# 高次弱近似法による Neural SDE の学習

西野 颯馬 (東京都立大学大学院理学研究科)

本講演では，Neural SDE の学習について考える．Neural SDE とは，SDE のベクトル場をニューラルネットワークで置き換えたものである．特に，本公演では下記の $d_x$ 次元 Neural SDE を扱う：

$$\mathrm{d}X_t = \mu_\theta(X_t)\mathrm{d}t + \sigma_\theta(X_t) \circ \mathrm{d}W_t. \tag{1}$$

ここで，$\mu_\theta \colon \mathbb{R}^{d_x} \to \mathbb{R}^{d_x}$ と $\sigma_\theta \colon \mathbb{R}^{d_x} \to \mathbb{R}^{d_x} \times \mathbb{R}^{d_w}$ はニューラルネットワーク，$W = \{W_t\}_{t\in[0,T]}$ は $d_w$ 次元標準ブラウン運動，$\circ$ は Stratonovich 積分を表す．ニューラルネットワークのパラメータ $\theta$ が与えられるごとに，SDE (1) が一つ定まり，対応する解 $X = \{X_t\}_{t\in[0,T]}$ が得られる．このパラメータは，与えられた時系列データセットと Neural SDE の解との間の誤差を最小化するようにして決定される．本講演においては，誤差関数として

$$\mathrm{Loss}(X, Y) = \left\| E[S(X)^{\le M}_{[0,T]}] - E[S(Y)^{\le M}_{[0,T]}] \right\|_2^2$$

を用いる．ここで，$S(X)^{\le M}_{[0,T]}$ は $X = \{X_t\}_{t\in[0,T]}$ の $M$ 次までの打ち切りシグネチャである．このシグネチャは次の微分方程式を満たすことが知られている：

$$\mathrm{d}S(X)^{\le M}_{[0,t]} = S(X)^{\le M-1}_{[0,t]} \otimes \circ dX_t. \tag{2}$$

これに式 (1) を代入することで，

$$\mathrm{d}S(X)^{\le M}_{[0,t]} = S(X)^{\le M-1}_{[0,t]} \otimes (\mu_\theta(X_t)\mathrm{d}t + \sigma_\theta(X_t) \circ \mathrm{d}W_t) \tag{3}$$

を得る．よって，$S(X)^{\le M}_{[0,T]}$ は式 (1) と (3) の連立 SDE を解くことで得られる．誤差関数を計算する上で，$E[S(X)^{\le M}_{[0,T]}]$ を数値的に求める必要がある．本講演においては，Ninomiya–Victoir scheme や Ninomiya–Ninomiya scheme などの SDE の高次弱近似手法を (1) と (3) の連立 SDE に適用することで，$E[S(X)^{\le M}_{[0,T]}]$ を求める．これらの高次弱近似法を適用した場合に，シグネチャの期待値をどれくらいの精度で近似できるか，さらに Neural SDE の学習をどれくらい高速化できるかを議論する．

本講演の基づく研究において，有益な助言をいただいた篠崎裕司准教授 (一橋大学) に深く感謝申し上げます．