УДК 528.223
DOI 10.33764/2618-981X-2023-1-1-207-215

В. Ф. Канушин[1], Н. Н. Кобелева[1], Д. Н. Голдобин[1], И. В. Зверев[1]*

Обзор методов интерполирования функции двух переменных в узлы регулярной сети при обработке геофизической информации

[1] Сибирский государственный университет геосистем и технологий, г. Новосибирск, Российская Федерация
* e-mail: n.n.kobeleva@mail.ru

Аннотация. Обработка геофизической информации является одной из важных задач в таких отраслях науки, как геодезия, геофизика, геология. При изысканиях геофизических полей приходится сталкиваться с обработкой больших массивов данных. Эти данные являются исходным материалом для проведения многих исследований и представляют собой результаты наблюдений, полученных из разнородных источников. От качества обработки исходной информации в значительной мере зависит доверие к получаемым результатам. Одной из важных задач обработки исходного цифрового материала являются процедуры интерполяции и экстраполяции. Существует множество методов интерполяции, каждый из которых имеет свои преимущества и недостатки. Обзор этих методов позволит выбрать наиболее подходящий для решения конкретной задачи и обеспечить более точную интерполяцию функций, что в свою очередь способствует более эффективной обработке геофизической информации. В статье проанализированы наиболее известные методы и алгоритмы задач интерполирования результатов наблюдений в узлы регулярной сети. Каждый из рассмотренных методов приводит к различному представлению геофизических данных. Рассмотрена общая постановка задачи интерполирования в узлы равномерной сети значений аномалий силы тяжести в редукции Буге. Приведены результаты оценки точности аппроксимации поля аномалий силы тяжести на локальном участке Новосибирской области (НСО).

Ключевые слова: интерполирование, равномерная сеть, геофизическая информация, функции двух переменных, интерпретация результатов, аномалии силы тяжести, редукция Буге

V. F. Kanushin[1], N. N. Kobeleva[1], D. N. Goldobin[1], I. V. Zverev[1]*

Overview of methods for interpolating a function of two variables into nodes of a regular network in the processing of geophysical information

[1] Siberian State University of Geosystems and Technologies, Novosibirsk, Russian Federation
* e-mail: n.n.kobeleva@mail.ru

Abstract. The processing of geophysical information is one of the important tasks in such branches of science as geodesy, geophysics, and geology. When surveying geophysical fields, one has to deal with the processing of large data arrays. These data are the source material for many studies and are the results of observations obtained from heterogeneous sources. The credibility of the results obtained largely depends on the quality of processing of the initial information. One of the important tasks of processing the original digital material is the interpolation and extrapolation procedures. There are many interpolation methods, each with its own advantages and disadvantages. A review of these methods will allow you to choose the most suitable for solving a particular problem and provide more accurate interpolation of functions, which in turn contributes to more efficient processing of geophysical information. The article analyzes the most well-known methods and algorithms for the problems of interpolating the results of observations into the nodes of a regular network.

Each of the considered methods leads to a different representation of geophysical data. The general statement of the problem of interpolation into the nodes of a uniform network of values gravity anomalies in the Bouguer reduction is considered. The results of estimating the accuracy of approximating the field of gravity anomalies in a local area of the Novosibirsk Region (NSO) are presented.

Keywords: interpolation, uniform network, geophysical information, functions of two variables, interpretation of results, gravity anomalies, Bouguer reduction

Введение

При обработке гравиметрических наблюдений одним из важных моментов является отнесение значений поля, измеренных в произвольно расположенных точках, в узлы равномерной сети. Это обусловлено тем, что большинство разработанных методов интерполяции гравитационных наблюдений предполагают задание исходных значений поля по упорядоченной сети. При разработке машинных алгоритмов решения геолого-геофизических задач выявлено, что квадратная сеть исходных данных более удобна, так как позволяет упростить вычислительные процедуры. Матрицы значений аномалий силы тяжести, заданных в узлах квадратной сети, используются при автоматизированном построении графиков и карт в изолиниях при трансформациях аномальных полей и во многих других случаях.

Отнесение значений аномалий силы тяжести, измеренных в произвольно расположенных точках, в узлы равномерной сети связано с решением задачи интерполяции. Известные варианты решения данного вопроса в большинстве своем основаны либо на аппроксимации поля алгебраическими полиномами и потенциальными функциями, либо на вычислении среднего взвешенного из окрестных значений. При этом одни алгоритмы построены на условии совпадения искомой функции с измеренными значениями без искажений, другие допускают невязку ради получения более гладкого результирующего поля (с целью сглаживания помех) [1–11].

При достаточно густом и равномерном расположении точек наблюдения многие способы интерполяции могут дать близкие результаты. Если сеть точек наблюдений редка и неравномерна (например, профильная съемка), то операция интерполирования с помощью формально-математического аппарата становится малоэффективной или просто неудовлетворительной.

В данной статье рассмотрены современные системы и алгоритмы интерполирования функции двух переменных при обработке геофизической информации.

Общая постановка задачи интерполирования в узлы равномерной сети

Пусть функция аномалий силы тяжести $\Delta g(X,Y)$ нам известна и принадлежит к некоторому обширному классу функции (F).

С другой стороны, указан некоторый класс функций (β). Требуется выделить функцию $Q(X,Y)$ класса (β), которая, мало отличается от данной функции $\Delta g(X,Y)$ [12].

Такова в общих чертах задача приближения функции. Выбор приближающей функции $Q(X,Y)$ выполняется априори на основании эффективности и удобства вычислений. На выбор функции $Q(X,Y)$ влияют следующие факторы: «сложность» поля $\Delta g(X,Y)$ в области D, частота и равномерность расположения исходных данных, наличие дополнительных условий или связей, мощность ЭВМ.

Приемлемость выбранной априори приближающей функции $Q(X,Y)$ определяется, как правило, апостериорной оценкой приближения, например, по контрольным точкам.

После выбора приближающей функции задача заключается в оценке качества приближения. Для её решения необходимо выбрать некоторое функциональное пространство, которому должны принадлежать приближенная и приближающая функции, и качество приближения оценивать расстоянием между этими элементами пространства. Примеры применяемых пространств и соответствующих метрик приведены в таблице 1.

Таблица 1

Примеры метрик

Элементы пространства	Обозначения	Метрики
Непрерывные функции	L_T	$d(\bar{a}_1,\bar{a}_2)=\iint\limits_T \lvert\bar{a}_1(x,y)-\bar{a}_2(x,y)\rvert dxdy$
	L_T^2	$d(\bar{a}_1,\bar{a}_2)=\sqrt{\iint\limits_T \lvert\bar{a}_1(x,y)-\bar{a}_2(x,y)\rvert^2 dxdy}$
	M_T	$d(\bar{a}_1,\bar{a}_2)=\max\limits_{(x,y)}\lvert\bar{a}_1(x,y)-\bar{a}_2(x,y)\rvert$
Дискретные функции	L_N	$d(\bar{a}_1,\bar{a}_2)=\sum\limits_{n=0}^{N-1}\lvert\bar{a}_{1n}-\bar{a}_{2n}\rvert$
	l_N^2	$d(\bar{a}_1,\bar{a}_2)=\sqrt{\sum\limits_{n=0}^{N-1}\lvert\bar{a}_{1n}-\bar{a}_{2n}\rvert^2}$
	m_N	$d(\bar{a}_1,\bar{a}_2)=\max\limits_{(x,y)}\lvert\bar{a}_{1n}-\bar{a}_{2n}\rvert$

В таблице 1 $\bar{a}_1$ и $\bar{a}_2$ – векторы приближенной и приближающей функций соответственно – элементы функционального пространства, d – расстояние между элементами $\bar{a}_1$ и $\bar{a}_2$.

После решения задачи определения многочлена $Q(X,Y)$ оцениваем точность приближения и вычисляем значение функции $\Delta g(X,Y)$ на множестве точек $k \in D$, которые могут быть также и узлами квадратной сети.

Обзор методов интерполирования функции двух переменных в узлы регулярной сети

Построение сетки – это создание регулярного массива значений Z –координат узловых точек по нерегулярному массиву *(X, Y, Z)* – координат исходных точек. Термин «нерегулярный массив координат» означает, что координаты *X, Y* точек данных распределены по области карты неравномерно. Регулярный массив узловых точек требуется для создания цифровых моделей местности и их машинной реализации. Процедура построения сетки представляет собой интерполяцию или экстраполяцию значений исходных точек данных на равномерно распределенные узлы в исследуемой области.

За последние двадцать лет разработано много различных систем и алгоритмов интерполирования функции двух переменных [13 – 15].

Применимость метода интерполяции аномалий силы тяжести в редукции Буге характеризуется несколькими критериями, которые сводятся к качеству приближения аппроксимирующей функции. Ошибка аппроксимирующей функции должна быть минимальной, как показывает формула 1

$$Q(X,Y) - \Delta g(X,Y) = \min \tag{1}$$

Важным элементом при обработке геофизических полей является возможность экстраполяции – заполнение «белых пятен» и получение значений за пределами диапазона данных.

Основываясь на данных критериях, рассмотрим применимость современных систем и алгоритмов интерполирования функции двух переменных построения регулярной сети.

Метод обратных расстояний – метод построения сеточной функции, применяется при использовании функции «Степень обратного расстояния». Данные взвешиваются во время интерполяции таким образом, что влияние одной точки относительно другой уменьшается с расстоянием от узла сетки. Весовые коэффициенты присваиваются данным с помощью весовой силы, которая контролирует, как весовые коэффициенты уменьшаются по мере увеличения расстояния от узла сетки. Чем больше весовой коэффициент, тем меньше влияние точек, удаленных от узла сетки, при интерполяции. По мере увеличения расстояния, значение узла сетки приближается к значению ближайшей точки. При небольшом расстоянии весовые коэффициенты более равномерно распределяются между соседними точками данных. Метод работает быстро, но плохо справляется с областями исходных данных с высокими горизонтальными градиентами, не экстраполирует значения за пределы диапазона данных.

Метод Кригинга – это метод построения геостатистической сетки – полезный и популярный во многих областях науки, является одним из наиболее гибких комбинированных методов, применяющихся для представления геофизических измерений. Для большинства наборов данных он весьма эффективен. При этом для большого количества исходных данных может быть весьма медленным.

Метод выполняет экстраполирование значений сетки за пределы диапазона данных.

Метод минимальной кривизны – широко используется в науках о Земле. Поверхность, построенная с помощью этого метода, аналогична тонкой упругой пленке, проходящей через все экспериментальные точки данных с минимальным числом изгибов. Он генерирует наиболее гладкую поверхность, которая проходит настолько близко к экспериментальным точкам, насколько это возможно, но эти экспериментальные точки не обязательно принадлежат интерполяционной поверхности.

Метод создает сетку путем многократного применения уравнения к сетке в попытке сгладить ее. Каждый проход по сетке считается одной итерацией. Значения узлов сетки пересчитываются до тех пор, пока последовательные изменения значений не станут меньше, или пока не будет достигнуто максимальное количество итераций.

Метод минимальной кривизны создает гладкие поверхности и работает быстро для большинства наборов данных, может экстраполировать значения за пределы диапазона данных, но при этом могут возникать значительные ошибки. Позволяет контролировать степень сглаживания. Метод минимальной кривизны не является точным методом.

Метод «естественного соседа» интерполирует значения сетки путем взвешивания соседних точек данных на основе пропорциональных площадей. В основе используется алгоритм, находящий ближайшее подмножество входных выборок к точке запроса и применяющей к ним веса на основе пропорциональных площадей для интерполяции значения. Он также известен как интерполяция с захватом площади. Метод создает хорошие контуры из наборов данных, содержащих плотные значения величин в одних областях и разреженные в других, не генерирует информацию в областях без данных и не экстраполирует значения за пределы диапазона данных.

Метод полиномиальной регрессии – используется для оценки пространственного тренда в данных, то есть зависимости изучаемой пространственной переменной Z от координат X, Y. Он не является интерполяционным в прямом смысле, так как не пытается предсказать неизвестные значения Z. Метод предназначен для выявления глобальных пространственных трендов и теряет детальную локальную информацию, содержащуюся в данных. Обрабатывает информацию таким образом, чтобы отображались лежащие в их основе крупномасштабные тенденции и закономерности, что используется для анализа поверхности тренда. Полиномиальная регрессия выполняется очень быстро для любого объема данных, но локальные детали теряются в сгенерированной сетке. Данный метод может экстраполировать значения сетки за пределы диапазона данных.

Метод радиальных базисных функций – довольно гибкий метод. Его можно сравнить с кригингом, поскольку он дает наилучшие общие интерпретации большинства наборов данных. Данный метод представляет собой разнообразную группу методов интерполяции. С точки зрения способности соответствовать вводимым данным и создавать гладкую поверхность, многие исследователи считают его лучшим. Все алгоритмы радиальной базисной функции являются точными интерполяторами, независимо от значения принимаемого коэффициента, поэтому они пытаются достоверно соответствовать задаваемым данным. Во все алгоритмы можно ввести формирующий коэффициент, чтобы получить более гладкую поверхность.

Модифицированный метод Шепарда – подобен методу обратных расстояний, но позволяет уменьшить эффект «бычьих глаз» в генерируемых контурах. В качестве оценки переменной в произвольной точке области исследования используется среднее взвешенное значений аппроксимирующих функций в этой точке. Не имеет тенденции генерировать концентрических контуров вокруг точек данных, особенно при использовании коэффициента сглаживания. Модифицированный метод Шепарда может экстраполировать значения за пределы диапазона данных, может быть как точным, так и сглаживающим интерполятором.

Триангуляция с линейной интерполяцией – быстро выполняет задачу интерполирования, не экстраполирует значения за пределы диапазона данных. Метод использует оптимальную триангуляцию Делоне. Алгоритм создает треугольники, проводя линии между точками данных. Исходные точки соединяются таким образом, чтобы ни одна грань треугольника не пересекалась с другими треугольниками. В результате получается лоскутное одеяло из треугольных граней по всей площади сетки. Этот метод является точным интерполятором. Триангуляция с линейной интерполяцией лучше всего работает, когда данные равномерно распределены по области сетки. Наборы данных, содержащие разреженные области, приводят к появлению на карте отдельных треугольных граней.

Метод скользящего среднего – является полезным инструментом для характеристики и исследования больших и очень больших наборов пространственных данных (более 1000 наблюдений). Скользящее среднее извлекает тренды и вариации среднего масштаба из больших зашумленных выборок и присваивает значения узлам сетки путем усреднения данных в пределах эллипса поиска узла сетки. Не рекомендуется для создания карт на основе небольших и умеренных наборов измерений. Данный метод является разумной альтернативой «Методу естественного соседа» для создания сетки из больших регулярно расположенных наборов величин.

Метод локального полинома – наиболее применим к наборам данных, которые являются локально гладкими (т. е. относительно гладкими поверхностями в окрестностях поиска). Размер исходной выборки не влияет на скорость вычислений. Данный метод присваивает значения узлам сетки, используя взвешенную подгонку по методу наименьших квадратов с данными в пределах эллипса поиска узла сетки.

Результаты

Параметры метода построения сетки управляют процедурами интерполяции. Различия между приемами построения заключаются в математических алгоритмах, используемых для расчета весов при интерполяции узлов сетки. Каждый способ приводит к различному представлению данных [16 – 23]. Необходимо протестировать каждый метод на типичном наборе измеренных величин, чтобы определить алгоритм построения сетки, обеспечивающий наиболее удовлетворительную интерпретацию данных.

С целью демонстрации точностных характеристик рассмотренных аппроксимационных методов выполнена оценка точности аппроксимации участка поля силы тяжести в виде набора точек, полученных по результатам площадной гравиметрической съемки на территории НСО. Участок представляет собой территорию площадью около 5000 км^2 с количеством 1500 пунктов, средним расстоянием между ними 1 – 3 км. Рельеф выбранной территории преимущественно равнинный. Из набора исходных точек выделено 10 процентов, которые взяты как контрольные и не участвовали в аппроксимации.

С помощью каждого рассмотренного метода создана регулярная сетка точек (с шагом около 1 км) и выполнена оценка точности в исходных и контрольных точках.

Результаты оценки точности аппроксимации рассмотренных методов приведены в таблице 2.

Таблица 2

Средние квадратические погрешности в исходных и контрольных точках

Метод интеполирования	СКП в исходных точках (мГал)	СКП в контрольных точках, (мГал)
Кригинг	0,19	0,56
Метод минимальной кривизны	0,21	0,66
Метод обратных расстояний	0,35	0,82
Триангуляция с линейной интерполяцией	0,20	0,60
«Естественный сосед»	0.22	0,58
Метод локального полинома	1,25	1,17
Радиальные базисные функции	0,11	0,58
Модифицированный метод Шепарда	0,08	0,70
Метод скользящего среднего	3,18	3,43
Полиномиальная регрессия	3,12	3,30

Заключение

Основным показателем качества аппроксимации следует рассматривать среднеквадратическую погрешность (СКП) в контрольных точках построения сетки. Немаловажную роль играет и СКП в исходных точках.

Результаты исследований позволяют сделать вывод, что метод «Кригинга» является более точным методом построения геофизических полей, демонстрируя наименьшую СКП в контрольных точках на выбранной территории. При этом «Модифицированный метод Шепарда» показал наименьшую СКП в исходных точках. Использование того или иного метода интерполяции зависит от задач исследований, что в свою очередь способствует более эффективной обработке геофизической информации.

Следует подчеркнуть, что все рассмотренные в таблице 3 алгоритмы пригодны лишь в тех случаях, когда плотность исходной сети обеспечивает интерполирование с погрешностью не более нескольких процентов [1–3, 5]. Если же сеть редка и неравномерна, и ошибки интерполяции, составляют десятки процентов и более, то применение формальных математических методов приближения функции становится малоэффективным или просто недопустимым [1–3, 5].

Исследование выполнено в рамках СЧ НИР «ГЕОТЕХ-Квант» с целью создания высокоточных моделей геопотенциального поля Земли и его характеристик на территории РФ.

БИБЛИОГРАФИЧЕСКИЙ СПИСОК

1. Аронов, В. И. Методы построения карт геолого-геофизических признаков и геометризации залежей нефти и газа на ЭВМ. М.: Недра. 1990. 300 с.

2. Аронов В.И. Методы математической обработки геологических данных на ЭВМ. М.: Недра, 1977. 168 с

3. Бычков, С. Г. Особенности обработки результатов современной гравиметрической съемки // Геофизический вестник. 2005. №12. С. 9-13.

4. Бычков, С. Г. К вопросу о вычислении аномалий силы тяжести в редукции Буге // Вопросы теории и практики геологической интерпретации гравитационных, магнитных и электрических полей. Екатеринбург: ИГф УрО РАН:, 2006. С. 73-77.

5. Бычков, С. Г, Симанов, А. А., Хохлова, В. В. Опыт использования современных процедур обработки высокоточных гравиметрических наблюдений // Глубинное строение. Геодинамика. Тепловое поле Земли. Интерпретация геофизических полей. Екатеринбург: ИГф УрО РАН. 2013а. С. 44-46.

6. Долгаль, А. С. Аппроксимация геопотенциальных полей эквивалентными источниками при решении практических задач // Геофизический журнал. 1999. Т. 21, №4. С. 71-80.

7. Долгаль, А. С. Компьютерные технологии обработки и интерпретации данных гравиметрической и магнитной съемок в горной местности. Абакан: ООО Фирма-МАРТ. 2002. 188 с.

8. Керимов, И. А. Метод F-аппроксимаций при решении задач гравиметрии и магнитометрии. М.: Физматлит. 2011. 264 с.

9. Клейнер, М. В., Бычков, С. Г., Микрюков, Г. Л. Об интерполировании вычисленных на ЭВМ поправок за влияние рельефа // Вопросы обработки и интерпретации геофизических наблюдений. 1974. Пермь: Пермский гос. ун-т. №11. С. 157-158.

10. Степанова, И. Э. S-аппроксимация рельефа земной поверхности // Вопросы теории и практики геологической интерпретации гравитационных, магнитных и электрических полей: материалы. Пермь: ГИ УрО РАН. 2005. С. 265-266.

11. Хохлова, В. В., Симанов, А. А. Современные методы обработки высокоточных гравиметрических наблюдений // Уральская молодежная научная школа по геофизике. Пермь: ГИ УрОРАН. 2013. С. 261-263.

12. Беспрозванный, П. А., Чернов, А. А. Расчет значений полей в узлах квадратной сети// Разведочная геофизика №62 , М.: Недра. 1974. С. 108-117.

13. Мальцев, К. А., Мухарамова, С.С. Построение моделей пространственных переменных (с применением пакета Surfer): Учебное пособие. – Казань: Казанский университет, 2014. – 103 с.
14. К.Ю. Силкин, К. Ю. Геоинформационная система Golden Software Surfer 8: Учебное пособие. – Воронеж: Воронежский государственный университет, 2008. – 66 с.
15. Половко, А. М., Бутусов, П. Н. Интерполяция. Методы и компьютерные технологии их реализации. BHV, 2004 . - 320 с.
16. Cressie, N. A. C. (1990), "The Origins of Kriging", Mathematical Geology, v. 22, p. 239-252.
17. Powell, M.J.D. (1990), The Theory of Radial Basis Function Approximation in 1990, University of Cambridge Numerical Analysis Reports, DAMTP 1990/NA11.
18. Hardy, R. L. (1990), Theory and Applications of the Multiquadric-BiHarmonic Method, Computers Math. Applic, v. 19, n. 8/9, p. 163-208.
19. Cressie, N. A. C. (1991), Statistics for Spatial Data, John Wiley and Sons, Inc., New York, 900 pp.
20. Deutsch, C. V., and Journel, A. G. (1992), GSLIB - Geostatistical Software Library and User's Guide, Oxford University Press, New York, 338 pp.
21. Isaaks, E. H., and Srivastava, R. M. (1989), An Introduction to Applied Geostatistics, Oxford University Press, New York, 561 pp.
22. Watson, Dave (1994), Nngridr - An Implementation of Natural Neighbor Interpolation, David Watson, P.O. Box 734, Clarement, WA 6010, Australia.
23. Carlson, R.E., and Foley, T.A. (1991a), Radial Basis Interpolation Methods on Track Data, Lawrence Livermore National Laboratory, UCRL-JC-1074238.