

変分推論・非線形フィルタリングを駆使した 時系列データの潜在モデルの推論・予測

石曽根毅[†] 樋口知之[‡] 中村和幸^{†§}

明治大学[†] 中央大学[‡] JST さきがけ[§]

1 はじめに

時系列データが従うモデルをその時系列データのみから推論することは、モデル化が困難な対象における予測・制御等につながるため極めて重要な意義を持つ。この目的を達成するために、近年、Sequential VAE (SVAE) と粒子フィルタを活用した手法である FIVO[1, 2, 3] 及びその発展手法 [4, 5] (FIVOs) が注目を集めている。しかし、FIVOs は粒子の退化による学習効率の低下・学習コストの増大といった問題を抱えている。

本研究では、これらを解決する新しい潜在モデルの推論手法として、SVAE と Ensemble Kalman Filter (EnKF) を組み合わせた手法である Ensemble Kalman Variational Objective (EnKO) を提案する。提案手法は、人工データ・実データに対し、既存手法を上回る予測精度を達成し、解釈可能性のある結果が提供できたことを報告する。

2 Sequential VAE (SVAE)

VAE は、観測データ x が潜在変数 z から生成される過程をニューラルネットワークを用いて学習する手法である [6]。VAE の潜在変数に再帰構造を組み込んだ時系列データの生成過程を学習するモデルのことを SVAE と呼び、図1のような構造となる。SVAE の代表的な手法としては、VRNN[7], SRNN[8] があり、音楽・文字・生理学データなどへの適用が研究されている。

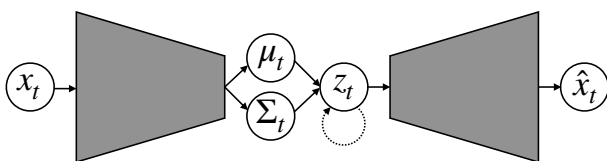


図 1: SVAE の構造。灰色の符号化器・復号化器、破線の再帰部はニューラルネットワークで表現される

3 Ensemble Kalman Filter (EnKF)

SVAE において、潜在変数 $\{z_t\}_{t=1}^T$ から観測変数 $\{x_t\}_{t=1}^T$ を生成するモデルは、潜在変数の遷移モデル $z_t \sim p(z_t|z_{t-1})$ と観測変数の出力モデル $x_t \sim p(x_t|z_t)$ で表すことができる。このモデルは、時系列解析で用いられる状態空間モデルの形式であり、状態空間モデルに対するフィルタリング手法を適用することができる。フィルタリングとは、条件付き分布 $p(z_t|x_{1:t-1})$ に従う z_t を時刻 t における観測 x_t を取り入れた分布 $p(z_t|x_{1:t})$ に従うように更新することであり、粒子フィルタ (PF) や EnKF が代表的な手法である。

PF や EnKF では、複数の粒子によるアンサンブルで条件付き分布を表現する。PF では、各粒子の観測データに対する当てはまり度合い (尤度) に従って粒子を複製する一方、EnKF では、全粒子を観測データをよく説明するように移動させる (図2)。PF は、同じ粒子が原理的に複製されてしまうことにより、分布を適切に表現できなくなる粒子の退化という問題を抱えているが、EnKF ではこの問題が生じ得ない。そのため、粒子の多様性を一定程度確実に担保するという観点からは、EnKF の方が適したフィルタリング手法である。

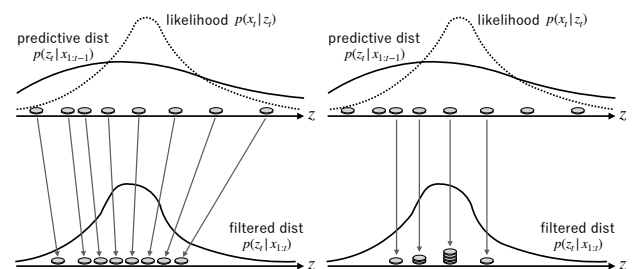


図 2: EnKF(左), PF(右) の更新方法

4 提案手法

SVAE に潜在変数のフィルタリングステップとして EnKF を取り入れた手法として Ensemble Kalman Variational Objective (EnKO) を提案する。既存手法との比較を表 1 に示す。PF を取り入れた FIVOs や

Latent Model Inference and Prediction based on Variational Inference and Nonlinear Filtering
Tsuyoshi Ishizone[†], Tomoyuki Hicuhi[‡], Kazuyuki Nakamura^{†§}
Meiji University[†], Chuo University[‡], JST, PRESTO[§]

表 1: 手法間の比較

手法	粒子効率	勾配推定量の不偏性
IWAE[9]	×	○
FIVO[1]	×	×
PSVO[5]	△	×
EnKO	○	△(2次モーメントまで○)

IWAE[9]では粒子の多様性が低いのに対し、EnKOでは粒子の多様性を確保することができ、結果として学習コストの削減・学習効率の増大に繋がる。また、最大化する目的関数として

$$\mathcal{L}_{\text{EnKO}} := \mathbb{E}_q \left[\log \left(\prod_{t=1}^T \frac{1}{N} \sum_{i=1}^N \frac{p(z_t^{(i)} | z_{1:t-1}^{(i)}) p(x_t | z_t^{(i)})}{q(z_t^{(i)} | x_{1:T}, z_{1:t-1}^{(i)})} \right) \right] \quad (1)$$

を採用することで、周辺尤度・勾配推定量の2次モーメントまでの不偏性を理論的に保証できる。ここで、 q はVAEの符号化器に相当する変分事後分布、 N は粒子数、添字の i は粒子のインデックスを表す。

5 数値実験

数理モデルに従って人工的に作成したデータとマウスの一次視覚野の細胞で観測された細胞内電圧データ[10]に対して、EnKO, FIVO, IWAEを適用した。

数理モデルとして2変数系であるFitzHugh-Nagumo (FHN) モデル

$$\dot{V} = V - \frac{V^3}{3} - W + I_{\text{ext}}, \quad (2)$$

$$\dot{W} = a(bV + d - cW), \quad (3)$$

を用い、 $x_t \sim N(V_t, 0.1^2)$ として1次元の観測データを300個(学習用200, 検証用40, 評価用60)作成した。細胞内電圧データは、先行研究[5]と同様な前処理を通して得られる40データを、学習用30, 評価用10に分割して実験を行った。

学習したモデルに対し、評価用データで予測 R^2

$$\text{MSE}_k = \sum_{t=1}^{T-k} (x_{t+k} - \hat{x}_{t+k})^2, \quad (4)$$

$$R_k^2 = 1 - \frac{\text{MSE}_k}{\sum_{t=1}^{T-k} (x_{t+k} - \bar{x}_k)^2}, \quad (5)$$

を計算した結果が図3である。ここで、 \hat{x}_{t+k} は時刻 t からの k 期先予測値($1 \leq k \leq 20$)、 \bar{x}_k は $\{x_t\}_{t=k+1}^T$ の平均値である。図より、EnKOは既存手法を予測精度で上回っていることが分かる。

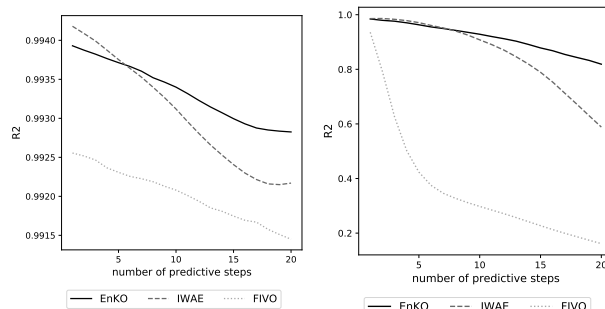


図 3: 予測 R^2 の結果 (左: FHN 系, 右: 電圧データ)

6 まとめ

本研究では、SVAEとEnKFを組み合わせた、時系列データの潜在的な生成モデルを推論する手法であるEnKOを提案した。提案手法は、先行研究であるFIVOsに見られた粒子の退化という問題点を解決し、実験的にも優れた予測精度を示したことから、時系列データの潜在変数モデルの推論における新しい標準的な手法となることが期待される。手法や実験の詳細については、当日報告する。

参考文献

- [1] C. J. Maddison, D. Lawson, G. Tucker, N. Heess, M. Norouzi, A. Mnih, A. Doucet, and Y. W. Teh. “Filtering Variational Objectives”, in *NeurIPS*, 2017.
- [2] C. A. Naesseth, S. W. Linderman, R. Ranganath, and D. M. Blei. “Variational Sequential Monte Carlo”, in *AISTATS*, pp. 968-977, 2018.
- [3] T. A. Le, M. Igl, T. Rainforth, T. Jin, and F. Wood. “Auto-encoding Sequential Monte Carlo”, in *ICLR*, 2018.
- [4] A. K. Moretti, Z. Wang, L. Wu, and I. Pe’er. “Smoothing Nonlinear Variational Objectives with Sequential Monte Carlo”, in *ICLR*, 2019.
- [5] A. K. Moretti, Z. Wang, L. Wu, I. Drori, and I. Pe’er. “Variational Objectives for Markovian Dynamics with Backwards Simulation”, in *ECAI*, 2020.
- [6] D. P. Kingma, and M. Welling. “Auto-encoding Variational Bayes”, *arXiv preprint arXiv:1312.6114*, 2013.
- [7] J. Chung, K. Kastner, L. Dinh, K. Goel, A. Courville, and Y. Bengio. “A Recurrent Latent Variable Model for Sequential Data”, in *NeurIPS*, 2015.
- [8] M. Fraccaro, S. K. Sønderby, U. Paquet, and O. Winther. “Sequential Neural Models with Stochastic Layers”, in *NeurIPS*, 2016.
- [9] Y. Burda, R. B. Grosse, and R. Salakhutdinov, “Importance Weighted Autoencoders”, *arXiv preprint arXiv:1509.00519*, 2015.
- [10] A. R. Jones, C. C. Overly, and S. M. Sunkin. “The Allen Brain Atlas: 5 years and beyond”, *Nature Reviews Neuroscience*, 10, 821EP, 2009.