

# Оценка финансовых деривативов нейронной сетью на основе биномиального дерева

С. Г. Шорохов

**Аннотация**— На основе модели оценки опционов Кокса-Росса-Рубинштейна построены архитектуры нейронных сетей прямого распространения для аппроксимации стоимости европейских и американских опционов пут. Для европейского опциона пут первым скрытым слоем нейронной сети является полносвязный слой с функцией активации ReLU, последующие слои являются сверточными с 1D фильтром размера 2 и тождественной функцией активации. Для американского опциона пут первый скрытый слой нейронной сети также является полносвязным с функцией активации ReLU, последующие слои имеют функцию активации тахот и зависят как от выхода предыдущего слоя, так и от выхода входного слоя (цены-страйк). Слои обеих нейронных сетей имеют количество нейронов, уменьшающееся на единицу с каждым последующим слоем вплоть до одного нейрона в выходном слое. Показано, что нейронная сеть для европейского опциона пут может быть упрощена до нейронной сети прямого распространения из двух слоев.

**Ключевые слова**—оценка опционов, биномиальное дерево, модель CRR, сверточная нейронная сеть.

## I. ВВЕДЕНИЕ

Одной из наиболее важных задач в финансах является задача оценки справедливой стоимости опционов, для решения которой необходимо строить модель динамики цены базисного актива.

В наиболее известной модели оценки опционов – модели Блэка-Шоулза [1, 2] время изменяется непрерывно, а цена базисного актива моделируется стохастическим дифференциальным уравнением с начальным условием.

В модели Блэка-Шоулза справедливая цена дериватива, являющаяся функцией цены базисного актива и времени, должна удовлетворять уравнению с частными производными Блэка-Шоулза-Мертон [1, 2], при этом учитываются граничные (терминальные) условия в зависимости от вида дериватива. Для европейских опционов колл и пут краевая задача для уравнения Блэка-Шоулза-Мертон имеет точные аналитические решения, определяемые формулами Блэка-Шоулза [1].

Кроме точных аналитических методов, задача оценки опционов может решаться при помощи численных методов, к которым относятся, в частности, методы конечных разностей [3], преобразование Фурье [4], метод Монте-Карло [5] и пр.

В настоящее время основным трендом в финансах является применение глубокого обучения для решения широкого круга задач [6]. Для аппроксимации цен

опционов нейронные сети применяются достаточно давно, начиная с работ [7, 8], обзор подходов к нейросетевой аппроксимации цен опционов приводится в [9]. Нейронные сети могут также применяться для аппроксимации цен опционов в моделях локальной волатильности [10], когда волатильность  $\sigma$  является функцией цены базисного актива  $S$  и времени  $t$  [11].

Альтернативой модели Блэка-Шоулза является биномиальная модель оценки опционов Кокса-Росса-Рубинштейна [12] (далее модель CRR), в которой время и цена базисного актива изменяются дискретно.

## II. БИНОМИАЛЬНОЕ ДЕРЕВО ДОПУСТИМЫХ ЦЕН

Пусть  $t_0$  и  $S_0$  – текущие время и цена базисного актива соответственно. Рассматривается европейский опцион пут с ценой-страйк  $K$ , погашаемый в момент времени  $T > t_0$ . Платежная функция такого опциона в момент погашения равна

$$u(S, T) = \max(K - S, 0), \quad (1)$$

где  $S$  – цена базисного актива.

Пусть срок до истечения опциона  $\tau = T - t_0$  разделен на  $n$  равных частей (периодов) продолжительностью  $\Delta t = \tau/n$ , таким образом, время  $t$  принимает дискретные значения  $t_i = t_0 + i \Delta t$ ,  $i = \overline{0, n}$ .

В модели CRR предполагается, что цена базисного актива  $S(t_i)$  в течение периода времени с началом в момент  $t_i$  либо увеличивается до значения  $u \cdot S(t_i)$ ,  $u > 1$ , либо уменьшается до значения  $d \cdot S(t_i)$ ,  $d < 1$ . Обозначая допустимые значения цены базисного актива в момент  $t_i$  через  $S_j^i$ , получим

$$S_j^i = S_0 u^j d^{i-j}, \quad i = \overline{0, n}, \quad j = \overline{0, i}. \quad (2)$$

Допустимые цены базисного актива  $S_j^i$  формируют биномиальное дерево (см. рис. 1).

Здесь для узлов биномиального дерева также использовано обозначение  $S_j^i$ .

В случае риск-нейтральности финансового рынка [13] указанные выше множители  $u$  и  $d$  равны [12]

$$u = e^{\sigma\sqrt{\Delta t}}, \quad d = e^{-\sigma\sqrt{\Delta t}}, \quad (3)$$

где  $\sigma$  – постоянная волатильность базисного актива, при этом (псевдо)вероятности увеличения и уменьшения цены базисного актива равны [12]

$$\pi_u = \frac{e^{r\Delta t} - d}{u - d}, \quad \pi_d = 1 - \pi_u = \frac{u - e^{r\Delta t}}{u - d}. \quad (4)$$

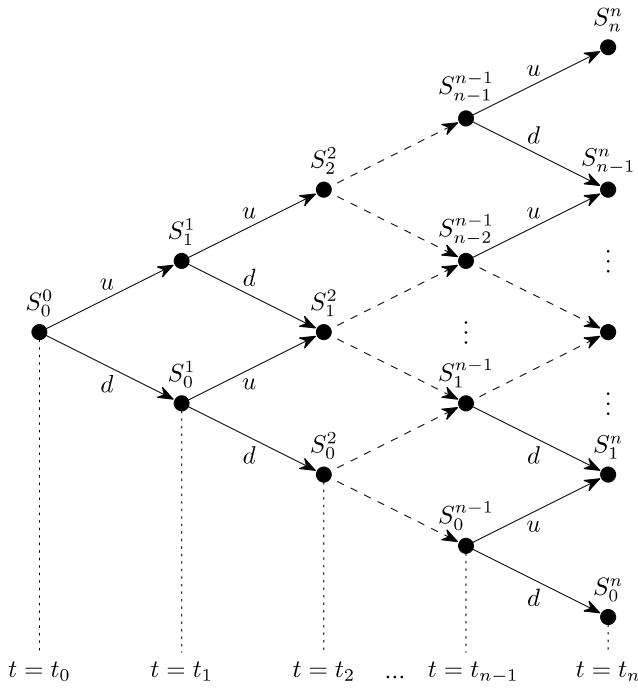


Рисунок 1. Допустимые цены базисного актива в биномиальной модели.

Построение биномиального дерева на рис. 1 производится начиная с корневого узла  $S_0^0$ , соответствующего текущей цене базисного актива ( $S_0^0 = S_0$ ) и моменту времени  $t_0$ , по направлению к листьям  $S_j^n, j = \overline{0, n}$ , соответствующим значениям базисного актива в момент погашения опциона  $T$ .

Биномиальное дерево также может быть построено и для моделей локальной волатильности, однако в этом случае процедура построения дерева является более сложной и неоднозначной [14].

### III. ОЦЕНКА ЕВРОПЕЙСКОГО ОПЦИОНА В БИНОМИАЛЬНОЙ МОДЕЛИ

Рассмотрим европейский опцион пут с ценой-страйк  $K$  и сроком погашения  $T$  на базисный актив, не приносящий дохода (дивидендов) на временном интервале  $[0, T]$ .

Европейский опцион пут исполняется в момент погашения  $T$  с платежной функцией (1). Обозначим через  $V_j^i$  стоимость европейского опциона в узле  $S_j^i$  биномиального дерева, а именно, в момент времени  $t_i$  при цене базисного актива  $S_j^i$ . Стоимость опциона в узлах биномиального дерева в момент погашения опциона (при  $t = T$ ) равна

$$V_j^n = \max(K - S_j^n, 0), j = \overline{0, n}. \quad (5)$$

Для момента времени  $t_i$ , предшествующего моменту погашения опциона (при  $i < n$ ), стоимость опциона  $V_j^i$  равна дисконтированной ожидаемой стоимости опциона в следующий момент времени  $t_{i+1}$ , а именно:

$$V_j^i = e^{-r\Delta t} (\pi_u V_{j+1}^{i+1} + \pi_d V_j^{i+1}), i = \overline{0, n-1}, j = \overline{0, i}. \quad (6)$$

Обратный метод вычисления текущей стоимости европейского опциона [12] состоит в том, чтобы, зная стоимость опциона в момент погашения для различных

допустимых значений цены базисного актива согласно формуле (5), находить дисконтированную ожидаемую стоимость опциона для предыдущих периодов времени, применяя формулу (6) достаточное число раз. В результате получим стоимость опциона  $V_0^0$  для начального момента времени  $t_0$  и текущей цены базисного актива  $S_0$ .

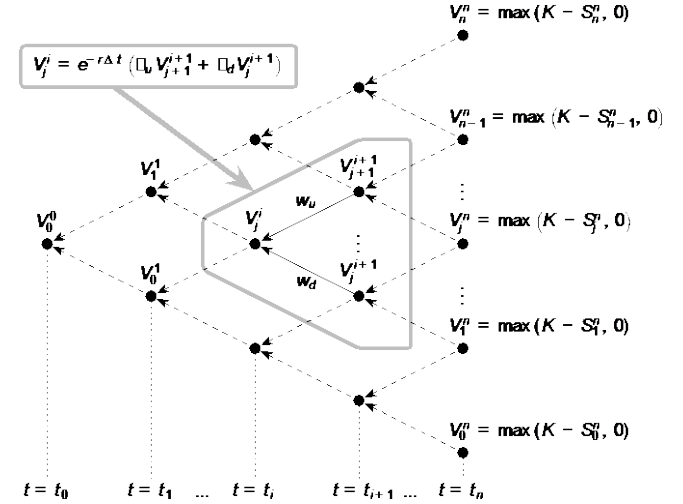


Рисунок 2. Расчет цены европейского опциона пут в биномиальной модели.

Таким образом, при вычислении текущей стоимости европейского опциона пут биномиальное дерево проходит в обратном порядке от листьев к корню дерева, как показано на рис. 2, с вычислением стоимости опциона в листьях дерева по формулам (5) и во внутренних узлах по формулам (6).

### IV. НЕЙРОННАЯ СЕТЬ ДЛЯ ОЦЕНКИ ЕВРОПЕЙСКИХ ОПЦИОНОВ

Построим на основе биномиального дерева цен (рис. 1) и алгоритма оценки европейского опциона пут (рис. 2) следующую глубокую нейронную сеть прямого распространения со сверточными слоями [15].

Линейная свертка (свертка 1D) [16] вектора  $x \in \mathbb{R}^m$  и фильтра  $w \in \mathbb{R}^k (m \geq k)$ , обозначаемая  $x * w$ , представляет собой вектор размерности  $n - k + 1$  с компонентами

$$(x * w)_i = \sum_{j=1}^k x_{i+j-1} w_j, i = \overline{1, m - k + 1}.$$

При  $m = k$  значением линейной свертки  $x * w$  является скалярное произведение векторов  $x$  и  $w$ .

Пусть входной слой  $\mathbb{L}^{n+1}$  состоит из одного узла, получающего на входе цену-страйк  $K$  оцениваемого опциона. Первый скрытый слой  $\mathbb{L}^n$  с  $n + 1$  нейроном получает от входного слоя  $\mathbb{L}^{n+1}$  в качестве входных данных цену-страйк  $K$ , которая затем преобразуется в выходной вектор  $\mathbf{V}^n \in \mathbb{R}^{n+1}$  при помощи нелинейной функции активации ReLU (Rectified Linear Unit) [17] и линейного преобразования с вектором весов  $\mathbf{w}^n$  и вектором смещений  $\mathbf{b}^n$

$$\mathbf{V}^n = \text{ReLU}(\mathbf{w}^n K + \mathbf{b}^n), \quad (7)$$

где  $\text{ReLU}(x) = \max(x, 0)$ ,  $\mathbf{w}^n \in \mathbb{R}^{n+1}$ ,  $\mathbf{b}^n \in \mathbb{R}^{n+1}$ .

В скрытых слоях  $\mathbb{L}^i, i = \overline{1, n-1}$ , содержащих  $i+1$  нейрон, выход  $\mathbf{V}^{i+1} \in \mathbb{R}^{i+2}$  предыдущего слоя  $\mathbb{L}^{i+1}$  преобразуется в выход  $\mathbf{V}^i \in \mathbb{R}^{i+1}$  слоя  $\mathbb{L}^i$  при помощи линейной свертки вектора  $\mathbf{V}^{i+1}$  и фильтра  $\mathbf{W} \in \mathbb{R}^2$

$$\mathbf{V}^i = \mathbf{V}^{i+1} * \mathbf{W} \quad (8)$$

и тождественной функции активации (id).

В выходном слое  $\mathbb{L}^0$  выход  $\mathbf{V}^1 \in \mathbb{R}^2$  последнего скрытого слоя  $\mathbb{L}^1$  преобразуется в выход  $y \in \mathbb{R}$  нейронной сети при помощи линейной свертки с фильтром  $\mathbf{W}$  и тождественной функции активации

$$y = \mathbf{V}^1 * \mathbf{W}. \quad (9)$$

Построенная таким образом нейронная сеть  $\mathcal{L}$  изображена на рис. 3 и содержит (кроме входного слоя  $\mathbb{L}^{n+1}$ , не производящего каких-либо действий над данными) один полносвязный слой  $\mathbb{L}^n$  с функцией активации ReLU и  $n$  сверточных слоев  $\mathbb{L}^{n-1}, \dots, \mathbb{L}^0$  с тождественной функцией активации.

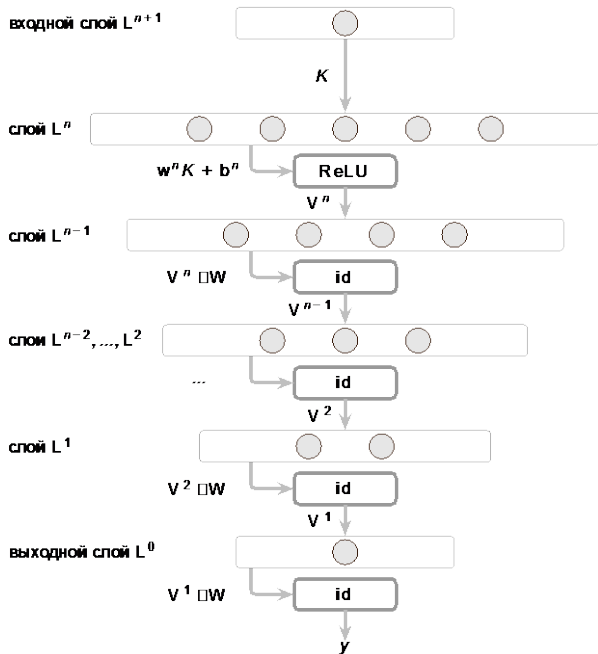


Рисунок 3. Структура нейронной сети для оценки европейского опциона пут.

Параметрами нейронной сети  $\mathcal{L}$  являются веса  $\mathbf{w}^n$  и смещения  $\mathbf{b}^n$  полносвязного слоя  $\mathbb{L}^n$ , а также веса фильтра  $\mathbf{W}$ , применяемого во всех сверточных слоях  $\mathbb{L}^{n-1}, \dots, \mathbb{L}^0$ .

**Теорема 1.** Нейронная сеть прямого распространения  $\mathcal{L}$ , состоящая из входного слоя  $\mathbb{L}^{n+1}$  и  $n+1$  слоев  $\mathbb{L}^n, \mathbb{L}^{n-1}, \dots, \mathbb{L}^1, \mathbb{L}^0$ , где:

- первый скрытый слой  $\mathbb{L}^n$  – полносвязный слой, состоящий из  $n+1$  нейрона, с функцией активации ReLU, вектором весов  $\mathbf{w}^n \in \mathbb{R}^{n+1}$  и вектором смещений  $\mathbf{b}^n \in \mathbb{R}^{n+1}$ ,
  - скрытые и выходной слой  $\mathbb{L}^i, i = n-1, \dots, 0$  – сверточные слои, состоящие из  $i+1$  нейрона, с фильтром  $\mathbf{W} \in \mathbb{R}^2$ , нулевым смещением и тождественной функцией активации,
- при поступлении на вход цены-страйк  $K$  выдает на

выходе аппроксимацию стоимости опциона, совпадающую с оценкой европейского опциона пут в модели CRR, если параметры  $\mathbf{w}^n, \mathbf{b}^n, \mathbf{W}$  нейронной сети  $\mathcal{L}$  выбраны (обучены) следующим образом:

$$\mathbf{w}^n = (1, \dots, 1) \in \mathbb{R}^{n+1}, \quad (10)$$

$$\mathbf{b}^n = (-S_0^n, \dots, -S_0^n), S_j^n = S_0 e^{\sigma\sqrt{\Delta t}(2j-n)}, j = \overline{0, n}, \quad (11)$$

$$\mathbf{W} = \left( \frac{e^{-r\Delta t}u-1}{u-d}, \frac{1-e^{-r\Delta t}d}{u-d} \right), u = e^{\sigma\sqrt{\Delta t}}, d = e^{-\sigma\sqrt{\Delta t}}. \quad (12)$$

*Доказательство теоремы 1.* Если на вход нейронной сети  $\mathcal{L}$  поступает цена-страйк  $K$ , то результатом вычислений в слое  $\mathbb{L}^n$  будет вектор  $\mathbf{V}^n$ , определяемый согласно (7). Если векторы  $\mathbf{w}^n$  и  $\mathbf{b}^n$  задаются согласно (10) и (11) соответственно, то вектор  $\mathbf{V}^n$  будет иметь компоненты, равные стоимости опциона в узлах (листьях) биномиального дерева в момент исполнения опциона (при  $t = T$ ) согласно формулам (5).

В сверточных слоях  $\mathbb{L}^i, i = \overline{0, n-1}$  для входного вектора  $\mathbf{V}^{i+1}$  будет выполняться линейная свертка с фильтром  $\mathbf{W}$  согласно (8) и (9). Если фильтр  $\mathbf{W}$  задается формулой (12), то операция свертки в слое нейронной сети будет эквивалентна операции дисконтирования ожидаемой стоимости опциона в биномиальном дереве по формуле (6).

Таким образом, выход нейронной сети  $\mathcal{L}$  будет равен ожидаемой дисконтированной цене опциона для начального момента времени  $t_0$  и текущей цены базисного актива  $S_0$  в соответствии с моделью CRR.  $\square$

Замечание 1. Выход  $y \in \mathbb{R}$  нейронной сети  $\mathcal{L}$  можно связать с входом  $K \in \mathbb{R}$  при помощи следующей явной формулы

$$y = \text{ReLU}(\mathbf{w}^n K + \mathbf{b}^n) * \underbrace{\mathbf{W} * \dots * \mathbf{W}}_{n \text{ раз}} \quad (13)$$

Для фильтра  $\mathbf{W} = (w_1, w_2)$  по индукции можно доказать, что справедлива формула

$$\underbrace{\mathbf{W} * \dots * \mathbf{W}}_{n \text{ раз}} = (w_1^n, n w_1^{n-1} w_2, \dots, \binom{n}{j} w_1^{n-j} w_2^j, \dots, w_2^n),$$

где  $\binom{n}{j} = \frac{n!}{(n-j)! j!}$  – число сочетаний из  $n$  по  $j$ .

Поэтому нейронная сеть  $\mathcal{L}$  из  $n+1$  слоев может быть свернута в нейронную сеть  $\bar{\mathcal{L}}$  из двух слоев:

1) первого скрытого слоя  $\mathbb{L}^n$  – полносвязного слоя, состоящего из  $n+1$  нейрона, с функцией активации ReLU, векторами весов  $\mathbf{w}^n \in \mathbb{R}^{n+1}$  и смещений  $\mathbf{b}^n \in \mathbb{R}^{n+1}$ ,

2) второго (выходного) полносвязного слоя  $\bar{\mathbb{L}}$ , состоящего из одного нейрона, с тождественной функцией активации, вектором весов  $\bar{\mathbf{w}} \in \mathbb{R}^{n+1}$  и (нулевым) смещением,

в которой выход  $y \in \mathbb{R}$  связан с входом  $K \in \mathbb{R}$  соотношением

$$y = \bar{\mathbf{w}} \text{ReLU}(\mathbf{w}^n K + \mathbf{b}^n). \quad (14)$$

## V. ОЦЕНКА АМЕРИКАНСКОГО ОПЦИОНА В БИНОМИАЛЬНОЙ МОДЕЛИ

Биномиальная модель может использоваться для оценки

практически всех видов деривативов, включая американские опционы [12].

Если при помощи модели CRR оценивается американский опцион пут, то в методе оценки опциона на биномиальном дереве (рис. 4) произойдут следующие изменения [12]. Будем обозначать через  $\tilde{V}_j^i$  стоимость американского опциона пут в узле  $S_j^i$ .

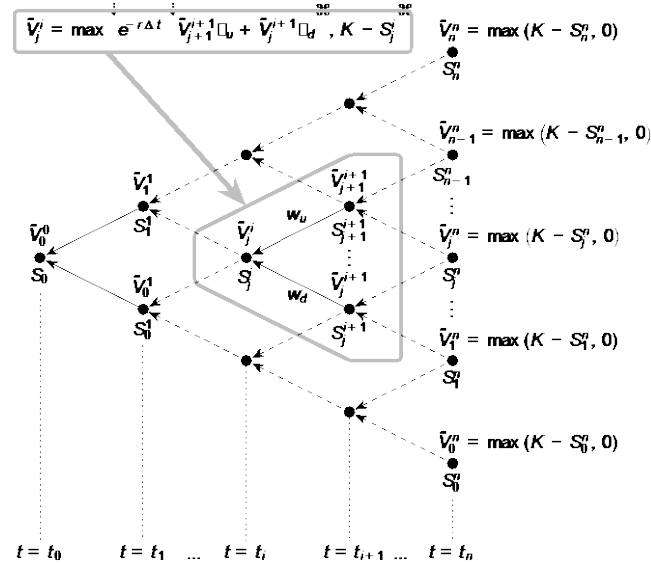


Рисунок 4. Расчет цены американского опциона пут в биномиальной модели.

Исполняя опцион в момент его погашения, держатель опциона получает платеж

$$\tilde{V}_j^n = \max(K - S_j^n, 0), j = \overline{0, n}, \quad (15)$$

равный стоимости опциона в соответствующем узле  $S_j^n$  биномиального дерева.

Для значений времени  $t_i$ , предшествующих сроку погашения опциона  $T$  (при  $i < n$ ), у держателя американского опциона есть две альтернативы: досрочно исполнить опцион или продолжать его удерживать.

Досрочное исполнение американского опциона пут имеет экономический смысл только в том случае, когда цена-страйк  $K$  больше текущей стоимости базисного актива  $S_j^i$  ( $K > S_j^i$ ), при этом держатель опциона получит платеж  $K - S_j^i$ , т.е.

$$\tilde{V}_j^i = K - S_j^i.$$

При удержании опциона согласно формуле (6) его стоимость в узле  $S_j^i$  будет равна дисконтированной ожидаемой стоимости опциона в момент времени  $t_{i+1}$ :

$$\tilde{V}_j^i = e^{-r\Delta t}(\tilde{V}_{j+1}^{i+1}\pi_u + \tilde{V}_{j+1}^{i+1}\pi_d).$$

Рациональный держатель американского опциона будет выбирать стратегию, максимизирующую финансовый результат, поэтому стоимость опциона в узле  $S_j^i$  составит

$$\tilde{V}_j^i = \max(e^{-r\Delta t}(\tilde{V}_{j+1}^{i+1}\pi_u + \tilde{V}_{j+1}^{i+1}\pi_d), K - S_j^i). \quad (16)$$

Обратный метод вычисления цены американского

опциона пут [12] состоит в том, чтобы, зная стоимость опциона в момент погашения согласно (15), в каждом промежуточном узле  $S_j^i$  находить максимум из дисконтированной ожидаемой стоимости опциона в следующий период времени и стоимости опциона при досрочном исполнении согласно (16). Применяя формулу (16) достаточное число раз, получим текущую цену американского опциона пут  $\tilde{V}_0^0$  для начального момента времени  $t_0$  и текущей цены базисного актива  $S_0$ .

## VI. НЕЙРОННАЯ СЕТЬ ДЛЯ ОЦЕНКИ АМЕРИКАНСКИХ ОПЦИОНОВ

Метод оценки американских опционов в модели CRR можно также интерпретировать в терминах глубокого обучения с определенными изменениями в архитектуре нейронной сети.

Нейронная сеть для оценки американского опциона пут  $\tilde{L}$  (рис. 5) будет состоять из входного слоя с одним узлом,  $n$  скрытых слоев с количеством нейронов, уменьшающимся от  $n+1$  до двух, и выходного слоя с одним нейроном. Обозначим слои нейронной сети символами  $\tilde{L}^i, i = \overline{0, n+1}$  с индексом  $i = n+1$  для входного слоя и индексом  $i = 0$  для выходного слоя.

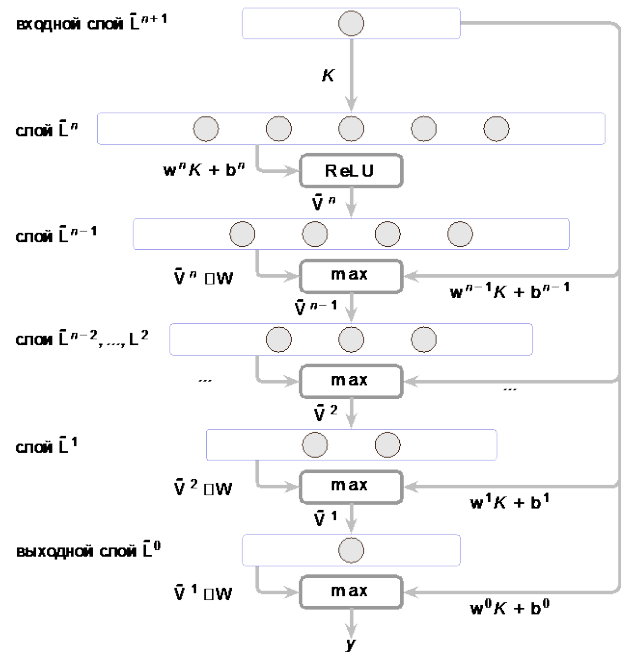


Рисунок 5. Структура нейронной сети для оценки американского опциона пут.

Первый скрытый слой  $\tilde{L}^n$  нейронной сети  $\tilde{L}$  аналогичен первому скрытому слою нейронной сети  $\tilde{L}$  и получает на входе значение цены-страйк  $K$ , которое преобразуется в выходной вектор  $\tilde{V}^n \in \mathbb{R}^{n+1}$  при помощи нелинейной функции активации ReLU и линейного преобразования с вектором весов  $\mathbf{w}^n \in \mathbb{R}^{n+1}$  и вектором смещений  $\mathbf{b}^n \in \mathbb{R}^{n+1}$

$$\tilde{V}^n = \text{ReLU}(\mathbf{w}^n K + \mathbf{b}^n). \quad (17)$$

В скрытом слое  $\tilde{L}^i, i = \overline{1, n-1}$ , выход  $\tilde{V}^{i+1} \in \mathbb{R}^{i+2}$  предыдущего слоя  $\tilde{L}^{i+1}$  и выход входного слоя  $K$

преобразуются в выходной вектор  $\tilde{\mathbf{V}}^i \in \mathbb{R}^{i+1}$  слоя  $\tilde{\mathbb{L}}^i$  по формуле

$$\tilde{\mathbf{V}}^i = \max(\tilde{\mathbf{V}}^{i+1} * \mathbf{W}, \mathbf{w}^i K + \mathbf{b}^i) \quad (18)$$

с применением линейной свертки вектора  $\tilde{\mathbf{V}}^{i+1}$  и фильтра  $\mathbf{W} \in \mathbb{R}^2$  и нелинейной функции активации maxout [18], где  $\mathbf{w}^i \in \mathbb{R}^{i+1}$ ,  $\mathbf{b}^i \in \mathbb{R}^{i+1}$ .

В выходном слое  $\tilde{\mathbb{L}}^0$  выход нейронной сети  $y \in \mathbb{R}$  формируется по формуле, аналогичной (22), а именно:

$$y = \max(\tilde{\mathbf{V}}^1 * \mathbf{W}, \mathbf{w}^0 K + \mathbf{b}^0), \quad (19)$$

где  $\mathbf{w}^0 \in \mathbb{R}$ ,  $\mathbf{b}^0 \in \mathbb{R}$ .

Построенная нейронная сеть  $\tilde{\mathcal{L}}$  содержит (кроме входного слоя  $\tilde{\mathbb{L}}^{n+1}$ , не производящего какие-либо действия над данными) один полносвязный слой  $\tilde{\mathbb{L}}^n$  с функцией активации ReLU и  $n$  слоев  $\tilde{\mathbb{L}}^{n-1}, \dots, \tilde{\mathbb{L}}^0$  с линейной сверткой и функцией активации maxout. Нейронная сеть  $\tilde{\mathcal{L}}$  не имеет последовательной структуры, так как слои  $\tilde{\mathbb{L}}^{n-1}, \dots, \tilde{\mathbb{L}}^0$  зависят как от выхода предыдущего слоя, так и от выхода входного слоя (значения цены-страйк  $K$  опциона). Параметрами нейронной сети  $\tilde{\mathcal{L}}$  являются веса  $\mathbf{w}^n, \dots, \mathbf{w}^0$  и смещения  $\mathbf{b}^n, \dots, \mathbf{b}^0$  слоев  $\tilde{\mathbb{L}}^n, \dots, \tilde{\mathbb{L}}^0$ , а также веса фильтра  $\mathbf{W}$ , применяемого в слоях  $\tilde{\mathbb{L}}^{n-1}, \dots, \tilde{\mathbb{L}}^0$ .

**Теорема 2.** Нейронная сеть прямого распространения  $\tilde{\mathcal{L}}$ , состоящая из входного слоя  $\tilde{\mathbb{L}}^{n+1}$  и  $n+1$  слоя  $\tilde{\mathbb{L}}^n, \dots, \tilde{\mathbb{L}}^0$ , где:

- первый скрытый слой  $\tilde{\mathbb{L}}^n$  – полносвязный слой, состоящий из  $n+1$  нейрона, с функцией активации ReLU, вектором весов  $\mathbf{w}^n \in \mathbb{R}^{n+1}$  и вектором смещений  $\mathbf{b}^n \in \mathbb{R}^{n+1}$ ,
- скрытые и выходной слои  $\tilde{\mathbb{L}}^i$ ,  $i = n-1, \dots, 0$ , состоящие из  $i+1$  нейрона, с функцией активации maxout, применяемой к линейной свертке выхода  $\tilde{\mathbf{V}}^{i+1}$  предыдущего слоя  $\tilde{\mathbb{L}}^{i+1}$  с фильтром  $\mathbf{W}$  и линейному преобразованию выхода  $K$  входного слоя с вектором весов  $\mathbf{w}^i$  и вектором смещений  $\mathbf{b}^i$ , при поступлении на вход цены-страйк  $K$  выдает на выходе аппроксимацию стоимости опциона, совпадающую с оценкой американского опциона пут в модели CRR, если параметры  $\mathbf{w}^n, \dots, \mathbf{w}^0$ ,  $\mathbf{b}^n, \dots, \mathbf{b}^0$ ,  $\mathbf{W}$  нейронной сети  $\tilde{\mathcal{L}}$  выбраны (обучены) следующим образом:

$$\mathbf{w}^i = (1, \dots, 1) \in \mathbb{R}^{i+1}, \quad i = \overline{0, n}, \quad (20)$$

$$\mathbf{b}^i = (-S_0^i, \dots, -S_n^i), \quad S_j^i = S_0 e^{\sigma\sqrt{\Delta t}(2j-i)}, \quad j = \overline{0, i}, \quad (21)$$

$$\mathbf{W} = \left( \frac{e^{-r\Delta t}u-1}{u-d}, \frac{1-e^{-r\Delta t}d}{u-d} \right), \quad u = e^{\sigma\sqrt{\Delta t}}, \quad d = e^{-\sigma\sqrt{\Delta t}}. \quad (22)$$

*Доказательство теоремы 2.* Если на вход нейронной сети  $\tilde{\mathcal{L}}$  поступает цена-страйк  $K$ , то результатом вычислений в слое  $\tilde{\mathbb{L}}^n$  будет вектор  $\tilde{\mathbf{V}}^n$ , определяемый согласно (21). Если векторы  $\mathbf{w}^n$  и  $\mathbf{b}^n$  задаются согласно (20) и (21) соответственно, то вектор  $\tilde{\mathbf{V}}^n$  будет иметь компоненты, равные стоимости опциона в узлах (листьях) биномиального дерева в момент исполнения опциона (при  $t = T$ ) согласно формулам (15).

В слоях  $\tilde{\mathbb{L}}^i$ ,  $i = \overline{0, n-1}$  для входного вектора  $\tilde{\mathbf{V}}^{i+1}$

будет вычисляться линейная свертка  $\tilde{\mathbf{V}}^{i+1} * \mathbf{W}$ , для входного значения  $K$  будет выполняться линейное преобразование  $\mathbf{w}^i K + \mathbf{b}^i$  и затем вычисляться выход слоев по формулам (18) и (19). Если фильтр  $\mathbf{W}$  задается формулой (22), а  $\mathbf{w}^i$  и  $\mathbf{b}^i$  – формулами (20) и (21) соответственно, то применение слоя  $\tilde{\mathbb{L}}^i$  к входным значениям  $\tilde{\mathbf{V}}^{i+1}$  и  $K$  будет эквивалентно операции дисконтирования ожидаемой стоимости опциона в биномиальном дереве по формуле (16) для интервала времени  $[t_i, t_{i+1}]$ .

Таким образом, выходное значение нейронной сети  $\tilde{\mathcal{L}}$  получается последовательным применением слоев  $\tilde{\mathbb{L}}^n, \tilde{\mathbb{L}}^{n-1}, \dots, \tilde{\mathbb{L}}^1, \tilde{\mathbb{L}}^0$  к значению цены-страйк  $K$  и будет равно ожидаемой дисконтированной цене американского опциона пут в методе CRR для начального момента времени  $t_0$  и текущей цены базисного актива  $S_0$ .  $\square$

## VII. ОБУЧЕНИЕ НЕЙРОННОЙ СЕТИ

В качестве иллюстрации обучим нейронную сеть для оценки европейских опционов пут со следующими параметрами:

$$S_0 = 0.5, \quad t_0 = 0, \quad T = 1, \quad r = 0.05, \quad \sigma = 0.25.$$

Количество периодов времени  $n$  в модели CRR выбрано равным 9, поэтому нейронная сеть  $\tilde{\mathcal{L}}$  будет состоять из 11 слоев. Цены-страйк  $K$  будут выбираться случайным образом из диапазона  $[0.25, 0.75]$  пакетами по 500 значений. Соответствующие ценам-страйк цены европейских опционов пут получены при помощи формулы Блэка-Шоулза [1]. Обучение нейронной сети состоит из 200 эпох по 10 шагов градиентного спуска для каждого пакета значений цен-страйк и соответствующих цен опционов. В качестве функции потерь использована среднеквадратичная ошибка – среднее арифметическое квадратов разностей между ценами опционов, предсказанными нейронной сетью, и ценами опционов согласно формуле Блэка-Шоулза [1].

Для реализации нейронной сети использован фреймворк TensorFlow [19]. В качестве алгоритма оптимизации нейронной сети выбран оптимизатор Adam [20] с начальным шагом обучения  $\alpha = 0.01$ .

Результаты обучения нейронной сети показывают, что предсказанные нейронной сетью значения цен европейских опционов пут (метка NN) близки к ценам опционов согласно формуле Блэка-Шоулза (метка BS) и ценам опционов в модели CRR (метка BT) (см. рис. 6).

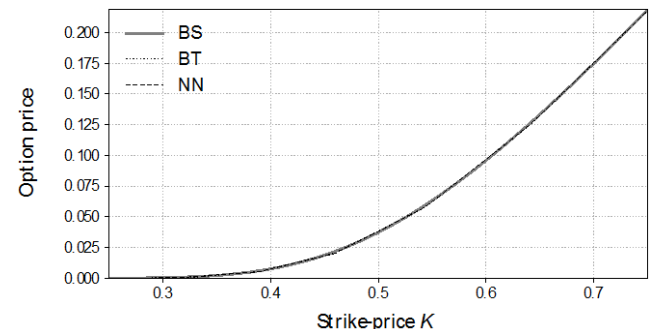


Рисунок 6. Цены европейских опционов пут в моделях Блэка-Шоулза, CRR и нейросетевой модели.



Однако при рассмотрении отклонений предсказанных цен опционов от точных значений согласно формуле Блэка-Шоулза (см. рис. 7) видно, что несмотря на обучение нейронной сети на точных теоретических значениях цен опционов (с точностью до ошибок при проведении операций с плавающей запятой) предсказанные нейронной сетью цены (метка NN) скорее соответствуют ценам опционов в модели CRR (метка BT), чем точным ценам опционов в модели Блэка-Шоулза (метка BS). Это связано с тем, что архитектура нейронной сети основана на биномиальном дереве и модели CRR, поэтому нейронная сеть склонна воспроизводить цены опционов в модели CRR.

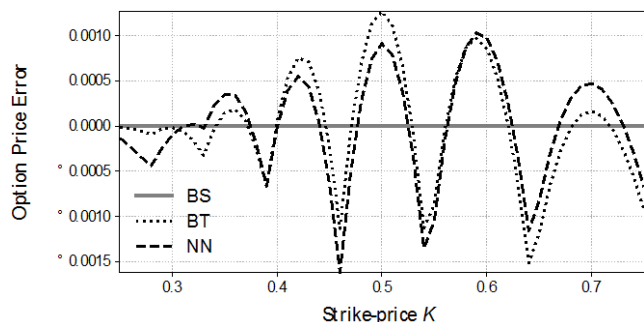


Рисунок 7. Отклонение цен европейских опционов пут в биномиальной и нейросетевой моделях.

## VIII. ЗАКЛЮЧЕНИЕ

В работе показано, что вычисление справедливых стоимостей европейских и американских опционов пут при помощи модели CRR может быть представлено в виде вычисления выхода нейронных сетей прямого распространения специальной архитектуры. Для европейского опциона пут нейронная сеть имеет последовательную структуру с первым скрытым полносвязным слоем с функцией активации ReLU и последующими сверточными слоями с 1D фильтром размера 2 и тождественной функцией активации. Для американского опциона пут нейронная сеть не имеет последовательной структуры, при этом в первом скрытом полносвязном слое также применяется функция активации ReLU, а в последующих слоях используется функции активации maxout, которая применяется к свертке выхода предыдущего слоя с 1D фильтром размера 2 и линейному преобразованию выхода входного слоя (цены-страйк опциона).

Теория опционов имеет обширные применения в финансах, начиная от оценки справедливой стоимости проектов, бизнеса и ценных бумаг и заканчивая оценкой и управлением рыночными [21] и кредитными [22] рисками. Предложенные в работе архитектуры нейронных сетей могут быть доработаны для использования в моделях локальной волатильности и применены при решении различных задач в финансах.

## БИБЛИОГРАФИЯ

[1] Black F., Scholes M. The Pricing of Options and Corporate Liabilities // Journal of Political Economy, vol. 81, pp. 637–654, 1973. doi:10.1086/260062  
 [2] Merton R. C. Theory of Rational Option Pricing // The Bell Journal of Economics and Management Science, vol. 4, pp. 141–183, 1973. doi:10.2307/3003143

[3] Schwartz E. The Valuation of Warrants: Implementing a New Approach // Journal of Financial Economics, vol. 4, pp. 79–94, 1977. doi:10.1016/0304-405X(77)90037-X  
 [4] Carr P., Madan D. Option valuation using the fast Fourier transform // The Journal of Computational Finance, vol. 2(4), pp. 61–73, 1999. doi:10.21314/jcf.1999.043  
 [5] Boyle P. Options: A Monte Carlo Approach // Journal of Financial Economics, vol. 4(3), pp. 323–338, 1977. doi:10.1016/0304-405X(77)90005-8  
 [6] Schmidhuber J. Deep learning in neural networks: An overview // Neural Networks, vol. 61, pp. 85–117, 2015. doi:10.1016/j.neunet.2014.09.003  
 [7] Malliaris M., Salchenberger L. A neural network model for estimating option prices // Applied Intelligence, vol. 3(3), pp. 193–206, 1993. doi:10.1007/BF00871937  
 [8] Hutchinson J. M. et al. A nonparametric approach to pricing and hedging derivative securities via learning networks // Journal of Finance, vol. 49(3), pp. 851–889, 1994. doi:10.3386/w4718  
 [9] Liu S., Oosterlee C. W., Bohte S. M. Pricing Options and Computing Implied Volatilities using Neural Networks // Risks, vol. 7, p.16, 2019. doi:10.3390/risks7010016  
 [10] Dupire B. Pricing with a smile // Risk Magazine, vol. 7(1), pp. 18–20, 1994.  
 [11] Shorokhov S. On Deep Learning for Option Pricing in Local Volatility Models // In: Proceedings of the 9th International Conference "Distributed Computing and Grid Technologies in Science and Education"(GRID'2021), Dubna, Russia, July 5-9, 2021. pp. 381–386. doi: 10.54546/mlit.2021.17.84.001.  
 [12] Cox J. C., Ross S. A., Rubinstein M. Option pricing: A simplified approach // Journal of Financial Economics, vol. 7(3), pp. 229–263, 1979. doi:10.1016/0304-405X(79)90015-1  
 [13] Cox J. C., Ross S. A. The Valuation of Options for Alternative Stochastic Processes // Journal of Financial Economics, vol. 3(1-2), pp. 145–166, 1976. doi:10.1016/0304-405X(76)90023-4  
 [14] Chriss N. Black Scholes and Beyond: Option Pricing Models. McGraw Hill Professional, 1997. 496 c.  
 [15] Lecun Y., Bottou L., Bengio Y., Haffner P. Gradient-based learning applied to document recognition // Proceedings of the IEEE, vol. 86(11), pp. 2278–2324, 1998. doi: 10.1109/5.726791  
 [16] Damelin S., Miller W. The Mathematics of Signal Processing. Cambridge University Press, 2011. doi:10.1017/cbo9781139003896  
 [17] Fukushima K. Visual feature extraction by a multilayered network of analog threshold elements // IEEE Transactions on Systems Science and Cybernetics, vol. 5(4), pp. 322–333, 1969. doi:10.1109/TSSC.1969.300225  
 [18] Goodfellow I., Warde-Farley D., Mirza M., Courville A., Bengio Y. Maxout Networks // Proceedings of the 30th International Conference on Machine Learning, PMLR, vol. 28(3), pp. 1319–1327, 2013.  
 [19] Abadi M. et al. TensorFlow: a system for large-scale machine learning // In: OSDI'16: Proceedings of the 12<sup>th</sup> USENIX conference on Operating Systems Design and Implementation, 2015. pp. 265–283.  
 [20] Kingma D. P., Ba J. Adam: A method for stochastic optimization // In: Y. Bengio, Y. LeCun (Eds.), 3rd International Conference on Learning Representations, ICLR 2015, San Diego, CA, USA, May 7-9, 2015, Conference Track Proceedings. 2015. URL <http://arxiv.org/abs/1412.6980>  
 [21] Шорохов С.Г. Введение в модели количественной оценки рыночных рисков // М.: РУДН, 2017.  
 [22] Шорохов С.Г. Введение в модели количественной оценки кредитных рисков // М.: РУДН, 2018.

Статья получена 15 ноября 2023. Публикация выполнена в рамках проекта №002092-0-000 Российского университета дружбы народов им. Патриса Лумумбы.

С. Г. Шорохов – Федеральное государственное автономное образовательное учреждение высшего образования "Российский университет дружбы народов имени Патриса Лумумбы, Москва, Россия (e-mail: shorokhov-sg@rudn.ru).  
 доцент кафедры математического моделирования и искусственного интеллекта, кандидат физико-математических наук, доцент

# Derivative pricing with a neural network based on the binomial tree

S. G. Shorokhov

**Abstract** — Based on the Cox-Ross-Rubinstein option pricing model, forward-propagation neural network architectures are built to approximate the value of European and American put options. For the European put option, the first hidden layer of the neural network is the dense layer with the ReLU activation function, the subsequent layers are convolutional with a 1D filter of dimension two and the identity activation function. For the American put option, the first hidden layer of the neural network is also dense with a ReLU activation function, subsequent layers have a maxout activation function and depend on both the output of the previous layer and the output of the input layer (the strike price value). The layers of both neural networks have the number of neurons decreasing by one with each subsequent layer up to one neuron in the output layer. It is shown that the neural network for the European put option can be simplified to a two-layer feed-forward neural network.

**Keywords** — option pricing, binomial tree, CRR model, convolution neural network.

## REFERENCES

- [1] Black F., Scholes M. The Pricing of Options and Corporate Liabilities // Journal of Political Economy, vol. 81, pp. 637–654, 1973. doi:10.1086/260062
- [2] Merton R. C. Theory of Rational Option Pricing // The Bell Journal of Economics and Management Science, vol. 4, pp. 141–183, 1973. doi:10.2307/3003143
- [3] Schwartz E. The Valuation of Warrants: Implementing a New Approach // Journal of Financial Economics, vol. 4, pp. 79–94, 1977. doi:10.1016/0304-405X(77)90037-X
- [4] Carr P., Madan D. Option valuation using the fast Fourier transform // The Journal of Computational Finance, vol. 2(4), pp. 61–73, 1999. doi:10.21314/jcf.1999.043
- [5] Boyle P. Options: A Monte Carlo Approach // Journal of Financial Economics, vol. 4(3), pp. 323–338, 1977. doi:10.1016/0304-405X(77)90005-8
- [6] Schmidhuber J. Deep learning in neural networks: An overview // Neural Networks, vol. 61, pp. 85–117, 2015. doi:10.1016/j.neunet.2014.09.003
- [7] Malliaris M., Salchenberger L. A neural network model for estimating option prices // Applied Intelligence, vol. 3(3), pp. 193–206, 1993. doi:10.1007/BF00871937
- [8] Hutchinson J. M. et al. A nonparametric approach to pricing and hedging derivative securities via learning networks // Journal of Finance, vol. 49(3), pp. 851–889, 1994. doi:10.3386/w4718
- [9] Liu S., Oosterlee C. W., Bohte S. M. Pricing Options and Computing Implied Volatilities using Neural Networks // Risks, vol. 7, p.16, 2019. doi:10.3390/risks7010016
- [10] Dupire B. Pricing with a smile // Risk Magazine, vol. 7(1), pp. 18–20, 1994.
- [11] Shorokhov S. On Deep Learning for Option Pricing in Local Volatility Models // In: Proceedings of the 9th International Conference "Distributed Computing and Grid Technologies in Science and Education"(GRID'2021), Dubna, Russia, July 5-9, 2021. pp. 381–386. doi: 10.54546/mlit.2021.17.84.001.
- [12] Cox J. C., Ross S. A., Rubinstein M. Option pricing: A simplified approach // Journal of Financial Economics, vol. 7(3), pp. 229–263, 1979. doi:10.1016/0304-405X(79)90015-1
- [13] Cox J. C., Ross S. A. The Valuation of Options for Alternative Stochastic Processes // Journal of Financial Economics, vol. 3(1-2), pp. 145–166, 1976. doi:10.1016/0304-405X(76)90023-4
- [14] Chriss N. Black Scholes and Beyond: Option Pricing Models. McGraw Hill Professional, 1997. 496 c.
- [15] Lecun Y., Bottou L., Bengio Y., Haffner P. Gradient-based learning applied to document recognition // Proceedings of the IEEE, vol. 86(11), pp. 2278–2324, 1998. doi: 10.1109/5.726791
- [16] Damelin S., Miller W. The Mathematics of Signal Processing. Cambridge University Press, 2011. doi:10.1017/cbo9781139003896
- [17] Fukushima K. Visual feature extraction by a multilayered network of analog threshold elements // IEEE Transactions on Systems Science and Cybernetics, vol. 5(4), pp. 322–333, 1969. doi:10.1109/TSSC.1969.300225
- [18] Goodfellow I., Warde-Farley D., Mirza M., Courville A., Bengio Y. Maxout Networks // Proceedings of the 30th International Conference on Machine Learning, PMLR, vol. 28(3), pp. 1319–1327, 2013.
- [19] Abadi M. et al. TensorFlow: a system for large-scale machine learning // In: OSDI'16: Proceedings of the 12<sup>th</sup> USENIX conference on Operating Systems Design and Implementation, 2015. pp. 265–283.
- [20] Kingma D. P., Ba J. Adam: A method for stochastic optimization // In: Y. Bengio, Y. LeCun (Eds.), 3rd International Conference on Learning Representations, ICLR 2015, San Diego, CA, USA, May 7-9, 2015, Conference Track Proceedings. 2015. URL <http://arxiv.org/abs/1412.6980>
- [21] Shorokhov S.G. Introduction into quantitative models of market risk valuation (in Russian) // Moscow: RUDN, 2017.
- [22] Shorokhov S.G. Introduction into quantitative models of credit risk valuation (in Russian) // Moscow: RUDN, 2018.