

Abstract

MPF 及び CD-method の objection function を整理したノートです。

1 CD-method

$$CD(p^{(0)}\|p^{(1)}) := KL(p^{(0)}\|p^{(\infty)}) - KL(p^{(1)}\|p^{(\infty)}) \quad (1)$$

$$= \sum_x p^{(0)}(x) \log\left(\frac{p^{(0)}(x)}{p^{(\infty)}(x|\theta)}\right) - \sum_x p^{(1)}(x|\theta) \log\left(\frac{p^{(1)}(x|\theta)}{p^{(\infty)}(x|\theta)}\right) \quad (2)$$

$$= - \sum_x p^{(0)}(x) \log(p^{(\infty)}(x|\theta)) + \sum_x p^{(1)}(x|\theta) \log\left(\frac{p^{(1)}(x|\theta)}{p^{(\infty)}(x|\theta)}\right) + const \quad (3)$$

ただし、 $p^{(0)}$ は経験分布。 $p^{(1)}$ は経験分布から 1MC-step で到達できる確率分とします。一次元 (サイト数 d) イジングモデル、かつ、パラメータが一樣 (一変数) の場合は分配関数を簡単に書くことができ、以下のような式になります。

$$Z(\theta) = (2 \cosh(\theta))^d + (2 \sinh(\theta))^d \quad (4)$$

$$p^{(0)}(x) = \frac{1}{|\mathcal{D}|} \sum_{y \in \mathcal{D}} \prod_{j=1}^d \delta(x_j, y_j) \quad (5)$$

従って、第 1 項目は

$$\sum_x p^{(0)}(x) \log\left(\frac{p^{(0)}(x)}{p^{(\infty)}(x|\theta)}\right) = \frac{1}{|\mathcal{D}|} \left\{ - \sum_{x \in \mathcal{D}} E(x|\theta) - \log((2 \cosh(\theta))^d + (2 \sinh(\theta))^d) \right\} - \log(|\mathcal{D}|) \quad (6)$$

1MC-step の遷移で得られる確率分布は以下のように考えました。

まずは、初期状態と次の遷移先の同時確率を、初期状態の変数を周辺化するとみなします ($p^{(1)}(x)$ は一度の遷移で到達でき変数の確率分布であるため)。

$$p^{(1)}(x) = \sum_{x': all} p(x|x') p(x')^{(0)} = \sum_{x' \in \mathcal{D}} p(x|x') \frac{1}{|\mathcal{D}|} \quad (7)$$

経験分布はデータに含まれている状態のみ有限の値を持つとします

$$p^{(0)}(x) = \begin{cases} \frac{1}{|\mathcal{D}|} & \text{for } x \in \mathcal{D} \\ 0 & \text{for } x \notin \mathcal{D} \end{cases} \quad (8)$$

また、 $p(x|x')$ は humming 距離 1,0 またはそれ以上かの 3 通り場合分けをして考えます。

$$p(x|x') = \begin{cases} d^{-1}(1 + \exp\{-E(x'|\theta) + E(x|\theta)\})^{-1} & \text{for } \|x - x'\| = 1 \\ 1 - \sum_{x: all, \|x-x'\|=1} d^{-1}\{1 + \exp(-E(x'|\theta) + E(x|\theta))\}^{-1} & \text{for } \|x - x'\| = 0 \\ 0 & \text{for } \|x - x'\| > 1 \end{cases} \quad (9)$$

$x = x'$ の場合は任意の確率変数 x, y に対して、 $\sum_{x: all} p(x|y) = 1$ であることから導かれます。

従って、1MC-step で遷移可能な確率分布は

$$p^{(1)}(x) = |\mathcal{D}|^{-1} \sum_{x' \in \mathcal{D}} p(x|x') \quad (10)$$

$$= |\mathcal{D}|^{-1} \sum_{x' \in \mathcal{D}, \|x-x'\|=1} p(x|x') + |\mathcal{D}|^{-1} \sum_{x' \in \mathcal{D}: x'=x} p(x|x') \quad (11)$$

この $p^{(1)}$ を CD の第 2 項目に代入すれば良いのだと思います。

2 MPF

1MC-step の遷移で得られる MPF の objective-function は

$$MPF(\theta) = KL(p^{(0)} \| p^{(1)}) = - \sum_{x:all} p^{(0)}(x) \log(p^{(1)}(x|\theta)) + const \quad (12)$$

CD の章で得られた $p^{(1)}$ を用いると以下の式を得られます

$$MPF(\theta) = - \sum_{x:all} p^{(0)}(x) \log(p^{(1)}(x|\theta)) \quad (13)$$

$$= - \frac{1}{|\mathcal{D}|} \sum_{x \in \mathcal{D}} \log(p^{(1)}(x|\theta)) \quad (14)$$

$$= - \frac{1}{|\mathcal{D}|} \sum_{x \in \mathcal{D}} \log \left\{ \sum_{x' \in \mathcal{D}, \|x-x'\|=1} \{d^{-1}(1 + \exp\{-E(x'|\theta) + E(x|\theta)\})^{-1}\} \right\} \quad (15)$$

$$+ 1 - \sum_{x':all, \|x-x'\|=1} d^{-1} \{1 + \exp(-E(x|\theta) + E(x'|\theta))\}^{-1} \} \quad (16)$$

式の意味としては、最初の状態点に 1MC-step 先でも止まる確率分布となっています。(たぶん、最後の行が間違っているのではないかと思います。)

一応 $MPF(\theta)$ の一次元 $d=16$, パラメータ数が 1 つ (真値 = 0.5) の場合で、プロットすると以下の様な図を得ました (本来局値に成るべき 0.5 に局値が現れません。何かが変です。)

縦軸は MPF の値で、横軸はパラメータです。 $p^{(1)}(x|\theta)$ をもう一度考え直して