混合分布モデルにおけるモンテカルロ法の非正則な振る舞い

東京大学 大学院数理科学研究科　　鎌谷研吾

序

本講演では，混合分布モデルについて，ギブスサンプラーの幾つかの非正則なふるまいを研究し，その対処法を考察する．[Kamatani(2009)] は正則なギブスサンプラーの漸近的な収束等の振る舞いを調べた．とくに正則なパラメトリック族については，ギブスサンプラーの振る舞いは AR 過程の分布に収束する事を示した．しかし，非正則なパラメトリック族では必ずしも AR 過程の分布には収束せず，以下の三つの場合がある．a) ギブスサンプラーの振る舞いはタイトネスを持ち，極限はエルゴード性を持つが，AR 過程ではない，b) タイトネスを持ち，極限はエルゴード性を持たない，c) タイトネスを持たない．強い識別条件のもとの混合分布モデルであっても，真値がパラメータ空間の境界上にある場合，ギブスサンプラーは正則なパラメトリック族と異なるふるまいをする事がある．識別条件の強さによって a) および b) の二つの場合が現れる．本講演ではこれらの収束，および b) の場合の対処法として，Metropolis-Hastings アルゴリズムを用いた方法について検討する．

1　ギブスサンプラーの空間と距離

本項目ではギブスサンプラーや Metropolis-Hastings algorithm の動く空間を考える．まず (S,d) は完備可分の距離空間で，$d(a,b)\le 1$ となるとする．可算個の S の直積を S^∞ と書き，その上での距離 d_∞ を

$$d_\infty(s,t)=\sum_{i=1}^{\infty}2^{-i}d(s(i),t(i)) \tag{1}$$

とさだめる．ただし $s=(s(1),s(2),\ldots)$ および $t=(t(1),t(2),\ldots)$ は S^∞ の元である．完備可分な距離空間 (S^∞,d_∞) 上の確率分布全体の集合を $\mathcal{P}(S^\infty)$ と書き，その上での Prohorov 距離を ρ と書く事にする．本論文で考える確率測度は，完備可分な距離空間 $(\mathcal{P}(S^\infty),\rho)$ の上の確率測度であり，その全体を $\mathcal{P}^*(S^\infty)$ と書き，またその上の Prohorov 距離を ρ^* とかく事にする．すると再び $(\mathcal{P}^*(S^\infty),\rho^*)$ 完備可分な距離空間である．ギブスサンプラーの振る舞いは $(\mathcal{P}^*(S^\infty),\rho^*)$ の元として表現できる．

Example 1.1 ギブスサンプラーのふるまいは $\mathcal{P}^*(\Theta^\infty)$ の元で表される．ただし，Θ はパラメータ空間である．ここで X および Y は状態空間で，X は観測の空間，Y はギブスサンプラーのための人工的な空間とする．$\Theta\times X\times Y$ 上の σ-有限な測度 p は分解

$$p(d\theta,dx,dy)=p_{X,Y}(dx,dy)p_{\Theta|X,Y}(d\theta|x,y)=p_{X,\Theta}(dx,d\theta)p_{Y|X,\Theta}(dy|x,\theta) \tag{2}$$

をもつとする．ただし，$p_{\Theta|X,Y}$ および $p_{Y|X,\Theta}$ は確率トランジションカーネルである．与えられた観測 $x\in X$ のもと，状態空間が $\Theta\times Y$ であるようなマルコフチェインが，ルール $y(i)\sim p_{Y|X,\Theta}(dy|x,\theta(i))$ および $\theta(i+1)\sim p_{\Theta|X,Y}(d\theta|x,y(i))$ かつ出発点 $\theta(1)\sim\nu_{\Theta|X}(d\theta|x)$ のもと生成されるとする．すると Θ へのマージナル $\theta(1),\theta(2),\ldots$ もやはりマルコフチェインで，その分布を，x に依存する事を明示して，$\omega(x)$ と書く．これは $\mathcal{P}(\Theta^\infty)$ の元である．さらに，$\omega(x)$ の法則は $\mathcal{P}^*(\Theta^\infty)$ の元である．

2 非正則なギブスサンプラーのふるまい

非正則なギブスサンプラーの一例を見る．まず $\Theta=[0,1], X=\mathbf{R}^{\mathbf{d}}, Y=\{0,1\}$ とし，$\mathcal{X}$ は X のボレル σ-代数で，$\mathcal{Y}=2^{\{0,1\}}$．また p_Θ は $\mathcal{P}(\Theta)$ の元であり，$F_0, F_1 \in \mathcal{P}(X)$ を固定すると $p_{X|Y,\Theta}(\cdot|y,\theta)=F_y\ (y=0,1)$ および $p_{Y|\Theta}(\{y\}|\theta)=(1-\theta)^{1-y}\theta^y$ であるものとする．独立同分布を考え，$p^n\in\mathcal{P}(\Theta\times X_n\times Y_n)$ を次のように定めるが，ここで X_n および Y_n は X および Y の n 回コピーである．

$$p^n(d\theta,dx_n,dy_n)=p^n_{X_n|Y_n,\Theta}(dx_n|y_n,\theta)p^n_{Y_n|\Theta}(dy_n|\theta)p_\Theta(d\theta) \tag{3}$$

ただし $x_n=(x^1,\ldots,x^n)$，$y_n=(y^1,\ldots,y^n)$ であり，$p^n_{X_n|Y_n,\Theta}(dx_n|y_n,\theta)=\prod_{i=1}^n p_{X|Y,\Theta}(dx^i|y^i,\theta)$, $p^n_{Y_n|\Theta}(dy^n|\theta)=\prod_{i=1}^n p_{Y|\Theta}(dy^i|\theta)$ である．

ここで

$$I:=\int (f_1(x)/f_0(x)-1)^2F_0(dx), \tag{4}$$

とする．

Proposition 2.1 真値が $\theta=0$ のとき，$p^n\in\mathcal{P}(\Theta\times X_n\times Y_n)$ から生成されたギブスサンプラーは，以下の条件のもと，一点にとどまり，動かないギブスサンプラーに収束する．

1. 強い識別条件を持つ: $\alpha=F_1((\mathrm{suppF_0})^{\mathrm{c}})=0$ かつ $I\in(0,\infty)$.
2. 事前分布 p_Θ は正則である．

3 Metropolis-Hastings algorithm

前節の p^n および人工的に作られた $\overline{p}^n_{\Theta|X_n}$ により生成される Metropolis-Hastings アルゴリズムについて，一般に次の命題が成り立つ．ここで，最尤推定量 $\hat{\theta}_n$ にたいして，$\varphi^n_S(\theta,x_n)=n^{1/2}(\theta-\hat{\theta}_n)$ とし，$p^n_{\Theta|X_n}(\cdot|x_n)$ および $\overline{p}^n_{\Theta|X_n}(\cdot|x_n)$ による φ^n_S の分布を $p^n_{S|X_n}(\cdot|x_n)$ および $\overline{p}^n_{S|X_n}(\cdot|x_n)$ と書く．また $q\in\mathcal{P}(\mathcal{S}\times\mathcal{T})$ および $\overline{q}\in\mathcal{P}(\mathcal{S}\times\mathcal{T})$ は disintegration を持つとする．

Proposition 3.1 ある可測関数の列 $\varphi^n_T: X_n\to T$ があって以下を満たすとき，*Metropolis-Hastings* アルゴリズムのふるまいは，ある *Metropolis-Hasings* アルゴリズムのふるまいに収束し，エルゴード性を持つ．

1. 以下の収束が成立する
$$\|p^n_{S|X_n}(\cdot|x_n)-q_{S|T}(\cdot|\varphi^n_T(x_n))\|=o_{p^n_{X_n}}(1),\ \|\overline{p}^n_{S|X_n}(\cdot|x_n)-\overline{q}_{S|T}(\cdot|\varphi^n_T(x_n))\|=o_{p^n_{X_n}}(1). \tag{5}$$
2. $\varphi^n_T(x_n)$ の $p^n_{X_n}$ による分布は r_T へ収束する．
3. $q_{S|T}(\cdot|t)$ および $\overline{q}_{S|T}(\cdot|t)$ は t について全変動ノルムに関して連続．

上の命題より，ギブスサンプラーが degenerate であっても，うまく $\overline{p}^n_{\Theta|X_n}$ を取れば良い収束をする事がわかる．本講演では適当な $\overline{p}^n_{\Theta|X_n}$ の取り方について検討する．

参考文献

[Kamatani(2009)] Kengo Kamatani. On Some Asymptotic Properties of the Gibbs Sampler (submitted). 2009.