PDE-Net: Deep Learning of Stochastic partial differential equation from data

Zhendong Shi and Ercan E. Kuruoglu

Tsinghua-Berkeley Shenzhen Institute

Abstract

We propose a new deep neural network to discover (time-dependent) PDEs from observed dynamic data with minor prior knowledge on the underlying mechanism that drives the dynamics even in a noisy environment.

Based on the model, we can easily quantify uncertainties of the neural network and prove the feasibility of the model through numerical studies on the Black-Scholes equation.

Motivation

This research intends to introduce the principle of PDE neural network, extend it to the stochastic process and unveil the connection between partial differential equations and neural network.

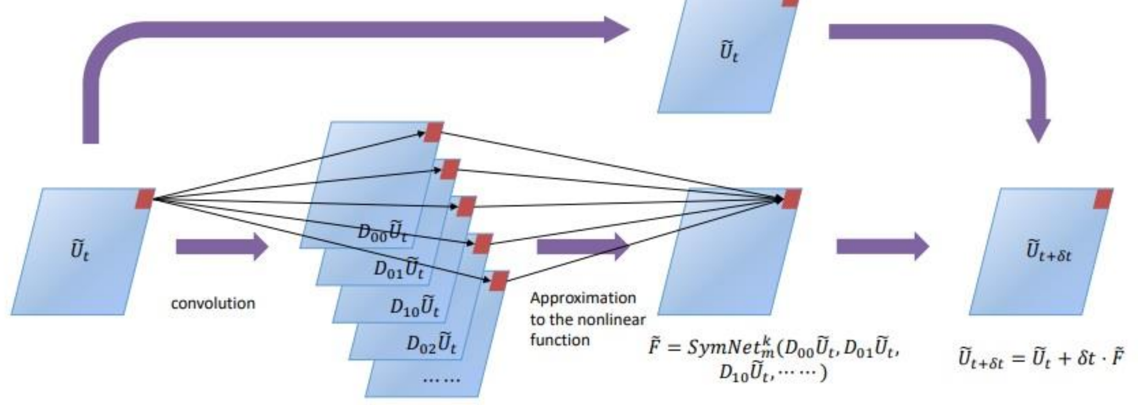
Principle

(1) Architecture of PDE-Net

$$U(t,x,y) = F(U, U_x, U_y, U_{xx}, U_{xy}, U_{yy}, \cdots)$$

$$\tilde{U}(t+\delta t,\cdot) \approx \tilde{U}(t,\cdot) + \delta t \cdot SymNet_m^k(D_{00}\tilde{U}, D_{01}\tilde{U}, D_{10}\tilde{U}, \cdots)$$

Here, the operators Dij are convolution operators with the underlying filters. The operators D10, D01, D11, etc. approximate differential operators, i.e. $D_{ij}u \approx \frac{\partial^{i+j}u}{\partial^i x \partial^j y}$.



(2) Convolutions and Differentiations

The relationship between convolutions and differentiations was presented by

$$(f \circ q)[l_1, l_2] = \sum_{k_1 k_2} q[k_1, k_2] f[l_1 + k_1, l_2 + k_2]$$

(3) Order of Sum Rules

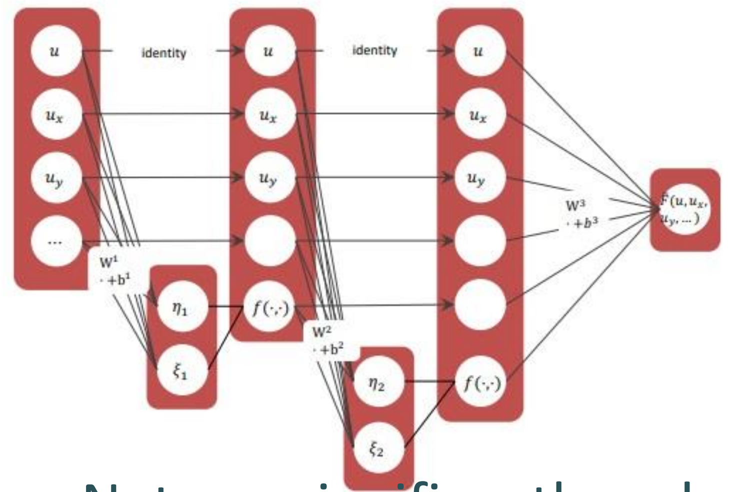
The order of sum rules is closely related to the order of vanishing moments in wavelet theory.

Let q be a filter with sum rules of order α∈Z^2+. Then for a smooth function F(x) on R^2,we have

$$\frac{1}{|\varepsilon|^\alpha} \sum_{k \in Z^2} q(k) F(x + \varepsilon k) = C_\alpha \frac{\partial^\alpha}{\partial x^\alpha} F(x) + O(\varepsilon), as\ \varepsilon \to 0$$

(4) Design of SymNet

The symbolic neural network is introduced to approximate the multivariate nonlinear response function F.



Algorithm 1 $SymNet_6^2$

Input: $u, u_x, u_y, v, v_x, v_y \in \mathbb{R}$,

$(\eta_1, \xi_1)^\top = W^1 \cdot (u, u_x, u_y, v, v_x, v_y)^\top + b^1, W^1 \in \mathbb{R}^{2\times 6}, b^1 \in \mathbb{R}^2$;

$f_1 = \eta_1 \cdot \xi_1$;

$(\eta_2, \xi_2)^\top = W^2 \cdot (u, u_x, u_y, v, v_x, v_y, f_1)^\top + b^2, W^2 \in \mathbb{R}^{2\times 7}, b^2 \in \mathbb{R}^2$;

$f_2 = \eta_2 \cdot \xi_2$;

Output: $F = W^3 \cdot (u, u_x, u_y, v, v_x, v_y, f_1, f_2)^\top + b^3 \in \mathbb{R}, W^3 \in \mathbb{R}^{1\times 8}, b^3 \in \mathbb{R}$.

SymNet can significantly reduce memory load and computation cost when input data is large.

(5) Loss Function and Regularization

We adopt the following loss function to train the proposed PDE-Net:

$$L = L^{data} + \lambda_1 L^{moment} + \lambda_2 L^{SymNet}$$

where the hyper-parameters λ1 and λ2 are chosen as λ1=0.001 and λ2=0.005.

(6) Uncertainty extension

$$\hat{U}(t,x,y) = U(t,x,y) + MW$$

A random item is added on the basis of the original PDE, here W is the standard Brownian motion.

Numerical result

Black Scholes equation is a fundamental partial differential equation in financial area.

$$\frac{\partial C}{\partial t} + \frac{\sigma^2}{2} S^2 \frac{\partial^2 C}{\partial S^2} + rS\frac{\partial C}{\partial S} - rC = 0$$

where C represents the option price, and asset price S is assumed to conform to geometric Brownian motion defined as equation:

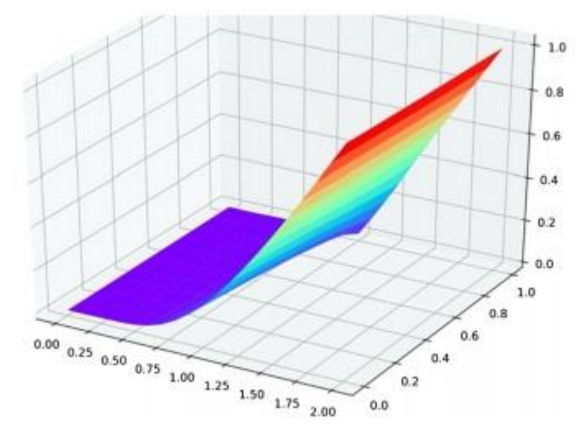
$$dS = \mu S dt + \sigma S dz$$

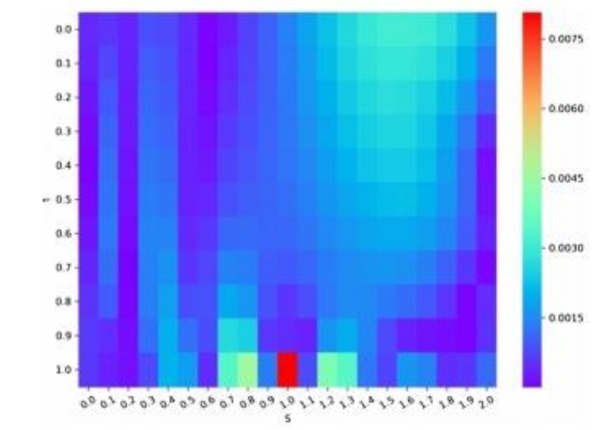
In this experiment, the classical Black Scholes Equation of European call options is tested. Mean square error (MSE) and maximum absolute error (MAE) are used to measure the accuracy and reliability of different methods.

Consider the following Black Scholes differential equation for the European call option:

$$\frac{\partial C}{\partial t} + 0.08 * S^2 \frac{\partial^2 C}{\partial S^2} + 0.04 S \frac{\partial C}{\partial S} - 0.04C = 0$$

$$(S,t) \in (0,2) * (0,1)$$





The left figure shows the numerical solution C obtained by PDE-net, and the right figure shows the absolute error comparing with the data generated by the above equation.

Correct PDE	$u_t = 0.06u + 0.04u_x - 0.08u_{xx}$
PDE – Net	$u_t = 0.0594u + 0.0389u_x - 0.0801u_{xx}$
Bernstein – net	$u_t = 0.0537u + 0.0422u_x - 0.0768u_{xx}$
Convolution – net	$u_t = 0.0576u + 0.0365u_x - 0.0782u_{xx}$

Method	MSE	MAE	Calculation time
PDE - Net	1.16×10^{-11}	1.74×10^{-5}	0.464
Bernstein - net	9.11×10^{-5}	2.79×10^{-2}	0.353
Convolution - net	8.19×10^{-8}	1.13×10^{-3}	0.404

Through above tables, we can find that PDE can obtain an approximation with a maximum error of less than 2 * 10^−5, only a few training points are required. Comparing with other neural network methods, PDE-Net performs much better except for calculation time.

Conclusion

1. According to principle metioned by this research, a numeric-symbolic hybrid deep network,called PDE-Net can be constructed and extended to the form of stochastic process.
2. This neural network can be applied to the solution of various fundamental equations such as the Black-Scholes equation.
3. The results confirm that PDE-Net performs much better than other neural network we mentioned in the experiment in terms of MSE and MAE.
4. The results confirm that PDE-Net has the potential to learn the stochastic equation through data.