变分推断 Background

数学角度 → 优化问题

回归 $\underbrace{f(w)=w^Tx}_{\text{model}}$ $\underbrace{\text{loss func}}_{\text{strategy}}$ $L(w)=\sum_{i=1}^{N}\|w^Tx_i-y_i\|^2$

$\hat{w}=\arg\min_{w} L(w)$

strategy 对应的 algorithm:

① 解析解 $\frac{\partial L(w)}{\partial w}=0 \Rightarrow w^*=(X^TX)^{-1}X^TY$

② 数值解: gradient descent, SGD, …

SVM: $f(w)=\text{sign}(w^Tx+b)$

Loss func $\min \frac{1}{2}w^Tw$ s.t. $y_i(w^Tx_i+b)\geq 1$, $i=1,2,\cdots,N$

有约束, 凸优化

Lagrange Algorithm

EM $\hat{\theta}=\arg\max \log P(X|\theta)$

$\theta^{(t+1)}=\arg\max_{\theta}\int \log P(X,Z|\theta)\,P(Z|X,\theta^{(t)})\,dZ$

贝叶斯角度 → 积分问题

$$P(\theta|X)=\frac{P(X|\theta)\cdot P(\theta)}{P(X)}$$

后验 = $\frac{\text{似然}\cdot\text{先验}}{\text{标准化}}$

其中 $P(X)=\int_\theta P(X|\theta)P(\theta)d\theta$

贝叶斯 inference: 求 $P(\theta|X)$

贝叶斯 decision: $X\to N$个样本

$\tilde{x}$ 新样本, 求 $P(\tilde{x}|X)$

$$P(\tilde{x}|X)=\int_\theta P(\tilde{x},\theta|X)d\theta$$
$$=\int_\theta P(\tilde{x}|\theta)P(\theta|X)d\theta$$
$$=E_{\theta|X}[P(\tilde{x}|\theta)]$$

公式推导

X: [redacted] observed data

Z: latent variable + parameter

(对于 EM, 隐变量是 Z, 参数是 θ)

(X,Z): complete data

$\log P(X)=\log P(X,Z)-\log P(Z|X)$ 对 q 求期望

$=\log\frac{P(X,Z)}{q(Z)}-\log\frac{P(Z|X)}{q(Z)}$ $\Rightarrow$ 左 $\int_Z \log P(X)\,q(Z)dZ=\log P(X)$

右 $\int_Z q(Z)\left(\log\frac{P(X,Z)}{q(Z)}-\log\frac{P(Z|X)}{q(Z)}\right)dZ$

DATE

$$左式 = \underbrace{\int_Z q(Z)\log\frac{P(X,Z)}{q(Z)}dZ}_{\text{ELBO (evidence lower bound)}} \underbrace{\int_Z q(Z)\log\frac{P(Z|X)}{q(Z)}dZ}_{KL(q||p)}$$

$$= \mathcal{L}(q) \quad + \quad KL(q||p) \geq 0$$

变分 加号,注意KL的定义

$$\hat{q}(Z) = \underset{q(Z)}{\arg\max}\ \mathcal{L}(q) \Longrightarrow \hat{q}(Z) \approx P(Z|X)$$

假设 $q(Z) = \prod_{i=1}^{M} q_i(Z_i)$ 统计学习中的 mean theory

把 $q(Z)$ 分成M份, M个份之间相互独立

$$\mathcal{L}(q) = \underbrace{\int_Z q(Z)\log P(X,Z)dZ}_{①} - \underbrace{\int_Z q(Z)\log q(Z)dZ}_{②}$$

$$① = \int_Z \prod_{i=1}^{M} q_i(Z_i)\log P(X,Z)\,dZ_1 dZ_2 dZ_3 \cdots dZ_M$$

$$= \int_Z q_j(Z_j)\Big(\underbrace{\prod_{i\neq j}^{M} q_i(Z_i)\log P(X,Z)\,dZ_1\cdots dZ_{j-1}\,dZ_{j+1}\cdots dZ_M}_{\log P(X,Z)\prod_{i\neq j}^{M} q_i(Z_i)\,dZ_i}\Big)dZ_j$$

$$= \int_{Z_j} q_j(Z_j)\underbrace{\Big(\int_{Z_i\neq Z_j}\log P(X,Z)\prod_{i\neq j}^{M} q_i(Z_i)\,dZ_i\Big)}\,dZ_j$$

$$= \int_{Z_j} q_j(Z_j)\quad E_{\prod_{i\neq j}^{M} q_i(Z_i)}[\log P(X,Z)]\quad dZ_j$$

$$② \int_Z q(Z)\log q(Z)dZ$$

$$= \int_Z \prod_{i=1}^{M} q_i(Z_i)\log\prod_{i=1}^{M} q_i(Z_i)dZ = \int_Z \Big(\prod q_i(Z_i)\Big)\Big(\sum \log q_i(Z_i)\Big)dZ$$

$$= \int_Z \prod_{i=1}^{M} q_i(Z_i)\,[\log q_1(Z_1) + \log q_2(Z_2) + \cdots + \log q_M(Z_M)]\,dZ$$

对于分配律其中一项 $\int_Z \prod_{i=1}^{M} q_i(Z_i) \log q_1(Z_1)\, dZ$

$$= \int_{Z_1, Z_2 \cdots Z_M} q_1 q_2 \cdots q_M \log q_1 \, dZ_1 dZ_2 \cdots dZ_M$$

$$= \int_{Z_1} q_1 \log q_1 \, dZ_1$$

因此 ② $= \int_Z \prod_{i=1}^{M} q_i(Z_i) \sum_{i=1}^{M} \log q_i(Z_i)\, dZ$

$$= \sum_{i=1}^{M} \int_{Z_i} q_i(Z_i) \log q_i(Z_i)\, dZ_i$$

$$= \int_{Z_j} q_j(Z_j) \log q_j(Z_j)\, dZ_j + C$$

$$\mathcal{L}(q) = \int_{Z_j} q_j(Z_j)\, E_{\prod_{i \neq j}^{M} q_i(Z_i)} [\log P(X, Z)]\, dZ_j - \int_{Z_j} q_j(Z_j) \log q_j + C$$

$$= \int_{Z_j} q_j(Z_j) \log \frac{\hat{P}(X, Z_j)}{q_j(Z_j)}\, dZ_j$$

$$= -KL\left(q_j \,\|\, \hat{P}(X, Z_j)\right) \leq 0$$

结论：

$$q_j(Z_j) \approx \hat{P}(X, Z_j)$$