2022/08/18

Kullback-Leiblerダイバージェンスに基づく不確実性集合

hdk

定義 1(Kullback-Leiblerダイバージェンスに基づく不確実性集合)

$$\mathcal{Q}(\epsilon) = \{q \in \mathbb{R}^{ ext{n}}: 1^Tq = 1, q \geq 0, \sum_{i=1}^n q_i \ln rac{q_i}{p_i} \leq \epsilon \}$$

定理1
$$egin{array}{c} egin{array}{c} \egin{array}{c} egin{array}{c} egin{array}{c} egin{a$$

は以下のように簡約化される。

$$\max_{\mathbb{Q} \in \mathcal{Q}} \mathbb{E}_{\mathbb{Q}}[f] = \min_{\lambda > 0, \eta} \{ \epsilon \lambda + \eta + \lambda \sum_{i=1}^n p_i \expigg(rac{f_i - \eta}{\lambda} - 1igg) \}$$

https://mathlog.info/articles/3430 1/3 証明

$$egin{array}{ll} \max_q & f^T q \ ext{s.t.} & \sum_{i=1}^n q_i \ln rac{q_i}{p_i} \leq \epsilon \ & 1^T q = 1 \ q_i > 0, i = 1, 2, \cdots, n \end{array}$$

主問題の標準形式に書き直すと

$$egin{array}{ll} \min_{q} & -f^T q \ \mathrm{s.t.} & \sum_{i=1}^n q_i \ln rac{q_i}{p_i} - \epsilon \leq 0 \ 1^T q - 1 = 0 \ q_i > 0, i = 1, 2, \cdots, n \end{array}$$

式(1)に対するLagrange関数は

$$L(q,\lambda,\eta) = -f^T q + \lambda \left(\sum_{i=1}^n q_i \ln rac{q_i}{p_i} - \epsilon
ight) + \eta \left(1^T q - 1
ight)$$

である。ただし、双対変数を (λ,η) としている。 λ は不等式制約に対応し、 $\lambda \geq 0$ である。 η は等式制約に対応する。

双対問題は

$$\max_{\lambda > 0} \min_{\eta} L(q, \lambda, \eta) \tag{3}$$

と定義される。

 $\min_q L(q,\lambda,\eta)$ は q に関する無制約最小化であるから、

$$rac{\partial L(q,\lambda,\eta)}{\partial a}=0$$

以下細かい計算をしていく。

$$egin{aligned} rac{\partial L(q,\lambda,\eta)}{\partial q} &= rac{\partial}{\partial q} igg(-f^T q + \lambda \left(\sum_{i=1}^n q_i \ln rac{q_i}{p_i} - \epsilon
ight) + \eta \left(1^T q - 1
ight) igg) \ &= -f + \lambda \left(egin{aligned} dots \ rac{q_i}{p_i} + 1 \ dots \end{aligned}
ight) + \eta \ &= 0 \end{aligned}$$

よって、

$$-f_i + \lambda \left(\ln rac{q_i}{p_i} + 1
ight) + \eta = 0$$

式変形すると、

$$\ln rac{q_i}{p_i} + 1 = rac{f_i - \eta}{\lambda}$$

よって、

$$\ln \frac{q_i}{p_i} = \frac{f_i - \eta}{\lambda} - 1 \tag{4}$$

$$q_i = p_i \exp\left(rac{f_i - \eta}{\lambda} - 1
ight)$$
 (5)

$$\lambda \neq 0$$
 (6)

式(2)Lagrange関数を式(3)の双対問題に代入すれば、

$$egin{aligned} \max_{\lambda \geq 0, \eta} \min_{q} L(q, \lambda, \eta) &= \max_{\lambda \geq 0, \eta} \min_{q} -f^T q + \lambda \left(\sum_{i=1}^n q_i \ln rac{q_i}{p_i} - \epsilon
ight) + \eta \left(1^T q - 1
ight) \ &= \max_{\lambda \geq 0, \eta} \min_{q} \left(\cdots, -f_i + \lambda \ln rac{q_i}{p_i} + \eta, \cdots
ight) q - \epsilon \lambda - \eta \ &= \max_{\lambda \geq 0, \eta} - \lambda \sum_{i=1}^n p_i \exp \left(rac{f_i - \eta}{\lambda} - 1
ight) - \epsilon \lambda - \eta \ \left(\because \vec{\pi}(4), \vec{\pi}(5)
ight) \ &= \min_{\lambda \geq 0, \eta} \epsilon \lambda + \eta + \lambda \sum_{i=1}^n p_i \exp \left(rac{f_i - \eta}{\lambda} - 1
ight) \ &= \min_{\lambda > 0, \eta} \epsilon \lambda + \eta + \lambda \sum_{i=1}^n p_i \exp \left(rac{f_i - \eta}{\lambda} - 1
ight) \ \left(\because \vec{\pi}(6)
ight) \end{aligned}$$

参考文献

[1] S.Boyd and L.Vandenberghe, Convex Optimization