

Subject: Stanford CS229 Machine Learning, Lecture 12, K-Means, GMM (non EM), Expectation Maximization

Date: from August 26, 2025 to August 28, 2025

Contents

A	Proof of Eq. (1)	7
B	Proof of Theorem 1	7

Stanford CS229 Machine Learning, GMM (EM), 2022, Lecture 13

Link on YouTube: Stanford CS229 Machine Learning, GMM (EM), 2022, Lecture 13

Introduction

Introduction

本节内容介绍聚类算法中 GMM 算法及 EM 算法. 首先介绍两个基本工具: 凸性和 Jensen 不等式, 在此基础上, 通过 EM algorithm as MLE 介绍 EM 算法原理, 进一步通过 GMM as EM, 将 GMM 作为 EM 算法的实例介绍 GMM.

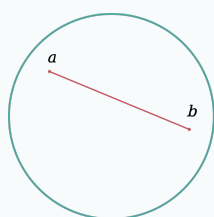
Convexity and Jensen's inequality

Convexity

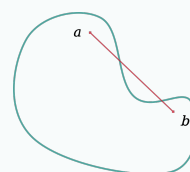
Definition 1. Convex and Nonconvex. 称一个集合 Ω 是凸的(convex), 若对于集合 Ω 中任意两点 $a, b \in \Omega$, 其二者之间的连线也均在该集合中. 数学上即:

$$\forall a, b \in \Omega, \quad (1 - \lambda)a + b \in \Omega, \quad \lambda \in [0, 1].$$

如图2所示, 图1a 中集合为凸的, 但图1b 中集合为非凸的.



(a) Convex



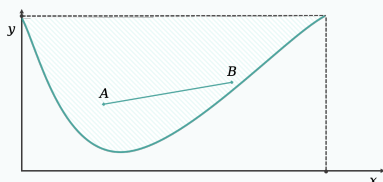
(b) Nonconvex

Figure 1: Comparison between convex and nonconvex.

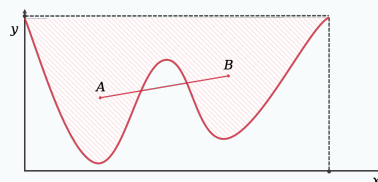
Definition 2. Convex Function. 对于函数 f , 定义其 graph 为 $G_f = \{(x, y) : y \geq f(x)\}$, 则称 f 为凸函数(convex function) 若其 graph 是凸的. 数学上即 (推导见附录A) :

$$\forall \lambda \in [0, 1], \lambda f(a) + (1 - \lambda)f(b) \geq f(z), \quad z = \lambda a + (1 - \lambda)b \quad (1)$$

如图2a 中函数为凸函数, 其中每一条弦 (即两点连线) 均在函数之上(every chord is beyond the function); 图2b 中为非凸函数, 不能保证每一条弦均在函数之上.



(a) Convex function



(b) Nonconvex function

Figure 2: Comparison between convex and nonconvex function

Theorem 1. 若函数 f 的二阶导恒大于等于 0, 即 $f''(x) \geq 0, \forall x$, 则 f 为凸函数.

Proof. 见附录B □

Definition 3. Strictly Convex. 称 f 是严格凸函数(*strictly convex*), 若 $f''(x) > 0, \forall x \in \text{dom}$.

Definition 4. Concave Function. 相对于凸函数, 称函数 g 为凹函数(*concave function*) 若 $-g$ 为凸函数.

Jensen's inequality

Theorem 2. Jensen's Inequality. 若 f 为凸函数, 则

$$\mathbb{E}[f(x)] \geq f(\mathbb{E}(x)). \quad (2)$$

例如, $P(x = a) = \lambda, P(x = b) = 1 - \lambda$, 则 $\mathbb{E}[f(x)] = \lambda f(a) + (1 - \lambda)f(b), f(\mathbb{E}[x]) = f(\lambda a + (1 - \lambda)b)$, 因此结合凸函数定义, 显然 Jensen's inequality 成立.

EM Algorithm as MLE

EM Algorithm as MLE

给定数据 $\{x_i\}_{i=1}^n$ 和参数 θ , 标准的最大似然估计(maximum likelihood estimation, MLE) 形式为:

$$l(\theta) = \sum_{i=1}^n \log P(x_i; \theta), \quad (3)$$

在已知信息的基础上, 若希望对问题引入一些特定 (但可能未知) 的结构时就会引入 **latent variable**, 例如 Lecture 12 的光子接收器中对每一个光子的来源比例做出 Θ 的结构性质假设. 此时 $P(x_i; \theta)$ 变为:

$$P(x; \theta) = \sum_z P(x, z; \theta). \quad (4)$$

EM Algorithm: Idea

如图3 所示, 由于原目标函数 $l(\theta)$ 较难优化 (一般无凸、凹性), 因此 EM 算法首先在 $\theta^{(t)}$ 处使用存在凸/凹性的新目标函数 $L_t(\theta)$ 替换 $l(\theta)$, 因此比 $l(\theta)$ 更容易优化, 接着再更新 $\theta^{(t+1)} = \arg \max_{\theta} L_t(\theta)$. 其中替代函数(surrogate) $L_t(\theta)$ 需要满足两点性质:

1. *Low bound:* $L_t(\theta) \leq l(\theta)$, 即替代函数是原目标函数的下界
2. *Tight:* $L_t(\theta^{(t)}) = l(\theta^{(t)})$, 即替代函数与原目标函数在当前迭代点的函数值相等

因此, EM 算法简单来说分为两步:

1. (E-step): 给定 $\theta^{(t)}$ 寻找 $L_t(\theta)$
2. (M-step): 给定 $L_t(\theta)$, 更新 $\theta^{(t+1)} = \arg \max_{\theta} L_t(\theta)$

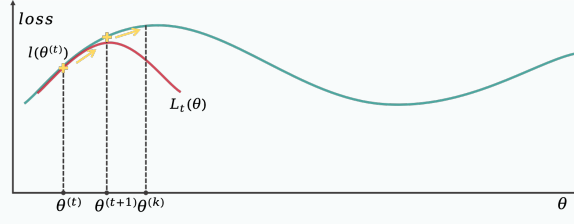


Figure 3: EM algorithm.

EM Algorithm: Select L_t

Lower bounds: 在引入 latent variable z 后, $\log P(x; \theta)$ 变为:

$$\log P(x; \theta) = \log \sum_z P(x, z; \theta) = \log \sum_z \frac{Q(z) \cdot P(x, z; \theta)}{Q(z)}, \forall Q(z), \quad (5)$$

其中 $Q(z)$ 为一离散分布, 满足 $\sum_z Q(z) = 1, Q(z) \geq 0$.

式(5) 可以进一步改写, 并应用 Jensen's inequality 可得:

$$\begin{aligned} \log \sum_z \frac{Q(z) \cdot P(x, z; \theta)}{Q(z)} &= \log \mathbb{E}_{z \sim Q(z)} \frac{P(x, z; \theta)}{Q(z)} \geq \mathbb{E}_{z \sim Q(z)} \left[\log \frac{P(x, z; \theta)}{Q(z)} \right] \\ &= \sum_z Q(z) \cdot \log \frac{P(x, z; \theta)}{Q(z)} \end{aligned} \quad (6)$$

Note 1. 注意此时满足 $\sum_z Q(z) = 1, Q(z) \geq 0$ 的 $Q(z)$ 都满足式(6), 这样我们就获得了 $l(\theta)$ 的下界集合, 集合中的函数 L 均满足 $L(\theta) \leq l(\theta)$.

Note 2. 式(6)中 $\sum_z Q(z) \cdot \log \frac{P(x, z; \theta)}{Q(z)}$ 被称为 **Evidence-based lower bound (ELBO)**, 即

$$ELBO(x, Q, z) = \sum_z Q(z) \cdot \log \frac{P(x, z; \theta)}{Q(z)}. \quad (7)$$

因此有

$$l(\theta) \geq \sum_{i=1}^n ELBO(x^{(i)}, Q^{(i)}; \theta). \quad (8)$$

在获得了 $l(\theta)$ 的下界集合后, 符合要求的 L_t 还需要其满足 *Tight* 的性质.

Make it tight. 从式(6)可以看到, 不等号是由求 $\log \frac{P(x, z; \theta)}{Q(z)}$ 的期望产生的, 因此为消除不等号可以令其为独立于 z 的常数, 即令 $\log \frac{P(x, z; \theta)}{Q(z)} = C$. 此时令 $Q(z) = P(z|x; \theta)$ 就有:

$$\log \frac{P(x, z; \theta)}{Q(z)} = \log \frac{P(z|x; \theta) \cdot P(x; \theta)}{P(z|x; \theta)} = \log P(x; \theta). \quad (9)$$

由于消除了关于 z 的随机性, 此时就有

$$l(\theta^{(t)}) = \sum_{i=1}^n \log P(x_i; \theta^{(t)}) = \sum_{i=1}^n ELBO(x^{(i)}, Q^{(i)}; \theta^{(t)}), \quad (10)$$

其中 $Q^{(i)} = P(z^{(i)}|x^{(i)}; \theta^{(t)})$.

Note 3. 由于 $Q^{(i)}$ 的选取与 x, θ 相关, 因此针对 $\theta^{(i)}$ 点处的替代函数而言, 因为选取了 $Q^{(i)} = P(z|x^{(i)}; \theta^{(i)})$, 因此只有 $\theta^{(i)}$ 处满足等式, 而其他 θ_β 处只有 $Q = P(z|x^{(i)}; \theta_\beta)$ 才满足等式, 因此只能满足不等式 $l(\theta_\beta) \geq \sum_{i=1}^n ELBO(x^{(i)}, Q^{(i)}; \theta^{(i)})$.

Note 4. $z^{(i)}$ 是对问题作一个潜在的结构假设, 例如在 Lecture 12 的光子实验中, 对于每个 sample i , $z^{(i)}$ 是对其来源的假设, 例如如果目前有两种来源, 那么可以假设 $P(z^{(i)} = 1) = \phi_1, P(z^{(i)} = 2) = \phi_2$, 这就是对此问题的一种潜在的结构假设.

EM Algorithm: Conclusion

综上所述 EM 算法分为如下两步:

1. (E-step): 给定 $\theta^{(t)}$, 令 $Q^{(i)} = P(z^{(i)}|x^{(i)}; \theta), i = 1, \dots, n$
2. (M-step): 给定 $L_t(\theta)$, 更新 $\theta^{(t+1)} = \arg \max_{\theta} L_t(\theta) = \sum_{i=1}^n ELBO(x^{(i)}, Q^{(i)}; \theta)$
1. 由于 $\theta^{(t+1)} = \arg \max_{\theta} L_t(\theta) = \sum_{i=1}^n ELBO(x^{(i)}, Q^{(i)}; \theta)$, 因此 $l(\theta^{(t+1)}) \geq l(\theta^{(t)})$, 故 EM 算法一定会收敛/ 终止.
2. 但是由于每次更新都选择了更优的参数, 因此并不能够跳出局部最优点, 故不能保证最终能收敛至全局最优.

EM for GMM

EM for GMM

Recall. 根据 Lecture 12 中 GMM 模型:

$$P(x^{(i)}, z^{(i)}) = P(x^{(i)}|z^{(i)}) \cdot P(z^{(i)}), \quad (11)$$

其中 $z^{(i)} \sim \text{Multinomial}(\Phi), \phi_i \geq 0, \sum_i \phi_i = 1, P(x^{(i)}|z^{(i)}) \sim \mathcal{N}(\mu_j, \sigma_j^2)$.

记 $Q^{(i)}(z^{(i)}) = P(z^{(i)}|x^{(i)}; \theta)$, 则此时 EM 算法中 M-step 为:

$$\max_{\Theta} \sum_{i=1}^n \sum_{z^{(i)}} Q^{(i)}(z^{(i)}) \log \frac{P(x^{(i)}, z^{(i)}; \theta)}{Q^{(i)}(z^{(i)})}, \quad (12)$$

其中 $\Theta = \{\Phi, \mu, \Sigma\}$.

Derivative for μ_j . 记 $f_i(\theta) = \sum_{z^{(i)}} Q^{(i)}(z^{(i)}) \log \frac{P(x^{(i)}, z^{(i)}; \theta)}{Q^{(i)}(z^{(i)})}$, $w_j^{(i)} = Q^{(i)}(z^{(i)} = j) = P(z^{(i)} = j|x^{(i)}; \theta)$, 则:

$$f_i(\theta) = \sum_j w_j^{(i)} \log \frac{\frac{1}{2\pi|\Sigma|^{1/2}} \exp\{-\frac{1}{2}(x^{(i)} - \mu_j)^T \Sigma_j^{-1} (x^{(i)} - \mu_j)\} \phi_j}{w_j^{(i)}}, \quad (13a)$$

$$\nabla_{\mu_j} \sum_{i=1}^n f_i(\theta) = \sum_{i=1}^n \nabla_{\mu_j} (w_j^{(i)} - \frac{1}{2}(x^{(i)} - \mu_j)^T \Sigma_j^{-1} (x^{(i)} - \mu_j)) = -\frac{1}{2} \sum_{i=1}^n w_j^{(i)} \Sigma_j^{-1} (x^{(i)} - \mu_j). \quad (13b)$$

令梯度 $\nabla_{\mu_j} \sum_{i=1}^n f_i(\theta) = 0$ ，即可以得到：

$$\mu_j = -\frac{\sum_i w_j^{(i)} (x^{(i)} - \mu_j)}{\sum_i w_j^{(i)}}. \quad (14)$$

Derivative for ϕ_j . 由于 ϕ_j 需要满足约束条件 $\sum_j \phi_j = 1, \phi_j \geq 0$ ，因此根据 Lagrangian，重新构造损失函数为^a：

$$\sum_{i=1}^n f_i(\theta) = \sum_{i=1}^n w_j^{(i)} \log \phi_j + \lambda (\sum_j \phi_j - 1) \quad (15)$$

再对 ϕ_j 求导可得：

$$\nabla_{\phi_j} \sum_{i=1}^n f_i(\theta) = \sum_{i=1}^n w_j^{(i)} \nabla_{\phi_j} \log \phi_j + \nabla_{\phi_j} \lambda (\sum_j \phi_j - 1) = \sum_{i=1}^n \frac{w_j^{(i)}}{\phi_j} + \lambda. \quad (16)$$

令 $\nabla_{\phi_j} \sum_{i=1}^n f_i(\theta) = 0$ ，得

$$\sum_j \phi_j = 1 = -\frac{1}{\lambda} \sum_{i,j} w_j^{(i)} = -\frac{n}{\lambda} \Rightarrow \lambda = -n. \quad (17)$$

^a这里为方便书写，省略了一些与求导无关的项。

A Proof of Eq. (1)

Proof of Eq. (1)

Proof. 根据凸函数定义, 有:

$$\forall \lambda \in [0, 1], \lambda(a, f(a)) + (1 - \lambda)(b, f(b)) \in \Omega \quad (18)$$

因此定义 $z = \lambda a + (1 - \lambda)b$, 则有

$$(z, \lambda f(a) + (1 - \lambda)f(b)) \in \Omega \Rightarrow \lambda f(a) + (1 - \lambda)f(b) \geq f(z) \quad (19)$$

□

B Proof of Theorem 1

Proof of Theorem 1

Proposition 1.

$$f''(x) \geq 0, \forall x \Rightarrow f \text{ is convex.} \quad (20)$$

Proof. 由 Taylor expansion 可得, $\forall z \in [a, b]$:

$$\begin{aligned} f(a) &= f(z) + f'(\alpha)(a - z) + f''(\alpha)(a - z)^2, \alpha \in [a, z] \\ f(b) &= f(z) + f'(\beta)(b - z) + f''(\beta)(b - z)^2, \beta \in [z, b], \end{aligned} \quad (21)$$

因此有

$$\lambda f(a) + (1 - \lambda)f(b) = f(z) + O(z) + C \geq f(z) \quad (22)$$

□

References