□□□□□

□□□□□□□□□□□□□□□

□□□□□

□□□□□□ □□□□N□□□□□□□□ $$ P(Y=c_k) = \frac{1}{K} \ PK - 1 = 0 $$ □□□□□□□□□□□□□ $$ P(Y=c_k) = \frac{\sum_{i=1}^{N}{I(y_i = c_k)}}{N} $$ □□ $$ P(Y=c_k)N - \sum_{i=1}^{N}{I(y_i = c_k)} = 0 $$ □ $$ \lambda(P(Y=c_k)K - 1) + P(Y=c_k)N - \sum_{i=1}^{N}{I(y_i = c_k)} = 0 $$ □□ $$ P(Y=c_k) = \frac{ \sum_{i=1}^{N}{I(y_i = c_k)} + \lambda}{K\lambda + N} $$ □□□□□□

- S_j□□j□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□$\frac{1}{S_j}$ $$ P(X^{(j)} = a_{jl} | Y =c_k) = \frac{1}{S_j} \ P(X^{(j)} = a_{jl} | Y =c_k) = \frac{\sum_{i=1}^{N}{I(X^{(j)} = a_{jl} , Y =c_k)}}{\sum_{i=1}^{N}{I(Y=c_k)}} $$ □□ $$ \lambda{PS_j - 1} + P\sum_{i=1}^{N}{I(Y=c_k)} - \sum_{i=1}^{N}{I(X^{(j)} = a_{jl} , Y =c_k)} = 0 $$ □□ $$ P(X^{(j)} = a_{jl} | Y =c_k)= \frac{\lambda +\sum_{i=1}^{N}{I(X^{(j)} = a_{jl} , Y =c_k)}}{\sum_{i=1}^{N}{I(Y=c_k)} + S_j\cdot \lambda} $$

□□□□□□□□□□

□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□

□□□□□□□□□□□□□ $$ \begin{aligned} R_{exp} &= \int{\int{L(y,f(\vec{x}))P(\vec{x},y)d\vec{x}}dy}\ &= \int_x{\int_y{L(y,f(\vec{x}))P(y|\vec{x})dy}P(\vec{x})d\vec{x}}\ &= \mathbb{E}_x{[\int_y{L(y,f(\vec{x}))P(y|\vec{x})dy}]} \ &= \mathbb{E}x{[\sum{k=1}^{K}{L(c_k,f(\vec{x}))P(c_k|\vec{x})}]}

\end{aligned} $$ □□□□□□□x□□□□$L(y,f(\vec{x}))P(y|\vec{x})$□□□□□ $$ \begin{aligned} f(x) &= \argmin_{y\in Y}{\sum_{k=1}^KL(c_k,y)P(c_k|X=x)}\ &= \argmin_{y\in Y}{\sum_{k=1}^KP(y\not = {c_k}|X= x)} \ &= \argmin_{y\in Y}({1 - P(y={c_k}|X= x)}) \quad \text{□□□□□□1□□□□1□□□□□□□□□}\ &= \argmax_{y\in Y}{P(y ={c_k}|X= x)} \ \end{aligned} $$

□□□□□

□□□□□□□□□

□□□□□□□□□□□□□□□ξ_i

□□□□□ $$ L = \frac{1}{2}||w||^2 + C\sum_{i = 1}^N{\xi_i} $$ □□□□ $$ \begin{aligned} y_i(w^Tx_i + b) &\ge 1 - \xi_i \ \xi_i \ge 0 \end{aligned} $$

- $y_i(w^Tx_i + b) \ge 1$□□□□□□□□□□□□□□
  - □□$1 - y_i(w^Tx_i + b)\le 0$□□□□□□□$\xi_i \ge 1 - y_i(w^Tx_i + b)$□$\xi_i$□□□□□□□□□□□□0
  - □□□□□□□□□□□□$\xi_i = 0$
- $y_i(w^Tx_i + b) < 1$□□□□□□□□□□□□□
  - □□ $1 - y_i(w^Tx_i + b)> 0$□□□□□□□□□□□$\xi_i \ge 1 - y_i(w^Tx_i + b)$□$\xi_i$□□□□□□□□□□□0
  - □□□□□□□□□□□□$\xi_i = 1 - y_i(w^Tx_i + b)$

□□□□ $$ \xi_i = \max{(0,1 - y_i(w^Tx_i + b))} $$

□□□□□□□$\lambda = \frac{1}{2C}$□□□□□□□ $$ \sum_{i=1}^N[1 - y_i(w^Tx_i + b)]+ + \lambda||w||^2 $$ □□ $$ [z]+ = \begin{cases} z , \quad z \ge 0 \ 0, \quad z < 0 \end{cases} $$

EM□□

EM□□□□□

□□□□□VAE□□□□□□□□□□Diffusion Model□□□
□□□□□□□VAE□Diffusion Model□□□

□□□□□EM□□□□□□**EM**□□□□□□□□□□□□**Z**□□
□**θ**

□□□□□□**KL**□□ $$ KL(q||p) = \sum_{Z}q(Z)\log{\frac{q(Z)}{p(Z)}} $$ □□□□q□□□p□□□□□□□□□□□□□□□□□□□0□

□□□□□□EM□□□□□□□□□□□□□□**Q**□□ $$ \begin{aligned} Q(\theta,\theta^{(i)})&=\mathbb{E}{Z}(\log{P(Y,Z|\theta)}|Y,\theta^{(i)}) \ &= \sum{Z}\log{P(Y,Z|\theta)}P(Z|Y,\theta^{(i)}) \end{aligned} $$

□□□□□□□□□□□□□□Q□□

□□EM□□□□□□□□□□□□□□□Y□□□□θ□□□□□□□ $$ \begin{aligned} L(\theta) = \log{P(Y|\theta)} &= \log{\sum_{Z}{P(Y,Z|\theta)}} \ &=\log{\sum_{Z}{q(Z)\frac{P(Y,Z|\theta)}{q(Z)}}} \end{aligned} $$

□□**Jeason**□□□ $$ \begin{aligned} \log{\sum_{Z}{q(Z)\frac{P(Y,Z|\theta)}{q(Z)}}} &\ge \sum_{Z}q(Z)\log{\frac{P(Y,Z|\theta)}{q(Z)}} \ &= \sum_{Z}q(Z)\log{P(Y,Z|\theta)} - \sum_{Z}q(Z)\log{q(Z)} \end{aligned} $$

□□□□ $$ \begin{aligned} \mathcal{{L}}(q,\theta) = \sum_{Z}q(Z)\log{P(Y,Z|\theta)} - \sum_{Z}q(Z)\log{q(Z)} \end{aligned} $$

□□ $$ \log{P(Y,Z|\theta)} = \log{P(Z|Y,\theta)P(Y|\theta)} $$

□□□ $$ \begin{aligned} \mathcal{{L}}(q,\theta) &= \sum_{Z}q(Z)\log{P(Y,Z|\theta)} - \sum_{Z}q(Z)\log{q(Z)} \ &= \sum_{Z}q(Z)\log{P(Z|Y,\theta)P(Y|\theta)} - \sum_{Z}q(Z)\log{q(Z)} \ &= \sum_{Z}q(Z)(\log{P(Z|Y,\theta)}+\log{P(Y|\theta)}) - \log{q(Z)})\ &= \log{P(Y|\theta)}\sum_{Z}q(Z) + \sum_{Z}q(Z)\log{\frac{P(Z|Y,\theta)}{q(Z)}} \ &= \log{P(Y|\theta)} - KL(q(Z)||P(Z|Y,\theta)) \end{aligned} $$

□□□□ $$ \begin{aligned} \log{P(Y|\theta)} &= KL(q(Z)||P(Z|Y,\theta)) + \mathcal{{L}}(q,\theta)\ \end{aligned} $$ □ $$ \log{P(Y|\theta)} \ge \mathcal{{L}}(q,\theta) $$

□□□□□□□□□□□□$q(Z)$□□□□□□$P(Z|Y,\theta^{i})$□□□□□□□□

> □***E***□□□□□□□□□□□□□□□□□□□□□$q(Z)$□□□□□□$\mathcal{{L}}(q,\theta)$□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□ $$ q(Z) = P(Z|Y,\theta^{i}) $$ □□□□□□□□ □□□□□□□□□□□$\theta^{(i)}$□□□□□□□□□□□□□□□□ $$ \log{P(Y,Z|\theta^{(i)})} = \mathcal{{L}}(q,\theta^{(i)}) $$ □□□KL□□□0□
>
> □□□
>
> □□□□□□□□KL□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□
>
> □□□□□□$q(Z) = P(Z|Y,\theta^{(i)})$□□□□□□$E$□□□□□□$\theta^{(i)}$□□KL□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□
> □□□□□M□□□□□θ□□□□□□□□□□□$P(Z|Y,\theta)$□□□□□□KL □□□□□□□□□□□□□□□□□□□□□□□□□□□

□□ $$ \begin{aligned} \log{P(Y|\theta)} &\ge \mathcal{{L}}(q,\theta)\ &= \sum_{Z}P(Z|Y,\theta^{(i)})\log{P(Y,Z|\theta)} + const \ &= \underbrace{\mathbb{E}{Z}

(\log{P(Y,Z|\theta)}|Y,\theta^{(i)})}{\text{Q□□}} + \underbrace{\sum_{Z}{P(Z|Y,\theta^{(i)})\log{P(Z|Y,\theta^{(i)})}}}_{\text{□□}} \end{aligned} $$

□□□□□□

□E□□□□□□ $$ q(Z) = P(Z|Y,\theta) $$ □□M□□□□□□□□□□□ $$ \mathcal{{L}}(q,\theta^{(i + 1)}) \ge \mathcal{{L}}(q,\theta^{(i)}) $$ □□□□□□□□□ $$ \log{P(Y,Z|\theta^{i})} = \mathcal{{L}}(q,\theta^{(i)}) $$ □□□□□M□□KL□□□□0□□□ $$ \log{P(Y,Z|\theta^{i + 1})} \ge \mathcal{{L}}(q,\theta^{(i + 1)}) $$ □□□ $$ \log{P(Y,Z|\theta^{i + 1})} \ge \mathcal{{L}}(q,\theta^{(i + 1)})\ge\mathcal{{L}}(q,\theta^{(i)})=\log{P(Y,Z|\theta^{i})} $$ □□□**EM**□□□□□□

□□□□□□

□□□□ $$ \phi(y|\theta_{k}) = \frac{1}{\sqrt{2\pi}\sigma_{k}}exp(-\frac{(x-\mu_k)^2}{2\sigma_{k}^2}) $$

□□□□□□□□□□□□□□□□□□□□□□ $$ \begin{aligned} &P(y|\theta) = \sum_{k=1}^K \alpha_k \phi(y|\theta_k) \ &\text{s.t.} \quad \alpha_k \geq 0 \ (k=1,\dots,K), \quad \sum_{k=1}^K \alpha_k = 1 \end{aligned} $$

Q□□□□ $$ Q(\theta,\theta^{(i)}) = \mathbb{E}_{\gamma}[\log{P(y,\gamma|\theta)}|y,\theta^{(i)}] $$

□□□□□□□□□□□□□□□□□□□□□□ $$ \theta = (\alpha_{k},\mu_{k},\sigma_{k}^2) $$

□□□□□□ $$ \gamma_{jk} = \begin{cases} 1 &\text{□j□□□□□□□□□k□□□□□} \ 0 \end{cases} $$

□□□□□□□□□□□□□□□□□□□□□□□□□□□Q□□ $$ \begin{aligned} P(y,\gamma|\theta) &= \prod_{j = 1}^{N}{p(y_j,\gamma_{j1},\gamma_{j2},\dots,\gamma_{jk}|\theta)} \ &=\prod_{j = 1}^{N}{\prod_{k=1}^{K}{[\alpha_{k}\phi(y_j|\theta_k)]^{\gamma_{jk}}}} \end{aligned} $$

□ $$ n_k = \sum_{j=1}^{N}\gamma_{jk} $$

□ $$ \begin{aligned} \prod_{j = 1}^{N}{\prod_{k=1}^{K}{[\alpha_{k}\phi(y_i|\theta_k)]^{\gamma_{jk}}}} &= \prod_{k = 1}^K{\alpha_k^{n_k}}\prod_{j=1}^N[\phi(y_j|\theta_k)]^{\gamma_{jk}} \end{aligned} $$

□□□Q□□□ $$ \begin{aligned} Q(\theta,\theta^{(i)}) &= \mathbb{E}{\gamma}{\sum{k=1}^N{n_k\log{\alpha_{k} + \sum_{j=1}^N{\gamma_{jk}[\log{\phi(y_j|\theta_{k})}]}}}|y,\theta^{(i)}} \ &= \mathbb{E}{\gamma}{\sum{k=1}^N{n_k\log{\alpha_{k} + \sum_{j=1}^N{\gamma_{jk}[(\log{\frac{1}{\sqrt{2\pi}}}-\log{\sigma_{k}} - \frac{(y_j - \mu_k)^2}{2\sigma_k^2})]}}}|y,\theta^{(i)}} \

\end{aligned} $$

□□□□□□□□□□□□□□□□□□□□□$\mathbb{E}(\gamma_{jk}|y,\theta^{(i)})$□□□□□□□□□□□□□□□□□□□□□□□□□□□□ $$ \begin{aligned} \mathbb{E}(\gamma_{jk}|y,\theta^{(i)}) &= P(\gamma_{jk} = 1|y,\theta^{(i)}) \ &= \frac{P(\gamma_{jk} = 1|\theta^{(i)})P(y_j|\gamma_{jk} = 1,\theta^{(i)})}{P(y|\theta^{(i)})}\ &= \frac{\alpha_{k}\phi(y_j|\theta_k)}{\sum_{k=1}^K{\alpha_{k}\phi(y_j|\theta_k)}} \ &= \hat{\gamma}_{jk} \end{aligned} $$

□□□□□ $$ \begin{aligned} n_k &= \sum_{j=1}^{N}{\gamma_{jk}} \ \mathbb{E}[n_k|y,\theta^{(i)}] &= \sum_{j=1}^N{\mathbb{E}[\gamma_{jk}|y,\theta^{(i)}]} = \sum_{j=1}^{N}{\hat{\gamma}_{jk}} \end{aligned} $$

□□□□**Q**□□□□□□□ $$ \begin{aligned} Q(\theta,\theta^{(i)}) &= \sum_{k = 1}^{N}{\sum_{j=1}^{N}{(\mathbb{E}{\gamma_{jk}})}\log{\alpha_k}+ \sum_{j=1}^N{(\mathbb{E}{\gamma_{jk}})}[(\log{\frac{1}{\sqrt{2\pi}}}- \log{\sigma_{k}} - \frac{(y_j - \mu_k)^2}{2\sigma_k^2})]|y,\theta^{(i)}}\ &= \sum_{k = 1}^{N}{\sum_{j=1}^{N}{\hat{\gamma}*{jk}}\log{\alpha_k}+ \sum*{j=1}^N{\hat{\gamma}*{jk}}[(\log{\frac{1}{\sqrt{2\pi}}}- \log{\sigma*{k}} - \frac{(y_j - \mu_k)^2}{2\sigma_k^2})]|y,\theta^{(i)}} \end{aligned} $$

> □□□□□□□□□□□□□□□
>
> - *$\mathbb{E}(\gamma_{jk}|y,\theta^{(i)})$□□□□□□$P(\gamma_{jk} = 1|y,\theta^{(i)})$□□□□□□□□□□□□□□□1□□0□□□□□□□□□□*
> - *$\hat{\gamma}_{jk}$□□□□□□□y_j□□□□□k□□□□*
> - *$P(y|\theta^{(i)})$□□□□□□□□□□□□□□$\theta^{(i)}$□□□□□□□□□□□□y□□□□□□□□□□□□□□□□□□□□□□□*
> - *$P(\gamma_{jk} = 1|\theta^{(i)})$□□□□□□$\theta^{(i)}$□□□□□□□□□□□k□□□□□□□□□□□□□□α_k*

□□□□□□□α,μ,σ□□□□ $$ \begin{aligned} \hat{\alpha}*k &= \frac{n_k}{N} = \frac{\sum*{j=1}^{N}{\hat{\gamma}*{jk}}}{N}\ \hat{\mu}*k &= \frac{\sum{j = 1}^N{\hat{\gamma}*{jk}y_j}}{\sum_{j = 1}^N{\hat{\gamma}*{jk}}}\ \hat{\sigma}*k &= \frac{\sum{j = 1}^N{\hat{\gamma}*{jk}(y_j-\mu_k)^2}}{\sum_{j = 1}^N{\hat{\gamma}_{jk}}} \end{aligned} $$