ФЕДЕРАЛЬНОЕ ГОСУДАРСТВЕННОЕ БЮДЖЕТНОЕ
ОБРАЗОВАТЕЛЬНОЕ
УЧРЕЖДЕНИЕ ВЫСШЕГО ОБРАЗОВАНИЯ
«МОСКОВСКИЙ ГОСУДАРСТВЕННЫЙ УНИВЕРСИТЕТ
имени М.В.ЛОМОНОСОВА»

МЕХАНИКО-МАТЕМАТИЧЕСКИЙ ФАКУЛЬТЕТ

КАФЕДРА МАТЕМАТИЧЕСКОЙ СТАТИСТИКИ И СЛУЧАЙНЫХ
ПРОЦЕССОВ

ВЫПУСКНАЯ КВАЛИФИКАЦИОННАЯ РАБОТА
(ДИПЛОМНАЯ РАБОТА)
специалиста

Оценивание плотности размеров наночастиц на основе вейвлетов

Выполнил студент
603 группы
Купряков Василий Юрьевич

(подпись студента)

Научный руководитель:
с.н.с., к.ф.-м.н.
Шкляев Александр Викторович

(подпись научного руководителя)

Москва
2022

1 Введение

В работе мы изучим задачу, которая возникает при исследовании коллоидных примесей в жидкости.

Примеси в исследуемой жидкости — это движущиеся частицы с размерами порядка 10^{-8} м. Для исследования таких примесей используется анализ траекторий наночастиц.

Жидкость просвечивают лазером, когда луч попадает на чатицу, она рассеивает свет. К микроскопу присоединена камера, которая фиксирует рассеянный свет.

Получается последовательность изображений. Для каждой частицы эта последовательность является последовательностью проекций частиц на площадь камеры. Мы можем построить векторы перемещений частиц в проекции на плоскость камеры по этим снимкам. Для отдельной частицы такие перемещения образуют броуновское движение с нулевым сносом и дисперсией $\sigma^2 = c/d$, где c — некоторая константа, а d — размер частицы.

Проблема в том, что размер частицы не связан напрямую с размером ее изображения. Наша задача — оценить распределение истинных размеров частиц по размерам на снимках.

Будем изучать равносильную задачу: оценить распределение σ^2. Рассмотрим n случайно выбранных частиц $E_1, \ldots, E_n$. Обозначим дисперсии для их движения как $\sigma_1^2, \ldots, \sigma_n^2$. Для i-й частицы у нас есть два $k(i)$-мерных вектора перемещений: $A_i^1, \ldots, A_i^{k(i)}$ по оси x и $A_i^{k(i)+1}, \ldots, A_i^{2k(i)}$ по оси y. Мы будем рассматривать только частный случай, когда все $k(i)$ равны k, а σ_i^2 непрерывна.

$A_i^1, \ldots, A_i^{2k}$ условно независимы при условии σ_i^2 и имеют условное распределение $\mathcal{N}\left(0, \sigma_i^2\right)$. Дальше вместо выборки $A_i^1, \ldots, A_i^{2k}$ будем рассматривать достаточную статистику $Z_i = \sum\limits_{j=1}^{2k} \left(A_i^j\right)^2$. Заметим далее, что $Z_i = \sigma_i^2 Y_i$, где $Y_i \sim \chi_{2k}^2$. При этом, Y_i независимы и не зависят от дисперсии σ_i^2.

Обозначим $X_i = \sigma_i^2$. Тогда задачу можно сформулировать так: $X_1, \ldots, X_n$ — независимые одинаково распределенные непрерывные случайные величины с неизвестным распределением и положительным носителем; $Y_1, \ldots, Y_n$ — независимые от них н.о.р. с.в. с распределением χ_{2k}^2; $Z_1, \ldots, Z_n = X_1Y_1, \ldots, X_nY_n$ — наблюдаемые случайные величины; а сама задача — по наблюдениям $Z_1, ..., Z_n$ оценить распределение X_1.

2 Предварительные сведения

Мы будем рассматривать дискретное вейвлет-преобразование. Сформулируем основные определения и свойства.

Определение. Материнский вейвлет — это функция $\psi(t)$, которая отвечает следующему свойству:

$$\int\limits_0^\infty \frac{\left|\hat{\psi}(\xi)\right|^2}{|\xi|} d\xi = \int\limits_{-\infty}^0 \frac{\left|\hat{\psi}(\xi)\right|^2}{|\xi|} d\xi < \infty,$$

где $\hat{\psi(\xi)}$ – образ фурье $\psi(\xi)$.

Из этого свойства следует, что

$$\int\limits_{-\infty}^{\infty} \psi(t) dt = 0.$$

Из материнского вейвлета $\psi(t)$ строится система вейвлетов

$$\psi_{m,n}(t) = \frac{1}{\sqrt{a^m}} \psi \left(\frac{t}{a^m} - nb \right).$$

В дальнейшем мы будем испольвать вейвлеты с $a = 2$ и $b = 1$. Таким образом, мы будем использовать систему:

$$\psi_{m,n}(t) = \frac{1}{\sqrt{2^m}} \psi \left(\frac{t}{2^m} - n \right).$$

Часто система вейвлетов не является ортогональной, но образует фрейм.

Определение. Семейство ϕ_k является фреймом в $L_2(\mathbb{R})$, если существуют постоянные A и B такие, что

$$\forall f \in L_2(\mathbb{R}) \quad A\|f\|^2 \leqslant \sum_k |(f, \phi_k)|^2 \leqslant B\|f\|^2$$

Лемма. *Если семейство ϕ_k образует фрейм, то*

$$f = \frac{2}{A+B} \sum_k (f, \phi_k)\, \phi_k + Rf,$$

где

$$\|R\| \leqslant \frac{B-A}{B+A}$$

Мы будем использовать два вейвлета: Mexican hat и вейвлет Мейера. Приведем их материнские функции.

Определение. Материнская функция вейвлета Mexican hat:

$$\psi(t) = \frac{2}{\sqrt{3}} \pi^{-1/4} \left(1 - t^2\right) e^{-t^2/2}.$$

Определение. Материнская функция вейвлета Мейера:

$$\psi(t) = \psi_1(t) + \psi_2(t),$$

где

$$\psi_1(t) = \frac{\frac{4}{3\pi}\left(t-\frac{1}{2}\right)\cos\left(\frac{2\pi}{3}\left(t-\frac{1}{2}\right)\right) - \frac{1}{\pi}\sin\left(\frac{4\pi}{3}\left(t-\frac{1}{2}\right)\right)}{\left(t-\frac{1}{2}\right) - \frac{16}{9}\left(t-\frac{1}{2}\right)^3},$$

$$\psi_2(t) = \frac{\frac{8}{3\pi}\left(t-\frac{1}{2}\right)\cos\left(\frac{8\pi}{3}\left(t-\frac{1}{2}\right)\right) + \frac{1}{\pi}\sin\left(\frac{4\pi}{3}\left(t-\frac{1}{2}\right)\right)}{\left(t-\frac{1}{2}\right) - \frac{64}{9}\left(t-\frac{1}{2}\right)^3}.$$

Вейвлет Mexican hat образует фрейм с границами $A = 3{,}223$, $B = 3{,}596$, вейвлет Мейера является ортогональным, т. е. образует базис $L_2(\mathbb{R})$.

3 Сведение задачи к вычислению обратного преобразования Лапласа

Есть случайные величины X, Y, Z. Мы не знаем распределение X, знаем распределение Y и наблюдаем Z. Кроме того, известно, что $Z = XY$, и что все величины непрерывны. Нужно оценить распределение X.

Мы будем использовать вейвлет «Mexican hat», потому что он прост и непрерывен. Его формула:

$$\psi(t) = \frac{2}{\sqrt{3}\pi^{1/4}}(1-t^2)e^{-t^2/2}.$$

Определим элементы фрейма:

$$\psi_{m,n}(t) = \frac{1}{\sqrt{2^m}}\psi\left(\frac{t}{2^m} - n\right) = \frac{1}{\sqrt{2^m}}\frac{2}{\sqrt{3}\pi^{1/4}}\left(1 - \left(\frac{t}{2^m} - n\right)^2\right)e^{-\left(\frac{t}{2^m}-n\right)^2/2}.$$

Рассмотрим случай $Y \sim \chi^2_{2k}$; $X > \delta > 0$, абсолютно непрерывен. Плотность Y:

$$\chi^2_{2k} \sim \frac{1}{2^k}\frac{1}{\Gamma(k)}x^{k-1}e^{-x/2}.$$

Будем строить функции $g_{m,n}$ такие, что $\mathrm{E}\, g_{m,n}(Z) = \mathrm{E}\,\psi_{m,n}(X)$. Заметим, что достаточно выполнения:

$$\forall x \in \operatorname{Im} X \quad \mathrm{E}\, g_{m,n}(xY) = \psi_{m,n}(x).$$

Заменим мат. ожидание преобразованием Лапласа и раскроем $\psi_{m,n}$:

$$\left(\frac{1}{2x}\right)^k \frac{1}{\Gamma(k)} L_z\left[g_{m,n}(z)z^{k-1}\right]\left(\frac{1}{2x}\right) = \left(\frac{1}{\sqrt{2}}\right)^m \psi_{m,n}\left(\frac{x}{2^m} - n\right).$$

Сделаем замену $u = \frac{1}{2x}$:

$$u^k \frac{1}{\Gamma(k)} L_z\left[g_{m,n}(z)z^{k-1}\right](u) = \left(\frac{1}{\sqrt{2}}\right)^m \psi_{m,n}\left(\frac{1}{2^{m+1}u} - n\right).$$

Используя обратное преобразование Лапласа, найдем $g_{m,n}(t)$:

$$\begin{aligned} u^k \frac{1}{\Gamma(k)} L_z \left[g_{m,n}(z) z^{k-1} \right](u) &= \left(\frac{1}{\sqrt{2}} \right)^m \psi_{m,n} \left(\frac{1}{2^{m+1}u} - n \right) \\ L_z \left[g_{m,n}(z) z^{k-1} \right](u) &= \frac{\Gamma(k)}{u^k} \left(\frac{1}{\sqrt{2}} \right)^m \psi_{m,n} \left(\frac{1}{2^{m+1}u} - n \right) \\ g_{m,n}(t) t^{k-1} &= L_u^{-1} \left[\frac{\Gamma(k)}{u^k} \left(\frac{1}{\sqrt{2}} \right)^m \psi_{m,n} \left(\frac{1}{2^{m+1}u} - n \right) \right](t) \\ g_{m,n}(t) &= \frac{1}{t^{k-1}} \frac{\Gamma(k)}{\sqrt{2^m}} L_u^{-1} \left[\frac{1}{u^k} \psi_{m,n} \left(\frac{1}{2^{m+1}u} - n \right) \right](t). \end{aligned}$$

Таким образом мы получили выражение для $g_{m,n}(t)$. Далее мы выразим его через ряды, используя формулу Меллина и основную теорему о вычетах.

4 Вычисление обратного преобразования Лапласа с помощью формулы Меллина и основной теоремы о вычетах

Мы выразили $g_{m,n}(t)$ как:

$$g_{m,n}(t) = \frac{1}{t^{k-1}} \frac{\Gamma(k)}{\sqrt{2^m}} L_u^{-1} \left[\frac{1}{u^k} \psi_{m,n} \left(\frac{1}{2^{m+1}u} - n \right) \right](t). \tag{1}$$

Теперь нужно вычислить обратное преобразование Лапласа. Для этого мы будем использовать формулу Меллина:

$$L_s^{-1} \left[F(s) \right](t) = \frac{1}{2\pi i} \int\limits_{\alpha - i\infty}^{\alpha + i\infty} e^{ts} F(s) ds.$$

Итак, нам нужно вычислить:

$$L_u^{-1} \left[\frac{1}{u^k} \psi_{m,n} \left(\frac{1}{2^{m+1}u} - n \right) \right](t).$$

Подставим вместо ψ формулу нашего вейвлета:

$$\begin{aligned} &L_u^{-1} \left[\frac{1}{u^k} \psi_{m,n} \left(\frac{1}{2^{m+1}u} - n \right) \right](t) = \\ &\qquad L_u^{-1} \left[\frac{1}{u^k} \frac{2}{\sqrt{3}\pi^{1/4}} \left(1 - \left(\frac{1}{2^{m+1}u} - n \right)^2 \right) \exp\left(-\frac{1}{2} \left(\frac{1}{2^{m+1}u} - n \right)^2 \right) \right](t). \end{aligned}$$

Распишем множитель перед экспонентой:

$$1-\left(\frac{1}{2^{m+1}u}-n\right)^2=1-\left(\frac{1}{2^{2(m+1)}u^2}-2\frac{1}{2^{m+1}u}n+n^2\right)=$$
$$\left(1-n^2\right)+\frac{1}{u}\left(\frac{n}{2^m}\right)-\frac{1}{u^2}\left(\frac{1}{4^{m+1}}\right).$$

Введем обозначение:

$$r_{m,n}(u)=\frac{2}{\sqrt{3}\pi^{1/4}}\exp\left(-\frac{1}{2}\left(\frac{1}{2^{m+1}u}-n\right)\right).$$

Таким образом,

$$L_u^{-1}\left[\frac{1}{u^k}\psi_{m,n}\left(\frac{1}{2^{m+1}u}-n\right)\right](t)=$$
$$=\left(1-n^2\right)L_u^{-1}\left[\frac{1}{u^k}r_{m,n}(u)\right](t)+\left(\frac{n}{2^m}\right)L_u^{-1}\left[\frac{1}{u^{k+1}}r_{m,n}(u)\right](t)-$$
$$-\left(\frac{1}{4^{m+1}}\right)L_u^{-1}\left[\frac{1}{u^{k+2}}r_{m,n}(u)\right](t). \quad (2)$$

Отсюда видно, что достаточно найти $L_u^{-1}[\frac{1}{u^k}r_{m,n}(u)](t)$ для каждого k.

4.0.1 Находим $L_u^{-1}[\frac{1}{u^k}r_{m,n}(u)](t)$

Выше мы ввели

$$L_u^{-1}\left[\frac{1}{u^k}r_{m,n}(u)\right](t).$$

Подставим обратно $r_{m,n}(u)$:

$$L_u^{-1}\left[\frac{1}{u^k}r_{m,n}(u)\right](t)=L_u^{-1}\left[\frac{1}{u^k}\frac{2}{\sqrt{3}\pi^{1/4}}\exp\left(-\frac{1}{2}\left(\frac{1}{2^{m+1}u}-n\right)^2\right)\right](t).$$

Раскроем квадрат под экспонентой:

$$L_u^{-1}\left[\frac{1}{u^k}r_{m,n}(u)\right](t)=$$
$$=L_u^{-1}\left[\frac{1}{u^k}\frac{2}{\sqrt{3}\pi^{1/4}}\exp\left(-\frac{1}{2}\left(\left(n^2\right)-\frac{1}{u}\left(\frac{n}{2^m}\right)+\frac{1}{u^2}\left(\frac{1}{4^{m+1}}\right)\right)\right)\right](t).$$

Сгруппируем $2^{m+1}u$:

$$L_u^{-1}\left[\frac{1}{u^k}r_{m,n}(u)\right](t)=$$
$$=L_u^{-1}\left[\frac{2^{k(m+1)}}{\left(2^{m+1}u\right)^k}\frac{2}{\sqrt{3}\pi^{1/4}}\exp\left(-\frac{n^2}{2}+\frac{n}{2^{m+1}u}-\frac{1}{2\left(2^{m+1}u\right)^2}\right)\right](t).$$

Вынесем множители, не зависящие от u, за L_u^{-1}:

$$L_u^{-1}\left[\frac{1}{u^k}r_{m,n}(u)\right](t) = \\ = e^{-\frac{n^2}{2}}2^{k(m+1)}\frac{2}{\sqrt{3}\pi^{1/4}}L_u^{-1}\left[\frac{1}{\left(2^{m+1}u\right)^k}\exp\left(\frac{n}{2^{m+1}u}-\frac{1}{2\left(2^{m+1}u\right)^2}\right)\right](t).$$

Используя теорему о замене переменной в обратном преобразовании Лапласа, делаем замену $s = 2^{m+1}u$:

$$L_u^{-1}\left[\frac{1}{u^k}r_{m,n}(u)\right](t) = e^{-\frac{n^2}{2}}2^{(k-1)(m+1)}\frac{2}{\sqrt{3}\pi^{1/4}}L_s^{-1}\left[\frac{1}{s^k}\exp\left(\frac{n}{s}-\frac{1}{2s^2}\right)\right]\left(\frac{t}{2^{m+1}}\right).$$

4.0.2 Находим обратное преобразование Лапласа

В предыдущем разделе мы выразили:

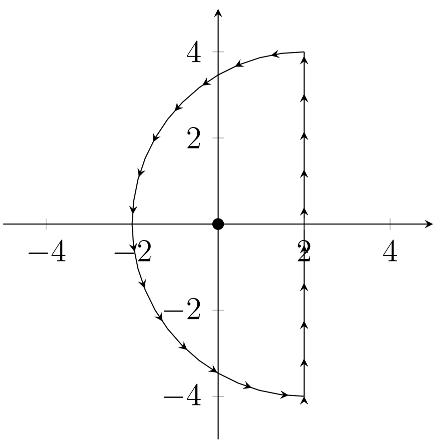
$$L_u^{-1}\left[\frac{1}{u^k}r_{m,n}(u)\right](t) = e^{-\frac{n^2}{2}}2^{(k-1)(m+1)}\frac{2}{\sqrt{3}\pi^{1/4}}L_s^{-1}\left[\frac{1}{s^k}\exp\left(\frac{n}{s}-\frac{1}{2s^2}\right)\right]\left(\frac{t}{2^{m+1}}\right). \tag{3}$$

Чтобы вычислить правую часть, найдем теперь

$$L_s^{-1}\left[\frac{1}{s^k}\exp\left(\frac{n}{s}\right)\exp\left(-\frac{1}{2s^2}\right)\right](\tau). \tag{4}$$

Воспользуемся формулой Меллина обратного преобразования Лапласа. У нас особенность только в нуле, поэтому можно взять любую $\alpha > 0$:

$$L_s^{-1}\left[\frac{1}{s^k}\exp\left(\frac{n}{s}\right)\exp\left(-\frac{1}{2s^2}\right)\right](\tau) = \frac{1}{2\pi i}\int\limits_{\alpha-i\infty}^{\alpha+i\infty} e^{s\tau}\frac{1}{s^k}e^{-1/(2s^2)}e^{n/s}ds.$$



Берем контур $C = C_1 + C_2$, где C_1 — искомый, а C_2 — дуга окружности (слева от C_1 с центром в $(\alpha, 0)$).

Оценим $F(s) := (1/s^k)e^{-1/(2s^2)}e^{n/s}$ на C_r, где $r > 4\alpha$. Для этого оценим каждый из множителей. Сначала $1/s$:

$$\left|\frac{1}{s}\right| = \left|\frac{1}{\alpha + re^{i\phi}}\right| = \frac{1}{\sqrt{(\alpha + r\cos\phi)^2 + (r\sin\phi)^2}} = \frac{1}{r\sqrt{\left(\frac{\alpha}{r} + \cos\phi\right)^2 + \sin^2\phi}} \leq$$

$$\leq \frac{1}{r\sqrt{1 + \frac{2\alpha\cos\phi}{r}}} \leq \frac{1}{r\sqrt{1 + \frac{\cos\phi}{2}}} \leq \frac{\sqrt{2}}{r}.$$

Теперь оценим $e^{-1/(2s^2)}$ на том же контуре. Известно, что $|e^z| = e^{|z|}$

$$\left|\exp\left(-\frac{1}{2s^2}\right)\right| \leq \exp\left(\left|-\frac{1}{2s^2}\right|\right) = \exp\left(\frac{1}{r^2}\right).$$

Аналогично оцениваем $e^{n/s}$:

$$\left|\exp\left(\frac{n}{s}\right)\right| \leq \exp\left(\left|\frac{n}{s}\right|\right) = \exp\left(\frac{|n|\sqrt{2}}{r}\right).$$

Объединяем оценки и получаем:

$$\left|\frac{1}{s^k}\exp\left(-\frac{1}{2s^2}\right)\exp\left(\frac{n}{s}\right)\right| \leq \left(\frac{\sqrt{2}}{r}\right)^k \exp\left(\frac{1}{r^2}\right)\exp\left(\frac{|n|\sqrt{2}}{r}\right) \xrightarrow[r\to\infty]{} 0.$$

А значит, по лемме Жордана $\int\limits_{C_r} e^{s\tau}F(s)ds$ стремится к нулю. Поэтому можем использовать основную теорему о вычетах:

$$\frac{1}{2\pi i}\int\limits_{\alpha-i\infty}^{\alpha+i\infty} e^{s\tau}\frac{1}{s^k}e^{-1/(2s^2)}e^{n/s}ds = \frac{1}{2\pi i}2\pi i \operatorname*{Res}_0\left(e^{s\tau}\frac{1}{s^k}e^{-1/(2s^2)}e^{n/s}ds\right).$$

У нас возникает два случая: $n = 0$ и $n \neq 0$

4.0.3 Случай $n = 0$

Воспользуемся теоремой о правильной части функции $e^{s\tau}e^{-1/(2s^2)}$. Нам нужен $k-1$-й член ряда Лорана. Получаем:

$$\frac{1}{2\pi i}\int\limits_{\alpha-i\infty}^{\alpha+i\infty} e^{s\tau}\frac{1}{s^k}e^{-1/(2s^2)}e^{n/s}ds = \sum_{j=0}^{\infty}\frac{\tau^{2j+k-1}/(-2)^j}{(2j+k-1)!\,j!}.$$

Таким образом, мы выразили вычислили **??**:

$$L_s^{-1}\left[\frac{1}{s^k}\exp\left(\frac{n}{s}\right)\exp\left(-\frac{1}{2s^2}\right)\right](\tau) = \sum_{j=0}^{\infty}\frac{\tau^{2j+k-1}/(-2)^j}{(2j+k-1)!\,j!}.$$

Подставим это выражение в **??**, заменяя τ на $t/2^{m+1}$:

$$L_u^{-1}\left[\frac{1}{u^k}r_{m,n}(u)\right](t) =$$
$$= e^{-\frac{n^2}{2}}2^{(k-1)(m+1)}\frac{2}{\sqrt{3}\pi^{1/4}}L_s^{-1}\left[\frac{1}{s^k}\exp\left(\frac{n}{s}-\frac{1}{2s^2}\right)\right]\left(\frac{t}{2^{m+1}}\right) =$$
$$e^{-\frac{n^2}{2}}2^{(k-1)(m+1)}\frac{2}{\sqrt{3}\pi^{1/4}}\sum_{j=0}^{\infty}\frac{\left(\frac{t}{2^{m+1}}\right)^{2j+k-1}/(-2)^j}{(2j+k-1)!\,j!}.$$

Наконец, подставим это в **??**:

$$L_u^{-1}\left[\frac{1}{u^k}\psi_{m,n}\left(\frac{1}{2^{m+1}u}-n\right)\right](t) =$$
$$= \left(1-n^2\right)L_u^{-1}\left[\frac{1}{u^k}r_{m,n}(u)\right](t) + \left(\frac{n}{2^m}\right)L_u^{-1}\left[\frac{1}{u^{k+1}}r_{m,n}(u)\right](t)+$$
$$+\left(\frac{1}{4^{m+1}}\right)L_u^{-1}\left[\frac{1}{u^{k+2}}r_{m,n}(u)\right](t) =$$
$$= \left(1-n^2\right)e^{-\frac{n^2}{2}}2^{(k-1)(m+1)}\frac{2}{\sqrt{3}\pi^{1/4}}\sum_{j=0}^{\infty}\frac{\left(\frac{t}{2^{m+1}}\right)^{2j+k-1}/(-2)^j}{(2j+k-1)!\,j!}+$$
$$+\left(\frac{n}{2^m}\right)e^{-\frac{n^2}{2}}2^{k(m+1)}\frac{2}{\sqrt{3}\pi^{1/4}}\sum_{j=0}^{\infty}\frac{\left(\frac{t}{2^{m+1}}\right)^{2j+k}/(-2)^j}{(2j+k)!\,j!}-$$
$$-\left(\frac{1}{4^{m+1}}\right)e^{-\frac{n^2}{2}}2^{(k+1)(m+1)}\frac{2}{\sqrt{3}\pi^{1/4}}\sum_{j=0}^{\infty}\frac{\left(\frac{t}{2^{m+1}}\right)^{2j+k+1}/(-2)^j}{(2j+k+1)!\,j!}. \quad (5)$$

Теперь получим выражение для $g(t)$, подставляя только что полученную формулу в **??**:

$$g_{m,n}(t) = \frac{1}{t^{k-1}}\frac{\Gamma(k)}{\sqrt{2^m}}L_u^{-1}\left[\frac{1}{u^k}\psi_{m,n}\left(\frac{1}{2^{m+1}u}-n\right)\right](t) =$$
$$= \frac{1}{t^{k-1}}\frac{\Gamma(k)}{\sqrt{2^m}}\left(1-n^2\right)e^{-\frac{n^2}{2}}2^{(k-1)(m+1)}\frac{2}{\sqrt{3}\pi^{1/4}}\sum_{j=0}^{\infty}\frac{\left(\frac{t}{2^{m+1}}\right)^{2j+k-1}/(-2)^j}{(2j+k-1)!\,j!}+$$
$$+\frac{1}{t^{k-1}}\frac{\Gamma(k)}{\sqrt{2^m}}\left(\frac{n}{2^m}\right)e^{-\frac{n^2}{2}}2^{k(m+1)}\frac{2}{\sqrt{3}\pi^{1/4}}\sum_{j=0}^{\infty}\frac{\left(\frac{t}{2^{m+1}}\right)^{2j+k}/(-2)^j}{(2j+k)!\,j!}-$$
$$-\frac{1}{t^{k-1}}\frac{\Gamma(k)}{\sqrt{2^m}}\left(\frac{1}{4^{m+1}}\right)e^{-\frac{n^2}{2}}2^{(k+1)(m+1)}\frac{2}{\sqrt{3}\pi^{1/4}}\sum_{j=0}^{\infty}\frac{\left(\frac{t}{2^{m+1}}\right)^{2j+k+1}/(-2)^j}{(2j+k+1)!\,j!}.$$

Начнем упрощать выражение. Для начала, вынесем из суммы, степень, не зависящую от переменной суммирования и сделаем замену $n = 0$ (так как рассматриваем именно этот случай):

$$g_{m,0}(t) = \frac{\Gamma(k)}{\sqrt{2^m}} \frac{2}{\sqrt{3}\pi^{1/4}} \left(\sum_{j=0}^{\infty} \frac{\left(\frac{t}{2^{m+1}}\right)^{2j}/(-2)^j}{(2j+k-1)!\,j!} - \left(\frac{t^2}{4^{m+1}}\right) \sum_{j=0}^{\infty} \frac{\left(\frac{t}{2^{m+1}}\right)^{2j}/(-2)^j}{(2j+k+1)!\,j!} \right).$$

4.0.4 Случай $n \neq 0$

Воспользуемся теоремой о правильной части функции $e^{s\tau}e^{-1/(2s^2)}e^{n/s}$. Нам нужен $k-1$-й член ряда Лорана. Получаем:

$$\frac{1}{2\pi i}\int_{\alpha-i\infty}^{\alpha+i\infty} e^{s\tau}\frac{1}{s^k}e^{-1/2s^2}e^{n/s}ds = \sum_{i=0}^{\infty}\sum_{j=0}^{\infty}\frac{\tau^{2j+k-1+i}/(-2)^j}{(2j+k-1+i)!\,j!}\frac{n^i}{i!}.$$

Таким образом, мы выразили вычислили **??**:

$$L_s^{-1}\left[\frac{1}{s^k}\exp\left(\frac{n}{s}\right)\exp\left(-\frac{1}{2s^2}\right)\right](\tau) = \sum_{i=0}^{\infty}\sum_{j=0}^{\infty}\frac{\tau^{2j+k-1+i}/(-2)^j}{(2j+k-1+i)!\,j!}\frac{n^i}{i!}.$$

Подставим это выражение в **??**, заменяя τ на $t/2^{m+1}$:

$$\begin{aligned}
&L_u^{-1}\left[\frac{1}{u^k}r_{m,n}(u)\right](t) = \\
&\qquad = e^{-\frac{n^2}{2}}2^{(k-1)(m+1)}\frac{2}{\sqrt{3}\pi^{1/4}}L_s^{-1}\left[\frac{1}{s^k}\exp\left(\frac{n}{s}-\frac{1}{2s^2}\right)\right]\left(\frac{t}{2^{m+1}}\right) = \\
&\qquad\qquad = e^{-\frac{n^2}{2}}2^{(k-1)(m+1)}\frac{2}{\sqrt{3}\pi^{1/4}}\sum_{i=0}^{\infty}\sum_{j=0}^{\infty}\frac{\left(\frac{t}{2^{m+1}}\right)^{2j+k-1+i}/(-2)^j}{(2j+k-1+i)!\,j!}\frac{n^i}{i!}.
\end{aligned}$$

Наконец, подставим это в **??**:

$$\begin{aligned}
&L_u^{-1}\left[\frac{1}{u^k}\psi_{m,n}\left(\frac{1}{2^{m+1}u}-n\right)\right](t) = \\
&\qquad = \left(1-n^2\right)L_u^{-1}\left[\frac{1}{u^k}r_{m,n}(u)\right](t) + \left(\frac{n}{2^m}\right)L_u^{-1}\left[\frac{1}{u^{k+1}}r_{m,n}(u)\right](t) + \\
&\qquad\qquad + \left(\frac{1}{4^{m+1}}\right)L_u^{-1}\left[\frac{1}{u^{k+2}}r_{m,n}(u)\right](t) = \\
&= \left(1-n^2\right)e^{-\frac{n^2}{2}}2^{(k-1)(m+1)}\frac{2}{\sqrt{3}\pi^{1/4}}\sum_{i=0}^{\infty}\sum_{j=0}^{\infty}\frac{\left(\frac{t}{2^{m+1}}\right)^{2j+k-1+i}/(-2)^j}{(2j+k-1+i)!\,j!}\frac{n^i}{i!} + \\
&\quad + \left(\frac{n}{2^m}\right)e^{-\frac{n^2}{2}}2^{k(m+1)}\frac{2}{\sqrt{3}\pi^{1/4}}\sum_{i=0}^{\infty}\sum_{j=0}^{\infty}\frac{\left(\frac{t}{2^{m+1}}\right)^{2j+k+i}/(-2)^j}{(2j+k+i)!\,j!}\frac{n^i}{i!} - \\
&\quad - \left(\frac{1}{4^{m+1}}\right)e^{-\frac{n^2}{2}}2^{(k+1)(m+1)}\frac{2}{\sqrt{3}\pi^{1/4}}\sum_{i=0}^{\infty}\sum_{j=0}^{\infty}\frac{\left(\frac{t}{2^{m+1}}\right)^{2j+k+1+i}/(-2)^j}{(2j+k+1+i)!\,j!}\frac{n^i}{i!}. \qquad (6)
\end{aligned}$$

Теперь получим выражение для $g_{m,n}(t)$, подставляя только что полученную формулу в **??**

$$g_{m,n}(t) = \frac{1}{t^{k-1}}\frac{\Gamma(k)}{\sqrt{2^m}}L_u^{-1}\left[\frac{1}{u^k}\psi_{m,n}\left(\frac{1}{2^{m+1}u}-n\right)\right](t) =$$
$$= \frac{1}{t^{k-1}}\frac{\Gamma(k)}{\sqrt{2^m}}\left(1-n^2\right)e^{-\frac{n^2}{2}}2^{(k-1)(m+1)}\frac{2}{\sqrt{3}\pi^{1/4}}\sum_{i=0}^{\infty}\sum_{j=0}^{\infty}\frac{\left(\frac{t}{2^{m+1}}\right)^{2j+k-1+i}/(-2)^j}{(2j+k-1+i)!\,j!}\frac{n^i}{i!}+$$
$$+\frac{1}{t^{k-1}}\frac{\Gamma(k)}{\sqrt{2^m}}\left(\frac{n}{2^m}\right)e^{-\frac{n^2}{2}}2^{k(m+1)}\frac{2}{\sqrt{3}\pi^{1/4}}\sum_{i=0}^{\infty}\sum_{j=0}^{\infty}\frac{\left(\frac{t}{2^{m+1}}\right)^{2j+k+i}/(-2)^j}{(2j+k+i)!\,j!}\frac{n^i}{i!}-$$
$$-\frac{1}{t^{k-1}}\frac{\Gamma(k)}{\sqrt{2^m}}\left(\frac{1}{4^{m+1}}\right)e^{-\frac{n^2}{2}}2^{(k+1)(m+1)}\frac{2}{\sqrt{3}\pi^{1/4}}\sum_{i=0}^{\infty}\sum_{j=0}^{\infty}\frac{\left(\frac{t}{2^{m+1}}\right)^{2j+k+1+i}/(-2)^j}{(2j+k+1+i)!\,j!}\frac{n^i}{i!}.$$

Упростим:

$$g_{m,n}(t) = \frac{\Gamma(k)}{\sqrt{2^m}}e^{-\frac{n^2}{2}}\frac{2}{\sqrt{3}\pi^{1/4}}\left(\left(1-n^2\right)\sum_{i=0}^{\infty}\sum_{j=0}^{\infty}\frac{\left(\frac{t}{2^{m+1}}\right)^{2j+i}/(-2)^j}{(2j+k-1+i)!\,j!}\frac{n^i}{i!}+\right.$$
$$+\left(\frac{nt}{2^m}\right)\sum_{i=0}^{\infty}\sum_{j=0}^{\infty}\frac{\left(\frac{t}{2^{m+1}}\right)^{2j+i}/(-2)^j}{(2j+k+i)!\,j!}\frac{n^i}{i!}-$$
$$\left.-\left(\frac{t^2}{4^{m+1}}\right)\sum_{i=0}^{\infty}\sum_{j=0}^{\infty}\frac{\left(\frac{t}{2^{m+1}}\right)^{2j+i}/(-2)^j}{(2j+k+1+i)!\,j!}\frac{n^i}{i!}\right).$$

4.1 Результат

Выпишем обе полученные формулы вместе:

$$g_{m,0}(t) = \frac{\Gamma(k)}{\sqrt{2^m}}\frac{2}{\sqrt{3}\pi^{1/4}}\left(\sum_{j=0}^{\infty}\frac{\left(\frac{t}{2^{m+1}}\right)^{2j}/(-2)^j}{(2j+k-1)!\,j!}-\left(\frac{t^2}{4^{m+1}}\right)\sum_{j=0}^{\infty}\frac{\left(\frac{t}{2^{m+1}}\right)^{2j}/(-2)^j}{(2j+k+1)!\,j!}\right).$$

$$g_{m,n}(t) = \frac{\Gamma(k)}{\sqrt{2^m}}e^{-\frac{n^2}{2}}\frac{2}{\sqrt{3}\pi^{1/4}}\left(\left(1-n^2\right)\sum_{i=0}^{\infty}\sum_{j=0}^{\infty}\frac{\left(\frac{t}{2^{m+1}}\right)^{2j+i}/(-2)^j}{(2j+k-1+i)!\,j!}\frac{n^i}{i!}+\right.$$
$$+\left(\frac{nt}{2^m}\right)\sum_{i=0}^{\infty}\sum_{j=0}^{\infty}\frac{\left(\frac{t}{2^{m+1}}\right)^{2j+i}/(-2)^j}{(2j+k+i)!\,j!}\frac{n^i}{i!}-$$
$$\left.-\left(\frac{t^2}{4^{m+1}}\right)\sum_{i=0}^{\infty}\sum_{j=0}^{\infty}\frac{\left(\frac{t}{2^{m+1}}\right)^{2j+i}/(-2)^j}{(2j+k+1+i)!\,j!}\frac{n^i}{i!}\right).$$

К сожалению, такой способ не привел к успеху из-за непригодности для численных методов.

5 Альтернативный подход к задаче

Вспомним наше изначальное интегральное уравнение:

$$\psi_{m,n}(x) = \int\limits_0^\infty g(xy) f_Y(y) dy.$$

Преобразуем интеграл, чтобы интегрирование было по xy:

$$\psi_{m,n}(x) = \int\limits_0^\infty g(xy) f_Y(\frac{xy}{x}) d\frac{xy}{x} = \int\limits_0^\infty \int\limits_0^\infty g(z)\frac{1}{x} f_Y(\frac{z}{x}) dz.$$

Таким образом мы получили интегральное уравнение Фредгольма первого рода:

$$\psi_{m,n}(x) = \int\limits_0^\infty K(x,z) g(z) dz.$$

Дальше мы будем использовать равномерную сетку $\left[\frac{1}{n_x}, \ldots, \frac{l_x n_x}{n_x}\right]$ для x, $\left[\frac{1}{n_z}, \ldots, \frac{l_z n_z}{n_z}\right]$ для z и дискретизируем наше уравнение. Получаем:

$$\psi_{m,n}[x] = \int\limits_0^\infty K[x,z] g[z] dz = \frac{1}{n_z} \sum_{p=1}^{l_z n_z} g\left(\frac{p}{n_z}\right) K\left[x, \frac{p}{n_z}\right].$$

Таким образом, мы получили систему линейных уравнений. Запишем их в матричном виде:

$$\boldsymbol{\psi}_{m,n} = \frac{1}{n_z} \boldsymbol{K}\boldsymbol{g}.$$

Увеличим матрицу K, чтобы добавить регуляризацию.

$$\tilde{\boldsymbol{K}} = \begin{pmatrix} \boldsymbol{K} \\ \alpha \boldsymbol{E} \end{pmatrix}.$$

И соответствующий $\tilde{\boldsymbol{f}}$:

$$\tilde{\boldsymbol{f}} = \begin{pmatrix} \boldsymbol{f} \\ \boldsymbol{0} \end{pmatrix}.$$

И будем использовать МНК-оптимизацию. Получаем:

$$\boldsymbol{g}_* = \arg\min_{\boldsymbol{g}} \|\tilde{\boldsymbol{K}}\boldsymbol{g} - \boldsymbol{f}\|.$$

5.1 Градиентный спуск

Вместо процедур для решения МНК-задачи мы можем использовать метод градиентного спуска. Будем использовать матричное представление

$$\boldsymbol{\psi}_{m,n} = \frac{1}{n_z}\boldsymbol{K}\boldsymbol{g}.$$

Тогда можно ввести функцию потери $L(\boldsymbol{\psi_{m,n}}, \boldsymbol{\hat{\psi}_{m,n}})$, где $\boldsymbol{\hat{\psi}_{m,n}} = \boldsymbol{K}\boldsymbol{\hat{g}_{m,n}}$, а $\boldsymbol{\hat{g}_{m,n}}$ — оценка для $\boldsymbol{g_{m,n}}$.

В частности, будем рассматривать следующие функции потерь:

- $l1$-потеря: $L(\boldsymbol{x}, \boldsymbol{y}) = \|x - y\|_1$;
- $l2$-потеря: $L(\boldsymbol{x}, \boldsymbol{y}) = \|x - y\|_2$;
- функция потери Хьюбера:

$$L(x, y) = \begin{cases} \frac{1}{2}(x-y)^2, \text{ при } |x-y| \leqslant 1 \\ |x-y| - \frac{1}{2}, \text{ при } |x-y| > 1 \end{cases}.$$

$$L(\boldsymbol{x}, \boldsymbol{y}) = \frac{1}{n}\sum_{i=1}^{k} L(x_i, y_i)$$

Для каждой из них будем использовать L_1- или L_2-регуляризацию:

$$\tilde{L}(\boldsymbol{\psi_{m,n}}, \boldsymbol{\hat{\psi}_{m,n}}) = L(\boldsymbol{\psi_{m,n}}, \boldsymbol{\hat{\psi}_{m,n}}) + \|\boldsymbol{g_{m,n}} - \boldsymbol{\hat{g}_{m,n}}\|$$

5.2 Итеративные методы

В статье [1] рассматриваются итеративные методы решения задачи Фредгольма первого рода: аддитивный и мультипликативный.

В приложении к задаче аддитивный метод использует следующие итерации:

$$g_{m,n;k}(z) = g_{m,n;k-1}(z) = \int\limits_0^\infty K(x,z)(\psi_{m,n;k}(x) - \psi_{m,n}(x))dx,$$

$$\psi_{m,n;k}(x) = \int\limits_0^\infty K(x,z)g_{m,n;k}(z)dz.$$

Для мультипликативного метода используются такие итерации:

$$g_{m,n;k}(z) = \frac{g_{m,n;k-1}(z)}{\int_0^\infty K(x,z)dx}\int\limits_0^\infty \frac{K(x,z)\psi_{m,n}(x)}{\psi_{m,n;k}(x)}dx,$$

$$\psi_{m,n;k}(x) = \int\limits_0^\infty K(x,z) g_{m,n;k}(z) dz.$$

Так как ψ и g могут принимать отрицательные значения, производится следующее преобразование: выбирается параметр t, $\psi_{m,n}$ заменяется на $\tilde{\psi}_{m,n} = \psi_{m,n} + t$, $f_{m,n;0}$ заменятся на $\tilde{f}_{m,n;0} = f_{m,n;0} + t$.

5.3 Поправка для оценок

Будем также использовать поправку, предложенную в статье [2] В ней рассматриваются два случая: когда интеграл

$$\int \max(\hat{f}(x), 0) dx$$

больше 1, и когда меньше единицы.

В первом случае оценка $\hat{f}$ заменяется на $\tilde{f}(x) = \max(0, \hat{f}(x) - \xi)$, где ξ выбирается так, чтобы выполнялось

$$\int \tilde{f}(x) dx = 1.$$

Во втором случае используется оценка

$$\tilde{f}(x) = \tilde{f}(x; M) = \begin{cases} \max(0, \hat{f}(x)) + \eta_M, \text{ для } |x| \leqslant M, \\ \max(0, \hat{f}(x)), \text{ для } |x| > M, \end{cases}$$

где

$$\eta_M = \frac{1}{2M} \left(1 - \int \max(0, \hat{f}(x)) dx \right).$$

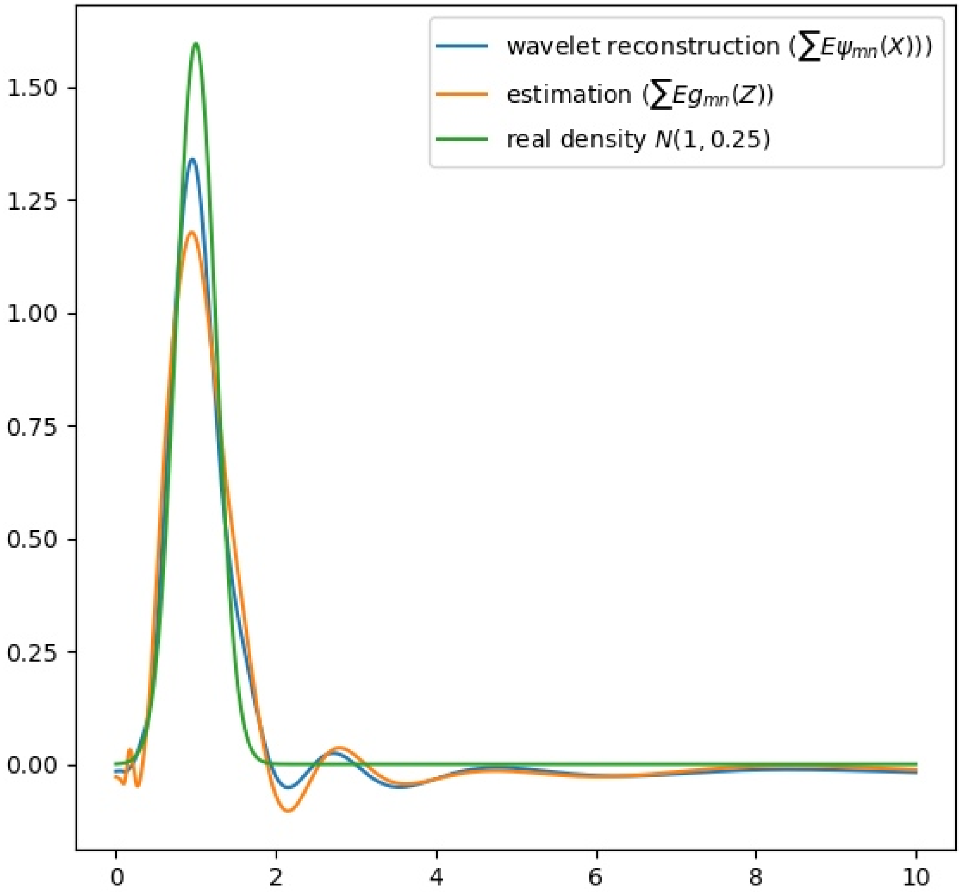
6 Эксперименты

Для аналитического способа.

Функция	Способ вычисления	Машинная точность (размер мантиссы),	Значение
$g_{0,0}(1)$	численно, интеграл, контур $[1-100i, 1+100i]$	100 десятичных знаков	0.864
	численно, ряд	256 двоичных знаков	0.864
$g_{0,0}(10)$	численно, интеграл контур $[1-100i, 1+100i]$	100 десятичных знаков	0.591
	численно, ряд	256 двоичных знаков	0.591
$g_{0,0}(100)$	численно, интеграл контур $[1-10i, 1+10i]$	100 десятичных знаков	-2×10^{19}
	численно, ряд	256 двоичных знаков	-0.440

Для численного вычисления интеграла. Мы использовали шаг 0,1 и $\alpha = 0{,}1$. И использовали $m = \{-5, \ldots, 5\}$, $n = \{-5, \ldots, 5\}$.

Рис. 1: $X \sim \mathcal{N}(0, 1)$



7 Обобщение на случай разных длин траекторий

Мы строили функции вида:

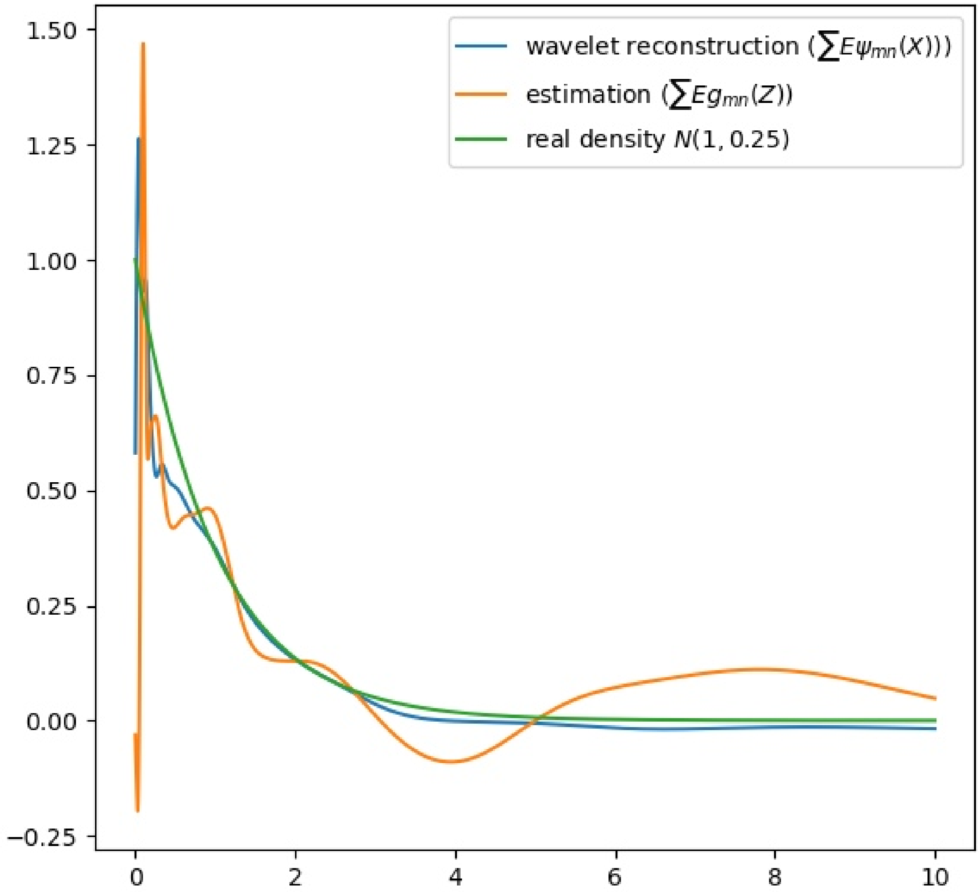
$$\mathbf{E} g_{m,n}(XY) = \mathbf{E} g_{m,n}(X) = c_{m,n}$$

и находили оценку плотности как

$$f_X(x) = c_{m,n}\psi_{m,n}(x).$$

Теперь рассмотрим случай, когда длины траекторий могут различаться. Для каждой длины k построим функции $g_{m,n,k}$ как описано выше и построим оценку $f_{X,k}(x)$

Рис. 2: $X \sim \exp(1)$



Пусть для длины траектории k у нас есть s_k наблюдений. И всего S наблюдений Тогда оценкой $f_X(x)$ будет

$$\sum_{k=1}^{K} \frac{s_k f_{X,k}(x)}{S}.$$

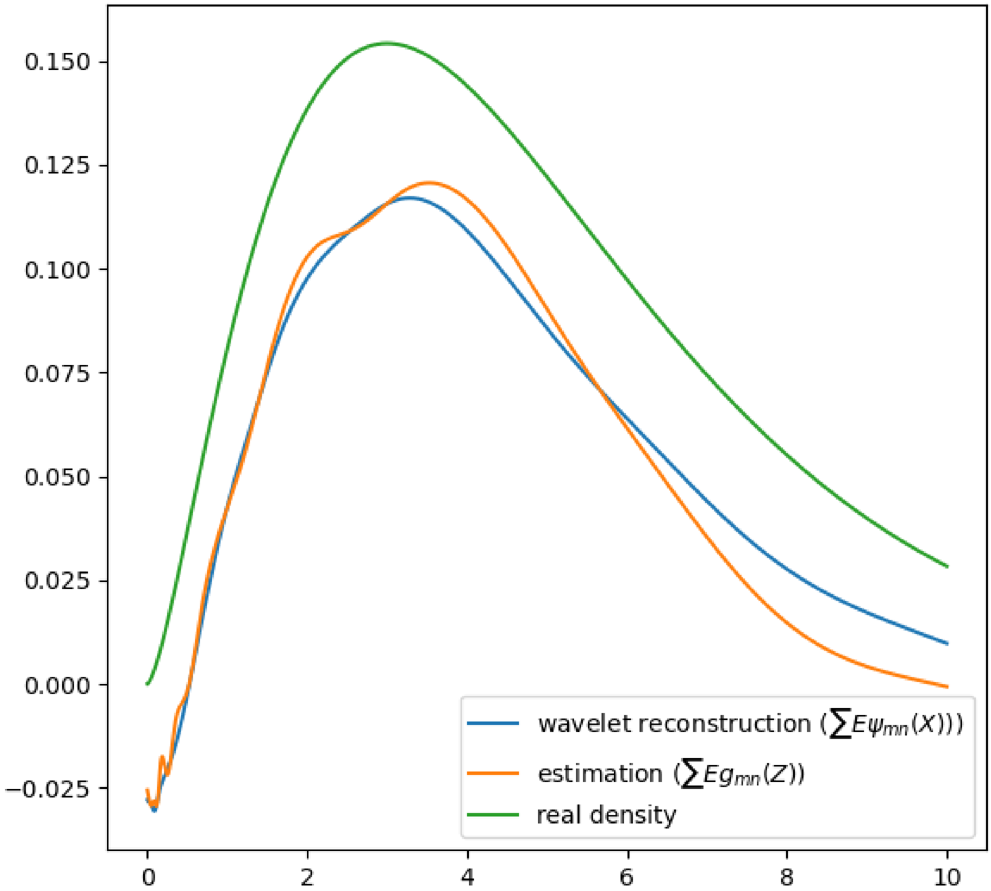
Докажем это. Разложим f_X в ряд по вейвлету:

$$f_X(x) = \sum_{m,n} c_{m,n} \psi_{m,n}(x).$$

Раскроем вейвлет-коэффициенты:

$$f_X(x) = \sum_{m,n} \mathbf{E} \psi_{m,n}(XY) \psi_{m,n}(x).$$

Рис. 3: $X \sim \chi_5^2$



Представим математическое ожидание в виде математического ожидания условного математического ожидания при условии длины траектории:

$$f_X(x) = \sum_{m,n} \mathbf{E}_k \left(\mathbf{E}\left(\psi_{m,n}(XY)|k\right)\right) \psi_{m,n}(x).$$

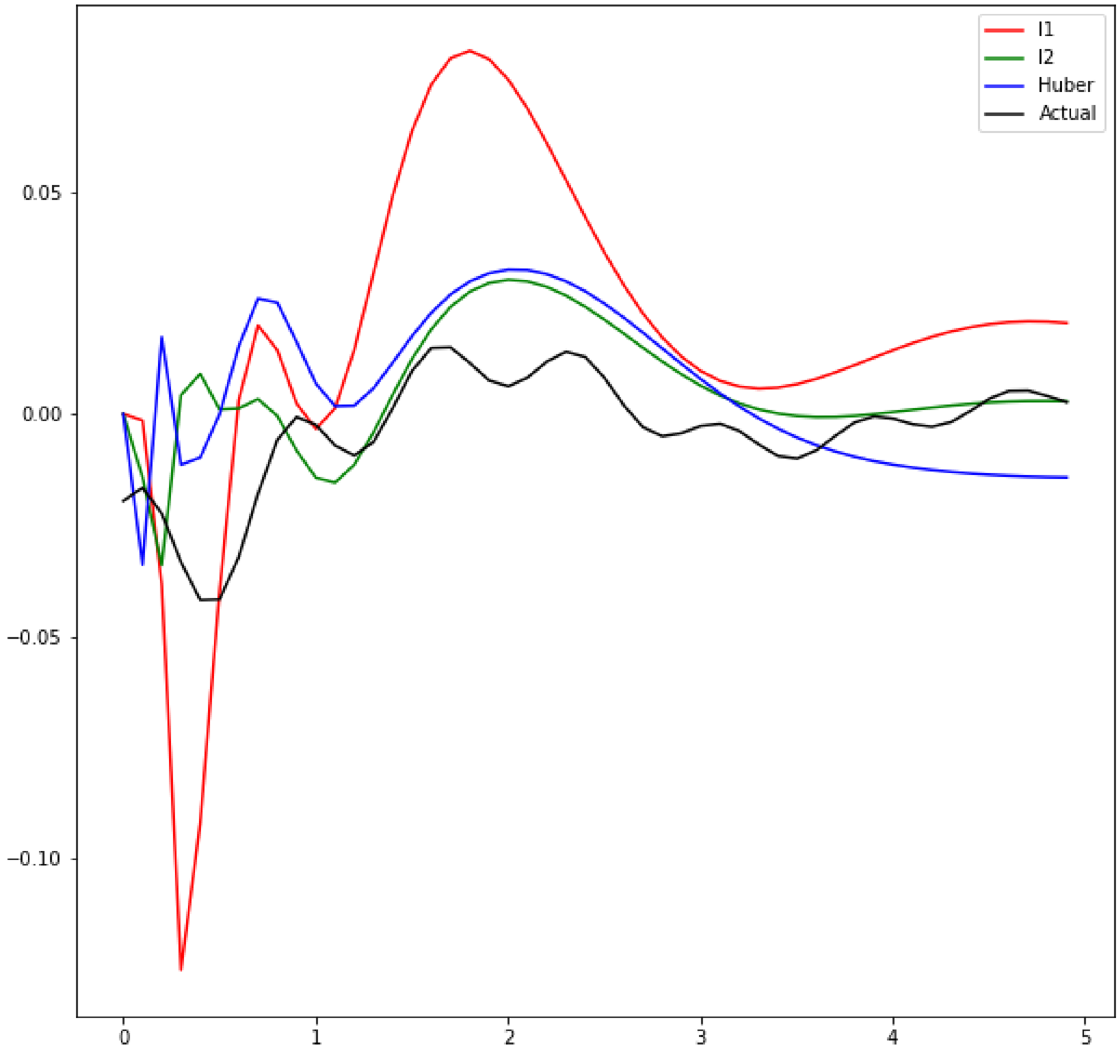
По линейности математического ожидания, можем внести сумму внутрь:

$$f_X(x) = \mathbf{E}_k \left(\sum_{m,n} \mathbf{E}\left(\psi_{m,n}(XY)|k\right) \psi_{m,n}(x) \right).$$

Вычислим вейвлет-коэффициенты:

$$f_X(x) = \mathbf{E}_k \left(\sum_{m,n} c_{m,n,k} \psi_{m,n}(x) \right).$$

Рис. 4: Сравнение функций ошибок для метода градиентного спуска



Заменим вейвлет-разложение на оригинальную функцию:

$$f_X(x) = \mathbf{E}_k f_{X,k}(x).$$

Получаем оценку:

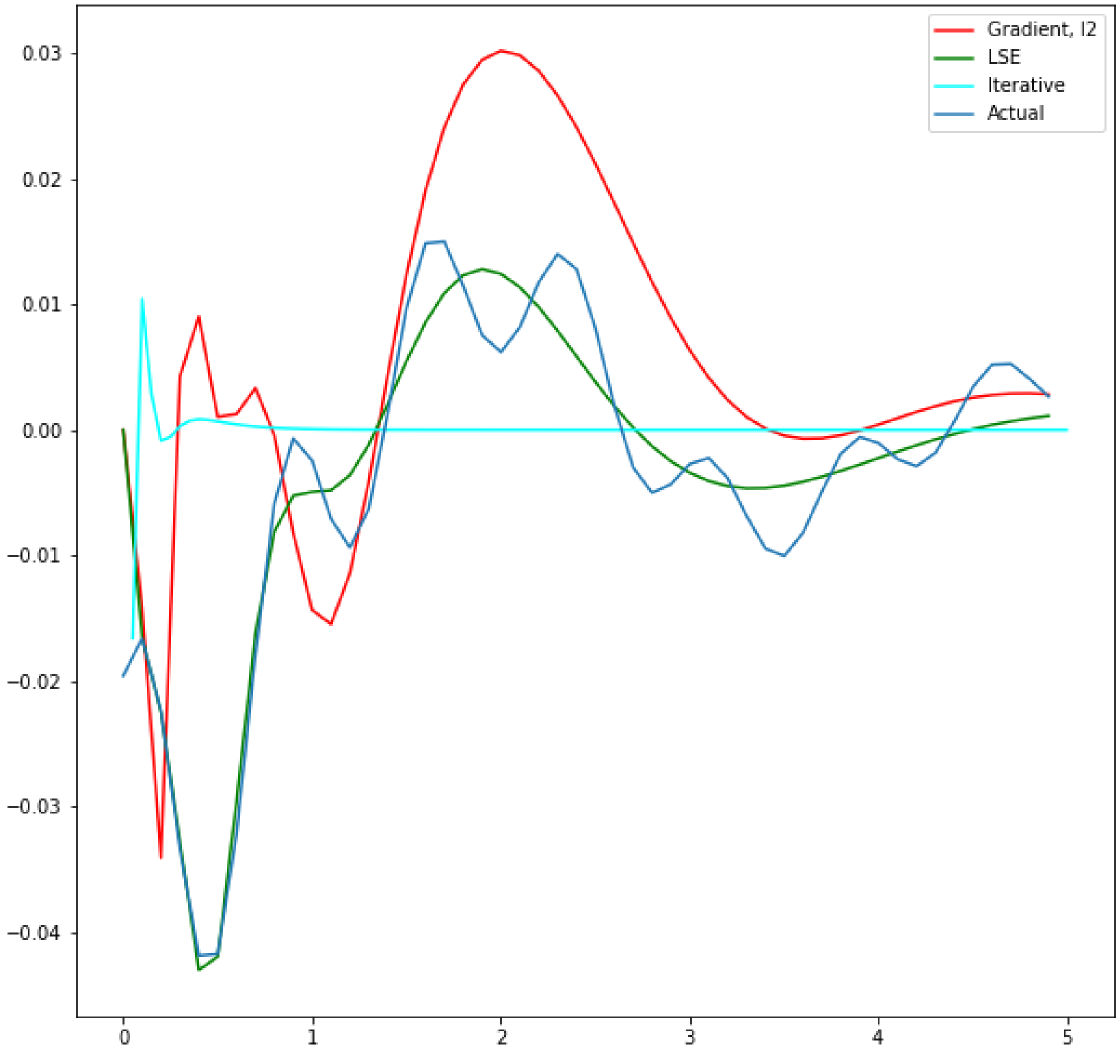
$$f_X(x) = \sum_{k=1}^{K} \frac{s_k f_{X,k}(x)}{S}.$$

8 Вывод

Лучшие результаты показывает МНК-оценка.

Оценка методом градиентного спуска более шумная, но позволяет использовать существенно более точный шаг дискретизации, так как возможно пожерт-

Рис. 5: Сравнение методов градиентного спуска, итеративного и МНК-оценки



вовать произодительностью и не вычислять матрицу K заранее, что существенно снижает требования к количеству видеопамяти.

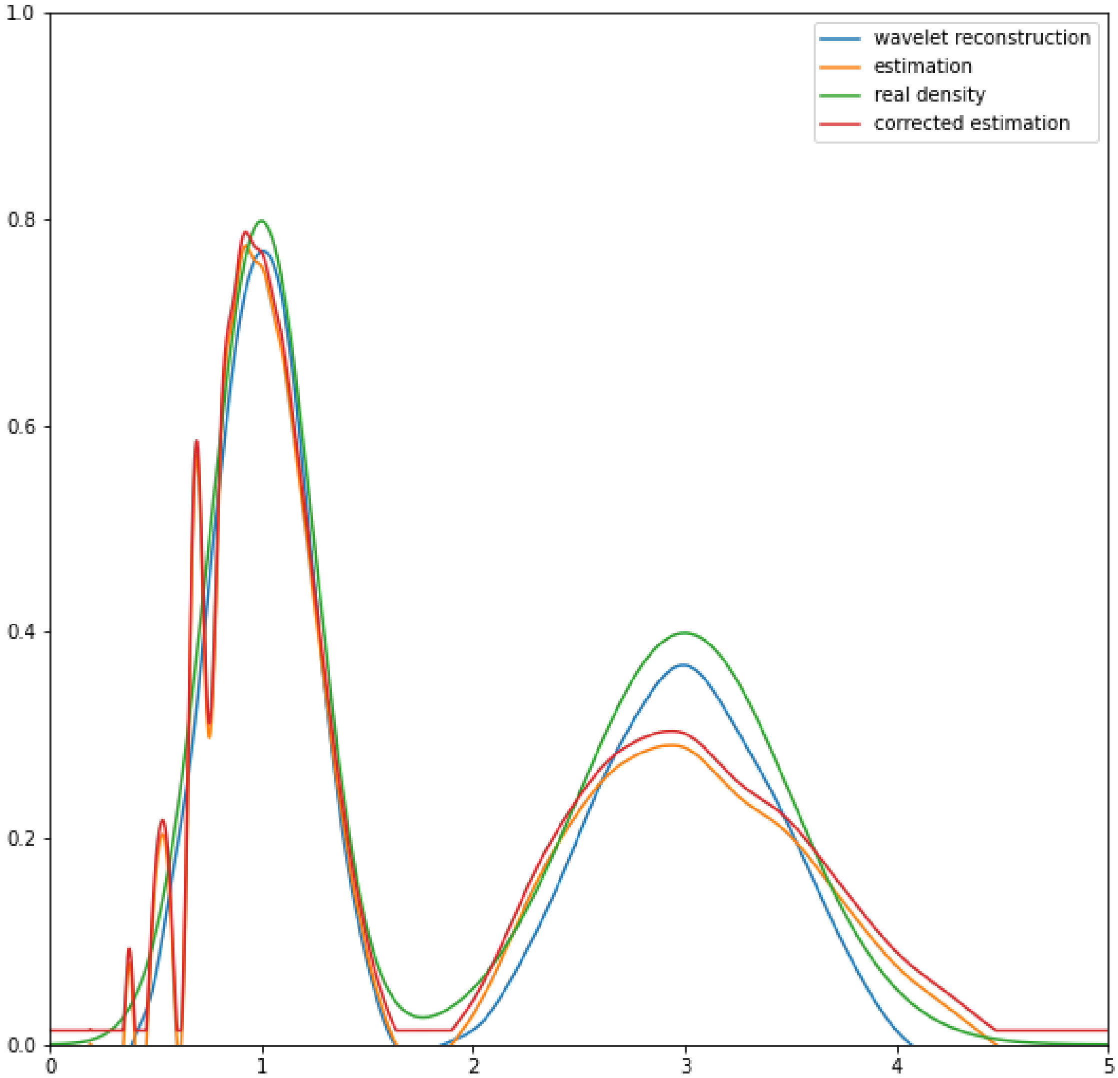
Итеративная оценка показывает неудовлетворительные результаты и сходится крайне медленно: разница между 1000 итераций и 10000 итераций несущественна.

Поправка для оценок плотностей несильно улучшает оценку.

Список литературы

[1] Minwoo Chae, Ryan Martin и Stephen G. Walker. “On an algorithm for solving Fredholm integrals of the first kind”. В: *Statistics and Computing* 29.4 (июль 2019), с. 645—654. ISSN: 1573-1375. DOI: 10.1007/s11222-018-9829-z. URL: https://doi.org/10.1007/s11222-018-9829-z.

Рис. 6: МНК-оценка для смеси нормальных распределений



[2] Ingrid K. Glad, Nils Lid Hjort и Nikolai G. Ushakov. "Correction of Density Estimators That Are Not Densities". В: *Scandinavian Journal of Statistics* 30.2 (2003), с. 415—427. ISSN: 03036898, 14679469. URL: http://www.jstor.org/stable/4616772.

Рис. 7: Оценка методом градиентного спуска для смеси нормальных распределений

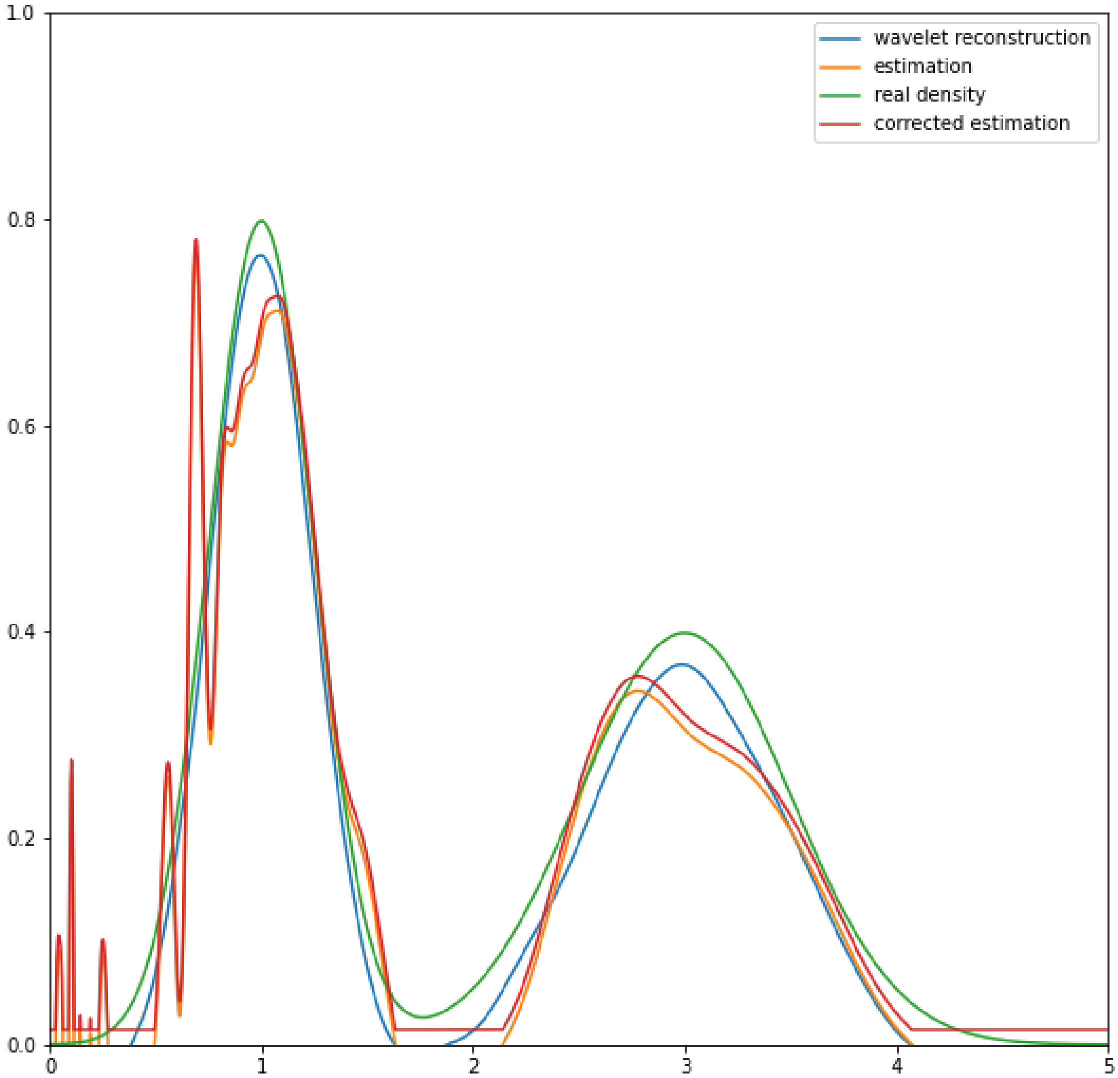


Рис. 8: Оценка итеративным методом для смеси нормальных распределений