

Neural SDE для нахождения моментов разладки временных рядов

Папай Иван Дмитриевич

Московский физико-технический институт

*Курс: Автоматизация научных исследований
(практика, В. В. Стрижов)/Группа 1286*

Эксперт: В. В. Стрижов

Консультант: Э. А. Владимиров

2024

Предмет исследования

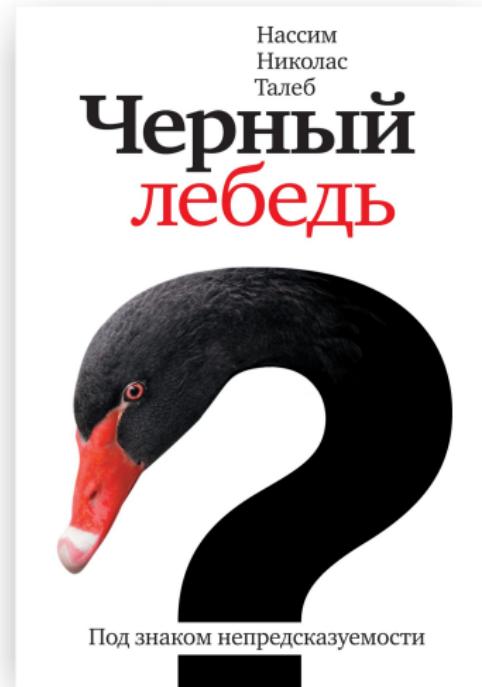


Рис.: первое упоминание термина

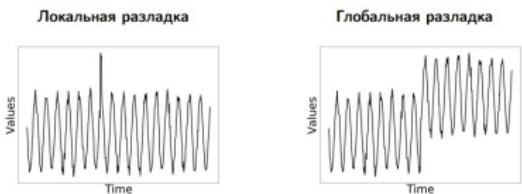


Рис.: случай временных рядов

Модель ResNet

Изначально Neural ODE был разработан как альтернатива методу Residual Networks(ResNet), состоящих из последовательности скрытых слоёв, значения на каждом из которых подчинялись следующей формуле:

$$h_{k+1} = h_k + f(h_k, w_k) \quad (1)$$

— где h_k — вход k -го слоя для $k \in [1, K]$, K -число слоёв и $f(h_k, w_k)$ — функция, параметризованная по w_k — динамический параметр, задающийся перед началом обучения модели.

Residual net

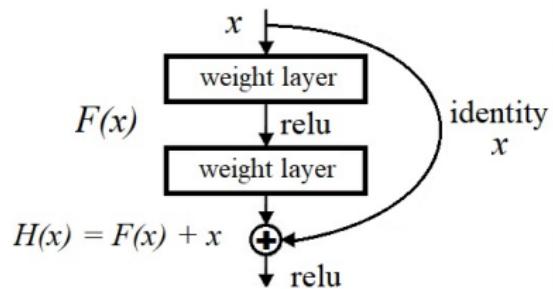


Рис.: схематичное описание работы

Модель Neural ODE

$$h_t = h_s + \int_s^t f(h_l, l; w) dl, \quad (2)$$

— вычисление такого дифференциального уравнения является задачей для Neural ODE. В данной работе в качестве числа слоев берется число элементов во временном ряду.

Algorithm 1 Neural ODE-solver

Require: динамические параметры w , начальное/конечное время t_0, t_1 , конечное значение $z(t_1)$, градиент функции потерь в конечной точке $\frac{\delta L}{\delta z(t_1)}$

$s_0 = [z(t_1), \frac{\delta L}{\delta z(t_1)}, 0_{[w]}]$

▷ Начальное состояние

$[z(t_0), \frac{\delta L}{\delta z(t_1)}, \frac{\delta L}{\delta w}] = ODESolve(s_0, [f(z(t), t, w), -a(t)^T \frac{\delta f}{\delta z}, -a(t)^T \frac{\delta f}{\delta w}], t_1, t_0, w)$

▷ Возвращаем градиенты

$return \frac{\delta L}{\delta z(t_0)}, \frac{\delta L}{\delta w}$

Рис.: псевдокод Neural ODE

Здесь $ODESolve$ — это черный ящик, возвращающий решение ODE — в качестве него предлагается использовать метод Рунге-Кутта, а L — MSE (mean squared error — среднее квадратов отклонения элементов выборки от их оценок).

Переход к Neural SDE

Для учёта шума в наше дифференциальное уравнение следует добавить недетерминированную компоненту, диффузию.

Получится следующее выражение, являющееся интегралом Стратоновича:

$$dX_t^w = h(t, X_t^w; w)dt + \sigma(X_t^w; w)dB_t \quad (3)$$

— где $B_t = [B_t^1 \dots B_t^K]$ - Винеровский процесс той же размерности, что и X_t , а $\sigma(X_t^w; w)$ - его матрица ковариаций в t -й момент времени

Обобщим это выражение для модели ResNet:

$$dh_t = f(h_t, t, w), \quad (4)$$

— таким образом выражение (1) для $(k+1)$ -го слоя изменится:

$$h_{k+1} = h_k + f(h_k, w_k) + \sigma(X_k^w; w)B_k, \quad (5)$$

— соответственно алгоритм остаётся тем же, что и для ODE с поправкой на вычисление матрицы ковариации элементов ряда.

Пример выборки с неоднородной стохастической природой

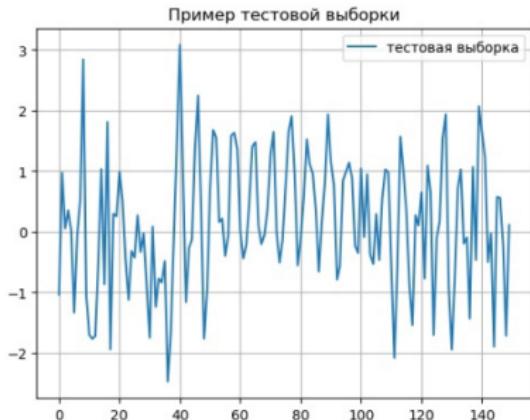


Рис.: выборка

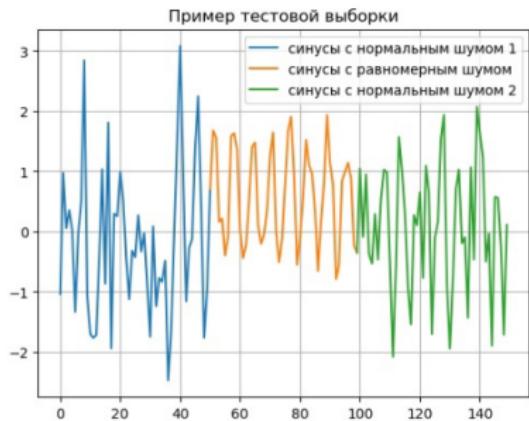


Рис.: происхождение разных отрезков ряда

Использование библиотеки torchSDE

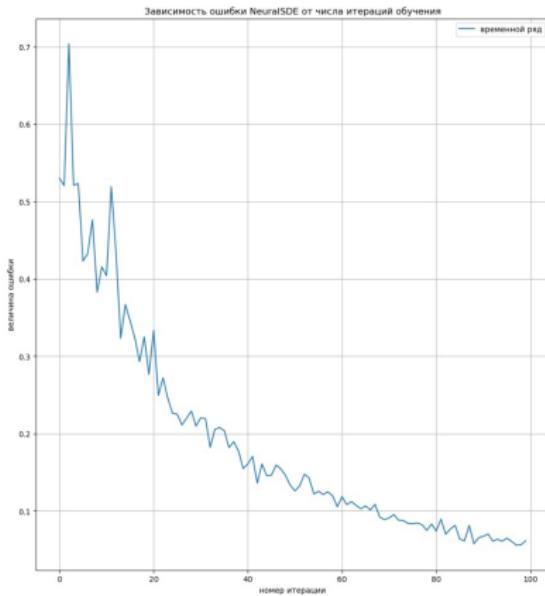


Рис.: процесс обучения

Строится матрица Ганкеля, описывающая фазовые траектории процесса, размера $(n+1)$ на w вида:

$$\begin{bmatrix} f_1 & f_2 & \cdots & f_w \\ f_2 & f_3 & \cdots & f_{w+1} \\ \vdots & \vdots & \ddots & \vdots \\ f_{n+1} & f_{n+2} & \cdots & f_{n+w} \end{bmatrix} \quad (6)$$

— где n — размер выборки, w — ширина окна.

Результаты поиска точек разладки

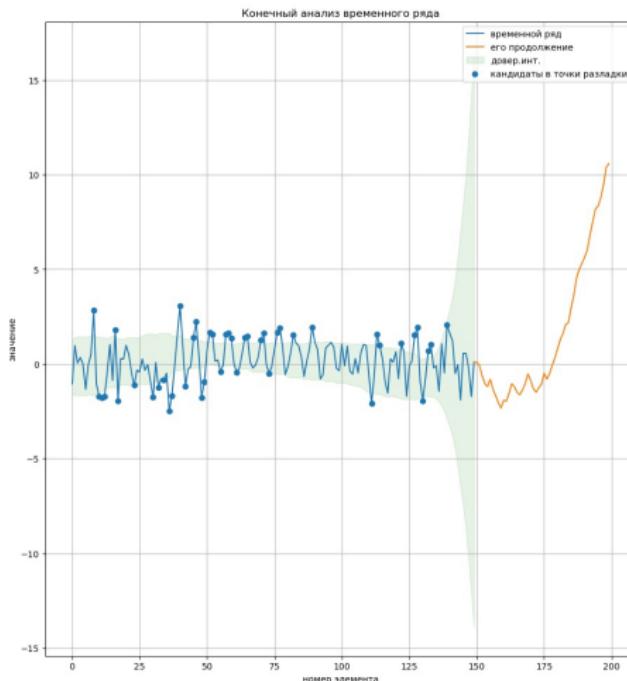


Рис.: воссоздание их фаз.траекторий с помощью Neural SDE

Источники

- [1] "Neural Ordinary Differential Equations Ricky" T. Q. Chen, Yulia Rubanova, Jesse Bettencourt, David Duvenaud
- [2] "Neural SDE: Stabilizing Neural ODE Networks with Stochastic Noise" Xuanqing Liu, Tesi Xiao, Si Si, Qin Cao, Sanjiv Kumar, Cho-Jui Hsieh
- [3] "Riemannian Neural SDE: Learning Stochastic Representations on Manifolds" Sung Woo Park , Hyomin Kim , Kyungjae Lee , Junseok Kwon
- [4] "Riemannian Diffusion Models" Chin-Wei Huang, Milad Aghajohari, Avishek Joey Bose, Prakash Panangaden, Aaron Courville
- [5] Tian Qi Chen, Yulia Rubanova, Jesse Bettencourt, and David K Duvenaud. Neural ordinary differential equations. In *Advances in Neural Information Processing Systems*, pages 6572–6583, 2018.
- [6] F.Yu. Yaushev, R. V. Isachenko, V. V. Strijov. Concordant models for latent space projections in forecasting, 2020.