УДК 519.642.2

О механизме оценки вероятностей пересечения блуждающей частицей поглощающего барьера в модели CGMY, основанном на применении нейронных сетей

О. Е. Кудрявцев

Ростовский Филиал РТА

В. В. Родоченко

Южный Федеральный Университет

Аннотация

Мы предлагаем новый метод вычисления вероятности перехода блуждающей частицы через фиксированный барьер. Рассматривая историю поведения блуждающей частицы во времени, мы строим гистограммы вероятностей распределения пересечений наперёд заданных барьеров, используем исходное предположение о наиболее подходящей модели Леви (конкретно – модели CGMY), которой подчиняется динамика случайного блуждания и, затем, тренируем искусственную нейронную сеть для калибровки параметров этой модели. Благодаря тому, что обучение нейронной сети можно проводить однажды, полученный подход (с предварительно обученной сетью) позволяет многократно превзойти традиционные механизмы калибровки в скорости, при этом незначительно уступая в точности. Метод имеет широкий ряд приложений в задачах, возникающих на финансовых рынках – в особенности если актив является одновременно высоколиквидным и высоковолатильным (например, криптовалюты).

Ключевые слова: нейронные сети, вероятность перехода через барьер, процессы Леви, численные методы.

1 Введение

Значительный прогресс, который наблюдается в сферах искусственного интеллекта и машинного обучения, оказывает влияние на методы, которые используются исследователями и инженерами для решения ежедневных задач. Настоящая статья посвящена применению искусственных нейронных сетей для решения одной из классических проблем управления риском – оценки вероятности перехода блуждающей частицы через фиксированный барьер. В финансовых приложениях такая задача решается при вычислении стоимости производной ценной бумаги – опциона. Точнее говоря, нас интересует 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, что позволяет естественным образом получать синтетические цены такого опциона из исторических данных даже на тех рынках, где торговля деривативами недостаточно развита, чтобы использовать данные о торгах.

Мы ограничиваем себя случаем, когда удаётся сделать предположение о наиболее подходящей модели, которой следует поведение частицы, заметив, что выбор такой модели может существенно влиять на выбор подходящего для работы инструмента (см, напр. [Madan]).

Для обеспечения работы финансовых рынков необходимо ежедневное решение весьма значительного количества задач, которые имеют существенную вычислительную сложностью. Среди них – вычисление торговых индикаторов, цен производных ценных бумаг, определений коэффициентов хеджирования. Для построения эффективной численной реализации методов их решения привлекается аппарат теории вероятностей и современные вычислительные методы. Тем не менее, в частности, методы калибровки моделей со скачками и/или стохастической волатильностью требуют огромных вычислительных мощностей, потому что методы вычисления производных бумаг в этих моделях работают сравнительно медленно, а само вычисление необходимо производить постоянно и многократно [Madan].

Идея о том, что для ускорения процесса искусственные нейронные сети могут быть эффективно использованы для решения задачи калибровки моделей, описывающих динамику случайного процесса, описана, в частности, в [Horvath]. Основа подхода состоит в следующем. При условии, что модель известна, задачей калибровки является найти оптимальные параметры модели – в том смысле, что они минимизируют некоторый функционал, понимаемый как разность между вектором цен, генерируемых рынком и вектором цен, предсказываемых моделью. Реализация задачи калибровки делится на два этапа. Сначала мы заменяем существующий, медленный способ калибровки, его приближением в виде метода машинного обучения, и обучаем модель на исторических данных. Затем – используем получившуюся модель вместо существующего метода и сравниваем их результаты с точки зрения скорости и точности.

Поскольку обучение нейронной сети можно проводить однажды, а затем использовать её для любого набора похожих данных, временем на обучение сети в вычислительных экспериментах пренебрегают (как пренебрегают временем, потраченным на кодирование и подготовку данных в прочих случаях). После обучения скорость решения задач калибровки может увеличиваться на порядки [Horvath, Itkin].

Заметим, что этот подход, хотя и является эффективным, не лишён некоторых принципиальных недостатков, подробно освещённых в [Itkin] и [Huh]. Среди них – неточность предсказаний при ограниченном наборе тестовых данных, способность генерировать скачкообразные изменения цен в непрерывных моделях, а также необходимость специальным образом учитывать условия безарбитражности.

Целью настоящей статьи является показать его применимость его для калибровки модели CGMY на исходных данных, содержащих историю переходов блуждающей частицы (цены базового актива) через ряд фиксированных барьеров.

Напомним, что процесс Леви – это стохастически непрерывный процесс с независимыми и стационарными приращениями (подробнее [Cont]). Известно, что процесс Леви X_t полностью определяется своей характеристической экспонентой ψ, которая находится из следующего 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.

Значительный интерес для финансовых приложений представляют процессы Леви, которые моделируют движение частицы как комбинацию линейного тренда, бесконечного числа малых скачков и редких больших скачков. Ярким представителем таких моделей являются умеренно устойчивые процессы Леви (англ. Tempered stable Lévy processes (TSL)), характеристическая экспонента которых 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 $\lambda_+ < -1 < 0 < \lambda_+$.

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

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

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

2 Архитектура сети и функция активации

Искусственные нейронные сети (ИНС) – это сети из искусственных нейронов [Bishop]. Связи между нейронами моделируются при помощи "весов". Нейрон может принимать на вход некоторое количество сигналов (вектор $x = [x_1, x_2, ...x_{d_0}]$, $x \in R^{d_0}$), которые затем изменяются при помощи весов $\omega_j = [\omega_{i,j}, ..., \omega_{i,m}], \omega_{i,j} \in \mathbb{R}^m$, соответствующих (скрытому) слою j, и суммируются как линейная комбинация. К результату применяется функция активации, которая окончательно определяет значение выходного сигнала $o_j = y$, $y \in R^{d_1}$. Если не применять функцию активации – нейронная сеть сможет решать только задачи линейной регрессии, что ограничит её способность к обучению.

Из практических соображений генерируемые результаты должны принадлежать классу C^2. В силу этого соображения, мы ограничимся случаем без циклов, аналогичным рассмотренному в [Itkin]. Для него действует теорема Хорника [Horkik], согласно которой функция активации должна принадлежать этому же классу:

Теорема 1. *Пусть $\mathcal{N}^{\sigma}_{d_0,d_1}$ – множество ИНС с функцией активации $\sigma : \mathbb{R} \to \mathbb{R}$, с размерностью входного вектора $d_0 \in N$ и выходного вектора $d_1 \in N$. Пусть $F \in C^n$ и $F_{ANN} : \mathbb{R}^{d_0} \to \mathbb{R}$. Тогда если функция активации $\sigma \in C^n(\mathbb{R})$ и не является константой, то $\mathcal{N}^{\sigma}_{d_0,d_1}$ приближает F и все её производные до порядка n.*

Заметим, что разнообразие архитектур нейронных сетей в настоящее время достаточно велико. Обзор наиболее широко применяемых архитектур можно найти в [Osterlee, Huh]. В статье [Hochreiter] представлены рекуррентные сети с памятью LSTM, которые успешно применяются для задач регрессии и прогнозирования ценовых рядов – например, в работе [Kim]. В [Huh] конструируются нейронные сети специального вида для прогнозирования параметров процессов Леви, а также приводится ряд идей для увеличения объёма тренировочной выборки. В работе [Osterlee] ИНН успешно используются для вычисления стоимости опционов в моделях Блэка-Шоулза и Хестона.

Рассмотрим для примера четырехпараметрическую модель CGMY [Carr]. В ней предполагается, что набор параметров C, G, M, Y (параметр сноса μ фиксируется условием мартингальности), является функцией от 5 параметров S, K, T, r, P – которые являются ценой актива, ценой исполнения европейского опциона, временем истечения контракта, безрисковой процентной ставкой и ценой опциона. Простейшая нейронная сеть, приближающая такую модель имела бы $d_0 = 5$ и $d_1 = 4$ и выполняла бы роль аппроксиматора, действующего $\mathbb{R}^5 \to \mathbb{R}^4$. Более точное приближение могло бы включать некоторую историю изменения цены базового актива и цен опционов, что увеличило бы размерность входного вектора.

Благодаря увеличению вычислительных мощностей, во-первых и наличию в открытом доступе библиотек Keras, Tensorflow и др., во-вторых, – решение многомерных задач такого рода стало достижимым.

3 Подготовка исходных данных

В качестве исходных данных мы использовали историю изменения цены биткоина (BTC/USD), полученную сбором минутных данных с торговой площадки GDAX при помощи API за 2017 и 2018 годы. Удобным для автоматизированного анализа является тот факт, что торговля ведётся круглосуточно, поэтому данные присутствуют за каждый из дней. Наиболее мелким периодом времени, который мы рассматривали, были интервалы по 5 минут – это позволило минимизировать количество синтетических данных. В случае, если 5 минут не проходило ни одной сделки – предполагалось, что цена актива осталась на прежнем уровне.

Традиционно данные предлагаются в формате $OHLCV$ (open, high, low, close, volume) и 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.

Для уменьшения размерности, мы формируем взвешенную цену закрытия за 5 минут как $S_m^w = S_i \cdot V_i / \sum(V_i), i = 1, 2, ..., 5, m \in N$, где среди $V_i, \quad i = 1, .., 5$ по крайней мере одна не равна нулю.

4 Результаты

5 Заключение

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 анализе поведения цен криптовалют и других высокорисковых высоковолатильных активов. Заметим, что предложенная методика расширения обучающей выборки, основанная на генерации большого числа траекторий процесса CGMY, исходит из предположения, что охватывачен весь диапазон изменения параметров валидационной выборки.

Рассматривая историю поведения блуждающей частицы во времени, мы строим гистограммы вероятностей распределения пересечений наперёд заданных барьеров, используем исходное предположение наиболее подходящей о модели Леви (конкретно – модели CGMY), которой подчиняется динамика случайного блуждания и, затем, тренируем искусственную нейронную сеть для калибровки параметров этой модели. Метод имеет широкий ряд приложений в задачах, возникающих на финансовых рынках – в особенности если актив является одновременно высоколиквидным и высоковолатильным (например, криптовалюты)

Литература

1. *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.

2. *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

3. *Huh, J.* Pricing Options with Exponential Lèvy Neural Network / J. Huh // Expert Systems with Applications. – 2019. – № 127. DOI: 10.1016/j.eswa.2019.03.008

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. *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.

6. *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. ISSN 194-0874.

7. *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

8. *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

9. *Bishop, C. M.* Neural Networks for Pattern Recognition. // C. M. Bishop / Oxford University Press. Inc. – New York, 1995. NY, USA.

10. *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

11. *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.

12. *Black, F.* The pricing of options and corporate liabilities / F. Black, M. Scholes // Journal of Political Economy. – 1973. – № 81(3) – P. 637–654.

13. *Cont R.*, Financial modelling with jump processes, R. Cont, P. Tankov / Second Edition, Chapman & Hall/CRC Press. – 2008. – 606 P.

14. *Kudryavtsev O.* Advantages of the Laplace transform approach in pricing first touch digital options in Lévy-driven models. Boletín de la Sociedad Matemática Mexicana, 2016, vol. 22No.2, P. 711–731.