

# 空間モンテカルロ積分法と焼なまし重点サンプリングを組み合わせた 高性能サンプリング近似法

関本快士<sup>†</sup>                      安田宗樹<sup>‡</sup>  
山形大学工学部<sup>†</sup>          山形大学大学院理工学研究科<sup>‡</sup>

## 1. はじめに

マルコフ確率場 (Markov random field (MRF)) は機械学習をはじめとする様々な分野において重要なモデルである。このモデルの学習・推論過程では期待値計算が必要となるが、これは組み合わせ爆発の問題を含んでいるため、実用的には近似が必要となる。空間モンテカルロ積分法 (spatial Monte Carlo integration (SMCI)) [1] は高い近似性能をもつことが知られているが、良質なサンプリングが難しい複雑な分布に対しては性能が大きく低下してしまう。本研究では SMCI と焼なまし重点サンプリング (Annealed importance sampling (AIS)) [2] を組み合わせた新しい近似法を提案する。AIS はこの性能低下を抑制することができるため、これを組み込むことで SMCI の近似性能を改善することができる。

## 2. マルコフ確率場

MRF は無向グラフ上で定義される確率モデルである。本研究では、 $n$  個の確率変数  $\mathbf{x} = \{x_i | i \in \mathcal{V} = \{1, 2, \dots, n\}\}$  から成る MRF を扱う。確率変数  $\mathbf{x}_{\mathcal{T}} = \{x_i | i \in \mathcal{T} \subseteq \mathcal{V}\}$  を対象としたある関数  $f(\mathbf{x}_{\mathcal{T}})$  の MRF 上での期待値は次の形で表される。

$$\mathbb{E}[f(\mathbf{x}_{\mathcal{T}})] = \sum_{\mathbf{x}} f(\mathbf{x}_{\mathcal{T}})P(\mathbf{x}) \quad (1)$$

ここで、 $P(\mathbf{x})$  は MRF の確率分布である。上式に含まれている多重和  $\sum_{\mathbf{x}}$  は、 $n$  の増加に伴って計算量が指数的に増加するため、一般的に厳密に計算することは難しい。よって、現実的な時間内で計算を行うために、近似を用いる必要がある。

## 3. サンプリング近似法

### 3.1. 空間モンテカルロ積分法

Effective sampling approximation based on spatial Monte-Carlo integration and Annealed importance sampling

<sup>†</sup> Kaiji Sekimoto; Faculty of Engineering, Yamagata University

<sup>‡</sup> Muneki Yasuda; Graduate School of Science and Engineering, Yamagata University

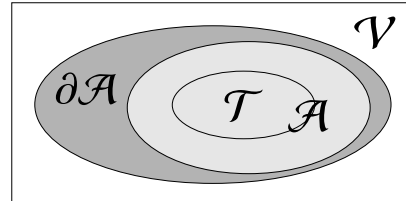


図1 SMCIで考慮される各領域の概念図。

SMCI は近似法として有名なモンテカルロ積分法よりも平均的に高い近似性能をもつ [1]。SMCI では、 $\mathcal{T}$  を含む部分集合  $\mathcal{A}$  と、 $\mathcal{A}$  の最近接集合  $\partial\mathcal{A}$  を考慮して近似する。 $\mathcal{T}$ 、 $\mathcal{A}$ 、 $\partial\mathcal{A}$  は、それぞれ目標領域、和領域、サンプル領域と呼ばれる。図1は各領域の関係を図式化したものである。

$P(\mathbf{x})$  から独立同分布で生成された  $N$  個のサンプル点の集合  $\mathbb{S} = \{\mathbf{s}^{(\mu)} | \mu = 1, 2, \dots, N\}$  を用いて、期待値は次のように近似される。

$$\mathbb{E}[f(\mathbf{x}_{\mathcal{T}})] \approx \frac{1}{N} \sum_{\mu=1}^N \sum_{\mathbf{x}_{\mathcal{A}}} f(\mathbf{x}_{\mathcal{T}})P(\mathbf{x}_{\mathcal{A}} | \mathbf{s}_{\partial\mathcal{A}}^{(\mu)}) \quad (2)$$

ここで、 $\mathbf{s}_{\partial\mathcal{A}}^{(\mu)}$  はサンプル領域の確率変数  $\mathbf{x}_{\partial\mathcal{A}}$  に対応するサンプル点である。SMCI の性能はサンプル点の質に強く依存するため、良質なサンプリングが難しい複雑な分布では近似性能が大きく低下してしまう。しかし、AIS はそのような性能低下を抑制することができる。

### 3.2. 焼なまし重点サンプリング

AIS は焼なまし法 (Simulated Annealing (SA)) 的なマルコフ連鎖の系列を用いた重点サンプリングの1種である [2]。AIS では、焼なましスケジュール  $0 = \beta_0 < \beta_1 < \dots < \beta_K = 1$  に従って、分布  $P_k(\mathbf{x}) \propto P_k^\dagger(\mathbf{x}) = P_0^\dagger(\mathbf{x})^{1-\beta_k} P^\dagger(\mathbf{x})^{\beta_k}$  の系列を構築する。ここで、 $P_0(\mathbf{x}) \propto P_0^\dagger(\mathbf{x})$  はサンプリングが容易な初期分布で、 $P^\dagger(\mathbf{x})$  は  $P(\mathbf{x})$  の非規格化分布である。これら分布について詳細つり合い条件を満たすマルコフ遷移を定義し、これを用いて SA 的に  $K$  個のサンプル点の系列  $\mathbb{X} = \{\mathbf{x}_k | k = 1, 2, \dots, K\}$  を生成する。取得したサンプル点  $\mathbf{x}_K$  を  $P(\mathbf{x})$  から生成されたサンプル点  $\mathbb{S}$  として採用する。そして、対応する

重点重みは次の形で得られる.

$$\omega = \prod_{k=1}^K \frac{P_k^\dagger(\mathbf{x}_k)}{P_{k-1}^\dagger(\mathbf{x}_k)} \quad (3)$$

このサンプリングを  $N$  回繰り返して得られたサンプル点の集合  $\mathbb{S}_{\text{AIS}} = \{\mathbf{s}^{(\mu)} \mid \mu = 1, 2, \dots, N\}$  と対応する重点重みの集合  $\Omega = \{\omega_\mu \mid \mu = 1, 2, \dots, N\}$  を用いて, 期待値は次のように近似される.

$$\mathbb{E}[f(\mathbf{x}_{\mathcal{T}})] \approx \frac{\sum_{\mu=1}^N \omega_\mu f(\mathbf{s}_{\mathcal{T}}^{(\mu)})}{\sum_{\mu=1}^N \omega_\mu} \quad (4)$$

#### 4. 提案法

本節では, SMCI と AIS を組み合わせた新しいサンプリング近似法を提案する. SMCI の近似式 (2) は関数

$$f_{(\mathcal{T}, \mathcal{A})}(\mathbf{x}_{\partial \mathcal{A}}) = \sum_{\mathbf{x}_{\mathcal{A}}} f(\mathbf{x}_{\mathcal{T}}) P(\mathbf{x}_{\mathcal{A}} \mid \mathbf{x}_{\partial \mathcal{A}}) \quad (5)$$

についての期待値  $\mathbb{E}[f_{(\mathcal{T}, \mathcal{A})}(\mathbf{x}_{\partial \mathcal{A}})]$  をモンテカルロ積分法を用いて近似したものと解釈できる. 提案法は  $\mathbb{E}[f_{(\mathcal{T}, \mathcal{A})}(\mathbf{x}_{\partial \mathcal{A}})]$  について AIS を用いて近似したものとなる.

提案法では,  $\mathbb{S}_{\text{AIS}}$  と  $\Omega$  を用いて  $\mathbb{E}[f(\mathbf{x}_{\mathcal{T}})]$  を次のように近似する.

$$\mathbb{E}[f(\mathbf{x}_{\mathcal{T}})] \approx \frac{\sum_{\mu=1}^N \omega_\mu f_{(\mathcal{T}, \mathcal{A})}(\mathbf{s}_{\partial \mathcal{A}}^{(\mu)})}{\sum_{\mu=1}^N \omega_\mu} \quad (6)$$

$N$  と  $K$  が十分に大きいと仮定すると, AIS と提案法の期待値近似の漸近分散  $\mathbb{V}_{\text{AIS}}[f(\mathbf{x}_{\mathcal{T}})]$  と  $\mathbb{V}_{\text{proposed}}[f(\mathbf{x}_{\mathcal{T}})]$  の間には, 近似的に次の不等式が成り立つ.

$$\mathbb{V}_{\text{proposed}}[f(\mathbf{x}_{\mathcal{T}})] \leq \mathbb{V}_{\text{AIS}}[f(\mathbf{x}_{\mathcal{T}})] \quad (7)$$

この不等式から, 提案法は標準的な AIS よりも平均的に高い近似性能をもつことが期待される.

#### 5. 数値実験

本節では, MRF の特殊形であるボルツマンマシンを用いて, 提案法の有効性を数値的に実証する. 接続確率  $p$  のランダムグラフ  $G(\mathcal{V}, \mathcal{E})$  上に,  $n = 20$  個の確率変数  $x_i \in \{-1, +1\}$  をもつボルツマンマシン [3]

$$Q(\mathbf{x}; \beta, \mathbf{b}, \mathbf{w}) \propto \exp \beta \left( \sum_{i \in \mathcal{V}} b_i x_i + \sum_{(i,j) \in \mathcal{E}} w_{ij} x_i x_j \right) \quad (8)$$

を定義し, パラメータ  $b_i, w_{ij}$  を区間  $[-1, +1]$  の一様乱数からランダムに生成した.  $\beta \geq 0$  は分布の複雑さを表すパラメータであり,  $\beta$  が大きいと複雑な分布になる. 本節の実験では, 近似精度を共分散  $\mathcal{X}_{ij}$  の平均絶対誤差

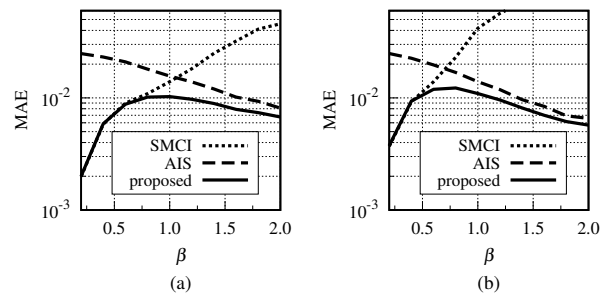


図2 (a)  $p = 0.4$ , (b)  $p = 0.8$  のランダムグラフ上での  $\beta$  に対する共分散  $\mathcal{X}_{ij}$  の MAE. 各プロットは 1000 回の実験の平均値である.

(mean absolute error (MAE))

$$\frac{1}{|\mathcal{E}|} \sum_{(i,j) \in \mathcal{E}} \left| \mathcal{X}_{ij}^{\text{exact}} - \mathcal{X}_{ij}^{\text{approx}} \right| \quad (9)$$

で評価した. ここで,  $\mathcal{X}_{ij}^{\text{exact}}$  は厳密な共分散で,  $\mathcal{X}_{ij}^{\text{approx}}$  は近似法から得られた共分散である. SMCI と提案法では,  $\mathcal{T} = \mathcal{A}$  とする 1 次の SMCI [1] を用いた. 焼なましスケジュールのサイズは  $K = 1000$  で,  $\beta_k = k/K$  とした.  $P_0(\mathbf{x})$  には一様分布を採用し, マルコフ遷移としてギブスサンプリングを用いた.  $\mathbb{S}$  は  $\mathbb{S}_{\text{AIS}}$  の生成方法と同一の方法で取得し, サイズは  $N = 1000$  とした.

図2に  $\beta$  に対する結果を示す. 提案法の近似精度は単純な分布 ( $\beta \leftarrow 0$ ) では SMCI と同程度であり, 複雑な分布 ( $\beta \gg 1$ ) では AIS よりも高いため, 提案法は  $\beta$  の広い範囲で高い近似性能をもつことが確認できた.

#### 6. まとめ

本研究では, SMCI と AIS を組み合わせた新しい高性能サンプリング近似法を提案した. AIS の重点重みは複雑な分布での SMCI の近似性能を大幅に改善することが実験的に示された. また, 提案法は標準的な AIS よりも平均的に高精度で近似できることが理論的かつ数値的に実証された.

#### 謝辞

本研究は科研費(18K11459, 18H03303), JST-CREST(JPMJCE1312) 及び JST COI プログラム (JPMJCE1312) の助成を受けたものである.

#### 文献

- [1] M. Yasuda, Monte Carlo Integration Using Spatial Structure of Markov Random Field, *Journal of the Physical Society of Japan*, Vol. 84, No. 3, 034001, 2015.
- [2] R. M. Neal, Annealed importance sampling, *Statistics and Computing*, Vol. 11, pp. 125–139, 2001.
- [3] D. H. Ackley, G. E. Hinton, and T. J. Sejnowski, A Learning Algorithm for Boltzmann Machines, *Cognitive Science*, Vol. 9, pp. 147–169, 1985.