int p_\theta(x|z)p(z)dz = \log \int \frac{q_\phi(z|x)}{q_\phi(z|x)} p_\theta(x|z)p(z)dz$$

這個等式成立，因為 $q_\phi(z|x)$ 是一個分佈，因此分母和分子相等。

$$\mathrm{E}[X] = \int_{-\infty}^{\infty} x f(x)\, dx.$$

f **: probability density function**

$$= \log \mathbf{E}_{q_\phi(z|x)}\left[\frac{p_\theta(x|z)p(z)}{q_\phi(z|x)}\right]$$

Jensen's inequality $\geq \mathbf{E}_{q_\phi(z|x)}\left[\log \frac{p_\theta(x|z)p(z)}{q_\phi(z|x)}\right]$ **稍後證明**

$$\underbrace{}_{\text{ELBO}}$$

Evidence Lower Bound (ELBO)

我們將右邊的期望進一步分解：

$$\mathbb{E}_{q_\phi(z|x)}\left[\log \frac{p_\theta(x|z)p(z)}{q_\phi(z|x)}\right] = \mathbb{E}_{q_\phi(z|x)}[\log p_\theta(x|z)] + \mathbb{E}_{q_\phi(z|x)}\left[\log \frac{p(z)}{q_\phi(z|x)}\right]$$

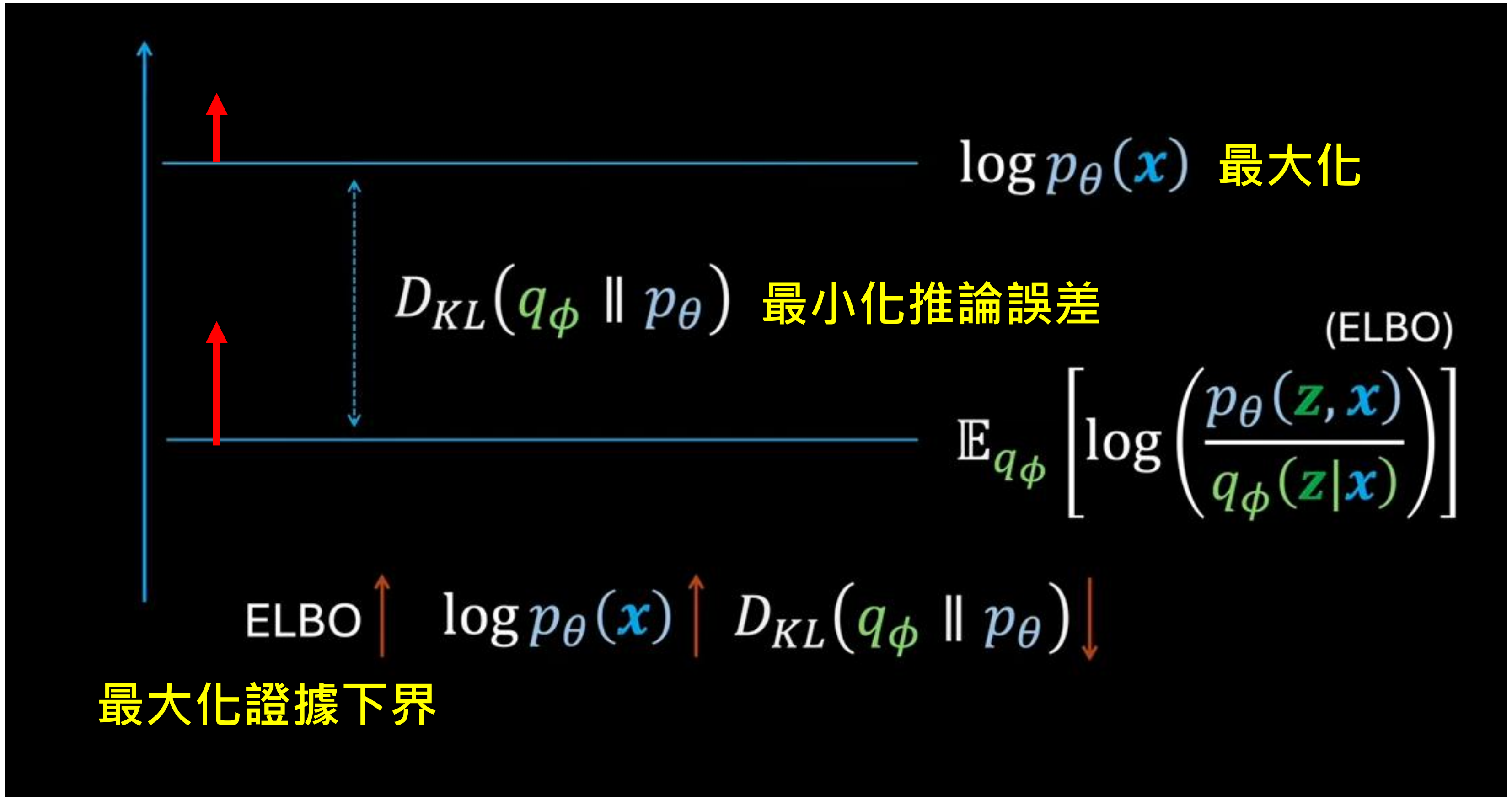
重構誤差×(–1) **變分推論誤差×(–1)**

其中：

- 第一項 $\mathbb{E}_{q_\phi(z|x)}[\log p_\theta(x|z)]$ 是重構誤差，表示模型能多好地重構輸入資料 x。
- 第二項 $\mathbb{E}_{q_\phi(z|x)}\left[\log \frac{p(z)}{q_\phi(z|x)}\right]$ 是 $q_\phi(z|x)$ 與 $p(z)$ 之間的 KL 散度。 **稍後證明**

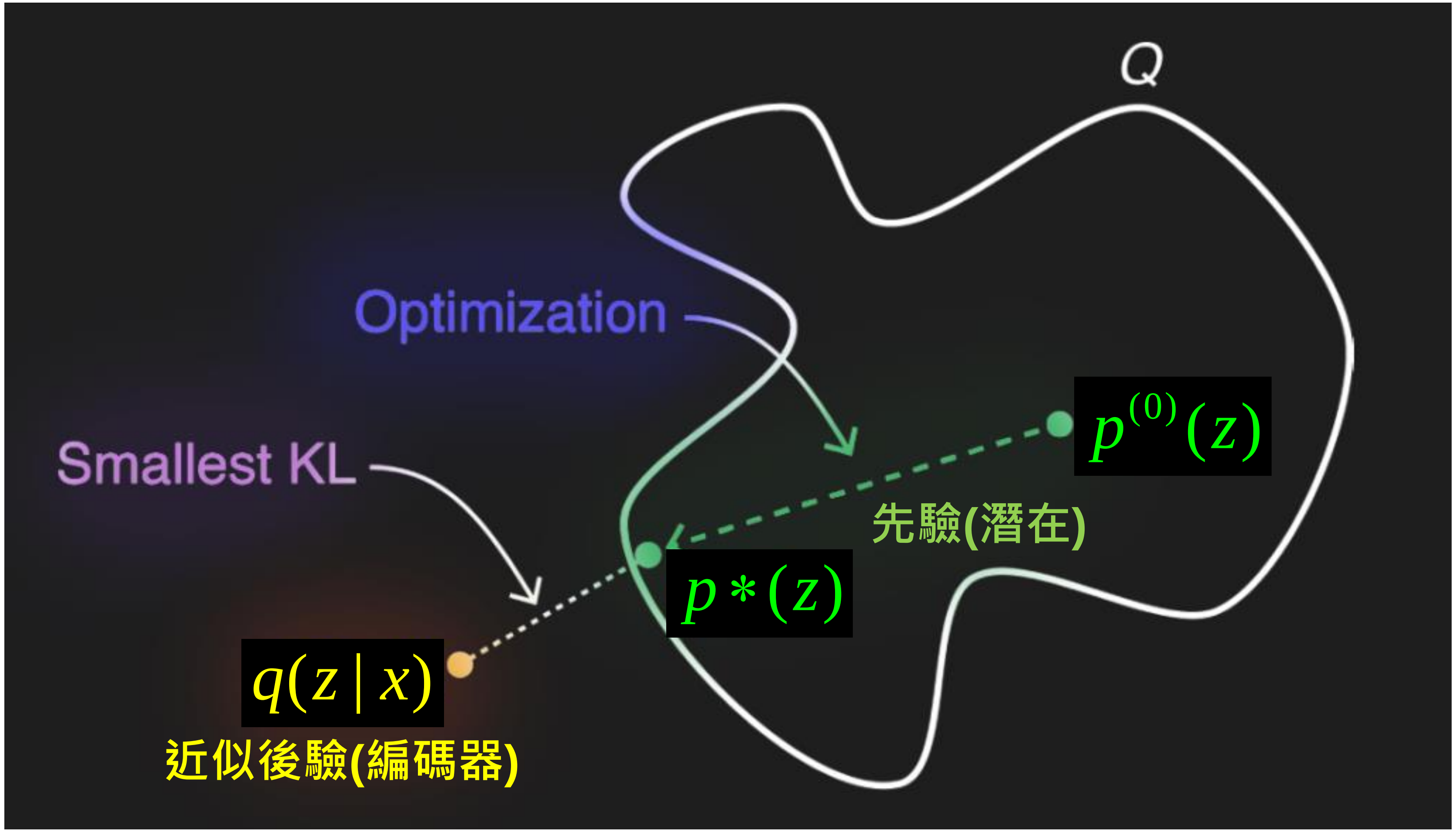
在高斯生成模型下，VAE 的重構誤差項 $\mathbb{E}_{q(z|x)}[\log p(x|z)]$ 的最大化等價於最小化均方誤差（MSE）。**稍後證明**

Evidence Lower Bound (ELBO)



最大化證據下界：最大化邊際似然、最小化推論誤差

Evidence Lower Bound (ELBO)



Evidence Lower Bound (ELBO)

在 VAE 中，要求近似後驗分佈與先驗分佈接近的原因可以從以下幾個方面來理解：

編碼器(後驗)的結果與潛在空間(先驗)分佈一致

(1) 保證生成樣本的合理性

若 $q(z|x)$ 與 $p(z)$ 接近，則表示在訓練過程中學到的潛在空間 z 與 $p(z)$ 的標準高斯分佈一致。這樣，當我們從 $p(z)$ 中採樣 z 並通過解碼器生成新數據時，生成的數據會更接近於訓練數據的分佈。如果 $q(z|x)$ 與 $p(z)$ 相差很大，解碼器將無法很好地解碼來自先驗分佈的樣本 z，從而生成的數據樣本可能無法擬合原始數據的分佈。

(2) 確保潛在空間的平滑性

當 $q(z|x)$ 與 $p(z)$ 接近時，不同數據樣本的潛在變量 z 分佈在潛在空間中會更加連續和平滑。這意味著潛在空間中的鄰近點對應著具有相似特徵的數據樣本。這樣的潛在空間平滑性在生成新樣本時會非常重要，因為從先驗分佈中隨機選取的潛在變量也將具有一致的生成效果。

Reconstruction Error

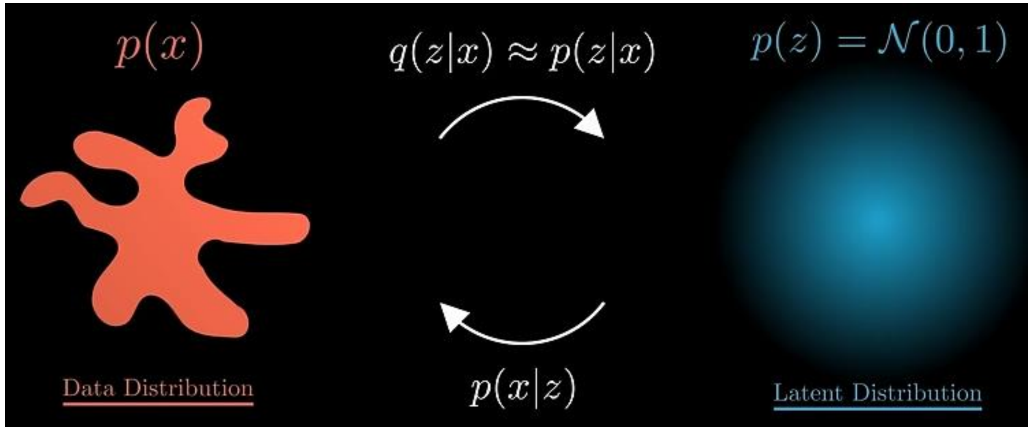
假設

假設生成分佈 $p(x|z)$ 為高斯分佈，均值由解碼器輸出 $f(z)$ 給出，方差為固定的 σ^2，即：

$$p(x|z) = \mathcal{N}(x; f(z), \sigma^2 I)$$

這裡 $f(z)$ 是解碼器給定 z 的條件下對 x 的預測值。

後驗(編碼)



似然(解碼)

重構誤差項的定義

在 VAE 的目標中，重構誤差項定義為：

對數似然在近似後驗分佈的期望值 $\mathbb{E}_{q(z|x)}[\log p(x|z)]$

這表示給定編碼分佈 $q(z|x)$ 下的期望，即期望重構樣本 x 的概率對數。

Reconstruction Error

計算 $\log p(x|z)$

將 $p(x|z)$ 代入得到 $\log p(x|z)$ 的具體表達式：

$$\log p(x|z) = \log\left(\frac{1}{\sqrt{(2\pi\sigma^2)^d}}\exp\left(-\frac{1}{2\sigma^2}\|x - f(z)\|^2\right)\right)$$

其中 d 是數據 x 的維度。

展開對數項，我們得到：

$$\log p(x|z) = -\frac{d}{2}\log(2\pi\sigma^2) - \frac{1}{2\sigma^2}\|x - f(z)\|^2$$

可以看到，這裡包含兩部分：常數項 $-\frac{d}{2}\log(2\pi\sigma^2)$ 和一項與 $\|x - f(z)\|^2$ 成比例的項。

Reconstruction Error

計算重構誤差的期望

重構誤差項的期望為：

$$\mathbb{E}_{q(z|x)}[\log p(x|z)] = \mathbb{E}_{q(z|x)}\left[-\frac{d}{2}\log(2\pi\sigma^2) - \frac{1}{2\sigma^2}\|x - f(z)\|^2\right]$$

由於常數項與 z 無關，我們可以將其直接提出期望之外，得：

$$\mathbb{E}_{q(z|x)}[\log p(x|z)] = -\frac{d}{2}\log(2\pi\sigma^2) - \frac{1}{2\sigma^2}\mathbb{E}_{q(z|x)}\left[\|x - f(z)\|^2\right]$$

得出均方誤差項

在這個表達式中，第二項 $\mathbb{E}_{q(z|x)}\left[\|x - f(z)\|^2\right]$ 是期望的均方誤差（MSE）：

$$\mathbb{E}_{q(z|x)}\left[\|x - f(z)\|^2\right]$$

編碼再解碼後的相似性

KL Divergence

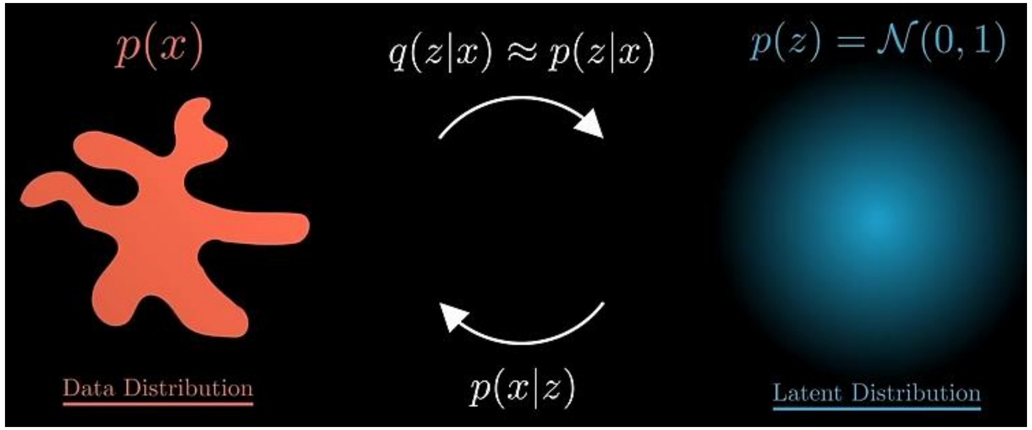
在**變分自編碼器（Variational Autoencoder, VAE）**中，KL 散度（Kullback-Leibler Divergence）度量的是潛在變量的近似分佈 $q(z|x)$ 與先驗分佈 $p(z)$ 的差異。VAE 通常選擇 $p(z)$ 為標準正態分佈 $\mathcal{N}(0,1)$，而 $q(z|x)$ 則是從編碼器得到的分佈 $\mathcal{N}(\mu,\sigma^2)$。以下是 KL 散度的推導過程：

近似後驗

1. 定義 KL 散度

KL 散度可以表達為：

$$D_{\mathrm{KL}}\left(q(z|x) \,\|\, p(z)\right) = \mathbb{E}_{q(z|x)}\left[\log \frac{q(z|x)}{p(z)}\right]$$



KL Divergence

2. 帶入具體分佈

假設 $q(z|x) = \mathcal{N}(\mu, \sigma^2)$ 和 $p(z) = \mathcal{N}(0, 1)$，根據高斯分佈的概率密度函數，我們可以表示這兩個分佈的概率密度為：

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

$$p(z) = \frac{1}{\sqrt{2\pi}} \exp\left(-\frac{z^2}{2}\right)$$

3. 代入 KL 散度公式

將 $q(z|x)$ 和 $p(z)$ 代入 KL 散度的表達式，得：

$$D_{\mathrm{KL}}\left(q(z|x) \parallel p(z)\right) = \mathbb{E}_{q(z|x)}\left[\log \frac{\frac{1}{\sqrt{2\pi\sigma^2}} \exp\left(-\frac{(z-\mu)^2}{2\sigma^2}\right)}{\frac{1}{\sqrt{2\pi}} \exp\left(-\frac{z^2}{2}\right)}\right]$$

KL Divergence

4. 展開對數項

簡化對數項：

$$D_{\mathrm{KL}}\left(q(z|x) \parallel p(z)\right) = \mathbb{E}_{q(z|x)}\left[\log \frac{1}{\sqrt{2\pi\sigma^2}} - \frac{(z-\mu)^2}{2\sigma^2} + \frac{z^2}{2}\right]$$

進一步分解得到：

$$= \mathbb{E}_{q(z|x)}\left[-\frac{1}{2}\log(2\pi\sigma^2) - \frac{(z-\mu)^2}{2\sigma^2} + \frac{z^2}{2}\right]$$

KL Divergence

5. 計算期望

$$\mathbf{E}_{q(z|x)}[f(z)] = \int_{-\infty}^{\infty} f(z) q(z|x) dz$$

我們現在可以對每一項分別求期望：

- **第一項**：$-\frac{1}{2}\log(2\pi\sigma^2)$，這項不依賴於 z，因此其期望值不變。
- **第二項**：$\mathbb{E}_{q(z|x)}\left[-\frac{(z-\mu)^2}{2\sigma^2}\right]$。因為 $z \sim \mathcal{N}(\mu, \sigma^2)$，所以 $\mathbb{E}_{q(z|x)}[(z-\mu)^2] = \sigma^2$，代入得：

稍後證明

$$-\frac{\sigma^2}{2\sigma^2} = -\frac{1}{2}$$

- **第三項**：$\mathbb{E}_{q(z|x)}\left[\frac{z^2}{2}\right]$。對於 $z \sim \mathcal{N}(\mu, \sigma^2)$，$\mathbb{E}[z^2] = \mu^2 + \sigma^2$，所以：

稍後證明

$$\frac{\mu^2 + \sigma^2}{2}$$

KL Divergence

6. 合併結果

將以上三項結果相加，得到 KL 散度：

$$D_{\mathrm{KL}}\left(q(z|x) \parallel p(z)\right) = -\frac{1}{2}\log(2\pi\sigma^2) - \frac{1}{2} + \frac{\mu^2 + \sigma^2}{2}$$

再進一步簡化得到：

$$= \frac{1}{2}\left(\mu^2 + \sigma^2 - 1 - \log \sigma^2\right)$$

KL Divergence

In[27]:= $\mathbf{EX}[f\_] := \mathbf{Normal}\left[\int_{-\infty}^{\infty} f \frac{1}{\sqrt{2\pi\sigma^2}} e^{-\frac{(z-\mu)^2}{2\sigma^2}} \, dz\right]$ // PowerExpand

$\mathbf{EX}[z]$

$\mathbf{EX}\left[(z-\mu)^2\right]$

$\mathbf{EX}\left[z^2\right]$

Out[28]= μ

Out[29]= σ^2

Out[30]= $\mu^2 + \sigma^2$

KL Divergence and Expected Value

$$D_{\mathrm{KL}}(P \parallel Q) = \sum_x P(x) \log \frac{P(x)}{Q(x)}$$ 離散

$$D_{\mathrm{KL}}(P \parallel Q) = \int P(x) \log \frac{P(x)}{Q(x)} \, dx$$ 連續

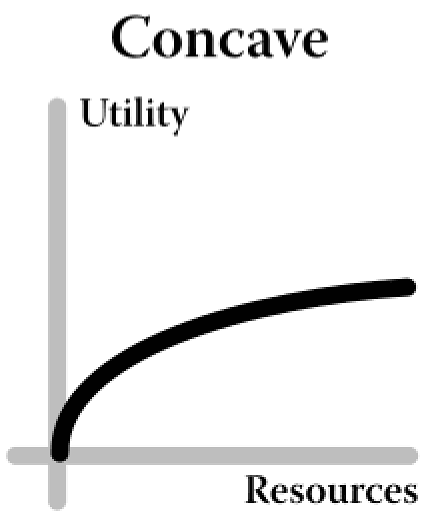
期望值 $\mathbb{E}[X] = \sum_x x \cdot P(x)$ 或 $\mathbb{E}[X] = \int x \cdot f(x) \, dx$

KL 散度可以看作是分佈 P 對於 $\log \frac{P(x)}{Q(x)}$ 的期望值：

$$D_{\mathrm{KL}}(P \parallel Q) = \mathbb{E}_P \left[\log \frac{P(x)}{Q(x)} \right]$$

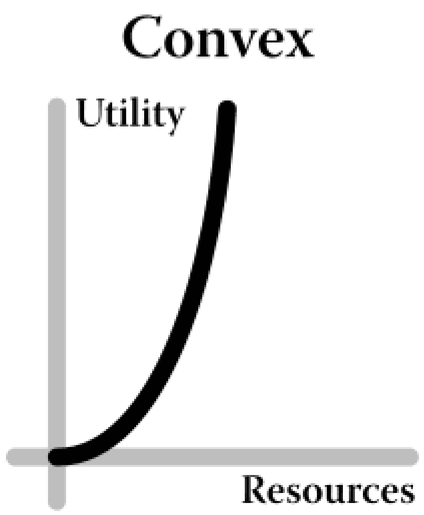
Concave and Convex

凹的

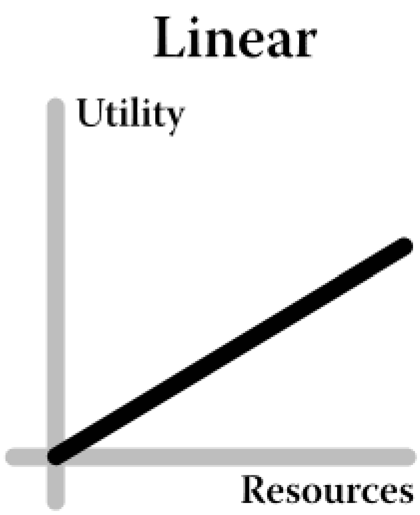


bends down

凸的



bends up

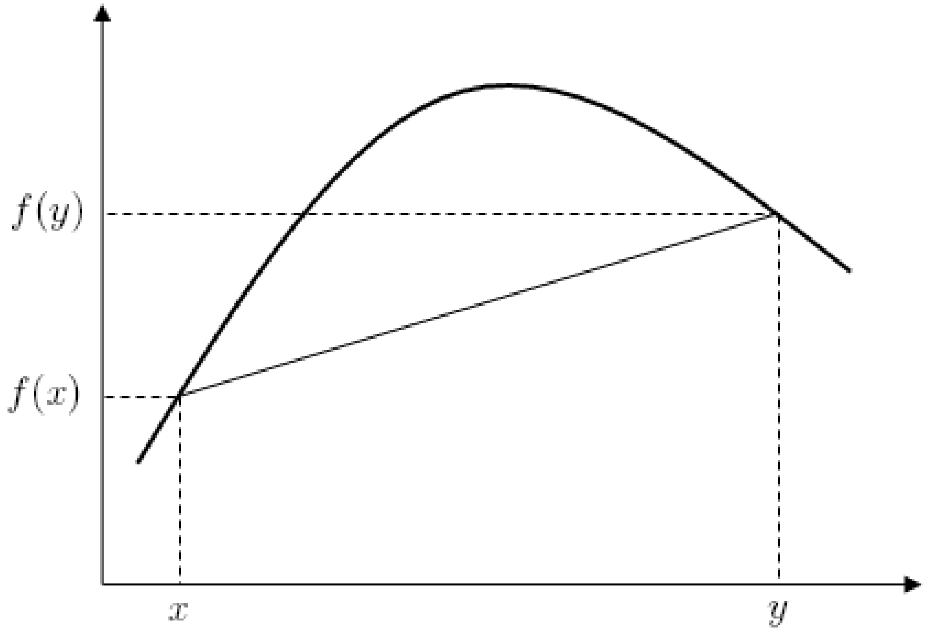


Concave Function

凹函數 f 的定義是：

$$f(\lambda x_1 + (1-\lambda)x_2) \geq \lambda f(x_1) + (1-\lambda) f(x_2)$$

對於所有 $x_1, x_2 \in \mathbb{R}$，以及 $\lambda \in [0,1]$。



Jensen's Inequality

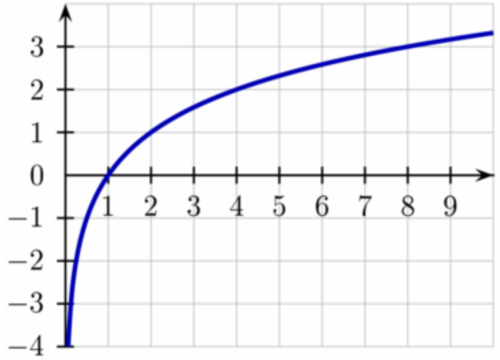
假設隨機變量 X 可以取 $x_1, x_2, \ldots, x_n$ 這些離散值，且對應的概率為 $p(x_1), p(x_2), \ldots, p(x_n)$，這些概率滿足 $\sum_{i=1}^{n} p(x_i) = 1$。

則 X 的期望值可以寫作：

$$\mathbb{E}[X] = \sum_{i=1}^{n} p(x_i) x_i$$

根據凹函數的定義，我們可以直接應用凹性的性質。由於 f 是凹函數，對於凸組合 $\sum_{i=1}^{n} p(x_i) x_i$ 而言，我們有：

Q：對數函數是凹的？



$$f\left(\sum_{i=1}^{n} p(x_i) x_i\right) \geq \sum_{i=1}^{n} p(x_i) f(x_i) \qquad f(\mathbb{E}[X]) \geq \mathbb{E}[f(X)],$$

對數 $\log \mathbb{E}[p_\theta(x|z)] \geq \mathbb{E}[\log p_\theta(x|z)]$

Reparameterization Trick

在變分自編碼器（Variational Autoencoder, VAE）中，**重參數化技巧（Reparameterization Trick）**的引入是為了解決隨機變量導致的反向傳播問題。這裡詳細說明為何需要重參數化技巧，以及它如何幫助進行反向傳播。

問題的核心：隨機樣本導致的反向傳播困難

VAE 的損失函數包含了隨機變量 z，它從編碼器生成的後驗分佈 $q(z|x)$ 中進行抽樣。具體來說，給定輸入 x，編碼器會生成潛在變量 z 的均值 μ 和標準差 σ：

$$z \sim \mathcal{N}(\mu, \sigma^2)$$

在這裡，z 的取值包含隨機性，導致以下問題：**隨機樣本 z 的引入使得損失函數相對於編碼器參數 μ 和 σ 的梯度無法直接計算**。這是因為隨機樣本的抽取過程會打斷梯度流，造成無法對編碼器進行端到端的優化。

Reparameterization Trick

重參數化技巧的解決方案

為了解決上述問題，重參數化技巧重新表示隨機變量 z 為一個**確定性函數**，使得我們可以繞開隨機性來計算梯度。

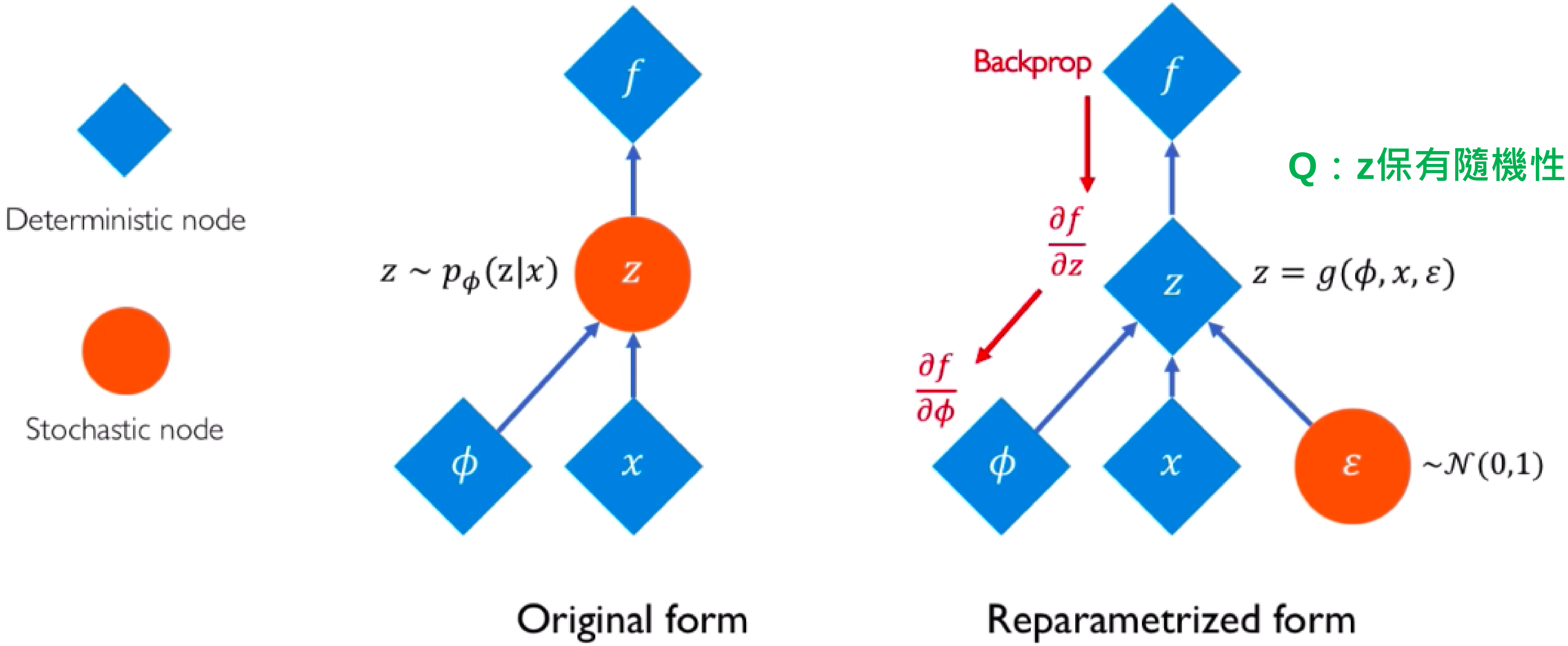
1. **引入隨機噪聲**：我們從標準正態分佈抽取隨機噪聲 $\epsilon \sim \mathcal{N}(0,1)$。
2. **重參數化表示**：將 z 表示為一個確定性函數：

$$z = \mu + \sigma \cdot \epsilon$$

   在這裡，μ 和 σ 是編碼器輸出的確定性變量，ϵ 是標準正態隨機變量。

這種重參數化表示將隨機性從 z 中移除，因為 z 現在只是 μ 和 σ 的確定性函數。

Reparameterization Trick

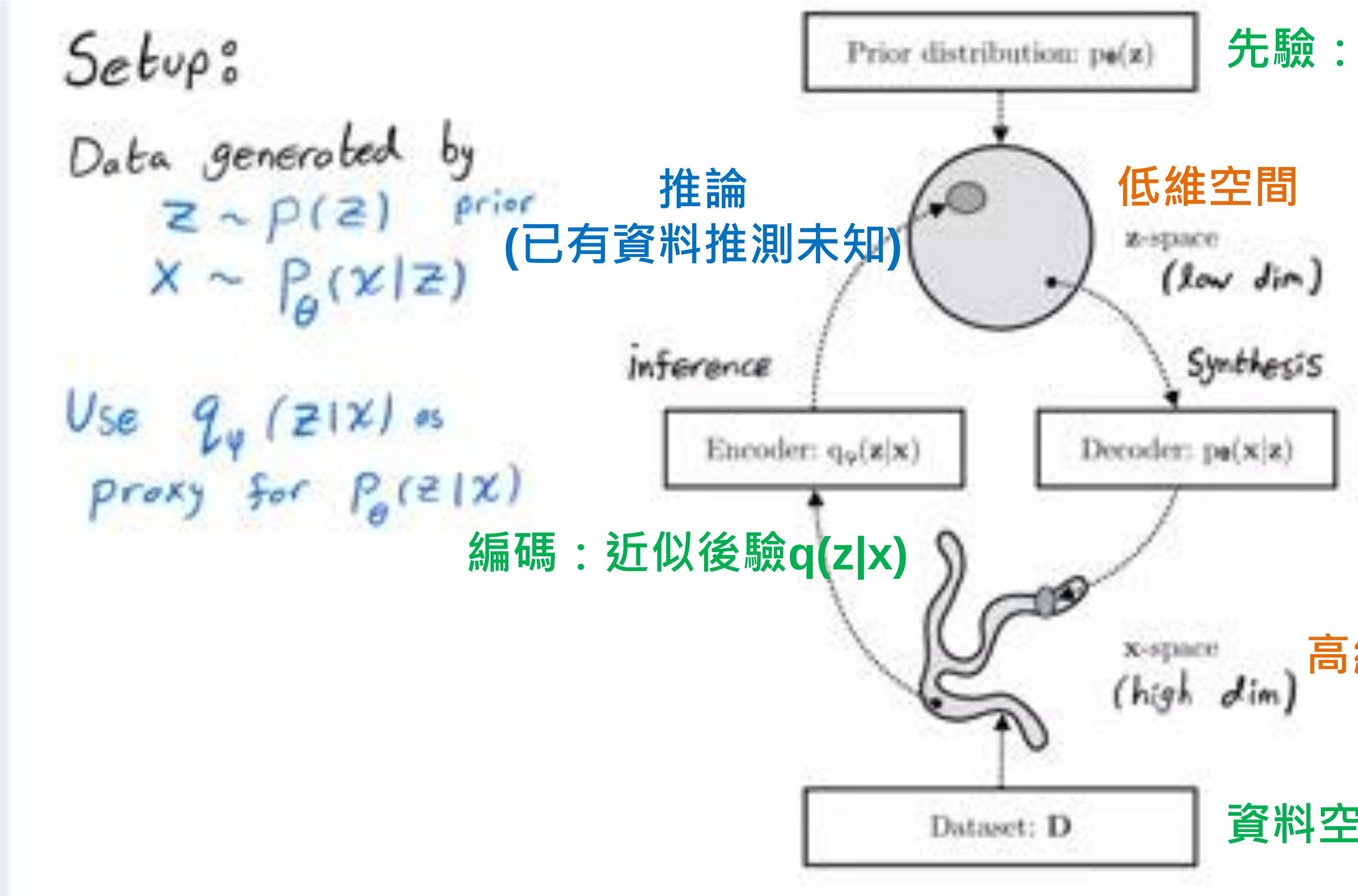


Q：z保有隨機性嗎？

z由取樣產生、無函數形式

z非取樣產生、有函數形式

Variational Autoencoder (VAE)



先驗：p(z)

推論
(已有資料推測未知)

低維空間

合成
(已知信息生成內容)

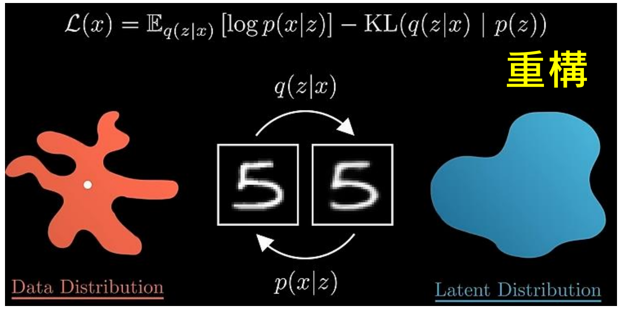
解碼：似然p(x|z)

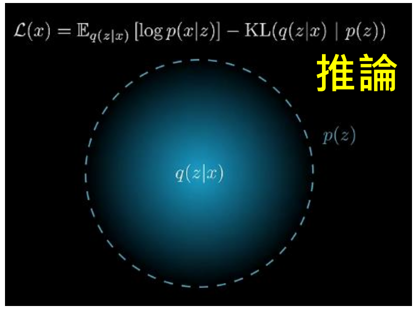
編碼：近似後驗q(z|x)

高維空間

資料空間

Variational Autoencoder (VAE)

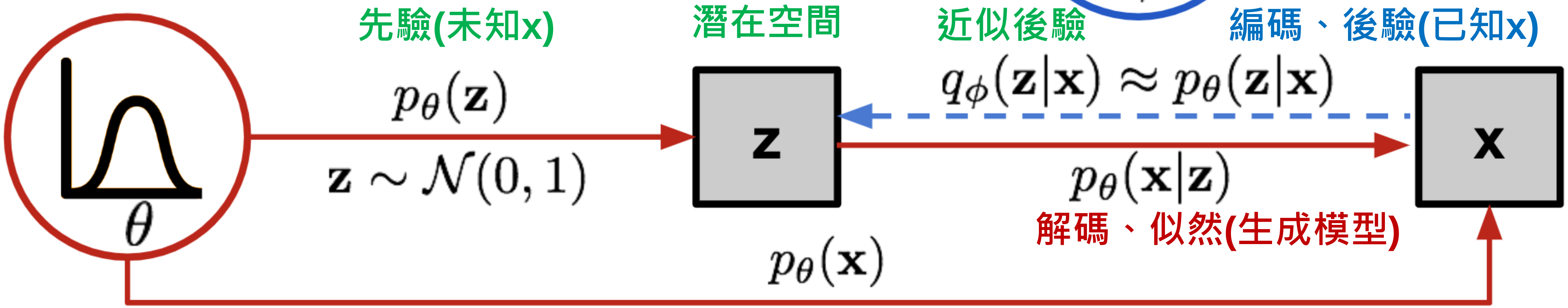




假設：高斯分佈

假設：單位高斯分佈

Q：為何需要近似後驗？

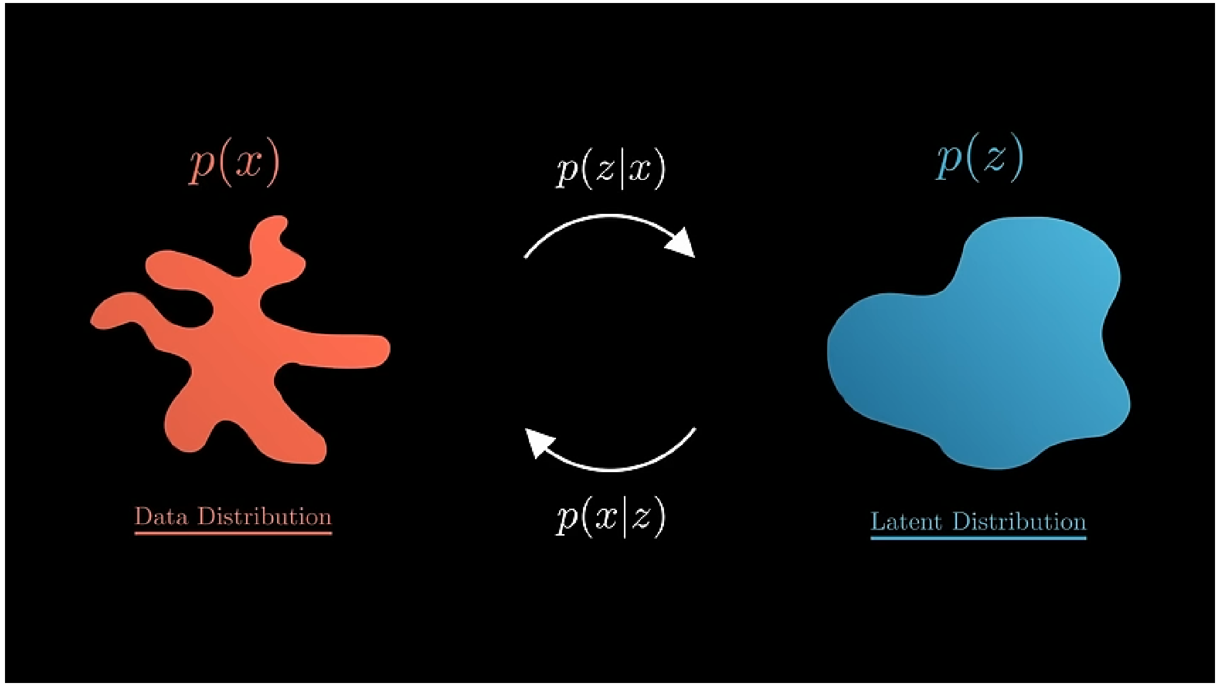


邊際似然、生成模型輸出x的機率

Q：為何不直接求p(x)的最大值？

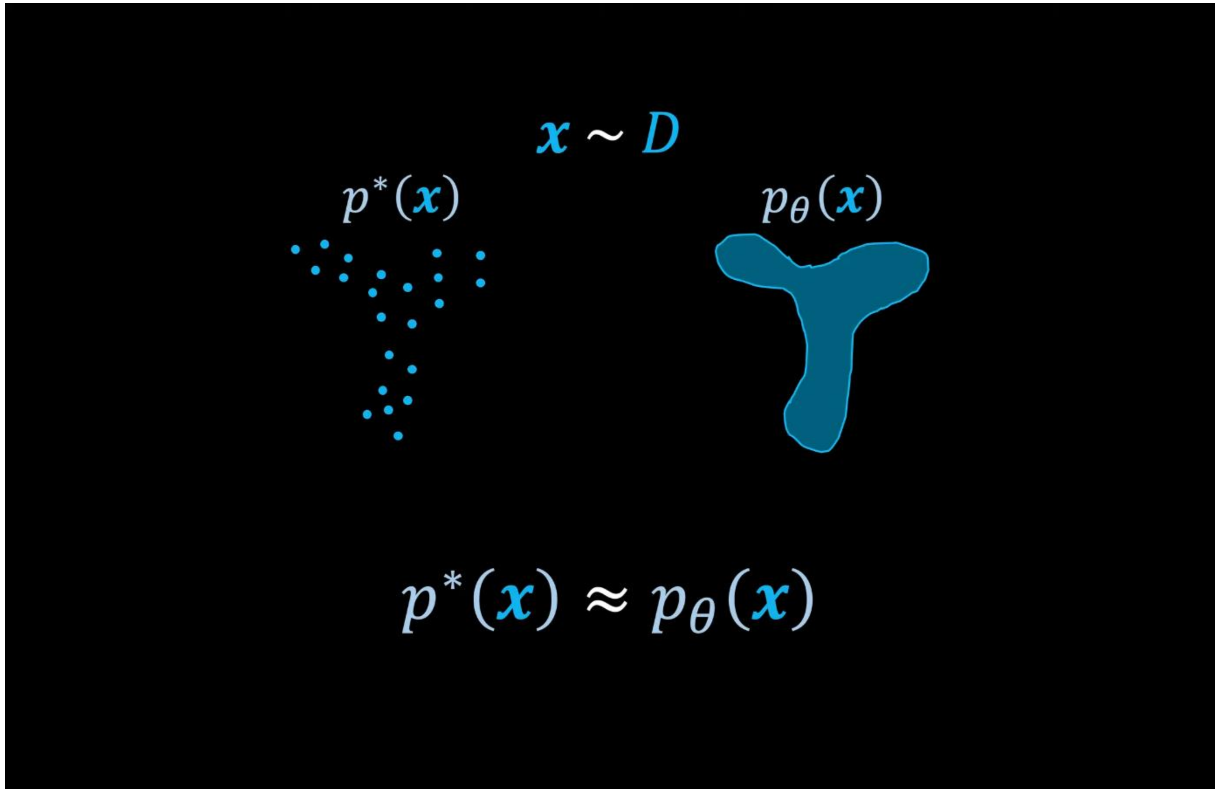
Variational Autoencoder (VAE)

🕗 **20:09**



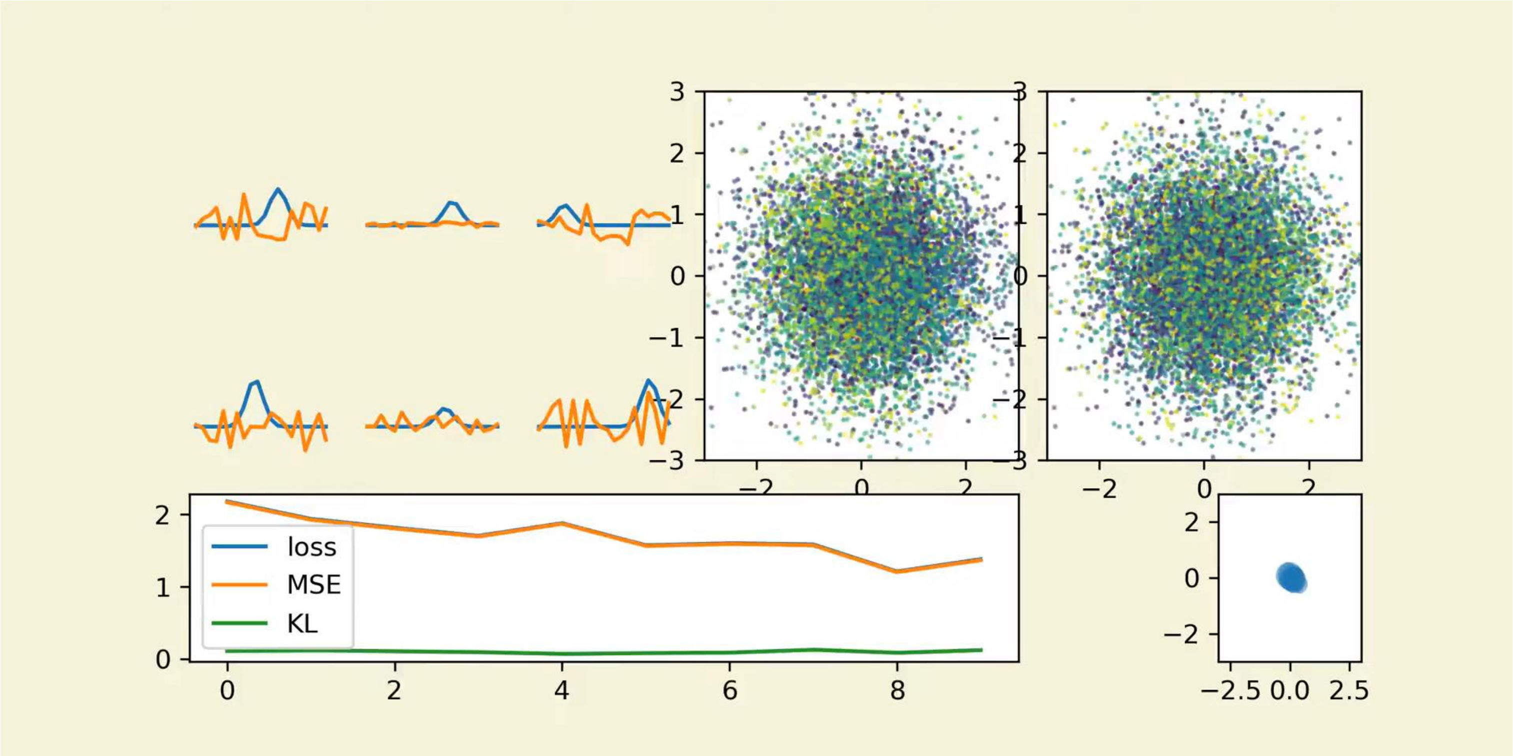
Variational Autoencoder (VAE)

🕖 29:53



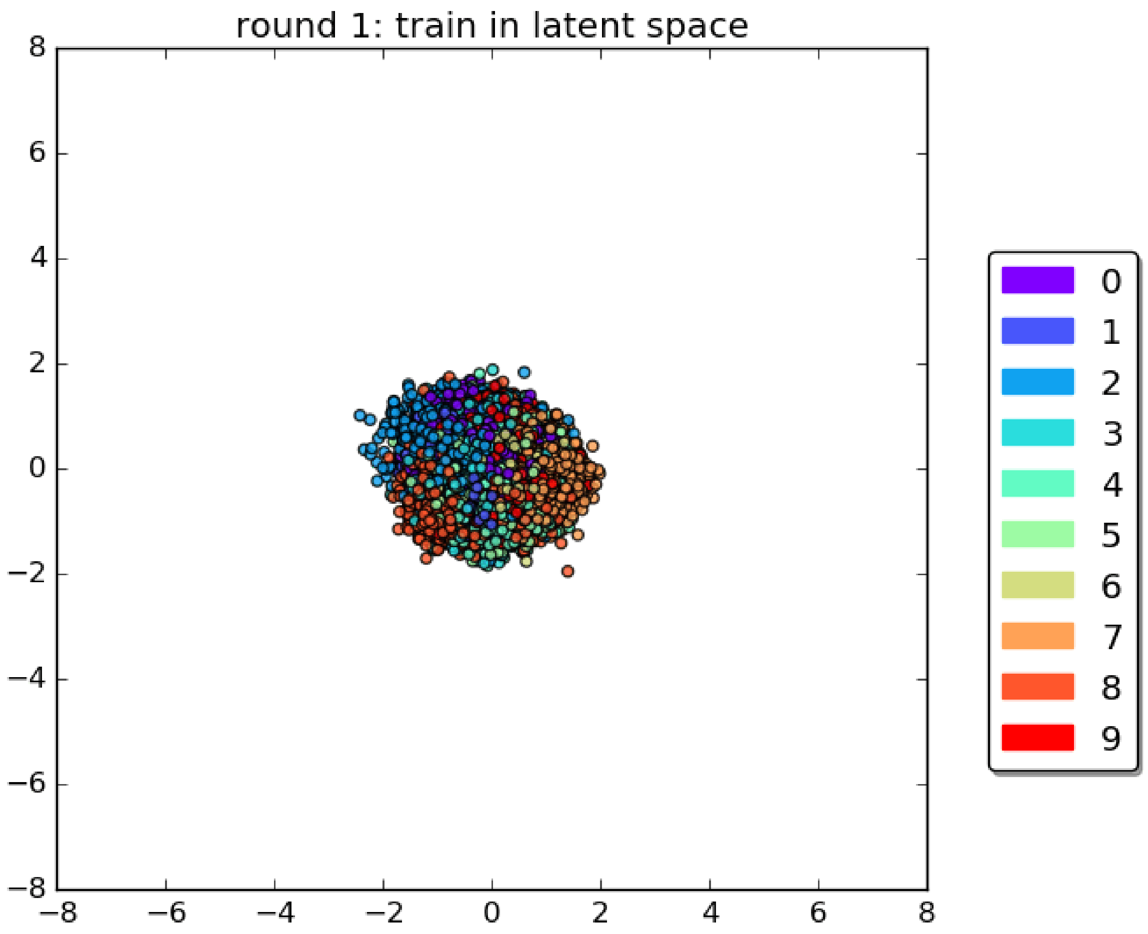
Variational Autoencoder (VAE)

資料
解碼

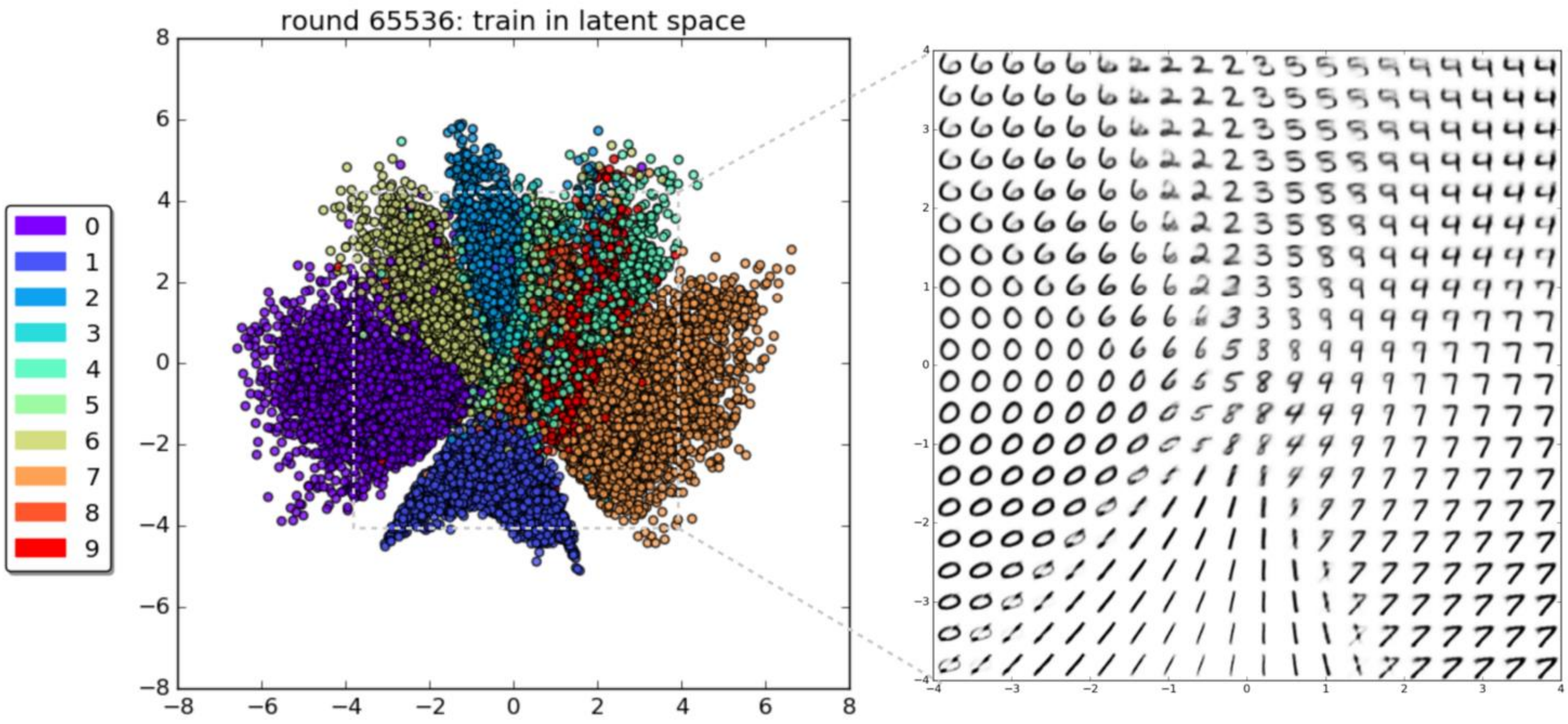


潛在空間

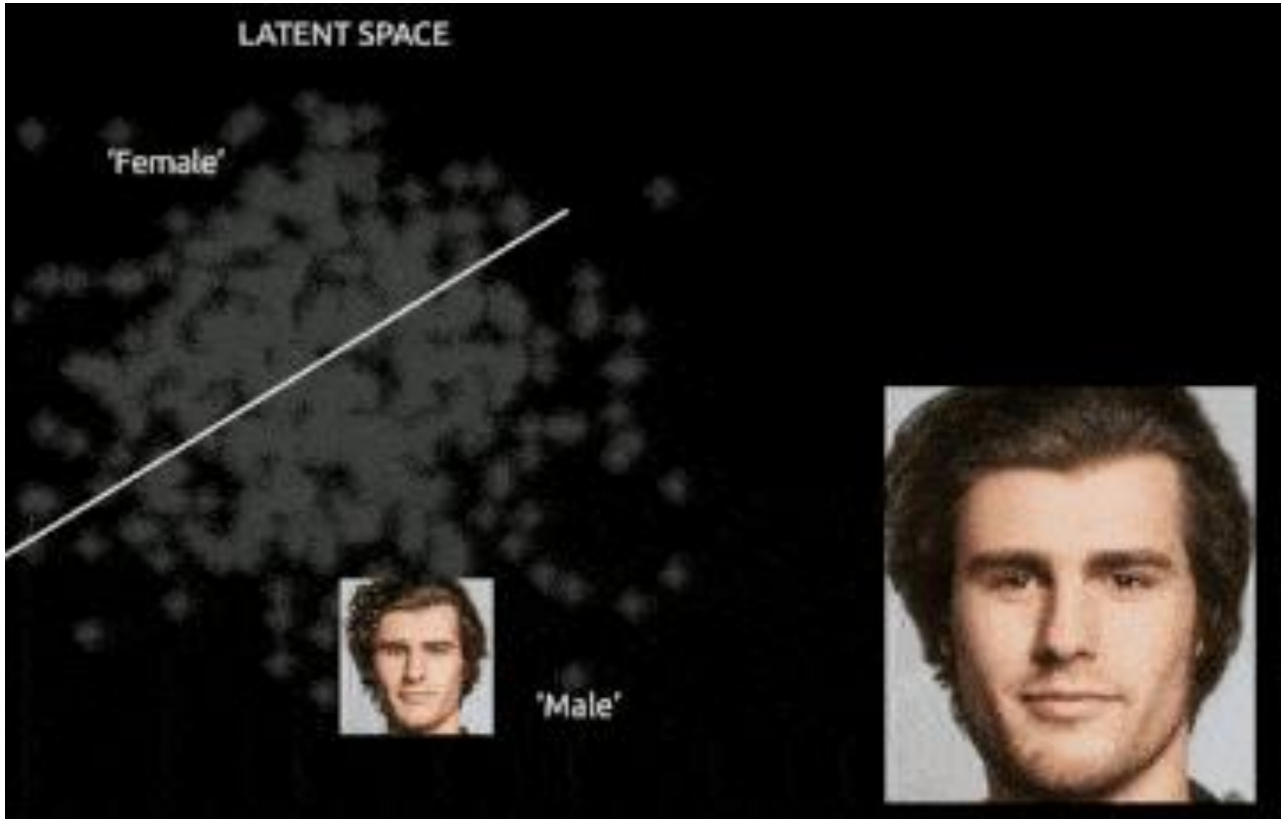
Latent Space



Latent Space



Latent Space



在變分自編碼器 (Variational Autoencoder, VAE) 中，更新的參數主要來自於編碼器 (Encoder) 和解碼器 (Decoder) 的神經網路權重。這些參數是通過反向傳播演算法，根據損失函數的梯度進行優化的。

1. 編碼器 (Encoder) 的參數

編碼器的目的是學習條件機率 $q(z|x)$，將輸入資料 x 映射到潛在變數空間 z。具體來說，編碼器輸出的是兩組參數：

- $\mu(x)$：潛在空間的均值向量。
- $\sigma^2(x)$：潛在空間的變異數或對數標準差。

這些參數由神經網路學習，因此編碼器包含以下需要更新的參數：

- **權重矩陣 (W_e)**：每層神經元的連接權重。
- **偏置項 (b_e)**：每層的偏置。

2. 解碼器 (Decoder) 的參數

解碼器的目的是學習條件機率 $p(x|z)$，將潛在變數 z 還原為與輸入資料 x 相似的重建樣本 $\hat{x}$。

解碼器的參數包括：

- **權重矩陣 (W_d)**：每層神經元的連接權重。
- **偏置項 (b_d)**：每層的偏置。

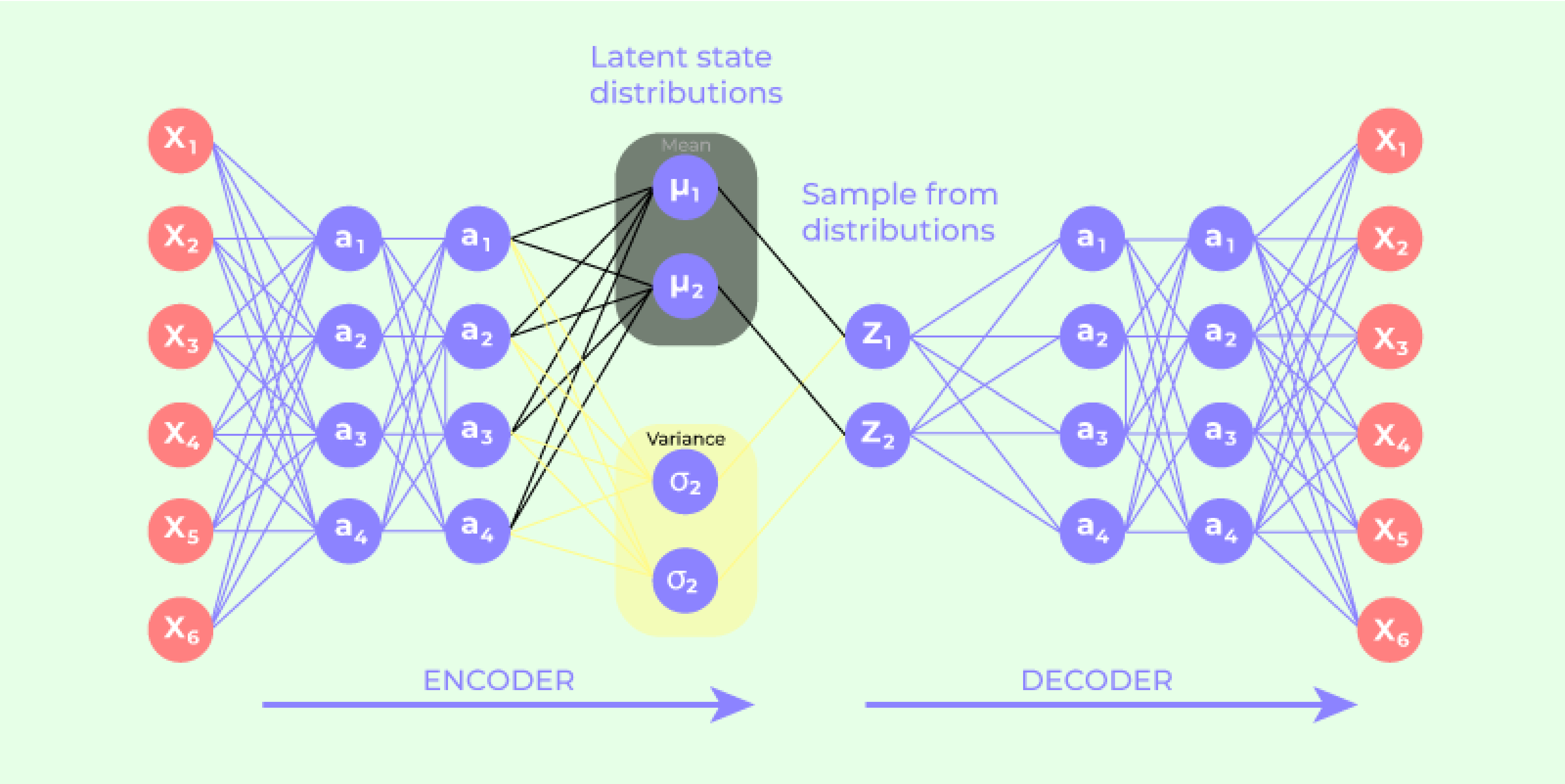
參數更新：使用梯度下降法或其變體（如 Adam 優化器）更新參數：

$$\theta \leftarrow \theta - \eta \cdot \nabla_\theta \mathcal{L}_{\text{VAE}}$$

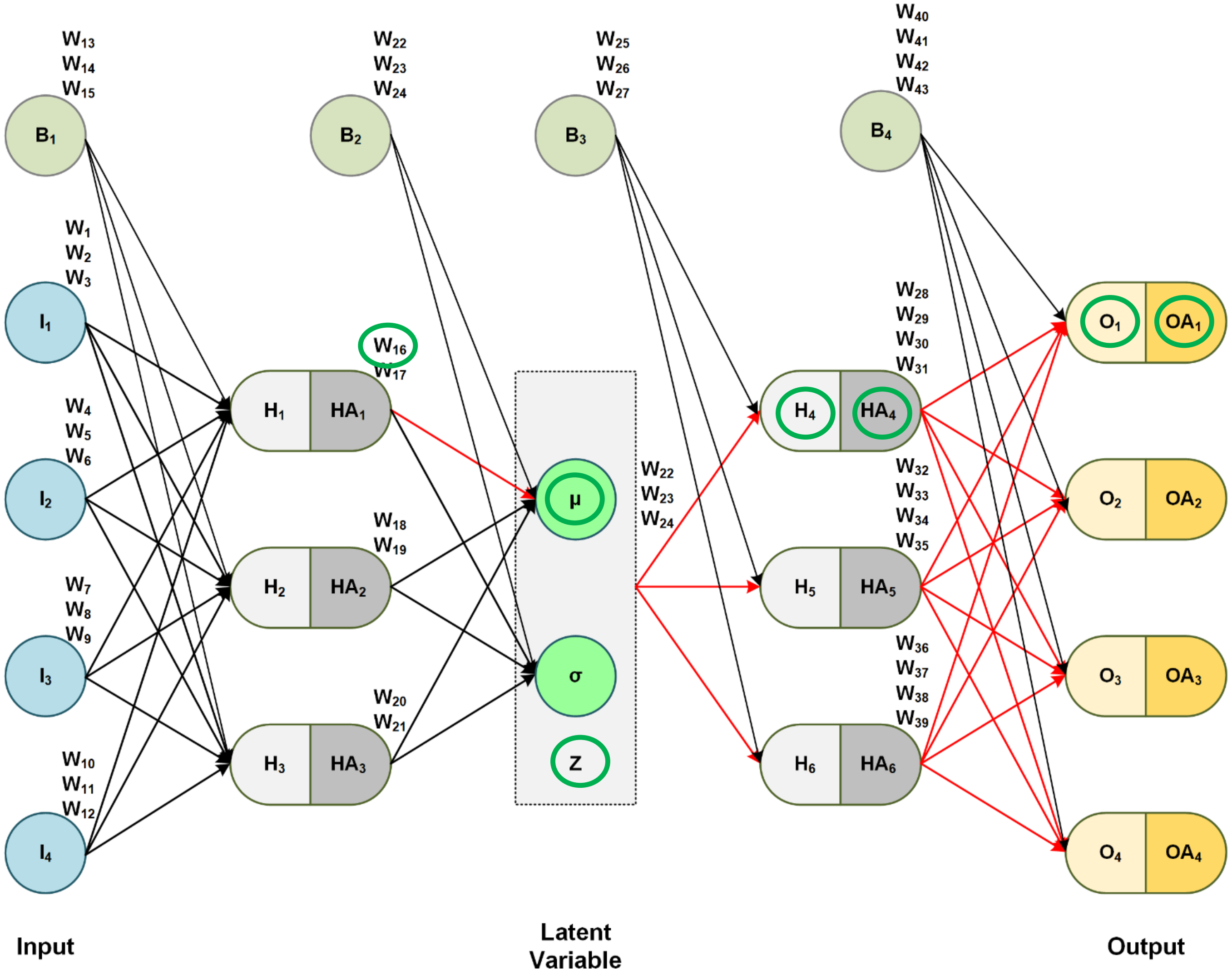
其中：

- θ：包括 W_e, b_e, W_d, b_d。
- η：學習率。

Variational Autoencoder



Variational Autoencoder



$$\frac{\partial E}{\partial w_{16}} = \frac{\partial E}{\partial OA_1}\frac{\partial OA_1}{\partial O_1}\frac{\partial O_1}{\partial HA_4}\frac{\partial HA_4}{\partial H_4}\frac{\partial H_4}{\partial Z}\frac{\partial Z}{\partial \mu}\frac{\partial \mu}{\partial w_{16}} + \frac{\partial E}{\partial OA_2}\cdots + \frac{\partial E}{\partial OA_3}\cdots + \frac{\partial E}{\partial OA_4}\cdots$$

...

Adversarial Autoencoder

對抗自編碼器（Adversarial Autoencoder, AAE） 是一種結合了自動編碼器（Autoencoder）和生成對抗網絡（GAN）的深度學習模型。它既繼承了自動編碼器用於數據壓縮和重建的特性，又利用 GAN 的對抗機制來對潛在空間進行正則化。這使得 AAE 在學習潛在分佈方面比傳統自動編碼器更加靈活和強大。

AAE 的結構和工作原理

AAE 由兩個主要部分組成：

1. **自動編碼器（Autoencoder）部分**：負責數據的編碼和解碼。
    - **編碼器（Encoder）**：將輸入數據 x 映射到潛在空間的潛在向量 $z = E(x)$。
    - **解碼器（Decoder）**：從潛在空間中的向量 z 重建輸入數據 $x' = D(z)$，以使重建誤差 $||x - x'||$ 最小化。

Adversarial Autoencoder

2. **生成對抗網絡（GAN）部分**：對潛在空間進行正則化，保證潛在向量符合特定的先驗分佈（如高斯分佈）。
   - **生成器（Generator）**：在 AAE 中，編碼器同時充當生成器，將輸入數據壓縮到潛在向量。
   - **判別器（Discriminator）**：用來區分來自先驗分佈（如正態分佈）的樣本和來自編碼器生成的潛在向量。

工作流程

1. **自動編碼器階段**：首先，編碼器將輸入數據 x 映射到潛在向量 z，然後解碼器根據潛在向量 z 嘗試重建輸入數據 x'。自動編碼器部分的目標是最小化重建誤差，使得輸入和重建結果越接近越好。這一過程類似於傳統的自動編碼器。

Adversarial Autoencoder

2. **對抗訓練階段**：接下來，對抗網絡開始工作。編碼器生成的潛在向量 z 與來自先驗分佈（通常為正態分佈）的樣本 z_{prior} 一起被送入判別器。判別器的目標是區分哪個樣本來自編碼器生成，哪個來自先驗分佈。編碼器的目標則是讓其生成的潛在向量儘可能接近先驗分佈，從而「騙過」判別器。

AAE 的損失函數

1. **重建誤差**：

$$\mathcal{L}_{\text{reconstruction}} = \mathbb{E}_{x \sim p_{\text{data}}(x)} ||x - D(E(x))||^2$$

這部分損失函數與普通自動編碼器一樣，目的是使重建的數據 x' 與輸入數據 x 盡可能相似。

Adversarial Autoencoder

2. **對抗損失**（來自 GAN 的對抗損失）：

潛在(先驗)：真　　編碼(後驗)：偽

$$\mathcal{L}_{\text{adversarial}} = \mathbb{E}_{z\sim p(z)}[\log D(z)] + \mathbb{E}_{x\sim p_{\text{data}}(x)}[\log(1 - D(E(x)))]$$

這部分損失來自生成對抗網絡，其中 $D(z)$ 是判別器的輸出，負責區分來自先驗分佈的樣本和來自編碼器的潛在向量。

編碼器的目標是最小化對抗損失，而判別器的目標則是最大化它。這樣，通過對抗訓練，編碼器會學習如何生成符合先驗分佈的潛在向量。

$$\min_G \max_D V(D, G) = \mathbb{E}_{x\sim p_{\text{data}}(x)}[\log D(x)] + \mathbb{E}_{z\sim p_z(z)}[\log(1 - D(G(z)))]$$

資料：真　　生成：偽

Adversarial Autoencoder

AAE 與 VAE 的區別

- VAE（變分自動編碼器）使用 KL 散度來將潛在空間約束為符合先驗分佈的形式（如正態分佈），這是通過一個解析表達式來實現的。
- AAE 則通過 GAN 的對抗訓練來正則化潛在空間，這使得 AAE 能夠更加靈活地適應不同的先驗分佈，並且不依賴於解析表達式。這意味著 AAE 能夠處理更加複雜的數據分佈。

AAE 的應用

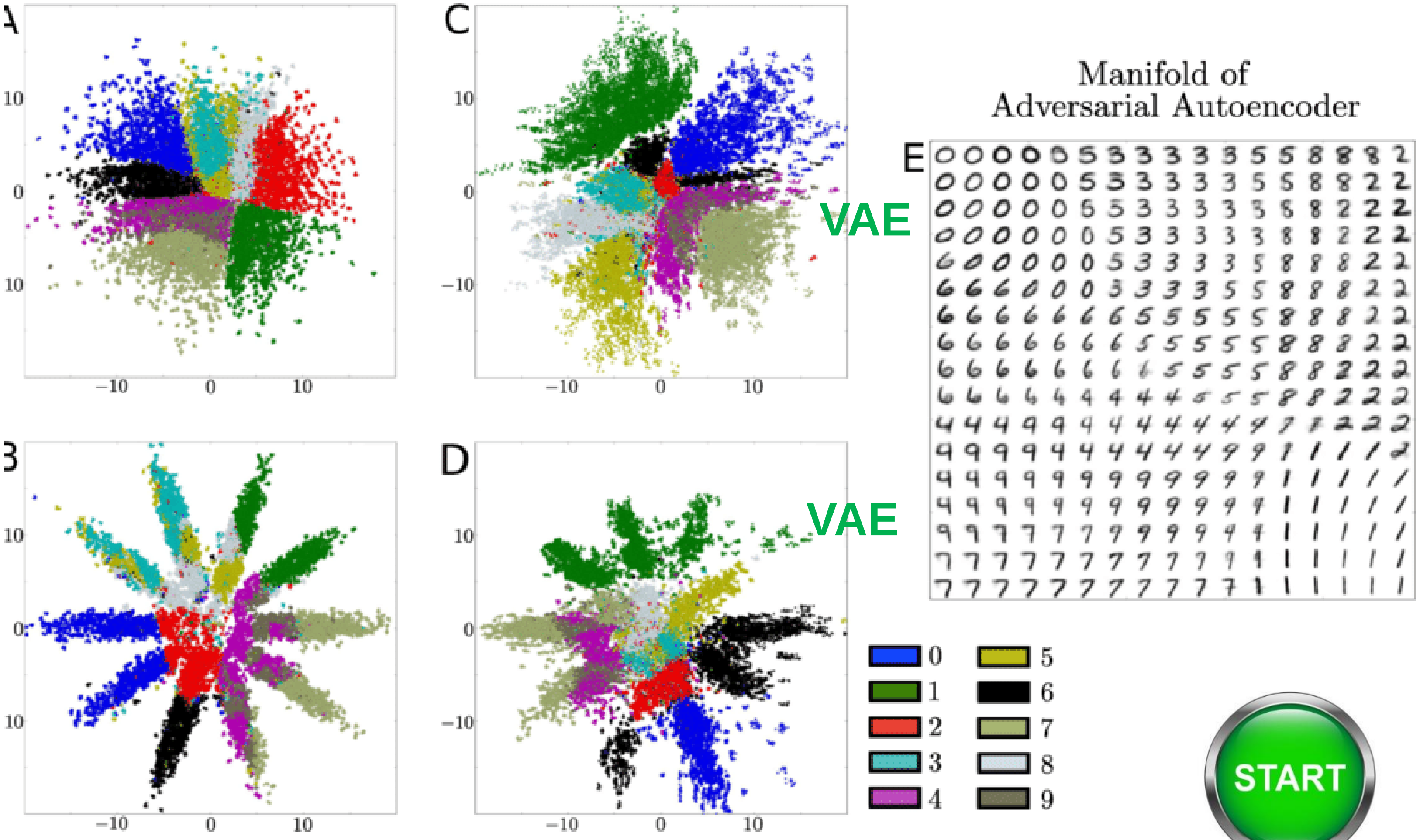
1. **生成圖像**：AAE 可以用來生成符合先驗分佈的圖像，並且通過對抗訓練，其生成的圖像質量可能比普通自動編碼器更好。
2. **半監督學習**：AAE 的潛在空間被正則化為具有良好結構，這有助於在半監督學習中進行分類任務，即即使只有少量標籤數據，模型也能有效學習。
3. **資料生成與降維**：AAE 可以用來生成數據樣本，或者用於資料降維和可視化，因為其潛在空間被正則化為易於分析的分佈。

Latent Space

A 2D Gaussian

AAE

VAE



VAE

AAE

10 2D Gaussian

START

Diffusion Model