
Neural SDE: phase trajectories of SDE in the action

A Preprint

Papay Ivan
MIPT University
papai.id@phystech.edu

Vladimirov Eduard
MIPT University

2024 год

Данная статья предлагает углубиться в математический аппарат, на котором строится модель Neural SDE. В ней будет рассмотрено, как вычисление фазовых траекторий СДУ обеспечивает качественный прогноз аномалий во временном ряду. Таким образом это предоставит как возможность эффективнее бороться с шумами, так и, в частности, полезный инструмент для упреждения "чёрных лебедей" которые могли бы нарушить корректную работу Neural SDE в виду высокой корреляции элементов анализируемой выборки между собой.

Keywords SDE · Stratonovich integral · More

1 Введение

Сбор данных и подготовка их к последующей обработке всегда были одной из важнейших задач машинного обучения. К сожалению не всегда исследователь может гарантировать их целостность и корректность, ведь для тренировки модели чаще всего требуются выборки из тысяч, а то и десятков тысяч элементов - не удивительно, что в данных допускается наличие "шума влияющего на работу обученной модели. Эта задача остаётся актуальной и для временных рядов, то есть данных, индексированных относительно временной координаты. Естественно желание - имея данные для начала временного ряда, проверить: возможно ли продолжить его новыми данными, насколько такое продолжение будет естественно, и не сломает ли это природу текущего временного ряда в стохастическом смысле?

Отсюда и далее мы сконцентрируемся на работе исключительно с временными рядами. В таком наши данные будут представлять из себя данные о некоем дискретном случайном процессе. Дискретном, потому как входная выборка, как множество, точно не будет континуально в силу естественной ограниченности анализируемых данных. Тем не менее корректно будет перейти к непрерывному случайному процессу в силу того факта, что он порождается сигма-алгеброй из конечномерных распределений, которые реально аппроксимировать с помощью данных, предоставленных для обучения модели.

Но пока что это всего лишь слова - как именно мы будем аппроксимировать искомые распределения? Если бы природа данных была бы строго детерминированной и мы бы изначально имели представление о распределении рассматриваемого случайного процесса, уместно было бы применить метод интерполяции или линейной регрессии. Но в условиях полной неопределённости по отношению и к характеру распределений, как функций, и к её параметрам - нам потребуется что-то другое. Как вариант: аппроксимировать ряд дифференциальными уравнениями.

Сама идея использования обыкновенных дифференциальных уравнений ("ОДУ"отсюда и далее) далеко не так нова, как могло бы показаться на первый взгляд. Так, примерно с 2017-го года она была использована для создания и теоретического обоснования корректности работы модели Neural ODE. Тем не менее, такой метод был всё ещё слаб в робастном смысле: то есть модель легко подпадала под влияние гауссовского шума, а также была уязвима к состязательным атакам. Модель Neural SDE уже строилась на использовании стохастических дифференциальных уравнений ("СДУ"отсюда и далее) и была в этом плане эффективнее своего предшественника. Математический аппарат требовался ещё

более серьезный, ведь для вычисления решения СДУ без знаний стохастического анализа, исчисления Ито и Стратоновича обойтись было нельзя.

Главной целью данного исследования является построение decision-rejection(принятие-отрицание) критерия корректности той или иной гипотезы о вероятностном распределении входных данных, как некоторого непрерывного случайного процесса. Таким образом для проверки фрагмента временного ряда на наличие аномалий достаточно применить этот критерий для проверки гипотезы о тождественности распределений для конкретного диапазона и для всего остального ряда - разумно будет заключить, что в ряду происходят аномалии, если природа данных в стохастическом смысле резко поменялась.

Вопрос состоит в том: как мы собираемся это делать? Ответ следующий - полагая, что временной ряд порождается определенными конечномерными распределениями, мы сможем приблизить его с помощью стохастических дифференциальных уравнений. То же, разумеется, применимо и к анализируемому диапазону, который требуется проверить на наличие аномалий. Если фазовые траектории полученных дифференциальных уравнений различаются, то есть происходит резкое их возмущение, то очевидно, что в ряду произошла аномалия.

В прошлых работах, связанных с Neural SDE, СДУ использовались только для построения доверительных интервалов для элементов временного ряда. Этот подход в статье предлагается развить посредством использования фазовых траекторий полученных СДУ. Таким образом, можно будет проверять большие массивы данных на корреляцию между собой. В том числе рассматривается конкретная задача - проверить, что два временных ряда обладают одинаковым вероятностным распределением. А именно проверяется соответствие видеоряда готовки еды и ряда данных, полученных с акселерометра, прикрепленного к его руке. Обладают ли эти данные одной и той же природой?

Список литературы

- [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