

**МИНИСТЕРСТВО НАУКИ И ВЫСШЕГО ОБРАЗОВАНИЯ
РОССИЙСКОЙ ФЕДЕРАЦИИ ФЕДЕРАЛЬНОЕ ГОСУДАРСТВЕННОЕ
АВТОНОМНОЕ ОБРАЗОВАТЕЛЬНОЕ УЧРЕЖДЕНИЕ ВЫСШЕГО
ОБРАЗОВАНИЯ «КАЗАНСКИЙ (ПРИВОЛЖСКИЙ) ФЕДЕРАЛЬНЫЙ
УНИВЕРСИТЕТ» ИНСТИТУТ МАТЕМАТИКИ И МЕХАНИКИ ИМ. Н.И.
ЛОБАЧЕВСКОГО КАФЕДРА АЭРОГИДРОМЕХАНИКИ**

Направление подготовки (специальность): 01.04.03 - Механика и математическое моделирование

Профиль (специализация, магистерская программа): Механика жидкости газа и плазмы

МАГИСТЕРСКАЯ ДИССЕРТАЦИЯ

**ГЕНЕРАЦИЯ ТУРБУЛЕНТНОГО ПОЛЯ СКОРОСТИ ДЛЯ ВХОДНЫХ
ГРАНИЧНЫХ УСЛОВИЙ**

На правах рукописи

УДК **xxx.xxx**

Фамилия Имя Отчество автора

НАЗВАНИЕ ДИССЕРТАЦИОННОЙ РАБОТЫ

Специальность **XX.XX.XX** —

«Название специальности»

Диссертация на соискание учёной степени

кандидата физико-математических наук

Научный руководитель:

уч. степень, уч. звание

Фамилия Имя Отчество

Город — 20XX

Оглавление

	Стр.
Глава 1. Метод стохастического моделирования	3
1.1 Обзор метода стохастического моделирования Гаусса	3
1.2 Математическая модель метода стохастического Гауссова моделирования	4
1.3 Метод кокригинга для генерации флуктуаций скоростей	7

Глава 1. Метод стохастического моделирования

1.1 Обзор метода стохастического моделирования Гаусса

Наиболее часто применения метода стохастического Гауссова моделирования можно найти в геофизических задачах для проведения интерполяции некоторых величин в межскважинном пространстве[0, с. 33]. Практика применения данного метода исходит из того факта, что большинство физических величин, которые простым или доступным способом могут быть измерены, не могут дать вполне четкой картины о распределении косвенно зависящих параметров. Поводом для этого может служить например небольшой набор исходных данных, сильная пространственная разреженность измерений. Интерполяция измеренных данных на протяжении интересующей области, в следствии выше сказанного, играет важную роль в процессе освоения нефтяных пластов, если касаться одного из возможных применений рассматриваемого метода.

В общем случае имеется достаточно большой набор методов, с помощью которых можно провести интерполяцию данных, помимо упомянутого метода стохастического моделирования, это, например, методы линейной интерполяции, полиномиальная интерполяция (Лагранжевы многочлены). Близко к полиномиальным методам интерполяции можно отнести сплайн-функции и их частный случай, довольно распространённые кубические сплайны. Примерами, относящимися к многомерной интерполяции могут служить: билинейная интерполяция, бикубическая интерполяция.

Имеется достаточной большой выбор разнородных методов интерполяции, но из них выделяется рассматриваемый в данной главе метод интерполяции — метод стохастического моделирования — одним из названий которого также является кригинг, а также метод Винера-Колмогорова. Основная идея данного метода состоит в использовании некоторого Гауссова процесса для интерполяции известных значений по критерию минимизации среднеквадратичного отклонения на основе матрицы ковариации, и также для оптимизации гладкости интерполируемых значений

Часто выделяют тот факт, что метод кригинга удовлетворяет условию наилучшего линейного несмещённого предсказания (BLUP - best linear unbiased prediction). Данный термин был введён для оценки случайных величин в линейных смешанных моделях. Есть близкий, практически эквивалентный термин - "лучшая линейная несмещенная оценка" (BLUE - best linear unbiased estimation) для заданных случайных величин, возникающий как часть известной теоремы Гаусса-Маркова[0, с. 64]. Данная теорема гласит, что обыкновенная оценка наименьших квадратов (OLS - ordinary least squares) имеет наименьшую дисперсию выборки в классе несмещённых оценок, если ошибки линейной регрессионной модели некоррелированы, имеют нулевое математическое ожидание и равные дисперсии. Различие BLUP и BLUE заключается в том, что в случае BLUP подразумевается оценка прогнозирования случайных величин, в случае BLUE, говорят об оценке случайной величины. Данный критерий (BLUP) может не выполняться для других методов описанных выше, например для метода сплайнов. Так как определение флуктуации компонент скоростей, из её физического смысла, подразумевает то, что её математическое ожидание равно 0, подразумевая осреднение как по времени, так и по пространству, автоматически удовлетворяется одно из условий теоремы. Удовлетворение другим условиям теоремы будет рассмотрено далее.

Метод кригинга, как и спектральный метод, относится к классу недетерминистических методов, то есть стохастических методов. Оба метода не используют в основе какую-то жесткую модель, в большей своей степени, вводя зависимость от некоторого случайного процесса, в предположении, характеризующего некоторый реальный процесс распределения физической величины. Метод кригинга имеет больший упор на удовлетворение пространственной ковариации, так как генерируемая случайная переменная Z описывает разброс вокруг неизвестного неизмеренного значения z .

1.2 Математическая модель метода стохастического Гауссова моделирования

В данной части рассмотрим основы метода кригинга. Как говорилось ранее метод кригинга является интерполяционным методом. Это означает, что перед на-

ми стоит задача интерполировать некоторую функцию $f_i = f(\vec{r}_i)$, где f_i — известный набор значений интерполируемой функции в точках \vec{r}_i . Проведение интерполяции подразумевает под собой нахождение функции F по определению:

$$\bar{f}(\vec{r}) = F(\vec{r}, \{f_i\}, \{\vec{r}_i\}), \quad (1.1)$$

где $\bar{f}(\vec{r})$ — проинтерполированной функции в точке пространства с радиус-вектором \vec{r} , F - функция интерполятор, определяющая правило интерполяции известных значений, f_i - набор известных значений, \vec{r}_i - точки в которых определены эти значения. Как видно из вводимого определения интерполятор F зависит от $2N + 1$ значений, где N — число известных значений. На F накладывается естественное условие точного совпадения значений в известных точках, или

$$\bar{f}_j = F(\vec{r}_j, \{f_i\}, \{\vec{r}_i\}) \quad (1.2)$$

В формуле (1.2): $\bar{f}_j = \bar{f}(\vec{r}_j)$ и $j = \overline{1..N}$, что представляет собой N условий налагаемых на функцию F .

В общем случае, вид функции F не известен, поэтому необходимо применить упрощения. Например, на порядок членов входящих в эту функцию. Наиболее часто встречаемый вид — это линейная зависимость. Как упоминалось выше, кригинг — линейный метод, поэтому применительно к нашему случаю мы будем строить интерполятор вида:

$$\bar{f} = \sum_i f_i w_i(\vec{r}), \quad (1.3)$$

где $f_i = f(\vec{r}_i)$ — уже ранее упомянутые известные значения функции, $w_i(\vec{r})$ - веса привязанные к значениям f_i .

Стоит отметить, что в более общей формулировке мы имеем:

$$\bar{f}(\vec{r}) - E(\bar{f}(\vec{r})) = \sum_i w_i(\vec{r})(f(\vec{r}_i) - E(f(\vec{r}_i))) \quad (1.4)$$

Введём общепринятое обозначения для средней величины: $m(\vec{r}) = E(\bar{f}(\vec{r}))$ и $m(\vec{r}_i) = E(f(\vec{r}_i))$, тогда:

$$\bar{f}(\vec{r}) - m(\vec{r}) = \sum_i w_i(\vec{r})(f(\vec{r}_i) - m(\vec{r}_i)) \quad (1.5)$$

Здесь $\bar{f}(\vec{r})$ — оценка значения неизвестной функции $f(\vec{r})$. По сравнению с (1.3) учитывается, что математическое ожидание для $\bar{f}(\vec{r})$ и для f_i могут быть не равны нулю. Но даже в таком случае обычно представляют функцию в виде $g(\vec{r}) = m(\vec{r}) + \xi(\vec{r})$, где $m(\vec{r})$ — среднее трендовое значение, $\xi(\vec{r})$ — случайная составляющая, таким образом приходя к уравнению (1.3). Здесь предполагается что $\xi(\vec{r})$ такая, что $E(\xi(\vec{r})) \equiv 0$, также стоит отметить что в таком случае: $E(g(\vec{r})) = m(\vec{r})$, то-есть среднее искомой функции есть трендовое значение.

Рассмотрим значение ошибки получаемое в случае использования (1.4):

$$R(\vec{r}) = \bar{f}(\vec{r}) - f(\vec{r}) \quad (1.6)$$

Как говорилось ранее метод удовлетворяет условию BLUE - условие не смещённости оценки, то-есть $E(R) = 0$. Это можно показать прямой подстановкой:

$$\begin{aligned} E(R(\vec{r})) &= E(\bar{f}(\vec{r}) - f(\vec{r})) = E(\bar{f}(\vec{r})) - E(f(\vec{r})) \\ &= E(m(\vec{r}) + \sum_i w_i(\vec{r})(f(\vec{r}_i) - m(\vec{r}_i))) - m(\vec{r}) = 0 \end{aligned} \quad (1.7)$$

Слагаемые отвечающие за математическое ожидание взаимоуничтожаются по определениям введённым выше, слагаемые внутри суммы также взаимоуничтожаются, что в конечном итоге приводит к нулевому математическому ожиданию оценки ошибки. Таким образом мы удовлетворили одному из пунктов теоремы Гаусса-Маркова строго. Далее приведём доказательство удовлетворению остальным пунктам теоремы.

Необходимо доказать, что построенный таким образом метод имеет наименьшую дисперсию в классе несмещённых оценок

Далее введём в рассмотрение функцию пространственной ковариации:

$$Cov(\xi(\vec{r}), \xi(\vec{r} + \vec{h})) = C_\xi(\vec{h}) = E((\xi(\vec{r}) - E(\xi(\vec{r}))) \cdot (\xi(\vec{r} + \vec{h}) - E(\xi(\vec{r} + \vec{h})))) \quad (1.8)$$

Данная функция нужна нам для внесения самосогласованности в метод. При рассмотрении формулы (1.3) можно заметить, что нет зависимости между соседними точками. Из-за этого возникает сильное изменение решения при появлении дополнительных точек. Но, если до этого момента не было необходимости рассматривать решение систем алгебраических уравнений, так как не присутствовало связанности между соседними узлами, то теперь внесём её за счёт учета некоторого

расстояния между точками.

$$\bar{f}(\vec{r}) = \sum_i \lambda_i \Omega(\vec{r} - \vec{r}_i). \quad (1.9)$$

Здесь λ_i – некоторые весовые коэффициенты, которые необходимо будет подобрать в ходе решения, $\Omega(\vec{r} - \vec{r}_i)$ – базисная функция монотонно убывающая с увеличением расстояния между точками. Как и ранее для известных значений мы должны получать тождество.

$$\bar{f}_k = \sum_i \lambda_i \Omega(\vec{r}_k - \vec{r}_i) \quad (1.10)$$

В итоге имеется система линейных уравнений для вычисления весовых коэффициентов λ . В матричном виде:

$$\bar{F} = \Omega \cdot \Lambda \quad (1.11)$$

Решением будет являться:

$$\Lambda = \Omega^{-1} \bar{F} \quad (1.12)$$

Если подставить (1.12) в (1.10) можно избавиться от параметра λ .

$$\bar{f}(\vec{r}) = \sum_i f_i \cdot \sum_n [\Omega_{in}^{-1} \cdot \omega(\vec{r}_n - \vec{r})] \quad (1.13)$$

Требование к выполнению условия, что в известных точках получаются точные известные значения, также выполняется.

$$\bar{f}(\vec{r}_i) = \sum_i f_i \cdot \sum_n [\Omega_{in}^{-1} \cdot \omega(\vec{r}_n - \vec{r}_i)] = f_i \quad (1.14)$$

1.3 Метод кокригинга для генерации флуктуаций скоростей

Метод кригинга изначально не учитывает возможные дополнительные корреляции с другими величинами. Это может сыграть роль для уточнения данных, так как зачастую при проведении физических экспериментов или измерений, из-

меряют сразу несколько величин. Так как кригинг наиболее часто используется в геологическом моделировании наиболее простым примером является измерение пористости и проницаемости или какого-нибудь другого набора параметров. Так мы можем дополнить модель используя ковариации между измеряемыми величинами, до этого момента мы рассматривали лишь корреляцию случайной величины сама с собой. Корреляция двух случайных величин представляется в виде:

$$\text{cov}(\xi, \zeta) = \frac{1}{N} \sum_{n=1}^N (\xi_n - \bar{\xi}) * (\zeta_n - \bar{\zeta}) \quad (1.15)$$

По аналогии с обычным методом кригинга, дополним модель дополнительным учётом зависимости между переменными. Оценим функцию z в окрестности точки \vec{r}_0 .

$$\overline{z}_{\alpha_0}(\vec{r}_0) = \sum_{\alpha=1}^K \sum_{i=1}^{n_{\alpha}} \lambda_i^{\alpha} z_{\alpha}(\vec{r}_i), \quad (1.16)$$

здесь $\overline{z}_{\alpha_0}(\vec{r}_0)$ – оценка в виде линейной комбинации различных переменных в окрестности точки \vec{r}_0 , λ_i^{α} – весовые коэффициенты. Как и обычный кригинг, ко-кригинг также должен удовлетворять условию несмещённости для построенной оценки (1.16).

$$\begin{aligned} E \{ \overline{z}_{\alpha_0}(\vec{r}_0) - z_{\alpha_0}(\vec{r}_0) \} &= E \left\{ \sum_{\alpha=1, \alpha \neq \alpha_0}^K \sum_{i=1}^{n_{\alpha}} \lambda_i^{\alpha} z_{\alpha}(\vec{r}_i) + \sum_{i=1}^{n_{\alpha_0}} \lambda_i^{\alpha_0} z_{\alpha_0}(\vec{r}_i) - z_{\alpha_0}(\vec{r}_0) \right\} \\ &= \sum_{\alpha=1, \alpha \neq \alpha_0}^K \left(m_{\alpha} \sum_{i=1}^{n_{\alpha_0}} \lambda_i^{\alpha} \right) + m_{\alpha_0} \left(\sum_{i=1}^{n_{\alpha_0}} \lambda_i^{\alpha_0} - 1 \right) = 0 \end{aligned} \quad (1.17)$$

Также дополним условие (1.2), веса должны быть подобраны таким образом, чтобы при основной переменной с индексом α_0 были равны 1, при любых других переменных, не соответствующих основной, коэффициенты должны быть равны нулю, то есть:

$$\sum_{i=1}^{n_{\alpha}} \lambda_i^{\alpha} = \delta_{\alpha\alpha_0} \quad (1.18)$$

При учёте ограничивающих выражений (1.3), получаем следующую систему уравнений для *обычного* кокригинга:

$$\begin{cases} \sum_{\beta=1}^K \sum_{j=1}^{n_{\beta i}} \lambda_j^{\beta} \gamma_{\alpha\beta}(\vec{r}_i - \vec{r}_j) + \mu_i = \gamma_{\alpha\alpha_0}(\vec{r}_i - \vec{r}_{i0}), \alpha = 1, \dots, K, i = 1, \dots, n_{\alpha} \\ \sum_{i=1}^{n_{\alpha}} \lambda_i^{\alpha} = \delta_{\alpha\alpha_0}, \alpha = 1, \dots, K, i = 1, \dots, n_{\alpha} \end{cases}$$

Таким образом используя систему 1.3 мы можем построить поле случайной величины удовлетворяющей некоторой заданной функции ковариаций. В качестве опорных точек для генерации можно взять, например, последовательность нормально распределённых чисел. Это не будет противоречить постановке задачи кригинга из-за в целом построения случайного процесса и в то же время внесёт дополнительную либо случайную составляющую, либо мы можем взять значения с некоторого поля и решить задачу аппроксимации кригингом для неё.

Задача кригинга и кокригинга требует в качестве входных данных либо аналитически заданную функцию ковариаций, либо её сеточную аппроксимацию. Основная проблема в том, чтобы каким-то образом перейти от энергетического спектра к ковариационной функции. Следуя из определения для энергетического спектра [0] можно обратными преобразованиями перейти от энергетического спектра к ковариационной функции:

$$\Phi_{ij}(\vec{k}) = \frac{1}{(2\pi)^3} \iiint_{-\infty}^{\infty} R_{ij}(\vec{r}) \exp(-i\vec{k} \cdot \vec{r}) d\vec{r}. \quad (1.19)$$

$$R_{ij}(\vec{r}) = \iiint_{-\infty}^{\infty} \Phi_{ij}(\vec{k}) \exp(i\vec{k} \cdot \vec{r}) d\vec{k}. \quad (1.20)$$

Функция ковариаций и тензор спектра скоростей связаны через преобразование Фурье. В связи с этим на функцию Φ накладываются некоторые ограничения связанные с определением функций ковариации. Так компоненты этого тензора с $i = j$ должны быть вещественными, то-есть $\Phi_{ii}(\vec{k}) = \Phi_{ii}^*(\vec{k})$. Также помимо этого в силу определения преобразования Фурье, компоненты тензора должны быть симметричными, то-есть $\Phi_{ij}(\vec{k}) = \Phi_{ji}^*(\vec{k}) = \Phi_{ji}(-\vec{k})$. Это также требует точной симметрии функции компонент тензора ковариаций, дабы удовлетворить сказанному выше условия вещественности его значений.

Энергетический спектр турбулентности выражается через диагональные компоненты тензора спектра скоростей следующей зависимостью:

$$E(k) = \oint \frac{1}{2} \Phi_{ii}(\vec{k}) d\Omega(k). \quad (1.21)$$

В данном случае переход от энергетического спектра к тензору не является явным и требует дополнительных предположений и условий. Переход от $\Phi_{ij}(\vec{k})$ можно осуществить численным интегрированием, либо при наличии функции $E(k) = F(\Phi_{ij}(\vec{k}))$ осуществив обратное разрешение относительно компонент тензора.

В конечном итоге, имеем алгоритм основными этапами которого является:

1. Построить тензор Φ_{ij} по заданной функции $E(k)$;
2. Перейти от тензора Φ_{ij} к тензору ковариаций R_{ij} ;
3. Построить матрицу ковариаций на основе полученного тензора спектра скоростей Φ_{ij} ;
4. Найти собственные значения и собственные вектора полученной матрицы ковариаций;
5. Построить поле скоростей базируясь на сгенерированной последовательности чисел и на полученных собственных значениях и векторах.

Наиболее длительным и затратным по вычислительным ресурсам и времени является процесс нахождения собственных чисел и векторов матрицы ковариаций. В среднем вычислительная сложность алгоритмов как итерационных, так и точных составляет $O(N^2)$. Если не использовать приемы для разряжения этой матрицы понадобится хранить $3 \cdot N^3$, где N – число узлов в сетке вычислительной области, не считая данных для сборки этой глобальной матрицы. Но стоит отметить, что сама процедура генерации является очень простой и происходит быстро, так как нам необходимо всего лишь просуммировать полученные значения. Если решить задачу один раз и сохранить полученные собственные вектора и числа можно достаточно эффективно генерировать турбулентные поля скоростей.

Для осуществления матричных преобразований используется библиотека "Armadillo" предоставляющая C++ API для работы с различными типами матриц, а также алгоритмов связанных с ними, например нахождение собственных чисел и значений для матриц [0]. Помимо этого, библиотека также имеет функционал

для генерации не только обычных последовательной случайных чисел, но также и коррелированных например по многомерному нормальному распределению.

Можно значительно ускорить описанный выше метод если использовать метод последовательного моделирования Гаусса. В общем, можно охарактеризовать метод последовательного моделирования, как разбиения рассматриваемой области на подобласти и проведение в ней стохастического моделирования. Основной сложностью является возникающие поверхности разделения на подобласти, необходимо учесть это, что бы не получать разрывы функций на этих поверхностях.

Как и в рассматриваемых не последовательных процедурах, в случае последовательного моделирования так же рассматривается набор точек \vec{p}_i , для которых известны значения v_i . Требуется рассчитать значение в некоторой точке \vec{q}_i при задаваемой пространственной ковариации $C(\vec{h})$ и среднем значении $\langle v \rangle$, в нашем случае, из физического смысла флуктуаций, задача упрощается, так как $\langle v' \rangle = 0$. Так же используя процедуру кригинга приходим к задаче:

$$\hat{A}\vec{x} = \vec{b} \quad (1.22)$$

$$A_{ij} = C(\vec{p}_i - \vec{q}_j) \quad (1.23)$$

$$b_i = C(\vec{p}_i - \vec{q}) \quad (1.24)$$

Как и в обычном подходе, СЛАУ также учитывает точное совпадение значений v_i в точках p_i . Из постановки выше следует что $E(u) = \sum_i x_i \cdot v_i + \langle v \rangle$, причём $\langle v \rangle$ - заданный параметр, и $Cov(v) = \sum_i x_i \cdot b_i + C(0)$, где $C(0)$ - также известный задаваемый параметр.

Как говорилось выше, вычислительная область подразделяется на подобласти. Так как расчёт ведётся на сетке, мы можем рассматривать не подобласть, а набор ближайших соседей узлов на сетке, тем самым, вводя новый параметр N - число ближайших соседей к рассматриваемой точке.

Так процедура последовательного моделирования может быть представлена в виде следующих шагов:

1. Выбирается случайная точка на сетке;
2. Ищется N ближайших к ней соседей с уже посчитанными значениями среднего и ковариации;
3. Строим матрицу $N \times N$ и решаем СЛАУ по методу описанному ранее методу, рассчитываем среднее и ковариацию в этой точке;

4. Зная среднее и ковариацию генерируем случайное число из Гауссова распределения $N(E(v), Cov(u))$;
5. Добавляем значение в массив известных.

Процедура может несколько отличаться для случаев если v – является вектором. В данном случае мы можем рассматривать два пути решения этой проблемы. Первый из них, не видоизменять алгоритм, работать с векторными и матричными величинами, в итоге среднее и ковариация являются вектором и матрицей соответственно. Окончательное значение для этих параметров находится из стохастического моделирования. Второй подход заключается в учёте каждой точки столько раз, сколько значений в векторе v , так например для поля флуктуаций u которого в каждой точке по 3 компоненты имеется по 3 одинаковые точки с различными значениями.