

ЗАДАЧА ОПТИМИЗАЦИИ ДЛЯ АППРОКСИМАЦИИ КРИВОЙ И ЕГО ПРИМЕНЕНИЕ

Ийри сызыктардын аппроксимациясын оптималдаштыруу маселеси жана анын колдонулушу

Curve approximation optimization problem and its applications

Аннотация: В данной статье предлагается алгоритм нахождения точек аппроксимации на контуре объекта, основанный на методах решения задач оптимизации. Основная идея в том, что масштабирование расстояний между точками дает взаимно-однозначное соответствие между контуром и точкой в векторном пространстве. Помимо задачи сопоставления объектов, данный алгоритм имеет широкий спектр применения, таких как численное интегрирование и вычисление производных с указанным числом узлов сетки. В работе показаны сравнения указанных задач на конкретных примерах.

Аннотация: Бул макалада, оптималдаштыруу маселелерди чечүү жолунун негизинде объекттерди контуру боюнча жакындаштыруу чекиттеринин табуу үчүн алгоритм сунуштайбыз. Иштин негизги идеясы, табылган жакындаштыруу чекиттердин аралыктары жана чекиттер менен контурдун ортосундагы аралыктар, вектордук мейкиндикте бирге-бир чагылтуу бергенинен чыгып келет. Объекттерди контуру боюнча дал келишин табуу жана таануудан тышкары, бул алгоритм, сандын туунду табуу жана сандык интегралдоо көйгөйлөрүндө колдонулушу мүмкүн. Бул макалада, алгоритмдин жыйынтыктары конкреттүү мисалдарды чечүү менен көрсөтүлгөн.

Annotation: In this paper, we propose to find approximation points on the object contour, based on methods for solving optimization problems. The basic idea is that scaling the distance between points gives a one-to-one mapping between the contour and a point in the vector space. In addition to matching objects, this algorithm has a wide range of applications, such as numerical integration and the calculation of derivatives with a specified number of grid nodes. The paper shows the comparison of solutions with known methods and specific examples.

Ключевые слова: упрощение контура объекта; линейное приближение; численные методы; численное интегрирование.

Урунттуу сөздөр: объектинин контурдун жөнөкөйлөтүү; сызыктуу жакындаштыруу; сандык ыкмалары; сандык интегралдоо.

Keywords: simplification of the contour of the object; linear approximation; numerical methods; numerical integration.

1. Введение

Как правило, при численном интегрировании используется приближение по формуле трапеции [5, 7] с различным разбиением интервала интегрирования. В этом методе для разделения на интервалы используются эквидистантные узлы [3]. Однако для лучшего приближения разница между заданной кривой $f(x)$ и ломанной кривой $L(x)$, полученной при разбиении исходной кривой, должна быть сведена к минимуму:

$$\min_{x_0 < x_1 < \dots < x_k} \sum_{i=0}^{k-1} \int_{x_i}^{x_{i+1}} [f(x) - L(x)] dx \quad (1)$$

для точек x_i , где $i = 0, 1, \dots, k$, значения чисел x_i лежат в интервале интегрирования.

Узлы, заданные на равном расстоянии, не всегда дают наилучшее приближение. Чтобы минимизировать ошибку, должны быть найдены значения точек сетки x_i , которые наилучшим образом аппроксимируют кривую. В литературе предложено много методов с неравномерными узлами [1-2].

Если говорить об области применения, расчет интеграла имеет множество применений. Например, в работе Дугласа и Пеукера, этот подход использовался для упрощения линий в картографии [4]. Султанов и соавт. предложил другую задачу минимизации которая обладает некими другими полезными свойствами [6]. Чтобы найти

наилучшее приближение, авторы предложили найти точки сетки $x_i, i = 0, 1, \dots, k$ на кривой $f(x)$, которые удовлетворяют следующему условию максимума:

$$\max_{x_0 < x_1 < \dots < x_k} \sum_{i=0}^{k-1} \sqrt{(f(x_{i+1}) - f(x_i))^2 + (x_{i+1} - x_i)^2} \quad (2)$$

Обе задачи аппроксимируют кривую через кусочно-линейные ломанные. В первой задаче площадь между кривой и ломанными минимизируется, а на второй задаче длина кривой и ломанных максимально приближается.

2. Алгоритм

Алгоритм для нахождения численного решения задачи (1) и (2) на основе метода покоординатного спуска можно сформулировать следующим образом:

1. Выбрать случайное начальное приближение: $\{x_i\}, i = 0, 1, \dots, n$.

2. Задать $y_0 = x_0, y_n = x_n$ и для всех $i = 1, \dots, n - 1$ вычислить

$$y_i = \frac{(x_{i-1} + x_{i+1})}{2}$$

3. Пока не выполнено условие $\sum_{i=0}^n |x_i - y_i| < \epsilon$, выполнить следующие шаги:

3.1 Задать

$$x_i = y_i$$

для всех $i = 0, 1, \dots, n$.

3.2 Вычислить

$$y_i = \operatorname{argmax} F(x_{i-1}, t, x_{i+1})$$

для всех $i = 1, \dots, n - 1$.

4. вычислить

$$x_i = \frac{(x_i + y_i)}{2}$$

для всех $i = 0, 1, \dots, n$.

Здесь, $\epsilon > 0$ – достаточно малое число; для задачи (1), функция $F(x, t, y)$ является площадью треугольника, заданного координатами $(x, f(x)), (t, f(t))$ и $(y, f(y))$; для задачи (2), функция $F(x, t, y)$ является общей длиной, определяемой суммой длин сторон треугольника от $(x, f(x))$ до $(t, f(t))$ и от $(t, f(t))$ до $(y, f(y))$.

3. Экспериментальные результаты

3.1. Применение алгоритма для численного интегрирования

Покажем результаты алгоритма при численном интегрировании. Рассмотрим функцию

$$f(x) = 10 * \exp(-10x)$$

на отрезке $[0, 1]$.

В качестве начального приближения взята равномерная сетка с 8 узлами. Расчеты проводились с использованием сетки 1, где узлы распределены равномерно; сетка 2, где предложенный алгоритм применяется для нахождения решения задачи (1); и сетка 3, где предложенный алгоритм применяется для нахождения решения задачи (2).

В таблице 1 представлены значения узлов сетки, рассчитанные с использованием правила трапеции с равномерным разбиением интервала интегрирования и с использованием вновь предложенного алгоритма.

Таблица 1. Значения узлов x_i , определенных с равномерными узлами сетки и узлами сетки, вычисленными предлагаемым алгоритмом

		Узлы сетки (i)							
		1	2	3	4	5	6	7	8
Сетка 1	x_i	0.000	0.143	0.286	0.429	0.571	0.714	0.857	1.000
	$f(x_i)$	10.000	2.393	0.573	0.137	0.033	0.008	0.002	0.0005
Сетка 2	x_i	0.000	0.045	0.098	0.163	0.245	0.359	0.545	1.000
	$f(x_i)$	10.000	6.376	3.753	1.959	0.863	0.276	0.043	0.0005
Сетка 3	x_i	0.000	0.180	0.290	0.373	0.444	0.518	0.624	1.000
	$f(x_i)$	10.000	1.653	0.550	0.240	0.118	0.056	0.019	0.0005

Если сравнить результаты алгоритмов графически, то видно, что предлагаемый алгоритм приводит к снижению ошибок. На рисунке 1(а), показаны результаты, производная функции в узлах и значение ошибки когда узлы сетки распределены равномерно. Аналогично, графический результат, производной функции по узлам сетки и значения ошибок, полученных путем вычисления узлов сетки с использованием нашего алгоритма, представлен на рисунке 1(б).

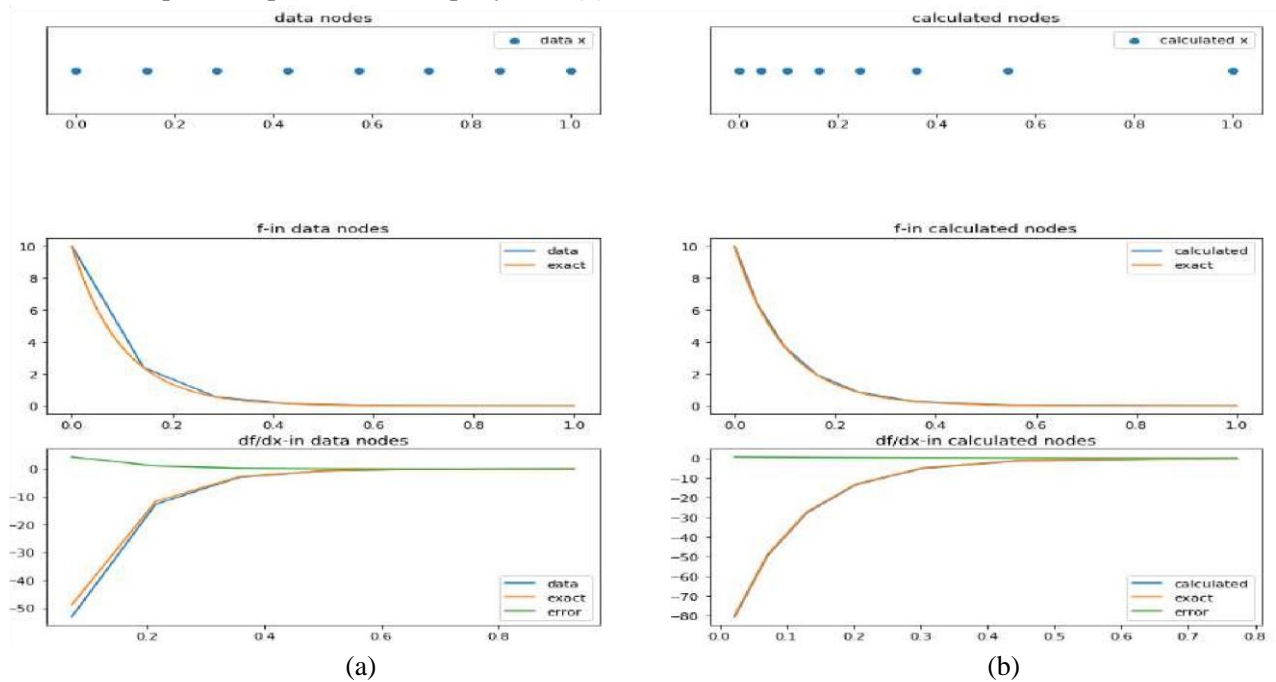


Рисунок 1. Результаты вычислений при равномерном распределении узлов сетки (а) и их вычисления с использованием нашего алгоритма (б) при решении задачи (1)

Как показывают результаты, когда используются узлы равномерной сетки, результат интегрирования равен 1,164, а при использовании нашего алгоритма он равен 1,040 (точное значение интеграла на отрезке $[0,1]$ равно 1.0). Таким образом, соотношение ошибок составляет 64% и 4% для формулы трапеции при равномерном разбиении и интервалом интегрирования полученным нашим алгоритмом задачи (1).

Что касается численного расчета производных по центральной разности, определяемой как

$$f'(x_i) = \frac{f(x_{i+1}) - f(x_{i-1}))}{x_{i+1} - x_{i-1}}$$

рассчитываем норму ошибок в пространстве C

$$\max |f'(x_i) - f'(\bar{x}_i)|$$

и в пространстве C_1

$$\frac{\sum_{i=1}^{n-1} |f'(x_i) - \overline{f'(x_i)}|}{n}$$

Для сетки 1, ошибки в норме C равно 4.27, а в норме C_1 равно 0.08. Для сетки 2, ошибки в норме C и C_1 равны 0.67 и 0.01, соответственно.

Далее мы рассчитали узлы, используя предложенный алгоритм для решения задачи (2). Значения узлов сетки в интервале $[0, 1]$ представлены в последних строках таблицы 1.

Аналогично сравнению при решении задачи (1), если сравнить результаты алгоритмов графически, то видно, что решение задачи (2) с использованием предложенного алгоритма приводит к снижению значений ошибок.

Результат численного интегрирования, полученный путем вычисления узлов сетки с использованием нашего алгоритма, производной функции по узлам и значений ошибок представлен на рисунке 2(б)

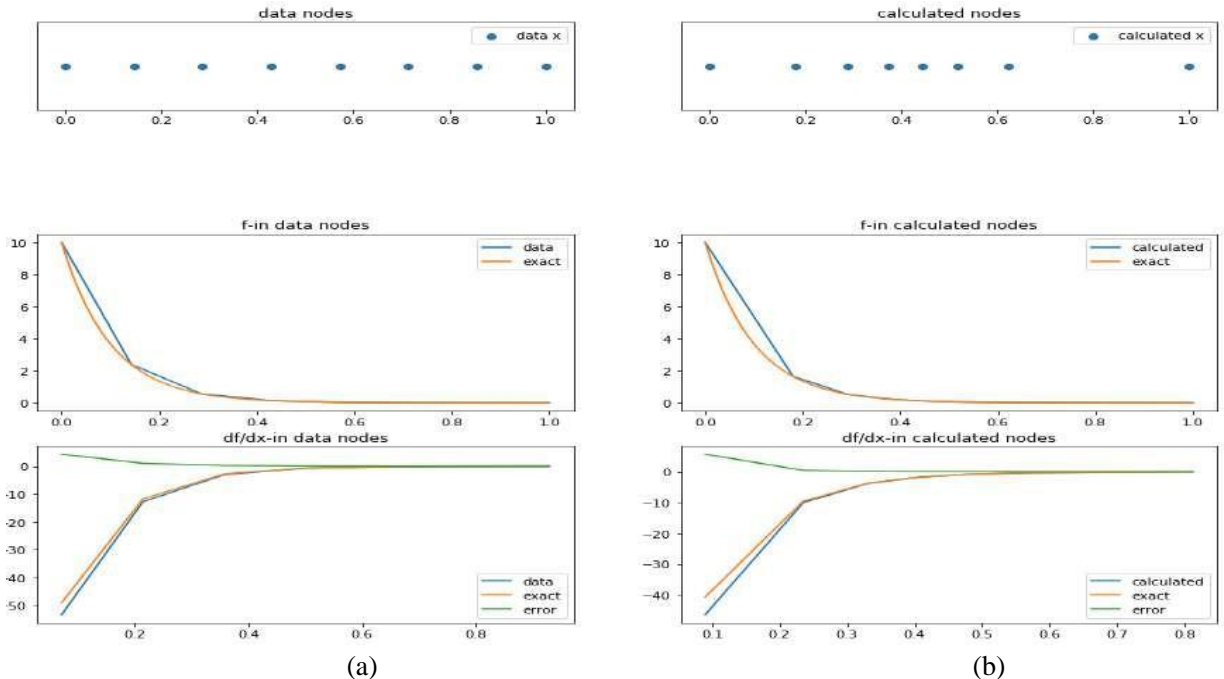


Рисунок 2. Результаты вычислений при равномерном распределении узлов сетки (а) и их вычислении с использованием нашего алгоритма (б) при решении задачи (2)

Результаты расчетов показывают меньшее значение ошибок числовых значений интеграла и производных на расчетных узлах. Ошибка интеграла на сетке 1 составила 64%, тогда как на сетке 3 ошибка составила 22%. Значение ошибки приближения производных в норме C равно 5.70 и норме C_1 равно 0.01.

Расчеты для решения задач (1) и (2) для численного интегрирования дало существенно разные результаты, хотя обе задачи аппроксимации кривой в заданном сегменте, но разными способами.

Если предположить, что заданная функция $f(x)$ является строго выпуклой на интервале $[a, b]$, то доказательство сходимости алгоритма единственному решению задачи (1) и (2), сводится к теореме сходимости и единственности монотонно возрастающей и ограниченной сверху последовательности. Отметим, что результат алгоритма найденные узлы x_i для нахождения узлов для задач (1) и (2) является инвариантным при преобразованиях, таких как равномерное сжатие, перенос и вращение. Это свойство, вместе со свойством инвариантности длины кривой, используется для идентификации форм объектов, представленных в виде кривых линий. Решения задач (1) и (2) проводились в разных начальных точках, а алгоритм (3) показал устойчивость и единственность решений. Данное свойство также верно не только для выпуклых функций, но и для широких класс кривых.

3.2. Применение алгоритма для сопоставления контуров объектов.

Применение данного алгоритма для распознавания изображений путем сопоставления их контуров был предложен в работе [6]. Для этого необходимо рассчитать координаты (\bar{x}, \bar{f}) геометрической середины узлов сетки 3 (Таблица 1, строка 3). Напомним, что данные узлы лежат на графике ломанной $f(x)$, которой аппроксимируется контур изображения. Найденная середина является также визуальным центром изображения. Затем рассчитываем длину ломанных, то есть расстояние между узлами $(x_i, f(x_i))$ и $(x_{i+1}, f(x_{i+1}))$, для всех $i = 0, \dots, 6$. Далее, вычисляем расстояние между узлами $(x_i, f(x_i))$ и центром

(\bar{x}, \bar{f}) для всех $i = 0, \dots, 7$. После, проводим масштабирование таким образом, чтобы сумма расстояний была равна 1. Результаты вычислений приведены в таблице 2.

Таблица 2. Масштабирование расстояния между узлами $(x_i, f(x_i))$ и $(x_{i+1}, f(x_{i+1}))$, и между узлами $(x_i, f(x_i))$ и центром (\bar{x}, \bar{f})

Расстояние	Узлы сетки (i)							
	1	2	3	4	5	6	7	8
между узлами	0.300	0.040	0.012	0.005	0.003	0.004	0.014	-
между узлами и центром	0.303	0.009	0.037	0.048	0.053	0.055	0.057	0.060

Расчитанные масштабированные расстояния не зависят от преобразований и являются инвариантами, таких как равномерное сжатие, вращение и перенос. Именно они могут быть взяты как координаты точек в многомерном векторном пространстве для заданной кривой. Численные эксперименты показали возможность применения решения задачи (2) для однозначной идентификации кривой на плоскости, поэтому они могут быть применены для распознавания объектов в изображении.

Заключение

В данной статье предлагается новая задача и алгоритм его решения. Найденные узлы сетки используется для численного интегрирования, численного дифференцирования и классификации кривой на плоскости. Предложенный алгоритм может быть использован для:

1. нахождения инвариантов кривой на плоскости, чтобы использовать решения задачи (2) для идентификации кривой в задачах распознавания изображений. В этом случае распознавание может быть выполнено путем сопоставления объектов с одинаковыми инвариантами.
2. получения более точного численного решения интеграла в заданном количестве узлов по правилу трапеции, используя разбиение, как решения в задаче (1).
3. получения более точной значений численных производных по центральной разности, используя узлы решения задачи (1)

Список цитируемых источников

1. Bryan, K. A scheme for numerical integration of the equations of motion on an irregular grid free of nonlinear instabilities // Journal of Computational Physics. – 1976. – № 94(1). – С. 38-40.
2. Clenshaw, C. W., Curtis, A. R. A method for numerical integration on an automatic computer // Numerische Mathematik. – 1960. – № 2(1). – С. 197-205.
3. Davis, P. J., Rabinowitz, P. Methods of numerical integration // Bull. Amer. Math. Soc. - 1976. - № 82. – С. 538-539.
4. Douglas, D.H. and Peucker, T.K. Algorithms for the reduction of the number of points required to represent a digitized line or its caricature // Cartographica: the international journal for geographic information and geovisualization. – 1973. - № 10(2). – С. 112-122.
5. Poisson, S.-D. Sur le calcul numérique des intégrales définies // Mémoires de l'Académie Royale des Sciences de l'Institut de France, – 1827. - № 4. – С. 571–602.
6. Sultanov, R., Atakan, A., Ismailova, R. An algorithm for line matching in an image by mapping into an n-dimensional // Turkish Journal of Electrical Engineering & Computer Sciences. – В печати.
7. Trefethen, L. N., Weideman, J. A. C. The exponentially convergent trapezoidal rule // SIAM Review. – 2014. - № 56(3). – С. 385-

Рецензенты: Омуралиев А. – доктор физико-математических наук, профессор кафедры прикладной информатики и математики КТУ «Манас»