

Санкт-Петербургский государственный университет  
Направление: 02.04.02 Фундаментальная информатика и информационные  
технологии  
ОП: Технологии искусственного интеллекта и Big Data

## ОТЧЕТ О НАУЧНО-ИССЛЕДОВАТЕЛЬСКОЙ РАБОТЕ

на тему: Модуль анализа вариограммы

Выполнил магистр: \_\_\_\_\_1\_\_\_\_\_ курса

Шэнь Жуйцзи

(Ф.И.О.)

\_\_\_\_\_  
(подпись)

Преподаватель-руководитель:

Митрофанова О. А., к.ф-м.н., доцент

(Ф.И.О., степень, звание, должность)

\_\_\_\_\_  
(подпись)

## СОДЕРЖАНИЕ

<b>Введение .....</b>	<b>3</b>
Цели исследования .....	4
Задачи исследования .....	4
<b>Теоретическая часть .....</b>	<b>4</b>
Основные концепции вариограммы .....	4
Структурная схема вариограммы .....	5
Модель вариограммы .....	6
Оптимальная аппроксимация вариограммы .....	7
<b>Практическая часть .....</b>	<b>8</b>

# Введение

Геостатистика находит все более широкое применение в таких областях, как оценка ресурсов, экологический мониторинг и планирование использования земли. Особенно важную роль в этом играет анализ и интерпретация пространственных данных. Анализ вариограммы, как ключевой инструмент оценки связанности пространственных данных, имеет важное значение для понимания и моделирования структуры пространственных переменных. Поэтому дальнейшие исследования и улучшения методов анализа вариограмм имеют важное практическое значение для повышения точности и эффективности обработки пространственных данных.

Учитывая быстрый рост объемов пространственных данных, традиционные методы их обработки уже не могут удовлетворить современные потребности. Разработка модуля анализа вариограммы может значительно повысить эффективность обработки данных и достоверность результатов, обеспечивая более точную поддержку в областях градостроительства, прогнозирования катастроф и управления ресурсами.

Данное исследование сосредоточено на модуле анализа вариограмм, в частности, на расчетах и аппроксимации, обучении

вариограмм. Оптимизация обучения вариограмм позволит улучшить интерпретируемость и точность прогнозов модели.

## **Цели исследования**

1. Разработать и обучить новые алгоритмы для повышения точности и эффективности оценки вариограмм.
2. Реализовать удобный в использовании модуль анализа вариограмм.

## **Задачи исследования**

1. Разработка веб-интерфейса (на основе Java);
2. Подключение к базе данных (на основе MySQL);
3. Вчисление и аппроксимация вариационных функций (на основе Python).

# **Теоретическая часть**

## **Основные концепции вариограммы**

Фон: Для понимания вариограммы необходимо сначала ознакомиться с методом Кригинга, который является крайне важным понятием в геостатистике. Метод Кригинга основан на аппроксимации данных известных точек для определения значений

в каждой позиции в пределах заданного радиуса или заданного количества точек. Это многошаговый процесс, включающий исследование статистики данных, моделирование вариограммы и создание поверхностей, изучение поверхностей дисперсии. Получение хороших результатов аппроксимации с помощью метода Кригинга всегда было предметом стремлений.

Концепция: Вариограмма — это прежде всего функция, подобно обычным одно- и двумерным функциям, которые мы знаем. Она также имеет свои независимые переменные, зависимые переменные и функциональные выражения. Вариограмма первоначально описывает характеристики и интенсивность пространственных изменений региональной переменной, определяется как математическое ожидание квадрата приращения региональной переменной, где зависимая переменная — это шаг ( $h$ ), а независимая переменная — значение изменения.

### **Структурная схема вариограммы**

Дальность: Размер области пространственной автокорреляции регионализированной переменной. Обратите внимание на приведенную ниже кривую: когда расстояние  $h$  увеличивается до определенного значения, модель на этом участке становится горизонтальной. Это расстояние называется дальностью. Образцы,

находящиеся на расстоянии меньшем, чем дальность, пространственно автокоррелированы, а образцы, находящиеся на большем расстоянии, не имеют пространственной автокорреляции.

Порог: Величина изменчивости регионализированной переменной (не все модели имеют порог, вариограммные модели можно разделить на модели с порогом и без порога).

Самородок: Величина случайности регионализированной переменной (отклонение порога + самородок).

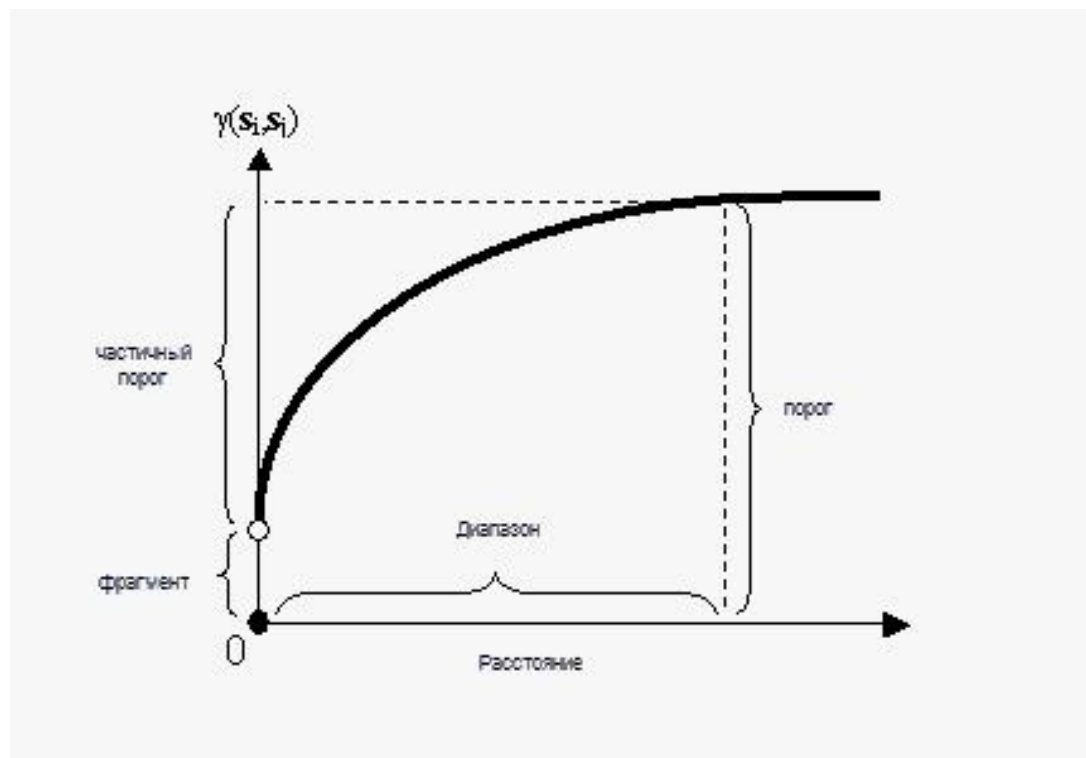


Рис. 1

## Модель вариограммы

Для получения вариограммы обычно необходимо выбрать теоретическую модель вариограммы и вычислить значения

самородка, порога и дальности. Большинство теоретических моделей вариограмм являются нелинейными (например, сферическая модель, экспоненциальная модель и гауссовская модель), и такие модели имеют множество параметров. Их аппроксимация по сути является методом оптимизации многопараметрической нелинейной функции. В настоящее время наиболее зрелыми методами являются взвешенный метод полиномиальной регрессии, взвешенный метод линейного программирования и метод целевого программирования.

Если в модели с порогом используется гауссовская модель, формула выглядит следующим образом:

$$\gamma(h) = \begin{cases} 0 & , \quad h = 0 \\ C_0 + C(1 - e^{-\frac{h^2}{a^2}}) & , \quad h > 0 \end{cases} ;$$

$C_0$  — значение самородка;  $C$  — высота порога;  $C_0 + C$  — значение порога.

### **Оптимальная аппроксимация вариограммы**

Оптимальная аппроксимация теоретической модели вариограммы включает три основных шага:

- 1.Определение формы модели вариограммы/типа кривой
- 2.Оптимальная оценка параметров модели
- 3.Оценка аппроксимации модели

## **Практическая часть**