論文の内容の要旨

論文題目: On some asymptotic properties of the Expectation-Maximization Algorithm and the Metropolis-Hastings Algorithm (EM アルゴリズムとメトロポリス-ヘイスティングスアルゴリズムの漸近的性質)

氏名: 鎌谷 研吾

本論文では, EM アルゴリズムと, メトロポリス-ヘイスティングスアルゴリズムの二種類のモンテカルロ法の漸近的性質を研究する. これら二つの手法は, ともに単純な計算の繰り返しにより, 求める統計量の近似計算をするという共通点がある.

まず最初に, 本論文では漸近的という言葉を二種類に用いることに注意する. 一つ目の意味は, 上記モンテカルロ法の繰り返し計算の回数を大きくしていくという意味である. 上記モンテカルロ法の漸近的性質を調べている殆どの既存研究は, この意味での漸近理論である. 本論文に置いても第三章は, この意味でのメトロポリス-ヘイスティングスアルゴリズムの収束を扱う. 二つ目の意味は, 繰り返しの数だけでなく, サンプルサイズも大きくなるという意味である. 第一章, 第二章ではこの意味での収束を考える. 必然的に, 後者の意味では, 大標本を仮定する.

最初の二つの章では, EM アルゴリズムとギブスサンプリングの漸近的理論の枠組みを作り上げる. ここで, ギブスサンプリングは, メトロポリス-ヘイスティングスアルゴリズムの主要な一種類である. この枠組みの作成には, 次のような三つの動機がある.

まず第一に, EM アルゴリズムとギブスサンプリングという二つの歴史のあるアルゴリズムを, 新しい意味で正当化することである. EM アルゴリズムについては, そのアルゴリズムの作る数列が, 正しく最尤推定量に収束することを示す. これは従来の, サンプルサイズを止めた考え方では難しい. なぜなら, 繰り返しの回数を増やすだけでは, 一般的には最尤推定量に収束するとは限らず, 局所解などに収束することがあるからである. 一方で, ギブスサンプリングにおいては, 既に繰り返しの回数を増やす考え方で, 収束理論は確立されていた. しかしながら, この収束理論はパラメータ空間の淵の部分での性質に強く影響されるものである. 一方, 今回の収束理論は, ギブスサンプリングのパラメータの真値周辺での性質であり, 中央値推定量などパラメータの空間の淵の影響をあまり重要視しない統計量を考えるのであれば, より本質的な収束理論と考えられる.

第二に, EM アルゴリズムやギブスサンプリングに関する, 様々な高速化手法の正当化を言うことである. これにはまず, 複数のアルゴリズムの比較法を確立することが必要である. 従来, EM アルゴリズムにおいての手法の比較は, 収束理論の改善は, レート行列という, フィッシャー情報行列と関連するある行列の比較をすることでなされていた. しかし, この比較の正当化は, 少々厳密性に欠ける方法でしかなされていなかった. 今回の研究ではこのレート行列の比較を正当化し, いくつかのアルゴリズムの比較をおこなった. ギブスサンプリングにおいては, フィッシャー情報行列に関連する行列をもちいて, 新しい比較法を確立し, 正当化した. この比較法の方が, 従来のスペクトル半径を比較する方法より, 実用的である.

第三にこれらの新しい漸近理論が, EM アルゴリズムやギブスサンプリングよりも複雑な理論に適用することへの期待である, 大標本理論を用いた近似の議論によって, 複雑な手法がシンプルな手法に近似され, 正当化に役立つことが期待される. たとえば Stochastic EM アルゴリズムや, 適合的モンテカルロ法に適用できるだろう.

第三章では, 繰り返しの数だけが大きくなる場合に, MH アルゴリズムの新しい近似を用いてアルゴリズムの評価と, より良いアルゴリズムの提案をする. ここで, MH アルゴリズムの説明のために, いくつかの記号を定義する.

まず, $(\mathbf{R}^d, \mathcal{B}(\mathbf{R}^d))$ 上の確率密度関数 $p(x)$ $(x \in \mathbf{R}^d)$ をもつ確率分布 $P(dx) = p(x)dx$ を考える. この P に関するある統計量, $\mu(P)$ の計算が目的であるとする. MH アルゴリズムは, あるシンプルなトランジションカーネル $Q.(\cdot)$ を少し変化させ, P を不変分布としてもつマルコフチェインの列 $x_1, \ldots, x_m$ を発生させる手法である. このような列が得られたら, その経験分布 P_m によって $\mu(P_m)$ として近似させる. 具体的には, $Q_x(dy) = q(x,y)dy$ なるカーネルを用いて, 次のようにマルコフチェイン $(M_n^x; n \in \mathbf{N}_0)$ を作る.

まず $M_0^x = x$ とする. 次に以下に従って順次数列を作る.

$$\left\{\begin{array}{lcl} Y_n^x & \sim & q(M_{n-1}^x, y)dy \ (n \in \mathbf{N}) \\ M_n^x & = & \left\{\begin{array}{ll} Y_n^x & \text{with probability } \alpha(M_{n-1}^x, Y_n^x) \\ M_{n-1}^x & \text{with probability } 1 - \alpha(M_{n-1}^x, Y_n^x). \end{array}\right. \end{array}\right.$$

ただし, $\alpha(x,y) = \min\{1, p(y)q(y,x)/(p(x)q(x,y))\}$ である.

この手法によってマルコフチェインは容易に作ることが出来るが, 問題は, $(M_n^x; n \in \mathbf{N}_0)$ の性質を調べることが比較的困難なことである. トランジションカーネル Q がシンプルであっても, $(M_n^x; n \in \mathbf{N}_0)$ は複雑なトランジションカーネルをもち, そのためとくに多次元での解析を複雑にしているのである.

ところが, メトロポリス-ヘイスティングスアルゴリズムの一手法の, ランジュバン法においては, t-分布のように $p(x)$ の裾が重い場合には, Q と, $(M_n^x; n \in \mathbf{N}_0)$ のカーネルの差が小さい. ここで, ランジュバン法とは, 次のような Q を用いる方法である.

$$Q(x, dy) \sim \mathrm{N}(x + \frac{1}{2}\langle h \nabla \log p(x) \rangle, h),$$

この単純なカーネル Q の解析に帰着させることで, 次のようなメリットがある. 第一に, 本質的仮定にしぼって, アルゴリズムの正当化が言えることである. 従来の多くのメトロポリス-ヘイス

ティングスアルゴリズムの解析においては, $(M_n^x; n \in \mathbf{N}_0)$ のカーネルの複雑さのために, やや強すぎる仮定をおいて, 収束性や収束レートの正当性を言うものが多い. しかしながら, この場合は Q の解析によって, このランジュバン法の解析に役立てることが出来て, 不必要な仮定が要らない. 第二に, 本論文ではランジュバン法を改善する新しい手法を提案するが, この正当化についても, さほど仮定を増やさずに議論が出来ることである.

実際の数値計算においては, いくつかの例で, 提案した新手法の優位性が確認されたが, パラメータの取り方によってはかえってうまく行かない例もある.

本論文の作成にあたっては, 修士博士通じて指導して頂いた東京大学数理科学研究科の吉田 朋広教授に感謝したい. 先生の有益なコメントや, 数々の問題点の指摘により本論文は大きく改善された. とくに, 独立同分布に限っていた本論文を, より一般的に拡張するようご指摘くださり, これによって論文全体がシンプルに, また一般的になった.

本研究は一部において学術振興会, および COE プログラムによってサポートされた.