Интерполяция и аппроксимация таблично заданной функции

Богачев Владимир Александрович

Декабрь 2022

1 Постановка задачи

Дана таблично заданная функция (x_i, f_i), $i = \overline{0, N}$. Необходимо построить интерполяционный многочлен Лагранжа и аппроксимировать данную функцию многочленом степени не выше n.

2 Построение приближений непрерывными функциями

В данном разделе будут описаны два алгоритма:

1. Алгоритм построения многочлена Лагранжа
2. Алгоритм аппроксимации функции многочленом заданной степени

В ходе описания алгоритмов будут показаны выявленные преимущества и недостатки. В следующей главе будет проведено подробное сравнение данных алгоритмов, в ходе которого будут изучены выявленные преимущества и недостаки.

2.1 Интерполяция многочленом Лагранжа

Интерполяционный многочлен Лагранжа строится по следующей формуле:

$$Q(x) = \sum_{i=0}^{N} f_i \cdot \prod_{j=0, j \neq i}^{N} \frac{x - x_j}{x_i - x_j} \tag{1}$$

Так как $\prod_{j=0,j\neq i}^{N} \frac{1}{x_i - x_j}$ не зависит от x для $i = \overline{0,\ N}$, то можно вычислить вспомогательные значения

$$q_i = \prod_{j=0, j \neq i}^{N} \frac{1}{x_i - x_j}$$

Это позволит переписать формулу 1 в виде

$$Q(x) = \sum_{i=0}^{N} f_i \cdot \prod_{j=0, j\neq i}^{N} q_i \cdot (x - x_i) \tag{2}$$

Для вычисления значения многочлена $Q(x)$ в точке x необходимо провести $N \cdot (N+1)$ операций. Следовательно, сложность данного алгоритма $\underline{O}(N^2)$

Далее, для построения оценок, будем предполагать, что таблично заданная функция (x_i, f_i) порождена некоторой неизвестной функцией $f(x)$, причем $f(x_i) = f_i \quad i = \overline{0, N}$

Теорема 1. *Теорема о сходимости: если $f(x) \in C^{N+1}([a, b])$, то*

$$|f(x) - Q(x)| < C \cdot M_{N+1} \cdot \left(\frac{b-a}{2N}\right)^{N+1}$$

Где $M_{N+1} = \max\limits_{[a,b]} |f^{N+1}(x)|$

Так как $x_0 < x_1 < \cdots < x_N$, то можно гарантировать, что $deg(Q) = N$. Следовательно, полученный многочлен может иметь довольно высокую амплитуду. Согласно теореме 1, если $f(x) \in C^{\infty}[a, b]$, то

$$\lim_{N\to\infty} (f(x) - Q(x)) = 0$$

Следовательно, необходимо выяснить, достаточно ли N точек для устойчивого поведения функции.

Стоит заметить, что по построению интерполяционный многочлен проходит через все заданные точки $(x_i,\, f_i) \quad i = \overline{0,\, N}$.

2.2 Апроксимация многочленом заданной степени

Рассмотрим многочлен степени n:

$$F(x) = a_0 + a_1 x + \ldots + a_n x^n$$

Где

$$(a_0, \ldots, a_n) = \underset{a_0, \ldots, a_n}{\arg\min} \left(\sum_{k=0}^{N} (F(x_k) - f_k)^2 \right) \tag{3}$$

Для решения поставленной оптимизационной задачи достаточно найти решение системы

$$A^* A\, \alpha = A^* f \tag{4}$$

Где

$$A = \begin{pmatrix} 1 & x_1 & x_1^2 & \ldots & x_1^n \\ 1 & x_2 & x_2^2 & \ldots & x_2^n \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 1 & x_N & x_N^2 & \ldots & x_N^n \end{pmatrix}_{N\times(n+1)}$$

$$\alpha = (a_0, a_1, \dots, a_n)^T, \quad f = (f_1, f_2, \dots, f_N)^T$$

Согласно теореме об альтернативе Фредгольма, система 4 разрешима для $\forall A \in \mathbb{R}^{N\times(n+1)}$

Для решения системы линейных алгебраических уравнений можно использовать метод Гаусса. Неравенство $[A^*A]_{i,i} = \sum_{k=1}^{n+1}[A]_{k,i}^2 \geq [A]_{1,i}^2 = 1$ гарантирует корректность работы алгоритма без перестановки строк.

Система 4 является плохо обусловленной, так как

$$||A|| \cdot ||A^{-1}|| \approx 3 \cdot 10^5 \gg 1$$

Следовательно при сравнении методов будет необходимо исследовать устойчивость коэффициентов при малых изменениях сетки x_i.

Сложность алгоритма вычисления коэффициентов $(a_0, \dots, a_n) - \underline{O}(n^3)$. Для вычисления значения многочлена $F(x)$ в точке x можно использовать схему Горнора со сложностью $\underline{O}(n)$.

3 Анализ полученных результатов

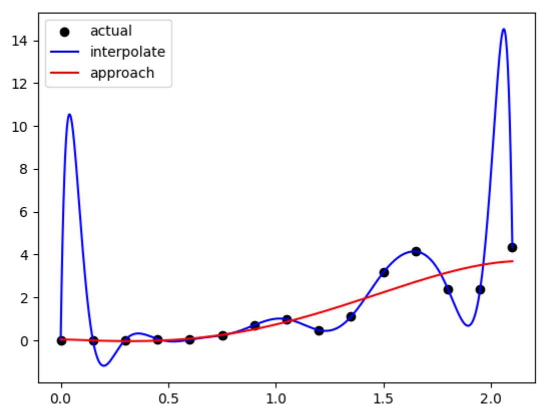

Рис. 1: Графики (x_i, f_i), $Q(x)$, $F(x)$

Для начала стоит понять, насколько сильно полученные многочлены $Q(x)$ и $F(x)$ отклоняются от заданных значений. Интерполяционный многочлен $Q(x)$ по построению проходит через все заданные точки (x_i, f_i). Аппроксимирующая функция не обязана проходить через заданные точки.

$$\sqrt{\sum_{k=1}^{N}(F(x_k) - f_k)^2} \approx 2.5$$

У обоих методов есть общий гипотетический недостаток — неустойчивость к малым изменениям исходной сетки. Далее будет проведен анализ данного гипотетического недостатка.

3.1 Анализ устойчивости к изменению количества точек сетки

Как было замечено в главе 2, многочлен $Q(x)$ может иметь достаточно высокую амплитуду. Данное предположение подтвердилость на практике, построенный многочлен Лагранжа действительно имеет значительно более высокую амплитуду, чем исходно заданная табличная функция. Это наводит на мысль, что N точек не хватает для устойчивого поведения функции, т.е. при удалении или добавлении новой точки интерполяционный многочлен может сильно измениться.

Рассмотрим таблично заданную функцию $g^{(i)}$ $\left(i = \overline{0, N}\right)$ с сеткой

$$(x_0, f_0),\, \ldots,\, (x_{i-1}, f_{i-1}),\, (x_{i+1}, f_{i+1}),\, \ldots,\, (x_N, f_N)$$

Пусть $Q_i(x)$ — интерполяционный многочлен Лагранжа для функции $g^{(i)}$, $F_i(x)$ — аппроксимация $g^{(i)}$ многочленом степени n.

Тогда

$$|Q_i(x) - Q_j(x)| \leq |f(x) - Q_i(x)| + |f(x) - Q_j(x)|$$

Следовательно, величина $|Q_i(x) - Q_j(x)|$ позволяет оценить снизу величину $|f(x) - Q_i(x)|$, а следовательно, и величину $|f(x) - Q(x)|$.

Построим матрицы устойчивости

$$[\mathcal{L}]_{i,j} = \sqrt{\int_0^2 (Q_i(x) - Q_j(x))^2\, dx}$$

$$[\mathcal{A}]_{i,j} = \sqrt{\int_0^2 (F_i(x) - F_j(x))^2\, dx}$$

$$\mathbb{E}([\mathcal{L}]_{i,j}) = \frac{1}{(N+1)^2} \sum_{i=0}^{N} \sum_{j=0}^{N} [\mathcal{L}]_{i,j} \approx 33.18$$

$$\mathbb{E}([\mathcal{A}]_{i,j}) = \frac{1}{(N+1)^2} \sum_{i=0}^{N} \sum_{j=0}^{N} [\mathcal{A}]_{i,j} \approx 0.28$$

$$\mathbb{E}([\mathcal{L}]_{i,j}) \gg \mathbb{E}([\mathcal{A}]_{i,j})$$

Следовательно, удаление одной точки из сетки значительно сказывается на интерполяционном многочлене Лагранжа и незначительно — на аппроксимирующей функции.

3.2 Анализ устойчивости к изменению x_i

Как было замечено в главе 2, решение задачи 3 может оказаться неустойчивым в силу плохой обусловленности матрицы 2.2. Следовательно, необходимо исследовать поведение интерполяционного и аппроксимирующего многочленов при малых изменениях сетки.

$\delta = (\delta_0, \ldots, , \delta_N)$ — (N+1)-мерный нормально распределенный случайный вектор с нулевым математическим ожиданием и дисперсией σ^2. F_δ — аппроксимация, Q_δ — интерполяционный многочлен Лагранжа таблично заданной функции $(x_i + \delta_i, f_i)$ $\quad i = \overline{0, N}$. F_δ, Q_δ являются случайными величинами.

$$\Delta_Q(\sigma) = \mathbb{E}\sqrt{\int_0^2 (Q_\delta(x) - Q(x))^2\, dx}$$

$$\Delta_F(\sigma) = \mathbb{E}\sqrt{\int_0^2 (F_\delta(x) - F(x))^2\, dx}$$

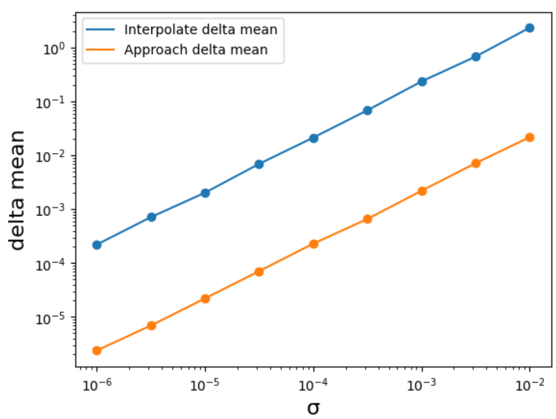

Рис. 2: Графики $\Delta_Q(\sigma)$, $\Delta_F(\sigma)$

Из полученных данных следует, что $\Delta_Q(\sigma) \gg \Delta_F(\sigma)$ для всех рассмотренных значений σ. Следовательно, аппроксимация многочленом степени не выше n, несмотря на плохую обусловленность системы 4, более устойчивая к малым изменениям сетки, чем интерполяция многочленом Лагранжа.

3.3 Анализ устойчивости к изменению f_i

В данном разделе будет проведено исследование устойчивости к изменению ординат сетки. План исследования полностью совпадает с планом исследования устойчивости к изменению абсцисс сетки.

$\delta = (\delta_0, \ldots, , \delta_N)$ — (N+1)-мерный нормально распределенный случайный вектор с нулевым математическим ожиданием и дисперсией σ^2. F_δ — аппроксимация, Q_δ — интерполяционный многочлен Лагранжа таблично заданной функции $(x_i,\ f_i + \delta_i)\quad i = \overline{0,\ N}$. F_δ, Q_δ являются случайными величинами.

$$\Delta_Q(\sigma) = \mathbb{E}\sqrt{\int_0^2 (Q_\delta(x) - Q(x))^2\,dx}$$

$$\Delta_F(\sigma) = \mathbb{E}\sqrt{\int_0^2 (F_\delta(x) - F(x))^2\,dx}$$

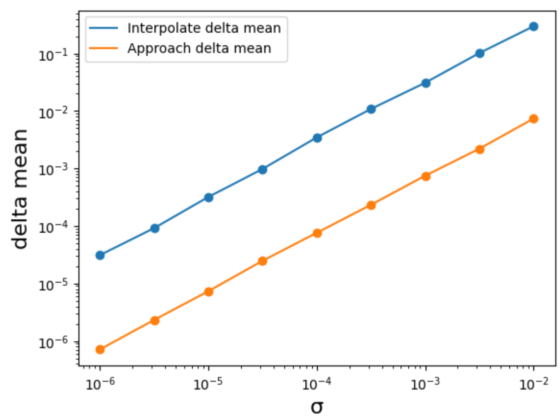

Рис. 3: Графики $\Delta_Q(\sigma)$, $\Delta_F(\sigma)$

Реакция алгоритмов на малые изменения ординат сетки совпадает с реакцией на малые изменения абсцисс сетки.

4 Выводы

В ходе работы было исследовано поведение интерполяционного многочлена Лагранжа и аппроксимирующей функции для таблично заданной функции $(x_i, f_i)\quad i = \overline{0,\ N}$. В ходе сравнения данных способов приближения было выяснено, что интерполяционный многочлен Лагранжа не является устойчивым при данном соотношении количества точек и длины отрезка. Малые изменения таблично заданной функции приводят к сильным изменениям интерполяционной функции. Особенно сильно сказывается удаление одной точки из сетки. Аппроксимирующая функция оказалась более устойчивой. Если предположить, что исходная таблично заданная функция описывает некоторый реальный физический процесс, то устойчивость аппроксимирующей функции даёт основания полагать, что данный способ приближения лучше соответствует описываемому физическому процессу.