4L-06

階層ベイズ学習に基づく組み合わせ最適化問題の統計的分析

石岡龍佑 † 関本快士 ‡ 安田宗樹 §
山形大学大学院理工学研究科 † 山形大学大学院理工学研究科 ‡ 山形大学大学院理工学研究科 §

1. はじめに

ネットショッピングの需要増加における配達経路最適化，働き方改革のためのスケジューリング問題といった現実社会における意思決定，問題解決を実現する有用な手段である組み合わせ最適化問題への期待・関心が高まっており，注目度の高い問題クラスである．組み合わせ最適化問題ではコスト関数が必須であるが，現実の問題に沿ったコスト関数の定式化は容易ではない．そこで，本稿では組み合わせ最適化問題をエネルギーベースモデル (energy-based model (EBM)) [1] と呼ばれる確率モデル（コスト関数をエネルギー関数に読み替えたギブス分布）を通して扱い，階層ベイズ学習に基づくコスト関数推定法を提案する．これによって，未定式のコスト関数をデータ集合から定式化できる．

2. 逆温度を導入したエネルギーベースモデル

EBM では n 次元データ $\boldsymbol{x} := \{x_i \mid i = 1, 2, \ldots, n\}$ を入力とし, スカラー値を出力とするエネルギー関数 $E(\boldsymbol{x} \mid \theta)$ に対して，以下の形で確率分布を定義する．

$$P(\boldsymbol{x} \mid \theta,\beta) := \frac{\exp(-\beta E(\boldsymbol{x} \mid \theta))}{Z(\theta,\beta)} \tag{1}$$

ここで，θ はエネルギー関数のパラメータ，β は最適解をシミュレーテッドアニーリング的に推定するための逆温度パラメータ，$Z(\theta,\beta)$ は分配関数である．エネルギー関数 $E(\boldsymbol{x} \mid \theta)$ を小さくする $\boldsymbol{x}$ が確率をより高くする．逆温度パラメータ β を変化させると図 1 に示すように，逆温度が β が小さい場合は平坦な分布となり，逆に大きい場合は高確率付近に鋭いピークを持つ分布となる．

最適解が複数ある場合や良質な準最適解が存在する場合 β を大きくすると，複数の鋭いピークが存在する．その場合，各要素 x_i に揺らぎ（分散）$V[x_i]$ が残る．この各要素の揺らぎを用いて最適解の頑健性や各要素の最適解に対する重要性を表す指標である重要度 τ_i を推定する方法を提案した [2].

$$\tau_i = \frac{V_{\max} - V[x_i]}{V_{\max}} \quad (0 \le \tau_i \le 1) \tag{2}$$

ここで，$V_{\max}$ は要素がとりうる揺らぎ $V[x_i]$ の最大値を表す．重要度の値が大きいほど，最適解の頑健性や最適解に対する重要性を示す要素であり，値が小さいほど良質な準最適解の発見に起因する要素といえる．組み合わせ最適化を現実社会で用いる際，最適性の担保，要望の即して柔軟な対応など要素ごとの働きを持つ．

3. エネルギー関数推定

本節では，入力データ点の集合と対応するエネルギー値の集合からエネルギー関数（問題に対応するパラメータ）を推定する手法を提案する．

訓練データ集合 $D = \{\mathrm{e}^{(\mu)}, \mathbf{x}^{(\mu)} \mid \mu = 1, 2, \ldots, N\}$ を用いてエネルギー関数を推定する．データ点の集合 $\mathrm{e}^{(\mu)}$ は未知のコスト関数から得られたコスト値であり $\mathbf{x}^{(\mu)}$ は $\mathrm{e}^{(\mu)}$ が得られた時の入力を表す．コスト値にノイズがまざる理由として，現実社会において環境や人為的な問題から同じ入力でも出力は違うことが想定される．例として最短経路探索における，天気や工事・事故による通行止めがある．そのため，$\mathrm{e}^{(\mu)} = E_\theta(\mathbf{x}^{(\mu)}) + \epsilon^{(\mu)}$ であり $\epsilon^{(\mu)}$ はノイズを表す．エネルギー関数推定として一般的に最尤推定を用いる．最尤推定では学習モデルが訓練データ集合を生成する確率を表す尤度関数を定義する．

$$L(D \mid \theta,\beta) := \prod_{\mu=1}^{N} P(\mathbf{x}^{(\mu)} \mid \theta,\beta) \tag{3}$$

$L(D \mid \theta,\beta)$ を最大化するパラメータを求める最尤推定の枠組みによってエネルギー関数推定ができる．

最尤推定ではデータ数が少ない場合推定したエネルギー関数の推定精度が低下してしまう．そこで，階層ベイズ学習を用いたエネルギー関数推定法を提案する．まず，パラメータについての事前分布 $P_{\mathrm{pri}}(\theta \mid \lambda)$ を導入し

Statistical analysis of combinatorial optimization problems based on hierarchical Bayesian learning

† Ryusuke Ishioka; Graduate School of Science and Engineering, Yamagata University
‡ Kaiji Sekimoto; Graduate School of Science and Engineering, Yamagata University
§ Muneki Yasuda; Graduate School of Science and Engineering, Yamagata University

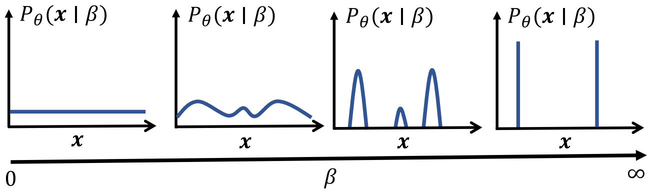

図 1: 確率分布と逆温度の関係図.

モデルを階層化する.

$$P_{\mathrm{pri}}(\theta \mid \lambda) := \prod_{\theta_c} \frac{1}{\sqrt{2\pi\lambda^{-1}}} \exp\left(-\frac{\theta_c^2}{2\lambda^{-1}}\right) \quad (4)$$

ここで，λ はパラメータのハイパパラメータである．次に，式 (3) の尤度関数と式 (4) の事前分布を用いて事後分布 $P_{\mathrm{post}}(\theta \mid D, \alpha, \lambda)$ を表す.

$$P_{\mathrm{post}}(\theta \mid D, \lambda, \beta) \propto L(D \mid \theta, \beta) P_{\mathrm{pri}}(\theta \mid \lambda) \quad (5)$$

式 (5) の事後分布と式 (1) の EBM を用いて階層ベイズ学習を行う新たな提案モデル $P(\boldsymbol{x} \mid D, \beta)$ を設計する.

$$P(\boldsymbol{x} \mid D, \beta) := \int P(\boldsymbol{x} \mid \theta, \beta) P_{\mathrm{post}}\,(\theta \mid D, \alpha, \lambda)\, d\theta \quad (6)$$

事後分布からパラメータのサンプル点を生成し，標本平均でモデル近似することですべてのパラメータについてあらゆるケースを考慮していると考えられる．そのため提案モデルは

$$P(\boldsymbol{x} \mid D, \beta) \approx \frac{1}{K} \sum_{k=1}^{K} P(x \mid \beta, \theta^{(k)}) \quad (7)$$

と近似できる.

4. 数値実験

本節では，エネルギー関数推定において最尤推定と階層ベイズ法を用いた提案モデルでエネルギー関数推定を行いそれぞれの推定精度を評価する．本実験のエネルギー関数は多くの組み合わせ最適化問題に用いられるイジング模型のエネルギー関数を用いた.

$$E_{\mathrm{IM}}(\boldsymbol{x} \mid \theta) = -\sum_{i \in \Omega} b_i x_i - \sum_{\{i,j\} \in \mathcal{E}} w_{ij} x_i x_j \quad (8)$$

$\theta = \{b_i, w_{ij}\}$ である．イジング模型は $n = 16$ 個のノードから成る完全無向グラフ上 $G(\Omega, \mathcal{E})$ に定義した．まず，訓練データ集合の入力 $\mathbf{x} = \{x_i \in \{0, 1\} \mid i = 1, 2, \ldots, 16\}$ を乱数生成し，N 点の入力データ点の集合を生成．次にデータ点 $\mathbf{x}$ に対応するエネルギー値 e を生成モデルから計算しノイズ ϵ をのせた．この時ノイズはガウス分布 $N(0, 0.01)$ で生成した．得られた訓練データ集合 $D = \{\mathrm{e}^{(\mu)}, \mathbf{x}^{(\mu)} \mid \mu = 1, 2, \ldots, N\}$ を用いて学習モデルを最尤推定または階層ベイズ学習（提案法）により学習し，エネルギー関数推定を行った．推定精度は，生成モデルの各要素の重要度と学習済みモデルのそれとの (mean absolute error(MAE)) で評価する.

図 2 より提案法は従来法より少ないデータ数においてMAE が小さく，データが豊富にある場合においても同程度の精度が得られることが確認できる.

5. まとめ

本稿では，組み合わせ最適化問題のエネルギー関数推定について階層ベイズ学習法という新たな手法の提案を行った．データ数が少ない場合，エネルギー関数推定の精度が低い問題を解消するためパラメータ θ についてあらゆるケースを考慮する階層ベイズ学習法を提案した．数値実験を通して、図 2 で示す通り提案法は最尤推定よりも高い精度でエネルギー関数推定が可能であることがわかった。

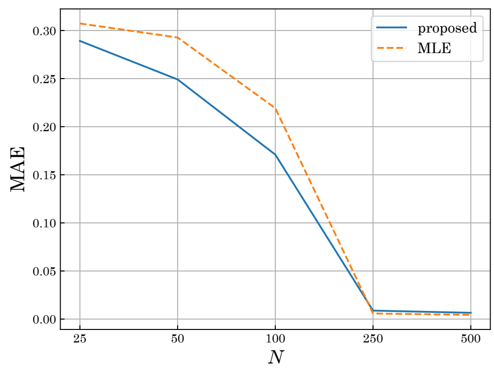

図 2: 重要度厳密値と最尤推定，提案法の平均絶対誤差.

謝辞

本稿は科研費（21K11778）の助成を受けたものである.

文献

[1] Y. LeCun, S. Chopra, R. Hadsell, M. Ranzato, and F. Huang: A tutorial on energy-based learning, MIT Press, 2006.

[2] 石岡龍佑，安田宗樹：最適化問題における統計的揺らぎの分析，第 85 回情報処理学会全国大会，2023.