

Применение нейронных сетей для калибровки
умеренно устойчивых моделей Леви
по вероятностям пересечения
блуждающей частицей поглощающего барьера

Родоченко Василий Владимирович

Южный Федеральный Университет

ноябрь 2019

Содержание

- 1 Введение
 - Процессы Леви
 - One touch digital
- 2 Архитектура сети и функция активации
- 3 Подготовка исходных данных
 - Исторические данные
 - Синтетические данные
- 4 Обучение сети
- 5 Заключение

Процессы Леви

Процесс Леви – это стохастически непрерывный процесс с независимыми и стационарными приращениями (подробнее [3]). Известно, что процесс Леви X_t полностью определяется своей характеристической экспонентой ψ , которая находится из представления характеристической функции: $E[e^{i\xi X_t}] = e^{-t\psi(\xi)}$. Характеристическая экспонента может быть записана по формуле Леви-Хинчина:

$$\psi(\xi) = \frac{\sigma^2}{2}\xi^2 - i\mu\xi + \int_{-\infty}^{+\infty} (1 - e^{i\xi y} + i\xi y \mathbf{1}_{|y|\leq 1})\Pi(dy),$$

где σ^2 – дисперсия гауссовой компоненты, а мера Леви $\Pi(dy)$ удовлетворяет условию $\int_{\mathbf{R}\setminus\{0\}} \min\{1, y^2\}\Pi(dy) < +\infty$. Если $\Pi(dx) = \pi(x)dx$, тогда $\pi(x)$ – плотность Леви.

Для умеренно устойчивых процессов Леви $\psi(\xi)$ определяется по формуле:

$$\psi(\xi) = -i\mu\xi + c_+\Gamma(-\nu_+)[\lambda_+^{\nu_+} - (\lambda_+ + i\xi)^{\nu_+}] + \\ + c_-\Gamma(-\nu_-)[(-\lambda_-)^{\nu_-} - (-\lambda_- - i\xi)^{\nu_-}],$$

где $\nu_+, \nu_- \in (0, 2)$, $\nu_+, \nu_- \neq 1$, $c_+, c_- > 0$, $\mu \in \mathbf{R}$, и $\lambda_- < -1 < 0 < \lambda_+$.

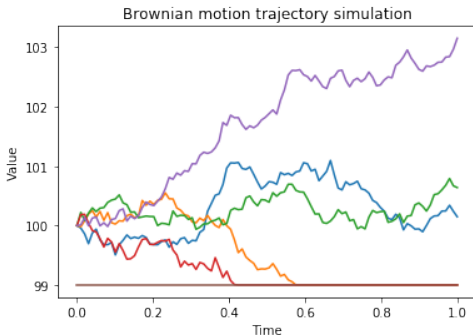
$$\pi(x) = c_+e^{\lambda_+x}|x|^{-\nu_+-1}\mathbf{1}_{\{x<0\}} + c_-e^{\lambda_-x}|x|^{-\nu_- -1}\mathbf{1}_{\{x>0\}}.$$

Случай CGMY

Если $c_- = c_+ = c$ и $\nu_- = \nu_+ = \nu$, тогда мы получаем модель KoBoL (CGMY).

В параметризации CGMY $C = c$, $G = \lambda_+$, $M = -\lambda_-$, $Y = \nu$.

Задача о пересечении блуждающей частицей поглощающего барьера



Приложения

В финансовых приложениях такого рода задача возникает при вычислении стоимости производной ценной бумаги – опциона.

Опцион one touch digital

Описание контракта

Цифровой опцион одного касания (**one touch digital**) – контракт, который платит 1\$ при условии, что за время его жизни $(0, T]$ цена базового актива S пересекает некоторый фиксированный барьер K . Его риск-нейтральная цена в момент времени 0 может быть интерпретирована как **вероятность пересечения фиксированного барьера** за время жизни T .

Полезное свойство

Синтетические цены такого опциона можно получить из **исторических данных** даже на тех рынках, где торговля деривативами недостаточно развита, чтобы использовать данные о торгах.

О задаче калибровки

Цель калибровки

- Определить модель, адекватно отражающую поведение блуждающей частицы
- Получить **подходящий** набор параметров этой модели

Описание задачи

Под "подходящим" набором понимаются **оптимальные параметры модели** – в том смысле, что они **минимизируют** некоторый **функционал ошибки**, понимаемый как разность между вектором цен (статистических вероятностей), генерируемых рынком и вектором, предсказываемым моделью.

Проблемы исчерпывающих методов калибровки

- Методы калибровки моделей со скачками и/или стохастической волатильностью требуют огромных вычислительных мощностей, потому что методы вычисления производных ценных бумаг в этих моделях работают сравнительно медленно [11].
- Вычисление необходимо производить постоянно и многократно

Цель работы

Предложить метод калибровки умеренно устойчивых процессов Леви с использованием ИНС по данным, содержащим историю переходов блуждающей частицы (цены базового актива) через ряд фиксированных барьеров, на примере модели умеренно-устойчивой модели Леви: CGMY.

Особенности применения нейронных сетей

Преимущество в скорости

Обучение нейронной сети можно проводить однажды, а затем использовать её для любого набора похожих данных.

Благодаря этому, после обучения скорость решения задач калибровки может увеличиваться на порядки [6, 8].

Этапы решения задачи калибровки и использованием ИНС

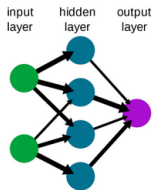
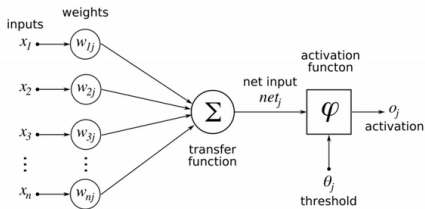
- Обучаем модель на исторических данных.
- Используем получившуюся модель вместо существующего метода калибровки и сравниваем их результаты с точки зрения скорости и точности.

Содержание

- 1 Введение
 - Процессы Леви
 - One touch digital
- 2 **Архитектура сети и функция активации**
- 3 Подготовка исходных данных
 - Исторические данные
 - Синтетические данные
- 4 Обучение сети
- 5 Заключение

Описание

Нейрон может принимать на вход некоторое количество сигналов (вектор $x = [x_1, x_2, \dots, x_{d_0}]$, $x \in R^{d_0}$), которые затем изменяются при помощи весов $\omega_j = [\omega_{i,j}, \dots, \omega_{i,m}]$, $\omega_{i,j} \in \mathbb{R}^m$, соответствующих (скрытому) слою j и суммируются как линейная комбинация. К результату применяется функция активации, которая окончательно определяет значение выходного сигнала $o_j = y$, $y \in R^{d_1}$.



Пример

Модель CGMY

Предполагается, что набор параметров C, G, M, Y (параметр сноса μ фиксируется условием мартингальности) является функцией от 5 параметров: S, K, T, r, P

Простейшая ИНС имела бы $d_0 = 5$ и $d_1 = 4$, $\mathbb{R}^5 \rightarrow \mathbb{R}^4$.

Более точное приближение могло бы включать некоторую историю изменения цены базового актива и цен опционов, что увеличило бы размерность входного вектора.

Содержание

- 1 Введение
 - Процессы Леви
 - One touch digital
- 2 Архитектура сети и функция активации
- 3 Подготовка исходных данных
 - Исторические данные
 - Синтетические данные
- 4 Обучение сети
- 5 Заключение

Источники данных

Мы собрали историю изменения цены биткоина (BTC/USD) с торговой площадки GDAX при помощи API за 2017 и 2018 годы. Наиболее мелким периодом времени, который мы рассматривали, были интервалы по 5 минут – это позволило минимизировать количество обработки данных, при сохранении достаточно мелкого шага по времени. В случае, если за 5 минут не проходило ни одной сделки – предполагалось, что цена актива осталась на прежнем уровне.

Подготовка данных

Традиционно данные предлагаются в формате *OHLCV* (open, high, low, close, volume) и поля имеют следующую интерпретацию:

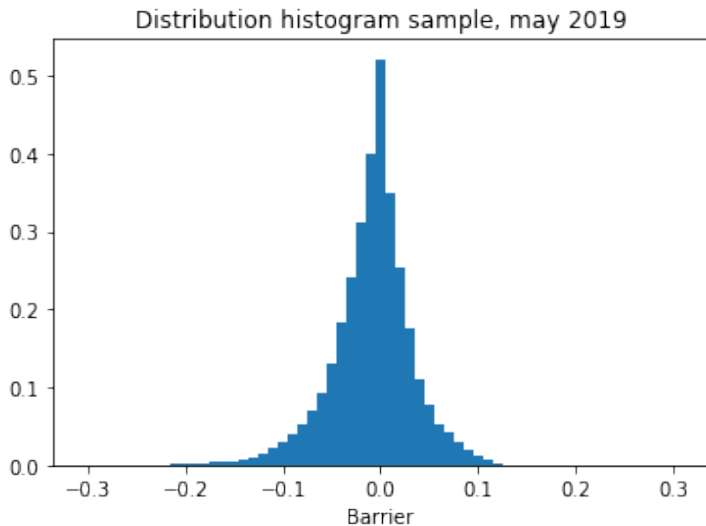
- open – цена первой сделки за торговый период,
- high – наивысшая цена сделки за торговый период,
- low – минимальная цена сделки за торговый период,
- close – цена последней сделки за торговый период,
- volume – объём сделок за торговый период.

Формирование входного вектора

Алгоритм

- 1 Формируем взвешенную цену закрытия за 5 минут как $S_m^w = S_i \cdot V_i / \sum(V_i), i = 1, 2, \dots, 5, m \in N$, где среди $V_i, i = 1, \dots, 5$ по крайней мере одна не равна нулю.
- 2 Переходим к последовательности логарифмов приращений: $x_0 = 0, x_{m+1} = \ln S_{m+1}^w - \ln S_m^w$
- 3 Фиксируем барьеры $h_k^+ = 10^{-2} \cdot k, k = 1, 2, \dots, 30$ и $h_k^- = 10^{-2} * k, k = 1, 2, \dots, 30$ и регистрируем их пересечения приращениями цены за один торговый день.
- 4 Строим гистограмму исторических (статистических) вероятностей пересечения барьеров по имеющимся исходным данным и получаем рыночную стоимость синтетического опциона one touch digital на логарифм цены базового актива, для каждой из цен исполнения $K = h_k^\pm$, сроком жизни в 1 день.

Статистические вероятности пересечения барьеров



"Исчерпывающая" калибровка

Алгоритм

- 1 Для некоторого наперёд заданного набора параметров модели мы вычисляем стоимость опциона one touch digital для всех $K = h_k^\pm$ и ошибку, понимаемую как среднеквадратическое отклонение от рыночной стоимости.
- 2 Для минимизации ошибки используем алгоритм Нелдера-Мида, со случайно выбранным симплексом.
- 3 Повторяем схему для нескольких наборов стартовых параметров, взятых на кубе Соболя.

Характеристика схемы

Хотя и не является оптимальной по времени, позволяет достаточно хорошо охватить весь диапазон значений параметров и приблизиться к глобальному минимуму функции ошибки.

Генерация синтетических данных

Алгоритм

- 1 Выберем достаточно большое количество случайных (не совпадающих) точек на кубе Соболя в d_1 - мерном пространстве параметров
- 2 Вычислим вероятности пересечения барьеров как цены опционов one touch digital, используя метод приближенной факторизации Винера-Хопфа, разработанный в [10].

Мы положили число таких точек равным 30 000, что позволило натренировать ИНС.

Содержание

- 1 Введение
 - Процессы Леви
 - One touch digital
- 2 Архитектура сети и функция активации
- 3 Подготовка исходных данных
 - Исторические данные
 - Синтетические данные
- 4 Обучение сети
- 5 Заключение

Функции активации

Отсутствие функции активации в структуре ИНС существенно снижает её способность к воспроизведению сложных закономерностей.

- Relu: $\phi(x) = \max(x, 0)$,
- Сигмоида: $\phi(x) = \frac{1}{1+e^{-x}}$,
- Leaky Relu: $\phi(x) = \max(x, \alpha x)$, $0 < \alpha < 1$
- Часто используется ELU

$$R(z) = \begin{cases} z, & z > 0, \\ \alpha(e^z - 1), & z \leq 0, \end{cases}$$

с гиперпараметром α

- Её аналог без разрыва производной в нуле – MELU [8]

$$R(z) = \begin{cases} \frac{\frac{1}{2}z^2 + az}{z+b}, & z > 0, \\ \alpha(e^z - 1), & z \leq 0, \end{cases}, \quad a = 1 - 2a, \quad b = -2 + \frac{1}{\alpha},$$

мы используем во внутренних слоях ИНС.

Сокращение диапазона параметров

Благодаря анализу рынка криптовалют в работе [13] нам удалось существенно сократить диапазон поиска параметров в пределах указанных дат до следующего:

- $\nu \in (0.5, 1.4)$,
- $\lambda_+ \in (2, 100)$,
- $-\lambda_- \in (2, 100)$,
- $c \in (0.1, 5.0)$.

Результаты калибровки традиционным способом на имеющихся исторических данных также показали, что значения параметров модели CGMY находятся внутри приведённых интервалов.

Слои сети

Следуя идее архитектуры сетей из [7, 8, 12], мы включили четыре скрытых слоя и $\alpha = 0.5$ и расположили функции активации следующим образом:

- 1 Relu, $\phi(x) = \max(x, 0)$,
- 2 Melu,

$$R(z) = \begin{cases} \frac{\frac{1}{2}z^2 + az}{z+b}, & z > 0, \\ \alpha(e^z - 1), & z \leq 0, \end{cases}, \quad a = 1 - 2a, \quad b = -2 + \frac{1}{\alpha},$$

- $\alpha = 0.5$,
- 3 Melu, $\alpha = 0.5$,
 - 4 Leaky Relu, $\phi(x) = \max(x, \alpha x)$, $\alpha = 10^{-3}$,

Обучение и результаты

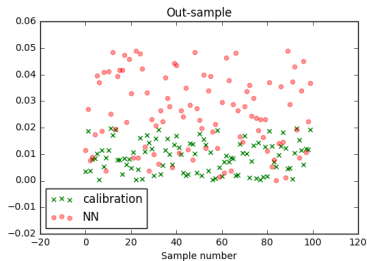
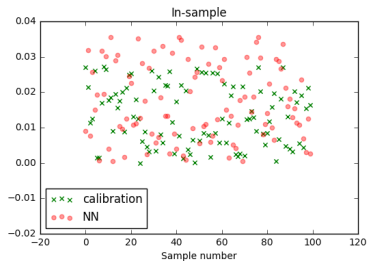
Методика обучения

- алгоритм стохастического градиентного спуска RMSprop (Keras)
- ошибка – нормализованное среднеквадратическое отклонение между прогнозируемыми ИНС оптимальными параметрами модели CGMY и параметрами, полученными методом традиционной калибровки.

Результаты калибровки

- OS Windows 10 64-bit, Intel Core i5-3470 CPU, 8 GB RAM).
- 870.0 секунд на образец – исчерпывающая калибровка
- 1.4 секунды на генерацию предсказания – ИНС

Значения ошибки для 100 случайных образцов



Содержание

- 1 Введение
 - Процессы Леви
 - One touch digital
- 2 Архитектура сети и функция активации
- 3 Подготовка исходных данных
 - Исторические данные
 - Синтетические данные
- 4 Обучение сети
- 5 Заключение

Вывод

Использование нейронных сетей позволяет быстро и без существенной потери точности проводить калибровку модели GCMY и оценивать в её рамках риски перехода блуждающей частицы через фиксированный барьер.

Полученный аппарат применим в финансовых приложениях – в частности, при вычислении цен опционов на криптовалюты и другие высокорисковые высоковолатильные активы.

Спасибо за внимание!

Литература

1. *Bishop, C. M.* Neural Networks for Pattern Recognition. // C. M. Bishop / Oxford University Press. Inc. – New York, 1995. NY, USA.
2. *Carr, P.* The fine structure of asset returns: An empirical investigation / P. Carr, H. Geman, D. Madan, M. Yor. // The Journal of Business. – 2002. – V.75, №2. – P. 305–333.
3. *Cont, R.* Financial modelling with jump processes / R. Cont, P. Tankov // Second Edition, Chapman & Hall, CRC Press. – 2008. – 606 P.
4. *Hochreiter, S.* Long short-term memory / J. Schmidhuber S. Hochreiter // Neural Computation. – 1997. – № 9(8). –P. 1735-1780.
DOI:10.1162/neco.1997.9.8.1735.
5. *Hornik, K.* Universal approximation of an unknown mapping and its derivatives using multilayer feedforward networks // K. Hornik, M. Stinchcombe, H. White / Neural Networks. – 1990. – №3. – P. 551–560.

Литература - 2

6. *Horvath, B.* Deep learning volatility a deep neural network perspective on pricing and calibration in (rough) volatility models // B. Horvath, A. Muguruza, M. Tomas / SSRN Electronic Journal. –2019. av. at. <http://dx.doi.org/10.2139/ssrn.3322085>
7. *Huh, J.* Pricing Options with Exponential Lévy Neural Network / J. Huh // Expert Systems with Applications. – 2019. – № 127.
8. *Itkin, A.* Deep learning calibration of option pricing models: some pitfalls and solutions / A. Itkin // (preprint) arXiv:1906.03507v1 [q-fin.CP] 8 Jun 2019.
9. *Kim, T.* Forecasting stock prices with a feature fusion LSTM-CNN model using different representations of the same data. // H. Y. Kim, T. Kim / PLoS ONE. – 2019. – V.14, №2. DOI: <https://doi.org/10.1371/journal.pone.0212320>
10. *Kudryavtsev O.* Advantages of the Laplace transform approach in pricing first touch digital options in Lévy-driven models / O. Kudryavtsev // Boletín de la Sociedad Matemática Mexicana. – 2016. –№22(2) –P. 711-731.

11. *Madan, D. B.* Machine Learning for Quantitative Finance: Fast Derivative Pricing, Hedging and Fitting // J. De Spiegeleer, D. B Madan, S. Reyners, W. Schoutens / Quantitative Finance. – 2018. – V.18, № 10. – P. 1-9. DOI: 10.1080/14697688.2018.1495335
12. *Osterlee, C. W.* A neural network-based framework for financial model calibration / S. Liu, A. Borovykh, L. Grzelak, C. W. Oosterlee // Journal of Mathematics in Industry. – 2019. – № 9(1):9. DOI: 10.1186/s13362-019-0066-7
13. *Кудрявцев, О. Е.* Статистические методы калибровки моделей цен криптовалют / О. Е. Кудрявцев, А. С. Гречко // Учет и статистика. –2018. –Т. 52, № 4. –С. 67-76.
- Исследование выполнено при поддержке РФФИ (проект 18-01-00910)