МИНОБРНАУКИ РОССИИ

ФЕДЕРАЛЬНОЕ ГОСУДАРСТВЕННОЕ АВТОНОМНОЕ ОБРАЗОВАТЕЛЬНОЕ УЧРЕЖДЕНИЕ ВЫСШЕГО ОБРАЗОВАНИЯ
«НОВОСИБИРСКИЙ НАЦИОНАЛЬНЫЙ ИССЛЕДОВАТЕЛЬСКИЙ ГОСУДАРСТВЕННЫЙ УНИВЕРСИТЕТ» (НОВОСИБИРСКИЙ ГОСУДАРСТВЕННЫЙ УНИВЕРСИТЕТ, НГУ)

Факультет информационных технологий

Кафедра Систем информатики

Направление подготовки: 230100 ИНФОРМАТИКА И ВЫЧИСЛИТЕЛЬНАЯ ТЕХНИКА

Магистерская программа: Компьютерное моделирование

ВЫПУСКНАЯ КВАЛИФИКАЦИОННАЯ РАБОТА

МАГИСТЕРСКАЯ ДИССЕРТАЦИЯ

Реализация алгоритма выбора параметра сглаживания с адаптивной регулировкой скорости сходимости

Мокшин Павел Владимирович

Тема диссертации утверждена распоряжением по НГУ №532 от «14» декабря 2012г.

«К защите допущена»	**Научный руководитель**
Заведующий кафедрой,	д.ф-м.н., с.н.с.,
д.ф-м.н., профессор	ИВМиМГ СО РАН
Лаврентьев М.М./…………..	Роженко А.И./………...
(фамилия , И., О.) / (подпись, МП)	(фамилия , И., О.) / (подпись, МП)
«……»………………2014г.	«……»………………2014г.

Дата защиты: «19»..июня…..………2014г.

Автор Мокшин П.В…………/………...

(фамилия , И., О.) / (подпись)

Новосибирск, 2014 г.

МИНОБРНАУКИ РОССИИ

ФЕДЕРАЛЬНОЕ ГОСУДАРСТВЕННОЕ АВТОНОМНОЕ ОБРАЗОВАТЕЛЬНОЕ УЧРЕЖДЕНИЕ ВЫСШЕГО ОБРАЗОВАНИЯ
«НОВОСИБИРСКИЙ НАЦИОНАЛЬНЫЙ ИССЛЕДОВАТЕЛЬСКИЙ ГОСУДАРСТВЕННЫЙ УНИВЕРСИТЕТ» (НОВОСИБИРСКИЙ ГОСУДАРСТВЕННЫЙ УНИВЕРСИТЕТ, НГУ)

Факультет информационных технологий

Кафедра Систем информатики

Направление подготовки: 230100 ИНФОРМАТИКА И ВЫЧИСЛИТЕЛЬНАЯ ТЕХНИКА

Магистерская программа: Компьютерное моделирование

УТВЕРЖДАЮ

Зав. кафедрой Лаврентьев М.М.
(фамилия, И., О.)
………………….
(подпись, МП)
«……»…………………………201..г.

ЗАДАНИЕ

НА ВЫПУСКНУЮ КВАЛИФИКАЦИОННУЮ РАБОТУ

МАГИСТЕРСКУЮ ДИССЕРТАЦИЮ

Студенту Мокшину Павлу Владимировичу

Тема: Реализация алгоритма выбора параметра сглаживания с адаптивной регулировкой скорости сходимости

Исходные данные (или цель работы): доработать и реализовать два алгоритма выбора параметра сглаживания – алгоритм, использующий только самое точное приближение невязки, и алгоритм, использующий несколько приближений функции невязки.

Структурные части работы: изучение теории сплайн-аппроксимации; доработка и реализация существующей версии алгоритма выбора параметра сглаживания; доработка и реализация алгоритма выбора параметра сглаживания, использующего несколько приближений функции невязки; встраивание алгоритмов в состав библиотеки SDM.

Научный руководитель
д.ф-м.н., с.н.с., ИВМиМГ СО РАН

Роженко А.И./…………..
(фамилия , И., О.) / (подпись)
«…»…………………201..г.

Задание принял к исполнению
Мокшин П.В./…………..
(ФИО студента) / (подпись)
«…»…………………201..г.

Оглавление

ВВЕДЕНИЕ

Во многих сферах человеческой деятельности (например, в моделировании поверхностей, физических процессов, САПР, ГИС) возникает задача построения приближения неизвестной функции по некоторым известным значениям в точках. Для приближения функций придумано довольно много различных методов (например, аппроксимация при помощи полиномов Лагранжа), одним из которых является сплайн-интерполяция.

Во многих случаях (когда имеют место неточности входных данных, либо точки находятся друг к другу слишком близко), вместо интерполяционного сплайна (который проходит через все заданные точки), используют сглаживающие сплайны, которые могут несколько отклоняться от заданных точек.

В вариационной теории сплайнов сглаживающий сплайн получается заменой условий интерполяции на добавление штрафа к минимизируемому функционалу интерполяционного сплайна. Величиной штрафа можно управлять при помощи параметра α, который называется параметром сглаживания.

При построении сглаживающего сплайна необходимо выбирать некоторое разумное значение параметра сглаживания α, чтобы отклонение сплайна в узлах сетки от условий интерполяции, было в пределах ошибок в исходных данных. Один из способов выбора параметра сглаживания основан на принципе невязки – параметр сглаживания α выбирается так, чтобы среднеквадратичное отклонение сглаживающего сплайна от значений в узлах сетки было равно заданному значению ε. Для того, чтобы по величине отклонения ε найти значение параметра сглаживания, необходимо решить нелинейное уравнение.

Для решения этого уравнения обычно используется итерационный метод Ньютона, который имеет квадратичную скорость сходимости и требует выполнения нескольких итераций для нахождения решения.

Проблема в том, что каждая итерация метода Ньютона эквивалентна по быстродействию построению интерполяционного сплайна. На каждой итерации производится сборка и декомпозиция матрицы (требует порядка $O(n^3)$ операций при построении сплайна многих переменных методом воспроизводящих ядер по значениям в n узлах сетки), а затем решение системы уравнений с двумя правыми частями, что выполняется существенно быстрее ($O(n^2)$ операций).

Ранее был разработан алгоритм с более высокой скоростью сходимости, основанный на разложении в ряд функции невязки [1, 2]. Суть метода сводится к решению системы уравнений со многими правыми частями на каждой итерации. За счет этого можно сократить количество итераций, что, несмотря на некоторое «утяжеление» каждой итерации, даст выигрыш во времени расчета.

Одной из целей дипломной работы является эффективная реализация данного алгоритма и встраивание его в библиотеку SDM, предназначенную для работы со сплайнами.

Библиотека SDM была разработана А.И. Роженко и включает в себя большое количество классов и функций для работы со сплайнами. В актуальной версии SDM на время начала работы задача выбора параметра сглаживания решалась только методом Ньютона. В результате встраивания нового алгоритма в библиотеку SDM, у пользователя должна появиться возможность выбрать любой из этих алгоритмов.

Указанный алгоритм использует лишь самое точное приближение функции невязки, но используя информацию о поведении менее точных приближений функции невязки, можно построить алгоритм вычисления параметра сглаживания, сходящийся с еще более высокой скоростью, что позволит еще сократить время решения задачи сглаживания за счет уменьшения числа итераций [3].

Вторая цель данной дипломной работы – разработать и эффективно реализовать алгоритм, использующий информацию о нескольких приближениях функции невязки, а также встроить разработанный алгоритм в библиотеку SDM.

Код библиотеки SDM написан на языке C#, поэтому разработка алгоритма также велась на этом же языке. В качестве среды разработки использовалась Visual Studio 2012.

В ходе выполнения работы были выделены и решены следующие задачи:

- Изучение теории сплайн-аппроксимации
- Доработка и реализация существующей версии алгоритма выбора параметра сглаживания
- Доработка и реализация алгоритма выбора параметра сглаживания, использующего несколько приближений функции невязки

- Встраивание алгоритмов в состав библиотеки SDM

Глава 1. Выбор параметра сглаживания

1.1 Кубический сплайн

Первоначально, под термином сплайн понималась кусочно-кубическая функция одной переменной, которая достаточно гладко «подклеена» в узлах стыковки полиномов.

Пусть на отрезке $[a, b]$ вещественной оси задана сетка $a = t_0 < t_1 < \cdots < t_N = b$, и в ее узлах задана сеточная функция $f_i, i = 0, \dots, N$.

Кубическим сплайном называют функцию $\sigma(t)$, удовлетворяющую следующим условиям:

- на каждом интервале $(t_{i-1}, t_i), i = 1, \dots, N$, функция σ является кубическим многочленом
- функция σ дважды непрерывно дифференцируема на $[a, b]$, т.е. $\sigma \in C^2[a, b]$

Сплайн σ называется интерполяционным, если он совпадает с сеточной функцией в узлах сетки.

Для однозначного определения кубического сплайна требуется добавить два краевых условия. Если потребовать, что $\sigma''(a) = \sigma''(b) = 0$, то кубический сплайн будет минимизировать функционал $\int_a^b \sigma''(t)^2 dt$ на множестве всех функций, принадлежащих пространству Соболева $W_2^2[a, b]$ и удовлетворяющих условиям интерполяции в узлах сетки.

1.2 Кубический сглаживающий сплайн

Для простоты сначала рассмотрим задачу сглаживания при помощи кубического сплайна.

Задачу можно сформулировать следующим образом:

У нас имеется некоторая неизвестная функция, которую мы хотим аппроксимировать, и мы знаем значений этой функции в $N + 1$ узлах сетки: $z = (z_0, \dots, z_N)$.

При этом мы не хотим интерполировать функцию, то есть найти функцию, точно проходящую через заданные точки, а допускаем некоторое отклонение сплайна от заданных значений на сетке.

Кубическим сглаживающим сплайном σ_α называется решение следующей задачи минимизации:

$$\sigma_\alpha = \underset{x \in W_2^2[a,b]}{\operatorname{argmin}} \left\{ \alpha \|x''\|_{L_2[a,b]}^2 + \sum_{i=0}^{N} |x(t_i) - z_i|^2 \right\},$$

где $\alpha > 0$ – параметр сглаживания. При помощи этого параметра можно регулировать поведение сглаживающего сплайна – насколько он будет отклоняться от значений функции в сетке.

При $\alpha \to 0$ сплайн σ_α стремится к интерполяционному сплайну σ.

Данное отклонение называется невязкой сглаживающего сплайна – это корень из второго слагаемого в формуле. Обозначим невязку как $\varphi(\alpha)$.

Тогда уравнение невязки формулируется следующим образом:

$$\varphi(\alpha) := \sqrt{\sum_{i=0}^{N} |\sigma_\alpha(t_i) - z_i|^2} = \varepsilon$$

Где $\varepsilon > 0$ – требуемый уровень невязки.

1.3 Общая постановка задачи сглаживания

В общем виде задача формулируется схожим образом.

Сглаживающим сплайном называется решение следующей задачи минимизации:

$$\sigma_\alpha = \underset{x \in \mathbb{X}}{\operatorname{argmin}} \ \{\alpha \|Tx\|_{\mathbb{Y}}^2 + \|Ax - z\|_{\mathbb{Z}}^2\} \quad (1)$$

где $\mathbb{X}$, $\mathbb{Y}$, $\mathbb{Z}$ - вещественные гильбертовы пространства, $T: \mathbb{X} \to \mathbb{Y}$ и $A: \mathbb{X} \to \mathbb{Z}$ - ограниченные линейные операторы, $z \in \mathbb{Z}$ - заданный вектор и $\alpha > 0$ – параметр сглаживания.

Оператор T называется оператором энергии сглаживающего сплайна, а оператор A – оператором, задающим условия аппроксимации.

В практических случаях задача сглаживания однозначно разрешима, если образ $R(T)$ оператора T замкнут, его ядро $N(T)$ конечномерно и пересекается по нулю с ядром оператора A (см., например, [4]).

Можно заметить, что для задачи сглаживания кубическим сплайном, оператор T – это оператор двукратного дифференцирования D^2, оператор $A = (\delta_{t_0}, \dots, \delta_{t_N})^T$, где $\delta_t(f) = f(t)$ пространство $\mathbb{X}$ – это пространство Соболева $W_2^2[a,b]$, $\mathbb{Y}$ – это пространство $L_2[a,b]$, $\mathbb{Z} = R^{N+1}$.

Принцип невязки формулируется следующим образом: найти $\alpha > 0$, при котором

$$\varphi(\alpha) := \|A\sigma_\alpha - z\| = \varepsilon \quad (2)$$

Заметим, что при $\alpha \to 0$, сглаживающий сплайн σ_α стремится к интерполяционному сплайну σ_0, а при $\alpha \to \infty$, σ_α стремится к σ_∞, которое является элементом ядра оператора T, минимально квадратично отклоняющимся от вектора z.

Функция $\varphi(\alpha)$ является монотонной неубывающей, $\varepsilon_{min} := \varphi(0) = 0$, если условия аппроксимации непротиворечивы, то есть интерполяционный сплайн существует. Следовательно, уравнение невязки имеет решение при условии, что $\varepsilon \le \varepsilon_{\max} := \varphi(\infty)$. Далее предполагаем, что данное условие выполняется.

1.4 Оператор невязки

Алгоритмы, описанные в данной работе, основаны на разложении невязки в ряд [5], поэтому приведем соответствующие теоремы и формулировки.

Введем оператор невязки R_α, такой что $R_\alpha z = z - A\sigma_\alpha$, где σ_α – решение задачи сглаживания (1) для заданного вектора $z \in \mathbb{Z}$.

Оператор невязки R_α является самосопряженным, положительным и его норма не превосходит единицу, а в случае непротиворечивых условий интерполяции $\|R_\alpha z\| < \|z\|$.

1.5 Асимптотическое разложение невязки при $\alpha < \alpha_0$

Теорема 1 [5]**.** Пусть $\alpha_0 > 0$. Тогда невязка $R_\alpha z = z - A\sigma_\alpha$ представима в виде суммы ряда

$$z - A\sigma_\alpha = z - \sum_{k=0}^{\infty} \left(\frac{\alpha_0 - \alpha}{\alpha_0}\right)^k R_{\alpha_0}^k A\sigma_{\alpha_0}$$

абсолютно сходящегося на отрезке $[0, 2\alpha_0]$.

Введем следующие обозначения:

$$\gamma := \frac{\alpha_0 - \alpha}{\alpha_0}, \qquad R(\gamma) := R_\alpha$$

Тогда, взяв две частичные суммы от разложения оператора R_α, можно получить:

$$R_n^-(\gamma) := (1-\gamma)\sum_{k=0}^{n-1} \gamma^k R_{\alpha_0}^{k+1}$$

$$R_n^+(\gamma) := R_{\alpha_0} - \sum_{k=1}^{n-1} \gamma^k R_{\alpha_0}^k (I - R_{\alpha_0})$$

Теорема 2 [2]. $\forall n \geq 0, z \in \mathbb{Z}$ и $\forall \gamma \in [0,1]$ верно:

1) $\|R(\gamma)z\|$ и $\|R_n^\pm(\gamma)z\|$ являются невозрастающими при $\gamma \in [0,1]$
2) $\|R_n^\pm(\gamma)z\|$ монотонны по n:
$$\|R_n^-(\gamma)z\| \leq \|R_{n+1}^-(\gamma)z\| \leq \|R(\gamma)z\| \leq \|R_{n+1}^+(\gamma)z\| \leq \|R_n^+(\gamma)z\|$$
3) $|\|R(\gamma)z\| - \|R_n^\pm(\gamma)z\|| \leq \gamma^n \|z\|$

Таким образом, $R_n^+(\gamma)$ и $R_n^-(\gamma)$ дают двустороннее приближение для $R(\gamma)$ и имеют порядок точности n при малых γ.

1.6 Асимптотическое разложение невязки для случая $\alpha > \alpha_0$

Чтобы обойти ограничение $\alpha \leq 2\alpha_0$, производится замена $\beta \coloneqq 1/\alpha$ и вводятся два новых оператора:

$$\hat{\hat{R}}_\beta \coloneqq R_{1/\beta}, \quad Q_\beta \coloneqq I - R_{1/\beta}$$

Для оператора $\hat{R}_\beta$ имеет место аналогичное разложение, что и для случая $\alpha^* \leq 2\alpha_0$.

Теорема 3 [5]. Пусть $\beta_0 > 0$. Тогда невязка $R_{1/\beta} z = z - A\sigma_{1/\beta}$ представима в виде суммы ряда

$$z - A\sigma_{1/\beta} = \sum_{k=0}^{\infty} \left(\frac{\beta_0 - \beta}{\beta_0}\right)^k Q_{\beta_0}^k (z - A\sigma_{1/\beta_0})$$

абсолютно сходящегося на отрезке $[0, 2\beta_0]$.

Введем следующие обозначения:

$$\delta \coloneqq \frac{\beta_0 - \beta}{\beta_0}, \qquad \hat{R}(\delta) \coloneqq \hat{R}_\beta$$

Тогда, взяв две частичные суммы от разложения оператора, получаем:

$$\hat{R}_n^-(\delta) := \sum_{k=0}^{n-1} \delta^k\, Q_{\beta_0}^k (I - Q_{\beta_0})$$

$$\hat{R}_n^+(\delta) := I - (1-\delta) \sum_{k=0}^{n-1} \delta^k\, Q_{\beta_0}^{k+1}$$

Имеет место следующее утверждение:

Теорема 4 [2]**.** $\forall n \geq 0,\ \ z \in \mathbb{Z}$ и $\ \delta \in \left[0,\ 1\right]$ верно:

1) $\left\|\hat{R}(\delta)z\right\|$ и $\left\|\hat{R}_n^{\pm}(\delta)z\right\|$ являются неубывающими при $\delta \in [0,1]$
2) $\left\|\hat{R}_n^{\pm}(\delta)z\right\|$ монотонны по n:

$$\left\|\hat{R}_n^-(\delta)z\right\| \leq \left\|\hat{R}_{n+1}^-(\delta)z\right\| \leq \left\|\hat{R}(\delta)z\right\| \leq \left\|\hat{R}_{n+1}^+(\delta)z\right\| \leq \left\|\hat{R}_n^+(\delta)z\right\|$$

3) $\left|\left\|\hat{R}(\delta)z\right\| - \left\|\hat{R}_n^{\pm}(\delta)z\right\|\right| \leq \delta^n \|z\|$

Теоремы 1 и 3 позволяют представить невязки в виде суммы ряда без ограничений на параметр α: при $0 \leq \alpha \leq \alpha_0$ используем первое разложение, а при $\alpha_0 \leq \alpha \leq \alpha_\infty$- второе.

Замечание 1. Если $\varepsilon_{min} \neq \varepsilon_{max}$, то фразы «являются неубывающими» и «являются невозрастающими» в утверждениях теорем 2 и 4 следует заменить соответственно на «строго возрастают» и «строго убывают», т.е. уравнение $\|P(t)z\| = \varepsilon$ имеет не более одного решения на интервале (0, 1), где P – любой из операторов $R, R_n^{\pm}, \hat{R}, \hat{R}_n^{\pm}$.

Далее будем предполагать, что $\varepsilon_{min} \neq \varepsilon_{max}$, что гарантирует единственность решения уравнения $\varphi(\alpha) = \varepsilon$ при $\varepsilon \in (\varepsilon_{min}, \varepsilon_{max})$.

Глава 2. Алгоритм выбора параметра сглаживания

2.1. Выбор параметра сглаживания методом Ньютона

Для решения уравнения (2) используют метод Ньютона, выполнив замену $\beta := \alpha^{-1}$ и преобразовав уравнение к следующему виду:

$$\frac{1}{\varphi(1/\beta)} = \frac{1}{\varepsilon} \quad (3)$$

В данном виде скорость сходимости метода Ньютона максимальна [6].

В терминах оператора невязки $R_\alpha z := z - A\sigma_\alpha$, формула шага метода Ньютона имеет вид [4]:

$$\alpha_{k+1} = \alpha_k \frac{1 - \omega(\alpha_k)}{\frac{\varphi(\alpha_k)}{\varepsilon} - \omega(\alpha_k)} \quad (4)$$

$$\omega(\alpha) = \frac{(R_\alpha z,\ R_\alpha^2 z)}{\varphi^2(\alpha)}$$

Можно заметить, что на каждом шаге требуется решить две задачи сглаживания с векторами z и $R_\alpha z$.

Итерации метода Ньютона для решения уравнения (3) следует начинать с достаточно малых значений β, при которых $\varphi\left(\frac{1}{\beta}\right) \geq \varepsilon$. Однако начинать итерации с $\beta = 0$ не всегда возможно, так как не всегда возможно вычислить производную функции $\varphi(1/\beta)$ при $\beta = 0$.

Поэтому на практике будем использовать гибридную схему [4]: если $\varphi(\alpha_k) < \varepsilon$, то расчет выполняем по формуле (4), а в противном случае используем дробно-рациональное приближение функции $\varphi(\alpha)$, которое производится по формулам:

$$\varepsilon_{max} := \varphi(\infty)$$

$$\alpha_{k+1} = \alpha_k \frac{\left(1 - \frac{\varphi(\alpha_k)}{\varepsilon_{max}}\right) - \left(1 - \frac{\varphi(\alpha_k)}{\varepsilon}\right) d(\alpha_k)}{\frac{\varphi(\alpha_k)}{\varepsilon} - \varphi(\alpha_k)/\varepsilon_{max}}$$

$$d(\alpha) = \frac{\frac{\varphi(\alpha)}{\varepsilon_{max}} - \omega(\alpha)}{1 - \omega(\alpha)}$$

Коэффициенты рациональной дроби выбраны так, чтобы ее асимптотика в бесконечности совпадали с ε_{max}, значение и производная в точке α_k совпадали со значением и производной функции невязки в этой точке.

Как и для формул метода Ньютона, здесь также требуется решить две задачи сглаживания с векторами z и $R_\alpha z$. Дополнительно необходимо еще вычислить ε_{max}, которое достаточно легко находится методом наименьших квадратов.

2.2 Метод решения, основанный на разложении невязки в ряд

Используя асимптотическое разложение невязки, можно построить алгоритм выбора параметра сглаживания.

Данный алгоритм является доработанной версией алгоритма, разработанного Е.Д. Ивановой и моим научным руководителем А.И. Роженко [1].

В предыдущей версии алгоритма, разработанного Е.Д. Ивановой, была рассмотрена только одна ситуация – когда начальное приближение параметра сглаживания находилось справа от оптимального значения, т.е. $\alpha_0 > \alpha^*$.

В текущей версии реализован и случай, когда начальное значение параметра сглаживания находится слева от оптимального, т.е. $\alpha_0 < \alpha^*$.

2.3 Шаг алгоритма при $0 < \alpha < \alpha_k$

Из замечания 1 следует, что уравнение $\|R_n^-(\gamma)z\| = \varepsilon$ имеет единственное решение γ_* на интервале (0, 1) при $\varepsilon < \varphi(\alpha_0)$. Причем при $\alpha_1 := \alpha_0(1-\gamma_*)$ будет выполняться неравенство $\varepsilon < \varphi(\alpha_1)$. Таким образом, на следующем шаге снова можно будет применить нижнюю оценку функции невязки для $0 < \alpha < \alpha_1$. Т.е. последовательность приближений к оптимальному значению параметра сглаживания будет при таком выборе монотонно убывать.

Если же для выбора приближения использовать уравнение $\|R_n^+(\gamma)z\| = \varepsilon$, то имеют место следующие проблемы:

- Если $\|R_n^+(1)z\| \geq \varepsilon$, то уравнение $\|R_n^+(\gamma)z\| = \varepsilon$ не имеет решения на интервале (0, 1), хотя уравнение $\|R(\gamma)z\| = \varepsilon$ может иметь решение
- Точность данной оценки будет невысокой, если γ_* близко к 1

Ввиду вышеизложенного, будем использовать уравнение $\|R_n^-(\gamma)z\| = \varepsilon$ для выбора следующего приближения параметра сглаживания.

Опишем шаг алгоритма $\alpha_k \to \alpha_{k+1}$ при $\varphi(\alpha_k) > \varepsilon$:

1. Вычисляем $R_{\alpha_k}^i z,\ i = 1,2,\dots,n$
2. Решаем уравнение $\|R_n^-(\gamma)z\| = \varepsilon$ при $\alpha = \alpha_k$ на интервале (0, 1), например, методом секущих. Данное уравнение имеет единственное решение $\gamma_* \in (0,1)$ в силу строгого убывания $\|R_n^-(\gamma)z\|$ в силу замечания 1
3. Полагаем $\alpha_{k+1} = \alpha_k(1-\gamma_*)$

2.4 Шаг алгоритма при $\alpha_k < \alpha < \infty$

По аналогии со случаем $0 < \alpha < \alpha_k$, при выборе α_1 при помощи уравнения $\|\hat{R}_n^+(\delta)z\| = \varepsilon$, получим $\varphi(\alpha_1) > \varepsilon$, т.е. последовательность α_k будет монотонно сходиться к оптимальному значению параметра сглаживания. Однако скорость сходимости этого процесса оказывается линейной, если начальное приближение выбрано достаточно далеко от решения.

Причина в том, что точность верхней оценки катастрофически падает при отдалении δ от нуля. Несложно заметить, что $\|\hat{R}_n^+(1)z\| = \|z\| \gg \varepsilon_{max}$.

Проводились численные расчеты поведения двусторонних приближений $\hat{R}_n^\pm(\delta)z$ и $R_n^\pm(\gamma)z$. В таблице 1 приведены результаты, полученные в тестовом расчете для 100 узлов сетки, описанном в главе 3. Параметры расчетов: начальное приближение $\alpha_0 = 0.001$, количество невязок $n = 3$

Таблица 1. Сравнение оценок функции невязки

$\boldsymbol{\delta}$	$\|\hat{R}_n^-(\delta)z\|$	$\|\hat{R}(\delta)z\|$	$\|\hat{R}_n^+(\delta)z\|$
0.0	1.921E-4	1.921E-4	1.921E-4
0.1	2.080E-4	2.081E-4	2.182E-3
0.2	2.264E-4	2.274E-4	1.735E-3
0.3	2.474E-4	2.514E-4	5.857E-2
0.4	2.710E-4	2.820E-4	1.388E-1
0.5	2.973E-4	3.227E-4	2.711E-1
0.6	3.263E-4	3.800E-4	4.685E-1
0.7	3.579E-4	4.681E-4	7.440E-1
0.8	3.922E-4	6.265E-4	1.111E0
0.9	4.293E-4	1.029E-3	1.581E0
0.99	4.649E-4	5.617E-3	2.104E0
1.0	4.690E-4	1.259E-1	2.169E0

Легко заметить, что нижняя оценка существенно точнее верхней. Поэтому в дальнейшем мы используем уравнение $\left\|\hat{R}_n^-(\delta)z\right\| = \varepsilon$ для получения приближений.

Отметим, что при $\left\|\hat{R}_n^-(1)z\right\| < \varepsilon$ данное уравнение не имеет решения, хотя уравнение невязки разрешимо вплоть до $\left\|\hat{R}(1)z\right\| \equiv \varepsilon_{max}$.

Результаты численных экспериментов показали, что на первом шаге дробно-рациональное приближение дает лучший результат по сравнению с разложением в ряд, даже при достаточно больших значениях $n > 10$. А при следующих итерациях, приближение при помощи решения уравнения $\left\|\hat{R}_n^-(\delta)z\right\| = \varepsilon$ уже более эффективно. Поэтому на первом шаге будем использовать дробно-рациональное приближение.

С учетом вышеизложенного, шаг алгоритма $\alpha_k \to \alpha_{k+1}$ для случая $\varphi(\alpha_k) < \varepsilon$ будет выглядеть следующим образом:

0. При k = 0, выполняем шаг дробно-рационального приближения и находим α_1.
1. Иначе, вычисляем $Q_{\beta_k}^i z,\ i = 1, \dots, n$
2. Решаем уравнение $\left\|\hat{R}_n^-(\delta)z\right\| = \varepsilon$ при $\beta = \beta_k$ на интервале (0, 1), например, методом секущих. Данное уравнение имеет не более одного решения $\delta_* \in (0, 1)$ в силу строгого возрастания $\hat{R}_n^-(\delta)$ в силу замечания 1. Если $\left\|\hat{R}_n^-(0.9)z\right\| < \varepsilon$, то строим линейную интерполяцию на интервале $[0.9, 1]$ со значениями $f(0.9) = \left\|\hat{R}_n^-(0.9)z\right\|$, $f(1) = \varepsilon_{max}$ и решаем уравнение $f(\delta) = \varepsilon$.
3. Полагаем $\alpha_{k+1} = \alpha_k/(1 - \delta_*)$

Особенности данного метода:

- Порядок скорости сходимости алгоритма равен числу используемых невязок - n.
- В данном алгоритме используется информация лишь о самом точном приближении невязки – n-м члене, а хотелось бы использовать информацию обо всех приближениях невязки.

2.5 Модификация метода

Моя работа также заключалась в доработке модификации этого алгоритма, которая использует несколько приближений функции невязки [3].

Идея алгоритма состоит в том, чтобы выявить характер зависимости функций $\left\|\hat{R}_n^-(\delta)z\right\|$ и $\|R_n^-(\gamma)z\|$ от n, а затем, используя данную асимптотику, получить более точное приближение.

При использовании метода, описанного выше, мы уже вычисляем все промежуточные приближения невязки, поэтому данная модификация не усложняет каждую итерацию.

Важное отличие от алгоритма Ивановой состоит в том, что мы используем два последних приближений, а не три. Кроме того, в данной работе алгоритм распространен на случай $\alpha_k < \alpha < \infty$.

В описанном в предыдущей части главы алгоритме используются нижние приближения:

$$R_n^-(\gamma)z{:=}(1-\gamma)\sum_{k=0}^{n-1}\gamma^k\, R_{\alpha_0}^k\left(R_{\alpha_0}z\right)$$

$$\hat{R}_n^-(\delta)z{:=}\sum_{k=0}^{n-1}\delta^k\, Q_{\beta_0}^k(R_{\alpha_0}z)$$

Можно заметить, что нижние приближения имеют схожий вид, поэтому приведем их к единому виду:

$$S_n = \left\|\sum_{k=0}^{n-1} t^k P^k y\right\| \quad (5)$$

где $P = R_{\alpha_0}$ для $R_n^-(\gamma)z$, $P = Q_{\beta_0}$ для $\hat{R}_n^-(\delta)z$; t – фиксированное число из интервала (0, 1).

Отметим, что оператор R_{α_0} самосопряжен и его спектр принадлежит отрезку [0, 1]. Поэтому получаем, что оператор tP – самосопряженный оператор, спектр которого принадлежит отрезку $[0, t], t < 1$.

Обозначим как S предельное значение S_n:

$$S = \left\|\sum_{k=0}^{\infty} t^k P^k\, y\right\|$$

Предположим, что y – собственный вектор оператора tP, а λ – соответствующее ему собственное значение, тогда:

$$S_n = \left\|\sum_{k=0}^{n-1} t^k P^k y\right\| = \left\|\sum_{k=0}^{n-1} (t\lambda)^k y\right\| = \frac{1-(t\lambda)^n}{1-t\lambda}\,\|y\| = S - cq^n, \text{где } q := t\lambda$$

В общем виде будем считать, что $S_n \approx S - cq^n$, где $q = t\frac{\|P^{n-1}y\|}{\|P^{n-2}y\|}$ – приближение для максимального собственного значения оператора tP по формуле степенного метода.

Тогда S можно приблизить следующим образом через два последних приближения:

$$\begin{cases} S_n \approx S - cq^n \\ S_{n-1} \approx S - cq^{n-1} \end{cases} \Rightarrow \quad S \approx \frac{S_n - qS_{n-1}}{1-q} \quad (6)$$

Таким образом, расчет уточненного приближения в точке выполняется без существенных дополнительных операций по формуле (6). Схема алгоритма остается прежней за исключением того, что мы на каждом шаге вычисляем уточнение по этой формуле.

Глава 3. Численные расчеты

3.1 Численные расчеты

Зададим на области $[0,1] \times [0,1]$ сетку из N узлов, выбранных при помощи счетчика псевдослучайных чисел с некоторым фиксированным стартовым значением, чтобы обеспечить повторяемость результатов.

Будем строить на данной сетке приближение функции

$$F(x,y) = \sin x + \exp(-(x-0.5)^2 - (y-0.5)^2) + \cos y$$

при помощи псевдокубического сплайна Дюшона в $\mathbb{R}^2$:

$$\sigma(x,y) = \sum_{i=1}^{N} \lambda_i((x-x_i)^2 + (y-y_i)^2)^{\frac{3}{2}} + a + bx + cy$$

Параметр α выбираем из решения взвешенного уравнения невязки, в котором функция невязка поделена на $\sqrt{N}$:

$$\tilde{\varphi}(\alpha) := \left(\frac{1}{N}\sum_{i=1}^{N}\left(\sigma_\alpha\left(t^{(i)}\right) - z_i\right)^2\right)^{1/2} = \varepsilon$$

В тестовых расчетах оценивалась относительная погрешность решения взвешенного уравнения невязки на шаге: $e_k = (\tilde{\varphi}(\alpha_k) - \varepsilon)/\varepsilon$.

Знак погрешности определяет метод расчета, используемый на следующем шаге: если $e_k > 0$, то используется ветвь 1 алгоритма, если же $e_k < 0$, то применяется ветвь 2.

Приведем результаты расчетов для различных значений параметров α_0, ε, N.

В заголовках столбцов указаны условные обозначения алгоритмов, использованных при расчете: метод Ньютона, n – алгоритм без уточнения, использующий n невязок, n+ – модифицированный алгоритм с уточнением, использующий n невязок.

Итерации прекращались при достижении оценки $|e_k| \leq 10^{-6}$

В таблице 2 отображены результаты расчетов для $\alpha_0 = 1, \varepsilon = 0.001$.

Таблица 2. Результаты расчетов при $\alpha_0 = 1, \varepsilon = 0.001$

Номер итерации	**Метод Ньютона**	**2**	**2+**	**5+**	**10**
$N = 100$					
1	7.15E-01	8.21E-01	7.03E-01	3.90E-01	4.99E-01
2	4.53E-02	6.68E-02	2.90E-02	-3.36E-05	8.31E-06
3	3.93E-04	1.14E-03	6.68E-05	-7.43E-15	8.21E-15
4	3.16E-08	3.84E-07	3.53E-10		
$N = 900$					
1	3.89E-01	4.06E-01	3.26E-01	6.28E-02	8.08E-02
2	1.92E-02	2.64E-02	8.26E-03	-1.88E-08	4.40E-12
3	7.98E-05	2.08E-04	5.88E-06		
4	1.42E-09	1.36E-08	2.98E-12		

Для значений параметров $\alpha_0 = 0.001, \varepsilon = 0.001$ на первой итерации применялась формула шага дробно-рационального приближения, поэтому в таблице 3 итерации показаны, начиная с номера 2.

Таблица 3. Результаты расчетов при $\alpha_0 = 0.001, \varepsilon = 0.001$

Номер итерации	**Метод Ньютона**	**2**	**2+**	**5+**	**10**
$N = 100$					
2	-1.19E-02	1.40E-01	9.32E-03	-3.67E-05	3.52E-06
3	-2.59E-05	4.45E-03	6.88E-06	-6.88E-15	-1.33E-15
4	-1.21E-10	5.74E-06	3.75E-12		
5		9.62E-12			
$N = 900$					
2	-6.70E-02	1.58E00	1.54E-01	6.33E-04	5.73E-03
3	-8.18E-04	1.44E-01	2.02E-03	-6.66E-16	2.22E-16
4	-1.18E-07	4.93E-03	3.51E-07		
5		7.58E-06			
6		1.81E-11			

Видно, что алгоритм 5+ примерно эквивалентен алгоритму 10, т.е., используя 5 невязок с уточнением, получаем алгоритм, сходящийся с той же скоростью, что и алгоритм с 10 невязками без уточнения.

Также видно, что среди методов, использующих две невязки, метод 2+ выигрывает у метода Ньютона, а метод 2 проигрывает.

Видно, что использование уточнения дает эффект ускорения сходимости метода. Отметим, что в практических расчетах не требуется решать уравнение невязки с высокой точностью. Обычно результат уже приемлем, если относительная погрешность не превышает 10 %. При таком критерии останова методы 5+ и 10 сходятся в данных тестах за две итерации, т.е. решение одной задачи сглаживания с выбором параметра будет примерно втрое дороже, чем решение задачи интерполяции.

3.2 Тестирование алгоритма на случайных значениях в узлах сетки

Кроме того, проводилось тестирование вышеописанных алгоритмов на случайных значениях из интервала (0, 1) на той же сетке, что в предыдущем пункте.

Цель данного эксперимента – не сгладить тестируемую функцию, а проверить работу алгоритма на плохих условиях – когда в разложении вектора z существенными являются сразу большое количество компонент, соответствующих разным собственным значениям. А для достаточно гладких функций существенных компонент всего несколько.

Набор алгоритмов и критерий завершения работы выбраны теми же, что и в предыдущем пункте.

В таблице 4 приведены результаты расчетов при значениях параметров $\alpha_0 = 1, \varepsilon = 0.001$.

Таблица 4. Результаты расчетов при $\alpha_0 = 1, \varepsilon = 0.001$

Номер итерации	Метод Ньютона	2	2+	5+	10
$N = 100$					
1	6.30E01	1.71E02	5.97E01	3.65E01	1.20E02
2	2.34E-01	7.92E00	1.86E-01	1.79E-03	3.71E-01
3	1.64E-05	2.27E-02	3.94E-06	1.92E-12	1.62E-12
4	-9.14E-13	2.48E-07	-1.05E-12		

Для значений параметров $\alpha_0 = 10^{-7},\ \varepsilon = 0.01$ на первой итерации применялась формула шага дробно-рационального приближения, поэтому в таблице 5 итерации показаны, начиная с номера 2.

Таблица 5. Результаты расчетов при $\alpha_0 = 10^{-7}, \varepsilon = 0.01$

Номер итерации	Метод Ньютона	2	2+	5+	10
$N = 100$					
2	-7.97E-04	1.15E-02	7.80E-05	-1.02E-07	8.90E-11
3	-5.48E-08	4.31E-06	4.39E-11		
4		1.74E-12			

Как видно, для данного теста имеют место те же самые выводы, что и для предыдущего эксперимента: алгоритм 5+ работает примерно так же, как и 10; скорость работы методы Ньютона выше, чем скорость алгоритма 2 и ниже, чем алгоритма 2+.

Глава 4. Программная реализация

Все описанные в данной работе алгоритмы были реализованы в рамках библиотеки программ SDM.NET (Splines for data mining) на языке C#.

Ранее в библиотеке SDM.NET был реализован алгоритм выбора параметра сглаживания только при помощи метода Ньютона.

4.1 Иерархия классов

Для обозначения абстракции алгоритма выбора параметра сглаживания, был выделен интерфейс ISplineFitting.

Данный интерфейс содержит единственный метод SelectSmoothingParameter, который возвращает найденное значение параметра сглаживания и принимает следующие аргументы:

- Начальное приближение $\alpha_0 > 0$ параметра сглаживания
- Требуемый уровень невязки $\varepsilon > 0$
- Значения в узлах сетки, вектор z
- Экземпляр класса, реализующего интерфейс библиотеки SDM.NET ISplineFitter. Данный интерфейс представляет собой абстракцию решателя, который умеет подготавливать и решать систему линейных уравнений для задачи сплайн-аппроксимации и вычислять вектор невязки.

Интерфейс ISplineFitting реализуется абстрактным классом SplineFitting, от которого предполагается наследование всех конкретных классов-алгоритмов выбора параметра сглаживания. В данном базовом классе определено несколько protected методов и свойств, которые могут быть полезны производным классам:

- Свойство AccuracyLevel позволяет задать требуемый уровень точности для прекращения работы алгоритма
- Метод NeedStopIterations, который проверяет условие остановки итераций
- Метод DoRationalIteration, который выполняет итерацию при помощи дробно-рационального приближения. Данная функция используется как в методе Ньютона, так и алгоритмах, основанных на разложении невязки в ряд.
- Метод EnsureEpsilonInRange, который проверяет, что требуемый уровень невязки ε не превышает ε_{max}.

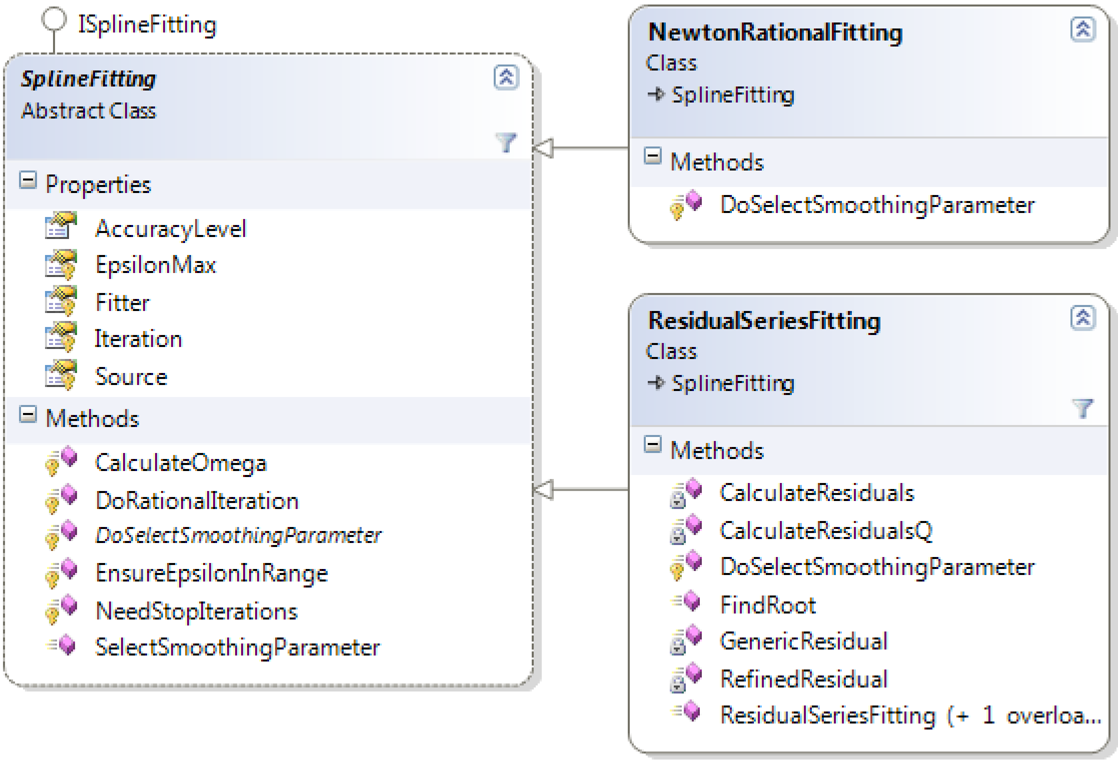

Рисунок 1. Диаграмма классов

Для метода Ньютона и алгоритмов, основанных на разложении невязки в ряд, были созданы классы NewtonRationalFitting и ResidualSeriesFitting соответственно, наследующиеся от абстрактного класса SplineFitting.

Существующая в SDM.NET реализация метода Ньютона была перенесена в новый класс, но теперь большая часть логики выполняется при помощи protected методов класса SplineFitting.

Остановимся на реализации класса ResidualSeriesFitting подробнее.

4.2 Реализация алгоритмов, основанных на разложении невязки в ряд

Класс ResidualSeriesFitting содержит в себе несколько вспомогательных функций:

- CalculateResiduals и CalculateResidualsQ – это функции, которые вычисляют $R^i_{\alpha_k}z,\ i = 1,2,\dots,n$ и $Q^i_{\beta_k}z,\ i = 1,\dots,n$ соответственно
- FindRoot – это реализация метода секущих для строго монотонных функций
- GenericResidual – функция, вычисляющая приближение (5) в заданной точке t. Используется для алгоритмов без уточнения

- RefinedResidual – функция, вычисляющая приближение (5) в заданной точке t, а затем применяющая уточнение по формуле (6) с использованием двух приближений. Используется для алгоритмов с уточнением

Сам алгоритм выбора параметра сглаживания реализован так, как описано в главе 2 с использованием перечисленных функций.

Заключение

Были выполнены все поставленные задачи:

- Доработан существующий алгоритм выбора параметра сглаживания, основанный на разложении невязки в ряд и использующий только самое точное приближение невязки: алгоритм распространен на случай когда $\alpha_0 < \alpha^*$
- Доработана модификация алгоритма, использующая несколько вычисленных приближений невязки
- Проведены численные расчеты, показавшие, что модифицированный алгоритм с уточнением сходится за меньшее количество итераций при достаточном количестве невязок
- Разработанные алгоритмы включены в библиотеку SDM

СПИСОК ЛИТЕРАТУРЫ

1. Иванова Е.Д. Выбор параметра сглаживания методом замораживания коэффициента: выпускная квалификационная работа бакалавра / Новосибирский государственный университет. Механико-математический факультет. Кафедра вычислительной математики. – Новосибирск, 2009.

2. Rozhenko A.I. A new method for finding an optimal smoothing parameter of the abstract smoothing spline // J. Approx. Theory. – 2010. – Vol. 162. – P. 1117–1127. – doi:10.1016/j.jat.2009.08.002.

3. Иванова Е.Д. Ускорение сходимости алгоритма выбора параметра сглаживания методом асимптотических оценок: выпускная квалификационная работа магистра / Новосибирский государственный университет. Механико-математический факультет. Кафедра вычислительной математики. – Новосибирск, 2011.

4. Роженко А.И. Вариационные сплайны: Учеб. Пособие: В 3 ч. / Новосиб. гос. ун-т. Новосибирск, 2009. Ч. 1. Основы теории.

5. Rozhenko A.I. On optimal choice of spline-smoothing parameter // Bull. of Novosibirsk Computing Center. Series Num. Anal. – Novosibirsk: NCC Publisher, 1996. – Iss. 7. – P. 79–86.

6. Гордонова В.И., Морозов В.А. Численные алгоритмы выбора параметра в методе регуляризации // Журн. вычисл. математики и мат. физики. – 1973. –Т. 13, № 3. – С. 539–545.