メトロポリス的モンテカルロ法を利用したハイパーパラメータの推定（イジング模型の場合）

伊庭幸人

非ガウス的な大規模モデルへのABIC法ないしTYPE II最尤法の適用においては，多重積分（多重和)をどう計算するかが重要な問題である．これに対する有力なアプローチのひとつとして，ABIC～自由エネルギーという統計物理アナロジーにもとづいて，メトロポリス的モンテカルロ法を応用することが考えられる．これにはGemanらによる"微分型"の方法とOgataによる"積分型"の方法がある．

ここでは，イジング模型のパターン（snapshot）に雑音を加えたものから結合定数と雑音の強さを同時推定する問題を例として，"微分型"の方法を適用した結果を報告する．ここで，snapshotはメトロポリス的モンテカルロ法で生成したものを用いる（"答のわかっている問題"でのテスト）．

イジング模型とは，多数の離散変数 $\{x_i\}$ の上の次のような確率分布である：

$$x_i \in \{1, -1\} \tag{1}$$

$$\pi(\{x_i\}) = \frac{1}{Z_\pi} \exp\left(\frac{1}{2} J \sum_{j \in N(i), i} x_i x_j\right) \tag{2}$$

$$Z_\pi = \sum_{\text{config.}} \exp\left(\frac{1}{2} J \sum_{j \in N(i), i} x_i x_j\right) \tag{3}$$

$\sum_{\text{config.}}$ は $\{x_i\}$ の 2^M 通りの組合せに関する和，M はパラメータ（画素ないしスピン）$\{x_i\}$ の数である．一般に，$\{x_i\}$ は格子またはネットワーク上に配置され，$N(i)$ はその上での i の近傍を示す．ここでは正方格子上のイジング模型で，$N(i)$ が i の4個の最隣接点の場合を考える．この模型は（無限系の極限で）いわゆる相転移を示し，J のある値（臨界点）を境に性質が急に変化することが知られている．

このような離散モデルは連続モデルの極限として見たとき，強い非ガウス性を持っていることを注意しておく．たとえば，

$$P(\phi) = \frac{1}{Z_p} \exp\left(\phi^2 - \lambda(\phi^2 - 1)^2\right) \tag{4}$$

という分布を考えると，これは非ガウス性の強い極限 $\lambda \to \infty$ で離散変数 $(\phi = \pm 1)$ の分布と見なせる．

データ（snapshotに雑音を加えたもの）を $\{y_i\}$ とし，雑音"$1 \to -1$"と雑音"$-1 \to 1$"が等確率 p で起こるとする．この場合，尤度関数は

$$L(\{y_i\} | \{x_i\}) = \frac{\exp\left(h \sum_i y_i x_i\right)}{(2\cosh(h))^M} \tag{5}$$

$$h = \frac{1}{2} \log \frac{1-p}{p} \tag{6}$$

となり，ABICは，

$$\log \sum_{\text{config.}} \exp(JA_\pi + hA_L) - \log \sum_{\text{config.}} \exp(JA_\pi) - M \log(2\cosh(h)) \tag{7}$$

となる．ただし，

$$A_\pi = \frac{1}{2} \sum_{j \in N(i), i} x_i x_j \tag{8}$$

$$A_L = \sum_i y_i x_i \tag{9}$$

である．

(7)式の第1項は,データ $\{y_i\}$ に由来する非一様磁場中のイジング模型の自由エネルギーに,第2項は磁場なしのイジング模型の自由エネルギーにそれぞれ相当する.

ABICを極小化する微分方程式

$$c_J \frac{dJ}{dt} = -\frac{1}{2}\frac{\partial(\mathrm{ABIC})}{\partial J} \tag{10}$$

$$c_h \frac{dh}{dt} = -\frac{1}{2}\frac{\partial(\mathrm{ABIC})}{\partial h} \tag{11}$$

は,(7),(8),(9)式より,

$$c_J \frac{dJ}{dt} = \langle A_\pi \rangle_{\mathrm{pos}} - \langle A_\pi \rangle_\pi \tag{12}$$

$$c_h \frac{dh}{dt} = \langle A_L \rangle_{\mathrm{pos}} - M \tanh(h) \tag{13}$$

となる.ここで,

$\langle\ \rangle_\pi$ 事前分布での期待値

$\langle\ \rangle_{\mathrm{pos}}$ 事後分布での期待値

であり,t は仮想的な時間,c_J, c_h は適当な時定数である.

実際にこれを解くときには,仮想的な時間 t に関して差分化する.t はモンテカルロステップ(MCS)とは別物であって,理想的には,t を $t+dt$ に動かすたびに期待値が収束するまでモンテカルロ法を十分長く走らせることになる.また,事後分布での期待値の計算と事前分布での期待値の計算は別の simulation を要するので,2つの simulation を並行して走らせなくてはならない.

数値実験の結論としては,臨界点より弱結合(無秩序相)ではすべて順調にいくことがわかった.この例題では (h, J) の空間に多くの局所的極小があってそこにトラップされるなどということは起こらなかった.

イジング模型やそれに関連した分布の場合,texture(肌理)等への応用では無秩序相も意味があるが,画像処理の事前分布として応用するには,非ガウス性が強くあらわれる秩序相が重要である.以上述べたような方法が,秩序相でうまくいくかは不明であって,テストの方法も含めて今後に残された問題である.

推定方程式の不偏性

柳 本 武 美

1. 序

標本 $\boldsymbol{x}=(x_1,\ldots,x_n)$ から母数 θ を推定するときに,推定方程式 $g(\boldsymbol{x};\theta)=0$ の解で推定量を定義することが多い.尤度推定方程式の解として得られる最尤推定量はその1例である.推定方程式が不偏である,即ち,

$$E(g(\boldsymbol{x};\theta))=0$$

を要請することは自然である.

2. 例

実際の推定量では上の要請は必ずしも守られていない.

例1. $X\sim N(\mu,\theta)$ とすると,θ の最尤推定量は