The 32nd Annual Conference of the Japanese Society for Artificial Intelligence, 2018

時系列生成モデルのためのリサンプリング学習

Learning to resample for time-series generative model

金子 貴輝 *1　大澤 昇平 *2　松尾 豊 *3
Takaaki Kaneko　Shohei Ohsawa　Yutaka Matsuo

*1*2*3東京大学大学院 工学系研究科 技術経営戦略学専攻
Department of Technology Management for Innovation, The University of Tokyo

Sequential Monte Carlo (SMC) is a typical sampling method that can be sampled in order from a sequential probabilistic model. However, due to degeneration of the sample, SMC may produce samples with low likelihood with a small number of particles. In our study, we focus on the fact that the same resampling targets of SMC for each sample cause degenerating samples. We want to relax this constraint, but analytically deriving asymmetric sequential resampling targets is difficult. Therefore, we expand resampling strategy of SMC asymmetrically by learning the sequential resampling target from the target of the whole series approximated to the lower bound. By this, by learning to resample, it is expected that accurate estimation of latent variable can be realized with the same particle number as SMC.

1. はじめに

深層学習において，観測変数 x の背後の依存関係を表現するために，潜在変数 z で同時分布 $p(x,z)$ をモデル化すると，推論分布 $p(z|x)$ の扱い方がモデルの学習のために重要になる．例えば変分オートエンコーダ（VAE）[Kingma 13] では，モデルの学習基準である周辺尤度を表現するために，この分布を必要とする．そのためエンコーダで定義される別の分布に置き換えて基準を近似し，KL divergence 最小化を用いてその分布を推論分布に近づける．

特にマルコフ性を持つ時系列の潜在変数における推論分布 $p(z_{1:T}|x_{1:T})$ は，その扱い方が時間の長さ T にスケールするかが問われる基本的な例であり，汎用なアルゴリズムを目指す上で重要である．潜在変数の状態が離散的で計算機で扱えないほど多くなければ，隠れマルコフモデルのような解析的な扱いが可能な場合もある．しかし動画や音声情報を潜在変数に含める場合，密な表現として出現する状態の数は特徴の数に対して指数的に多くなりうる．この場合，解析的に扱うことも直接全体をエンコーダで置き換えることも，時間の長さにスケールしない．このような推論分布の深層学習による近似は研究途上の課題である．

時系列の推論分布に関する問題として，各時刻の潜在変数を推定する逐次推定の問題（2 章）があり，従来のアプローチとして代表的なものに Sampling Importance Resampling（SIR）の逐次適用を行う，逐次モンテカルロ（SMC）[Gordon 93] がある．SIR とは，粒子と呼ばれる確率変数の値の候補を複数持ち，途中までの尤度を使って粒子を選び直す手法である．SMC は将来的に得られるサンプルの尤度が予測できない問題を SIR で解決する手法であり，粒子フィルタとも呼ばれる．SMC によって時系列の推論分布の近似的なサンプルが得られるので，期待値をモンテカルロ近似することができる．3 章で述べるように，SMC の提案分布にニューラルネットを用いて推論分布を近似する研究も出てきている．

しかし SMC では途中までの分布での尤度が低い粒子は捨てられやすく，後から尤度が大きくなる粒子を本来より低い確率でサンプリングしてしまう．これは縮退問題として知られている．特に明らかな問題だと思われるのは，同程度の尤度の粒子を選び直すことで一部無くしてしまう傾向である．これは図 1（左）で示されるように，2 粒子 2 状態の単純な例（5.2 章参照）で顕著である．これは潜在変数が連続値でも起こりうる．

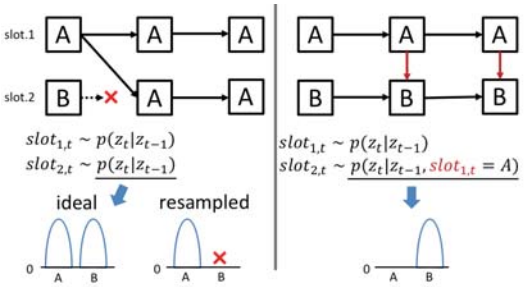

図 1: SMC と提案手法のリサンプリング方式の比較．観測変数は省略している．黒矢印はリサンプリングの流れを，アルファベットは潜在変数の値を表す．(左）SMC でのリサンプリング．独立同分布を用いてサンプルしている．(右）提案手法でのリサンプリング．リサンプリング結果を順に依存させて次のリサンプリングを行うことができる．

本研究では縮退問題のうち，この同程度の粒子をなくす問題に注目する．原因の一つは，リサンプリング目標がどの粒子でも同じであること，つまりリサンプリング目標の対称性である．多様なサンプルを保持できるように，本研究では図 1（右）のように，複数の粒子のリサンプリング目標を非対称にする．

最適な非対称リサンプリング目標は，未来の変数についての期待値の式になるため，一般には解析的に計算できない．学習によって逐次目標を獲得する方針は，計算時間などの有用性が乏しいため注目されてこなかった．しかし潜在変数系列の推論分布に対する新たなアプローチとして重要なアルゴリズムであると考えたため，本研究では逐次目標を学習によって獲得する新しい逐次リサンプリング手法を検討する．

提案手法では，非対称なリサンプリング目標を実現するため，リサンプリング重みに学習パラメータを追加して拡張する．これは従来の SIR のリサンプリング重みに，過去のサンプルを引数に取る新たな重みを掛け合わせることで実現できる．

さらに拡張した重みが縮退を軽減させる適切な値になるよ

連絡先: kaneko@weblab.t.u-tokyo.ac.jp

うに，最終目標 $p(z_{1:T}|x_{1:T})$ との KL 誤差の上界を学習基準として，勾配法を用いて重みを学習する．本研究では，この学習とそれによって獲得された非対称なリサンプリング目標によるリサンプリングのアルゴリズム（4 章）を提案手法とする．学習基準は上界であるため，必ず良い方向に学習が進むとは限らないのは問題であるが，5.1 章で示すように，過去の情報の解釈が未来の尤度に大きく寄与するのに，長期に渡って複数の解釈が尤もらしい状況を簡略化した問題においては，SMC よりも提案手法が正確なサンプリングを目指す学習基準になることを保証できる．

時系列確率モデルでの縮退の軽減を検証するため，本研究では前述の問題設定で実験（5 章）を行い，学習基準と正答率が連動し，SMC より提案手法が正確なリサンプリングが行えることを示した．

2. 問題設定

時系列マルコフモデルの同時分布を以下の式で定義する．

$$p(z_{1:T}, x_{1:T}) = \prod_{t=1}^{T} p(z_t, x_t|z_{t-1}),$$
$$p(z_1, x_1|z_0) = p(z_1, x_1)$$

このモデルからサンプルされた系列 $(x_{1:T}, z_{1:T})$ の $x_{1:T}$ から $z_{1:T}$ を推定し，その誤差の期待値を最小化する問題を考える．

$$\mathbb{E}_{p(z_{1:T}, x_{1:T})}\left[L(z_{1:T}, \hat{z}_{1:T})\right]$$

ただし L は適当な損失関数，$\hat{z}_{1:T}$ は $x_{1:T}$ を用いた推定値とする．

3. 関連研究

2 章の問題には，SMC を基本とした様々なアルゴリズムが提案されてきたが，近年，時系列生成モデルの学習の文脈から，深層学習と SMC を組み合わせる新しい手法が提案されている．SMC とこの新しい手法の二つについて説明する．

3.1 SMC

SMC は時刻 $t = 1$ から $t = T$ まで以下の式で表されるサンプリングを繰り返すアルゴリズムである．

$$\begin{aligned}
&propose && z_t^j \sim r(z_t|z_{1:t-1}^{a_{t-1}^j}),\\
&append && z_{1:t}^j = \{z_{1:t-1}^{a_{t-1}^j}, z_t^j\},\\
&reweight && w_t^j = \frac{p(z_t^j, x_t|z_{t-1}^{a_{t-1}^j})}{r(z_t^j|z_{1:t-1}^{a_{t-1}^j})},\\
&resample && a_t^i \sim \mathrm{Categorical}(w_t^i / \sum_l w_t^l)
\end{aligned}$$

$z_{1:t-1}^{a_{t-1}^j}$ は $t-1$ での j 番目の粒子であり，提案分布 r によって t での潜在変数の値 z_t^j が加わり $z_{1:t}^j$ となる．これを SIR のリサンプリング重み w_t^j にしたがってリサンプリングを行い t での i 番目の粒子 $z_{1:t}^{a_t^i}$ を得る．

SMC では，粒子間でリサンプリング重みを比較し a_t^i をリサンプリングすることで，尤度 $p(z_{1:t}, x_{1:t})$ を正規化して推論分布 $p(z_{1:t}|x_{1:t})$ を計算することなく，近似的なサンプリングを可能としている．しかし全体での目標分布は $p(z_{1:T}|x_{1:T})$ であるため，このリサンプリングが最適であるとは限らない．

3.2 Variational SMC

Variational SMC（VSMC）[Naesseth 17] は VRNN[Chung 15] のような時系列生成モデルを拡張して，SMC と組み合わせたエンコーダとそれに対応した学習基準からなる．SMC における提案分布をニューラルネットで定義できるので，長期的な依存関係を学習することができ，エンコーダで推論分布をさらにうまく近似できるようになった．

本研究の式導出はこれらの研究を参考にしているが，学習させるパラメータは異なる．これらの研究ではエンコーダのパラメータを学習することが目標であり，リサンプリング重みにパラメータは置いていない．

4. 非対称逐次リサンプリング

SMC を非対称なリサンプリングに拡張した提案手法について説明する．まず重みの定義とリサンプリングアルゴリズムを定義したのち，パラメータを学習するための同時分布と学習目標の定義，学習のための勾配法の利用について説明する．

4.1 非対称逐次リサンプリングの定義

時刻 t での，SMC でのリサンプリングを以下の拡張したリサンプリング重みの式で置き換える．

$$\begin{aligned}
&reweight && w_t^{i,j} = \frac{p(z_t^j, x_t|z_{t-1}^{a_{t-1}^j})}{r(z_t^j|z_{1:t-1}^{a_{t-1}^j})} \cdot w_{t,i}^{\Delta}(z_t^j; z_{1:t-1}^{a_{t-1}^j}, z_t^{a_t^{<i}}),\\
&resample && a_t^i \sim w_t^{i,\cdot} / \sum_l w_t^{i,l}
\end{aligned}$$

拡張重み $w_{t,i}^{\Delta}$ は多段のニューラルネットを用いてモデル化する．$z_t^{a_t^{<i}}$ は i 番目より前のリサンプリングされた粒子の組 $\{z_t^{a_t^k}|0 \le k < i\}$ を表す．リサンプリングの計算順序は図 2 のグラフで示すようになる．

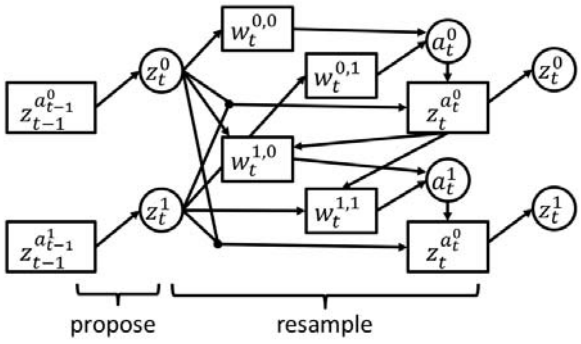

図 2: 提案手法の各時間ステップの計算グラフ．円形のノードが確率変数を，正方形のノードが決定論的な変数を表す．SMC と同様に提案分布（propose）からのサンプリングの後，拡張したリサンプリング重み $w^{i,*}$ と選ばれる粒子 z^{a^i} を粒子のスロット i ごとに繰り返し計算してリサンプリングを行う（resample）．

4.2 同時分布

棄却される粒子も含めた全ての同時分布 ϕ は以下の式で表される．

$$\phi(z, b, a) = \left[\prod_i r(z_1^i)\right] \prod_{t=2}^{T} \prod_i \left[\frac{w_{t-1}^{i,a_{t-1}^i}}{\sum_l w_{t-1}^{a_{t-1}^i}} r(z_t^i|z_{t-1}^{a_{t-1}^i})\right] \left[\frac{w_T^{1,b_T}}{\sum_l w_T^{1,l}}\right]$$

ここで $z = \{z_t^i | t = 1 \dots T, i = 1 \dots N\}, a = \{a_t^i | t = 1 \dots T, i = 1 \dots N\}, b = b_T$ とした．これを周辺化することで提案手法のサンプリング分布が表される．

$$q(z_{1:T}) = q(z_{1:T}^{b_{1:T}}) = \sum_{b_{1:T}} \sum_{a_{1:T}^{\neg b_{1:T}}} \sum_{z_{1:T}^{\neg b_{1:T}}} \phi(z, b, a)$$

ただし $b_t = a_t^{b_{t+1}}, z_{1:T}^{\neg b_{1:T}} = \{z_t^{\neg b_{t+1}} | 1 \leq t \leq T\}, z_t^{\neg b_{t+1}} = \{z_t^i | i \neq b_{t+1}\}, a_{1:T}^{\neg b_{1:T}} = \{a_t^{\neg b_{t+1}} | 1 \leq t \leq T\}, a_t^{\neg b_{t+1}} = \{a_t^i | i \neq b_{t+1}\}$ とする．

4.3 学習基準の導出

目標分布とサンプリング分布で定義される次の KL divergence を最小化することを考える．

$$\mathbb{E}_{p(x_{1:T})}\left[D_{\mathrm{KL}}(q \| p)\right] = \mathbb{E}_{p(x_{1:T})q(z_{1:T}|x_{1:T})}\left[\log \frac{q(z_{1:T}|x_{1:T})}{p(z_{1:T}|x_{1:T})}\right]$$

サンプリング分布に含まれる周辺化計算が解析的に行えないため，以下でイェンセンの不等式を用いた学習基準の近似を行う．

KL divergence の最小化は VAE のエビデンス下界 [Kingma 13] の最大化と一致する．

$$\begin{aligned}
&- \mathbb{E}_{p(x_{1:T})}\left[D_{\mathrm{KL}}(q \| p)\right] \\
&= \mathbb{E}_{p(x_{1:T})q(z_{1:T}|x_{1:T})}\left[\log \frac{p(z_{1:T}, x_{1:T})}{q(z_{1:T}|x_{1:T})}\right] - \mathbb{E}_{p(x_{1:T})}\left[\log p(x_{1:T})\right]
\end{aligned}$$

この式のサンプリング分布の対数尤度を計算し，真の対数尤度と残りに分解する．

$$\begin{aligned}
\phi(z, b, a) =& \prod_{t=1}^{T}\left[\frac{w_t^{b_{t+1}, b_t}}{\sum_l w_t^{b_{t+1}, l}}\right] \prod_{i \neq b_1} r(z_1^i) \\
& \cdot \prod_{t=2}^{T} \prod_{i \neq b_t} \left[\frac{w_{t-1}^{i, a_{t-1}^i}}{\sum_l w_{t-1}^{i,l}}\right] \prod_{t=1}^{T}\left[r(z_t^{b_t} | z_{t-1}^{b_{t-1}})\right] \\
=& \prod_{t=1}^{T}\left[\frac{w_t^{b_{t+1}, b_t}}{\sum_l w_t^{b_{t+1}, l}} r(z_t^{b_t} | z_{t-1}^{b_{t-1}})\right] \cdot \phi(x^{\neg b}, a^{\neg b} | x^b, b)
\end{aligned}$$

ただし，以下のように置いた．

$$\phi^{\neg b} = \phi(z^{\neg b}, a^{\neg b} | z^b, b) = \prod_{i \neq b_1} r(z_1^i) \prod_{t=2}^{T} \prod_{i \neq b_t} \left[\frac{w_{t-1}^{i, a_{t-1}^i}}{\sum_l w_{t-1}^{i,l}}\right]$$

SIR のリサンプリング重みから目標分布の尤度を取り出して，次のように式変形する．

$$\begin{aligned}
&\phi(z, b, a) \\
&= \prod_t \left[\tilde{p}(z_t^{b_t}, z_{1:t-1}^{b_{t-1}})\right] \cdot \prod_t \left[\frac{w_{t, b_{t+1}}^{\Delta}(z_t^{b_t}; x_t, z_{t-1}^{b_{t-1}}, z_t^{a_t^{<b_{t+1}}})}{\sum_l w_t^{b_{t+1}, l}}\right] \\
&\quad \cdot \phi(x^{\neg b}, a^{\neg b} | x^b, b) \\
&= p(z^b, x) \cdot W \cdot \phi(z^{\neg b}, a^{\neg b} | z^b, b)
\end{aligned}$$

ただし，以下のように置いた．

$$W = \prod_t \left[\frac{w_{t, b_{t+1}}^{\Delta}(z_t^{b_t}; x_t, z_{t-1}^{b_{t-1}}, z_t^{a_t^{<b_{t+1}}})}{\sum_l w_t^{b_{t+1}, l}}\right]$$

すると $q(z_{1:T}) = p(z) \sum_{b_{1:T}} \mathbb{E}_{\phi^{\neg b}}\left[W\right]$ となるので，真の対数尤度がキャンセルする．

$$\begin{aligned}
&\mathbb{E}_q\left[\log p - \log q\right] \\
&= \sum_z q(z)(\log p - \log p - \log \sum_b \mathbb{E}_{\phi^{\neg b}}\left[W\right]) \\
&= - \sum_z p(z) \sum_b \mathbb{E}_{\phi^{\neg b}}\left[W\right] \log \sum_b \mathbb{E}_{\phi^{\neg b}}\left[W\right]
\end{aligned}$$

次に期待値と対数尤度に使われているサンプリング分布の b についての周辺化を人工的な一様分布 $p(b_{1:T}) = \frac{1}{N^T}$ で置き換え，式に含まれる凸関数 $f(x) = -x \log(x)$ に対してイェンセンの不等式を適用する．

$$\begin{aligned}
&\mathbb{E}_q\left[\log p - \log q\right] \\
&= - \sum_z p(z) \sum_b \frac{1}{N^T} \mathbb{E}_{\phi^{\neg b}}\left[N^T W\right] \log \sum_b \frac{1}{N^T} \mathbb{E}_{\phi^{\neg b}}\left[N^T W\right] \\
&\geq \mathbb{E}_{p(z)}\left[\mathbb{E}_{p(b)}\left[\mathbb{E}_{\phi^{\neg b}}\left[N^T W \log(N^T W)^{-1}\right]\right]\right] \\
&= \sum_z \sum_b \sum_{a^{\neg b}, z^{\neg b}} p(z) W \phi^{\neg b} \cdot \left[\log(N^T W)^{-1}\right]
\end{aligned}$$

すると得られる下界はサンプリングの同時分布の期待値と，計算可能なその中身によって成り立っているため，モンテカルロ近似で算出可能な学習目標となる．

$$L(\lambda) = \mathbb{E}_\phi\left[\log(\frac{1}{N} W^{-1})\right]$$

この学習目標は重みが対称なとき VSMC の学習基準と一致する．ただし VSMC ではこの学習基準は提案分布の学習に用いられるのに対し，本手法では SMC を拡張する重みの学習に用いられている．またこの学習目標は下界であるため，この最大化は KL divergence を最小化することを保証しないが，長期的な依存関係が重視される場合においては，下界を高めることが KL divergence を下げることに繋がることが言える．(5.3 章で後述)

4.4 勾配法

分布にパラメータを含むため，勾配を計算すると方策勾配の項が導出される．VSMC では，この項は高分散であるとして実装上は無視されていた．しかし今回の導出ではリサンプリング目標を学習する上で，離散な確率変数 a_t^i が重要になるため，項を無視しなかった．

5. 実験

提案手法を調べるため，実世界の一部の問題を簡略化したトイプロブレムを構成する．

この構成とそれが簡略化していると考えたものについてまず説明し，次に SMC を使用すると起きる問題について説明する．このとき提案手法の学習基準はより良いパラメータを測ることができることを確認し，実際の実験設定について説明する．得られた結果から，提案手法がトイプロブレムで SMC より正確に潜在変数を推定できることを主張する．

5.1 保留期間の長いトイプロブレム

潜在変数も観測変数も 2 状態しかない時刻 $1 \leq t \leq T$ の隠れマルコフモデルを考える．$t = 1$ で潜在変数は一様な確率で 2 状態を取り，時刻 T まで観測変数は潜在変数と独立して決まり，潜在変数は確率 ε_τ で別状態に変化して遷移し，時刻 T で確率 $1 - \varepsilon$ で潜在変数の状態に対応する観測変数が得られる生成モデルを考える．このモデルでの時刻 T までを保留期間とし，時刻 1 で保留された 2 状態を T まで縮退せずに保持できるかを見ることができる．

保留されるべき 2 状態は，動画や音声情報などを潜在変数に符号化した場合の扱いきれない状態数のうち，ある時刻で同程度の尤度を持つものの簡略表現とした．また，確率 ε_τ と ε が小さいほど潜在変数は確定的に遷移して T で観測されることは，同程度の尤度を持つ候補の情報が T ステップ後の状態推定に重要な意味を持つ状況の表現とした．最後に時刻 T まで潜在変数と観測変数が独立していることは，保留期間中に候補が同程度の尤度を持つ状況が変化しない状況の表現とした．

5.2 SMC による縮退問題

SMC を用いてこの 2 状態を保持した粒子をリサンプリングすると，各リサンプリングごとに 1/2 の確率で片方を取りこぼす．これはサンプルごとに対称なリサンプリング目標を使う限り避けられない．取りこぼした片方が時刻 T に観測された場合，ε_τ と ε が小さいほど，尤度の低いサンプルとなる．

5.3 下界での評価

このトイプロブレムでは，ε を限りなく 0 に近づければ，誤ったサンプルに対する対数尤度は $-\infty$ に発散する．すると提案手法の学習基準は，非対称なリサンプリングが最後が誤ったサンプルを生成する確率を低くする限り，下界を限りなく上昇させる．そしてそのような非対称なリサンプリングを実現するパラメータは拡張重み w^Δ に十分な自由度があれば，存在することがわかる．よって，学習過程でそのようなパラメータにたどり着ける限り，下界評価によって学習の進捗を評価できる．

5.4 実験設定

このトイプロブレムを使って提案手法が長期的なサンプリング目標に備えて SMC より縮退を減少させられることを確認した．状態数が多く，取りこぼした粒子の情報を回復できない状況を表すため，トイプロブレムの遷移分布 $p(z_t|z_{t-1})$ を提案分布として固定した．トイプロブレムのパラメータは $T = 4, \varepsilon_\tau = \varepsilon = 0.0005$ とした．保留期間中は観測変数から潜在変数を推定できないため，時刻 T での真の潜在変数と推定した潜在変数を比較し，正答率を評価基準とした．比較対象として毎回リサンプリングを行う SMC と，時刻 T 以外はリサンプリングを行わない SMC を比較対象とした．

モデルのパラメータは，拡張重みのネットワークは 2 層で 1 層目の活性化関数は ELU，2 層目は線形のままにした．2 層目のユニット数は 2 とした．勾配法には Adam を使用した．バッチサイズは 32，学習率は 0.001 とした．粒子数=2 とした．

5.5 実験結果

学習ステップの間の正答率の変化をグラフにすると図 3 のようになった．初期のパラメータでは拡張重み $w^\Delta \approx 1$ なので SMC と同じ正答率だが，学習を進めるにつれてリサンプリングしない場合の正答率に近づいた．このトイプロブレムにおいて提案手法が非対称なリサンプリング目標を学習によって獲得し，SMC よりも正確に潜在変数を推定できることを示した．

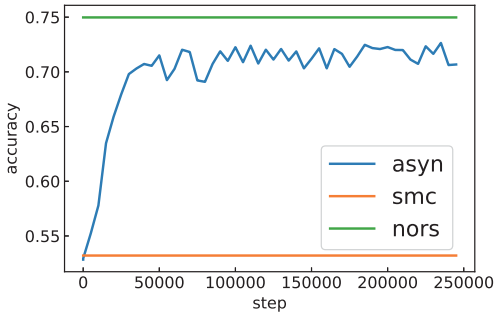

図 3: 学習に伴う最終時間でのサンプル状態の正答率の変化．橙線は SMC，緑線は最終時間のみリサンプリングした場合，青線が提案手法，のそれぞれの正答率．縮退を抑えるよう学習できている．

6. まとめ

本研究では，SMC において粒子を縮退させる原因としてリサンプリングの目標の対称性に注目し．下界によってこのリサンプリング目標を近似して学習することで，SMC のリサンプリング戦略を非対称に拡張し，トイプロブレムでの縮退の軽減を検証した．これにより SMC と同じ粒子数で，正確な潜在変数の推定を実現できる場合が存在することが示された．

一方，今後の課題として，縮退の軽減によるサンプル効率が計算効率に繋がるかの解明が挙げられる．非対称なサンプリングの内，リサンプリング結果に依存する部分は並列化できないため，計算効率が上がるとは限らないからである．今回は明示的に非対称性を学習するためにリサンプリング結果を引数としたが，これは非対称性からの必要条件かは不明である．これについては今後検証していきたい．

謝辞

本研究は JSPS 科研費 JP15H05327,JP16H06562 の助成を受けたものです。

参考文献

[Chung 15] Chung, J., Kastner, K., Dinh, L., Goel, K., Courville, A. C., and Bengio, Y.: A Recurrent Latent Variable Model for Sequential Data, in Cortes, C., Lawrence, N. D., Lee, D. D., Sugiyama, M., and Garnett, R. eds., *Advances in Neural Information Processing Systems 28*, pp. 2980–2988, Curran Associates, Inc. (2015)

[Gordon 93] Gordon, N. J., Salmond, D. J., and Smith, A. F. M.: Novel approach to nonlinear/non-Gaussian Bayesian state estimation, *IEE Proceedings F - Radar and Signal Processing*, Vol. 140, No. 2, pp. 107–113 (1993)

[Kingma 13] Kingma, D. P. and Welling, M.: Auto-encoding variational bayes, *arXiv preprint arXiv:1312.6114* (2013)

[Naesseth 17] Naesseth, C. A., Linderman, S. W., Ranganath, R., and Blei, D. M.: Variational Sequential Monte Carlo, *ArXiv e-prints* (2017)