10.3

ノートのタイトル 2016/05/15

(10.16) $KL(p\|q) = -\int p(z)\left[\sum_{i=1}^{M} \ln q_i(z_i)\right] dz + c$

q_j についての制約条件は

$$\int q_j \, dz_j = 1$$

なので、KLの極値を求める為のラグランジュ関数は

$$\frac{\partial KL}{\partial q_i} = -\int p(z)\frac{1}{q_i} dz$$

$$\frac{\partial^2 KL}{\partial q_i^2} = \int p(z)\frac{1}{q_i^2} dz > 0$$

なのでKLは下に凸なので極値は極小になる。

$$L = KL(p\|q) + \lambda\left(\int q_j \, dz_j - 1\right)$$

となる。

ラグランジ乗数法の停留条件(E.3)より

$$\frac{\partial L}{\partial q_j} = \frac{\partial L}{\partial q_j} KL + \lambda \frac{\partial}{\partial q_j}\int q_j \, dz_j = 0 \quad \cdots ①$$

$$\frac{\partial L}{\partial \lambda} = \int q_j \, dz_j - 1 = 0 \quad \cdots ②$$

注 ~~汎ぎ積分で $\int_{-\infty}^{\infty}\frac{\partial}{\partial q_j} q_j dz_j$ が発散するので微分と積分は交換できない~~

そもそも q_j と z_j は独立でないので微分と積分は交換できない

となるが、①については $\frac{\partial}{\partial q_j}\int q_j \, dz_j$ をこれ以上変形することができず q_j について解くことができない。

なので q_j について別の停留条件を探す。

ここで、

$$L = KL(p\|q) + \lambda\left(\int q_j \, dz_j - 1\right)$$

$$= -\int p(z)\left[\sum_{i=1}^{M} \ln q_i\right] dz + c + \lambda\left(\int q_j \, dz_j - 1\right)$$

$$= -\int\left[\int p(z)\left[\sum_{i=1}^{M} \ln q_i\right]\prod_{i\neq j} dz_i\right] dz_j + \lambda\int q_j \, dz_j + (c+\lambda)$$

$$= \int\left[-\int p(z)\left[\sum_{i=1}^{M} \ln q_i\right]\prod_{i\neq j} dz_i + \lambda q_j + (c+\lambda)\delta(z_j)\right] dz_j$$

である。Lは q_i の汎関数になっている。

なのでLが停留する条件はオイラーラグランジュ方程式(D.8)で与えられる

(D.8)より

$$-\int p(z)\frac{1}{q_j}\prod_{i\neq j} dz_j + \lambda = 0$$

$$\frac{\partial}{\partial q_j}\int p(z)\sum \ln q_i \prod_{i\neq j} dz_i = \int p(z)\left(\frac{\partial}{\partial q_j}\sum \ln q_i\right)\prod_{i\neq j} dz_i = \int p(z)\frac{1}{q_j}\prod_{i\neq j} dz_i = \underbrace{\frac{1}{q_j}\int p(z)\prod_{i\neq j} dz_i}$$

この汎ぎ積分は収束するので最後の積分は実行可能

$$\therefore -\frac{1}{q_j}\int p(z)\prod_{i\neq j} dz_i + \lambda = 0$$

$p(z_1, z_2 \cdots)$ の $\{z_i\}_{i\neq j}$ についての周辺化になっているので

$$\therefore \lambda q_j = \int p(z)\prod_{i\neq j} dz_i = p(z_j) \quad \cdots ③$$

③の両辺を z_j で周辺化して②を代入する

$$\lambda\int q_j \, dz_j = \int p(z_j) dz_j$$

$$\therefore \lambda = 1$$

③に戻して

$$q_j = p(z_j)$$

を得る。