



ИПМ им.М.В.Келдыша РАН • Электронная библиотека

Препринты ИПМ • Препринт № 79 за 2011 г.



Миннихметов И.Р.
Стохастическое
моделирование условных
гауссовских процессов

Рекомендуемая форма библиографической ссылки: Миннихметов И.Р. Стохастическое моделирование условных гауссовских процессов // Препринты ИПМ им. М.В.Келдыша. 2011. № 79. 18 с. URL: <http://library.keldysh.ru/preprint.asp?id=2011-79>

РОССИЙСКАЯ АКАДЕМИЯ НАУК
ОРДЕНА ЛЕНИНА
ИНСТИТУТ ПРИКЛАДНОЙ МАТЕМАТИКИ
имени М. В. КЕЛДЫША

И. Р. Миннихметов

**Стохастическое моделирование
условных гауссовских процессов**

Москва 2011

И. Р. Миннихметов, Стохастическое моделирование условных гауссовских процессов

Аннотация. Стохастическое моделирование 2D и 3D условных гауссовских полей необходимо для моделирования углеводородных месторождений. Основные методы генерации реализаций гауссовских полей можно условно разделить на три группы: коррекция безусловных гауссовских полей посредством учета невязок на скважинах, последовательная гауссовская симуляция, разложение Холецкого матрицы ковариации. Однако все методы имеют существенные ограничения. Первая группа методов некорректно отражает заданную корреляционную функцию. При моделировании методом последовательной гауссовской симуляции используются ограничения на число соседей, что приводит к завышенной дисперсии. Третья группа методов, основанная на разложении Холецкого матрицы ковариации, ввиду большой вычислительной сложности алгоритма не применима для решения задач геологического моделирования. В данной работе предложен метод, основанный на спектральном представлении условного процесса. Используя преобразование Фурье матрицы ковариации, алгоритм разложения Холецкого может быть существенно упрощен. Предложенный алгоритм является достаточно быстрым и точным, что позволяет эффективно использовать его в задачах геологического моделирования углеводородных месторождений.

I.R. Minniakhmetov, Conditional simulation of Gaussian stochastic processes

Abstract. Generation of 2D and 3D normally distributed random fields conditioned on well data is often required in reservoir modeling. Such fields can be obtained by using three groups of methods: unconditional simulation with kriging interpolation (turning band or spectral methods), Sequential Gaussian Simulation (SGS) and Cholesky factorization of the covariance matrix. However, all these methods have limitations. First, it is known, that the second moment of the stochastic process conditionally simulated with the help of the kriging method is not identical to the target second moment (a priori known statistics). Second, SGS can't be calculated without limitation on the number of neighbors. As a result, SGS is only asymptotically exact. Third approach, which has the advantage of being general and exact, is to use a Cholesky factorization of the covariance matrix representing grid point's correlation. However, for the large fields the Cholesky factorization can be computationally expensive. In this work we present an alternative approach, based on the usage of spectral representation of a conditional process. It is shown that covariance of two arbitrary spectral components could be factorized into functions of corresponding harmonics. In this case the Cholesky decomposition could be considerably simplified. The advantage of the presented approach is its accuracy and computational simplicity.

Содержание

Введение	3
1. Постановка задачи	4
2. Построение условного стохастического процесса.....	8
3. Алгоритм моделирования условного гауссовского процесса.....	13
3.1 Одномерный случай	13
3.2 Многомерный случай.....	14
4. Оценка параметров условного распределения	15
Заключение.....	17
Литература.....	17

Введение

В настоящее время при решении задач оптимизации разработки углеводородных месторождений, т.е. при определении режимов заводнения, выбора методов увеличения нефтеотдачи, проведении гидроразрыва пласта и других геолого-технологических мероприятий, широко используется математическое моделирование. Оно может быть выполнено на основе постоянно действующих геолого-гидродинамических технологических моделей месторождений (ПДГТМ). Основные этапы построения ПДГТМ:

- Геологическое моделирование
- Ремасштабирование
- Гидродинамическое моделирование

Исходные данные для решения задачи геологического моделирования – это данные сейсмических измерений и результаты геофизических исследований на скважинах (ГИС: электрический, акустический, нейтронный, гамма и др. виды каротажа, результаты интерпретации ГИС, результаты опробования, исследования керна). Сейсмические данные позволяют определить характер напластования, выявить разломы и другие крупномасштабные структурные особенности, в то время как скважинная корреляция дает возможность выделить мелкомасштабные элементы (пласты и пропластки). Совокупность этих данных создает основу для решения задачи структурного моделирования, т.е. построения геометрии пласта с учетом разномасштабных структурных особенностей. В результате формируется сетка, обычно называемая геологической.

Как правило (С.Комаров, 1963), данные в каротажных кривых записаны с шагом, составляющим примерно 20 см, при этом размеры ячеек геологической сетки по вертикали могут варьироваться от 10см до нескольких метров (К.Закревский, 2009). Следовательно, для последующего распределения свойств среды необходимо осреднить скважинные кривые в ячейки геологической сетки. Ячейки, в которых определены свойства среды,

будем называть оцененными, при этом остальные ячейки сетки считаются неоцененными.

Одним из основных этапов геологического моделирования является построение пространственного распределения фильтрационно-емкостных свойств среды (пористость, проницаемость; далее ФЕС). Задача определения ФЕС в неоцененных ячейках – это задача статистического оценивания функций по данным, определенным в некотором наборе точек. При такой постановке задачи естественно использовать различные алгоритмы интерполяции (К. де Бор, 1985; M.Buhmann, M.Powell, 1900; А.Самарский, 1989, и др.). Однако, классические интерполяционные алгоритмы эффективны для достаточно гладких функций, в то время как реализация случайного стационарного процесса с вероятностью принадлежит пространству L2 (И.И. Гихман, А.В.Скорход, 1977; Д.А.Лаврик, И.Р.Миннихметов, А.Х.Пергамент, 2011). Поэтому многими авторами (N.Cressie, 1990; D.Krige, 1951; G.Matheron, 1963, и др.) был развит алгоритм kriging (Best Linear Unbiased Estimation). В неоцененных ячейках этот алгоритм определяет значения случайного процесса в виде линейной комбинации известных значений, имеющих минимальную дисперсию, при условии, что известна корреляционная функция. В результате полученные значения представляют собой условное мат. ожидание процесса, т.к. значения случайного процесса в оцененных точках являются известными.

Помимо задачи оценивания существует проблема построения реализаций случайного процесса, значения которого в оцененных точках совпадают с исходными данными. Такие случайные процессы можно назвать условными случайными процессами (А.В.Булинский, А.Н.Ширяев, 2005).

Данная работа посвящена исследованию задачи построения реализаций случайного процесса и разработке эффективных алгоритмов ее решения.

1. Постановка задачи

Пусть имеется случайный процесс $X(\vec{r})$, представляющий собой сумму тренда $a(\vec{r})$ и гауссовского стационарного случайного процесса $\xi(\vec{r})$ с нулевым средним и корреляционным оператором $B(\vec{r})$:

$$X(\vec{r}) = a(\vec{r}) + \xi(\vec{r}),$$

(1)

где $B(\vec{r} - \vec{r}') = E(\xi(\vec{r})\xi(\vec{r}')) = E(X(\vec{r}) - a(\vec{r}), X(\vec{r}') - a(\vec{r}'))$,

E - операция вычисления математического ожидания.

Также задана ортогональная сетка $\Omega_{\vec{h}}$ с постоянным шагом \vec{h} . Координаты произвольного узла сетки $\vec{r}_{i,j,k} = (x_i, y_j, z_k)$ выражаются следующим образом:

$$\begin{aligned}x_i &= x_0 + ih_x, i = 0..M_x - 1 \\y_j &= y_0 + jh_y, j = 0..M_y - 1 \\z_k &= z_0 + kh_z, k = 0..M_z - 1\end{aligned}$$

где x_0, y_0, z_0 – координаты первого узла сетки, M_x, M_y, M_z – число узлов в направлении x, y, z соответственно.

Значения тренда $a(\vec{r})$ известны в узлах сетки $\Omega_{\vec{h}}$. Также задан корреляционный оператор $B(\vec{r})$, который является положительно определенным и ядерным.

Как известно (Крамер-Лидбеттер, Ибрагимов-Розанов)

$$\begin{aligned}\xi(\vec{r}) &= \int e^{i\vec{k}\vec{r}} dy(\vec{k}) \\dy(\vec{k}) &= g(\vec{k})dW_{\vec{k}},\end{aligned}$$

где i – мнимая единица, $y(\vec{k})$ – процесс с ортогональными приращениями и конечным вторым моментом, $W_{\vec{k}}$ – гауссовский случайный процесс с независимыми приращениями (Броуновское движение), $g(\vec{k}) \in L_2$

$$\begin{aligned}E | dy(\vec{k}) dy^*(\vec{k}') | &= g(\vec{k}) g^*(\vec{k}') \delta(\vec{k} - \vec{k}') dk dk' \\B(\vec{r}) &= \int |g(\vec{k})|^2 e^{i\vec{k}\vec{r}} d^3\vec{k},\end{aligned}$$

где $\delta(\vec{k} - \vec{k}')$ – δ -функция Римана, $dy^*(\vec{k}'), g^*(\vec{k}')$ – комплексно-сопряженные величины.

Значения процесса $X(\vec{r})$ известны в произвольных m узлах сетки $\Omega_{\vec{h}}$:

$$\begin{aligned}X(\vec{r}_{n_1}) &= y_1, X(\vec{r}_{n_2}) = y_2, \dots X(\vec{r}_{n_m}) = y_m,\end{aligned}\tag{2}$$

где n_1, n_2, \dots, n_m – индексы узлов сетки, в которых заданы значения процесса $X(\vec{r})$.

Так как мы рассматриваем случайный процесс в конечном числе точек, то далее будем рассматривать вместо процесса (1) случайный вектор \vec{X} с коррелированными компонентами:

$$\begin{aligned}X_k &= X(\vec{r}_k) = a(\vec{r}_k) + \xi(\vec{r}_k) = a_k + \xi_k, k = 1..N\end{aligned}\tag{3}$$

$$C = E(\vec{X}\vec{X}^+),$$

где C – матрица ковариации вектора \vec{X} .

Необходимо построить реализации вектора \vec{X} , проходящие через заданные точки (2).

На сегодняшний день, основные методы генерации реализаций гауссовских полей можно условно разделить на три группы: коррекция безусловных гауссовских полей посредством учета невязок на скважинах, последовательная гауссовская симуляция, разложение Холецкого матрицы ковариации.

Существенным элементом первой группы методов является моделирование безусловных процессов (turning bands (G.Matheron, 1973), спектральный метод (A.Yaglom, 1962), и др.), т.е. при генерации реализаций не учитываются значения в ячейках, через которые проходят скважины. Для того чтобы реализации процесса удовлетворяли скважинным данным, определяются невязки Δ между значениями на скважинах и значениями полученной реализации $Z(\vec{r})$. Далее проводится интерполяция невязок методом ordinary kriging (X.Emery, 2005). Результирующий процесс представляется в виде суммы реализации безусловного процесса $Z(\vec{r})$ и функции невязки $\Delta(\vec{r})$:

$$\xi(\vec{r}) = \Delta(\vec{r}) + Z(\vec{r})$$

Однако, как известно (M.Shinozuka, R.Zhang, 1996), корреляционная функция полученного процесса не совпадает с исходной корреляционной функцией, что является существенным недостатком данных методов.

Следующий метод генерации, предложенный C.Deutsch, A.Journel (1997) - последовательная гауссовская симуляция (ПГС). Этот алгоритм является основным стохастическим методом моделирования ФЕС в современных геологических пакетах (Schlumberger «Petrel», ROXAR «Irap RMS»). В методе ПГС проводится последовательный розыгрыш гауссовских случайных величин во всех не оцененных ячейках. В каждой последующей точке значения случайного процесса – гауссовская случайная величина с математическим ожиданием и дисперсией, которые определяются методом kriging по всем предшествующим значениям (M.Shinozuka, R.Zhang, 1996). При этом приходится решать СЛАУ, размерность которой равна числу оцененных ячеек. В виду больших размеров исходной сетки (количество ячеек может достигать 10^9), на последних шагах моделирования метод требует огромных вычислительных ресурсов. В качестве упрощения, в коммерческих пакетах используют ограничение системы до заданного числа уравнений (обычно от 16 до 128). Как следствие, построенные реализации процесса имеют приближенную корреляционную функцию и завышенную дисперсию (А.Х.Пергамент, И.Р.Миннихметов, А.Р.Ахметсафина, 2009), т.е. определяются с большой погрешностью. Впоследствии, данные расхождения могут привести к некорректным значениям дебитов добычи углеводородов.

Еще один метод, который является наиболее точным и широко применяемым в генерации последовательностей случайных величин – метод, основанный на разложении Холецкого матрицы ковариации Σ случайного вектора \vec{X} , компоненты которого удовлетворяют условию (2). Как известно (K.Karhunen 1946, У.Гренандер, 1961), последовательность случайных величин с заданной корреляционной функцией можно представить в виде $\vec{x} = L\vec{\eta}$, где $\vec{\eta}$ – вектор из независимых стандартных (с единичной дисперсией и нулевым мат. ожиданием) нормальных случайных величин, а L – разложение Холецкого матрицы ковариации $\Sigma = LL^*$. Однако размерность задач,

решенных с помощью этого алгоритма (M.Naugh, 2004; J.Wang, C.Liu 2006; R.Halbleib, V.Voev 2011), намного меньше размерности задач геологического моделирования. В силу этого генерация пространственных распределений фильтрационно-емкостных параметров с помощью разложения Холецкого труднореализуема даже на современных машинах. Несомненно, существуют различные способы приближенного представления матрицы ковариации для ускорения разложения Холецкого, но в этом случае приходится искать компромисс между скоростью работы метода и точностью расчета.

На идеях сформулированных выше, основан один из современных методов генерации реализаций случайного процесса. Суть алгоритма заключается в моделировании Фурье-образов реализаций условного гауссовского процесса с помощью разложения Холецкого. Так как мы имеем дело с процессом в конечном числе точек, то далее под Фурье-преобразованием вектора \vec{x} будем понимать следующее:

$$\vec{x}^\omega = F\vec{x},$$

где \vec{x}^ω – Фурье-образ вектора \vec{x} , F – матрица дискретного преобразования Фурье (ДПФ):

$$F_{kl} = \exp \left[-2\pi i \left(\frac{(k^x - 1)(l^x - 1)}{M_x} + \frac{(k^y - 1)(l^y - 1)}{M_y} + \frac{(k^z - 1)(l^z - 1)}{M_z} \right) \right]$$

где k, l – порядковые номера произвольных узлов сетки $\Omega_{\vec{h}}$, $k^x, k^y, k^z, l^x, l^y, l^z$ – индексы узлов k, l соответственно.

Таким образом, построение случайного вектора \vec{x} сводится к генерации и обратному Фурье-преобразованию вектора \vec{x}^ω . Согласно результатам (M.Naugh, 2004) для этого необходимо иметь представление матрицы ковариации в Фурье пространстве Σ^ω :

$$\Sigma^\omega = F \Sigma F^+,$$

где Σ^ω – Фурье-образ матрицы Σ , F^+ – эрмитово сопряжение матрицы F .

Нетрудно показать, что матрица Σ^ω эрмитова и положительно определенная:

$$\Sigma^\omega = F \Sigma F^+ = F L L^+ F^+ = F L (F L)^+ = L^\omega L^{\omega+}$$

$$\Sigma^{\omega*} = (L^\omega L^{\omega+})^+ = L^\omega L^{\omega+} = \Sigma^\omega,$$

(4)

Следовательно, Фурье-образ матрицы ковариации Σ можно разложить с помощью алгоритма Холецкого (D.Vau III, L.Trefethen, 1997) и, тем самым, построить Фурье-образ реализации условного гауссовского процесса:

$$\vec{x}^\omega = L^\omega \vec{\eta},$$

где $\vec{\eta}$ – вектор из независимых стандартных нормально распределенных величин.

Особенность данного подхода в том, что для процесса, заданного на равномерной сетке $\Omega_{\vec{r}}$ матрица Σ^ω представима в виде:

$$(5) \quad \Sigma^\omega = D - SC_m^{-1}S^+,$$

где D – диагональная матрица, отвечающая за безусловное распределение, C_m^{-1} – обратная матрица ковариации случайных величин, заданных в точках (2), S – Фурье-образ матрицы ковариации вектора \vec{X} и случайных величин, заданных в точках (2).

Используя представление (5), алгоритм разложения Холецкого для матрицы Σ^ω может быть существенно упрощен. Благодаря этому вычислительная сложность алгоритма значительно уменьшается, что позволяет эффективно строить реализации условного гауссовского процесса без потери точности.

Также в данной статье представлены результаты моделирования реализаций условного гауссовского процесса для одномерных и двумерных задач, оценки корреляционной функции и среднего значения полученных реализаций, а также исследована область применимости алгоритма.

2. Построение условного стохастического процесса

Проведем бесконечное множество реализаций стационарного случайного процесса $\xi(\vec{r})$, проходящего через заданные точки. В результате мы получим пучок кривых. В связи с имеющейся корреляцией между значениями, условное математическое ожидание и условная дисперсия в произвольной точке \vec{r}_0 будет зависеть от расстояния до исходных данных:

$$\begin{aligned} E(\xi_0 | \xi_1 = y_1, \xi_2 = y_2, \dots, \xi_n = y_n) &\rightarrow a, R_{k0} \rightarrow \infty, k = 1..n \\ D(\xi_0 | \xi_1 = y_1, \xi_2 = y_2, \dots, \xi_n = y_n) &\rightarrow B(0) = \sigma^2, R_{k0} \rightarrow \infty, k = 1..n \\ E(\xi_0 | \xi_1 = y_1, \xi_2 = y_2, \dots, \xi_n = y_n) &\rightarrow y_k, R_{k0} \rightarrow 0 \\ D(\xi_0 | \xi_1 = y_1, \xi_2 = y_2, \dots, \xi_n = y_n) &\rightarrow 0, R_{k0} \rightarrow 0 \end{aligned}$$

где $\xi_k = \xi(\vec{r}_k)$, R_{k0} – расстояние от k -ого узла до точки \vec{r}_0 .

Т.е. при удалении от оцененных точек, сечение процесса имеет безусловное распределение, при приближении к точкам данных сечение процесса стремится к заданному значению (Рис.1).

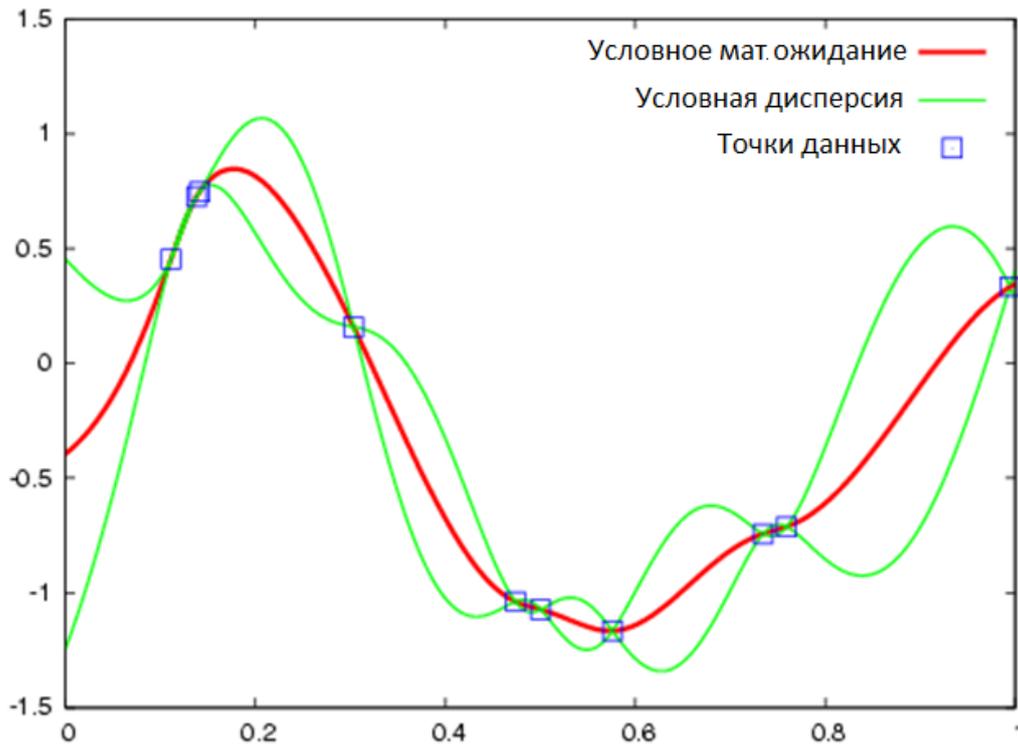


Рис.1. Условные мат. ожидание и дисперсия стационарного процесса.

Используя матричные преобразования и теорию случайных процессов можно показать справедливость следующей теоремы (Anderson, 1984):

Теорема 1. Пусть компоненты вещественного случайного гауссовского вектора \vec{Z} разделены на две группы, составляющие векторы \vec{Z}_1 и \vec{Z}_2 . При этом:

$$E(\vec{Z}) = E \begin{pmatrix} \vec{Z}_1 \\ \vec{Z}_2 \end{pmatrix} = \begin{pmatrix} \vec{\mu}_1 \\ \vec{\mu}_2 \end{pmatrix}$$

$$Cov(\vec{Z}) = \begin{pmatrix} \Sigma_{11} & \Sigma_{12} \\ \Sigma_{21} & \Sigma_{22} \end{pmatrix},$$

где Σ_{11} — матрица ковариации компонент вектора \vec{Z}_1 , Σ_{22} — матрица ковариации компонент вектора \vec{Z}_2 , Σ_{12} — матрица ковариации компонент вектора \vec{Z}_1 и \vec{Z}_2 . Тогда, если компоненты \vec{Z} имеют гауссовское распределение, то условное распределение \vec{Z}_1 при условии $\vec{Z}_2 = \vec{z}_2$, где \vec{z}_2 — реализация вектора \vec{Z}_2 , также является гауссовским с мат. ожиданием $\mu_1 + \Sigma_{12} \Sigma_{22}^{-1} (\vec{z}_2 - \vec{\mu}_2)$ и матрицей ковариации $\Sigma_{11} + \Sigma_{12} \Sigma_{22}^{-1} \Sigma_{21}$.

Из этой теоремы непосредственно следует:

Следствие 1. Пусть \vec{X} – случайный гауссовский вектор, с нулевым мат. ожиданием и матрицей ковариации C . Положим n_1, n_2, \dots, n_m – индексы компонент вектора \vec{X} с заданными значениями $X_{n_j} = y_j, j=1..m$. Тогда параметры условного распределения для произвольных компонент вектора \vec{X} представимы в виде:

$$E(X_k | X_{n_1} = y_1, X_{n_2} = y_2, \dots, X_{n_m} = y_m) = \vec{b}_k^T C_m^{-1} \vec{y}$$

$$\Sigma_{kl} = \text{Cov}(X_k X_l | X_{n_1} = y_1, X_{n_2} = y_2, \dots, X_{n_m} = y_m) = C_{kl} - \vec{b}_k^T C_m^{-1} \vec{b}_l$$

(6)

$$k = 1..N, l = 1..N$$

Здесь Σ – матрица ковариации вектора \vec{X} ,

C_m – матрица ковариации n_1, n_2, \dots, n_m компонент вектора \vec{X} :

$$C_m = \begin{pmatrix} C_{n_1 n_1} & C_{n_1 n_2} & \dots & C_{n_1 n_m} \\ C_{n_2 n_1} & C_{n_2 n_2} & \dots & C_{n_2 n_m} \\ \vdots & \vdots & \ddots & \vdots \\ C_{n_m n_1} & C_{n_m n_2} & \dots & C_{n_m n_m} \end{pmatrix}$$

$$\vec{b}_k = \begin{pmatrix} C_{kn_1} \\ C_{kn_2} \\ \vdots \\ C_{kn_m} \end{pmatrix}$$

Как было упомянуто выше, реализации условного процесса будут строиться с помощью преобразования Фурье. Известно (R.Lidl, G.Pilz., 1999), что ДПФ рассматривает периодические сигналы. Первоначально рассмотрим одномерный случай. Обобщение на многомерное пространство будет сделано в дальнейшем. Продолжение ковариационной матрицы C должно быть таким, чтобы каждая компонента вектора \vec{X} имела симметричную ковариационную функцию:

$$C'_{k,1} = \begin{cases} C_{k,1}, k \leq N \\ C'_{2N-k+2,1}, k > N \end{cases}$$

$$C'_{kl} = \begin{cases} C'_{l-k+1,1}, k \leq l \\ C'_{2N+(l-k)+1,1}, k > l \end{cases}$$

(7)

$$k = 1..2N, l = 1..2N$$

Для наглядности ниже показаны продолжения нескольких столбцов матрицы C .

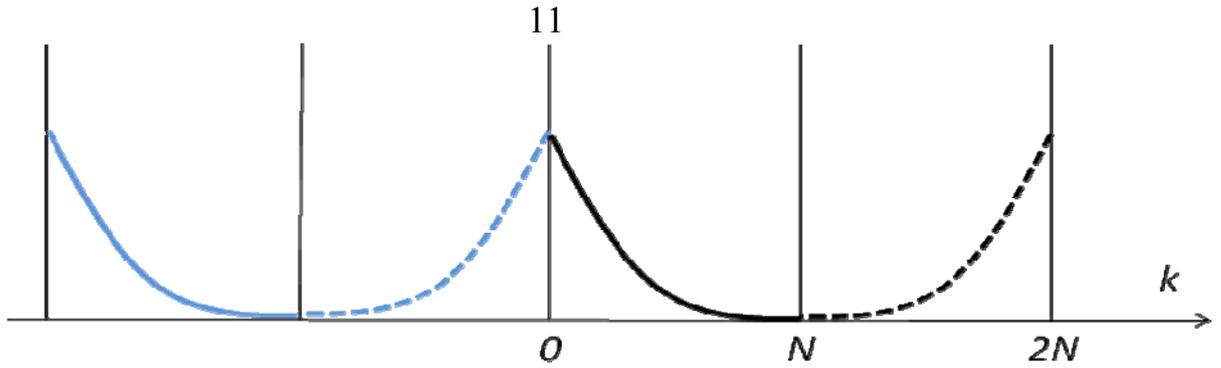


Рис.2.Первый столбец матрицы C' .

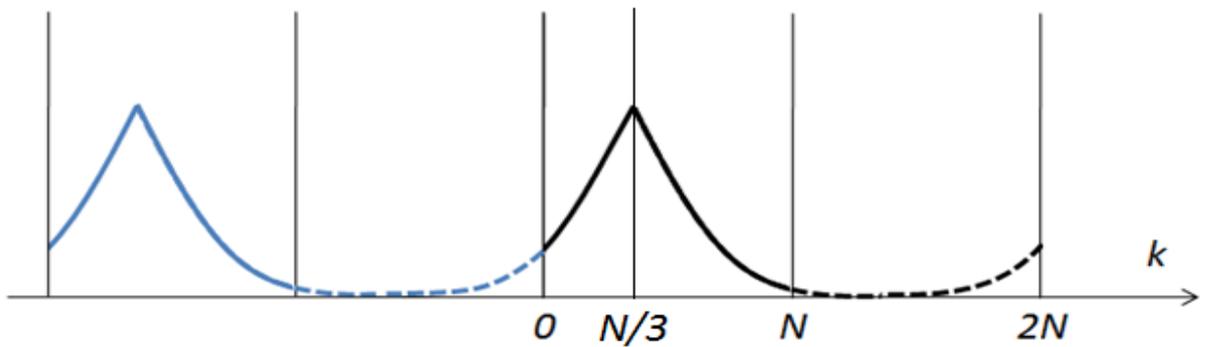


Рис.3.Столбец $N/3$ матрицы C' .

Если продолжение матрицы C обладает вышеупомянутыми свойствами, то справедлива лемма:

Лемма 1. Фурье-образ матрицы C' представим в виде:

$$FC'F^+ = 2N \text{diag}\{FC^{(1)}\},$$

где $C^{(1)}$ – первый столбец матрицы C' , $FC^{(1)}$ – Фурье-образ первого столбца матрицы C' , $2N$ – число компонент вектора $FC^{(1)}$, $\text{diag}\{FC^{(1)}\}$ – диагональная матрица, элементами которой являются компоненты вектора $FC^{(1)}$.

Доказательство:

► Матрица ДПФ имеет вид:

$$F = \begin{pmatrix} 1 & 1 & 1 & \dots & 1 \\ 1 & e^{i\frac{2\pi}{2N}} & e^{2i\frac{2\pi}{2N}} & \dots & e^{(N-1)i\frac{2\pi}{2N}} \\ 1 & e^{2i\frac{2\pi}{2N}} & e^{4i\frac{2\pi}{2N}} & \dots & e^{2(N-1)i\frac{2\pi}{2N}} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 1 & e^{(2N-1)i\frac{2\pi}{2N}} & e^{2(2N-1)i\frac{2\pi}{2N}} & \dots & e^{(2N-1)(2N-1)i\frac{2\pi}{2N}} \end{pmatrix}$$

Так как из соотношения (7) следует, что l -ый столбец матрицы C' есть не что иное, как сдвиг первого столбца на $(l-1)$ элементов, то из свойств ДПФ имеем:

$$(FC^{(l)})_k = (FC^{(1)})_k e^{\frac{2\pi i(k-1)(l-1)}{(2N)}}$$

$$k = 1..2N, l = 1..2N.$$

Отсюда

$$FC'F^+ = \begin{pmatrix} (FC^{(1)})_1 & 0 & \dots & 0 \\ 0 & (FC^{(1)})_2 & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & (FC^{(1)})_{2N} \end{pmatrix} \begin{pmatrix} 1 & 1 & \dots & 1 \\ 1 & e^{\frac{2\pi i}{(2N)}} & \dots & e^{\frac{2\pi i(2N-1)}{(2N)}} \\ \vdots & \vdots & \ddots & \vdots \\ 1 & e^{\frac{2\pi i(2N-1)}{(2N)}} & \dots & e^{\frac{2\pi i(2N-1)(2N-1)}{(2N)}} \end{pmatrix} F^+$$

$$FC'F^+ = \text{diag}\{FC^{(1)}\}FF^+ = \text{diag}\{FC^{(1)}\}F(2NF^{-1}) = 2N\text{diag}\{FC^{(1)}\}$$

Лемма доказана. ◀

Используя следствие 1 и лемму 1, докажем следующую теорему:

Теорема 2. Пусть \vec{X} – случайный гауссовский вектор с одинаково распределенными компонентами, с нулевым мат. ожиданием и матрицей ковариации C . Положим n_1, n_2, \dots, n_m – индексы компонент вектора \vec{X} с заданными значениями $X_{n_j} = y_j, j = 1..m$. Тогда если матрица Σ' – матрица ковариации (6) для продолжения \vec{X}' , то ее Фурье-образ Σ'^{ω} представим в виде:

$$\Sigma'^{\omega} = D(2ND^{-1} - WC_m^{-1}W^+)D,$$

(8)

где $D = \text{diag}\{FC^{(1)}\}$

$$W = \begin{pmatrix} e^{\frac{2\pi i 0(n_1-1)}{2N}} & e^{\frac{2\pi i 0(n_2-1)}{2N}} & \dots & e^{\frac{2\pi i 0(n_m-1)}{2N}} \\ e^{\frac{2\pi i 1(n_1-1)}{2N}} & e^{\frac{2\pi i 1(n_2-1)}{2N}} & \dots & e^{\frac{2\pi i 1(n_m-1)}{2N}} \\ \vdots & \vdots & \ddots & \vdots \\ e^{\frac{2\pi i (2N-1)(n_1-1)}{2N}} & e^{\frac{2\pi i (2N-1)(n_2-1)}{2N}} & \dots & e^{\frac{2\pi i (2N-1)(n_m-1)}{2N}} \end{pmatrix}$$

(9)

Доказательство.

► Используя результаты следствия 1, построим Фурье-образ матрицы

Σ' :

$$\Sigma'^{\omega} = F(C' - B'C_m^{-1}B^+)F^+,$$

$$\text{где } B' = \begin{pmatrix} C'_{1n_1} & C'_{1n_2} & \dots & C'_{1n_m} \\ C'_{2n_1} & C'_{2n_2} & \dots & C'_{2n_m} \\ \vdots & \vdots & \ddots & \vdots \\ C'_{2Nn_1} & C'_{2Nn_2} & \dots & C'_{2Nn_m} \end{pmatrix}$$

Из леммы 1

$$\Sigma'^{\omega} = 2ND - FB'C_m^{-1}B^+F^+$$

Далее из свойств ДПФ и соотношения (7):

$(FB^{(k)})_j = FC^{(1)} e^{\frac{2\pi i(n_k-1)(l-1)}{(2N)}}$, отсюда

$$FB' = \text{diag}\{FC^{(1)}\}W$$

В результате получаем:

$$\Sigma^{\omega} = 2ND - DWC_m^{-1}W^+D^+ = D(2ND^{-1} - WC_m^{-1}W^+)D$$

Теорема доказана. ◀

3. Алгоритм моделирования условного гауссовского процесса

3.1 Одномерный случай

Не ограничивая общности, положим $\vec{a} \equiv \vec{0}$. Тогда для случайного вектора (3) справедливы результаты теоремы 2. Очевидно, для разложения Холецкого матрицы Σ^{ω} достаточно получить разложение матрицы $G = 2ND^{-1} - WC_m^{-1}W^+$. Тогда матрица Σ^{ω} может быть представлена в виде:

$$\Sigma^{\omega} = D(2ND^{-1} - WC_m^{-1}W^+)D = DL_G L_G^+ D^+ = (DL_G)(DL_G)^+,$$

где L_G – разложение Холецкого матрицы $G = L_G L_G^+$.

Реализация Фурье-образа случайного вектора \vec{X}' имеет вид:

$$\vec{x}^{\omega} = F\vec{X}' = (DL_G)\vec{\eta}, \quad (10)$$

где $\vec{\eta}$ – случайный вектор с независимыми стандартными нормально распределенными компонентами.

Методом математической индукции можно показать, что алгоритм разложения Холецкого матрицы G сводится к простому алгоритму:

1. Пусть $A = C_m^{-1}$

2. Для всех $j = 1..2N$

а. Положим $\vec{w}_j^+ = \left(e^{\frac{2\pi i 0(n_1-1)}{2N}} \quad e^{\frac{2\pi i 1(n_2-1)}{2N}} \quad \dots \quad e^{\frac{2\pi i (2N-1)(n_m-1)}{2N}} \right)$ и $\vec{c} = (FC^{(1)})_j$

б. Определим вектор $\vec{g}_j = \frac{A^j \vec{w}_j}{L_{jj}}$

в. Коэффициенты разложения Холецкого матрицы L_G :

$$(L_G)_{jj} = \sqrt{\frac{2N}{c_j} - \vec{w}_j^+ A^j \vec{w}_j} \quad (11)$$

$$(L_G)_{ij} = -\vec{w}_i^+ \vec{g}_j, i < j$$

г. Для следующего столбца матрица A преобразуется к виду

$$A^{j+1} = A^j + \vec{g}_j \vec{g}_j^+$$

Далее, учитывая (10), для Фурье-образа x^{ω} получаем:

$$x^{\omega}_j = c_j \left(\sum_{k=1}^{j-1} L_{jk} \eta_k + L_{jj} \eta_j \right) = c_j \left(-\sum_{k=1}^{j-1} \vec{w}_j \vec{g}_k \eta_k + L_{jj} \eta_j \right) = c_j \left(-\vec{w}_j^+ \vec{\theta}_j + L_{jj} \eta_j \right), \quad (12)$$

где $\vec{\theta}_{j+1} = \vec{\theta}_j + \vec{g}_j \eta_j$

Используя обратное преобразование Фурье, получаем реализацию случайного вектора \vec{X}' :

$$\vec{x}' = F^{-1} \vec{x}'^{\omega}$$

Интересующий вектор \vec{x} находится сужением вектора \vec{x}' на область сетки $\Omega_{\vec{h}}$:

$$x_i = x'_i, i = 1..N.$$

Различные реализации случайного вектора \vec{X} представлены ниже.

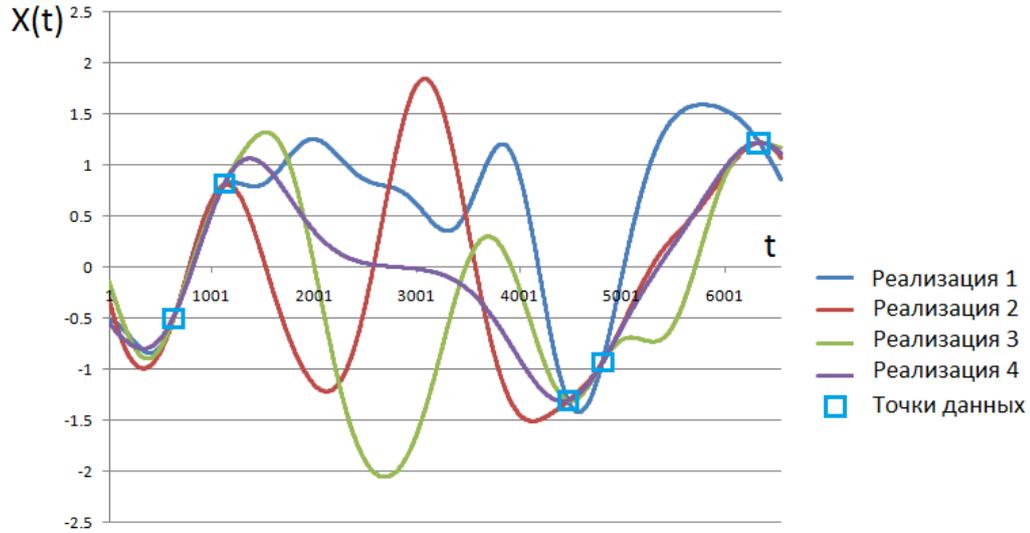


Рис.4. Реализации условного гауссовского процесса $X(\vec{r})$ в одномерном случае.

3.2 Многомерный случай

Обобщим результаты, полученные выше, на многомерный случай. Пусть k, l – порядковые номера произвольных узлов сетки $\Omega_{\vec{h}}$, $k^x, k^y, k^z, l^x, l^y, l^z$ – индексы узлов k, l соответственно. Тогда выражение (7) примет вид:

$$C'_{k,l} = C_{k',l'}$$

$$C'_{k,l} = C_{p+1,l}$$

$$k^{s'} = \begin{cases} k^s, k^s \leq M_s \\ 2M_s - k^s + 2, k^s > M_s \end{cases}$$

$$p^s = \begin{cases} l^s - k^s, k^s \leq l^s \\ 2M_s + l^s - k^s, k^s > l^s \end{cases}$$

$$s = x, y, z$$

При этом матрица W преобразуется к виду:

$$W_{kl} = \exp \left[-2\pi i \left(\frac{(k^x - 1)(n_l^x - 1)}{2M_x} + \frac{(k^y - 1)(n_l^y - 1)}{2M_y} + \frac{(k^z - 1)(n_l^z - 1)}{2M_z} \right) \right]$$

$$k_s = 1..2M_s, l = 1..m$$

$$s = x, y, z$$

Ниже представлены результаты моделирования в двумерном и трехмерном случае.

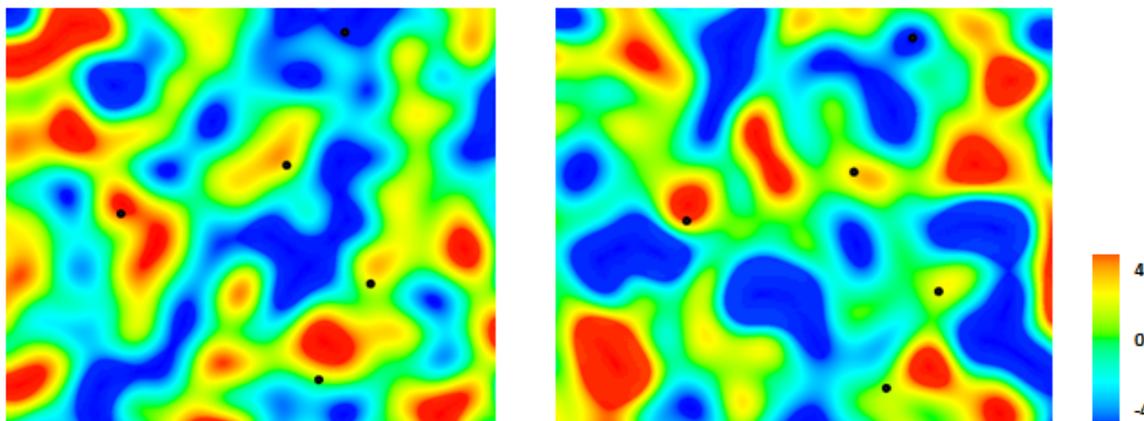


Рис.5. Две реализации условного гауссовского процесса $X(\vec{r})$ в двумерном случае. Точками показаны исходные данные.

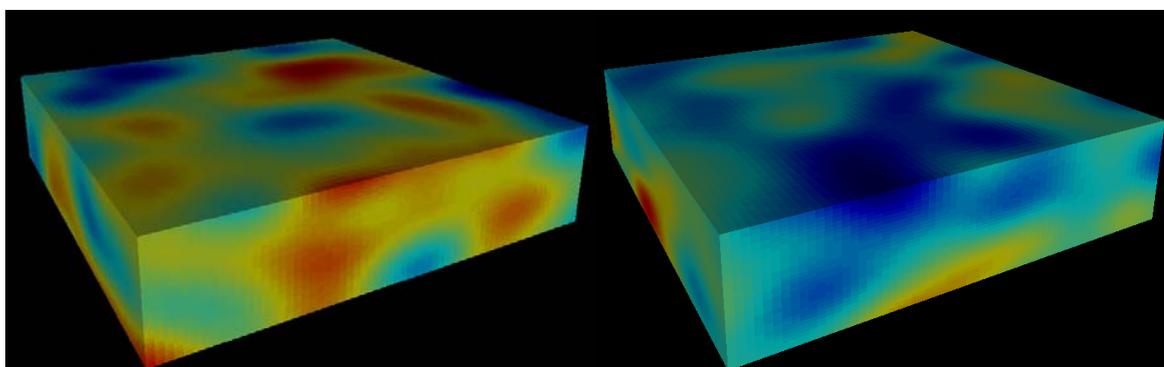


Рис.6. Две реализации условного гауссовского процесса $X(\vec{r})$ в трехмерном случае.

Однако, несмотря на быстрдействие, данный алгоритм обладает некоторыми недостатками:

1) Для корректной работы дискретного преобразования Фурье, необходимо, чтобы характерные масштабы изменения корреляционного оператора $B(\vec{r})$ были существенно меньше области моделирования.

2) Матрица ковариаций известных значений для гауссовской корреляционной функции $B(\vec{r}) = \sigma^2 e^{-r^2/2R^2}$ является плохообусловленной, в случае, когда расстояние между данными много меньше радиуса корреляции R .

3) Для дискретного преобразования Фурье требуется продолжение вектора \vec{X}' , что требует выделения в два раза большей памяти, чем при моделировании методом ПГС.

4. Оценка параметров условного распределения

Для оценки качества моделируемого процесса проведем оценку условного мат. ожидания, условной дисперсии и ковариационной функции процесса в произвольной точке, находящейся в зоне влияния точек исходных данных. Не ограничивая общности, положим $a(\vec{r}) \equiv \vec{0}$. Оценку параметров распределения будем проводить согласно формулам:

$$\hat{\mu}_i = \hat{E}(X_i) = \frac{1}{N_r} \sum_{k=1}^{N_r} x_i^k$$

$$\hat{\sigma}_i^2 = \hat{D}(X_i) = \frac{1}{N_r} \sum_{k=1}^{N_r} [x_i^k]^2, \quad ,$$

$$\hat{\Sigma}_{ij} = \hat{Cov}(X_i, X_j) = \frac{1}{N_r} \sum_{k=1}^{N_r} x_i^k x_j^k$$

(13)

где x_i^k – значение k -ой реализации i -ой компоненты вектора \vec{X} , N_r – число реализаций.

Оценки (13) для компоненты $N/2$ случайного вектора \vec{X} представлены ниже.

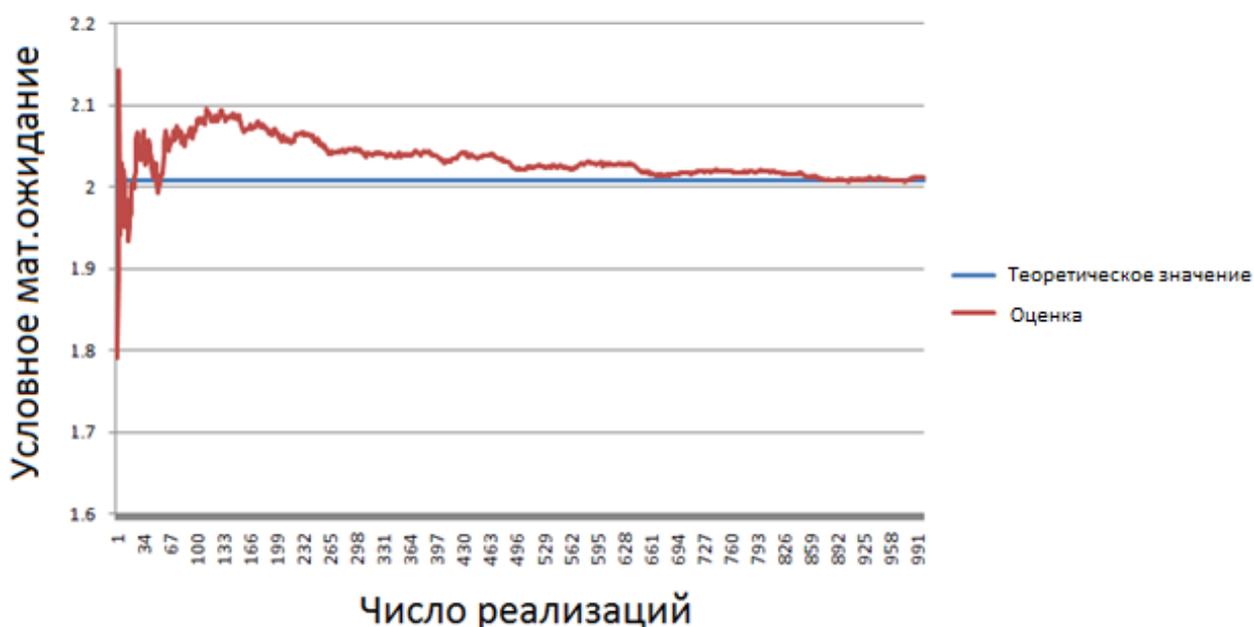


Рис.7. Оценка условного мат ожидания в зависимости от числа реализаций



Рис.8. Оценка условной дисперсии в зависимости от числа реализаций

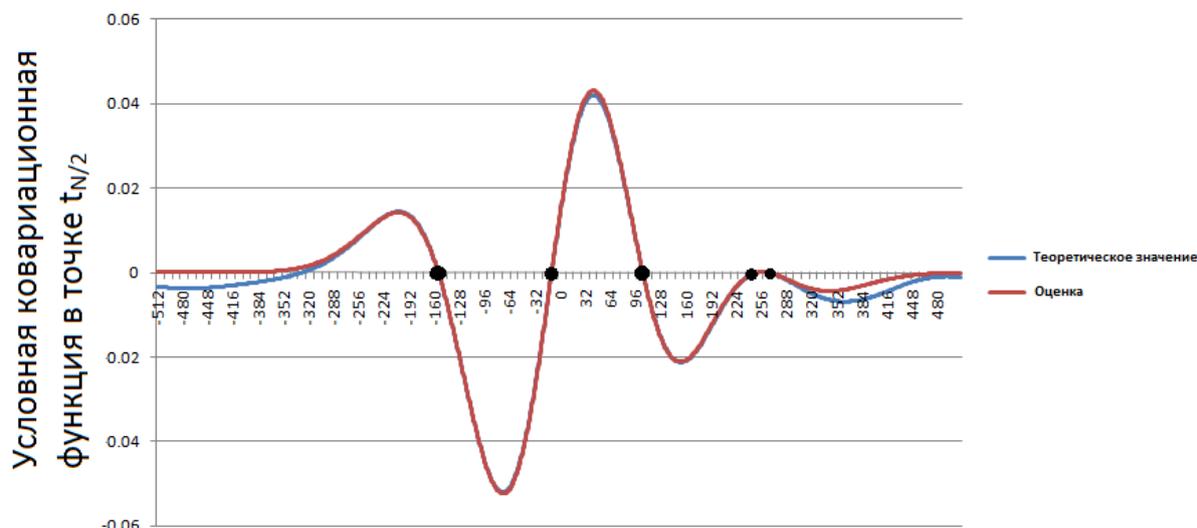


Рис.9. Оценка ковариационной функции для 2000 реализаций для точки $t_{N/2}$. Точками показаны координаты исходных данных.

Заключение

В отличие от общеизвестных методов генерации условных гауссовских процессов (ПГС, разложение Холецкого), предложенный в данной работе алгоритм обладает сложностью $O(N^2+m^3+Nm^2)$ (в случае использования быстрого преобразования Фурье $O(N\log(N)+m^3+Nm^2)$), что является приемлемым для современного уровня вычислительной техники.

В целом, алгоритм является достаточно быстрым и точным, что позволяет эффективно использовать его в задачах геологического моделирования как альтернативу ПГС. Благодаря быстродействию, а также корректной передаче корреляционных зависимостей, алгоритм будет незаменим и в других областях применения стационарных случайных процессов: моделирование процессов в области радиоэлектроники, финансов, теории игр и т.д.

Дальнейшие исследования предложенного метода должны быть направлены на ускорение и оптимизацию алгоритма, моделирование процессов со стационарными приращениями и фрактальных процессов.

Литература

1. T. Anderson, An Introduction to Multivariate Statistical Analysis, 2nd ed.: Wiley, New York, 1984;
2. D. Bau III, L. Trefethen, Numerical linear algebra, Philadelphia: Society for Industrial and Applied Mathematics, pp. 172-180, 1997
3. M. Buhman, M. Powell, Radial basis function interpolation on an infinite regular grid, in Algorithms for approximation II (Mason & Cox, eds.), Chapman & Hall, London, 1990. Hall, London, 1990, 146-169;
4. N. Cressie, The Origins of Kriging, Mathematical Geology, v. 22, pp 239–252, 1990;

5. X. Emery, Simple and Ordinary Kriging Multigaussian Kriging for Estimating recoverable Reserves, *Mathematical Geology*, v. 37, n. 3, pp. 295-319, 2005;
6. G. Fenton, Error Evaluation of Three Random Field Generators *J. Eng. Mech.* 120, 1994;
7. R. Halbleib, V. Voev, Forecasting Multivariate Volatility using the VARFIMA Model on Realized Covariance Cholesky Factors, *Journal of Economics and Statistics*, vol. 231(1), pp.134-152, 2011;
8. M. Haugh, The Monte Carlo framework, examples from finance and generating correlated random variables, IEOR E4703: Monte Carlo Simulation Course Notes, 2004;
9. A. Journel, C. Huijbregts, *Mining Geostatistics*. Academic Press, New York. 1978.
10. D. Krige, A statistical approach to some mine valuations and allied problems at the Witwatersrand, Master's thesis of the University of Witwatersrand, 1951;
11. R. Lidl, G. Pilz. *Applied Abstract Algebra*, 2nd edition. Wiley, pp. 217-219, 1999;
12. G. Matheron, Principles of geostatistics, *Economic Geology*, 58, pp. 1246–1266, 1963;
13. G. Matheron, The intrinsic random functions and their applications. *Adv. Appl. Probability*, n.5, pp. 439-468, 1973;
14. M. Shinozuka, R. Zhang, Equivalence Between Kriging And Cpdf Methods For Conditional Simulation *J. Eng. Mech.* 122, 1996;
15. D. Terzopoulos. The computation of visible-surface representations. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, PAMI-10(4): pp. 417– 438, 1988;
16. J. Wang, C. Liu, Generating multivariate mixture of normal distributions using a modified Cholesky decomposition, *Winter Simulation Conference*, p. 342, 2006;
17. A. Yaglom, *An Introduction to the Theory of Stationary Random Functions*, Englewood Cliffs, NJ: Prentice-Hall, 1962;
18. А. В. Булинский, А.Н.Ширяев, *Теория случайных процессов*, М.: ФИЗМАТЛИТ, 408 с, 2005;
19. И. И. Гихман, А.В.Скорород, *Введение в теорию случайных процессов*, М.: Наука, 1977;
20. К. Е. Закревский, *Геологическое 3D моделирование*. М.:ООО "ИПЦ Маска", 2009;
21. С. Г. Комаров, *Геофизические методы исследования скважин*, М.: Недра, 1963;
22. Д. А. Лаврик, И. Р. Миннихметов, А. Х. Пергамент, Регуляризованные алгоритмы статистического оценивания функций в задачах геологического моделирования, *М.:Матем. моделирование*, 23:4,23–40с,2011;