$$X=\{x_1, x_2 \cdots x_n\}$$

$$\hat{\theta} = \arg\max_\theta p(X|\theta) = p(x_1, x_2, x_3 \cdots | \theta)$$

$$= \prod_{i=1}^{N} p(x_i|\theta)$$

极大似然估计

$$l(\theta) = \log\left(\prod_{i=1}^{N} p(x_i|\theta)\right) = \sum_{i=1}^{N} \log p(x_i|\theta)$$

$$\nabla_\theta l(\theta) = \sum_{i=1}^{N} \nabla_\theta \log p(x_i|\theta) = 0 \quad \text{求解极值点}$$

单高斯分布 $= N \sim (\mu, \sigma^2)$

$$\mu = \frac{1}{N}\sum_{i=1}^{N} x_i \quad , \quad \sigma^2 = \frac{1}{N}\sum_{i=1}^{N}(x_i-\mu)^2$$

多模态拟合：随机数 $j \in [1, M]$

(随机分布) $\alpha_j = p(y=j|\theta) \left(\sum_{j=1}^{M} \alpha_j = 1\right)$

(概率密度) $p_j(x|\theta_j) = p(x|y=j, \theta)$

(似然估计) $\Theta = \theta_1 \cup \theta_2 \cup \cdots \cup \theta_m$

$$p(x|\Theta) = \sum_{j=1}^{m} p(x, \theta_j)$$

$$= \sum_{j=1}^{m} p(x|\theta_j) p(\theta_j)$$

$$= \sum_{j=1}^{M} \alpha_j p(x|\theta_j)$$

GMM

极大似然: $l(\Theta) = \sum_{i=1}^{N} \log\left(\sum_{j=1}^{M} \alpha_j p(x|\theta_j)\right)$

单类别后验概率: (天池多数据)

$$p(y_i|x_i, \theta) = \frac{p(x_i|y_i, \theta) p(y_i|\theta)}{p(x_i|\theta)}$$

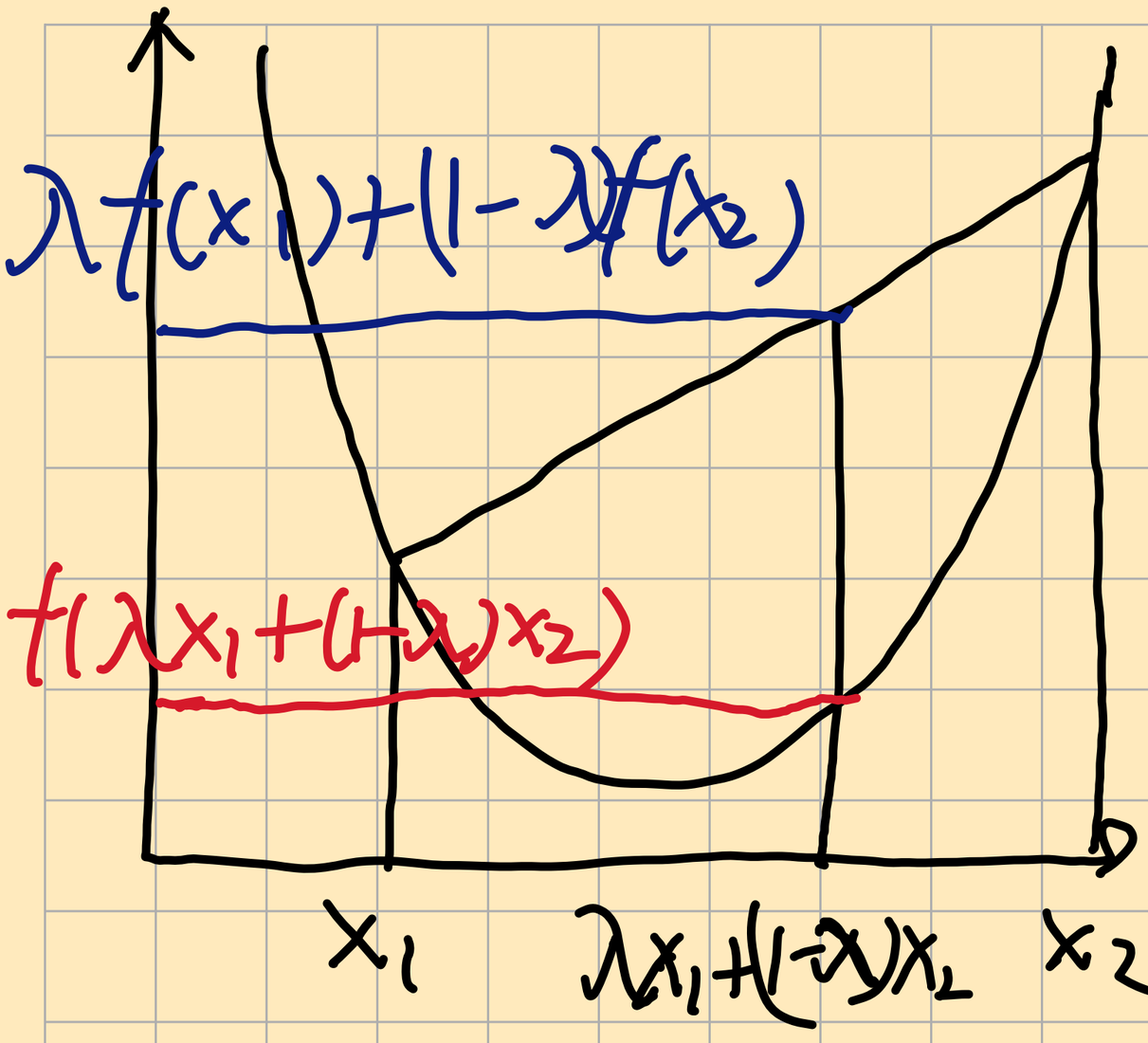
GMM极大值定理数f(x)

$\forall x \rightarrow f(x) \geq A(x, x^t)$, $A(x^t, x^t) = f(x^t)$

x^t: t时刻最优参数值

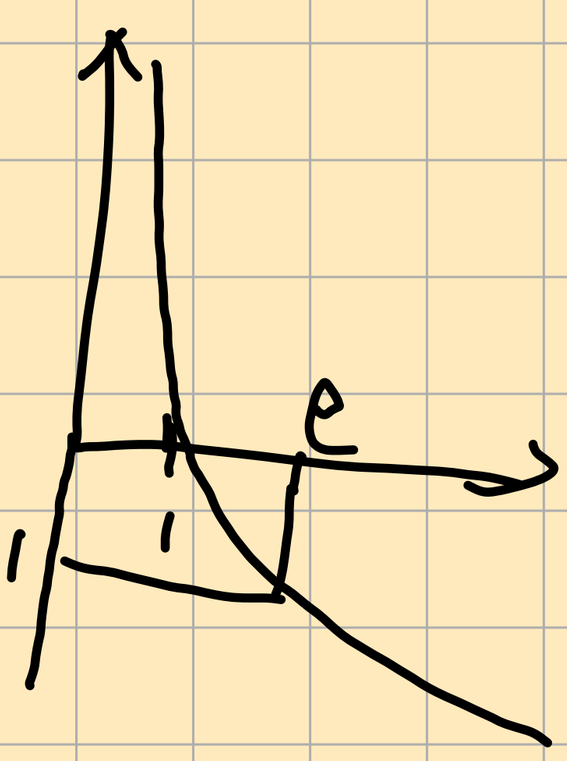
$h(x) = f(x) - A(x, x^t) \geq 0$

Jensen不等式 (马休农)



$$f\left(\sum_{i=1}^{n}\lambda_i x_i\right) \le \sum_{i=1}^{n}\lambda_i f(x_i)$$

$\because -\log(x)$ 是凸函数



$$\therefore \log\sum_{i=1}^{n}\lambda_i x_i \ge \sum_{i=1}^{n}\lambda_i \log(x_i)$$

$\log p(x|\theta)$

$$= \log\sum_{j=1}^{M} p(x, y_j|\theta)$$

新分布

$$= \log\sum_{j=1}^{M} q(y_j)\frac{p(x, y_j|\theta)}{q(y_j)}$$

$$\geq \sum_{i=1}^{n} q(y_i) \log\left[\frac{p(x, y_i|\theta)}{q(y_i)}\right]$$

故：$h(x)$

$$= \log p(x|\theta) - \sum_{j=1}^{n} q(y_j) \log\left[\frac{p(xy_j|\theta)}{q(y_j)}\right]$$

$$= \sum q(y)\left[\log p(x|\theta) - \log \frac{p(x, y_j|\theta)}{q(y_j)}\right]$$

$$= \sum q(y) \log \frac{p(x|\theta)\, q(y_j)}{p(x, y_j|\theta)}$$

$$= \sum q(y) \log \frac{q(y_j)}{p(y_j|x, \theta)}$$

$$= KL\left(q(y) \,||\, p(y_j|x, \theta)\right) \; ??$$

$$\alpha_j = = = q(y_j)$$

$$= = = p(y_j \mid x, \theta^t)$$

$$L(\theta) = \sum_{i=1}^{N} \log p(x_i \mid \theta)$$

$$\geq \sum_{i}^{N} \sum_{j}^{M} p(y_j \mid x_i, \theta^t) \log \left(\frac{p(x_i, y_j \mid \theta)}{p(y_j \mid x_i, \theta^t)} \right.$$

$$= A(\theta, \theta^t)$$

求其极值点

GMM：

① initialize θ^t

E ② calculate $p(y_j \mid x_i, \theta^t)$后验概率

$A(\theta, \theta^t)$

M ③ $\theta = \arg\max A(\theta, \theta^t)$

④ inspect $\|\theta^t - \theta\|$

结束 计收敛 else 更新

$KL(p(x) \| Q(x))$

$= \sum_{x \in X} \left[p(x) \log \frac{p(x)}{Q(x)} \right]$

相对熵？