

机器学习复习8

2022年7月7日 星期四

21:55

The EM Algorithm

1. Jensen 不等式

若 hessian 矩阵 H 为半正定, ($H \geq 0$),

则 f 为凸函数, 若 $f''(x) > 0$ 或 $H > 0$

则称 f 为严格凸函数

有: $E[f(x)] \geq f(EX)$

对于凹函数, 有: $E[f(x)] \leq f(EX)$

2. EM Algorithm

sample: $\{x^{(1)}, \dots, x^{(m)}\}$

find $z, P(x, z) \max$.

$$\ell(\theta) = \sum_{i=1}^m \log \sum_z P(x, z; \theta)$$

$$\circ \sum_i \log P(x^{(i)}; \theta) = \sum_i \log \sum_{z^{(i)}} P(x^{(i)}, z^{(i)}; \theta)$$

$$= \sum_i \log \sum_{z^{(i)}} Q_i(z^{(i)}) \frac{P(x^{(i)}, z^{(i)}; \theta)}{Q_i(z^{(i)})}$$

$$\geq \sum_i \sum_{z^{(i)}} Q_i(z^{(i)}) \log \frac{P(x^{(i)}, z^{(i)}; \theta)}{Q_i(z^{(i)})}$$

(利用了 Jensen 不等式)

令随机变量为常数 θ .

$$\frac{P(x^{(i)}, z^{(i)}; \theta)}{Q_i(z^{(i)})} = c$$

且有 $Q_i(z^{(i)}) = P(z^{(i)} | x^{(i)}; \theta)$

故 EM 算法步骤如下:

{

$$Q_i(z^{(i)}) := P(z^{(i)} | x^{(i)}; \theta)$$

$$\theta := \arg \max_{\theta} \sum_i \sum_{z^{(i)}} Q_i(z^{(i)}) \log \frac{P(x^{(i)}, z^{(i)}; \theta)}{Q_i(z^{(i)})}$$

}

$$\ell(\theta^{(t+1)}) = \sum_i \sum_{z^{(i)}} Q_i^{(t+1)}(z^{(i)}) \log \frac{P(x^{(i)}, z^{(i)}; \theta^{(t+1)})}{Q_i^{(t+1)}(z^{(i)})}$$

且有 $\ell(\theta^{(t+1)}) \geq \ell(\theta^{(t)})$

• Review of Gauss model

$$\mu_1 = \frac{\sum_{i=1}^m w_1^{(i)} x^{(i)}}{\sum_{i=1}^m w_1^{(i)}}$$

$$\text{同理可推导: } \phi_j = \frac{1}{m} \sum_{i=1}^m w_j^{(i)}$$