Проект 6. Вокруг MLMC (Multilevel Monte Carlo). Приложения MLMC в финансовой математике

I. Постановка базовой задачи

В данной работе предлагается рассмотреть различные подходы к нахождению численных решений стохастических дифференциальных уравнений. Базовым является следующее дифференциальное уравнение с начальным условием

$$dS(t) = a(S(t), t)dt + b(S(t), t)dW(t), \quad 0 \leqslant t \leqslant T, \quad S(0) = S_0. \tag{1}$$

Здесь $W(t)$ – винеровский процесс.

Помимо нахождения численного решения дифференциальной задачи (которое, вообще говоря, зависит от реализации), бывает необходимо предложить оценку для следующей величины

$$Y = \mathbb{E}[f(S(T))],$$

где $f(S)$ – липшицева функция. Данная задача оценки возникает в различных приложениях (в частности, в теории ценообразования опционных контрактов).

Для начала рассмотрим самый очевидный подход к решению дифференциальной задачи. Введем равномерную сетку

$$t_n = n\tau, \quad n = \overline{0, N}, \quad T = N\tau.$$

Каждому t_n сопоставим значение сеточной функции S_n. Запишем явную схему (*Euler-Maruyama method*), "заменив" дифференциалы конечными приращениями.

$$\widehat{S}_{n+1} = \widehat{S}_n + a(\widehat{S}_n, t_n)\tau + b(\widehat{S}_n, t_n)\Delta W_n. \tag{2}$$

В данной работе мы будем оперировать некоторыми важными терминами [4].

Определение. Говорят, что численное решение $\widehat{S}$ стохастического дифференциального уравнения (1) *сходится строго к точному решению* S в момент T, если выполнено следующее предельное соотношение

$$\lim_{\tau \to 0} \mathbb{E}(|S(T) - \widehat{S}_N|) = 0.$$

Определение. Говорят, что численное решение $\widehat{S}$ стохастического дифференциального уравнения (1) *сходится строго к точному решению* S в момент T *с порядком аппроксимации* γ, если найдется такое положительное число C, не зависящее от τ, и такое положительное число τ_0, что выполнено следующее неравенство

$$\mathbb{E}(|S(T) - \widehat{S}_N|) \leqslant C\tau^{\gamma}$$

для любого $\tau \in (0, \tau_0)$.

Определение. Говорят, что численное решение $\widehat{S}$ стохастического дифференциального уравнения (1) *сходится слабо к точному решению* S в момент T в классе функций $\mathcal{K}$, если для любой $f \in \mathcal{K}$ выполнено следующее предельное соотношение

$$\lim_{\tau \to 0} \left| \mathbb{E}\left[f(S(T)\right] - \mathbb{E}\left[f(\widehat{S}_N)\right] \right| = 0.$$

Определение. Говорят, что численное решение $\widehat{S}$ стохастического дифференциального уравнения (1) *сходится слабо к точному решению* S в момент T *с порядком аппроксимации* γ в классе функций

$\mathcal{C}_P^{2(\gamma+1)}(\mathbb{R})$, если для любой $f \in \mathcal{C}_P^{2(\gamma+1)}(\mathbb{R})$ найдется такое положительное число C, не зависящее от τ, и такое положительное число τ_0, что выполнено следующее неравенство

$$\left| \mathbb{E}\left[f(S(T)\right] - \mathbb{E}\left[f(\widehat{S}_N)\right] \right| \leqslant C\tau^{\gamma}$$

для любого $\tau \in (0, \tau_0)$.

Замечание: $\mathcal{C}^m(\mathbb{R})$ — пространство m раз непрерывно дифференцируемых функций, заданных на вещественной прямой, а $\mathcal{C}_P^m(\mathbb{R})$ — подпространство функций $f \in \mathcal{C}^m(\mathbb{R})$, для которых все частные производные до порядка m включительно имеют полиномиальный рост. То есть найдется число $K > 0$ и число $r \in \mathbb{Z}^+$, зависящие, вообще говоря, от f, что

$$\left| \frac{\partial^j f(y)}{\partial y^j} \right| \leqslant K(1 + |y|^{2r}) \quad \forall y \in \mathbb{R}, \quad \forall j \leqslant m.$$

II. Явные методы аппроксимации стохастических дифференциальных уравнений. Monte Carlo подход к оценке целевой функции

В качестве оценки Y возьмем среднее по нескольким реализациям процессов $\{\widehat{S}_n^{(i)}\}$

$$\overline{Y} = \frac{1}{M} \sum_{i=1}^{M} f(\widehat{S}_N^{(i)}). \tag{3}$$

Возникает естественный вопрос: насколько оценка $\overline{Y}$ аппроксимирует точную величину Y? В качестве метрики качества аппроксимации будем использовать mean-squared error

$$\mathrm{MSE} = \mathbb{E}|Y - \overline{Y}|^2.$$

Кажется понятным, что функция MSE должна зависеть от используемых параметров конечно-разностной задачи, то есть от шага временной сетки τ и количества различных реализаций процессов M.

Для исследования асимптотики MSE в зависимости от параметров τ и M рассмотрим величину

$$\widehat{Y} = \mathbb{E}\left[f(\widehat{S}_N)\right].$$

Теперь распишем величину MSE

$$\mathrm{MSE} = \mathbb{E}\left[(\overline{Y} - Y)^2\right] = \mathbb{E}\left[(\overline{Y} - \widehat{Y} + \widehat{Y} - Y)^2\right] = \mathbb{E}\left[(\overline{Y} - \widehat{Y})^2\right] + 2\mathbb{E}\left[(\overline{Y} - \widehat{Y})(\widehat{Y} - Y)\right] + \mathbb{E}\left[(\widehat{Y} - Y)^2\right].$$

Поскольку $\widehat{Y} - Y = \mathrm{const}$, то

$$\mathrm{MSE} = \mathbb{E}\left[(\overline{Y} - \widehat{Y})^2\right] + 2(\widehat{Y} - Y)\mathbb{E}\left[(\overline{Y} - \widehat{Y})\right] + \mathbb{E}\left[(\widehat{Y} - Y)^2\right] = \mathbb{E}\left[(\overline{Y} - \widehat{Y})^2\right] + 2(\widehat{Y} - Y)\left(\mathbb{E}\overline{Y} - \widehat{Y}\right) +$$

$$+ \left(\widehat{Y} - Y\right)^2 = \mathbb{E}\left[(\overline{Y} - \widehat{Y})^2\right] + \left(\widehat{Y} - Y\right)^2.$$

Здесь мы учли, что на независимых реализациях

$$\mathbb{E}\overline{Y} = \frac{1}{M}\mathbb{E}\left[\sum_{i=1}^{M} f(\widehat{S}_N^{(i)})\right] = \frac{1}{M}\sum_{i=1}^{M}\mathbb{E}\left[f(\widehat{S}_N^{(i)})\right] = \mathbb{E}\left[f(\widehat{S}_N^{(1)})\right] = \widehat{Y}.$$

Осталось преобразовать первое слагаемое

$$\mathbb{E}\left[(\overline{Y} - \widehat{Y})^2\right] = \mathbb{E}\overline{Y}^2 - 2\mathbb{E}(\overline{Y}\widehat{Y}) + \mathbb{E}\widehat{Y}^2 = \mathbb{E}\overline{Y}^2 - (\mathbb{E}\overline{Y})^2 = \mathbb{D}\overline{Y} = \frac{1}{M}\mathbb{D}\left[f(\widehat{S}_N^{(1)})\right].$$

Итого,

$$\mathrm{MSE} = \frac{1}{M}\mathbb{D}\left[f(\widehat{S}_N^{(1)})\right] + \left(\widehat{Y} - Y\right)^2. \tag{4}$$

Мы получили очень важный результат: первое слагаемое в MSE возникает из-за реализации эксперимента по методу Monte Carlo. Второе слагаемое возникает из-за дискретизации исходной дифференциальной задачи явной схемой Euler-Maruyama. Распишем его подробнее. Чтобы это сделать, воспользуемся следующим результатом

Теорема 1. *Пусть $a(x,t), b(x,t)$ – липщицевы функции из класса $\mathcal{C}_P^4(\mathbb{R}^2)$. Для решения стохастического дифференциального уравнения*

$$dS(t) = a(S(t),t)dt + b(S(t),t)dW(t), \quad 0 \leqslant t \leqslant T, \quad S(0) = S_0.$$

предлагается использовать разностную схему

$$\widehat{S}_{n+1} = \widehat{S}_n + a(\widehat{S}_n, t_n)\tau_n + b(\widehat{S}_n, t_n)\Delta W_n, \quad \widehat{S}_0 = S_0.$$

с максимальным шагом по времени $\tau = \max\limits_n \tau_n$. Тогда для любой липшицевой функции $f \in \mathcal{C}_P^4(\mathbb{R})$ найдется положительное число C_f, не зависящее от τ, такое, что

$$\left|\mathbb{E}\left[f(S(T)\right] - \mathbb{E}\left[f(\widehat{S}_N)\right]\right| \leqslant C_f\tau. \tag{5}$$

Доказательство. Нам придется применить инструмент условного математического ожидания. Рассмотрим случайный процесс

$$u(x,t) = \mathbb{E}\left[f(S(T))|S(t) = x\right]. \tag{6}$$

Согласно формуле Feynman-Kac [4], $u(x,t)$ является решением дифференциального уравнения

$$\frac{\partial u}{\partial t} + a\frac{\partial u}{\partial x} + \frac{1}{2}b^2\frac{\partial^2 u}{\partial x^2} = 0. \tag{7}$$

Применим формулу Ито для функции $u(\widehat{S}(t),t)$

$$u(\widehat{S}(T),T) - u(\widehat{S}(0),0) = \int\limits_0^T du(\widehat{S}(t),t) = \int\limits_0^T \left(\frac{\partial u(\widehat{S}(t),t)}{\partial t} + \widehat{a}(t)\frac{\partial u(\widehat{S}(t),t)}{\partial x} + \frac{1}{2}\widehat{b}^2(t)\frac{\partial^2 u(\widehat{S}(t),t)}{\partial x^2}\right)dt +$$

$$+\int\limits_0^T \frac{\partial u(\widehat{S}(t),t)}{\partial x}\widehat{b}(t)dW.$$

Теперь сделаем замену, используя (7)

$$\frac{\partial u}{\partial t} = -a\frac{\partial u}{\partial x} - \frac{1}{2}b^2\frac{\partial^2 u}{\partial x^2}.$$

Подставим в предыдущую формулу и получим

$$\int\limits_0^T \left(\left(\widehat{a}(t) - a(\widehat{S}(t),t)\right)\frac{\partial u(\widehat{S}(t),t)}{\partial x} + \frac{1}{2}\left(\widehat{b}^2(t) - b^2(\widehat{S}(t),t)\right)\frac{\partial^2 u(\widehat{S}(t),t)}{\partial x^2}\right)dt + \int\limits_0^T \frac{\partial u(\widehat{S}(t),t)}{\partial x}\widehat{b}(t)dW.$$

Математическое ожидание от последнего интеграла обращается в нуль

$$\mathbb{E}\left[\int\limits_0^T \frac{\partial u(\widehat{S}(t),t)}{\partial x}\widehat{b}(t)dW\right] = 0.$$

Учтем также свойства введенного нами процесса, а именно

$$u(\widehat{S}(T),T) = f(\widehat{S}(t)), \quad u(\widehat{S}(0),0) = u(S_0,0) = \mathbb{E}\left[f(\widehat{S}(t))\right].$$

Далее, получаем

$$\mathbb{E}\left[f(S(T))\right] - \mathbb{E}\left[f(\widehat{S}(T))\right] = \int\limits_0^T \mathbb{E}\left[\left(\widehat{a}(t) - a(\widehat{S}(t),t)\right)\frac{\partial u(\widehat{S}(t),t)}{\partial x}\right]dt +$$

$$+\frac{1}{2}\int\limits_0^T \mathbb{E}\left[\left(\widehat{b}^2(t) - b^2(\widehat{S}(t),t)\right)\frac{\partial^2 u(\widehat{S}(t),t)}{\partial x^2}\right]dt = \sum_{n=0}^{N-1}\int\limits_{t_n}^{t_{n+1}} \mathbb{E}\left[\left(\widehat{a}(t) - a(\widehat{S}(t),t)\right)\frac{\partial u(\widehat{S}(t),t)}{\partial x}\right]dt +$$

$$+\frac{1}{2}\sum_{n=0}^{N-1}\int\limits_{t_n}^{t_{n+1}} \mathbb{E}\left[\left(\widehat{b}^2(t) - b^2(\widehat{S}(t),t)\right)\frac{\partial^2 u(\widehat{S}(t),t)}{\partial x^2}\right]dt.$$

Оставим без доказательства следующие верхние оценки

$$\sup_{t_n<t<t_{n+1}}\left|\mathbb{E}\left[\left(\widehat{a}(t) - a(\widehat{S}(t),t)\right)\frac{\partial u(\widehat{S}(t),t)}{\partial x}\right]\right| \leqslant C_1\tau, \tag{8}$$

$$\sup_{t_n<t<t_{n+1}}\left|\mathbb{E}\left[\left(\widehat{b}^2(t) - b^2(\widehat{S}(t),t)\right)\frac{\partial^2 u(\widehat{S}(t),t)}{\partial x^2}\right]\right| \leqslant C_2\tau. \tag{9}$$

Получаем первый порядок слабой сходимости численного решения $\widehat{S}(t)$ к точному решению $S(t)$ при $\tau \to 0$.

$$\left|\mathbb{E}\left[f(S(T))\right] - \mathbb{E}\left[f(\widehat{S}(T))\right]\right| \leqslant \left(C_1 + \frac{C_2}{2}\right)\tau\sum_{n=0}^{N-1}\int\limits_{t_n}^{t_{n+1}} dt = C(T)\tau. \tag{10}$$

■

Итак, мы выяснили, что схема Euler-Maruyama имеет первый порядок аппроксимации. Помимо точности, к численному решению предъявляются условия на вычислительные ресурсы, необходимые для получения данного решения. Ясно, что число вычислительных операций (total cost), необходимых для получения оценки $\overline{Y}$, равно $\mathrm{TC} \sim MN$, (напомним, что N – число подотрезков, на которые разбивается отрезок $[0,T]$, а M – число различных реализаций метода Monte Carlo). Оценка MSE приближенно равна

$$\mathrm{MSE}(M,N) \sim \frac{C_1}{M} + \frac{C_2}{N^2} = \frac{C_1}{M} + \frac{C_2 M^2}{\mathrm{TC}^2}.$$

Найдем минимум данной функции

$$\frac{\partial \mathrm{MSE}}{\partial M} = -\frac{C_1}{M^2} + \frac{2C_2 M}{\mathrm{TC}^2} = 0,$$

$$M = \left(\frac{C_1 \mathrm{TC}^2}{2C_2}\right)^{1/3}, \quad N = \left(\frac{2C_2 \mathrm{TC}}{C_1}\right)^{1/3}.$$

Поскольку

$$\frac{\partial^2 \mathrm{MSE}}{\partial M^2} = \frac{2C_1}{M^3} + \frac{2C_2}{\mathrm{TC}^2} > 0, \quad \forall M > 0,$$

то найденные значения M и N обеспечивают минимум MSE. При этом

$$\frac{C_1}{M} = \left(\frac{2C_2C_1^2}{\mathrm{TC}^2}\right)^2, \quad \frac{C_2}{N^2} = \left(\frac{C_2C_1^2}{4\mathrm{TC}^2}\right)^2,$$

то есть член, связанный с вариацией $\overline{Y}$ из-за реализации эксперимента по методу Monte Carlo в 2 раза больше, чем член, связанный с дискретизацией дифференциальной задачи (bias). Это довольно важный результат, который следует учитывать при подборе шага по времени τ и числа реализаций M по методу Monte Carlo.

Если же мы хотим минимизировать число вычислительных операций, то необходимо решить следующую оптимизационную задачу

$$\mathrm{TC} = MN = C_1 N \left(\mathrm{MSE} - \frac{C_2}{N^2}\right)^{-1} \longrightarrow \min.$$

$$\frac{\partial \mathrm{TC}}{\partial N} = \frac{3C_1N^2}{N^2\mathrm{MSE} - C_2} - \frac{2C_1N^4\mathrm{MSE}}{(N^2\mathrm{MSE} - C_2)^2} = 0,$$

Оптимальные значения числа временных подотрезков и числа реализаций метода Monte Carlo

$$N = \left(\frac{3C_2}{\mathrm{MSE}}\right)^{1/2}, \quad M = \frac{3C_1}{2\mathrm{MSE}}.$$

Количество операций

$$\mathrm{TC} = \left(\frac{3}{\mathrm{MSE}}\right)^{3/2} \frac{C_1\sqrt{C_2}}{2}.$$

Если от оценки требуется точность $\sqrt{\mathrm{MSE}} = O(\varepsilon)$, то с точки зрения оптимального количества вычислительных операций эффективно подбирать параметры следующим образом

$$N = O(\varepsilon^{-1}), \quad M = O(\varepsilon^{-2}),$$

при этом $\mathrm{TC} = O(\varepsilon^{-3})$.

Другой подход для построения конечной разностной схемы, аппроксимирующей исходную дифференциальную задачу (**Milstein method**)

$$\widehat{S}_{n+1} = \widehat{S}_n + a(\widehat{S}_n, t_n)\tau + b(\widehat{S}_n, t_n)\Delta W_n + \frac{1}{2} b(\widehat{S}_n, t_n)\frac{\partial b}{\partial t}(\widehat{S}_n, t_n)((\Delta W_n)^2 - \tau). \tag{11}$$

Пример 1. Рассмотрим стохастическое дифференциальное уравнение со следующими функциями

$$a(S(t), t) = \frac{2}{5}S^{3/5}(t) + 5S^{4/5}(t), \quad b(S(t), t) = S^{4/5}(t).$$

Для решения уравнения с заданным начальным условием

$$dS(t) = \left(\frac{2}{5}S^{3/5}(t) + 5S^{4/5}(t)\right)dt + \left(S^{4/5}(t)\right)dW(t), \quad 0 \leqslant t \leqslant T, \quad S(0) = 1 \tag{12}$$

воспользуемся формулой Ито. Рассмотрим функцию $\psi(x, t) = x^{1/5}$ и процесс $\psi(S(t), t)$. Стохастический дифференциал

$$d\psi(S(t), t) = \left[\frac{\partial \psi(S, t)}{\partial t} + \frac{\partial \psi(S, t)}{\partial x} a(S, t) + \frac{1}{2}\frac{\partial^2 \psi(S, t)}{\partial x^2} b^2(S, t)\right] dt + \frac{\partial \psi(S, t)}{\partial x} b(S, t) dW(t). \tag{13}$$

После подстановки

$$d\psi(S(t),t) = \left(\frac{2}{25}S^{-1/5}(t) + 1 - \frac{2}{25}S^{-1/5}(t)\right)dt + \frac{1}{5}dW(t) = dt + \frac{1}{5}dW(t).$$

Отсюда получаем точное решение стохастического дифференциального уравнения.

$$S(t) = \left(S_0 + t + \frac{1}{5}W(t)\right)^5 = \left(1 + t + \frac{1}{5}W(t)\right)^5. \tag{14}$$

Вычислим математическое ожидание $Y = \mathbb{E}S(T)$ при $T = 1$. Учтем, что $W(T) \sim \mathcal{N}(0,T) = \mathcal{N}(0,1)$. Моменты нормально распредленной случайной величины со средним математическим ожиданием

$$\mathbb{E}\left[W^p(T)\right] = \begin{cases} 0, & p - \text{нечетное}, \\ T^p(p-1)!!, & p - \text{четное}. \end{cases}$$

Таким образом, точное значение целевой переменной Y есть

$$\mathbb{E}\left(1 + T + \frac{1}{5}W(T)\right)^5 = (1+T)^5 + \frac{2}{5}(1+T)^3\mathbb{E}W^2(T) + \frac{1}{125}(1+T)\mathbb{E}W^4(T) = 32 + \frac{16}{5} + \frac{3}{125} = 35{,}224. \tag{15}$$

На рис. 1 представлены траектории процессов $S(t)$ для нескольких одинаковых реализаций винеровского процесса. Слева изображены траектории на основании точного решения стохастического дифференциального уравнения, а справа — траектории, рассчитанные Euler-Maruyama методом. Временной шаг сетки $\tau = 2^{-10}$. Число усреднений $M = 10^4$. На рис. 2 справа изображены траектории, рассчитанные Milstein методом.

Рис. 1: **Точные решения дифференциального уравнения (12) и численные решения, найденные методом Euler-Maruyama, при одинаковых реализациях винеровского процесса. Число реализаций $M = 10^4$, число реализаций $N = 2^{10}$**

Интересно исследовать зависимость величины ошибки MSE в зависимости от размера временного шага τ. В качестве оценки MSE будем использовать величину

$$\widehat{\chi} = \frac{1}{M}\sum_{i=1}^{M}(S^{(i)}(T) - \widehat{S}_N^{(i)})^2.$$

Рис. 2: **Точные решения дифференциального уравнения (12) и численные решения, найденные методом Milstein, при одинаковых реализациях винеровского процесса. Число реализаций $M = 10^4$, число реализаций $N = 2^{10}$**

Здесь индекс i относится к одной и той же реализации как численно найденного решения, так и точного решения дифференциальной задачи.

На рис. X представлена зависимость ошибки $\widehat{\chi}$ от значения N для двух методов дискретизации. В обоих случаях $M = 10^4$.

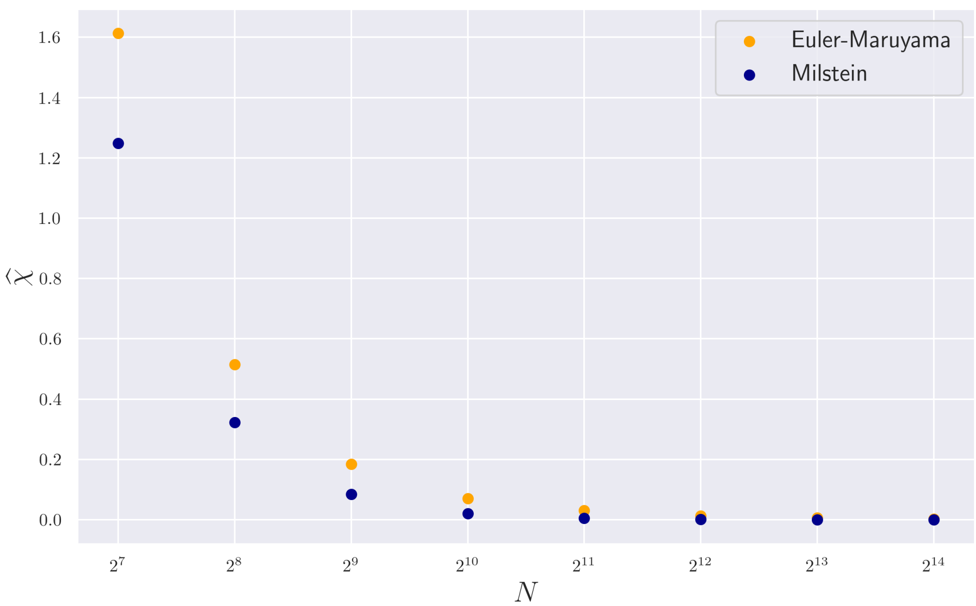

Рис. 3: **Оценка MSE для двух методов дискретизации дифференциальной задачи. Число реализаций $M = 10^4$**

III. Multilevel Monte Carlo подход к оценке целевой функции

Предлагается применить другую, куда более эффективную процедуру нахождения оценки величины Y. Данная процедура основана на multilevel Monte-Carlo (MLMC) методе.

Кратко будем обозначать (payoff function) величину

$$P = f(S(T)).$$

Данная терминология пришла из теории ценообразования опционов, к которой мы обратимся далее. Здесь как и раннее полагается, что f удовлетворяет условию Липшица на отрезке $[0, T]$.

Используя метод Euler-Maruyama, мы выяснили, что если мы хотим поддерживать значение MSE на уровне $O(\varepsilon^2)$, то оптимальный выбор параметров есть $\tau = O(\varepsilon)$ и $M = O(\varepsilon^{-2})$, а число вычислительных операций при таком подборе параметров $\text{TC} = O(\varepsilon^{-3})$. Наша задача состоит в оптимизации вычислительных ресурсов и приведении числа операций к асимптотике

$$\text{TC} = O(\varepsilon^{-2}(\log \varepsilon)^2),$$

при этом вычисления проводятся с различным количеством временных шагов N. Число изменений временных шагов будет подбираться таким образом, чтобы сохранялась требуемая точность численного решения.

Рассмотрим равномерные разбиения отрезка $[0, T]$ с варьируемым шагом $\tau_l = q^{-l}\tau_0$, где q – некоторое натуральное число, $l = \overline{0, L}$, а τ_0 берется таким, чтобы

$$N_0\tau_0 = T, \quad N_0 \in \mathbb{N}.$$

При этом число отрезков разбиения на уровне l есть $N_l = T/(\tau_0 q^l)$. Payoff на произвольной i-ой реализации численного решения на уровне l обозначим как

$$\widehat{P}_l^{(i)} = f(\widehat{S}_{N_l}^{(i)}).$$

Ясно, что

$$\mathbb{E}\left[\widehat{P}_L\right] = \mathbb{E}\left[\widehat{P}_0\right] + \sum_{l=1}^{L} \mathbb{E}\left[\widehat{P}_l - \widehat{P}_{l-1}\right].$$

Основная задача состоит в оценке $\mathbb{E}\left[\widehat{P}_0\right]$ и $\mathbb{E}\left[\widehat{P}_l - \widehat{P}_{l-1}\right]$ методом Monte Carlo

$$\overline{Z}_l = \frac{1}{M_l}\sum_{i=1}^{M_l}\left(\widehat{P}_l^{(i)} - \widehat{P}_{l-1}^{(i)}\right), \quad \overline{Z}_0 = \frac{1}{M_0}\sum_{i=1}^{M_0}\widehat{P}_0^{(i)}. \tag{16}$$

Используя независимые реализации для каждого уровня l, можно записать среднеквадратичную ошибку

$$\text{MSE} = \mathbb{D}\left[\sum_{l=0}^{L}\overline{Z}_l\right] = \sum_{l=0}^{L}\mathbb{D}\left[\overline{Z}_l\right] = \sum_{l=0}^{L}\frac{1}{M_l^2}\mathbb{D}\left[\sum_{i=1}^{M_l}\left(\widehat{P}_l^{(i)} - \widehat{P}_{l-1}^{(i)}\right)\right] = \sum_{l=0}^{L}\frac{V_l}{M_l},$$

где

$$V_0 = \mathbb{D}\left[\widehat{P}_0\right], \quad V_l = \mathbb{D}\left[\widehat{P}_l - \widehat{P}_{l-1}\right].$$

Количество вычислительных операций есть сумма числа операций на каждом уровне

$$\text{TC} = \sum_{l=0}^{L}\text{TC}_l = \sum_{l=0}^{L}M_l N_l.$$

Для нахождения минимума функции TC при фиксированной погрешности воспользуемся методом множителей Лагранжа, найдя оптимальные значения числа реализаций M_l метода Monte Carlo. Функцию Лагранжа представим в виде

$$\mathcal{L} = \sum_{l=0}^{L} \left(M_l N_l + \frac{V_l}{M_l} \mu^2 \right) \tag{17}$$

с множителем Лагранжа μ^2. Произведем дифференцирование функции Лагранжа по переменным M_l и найдем значения реализаций по методу Monte Carlo, при которых достигается оптимум

$$\frac{\partial \mathcal{L}}{\partial M_l} = N_l - \frac{V_l}{M_l^2} \mu^2 = 0, \quad M_l = \mu \sqrt{V_l / N_l}.$$

Если мы хотим поддерживать MSE на уровне ε^2, то необходимо удовлетворить

$$\mathbb{D} \left[\sum_{l=0}^{L} \overline{Z}_l \right] = \sum_{l=0}^{L} \frac{V_l}{\mu \sqrt{V_l / N_l}} = \frac{1}{\mu} \sum_{l=0}^{L} \sqrt{V_l N_l} \sim \varepsilon^2.$$

Отсюда находим оптимальное значение множителя Лагранжа

$$\mu = \varepsilon^{-2} \sum_{l=0}^{L} \sqrt{V_l N_l}.$$

Итоговое количество вычислительных операций

$$\mathrm{TC} = \varepsilon^{-2} \left(\sum_{l=0}^{L} \sqrt{V_l N_l} \right)^2. \tag{18}$$

Важно проследить за изменением произведения $V_l N_l$ при переходе с одного уровня на другой. Если $V_l N_l$ растет с увеличением уровня l, то наибольший вклад в total cost дает последний член с $l = L$, поэтому $\mathrm{TC} \approx \varepsilon^{-2} V_L N_L$. Если же наблюдается обратная ситуация, то есть произведение $V_l N_l$ падает с увеличением l, то доминирующим будет член $V_0 N_0$, и $\mathrm{TC} \approx \varepsilon^{-2} V_0 N_0$.

Попробуем получить асимптотику для V_l. Для липшицевой функции с константой Липшица K выполнено неравенство

$$|f(S_1) - f(S_2)| \leqslant K |S_1 - S_2|,$$

поэтому

$$\mathbb{D} \left[f(S(T)) - f(\widehat{S}_{N_l}^{(i)}) \right] = \mathbb{D} \left[P - P_l^{(i)} \right] \leqslant \mathbb{E} \left[(P - P_l^{(i)})^2 \right] \leqslant K^2 \mathbb{E} \left[(S(T) - \widehat{S}_{N_{l-1}}^{(i)})^2 \right], \tag{19}$$

$$V_l^{(i)} = \mathbb{D} \left[f(\widehat{S}_{N_l}^{(i)}) - f(\widehat{S}_{N_{l-1}}^{(i)}) \right] = \mathbb{D} \left[P_l^{(i)} - P_{l-1}^{(i)} \right] \leqslant 2 \left(\mathbb{D} \left[P - P_l^{(i)} \right] + \mathbb{D} \left[P - P_{l-1}^{(i)} \right] \right). \tag{20}$$

Поскольку

$$\mathbb{E} \left[f(S(T)) - f(\widehat{S}_N) \right] = O(\tau),$$

то $V_l = O(\tau_l)$. Поэтому

$$\mathrm{TC} = O(\varepsilon^{-2} L^2).$$

Если положить число уровней

$$L = -\frac{\log_2 \varepsilon}{\log_2 q} + O(1),$$

то при поддержании MSE $\sim \varepsilon^2$ число операций, необходимых для вычисления $\overline{Y}$ методом MLMC, имеет асимптотику

$$\mathrm{TC} = O(\varepsilon^{-2}L^2) = O(\varepsilon^{-2}(\log_2 \varepsilon)^2).$$

Далее нам понабится еще один крайне важный результат.

Теорема 2. *(**MLMC theorem**, [2]) Пусть $P = f(S)$ — payoff-функция решения $S(t)$ стохастического дифференциального уравнения (1), а $\widehat{P}_l$ — соответствующая ей оценка, полученная в результате численной аппроксимации (1) и использования метода MLMC на некотором уровне l.*

Если существуют независимые оценки $\overline{Z}_l$, рассчитанные на основе M_l реализаций по методу Monte Carlo, на каждую из которых было потрачено C_l операционного времени (объем вычислительных затрат) и положительные константы $\alpha, \beta, \gamma, c_1, c_2, c_3$ такие, что

$$\alpha \geqslant \frac{1}{2}\min(\beta, \gamma),$$

и

1. *для слабой сходимости $\widehat{P}_L$ выполнено*

$$\left| \mathbb{E}\left[\widehat{P}_l - P\right] \right| \leqslant c_1 2^{-\alpha l},$$

2. *математические ожидания оценок $\overline{Z}_l$ есть*

$$\mathbb{E}\left[\overline{Z}_l\right] = \begin{cases} \mathbb{E}\left[\widehat{P}_0\right], & l = 0, \\ \mathbb{E}\left[\widehat{P}_l - \widehat{P}_{l-1}\right], & l > 0, \end{cases}$$

3. *дисперсии оценок $\overline{Z}_l$ удовлетворяют неравенству*

$$\mathbb{D}\left[\overline{Z}_l\right] \leqslant c_2 M_l^{-1} 2^{-\beta l},$$

4. *математические ожидания вычислительных затрат удовлетворяют неравенству*

$$\mathbb{E}\left[C_l\right] \leqslant c_3 2^{\gamma l}.$$

Тогда существует такая положительная константа c_4, что для любого $0 < \varepsilon < 1$ найдутся число уровней L и число реализаций M_l на каждом уровне, для которых MLMC-оценка

$$\overline{Y} = \sum_{l=0}^{L} \overline{Z}_l$$

будет обладать MSE, удовлетворяющей неравенству

$$MSE = \mathbb{E}\left[(\overline{Y} - \mathbb{E}\left[P\right])^2\right] < \varepsilon^2$$

с полными вычислительными затратами

$$TC = \begin{cases} c_4\varepsilon^{-2}, & \beta > \gamma, \\ c_4\varepsilon^{-2}, (\log_2 \varepsilon)^2 & \beta = \gamma, \\ c_4\varepsilon^{-2-(\gamma-\beta)/\alpha}, & 0 < \beta < \gamma. \end{cases}$$

В приложениях оказывается крайне важным подобрать оптимальное значение параметра q (то, на сколько частей мы разбиваем временной отрезок длиной τ_l при переходе с уровня l на уровень $l+1$). Итак, мы выяснили, что

$$\mathbb{D}\left[\widehat{P}_l - P\right] \approx c_0 \tau_l.$$

Для разности payoff-функций на двух уровнях имеем

$$\widehat{P}_l - \widehat{P}_{l-1} = (\widehat{P}_l - P) - (\widehat{P}_{l-1} - P).$$

Дисперсию разности двух случайных величин можно представить в виде

$$V_l = \mathbb{D}\left[\widehat{P}_l - \widehat{P}_{l-1}\right] = \mathbb{D}\left[\widehat{P}_l - P\right] + \mathbb{D}\left[\widehat{P}_{l-1} - P\right] - 2\mathrm{cov}(\widehat{P}_l - P, \widehat{P}_{l-1} - P).$$

Отсюда получаем верхнюю и нижнюю границу для $\mathbb{D}\left[\widehat{P}_l - \widehat{P}_{l-1}\right]$, учитывая коллинеарность и антиколлинеарность величин $\widehat{P}_l - P$ и $\widehat{P}_{l-1} - P$

$$c_0 \tau_l (\sqrt{q} - 1)^2 \leqslant V_l \leqslant c_0 \tau_l (\sqrt{q} + 1)^2.$$

Теперь предположим, что значение V_l приближенно задается средним геометрическим верхней и нижней границ

$$V_l \approx c_0 \tau_l (q - 1).$$

Отсюда можно получить приближение для числа реализаций по методу multilevel Monte Carlo

$$M_l = \varepsilon^{-2} \sqrt{V_l / N_l} \sum_{l=0}^{L} \sqrt{V_l N_l} \approx 2\varepsilon^{-2}(L+1) V_l = 2\varepsilon^{-2}(L+1) c_0 \tau_l (q-1).$$

Число вычислительных операций для определения $\widehat{P}_l - \widehat{P}_{l-1}$ равно

$$M_l (N_l + N_{l-1}) = M_l N_l (1 + 1/q) \approx 2\varepsilon^{-2}(L+1) C_0 (q - 1/q).$$

Поскольку мы выбрали асимптотику для числа уровней следующим образом

$$L = -\frac{\log_2 \varepsilon}{\log_2 q} + O(1),$$

то, осуществляя суммирование по всем уровням, мы получим для total cost оценку

$$\mathrm{TC} = O(\varepsilon^{-2} (\log_2 \varepsilon)^2 \omega(q)), \quad \omega(q) = \frac{q - q^{-1}}{(\log_2 q)^2}.$$

График зависимости $\omega(q)$ представлен на рис. 4. Наименьшее значение $\omega(q)$ достигается при $q = 6{,}76 \approx 7$. Минимум $\omega(q)$ почти в 1,8 раз меньше значения $\omega(2)$. Поэтому надлежащий выбор фактора q позволяет значительно сэкономить вычислительные ресурсы. В дальнейших численных экспериментах будет браться значение $q = 4$.

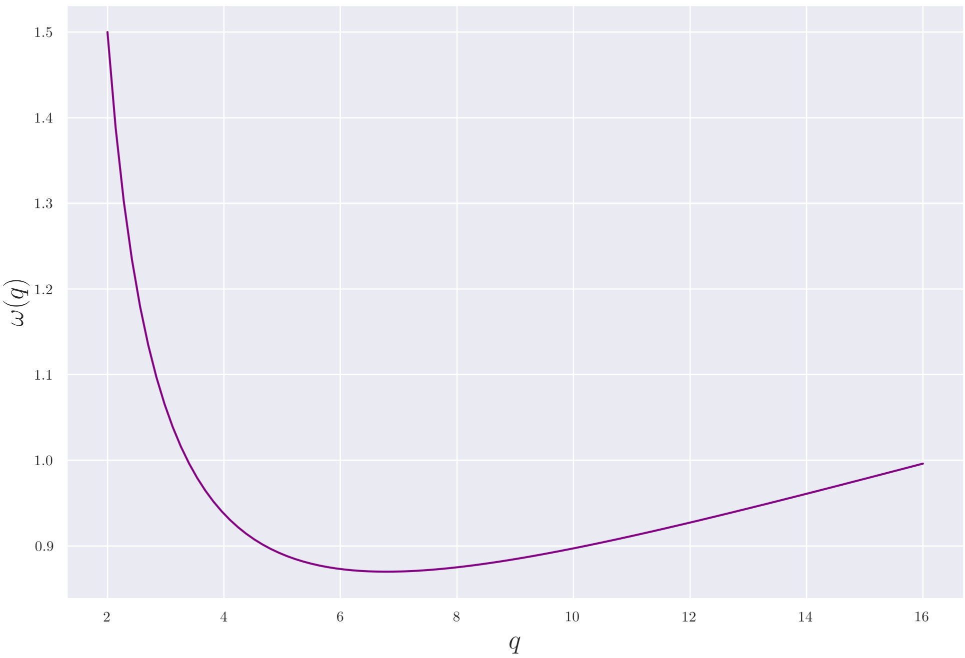

Рис. 4: **Зависимость $\omega(q)$ для нахождения оптимального значения параметра q**

Пример 2. Рассмотрим пример применения MLMC для расчета payoff европейского call-опциона на некоторый актив. Под call-опционом понимают финансовый инструемент, дающий право на покупку некоторого актива по заранее назначенной цене K (*цена исполнения опциона*) через время T (*время исполнения опциона*).

Пусть цена базового актива $S(t)$ описывается геометрическим броуновским движением

$$dS = rS(t)dt + \sigma S(t)dW, \quad 0 < t < T. \tag{21}$$

Начальная цена базового актива $S(0) = S_0 = 100$, время исполнения опциона $T = 1$. Также введены параметры $r = 0{,}05$ (*процентная ставка*), $\sigma = 0{,}2$ (*волатильность*).

Payoff-функция опциона

$$P(S(T)) = \exp(-rT)\max(S(T) - K, 0) = \exp(-rT)(S(T) - K)^+,$$

где *цену исполнения опциона* примем равной начальной цене базового актива $K = 100$.

Сперва найдем точное решение стохастического дифференциального уравнения (21). Используем формулу Ито для функции $\psi(x,t) = \ln x$. Стохастический дифференциал

$$d\psi(S(t),t) = d\ln S(t) = \left(r - \frac{\sigma^2}{2}\right)dt + \sigma dW.$$

Данное уравнение уже с легкостью интегрируется. Цена базового актива

$$S(t) = S_0 \exp\left(\left(r - \sigma^2/2\right)t + \sigma W(t)\right). \tag{22}$$

На рис. 5 представлена динамика цены базового актива $S(t)$, рассчитанная по формуле (22) для одной реализации, а также соответствующая payoff-функция $P(t)$. Как видно из рисунка, $P(t) = 0$ в тех случаях, когда цена базового актива $S(t)$ становится ниже цены исполнения опциона K (пунктирная линия).

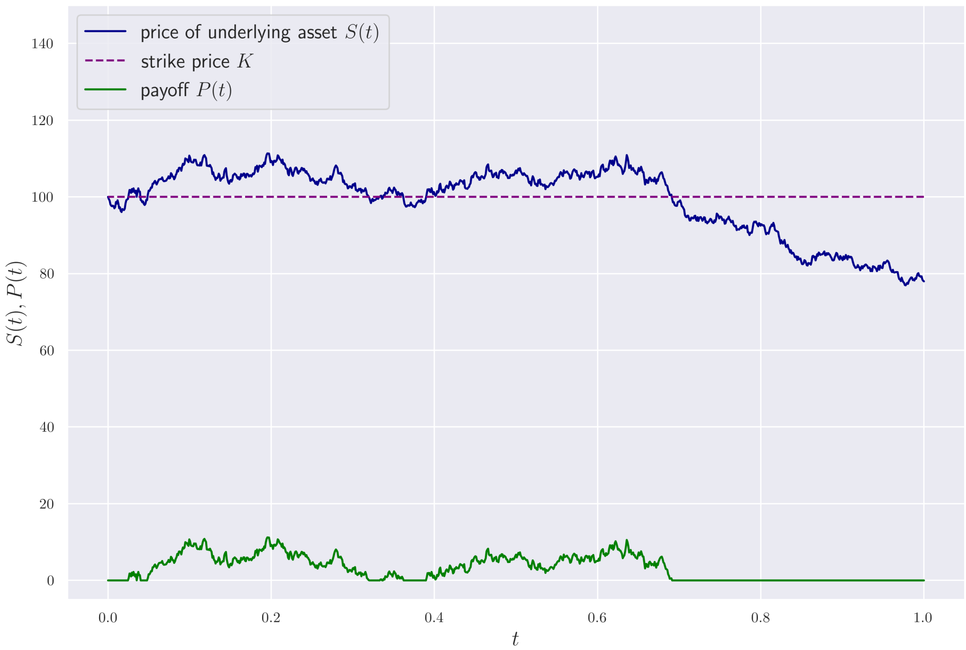

Рис. 5: **Динамика цены базового актива $S(t)$, рассчитанная по формуле для точного решения уравнения геометрического броуновского движения, и payoff-функции $P(t)$**

Для численного решения дифференциальной задачи предлагается использовать схему Euler-Maruyama

$$\widehat{S}_{n+1} = \widehat{S}_n + r\widehat{S}_n\tau_n + \sigma\widehat{S}_n\Delta W_n, \quad \overline{n = 0, N-1}, \quad \widehat{S}_0 = S(0).$$

На рис. 6 представлены пять реализаций цены базового актива $\widehat{S}(t)$, рассчитанные методом Euler-Maruyama (число шагов $N = 2^{10}$). Красным цветом обозначена средняя цена базового актива, усредненная по $M = 10^4$ реализациям метода Monte Carlo.

Сопоставим среднее значение payoff-функции, рассчитанное standard Monte Carlo методом, с формулой Black-Scholes

$$C(T) = S_0\Phi(y_+) - K\exp(-rT)\Phi\left(y_-\right), \tag{23}$$

где

$$y_\pm = \frac{1}{\sigma\sqrt{T}}\left(\ln\frac{S_0}{K} + \left(r \pm \frac{\sigma^2}{2}\right)T\right).$$

На рис. 7 желтым цветом обозначена (t), рассчитанная по формуле Black-Scholes. Зелеными точками обозначены усредненные по $M = 10^4$ реализациям метода Monte Carlo значения payoff.

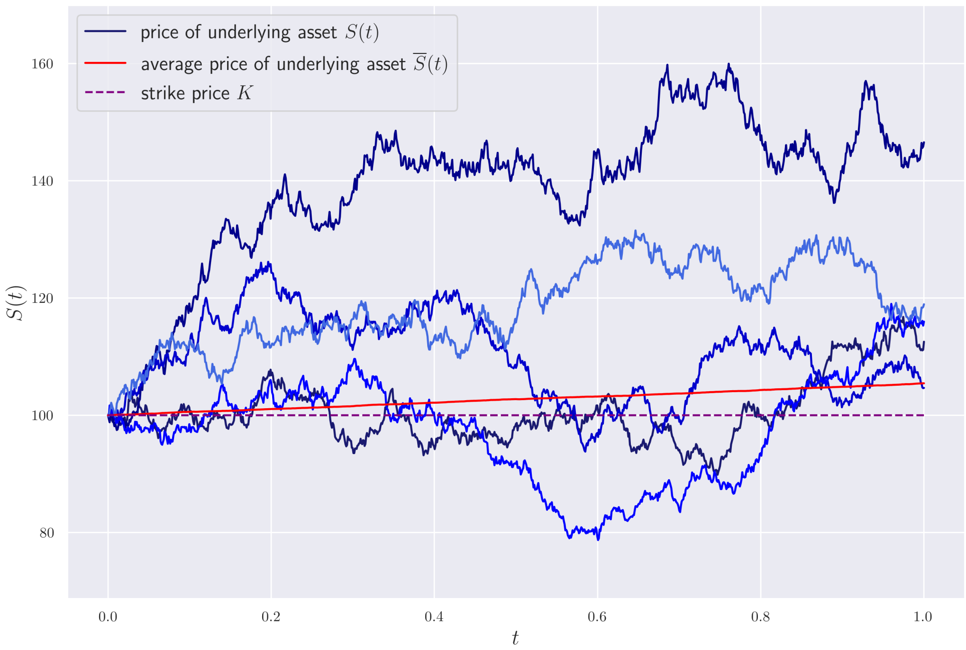

Рис. 6: **Динамика цены базового актива $\widehat{S}(t)$, рассчитанная методом Euler-Maruyama (standard MC), и усредненная цена базового актива для $M = 10^4$ реализаций. Количество шагов $N = 2^{10}$**

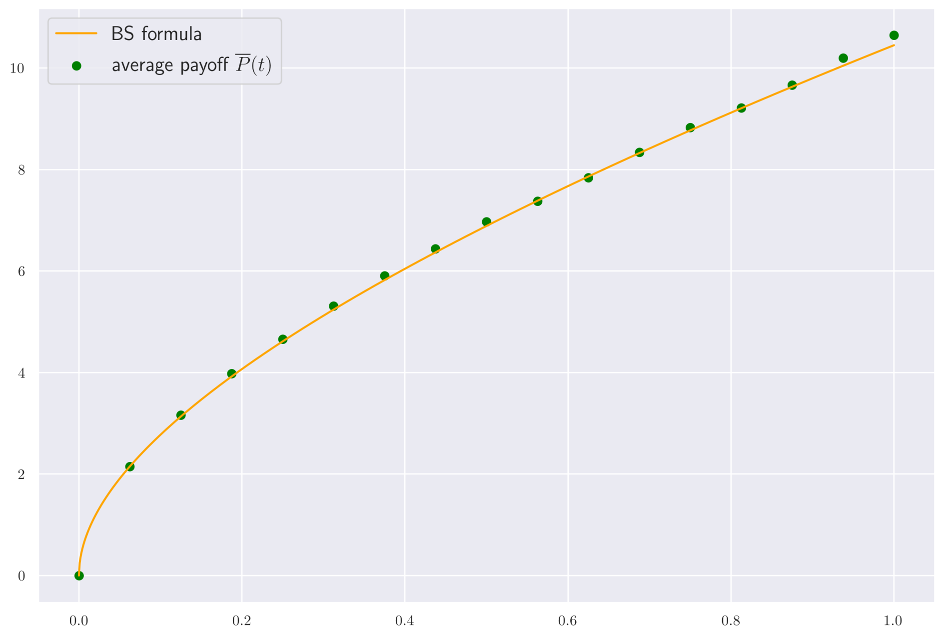

Рис. 7: **Сопоставление формулы Black-Scholes (23) с результатами численных расчетов методом Monte Carlo для $M = 10^4$ реализаций. Количество шагов $N = 2^{10}$**

Теперь приступим к реализации MLMC метода для изучения динамики цены опционного контракта. За основу взят модуль `pymlmc`, реализованный Patrick Farrell и Mike Giles на языке `Python`. MLMC-процедура состоит из нескольких этапов

1. На первом этапе предлагается реализовать проверку сходимости (convergence test). Начиная со значения $l = 0$ до значения $l = L$ находятся средние $\mathrm{Mean}(\widehat{P}_l - \widehat{P}_{l-1})$, $\mathrm{Mean}(\widehat{P}_l)$, а также отклонения $\mathrm{Var}(\widehat{P}_l - \widehat{P}_{l-1})$, $\mathrm{Var}(\widehat{P}_l)$.

   Сначала реализуется проверка сходимости на основе значения коэффициента эксцесса (kurtosis). Величина kurtosis рассчитывается по формуле

$$\varkappa = \frac{\eta_4 - 4\eta_3\eta_1 + 6\eta_2\eta_1^2 - 3\eta_1^4}{(\eta_2 - \eta_1^2)^2}, \quad \eta_p = \frac{1}{M}\sum_{i=1}^{M}(\widehat{P}_l^{(i)} - \widehat{P}_{l-1}^{(i)})^p.$$

   Здесь M — заданное пользователем число реализаций, выбранное для проверки сходимости. Если значение $\varkappa$ слишком велико, то программа выдает предупреждение: эмпирическая оценка дисперсии является неудовлетворительной.

   Также реализована проверка согласованности (consistency check). Если a, b, c есть суть оценки для величин $\mathbb{E}\left[\widehat{P}_l\right]$, $\mathbb{E}\left[\widehat{P}_{l-1}\right]$ и $\mathbb{E}\left[\widehat{P}_l - \widehat{P}_{l-1}\right]$ соответственно, то $a - b - c \approx 0$. Проверка согласованности подтверждает, что это равенство выполнено с точностью, которую можно было бы ожидать из-за ошибки по методу Monte Carlo. Для проверки согласованности рассчитывается статистика

$$F = \frac{|a - b - c|}{3(\sqrt{V_a} + \sqrt{V_b} + \sqrt{V_c})},$$

   где V_a, V_b, V_c — эмпирические дисперсии величин a, b, c. Вероятность того, что статистика $F > 1$ меньше 0,3%. Если окажется так, что $F > 1$, то программа также выдаст предупреждение о том, что нарушается равенство для двух аппроксимаций (грубой и точной) на уровне l

$$\mathbb{E}\left[\widehat{P}_l^c\right] = \mathbb{E}\left[\widehat{P}_l^f\right].$$

2. Далее осуществляется оценка параметров α, β, γ с помощью линейной регрессии.

3. По рассчитанным параметрам α, β, γ для погрешности ε, с которой мы хотим поддерживать величину $\sqrt{\mathrm{MSE}}$, находятся оптимальные значения числа реализаций M_l, а также рассчитываются вычислительные затраты C_l на каждом уровне l и итоговое значение оценки $\widehat{P}_L$ на последнем уровне L.

Результаты расчетов для европейского call-опциона с заданными параметрами представлены на рис. 8. На левом верхнем графике показано поведение дисперсий $\mathrm{Var}(\widehat{P}_l - \widehat{P}_{l-1})$, и $\mathrm{Var}(\widehat{P}_l)$ в зависимости от номера уровня l. Наклон графика $\mathrm{Var}(\widehat{P}_l - \widehat{P}_{l-1})$ в двойном логарифмическом масштабе равен $-\log_2 q = -2$, что соответствует раннее доказанному факту об асимптотике $V_l = O(\tau_l)$. При $l = 4$ дисперсия V_l почти в 2^{12} раза меньше дисперсии $\mathrm{Var}(\widehat{P}_l)$, которая рассчитывается в standard Monte Carlo методе.

На правом верхнем графике представлены средние значения $\mathrm{Mean}(\widehat{P}_l - \widehat{P}_{l-1})$, и $\mathrm{Mean}(\widehat{P}_l)$ в зависимости от номера уровня l.

На левом нижнем графике — реализация метода MLMC для различных значений ε, на уровне которых поддерживается погрешность $\sqrt{\mathrm{MSE}}$. Здесь представлены зависимости числа реализаций M_l в зависимости от номера уровня l. Уменьшение числа реализаций M_l с ростом l связано как с уменьшением дисперсий V_l, так и с уменьшением временного шага τ_l. Максимально необходимое число уровней L увеличивается с уменьшением допуска на величину MSE.

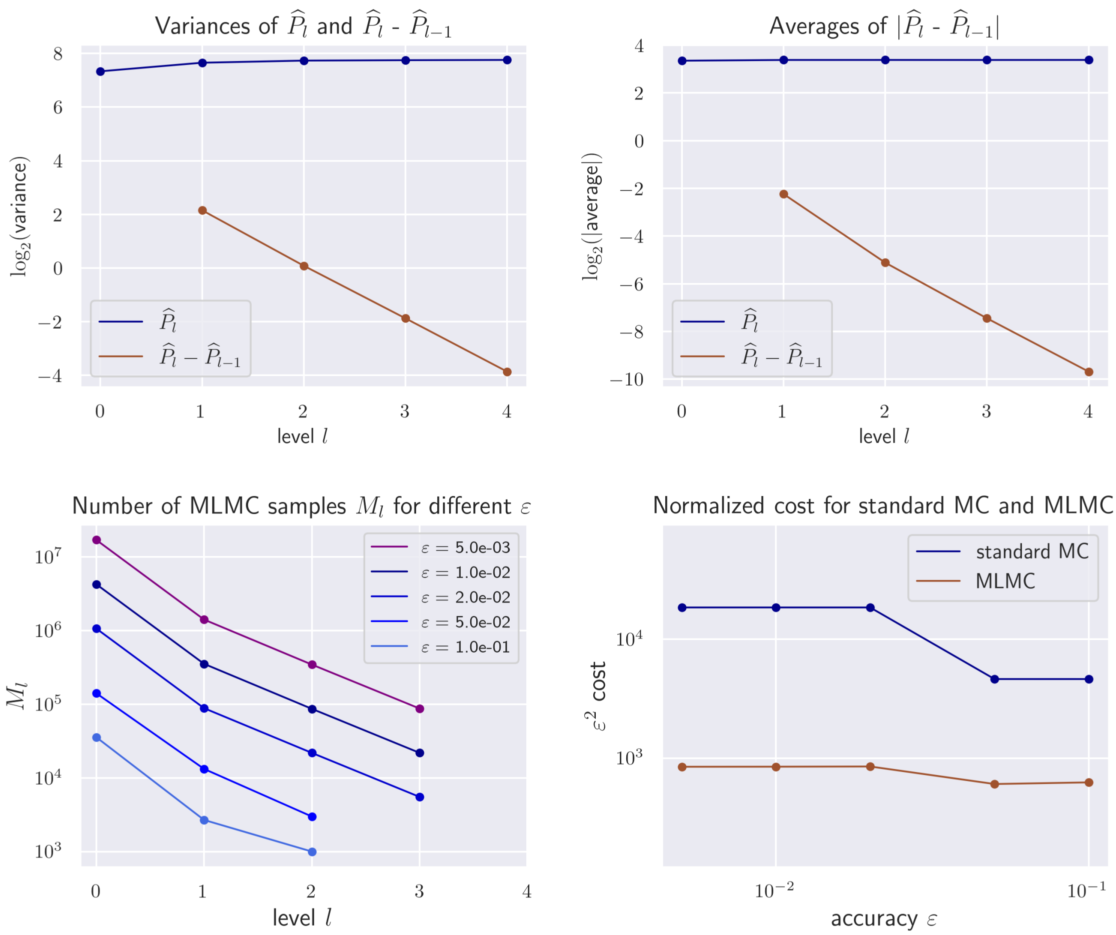

Рис. 8: **Результаты применения multilevel Monte Carlo для оценки payoff европейского call-опциона**

Последний правый нижний график показывает зависимость объема вычислительных ресурсов ТС, которые необходимо затратить для реализации MLMC, от требуемой точности ε.

Список литературы

[1] Гасников, А.В. *Стохастический анализ в задачах: Часть I* / А.В. Гасников, Н.О. Бузун, Ф.О. Гончаров, О.Г. Горбачев, С.А. Гуз, Е.А. Крымова, А.А. Натан, Е.О. Черноусова. – Москва: МФТИ, 2016. – С. 212.

[2] Giles, M.B. *Multilevel Monte Carlo methods* / M.B. Giles. – UK, Cambridge: Cambridge University Press, 2018. – P. 70.

[3] Giles, M.B. *Multilevel Monte Carlo Path Simulation* / M.B. Giles. // Operations Research. – 2008. – V. 56. – № 3. – P. 607-617.

[4] Kloeden, P.E. *Numerical solution of stochastic differential equations* / P.E. Kloeden, E. Platen. – Berlin: Springer-Verlag, 1992. – P. 666.

[5] Гасников, А.В. *Лекции по случайным процессам* / А.В. Гасников, Э.А. Горбунов, С.А. Гуз, Е.О. Черноусова, М.Г. Широбоков, Е.В. Шульгин. – Москва: МФТИ, 2019. – С. 285.

[6] Булинский, А.В. *Случайные процессы. Примеры, задачи и упражнения* / А.В. Булинский. – Москва: МФТИ, 2010. – С. 216.