

УДК 519.633.2

ЧИСЛЕННЫЙ АЛГОРИТМ РЕШЕНИЯ ЗАДАЧИ КОШИ ДЛЯ ДВУМЕРНОГО ПОЛНОСТЬЮ НЕЛИНЕЙНОГО ПАРАБОЛИЧЕСКОГО УРАВНЕНИЯ

© 2025 А.А. Чубатов

Автономная некоммерческая образовательная организация высшего образования
«Научно-технологический университет «Сириус» (Университет «Сириус»),
федеральная территория «Сириус», Россия

Статья поступила в редакцию 19.08.2025

Статья является продолжением исследования, представленного в работе [1]. Рассматривается задача Коши для двумерного полностью нелинейного уравнения, возникающего в задаче оптимизации инвестиционного портфеля на рынке Хестона. Сначала исходное уравнение включается в систему квазилинейных параболических уравнений. Решение задачи Коши для этой системы сводится к стохастической формулировке в терминах прямого-обратного стохастических дифференциальных уравнений (ПОСДУ). Затем ПОСДУ преобразуется к задаче стохастического оптимального управления, которая решается с использованием нейронной сети. Предложенный численный алгоритм протестирован на примере, для которого удается получить квази-явное решение.

Ключевые слова: полностью нелинейное параболическое уравнение, задача Коши, система квазилинейных параболических уравнений, прямое-обратное стохастические дифференциальные уравнения (ПОСДУ), задача стохастического оптимального управления, глубокое обучение, нейронные сети для прямых-обратных стохастических дифференциальных уравнений.

DOI: 10.37313/1990-5378-2025-27-5-230-239

EDN: ZJPVLI

Результаты получены при финансовой поддержке проекта, реализуемого в рамках государственной программы федеральной территории «Сириус» «Научно-технологическое развитие федеральной территории «Сириус» (Соглашение № 18-03 от 10.09.2024).

Благодарности. Выражаю особую благодарность д. ф.-м. н., проф. Белопольской Яне Исаевне за значимые замечания и советы при подготовке данной статьи.

ВВЕДЕНИЕ

В статье [2] описан подход, позволяющий построить вероятностную интерпретацию решения задачи Коши для полностью нелинейного параболического уравнения вида

$$\frac{\partial u}{\partial \tau} + f(x, u, \nabla u, \nabla^2 u) = 0, \quad u(T, x) = h(x), \quad (1)$$

где $\tau \in [T_1; T]$, $x \in R^d$, $u = u(\tau, x) \in R$, $\nabla u \in R^d$ – градиент u , $\nabla^2 u \in R^{d \times d}$ – гессиан u .

Мы рассматриваем локальные по времени решения ($\tau \in [T_1; T]$), т.к. в нелинейных задачах зачастую нет глобальных по времени решений и возникает «взрыв решения», при котором решение бесконечно возрастает за конечное время.

Подход, описанный в [2], использует возможность свести задачу Коши для полностью нелинейного уравнения к задаче Коши для системы квазилинейных параболических уравнений. Далее решение задачи Коши для этой системы сводится к решению системы прямого-обратного стохастических дифференциальных уравнений (ПОСДУ), которая позволяет получить вероятностное представление решения задачи Коши. Представленный подход является альтернативным по отношению к подходу, описанному в работе [3], в котором решение задачи Коши для полностью нелинейного уравнения сводится к обратному стохастическому дифференциальному уравнению (ОСДУ) 2-го порядка (так называемому 2BSDE).

В статье [2] были описаны общие идеи подхода и выведено двумерное полностью нелинейное уравнение, возникающее в задаче оптимизации портфельных инвестиций на рынке Хестона. В статье [1] подход был модифицирован в части архитектуры нейронной сети (использовалась сеть типа FBSNN).

Данная статья продолжает исследования, представленные в статье [1], в которой вероятностный подход применялся для решения одномерного полностью нелинейного уравнения, возникающего в задаче оптимизации инвестиционного портфеля на рынке Блэка-Шоулса.

Чубатов Андрей Алексеевич, младший научный сотрудник Научного центра информационных технологий и искусственного интеллекта. E-mail: chaa@inbox.ru

1. МАТЕМАТИЧЕСКИЙ АППАРАТ

Опишем кратко подход к построению вероятностного представления решения задачи Коши (1), разработанный в статье [2].

Предположим, что функция $u = u(t, x)$ трижды дифференцируема по x и функции $f(x, u, v, \nabla v)$, $h(x)$ C^1 -гладкие и ограниченные по все аргументам. Тогда задача Коши (1) сводится к задаче Коши для следующей системы квазилинейных параболических уравнений

$$\frac{\partial u}{\partial \tau} + \frac{1}{2} \text{Tr}[B \nabla^2 u] + \langle a, \nabla u \rangle + \psi_0(x, u, v, \nabla v) = 0, \quad u(T, x) = h(x), \quad (2)$$

$$\frac{\partial v_k}{\partial \tau} + \frac{1}{2} \text{Tr}[B \nabla^2 v_k] + \langle a, \nabla v_k \rangle + \psi_k(x, u, v, \nabla v) = 0, \quad v_k(T, x) = \frac{\partial h(x)}{\partial x_k}, \quad k = 1, \dots, d, \quad (3)$$

где Tr – след матрицы, $\langle \cdot, \cdot \rangle$ – скалярное произведение векторов, $v = \nabla u \in R^d$,

$$B = B(x, u, v, \nabla v), \quad B_{ji} = 2 \frac{\partial f}{\partial [\nabla v]_{ij}},$$

$$a = a(x, u, v, \nabla v), \quad a_i = \frac{\partial f}{\partial v_i},$$

$$\psi_0(x, u, v, \nabla v) = f(x, u, v, \nabla v) - \frac{1}{2} \text{Tr}[B \nabla v] - \langle a, v \rangle,$$

$$\psi_k(x, u, v, \nabla v) = \frac{\partial f(x, u, v, \nabla v)}{\partial x_k} + \frac{\partial f(x, u, v, \nabla v)}{\partial u} v_k.$$

Отметим, что система (2)–(3) записана в виде, отличающемся от ее представления в статье [1]. Отличие заключается в выделении слагаемых вида $\langle a, \cdot \rangle$, которые понадобятся нам при применении данного подхода.

Замечание 1. Для применения вероятностного подхода достаточно, чтобы симметричная матрица B была положительно определенной и тем самым допускала разложение

$$B = AA^+, \quad A = A(x, u, v, \nabla v), \quad (4)$$

где $^+$ – операция транспонирования матрицы.

Разложение положительно определенной матрицы B в виде (4) не единственно. Существует бесконечно много матриц A , удовлетворяющих этому равенству. Матрица A определяется с точностью до умножения справа на ортогональную матрицу. Матрицу A можно найти с помощью разложения Холецкого, спектрального разложения или сингулярного разложения (SVD).

Будем считать, что $\frac{\partial f}{\partial [\nabla v]} \succ 0$, т.е. матрица B положительно определена.

Рассмотрим F_t -измеримый случайный процесс $X(t)$, $t \in [\tau; T]$, $\tau \in [T_1; T]$, и случайные процессы

$$U(t) = u(t, X(t)), \quad V(t) = v(t, X(t)), \quad H(t) = \nabla u(t, X(t)), \quad G(t) = \nabla v(t, X(t)),$$

где $(u(\tau, x), v(\tau, x))$ – классическое решение (2)–(3).

Зададим процесс $X(t)$, как процесс, удовлетворяющий СДУ

$$dX(t) = a(X(t), U(t), V(t), G(t))dt + A(X(t), U(t), V(t), G(t))dW(t), \quad X(\tau) = x, \quad (5)$$

где $W(t) \in R^d$ – стандартный винеровский случайный процесс, (τ, x) – точка, в которой мы хотим найти решение $(u(\tau, x), v(\tau, x))$ системы квазилинейных параболических уравнений (2)–(3), задает значение x процесса $X(t)$ в начальный момент времени τ .

Согласно формуле Ито, процессы $U(t)$ и $V(t)$ удовлетворяют следующим обратным СДУ (ОСДУ)

$$dU(t) = -\psi_0(X(t), U(t), V(t), G(t))dt + H(t)^+ A(X(t), U(t), V(t), G(t))dW(t), \quad U(T) = h(X(T)), \quad (6)$$

$$dV(t) = -\psi(X(t), U(t), V(t), G(t))dt + G(t)^+ A(X(t), U(t), V(t), G(t))dW(t), \quad V(T) = \nabla h(X(T)). \quad (7)$$

Система уравнений (5)–(7) представляет собой прямое-обратные СДУ (ПОСДУ).

Решением ПОСДУ (5)–(7) называется набор процессов $(X(t), U(t), V(t), H(t), G(t))$, удовлетворяющих системе СДУ с заданными начально-терминальными условиями и согласованных с фильтрацией, порождённой винеровским процессом.

Отметим, что в общем случае система ПОСДУ (5)–(7) является сильно связанной, потому что коэффициенты в прямом СДУ (5) зависят от процессов $U(t)$, $V(t)$ и $G(t)$, что не позволяет отделить прямое СДУ в независимое уравнение и решить его отдельно от обратных СДУ.

Связь между решением ПОСДУ и решением задачи Коши для полностью нелинейного уравнения описывается следующей теоремой.

Теорема 1 [4]. Если существует единственное решение ПОСДУ (5)–(7) $(X(t), U(t), V(t), H(t), G(t))$, то $(u(\tau, x), v(\tau, x)) = (U(\tau), V(\tau))$ – вязкостное решение задачи Коши для системы (2)–(3), а $u(\tau, x) = U(\tau)$ – вязкостное решение (1).

Определение вязкостного решения дифференциального уравнения было дано в статье [5]. В статье [4] определение вязкостного решения было расширено на системы уравнений.

2. ПОЛНОСТЬЮ НЕЛИНЕЙНОЕ УРАВНЕНИЕ, ВОЗНИКАЮЩЕЕ НА РЫНКЕ ХЕСТОНА

Задача оптимизации портфельных инвестиций на рынке Хестона приводит к решению задачи Коши для полностью нелинейного уравнения [2]

$$\frac{\partial u}{\partial \tau} - \frac{1}{2} \frac{((\mu - r)u_x + \rho \sigma u_{xy})^2}{yu_{xx}^2} u_{xx} + \frac{1}{2} \sigma^2 u_{yy} + rxu_x + \kappa(\theta - y)u_y = 0, \quad u(T, x, y) = h(x), \quad (8)$$

где $u = u(\tau, x, y) : [0, T] \times R \times R_+ \rightarrow R$, $\rho \in (-1, +1)$, $\kappa > 0$, $\theta > 0$, $\sigma > 0$, нижний индекс обозначает частную производную.

Будем считать, что для коэффициентов κ , θ , σ выполняется условие Феллера [6]

$$2\kappa\theta > \sigma^2, \quad (9)$$

которое является условием корректности для модели Хестона [7].

2.1. Вероятностное представление решения задачи Коши

Применим методику построения вероятностного представления решения, описанную в Разделе 1, к задаче (8). В этом случае $d = 2$, $(x, y) \in R^2$. Введем новую векторную функцию $v = v(\tau, x, y) \in R^2$

$$v = (v_1, v_2) = \nabla u = (u_x, u_y). \quad (10)$$

Включим уравнение (8) в систему квазилинейных параболических уравнений, которая должна состоять из трех уравнений для функций u , $v_1 = u_x$, $v_2 = u_y$. Учитывая обозначения (10), представим уравнение (8) в квазилинейном (относительно функций u, v_1, v_2) виде

$$\frac{\partial u}{\partial \tau} - \frac{1}{2} \frac{\alpha^2(y, v_1, v_{1y})}{yv_{1x}^2} u_{xx} + \frac{1}{2} \sigma^2 u_{yy} + rxu_x + \kappa(\theta - y)u_y = 0, \quad u(T, x, y) = h(x), \quad (11)$$

где $\alpha(y, v_1, v_{1y}) = (\mu - r)v_1 + \rho \sigma v_{1y}$.

Уравнение (11) является квазилинейным, т.к. коэффициент при u_{xx} зависит от функции v_1 и ее производных $\nabla v_1 = (v_{1x}, v_{1y})$. Заметим, что уравнение (11) не зависит от функции v_2 .

Для получения замкнутой системы необходимо к уравнению (11) добавить уравнение для функции v_1 и возможно еще уравнение для функции v_2 , если коэффициенты уравнения для функции v_1 будут зависеть от функции v_2 .

Для того чтобы записать уравнение для функции v_1 продифференцируем по x уравнение и терминальное условие (11). В итоге, учитывая, что $v_1 = u_x$, получим уравнение вида (3)

$$\frac{\partial v_1}{\partial \tau} + \frac{1}{2} \text{Tr}[B \nabla^2 v_1] + \langle a, \nabla v_1 \rangle + \psi_1(y, v_1, v_{1y}) = 0, \quad v_1(T, x, y) = h_x(x), \quad (12)$$

где $\text{Tr}[B \nabla^2 v_1] = B_{11}v_{1xx} + 2B_{12}v_{1xy} + B_{22}v_{1yy}$, $B = B(y, v_1, \nabla v_1) = \begin{pmatrix} B_{11} & B_{12} \\ B_{21} & B_{22} \end{pmatrix}$,

$$B_{11} = \frac{\alpha^2(y, v_1, v_{1y})}{y v_{1x}^2}, \quad B_{12} = B_{21} = -\frac{\rho \sigma \alpha(y, v_1, v_{1y})}{v_{1x}}, \quad B_{22} = \sigma^2 y, \quad a = a(x, y) = (rx, \kappa(\theta - y)),$$

$$\psi_1(y, v_1, v_{1y}) = r v_1 - (\mu - r) \frac{\alpha(y, v_1, v_{1y})}{y}.$$

Перепишем уравнение (11) в виде (2)

$$\frac{\partial u}{\partial \tau} + \frac{1}{2} \text{Tr}[B \nabla^2 u] + \langle a, \nabla u \rangle + \psi_0(y, v_1, \nabla v_1) = 0, \quad u(T, x) = h(x), \quad (13)$$

$$\text{где } \psi_0(y, v_1, \nabla v_1) = -(\mu - r) \frac{\alpha(y, v_1, v_{1y})}{y} \frac{v_1}{v_{1x}}.$$

Поскольку уравнения (12) и (13) не содержат функцию v_2 , то мы получили замкнутую систему квазилинейных параболических уравнений, состоящую всего из двух (а не трех, как ожидалось в общем случае) уравнений для функций u , v_1 (уравнение для функции v_2 не понадобилось).

Итак, мы свели полностью нелинейное уравнение (8) к системе квазилинейных параболических уравнений (12)–(13). Для записи вероятностного представления решения задачи (12)–(13) нужно проверить положительную определенность матрицы B .

Теорема 2. Пусть существует решение задачи (12)–(13) такое что

$$v_{1x} = u_{xx} \neq 0, \quad (14)$$

и выполняются условия

$$y > 0, \quad |\rho| \neq 1, \quad (15)$$

тогда матрица B является положительно определенной и допускает разложение (4), в котором матрица A имеет вид

$$A = A(y, v_1, \nabla v_1) = \begin{pmatrix} \frac{\alpha(y, v_1, v_{1y})}{\sqrt{y} v_{1x}} & 0 \\ -\rho \sigma \sqrt{y} & \sqrt{1 - \rho^2} \sigma \sqrt{y} \end{pmatrix}, \quad (16)$$

где $\alpha(y, v_1, v_{1y}) = (\mu - r)v_1 + \rho \sigma y v_{1y}$.

Доказательство.

Условия (15) выполнены благодаря ограничениям на параметры в модели Хестона (8): $y \in R_+$, $\rho \in (-1, +1)$. Условие (14) означает, что мы будем искать решение $u(\tau, x, y)$ исходного уравнения (8) в классе выпуклых функций.

Т.к. $v_{1x} \neq 0$, $y \in R_+$ и $|\rho| \neq 1$, то угловые миноры матрицы B положительны $B_{11} = \frac{\alpha^2(y, v_1, v_{1y})}{y v_{1x}^2} > 0$,

$\det B = \frac{\alpha^2(y, v_1, v_{1y})}{v_{1x}^2} \sigma^2 (1 - \rho^2) > 0$. Тогда, согласно критерию Сильвестра, симметричная матрица B

является положительно определенной.

Найдя элементы матрицы A с помощью разложения Холецкого или методом неопределенных коэффициентов, получим (16).

Запишем ПОСДУ, соответствующее системе квазилинейных параболических уравнений (12)–(13).

Рассмотрим F_t -измеримые случайные процессы $X(t) = (X_1(t), X_2(t)) \in R^2$, $t \in [\tau; T]$, $\tau \in [T_1; T]$, $U(t) = u(t, X(t)) \in R$, $V(t) = v_1(t, X(t)) \in R$, $H(t) = (H_1(t), H_2(t)) = \nabla u(t, X(t)) \in R^2$, $G(t) = (G_1(t), G_2(t)) = \nabla v_1(t, X(t)) \in R^2$.

Зададим процесс $X(t)$ с помощью следующего СДУ

$$dX(t) = a(X(t))dt + A(X(t), V(t), G(t))dW(t), \quad X(\tau) = (x, y), \quad (17)$$

где $W(t) \in R^2$ – стандартный винеровский случайный процесс, $a(X(t)) = (rX_1(t), \kappa(\theta - X_2(t))) \in R^2$,

$$A(X(t), V(t), G(t)) = \begin{pmatrix} \frac{(\mu - r)V(t) + \rho\sigma X_2(t)G_2(t)}{\sqrt{X_2(t)}G_1(t)} & 0 \\ -\rho\sigma\sqrt{X_2(t)} & \sqrt{1 - \rho^2}\sigma\sqrt{X_2(t)} \end{pmatrix}.$$

Тогда процессы $U(t)$ и $V(t)$ удовлетворяют следующим обратным СДУ (ОСДУ)

$$dU(t) = -\psi_0(X(t), V(t), G(t))dt + H(t)^+ A(X(t), V(t), G(t))dW(t), \quad U(T) = h(X(T)), \quad (18)$$

$$dV(t) = -\psi(X(t), V(t), G(t))dt + G(t)^+ A(X(t), V(t), G(t))dW(t), \quad V(T) = h_x(X(T)), \quad (19)$$

где $\psi_0(X(t), V(t), G(t)) = -(\mu - r) \frac{(\mu - r)V(t) + \rho\sigma X_2(t)G_2(t)}{X_2(t)} \frac{V(t)}{G_1(t)}$,

$$\psi(X(t), V(t), G(t)) = rV(t) - (\mu - r) \frac{(\mu - r)V(t) + \rho\sigma X_2(t)G_2(t)}{X_2(t)}.$$

Компоненты $U(t)$, $V(t)$ решения ПОСДУ (17)–(19) позволяют получить вероятностное представление решения задачи (12)–(13)

$$(u(\tau, x, y), v_1(\tau, x, y)) = (U(\tau), V(\tau)). \quad (20)$$

2.2. Численное решение ПОСДУ

Отметим, что система ПОСДУ (17)–(19) не является замкнутой, т.к. содержит 3 уравнения и 5 иско- мых случайных процессов. Одним из подходов к решению ПОСДУ является замыкание системы, допи- сыванием еще двух уравнений, с помощью теоремы Ито о представлении квадратично интегрируемого мартингала. Мы применяем другой способ, основанный на переходе к задаче оптимального управления [8, 9], в котором численное решение ПОСДУ сводится к численному решению задачи стохастического оптимального управления с использованием нейронных сетей.

Рассмотрим задачу оптимального управления, ассоциированную с ПОСДУ (17)–(19). В качестве управления возьмем процессы $U(t), V(t), H(t), G(t)$. Оптимизационная задача заключается в миними- зации следующей функции потерь

$$L = E \left[|u_T - h(X^C(T))|^2 + \int_{\tau}^T |u_t - U^C(t)|^2 dt + |v_T - h_x(X^C(T))|^2 + \int_{\tau}^T |v_t - V^C(t)|^2 dt \right], \quad (21)$$

где $E[\cdot]$ – математическое ожидание, $C = (u_t, v_t, h_t, g_t)$ – управление, u_t, v_t, h_t, g_t – F_t -измеримые случайные процессы, процессы $X^C(t)$, $U^C(t)$ и $V^C(t)$ удовлетворяют СДУ

$$X^C(t) = X(\tau) + \int_{\tau}^t a(X^C(s))ds + \int_{\tau}^t A(X^C(s), v_s, g_s)dW(s),$$

$$U^C(t) = u_{\tau} - \int_{\tau}^t \psi_0(X^C(s), v_s, g_s)ds + \int_{\tau}^t h_s^+ A(X^C(s), v_s, g_s)dW(s),$$

$$V^C(t) = v_{\tau} - \int_{\tau}^t \psi(X^C(s), v_s, g_s)ds + \int_{\tau}^t g_s^+ A(X^C(s), v_s, g_s)dW(s).$$

Для решения оптимизационной задачи мы используем нейронную сеть для прямых-обратных стоха- стических дифференциальных уравнений (Forward-Backward Stochastic Neural Network, FBSNN) [10, 1].

Аппроксимируем $(u(t, X(t)), v(t, X(t)))$ с помощью нейронной сети

$$(U^{\theta}(t), V^{\theta}(t)) = (u^{\theta}(t, X(t)), v^{\theta}(t, X(t))) = \text{net}(t, X(t); \theta),$$

где θ – параметры (весовые коэффициенты) нейронной сети.

Процедура автоматического дифференцирования [11] позволяет вычислить процессы $H^{\theta}(t) = Du^{\theta}(t, X(t))$ и $G^{\theta}(t) = Dv^{\theta}(t, X(t))$. Символ D обозначает операцию автоматического диф- ференцирования.

Рассмотрим разбиение промежутка $[\tau; T]$: $\tau = t_0 < t_1 < \dots < t_n < \dots < t_N = T$, $\Delta t = t_{n+1} - t_n = \frac{T - \tau}{N}$.

Введем обозначения для аппроксимаций случайных процессов

$$\tilde{U}_n = U^\theta(t_n), \quad \tilde{V}_n = V^\theta(t_n), \quad \tilde{H}_n = H^\theta(t_n), \quad \tilde{G}_n = G^\theta(t_n).$$

Разностная аппроксимация $L(\theta; N)$ функции потерь (21) имеет вид

$$L(\theta; N) = E \left[\left| \tilde{U}_N - h(X_N) \right|^2 + \left| \tilde{V}_N - h_x(X_N) \right|^2 + \Delta t \sum_{n=0}^{N-1} \left(\left| \tilde{U}_{n+1} - \bar{U}_{n+1} \right|^2 + \left| \tilde{V}_{n+1} - \bar{V}_{n+1} \right|^2 \right) \right], \quad (22)$$

где $\bar{U}_{n+1}$ и $\bar{V}_{n+1}$ – контрольные значения, вычисляемые по схеме Эйлера-Маруямы для ОСДУ (18)–(19)

$$\begin{aligned} \bar{U}_{n+1} &= \tilde{U}_n - \psi_0(X_n, \tilde{V}_n, \tilde{G}_n) \Delta t + \tilde{H}_n^+ A(X_n, \tilde{V}_n, \tilde{G}_n) \Delta W_n, \\ \bar{V}_{n+1} &= \tilde{V}_n - \psi(X_n, \tilde{V}_n, \tilde{G}_n) \Delta t + \tilde{G}_n^+ A(X_n, \tilde{V}_n, \tilde{G}_n) \Delta W_n, \end{aligned}$$

значения $\tilde{U}_{n+1}$, $\tilde{V}_{n+1}$, вычисляемые с помощью нейронной сети $(\tilde{U}_{n+1}, \tilde{V}_{n+1}) = \text{net}(t_{n+1}, X_{n+1}; \theta)$,

X_{n+1} вычисляется по схеме Эйлера-Маруямы для СДУ (17)

$$X_{n+1} = X_n + a(X_n) \Delta t + A(X_n, \tilde{V}_n, \tilde{G}_n) \Delta W_n, \quad X_0 = (x, y).$$

Нейронная сеть использует функцию потерь (22). Близкое к нулю значение функции потерь обеспечивает выполнение терминальных условий и разностных соотношений, определяемых ОСДУ (18), (19).

Для вычисления математического ожидания воспользуемся методом Монте-Карло. Сгенерируем M траекторий процесса $X^m(t)$, $m = 1, \dots, M$, выходящих из одной точки $X^m(t_0) = X_0 = (x, y)$. Тогда

оценка $\hat{L}(\theta; N, M)$ для функции потерь (22) примет вид

$$\hat{L}(\theta; N, M) = \frac{1}{M} \sum_{m=1}^M \left(\left| \tilde{U}_N^m - h(X_N^m) \right|^2 + \left| \tilde{V}_N^m - h_x(X_N^m) \right|^2 + \Delta t \sum_{n=0}^{N-1} \left(\left| \tilde{U}_{n+1}^m - \bar{U}_{n+1}^m \right|^2 + \left| \tilde{V}_{n+1}^m - \bar{V}_{n+1}^m \right|^2 \right) \right),$$

где

$$\begin{aligned} \bar{U}_{n+1}^m &= \tilde{U}_n^m - \psi_0(X_n^m, \tilde{V}_n^m, \tilde{G}_n^m) \Delta t + \tilde{H}_n^{m+} A(X_n^m, \tilde{V}_n^m, \tilde{G}_n^m) \Delta W_n^m, \\ \bar{V}_{n+1}^m &= \tilde{V}_n^m - \psi(X_n^m, \tilde{V}_n^m, \tilde{G}_n^m) \Delta t + \tilde{G}_n^{m+} A(X_n^m, \tilde{V}_n^m, \tilde{G}_n^m) \Delta W_n^m, \\ X_{n+1}^m &= X_n^m + a(X_n^m) \Delta t + A(X_n^m, \tilde{V}_n^m, \tilde{G}_n^m) \Delta W_n^m, \quad X_0^m = (x, y). \end{aligned}$$

3. ТЕСТОВЫЙ ПРИМЕР

3.1. Квази-явное решение

В общем случае задача (8) не имеет аналитического решения, но для некоторых терминальных условий получается свести решение задачи (8) к решению задачи Коши для линейного параболического уравнения, решение которого можно получить с помощью формулы Фейнмана-Каца [12, 13].

Рассмотрим случай экспоненциального терминального условия

$$h(x) = -e^{-\gamma x}, \quad \gamma > 0, \quad x \in R. \quad (23)$$

В этом случае и при дополнительном условии $r = 0$ удастся мультипликативно разделить переменные в решении задачи (8) и искать решение в виде

$$u(\tau, x, y) = -e^{-\gamma x} f(\tau, y)^\delta, \quad (24)$$

где δ – параметр, значение которого будет найдено ниже (в формуле (27)).

Учитывая (10), из $v_1(\tau, x, y) = u_x(\tau, x, y)$ следует, что

$$v_1(\tau, x, y) = \gamma e^{-\gamma x} f(\tau, y)^\delta. \quad (25)$$

Подставляя (23), (24) в задачу (8) при $r = 0$, после несложных, но громоздких преобразований получим, что функция $f(\tau, y)$ является решением квазилинейного параболического уравнения

$$\frac{\partial f}{\partial \tau} + \frac{1}{2} \sigma^2 y f_{yy} + \kappa(\theta_2 - y) f_y - \frac{1}{2} \frac{\mu^2}{\delta y} f - \frac{1}{2} \sigma^2 (1 - \delta(1 - \rho^2)) y \frac{(f_y)^2}{f} = 0, \quad f(T, y) = 1, \quad (26)$$

где $\theta_2 = \theta - \frac{\mu \rho \sigma}{\kappa}$.

Уравнение (26) является квазилинейным из-за последнего слагаемого, содержащего квадрат производной f_y .

Заметим, что при

$$\delta = \frac{1}{1-\rho^2} \quad (27)$$

квазилинейное уравнение (26) становится линейным

$$\frac{\partial f}{\partial \tau} + \frac{1}{2} \sigma^2 y f_{yy} + \kappa(\theta_2 - y) f_y - \frac{1}{2} \frac{\mu^2(1-\rho^2)}{y} f = 0, \quad f(T, y) = 1. \quad (28)$$

Решение задачи Коши (28) с помощью формулы Фейнмана-Каца имеет вид

$$f(\tau, y) = E \left[\exp \left(-\frac{1}{2} \mu^2 (1-\rho^2) \int_{\tau}^T \frac{1}{\xi(t)} dt \right) \middle| \xi(\tau) = y \right], \quad (29)$$

где $E[\cdot]$ – условное математическое ожидание, процесс $\xi(t)$ задается СДУ

$$d\xi(t) = \kappa(\theta_2 - \xi(t))dt + \sigma \sqrt{\xi(t)} dW(t), \quad \xi(\tau) = y.$$

Процесс $\xi(t)$ является процессом Кокса-Ингерсолла-Росса (CIR) [14] и условие Феллера (9), обеспечивающее отделимость от нуля значений процесса $\xi(t)$, имеет вид

$$2\kappa\theta_2 > \sigma^2.$$

Квази-явное решение задачи (28) вычисляется методом Монте-Карло, используя формулу (29). Значение $f(\tau, y)$ позволяет (для частного случая $r = 0$ и $h(x) = -e^{-\gamma x}$) по формулам (24), (25), (27), (29) получить квази-явное решение задачи (12)–(13) и, соответственно, задачи (8).

3.2. Численные эксперименты

Фреймворк FBSNN_tf1_FnL [1], изначально написанный с использованием библиотеки TensorFlow1, переписан, используя библиотеку TensorFlow2. Это упрощает задание модели нейронной сети с помощью библиотеки TensorFlow.Keras, делает код и его отладку интуитивно более простыми за счет энергичной (eager execution) (а не ленивой, сессионной) модели вычислений. Кроме того, в TensorFlow2 улучшена оптимизация вычислительных графов и поддерживается автоматическая компиляция в граф с помощью декоратора @tf.function, что ускоряет выполнение программного кода.

Решение системы (12)–(13) $(u(\tau, x, y), v_1(\tau, x, y))$ аппроксимируется с помощью нейронной сети, содержащей 3 скрытых слоя и 128 нейронов в каждом слое. Входной слой сети содержит 3 нейрона, соответствующих координатам τ, x, y . Выходной слой, задающий пару функций $(u(\tau, x, y), v_1(\tau, x, y))$, содержит 2 нейрона.

Для минимизации функции потерь нейронной сети $\hat{L}(\theta; N, M)$ был применен оптимизатор AdaDelta – адаптивный алгоритм градиентного спуска, который автоматически настраивает скорость обучения.

Квази-явное решение $(u^{exact}(\tau, x, y), v_1^{exact}(\tau, x, y))$ задачи (12)–(13), полученное по формулам (24), (25), (27), (29) используется для оценки точности численного решения (20), полученного с помощью нейронной сети. Относительная погрешность компонент численного решения вычисляется по формулам

$$err_{rel}^u(\tau, x, y) = \frac{|u(\tau, x, y) - u^{exact}(\tau, x, y)|}{|u^{exact}(\tau, x, y)|}, \quad err_{rel}^v(\tau, x, y) = \frac{|v_1(\tau, x, y) - v_1^{exact}(\tau, x, y)|}{|v_1^{exact}(\tau, x, y)|}.$$

На рисунке 1 приведены численное и точное (квази-явное) решения системы (12)–(13) для экспоненциального терминального условия (23) и $r = 0$ при следующих значениях параметров $\gamma = 0,5$, $\rho = 0$, $\mu = 0,02$, $\sigma = 0,1$, $\kappa = 1$, $\theta = 0,1$. Решение системы (12)–(13) $(u(\tau, x, y), v_1(\tau, x, y)) = (U(\tau), V(\tau))$, обозначаемое на рисунке черным кружочком, приведено в точке $(\tau, x, y) = (0,8; 0,5; 0,2)$. На рисунке 1 слева и справа приведены по 10 случайных траекторий процессов $U(t)$, $V(t)$ (сплошные линии), соответственно, и значения $u^{exact}(\tau, X_1(t), X_2(t))$, $v_1^{exact}(\tau, X_1(t), X_2(t))$ (пунктирные линии), вычисленные по формулам (24)–(25). Относительная погрешность компонент решения составляет, соответственно, 1,59% и 2,97%.

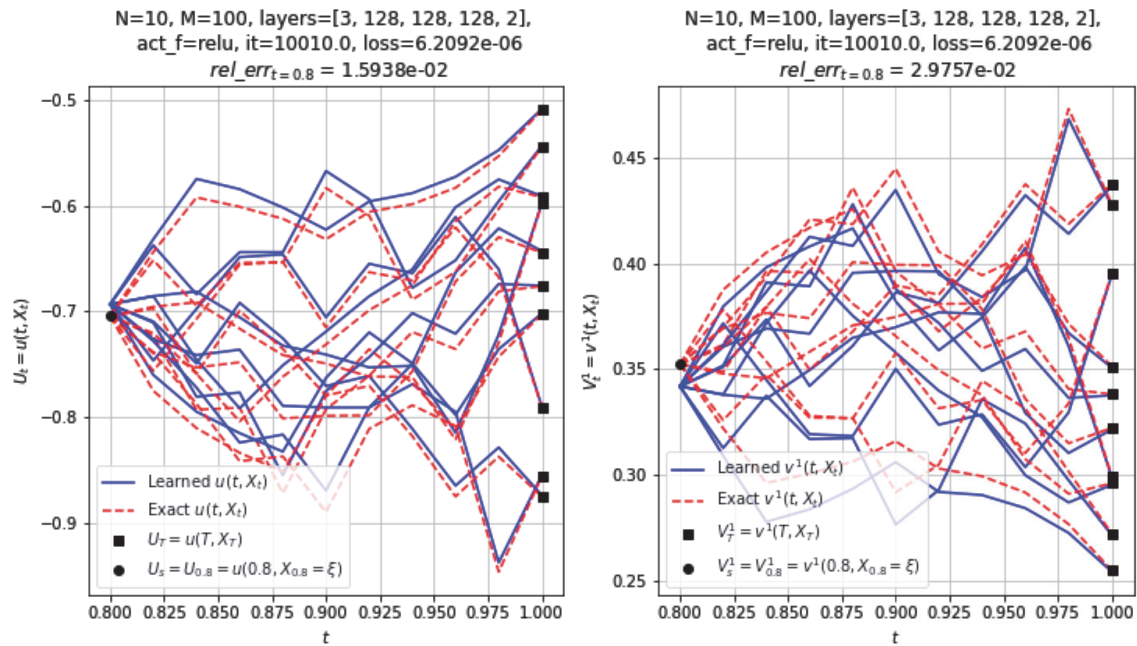


Рис. 1. Сравнение численного и точного решения системы (12)–(13) для терминального условия (23)

ЗАКЛЮЧЕНИЕ

В данной работе для решения задачи Коши для двумерного полностью нелинейного параболического уравнения (8) применен вероятностный подход, основанный на решении ПОСДУ. На первом этапе полностью нелинейное уравнение включается в систему квазилинейных параболических уравнений. Для полученной системы доказана положительная определенность матрицы коэффициентов перед старшими производными (матрицы B), что позволяет применить к решению задачи вероятностные методы. Далее решение задачи Коши для полученной системы сводится к решению системы прямого-обратного стохастических дифференциальных уравнений (ПОСДУ). Численное решение системы ПОСДУ в свою очередь сводится к решению задачи стохастического оптимального управления, для решения которой используется алгоритм, использующий нейронную сеть для аппроксимации решения. Для оценки точности полученного решения, алгоритм был протестирован на тестовом примере для частного случая уравнения (8): $h(x) = -e^{-\gamma x}$, $r = 0$, для которого существует квази-явное решение, вычисляемое с помощью формулы Фейнмана-Каца.

СПИСОК ЛИТЕРАТУРЫ

1. Чубатов, А.А. Численный алгоритм решения полностью нелинейных параболических уравнений на основе прямых-обратных стохастических дифференциальных уравнений и нейронных сетей / А.А. Чубатов // Известия Самарского научного центра Российской академии наук. – 2024. – Т. 26. – № 4. – С. 61–69.
2. Белопольская, Я.И. Оптимизация инвестиционного портфеля в модели Хестона / Я.И. Белопольская, А.А. Чубатов // Записки научных семинаров Санкт-Петербургского отделения Математического института им. В.А. Стеклова Российской академии наук. – 2023. – Т. 526. – С. 29–51.
3. Second order backward stochastic differential equations and fully non-linear parabolic PDEs / P. Cheridito, H. M. Soner, N. Touzi, N. Victoir // Communications on Pure and Applied Mathematics. 2007. V. 60. No. 7. Pp. 1081–1110.
4. Belopolskaya, Ya. I. Probabilistic counterparts of nonlinear parabolic partial differential equation systems // Modern stochastic and applications. Springer Optimization and Its Applications; vol. 90. / ed. by V. Korolyuk [et al.]. – Springer, Cham, 2014. – P. 71–94. – DOI: 10.1007/978-3-319-03512-3_5.
5. Crandall M. G., Lions P.-L. Viscosity solutions of Hamilton-Jacobi equations // Transactions of the American Mathematical Society. – 1983. – Vol. 277, no. 1. – P. 1–42. – DOI: 10.1090/S0002-9947-1983-0690039-8.
6. Feller W. Two singular diffusion problems // Annals of Mathematics. – 1951. – Vol. 54, No. 1. – P. 173–182. – DOI: 10.2307/1969318.
7. Heston S. L. A closed-form solution for options with stochastic volatility with applications to bond and currency options // The Review of Financial Studies. – 1993. – Vol. 6, No. 2. – P. 327–343. – DOI: 10.1093/rfs/6.2.327.
8. Ma J., Yong J. Forward-Backward stochastic differential equations and their applications. Springer. 2007. 270 p.

9. Three algorithms for solving high-dimensional fully-coupled FBSDEs through deep learning / S. Ji, S. Peng, Y. Peng, X. Zhang // IEEE Intelligent Systems. – 2020. – V. 35. – No. 3. – Pp. 71–84.
10. Raissi M. Forward-backward stochastic neural networks: Deep learning of high-dimensional partial differential equations // arXiv. – 2018. – arXiv:1804.07010. – URL: <https://arxiv.org/abs/1804.07010> (дата обращения: 01.07.2025).
11. Automatic differentiation in machine learning: a survey / A.G. Baydin, B.A. Pearlmutter, A.A. Radul, J.M. Siskind // arXiv. – 2018. – arXiv:1502.05767 – URL: <https://arxiv.org/abs/1502.05767> (дата обращения: 01.07.2025).
12. Kac M. On distributions of certain Wiener functionals // Transactions of the American Mathematical Society. – 1949. – Vol. 65. – P. 1–13. – DOI: 10.1090/S0002-9947-1949-0027960-X.
13. Shreve S. E. Stochastic Calculus for Finance II: Continuous-Time Models – New York: Springer, 2004. – 550 p.
14. Cox J. C., Ingersoll J. E., Ross S. A. A Theory of the Term Structure of Interest Rates // Econometrica. – 1985. – Vol. 53, No. 2. – P. 385–407. – DOI: 10.2307/1911242.

A NUMERICAL ALGORITHM FOR SOLVING THE CAUCHY PROBLEM FOR A TWO-DIMENSIONAL FULLY NONLINEAR PARABOLIC EQUATION

© 2025 A.A. Chubatov

Sirius University of Science and Technology, “Sirius” Federal Territory, Russia

The article continues the research presented in [1]. We solve the Cauchy problem for a two-dimensional fully nonlinear equation arising in the optimization of portfolio investments in the Heston market. First, the fully nonlinear equation is included in a system of quasilinear parabolic equations. The solution of the Cauchy problem for this system is reduced to a stochastic problem in terms of forward-backward stochastic differential equations (FBSDE). Next, the FBSDE is transformed into a stochastic optimal control problem. Finally, this problem is solved using a neural network approach. The proposed numerical algorithm was validated on an example for which it is possible to obtain a quasi-explicit solution.

Keywords: fully nonlinear parabolic equation, Cauchy problem, system of quasilinear parabolic equations, forward-backward stochastic differential equations (FBSDE), stochastic optimal control problem, deep learning, forward-backward stochastic neural network (FBSNN).

DOI: 10.37313/1990-5378-2025-27-5-230-239

EDN: ZJPVLI

*This work was supported by the grant of the state program of the «Sirius» Federal Territory
«Scientific and technological development of the «Sirius» Federal Territory»
(Agreement № 18-03 date 10.09.2024).*

REFERENCES

1. Chubatov, A. A. Chislennyi algoritm resheniya polnost'iu nelineinykh parabolicheskikh uravnenii na osnove primarnykh-obratnykh stokhasticheskikh differentsial'nykh uravnenii i neironnykh setei [A numerical algorithm for solving fully nonlinear parabolic equations based on forward-backward stochastic differential equations and neural networks]. *Izvestiya Samarskogo nauchnogo tsentra Rossiiskoi akademii nauk* [Proceedings of the Samara Scientific Center of the Russian Academy of Sciences]. 2024, vol. 26, no. 4, pp. 61–69. (in Russian)
2. Belopolskaia, Ya. I., Chubatov, A. A. Optimizatsiya investitsionnogo portfel'ia v modeli Hestona [Portfolio optimization in the Heston model]. *Zapiski nauchnykh seminarov Sankt-Petersburgskogo otdeleniya Matematicheskogo instituta im. V.A. Steklova Rossiiskoi akademii nauk* [Proceedings of the Scientific Seminars of the St. Petersburg Department of the Steklov Mathematical Institute of the Russian Academy of Sciences]. 2023, vol. 526, pp. 29–51. (in Russian)
3. Second order backward stochastic differential equations and fully non-linear parabolic PDEs / P. Cheridito, H. M. Sonner, N. Touzi, N. Victoir // Communications on Pure and Applied Mathematics. 2007. V. 60. No. 7. Pp. 1081–1110.
4. Belopolskaya, Ya. I. Probabilistic counterparts of nonlinear parabolic partial differential equation systems // Modern stochastics and applications. Springer Optimization and Its Applications; vol. 90. / ed. by V. Korolyuk [et al.]. – Springer, Cham, 2014. – P. 71–94. – DOI: 10.1007/978-3-319-03512-3_5.
5. Crandall M. G., Lions P.-L. Viscosity solutions of Hamilton-Jacobi equations // Transactions of the American Mathematical Society. – 1983. – Vol. 277, no. 1. – P. 1–42. – DOI: 10.1090/S0002-9947-1983-0690039-8.
6. Feller W. Two singular diffusion problems // Annals of Mathematics. – 1951. – Vol. 54, No. 1. – P. 173–182. – DOI: 10.2307/1969318.
7. Heston S. L. A closed-form solution for options with stochastic volatility with applications to bond and currency options // The Review of Financial Studies. – 1993. – Vol. 6, No. 2. – P. 327–343. – DOI: 10.1093/rfs/6.2.327.
8. Ma J., Yong J. Forward-Backward stochastic differential equations and their applications. Springer. 2007. 270 p.
9. Three algorithms for solving high-dimensional fully-coupled FBSDEs through deep learning / S. Ji, S. Peng, Y. Peng, X. Zhang // IEEE Intelligent Systems. – 2020. – V. 35. – No. 3. – Pp. 71–84.
10. Raissi M. Forward-backward stochastic neural networks: Deep learning of high-dimensional partial differential equations // arXiv. – 2018. – arXiv:1804.07010. – URL: <https://arxiv.org/abs/1804.07010> (Available 1.07.2025).

11. Automatic differentiation in machine learning: a survey / A.G. Baydin, B.A. Pearlmutter, A.A. Radul, J.M. Siskind // arXiv. – 2018. – arXiv:1502.05767 – URL: <https://arxiv.org/abs/1502.05767> (Available 1.07.2025).
12. *Kac M.* On distributions of certain Wiener functionals // Transactions of the American Mathematical Society. – 1949. – Vol. 65. – P. 1–13. – DOI: 10.1090/S0002-9947-1949-0027960-X.
13. *Shreve S. E.* Stochastic Calculus for Finance II: Continuous-Time Models – New York: Springer, 2004. – 550 p.
14. *Cox J. C., Ingersoll J. E., Ross S. A.* A Theory of the Term Structure of Interest Rates // Econometrica. – 1985. – Vol. 53, No. 2. – P. 385–407. – DOI: 10.2307/1911242.