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）

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

定理 1

$$\max_{\mathbb{Q} \in \mathcal{Q}} \mathbb{E}_{\mathbb{Q}}[f] = \left| \begin{array}{ll} \displaystyle\max_{q} & f^T q \\ \text{s.t.} & \displaystyle\sum_{i=1}^{n} q_i \ln \frac{q_i}{p_i} \leq \epsilon \\ & 1^T q = 1 \\ & q_i > 0, i = 1, 2, \cdots, n \end{array} \right.$$

は以下のように簡約化される。

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

証明

$$
\left|
\begin{array}{ll}
\max\limits_{q} & f^T q \\
\text{s.t.} & \displaystyle\sum_{i=1}^{n} q_i \ln \frac{q_i}{p_i} \le \epsilon \\
& 1^T q = 1 \\
& q_i > 0, i = 1, 2, \cdots, n
\end{array}
\right.
$$

主問題の標準形式に書き直すと

$$
\left|
\begin{array}{ll}
\min\limits_{q} & -f^T q \\
\text{s.t.} & \displaystyle\sum_{i=1}^{n} q_i \ln \frac{q_i}{p_i} - \epsilon \le 0 \\
& 1^T q - 1 = 0 \\
& q_i > 0, i = 1, 2, \cdots, n
\end{array}
\right. \tag{1}
$$

式(1)に対するLagrange関数は

$$
L(q, \lambda, \eta) = -f^T q + \lambda \left(\sum_{i=1}^{n} q_i \ln \frac{q_i}{p_i} - \epsilon \right) + \eta \left(1^T q - 1 \right) \tag{2}
$$

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、 $\lambda \ge 0$ である。 η は等式制約に対応する。

双対問題は

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

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、

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

以下細かい計算をしていく。

$$
\begin{aligned}
\frac{\partial L(q, \lambda, \eta)}{\partial q} &= \frac{\partial}{\partial q} \left(-f^T q + \lambda \left(\sum_{i=1}^{n} q_i \ln \frac{q_i}{p_i} - \epsilon \right) + \eta \left(1^T q - 1 \right) \right) \\
&= -f + \lambda \begin{pmatrix} \vdots \\ \ln \dfrac{q_i}{p_i} + 1 \\ \vdots \end{pmatrix} + \eta \\
&= 0
\end{aligned}
$$

よって、

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

式変形すると、

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

よって、

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

$$q_i = p_i \exp\left(\frac{f_i - \eta}{\lambda} - 1\right) \tag{5}$$

$$\lambda \neq 0 \tag{6}$$

式(2)Lagrange関数を式(3)の双対問題に代入すれば、

$$\begin{aligned}
\max_{\lambda \geq 0, \eta} \min_q L(q, \lambda, \eta) &= \max_{\lambda \geq 0, \eta} \min_q -f^T q + \lambda \left(\sum_{i=1}^n q_i \ln \frac{q_i}{p_i} - \epsilon \right) + \eta \left(1^T q - 1 \right) \\
&= \max_{\lambda \geq 0, \eta} \min_q \left(\cdots, -f_i + \lambda \ln \frac{q_i}{p_i} + \eta, \cdots \right) q - \epsilon\lambda - \eta \\
&= \max_{\lambda \geq 0, \eta} -\lambda \sum_{i=1}^n p_i \exp\left(\frac{f_i - \eta}{\lambda} - 1 \right) - \epsilon\lambda - \eta \;\; (\because \text{式}(4)\text{、式}(5)) \\
&= \min_{\lambda \geq 0, \eta} \epsilon\lambda + \eta + \lambda \sum_{i=1}^n p_i \exp\left(\frac{f_i - \eta}{\lambda} - 1 \right) \\
&= \min_{\lambda > 0, \eta} \epsilon\lambda + \eta + \lambda \sum_{i=1}^n p_i \exp\left(\frac{f_i - \eta}{\lambda} - 1 \right) \;\; (\because \text{式}(6))
\end{aligned}$$

参考文献

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