



ОЦЕНИВАНИЕ ПАРАМЕТРОВ СЕЙСМИЧЕСКИХ ВОЛН С ПРИМЕНЕНИЕМ ГЕНЕТИЧЕСКОГО АЛГОРИТМА

Р.Р. Султангалеев, В.Н. Троян

*Санкт-Петербургский государственный университет, 199034, Санкт-Петербург,
Университетская наб., 7/9, Россия, e-mail: ruslan.sultangaleev@gmail.com, v.troyan@spbu.ru*

Проблема быстрого и эффективного определения параметров сейсмических волн привлекает к себе внимание геофизиков-сейсмологов, поскольку ее решение позволяет повысить разрешенность сейсмических данных и увеличить скорость определения параметров волны. В данной статье в качестве параметра сейсмической волны используется угловой коэффициент в линейно-изменяющейся скорости. Работа включает в себя использование генетического алгоритма для оценки параметра в градиентной среде. Генетический алгоритм (ГА) – это способ решения глобальных нелинейных задач оптимизации. В основе лежит использование эволюционных принципов для поиска оптимального решения: кодирование, отбор, скрещивание, мутация и выбор. При помощи конечно-разностного метода были получены синтетические сейсмограммы. Полученные результаты показали высокую эффективность в оценке скорости распространения сейсмических волн в градиентных средах. ГА можно успешно использовать для оценивания скорости распространения сейсмических волн в градиентных средах.

Генетический алгоритм, синтетическая сейсмограмма, отношение сигнал/помеха, целевая функция

ESTIMATION OF SEISMIC WAVES PARAMETERS WITH APPLICATION OF A GENETIC ALGORITHM

R.R. Sultangaleev, V.N. Troyan

*Saint Petersburg State University, 7/9 Universitetskaya nab., St. Petersburg, 199034, Russia,
e-mail: ruslan.sultangaleev@gmail.com, v.troyan@spbu.ru*

The problem of fast and effective determination of seismic waves parameters draws to itself attention of geophysicists as the decision of problem allows to increase resolution of seismic data and to increase the speed of determination wave parameters. In this article as the parameter of a seismic wave the slope in linearly – the changing speed is used. Work includes use of a genetic algorithm for parameter assessment in the gradient environment. The Genetic Algorithm (GA) is a way of the solution of global nonlinear problems of optimization. Use of the evolutionary principles for search of an optimal solution is the cornerstone: coding, selection, crossing, mutation and choice. By means of a finite and differential method synthetic seismographic records have been received. The received results have shown high efficiency in assessment of speed of distribution of seismic waves in gradient environments. GA can be used successfully for estimation of speed of distribution of seismic waves in gradient environments.

ВВЕДЕНИЕ

Генетический алгоритм (ГА) – это способ решения глобальных нелинейных задач оптимизации [Holland, 1975; Reeves, Rowe, 2002; Troyan, Kiselev, 2010]. Особенность его в том, что он допускает использование любого типа объектного функционала без требования его гладкости. Алгоритм способен комбинировать локальный и глобальный поисковые механизмы в единый эффективный метод. Интерес к этому методу неуклонно растет, он находит все новые применения в различных областях физики, математики и других естественных наук, где требуется решать задачи оптимизации. Так как большинство обратных задач являются задачами оптимизации, то можно ожидать, что генетический подход найдет применение для решения многих геофизических проблем, в том числе для определения скорости распространения сейсмической волны в различных средах. Алгоритм нашел применение в ряде геофизических задач, таких как, например, нелинейное обращение морских рефрагированных сейсмических данных для построения одномерного скоростного разреза [Sambridge, Drijkoningen, 1992]. Другое применение алгоритм нашел при решении двумерной обратной задачи магнитотеллурического зондирования [Everett, Schultz, 1993]. Кроме того, в работе [Wilson, Vasudevan, 1991] осуждаются вопросы применения генетического алгоритма к задаче оценивания остаточной статистики при обработке сейсмических данных. Для нашего исследования мы рассмотрели градиентную среду, и в этом случае основным параметром является скорость распространения волны и её изменение с глубиной. Основные этапы работы алгоритма:

Кодирование → отбор → скрещивание → мутация → выбор.

Отбор наилучших моделей путем скрещивания и мутации производится при генетическом выборе наилучшей популяции с точки зрения выбранного критерия, поэтому алгоритм получил название генетического [Holland, 1975].

ОПИСАНИЕ АЛГОРИТМА И ЧИСЛЕННОЕ МОДЕЛИРОВАНИЕ

Рассматривается следующая скоростная модель среды: скорость меняется по линейному закону до границы раздела сред ($V_0 + \alpha Z$), где происходит скачок скорости, далее скорость остается постоянной. Первое слагаемое в формуле обозначает скорость распространения источника по дневной поверхности, т. е. по земной поверхности в данном случае. Второе слагаемое есть произведение углового коэффициента и глубины распространения в среде. С помощью алгоритма оценивается параметр α . При построении модели мы принимаем в качестве истинного значения параметра α , равное 0.3. Общей чертой всех ГА является двоичная кодировка параметров модели. Задается интервал изменения параметра α от 0 до 1.4 с шагом 0.05. На этапе кодирования алгоритм определяет пригодность модели, т. е. синтетические данные сравниваются с экспериментальными при помощи целевой функции. При выборе функции наряду с физическим обоснованием необходимо учитывать вычислительную сложность и практическую реализуемость выбранного подхода. В работе используется среднеквадратическая функция ошибок:

$$F(\alpha) = \sum_{j=1}^K \sum_{i=1}^n \sum_{i=1}^n (u_{ij0} - u_{ij}(\alpha))^2, \quad (1)$$

где K – число трасс сейсмограммы, n – число дискретных значений по времени, U_{ij0} и $U_{ij}(\alpha)$ соответствуют наблюдаемым данным и модельным данным для параметра α .

На следующем шаге необходимо оценить величину согласия моделей. В качестве оценки предлагается отношение целевой функции индивидуальной модели к сумме целевых функций всех моделей из рассматриваемого набора:

$$p_s(\alpha_i) = \frac{F(\alpha_i)}{\sum_{j=1}^L F(\alpha_j)}, \quad (2)$$

где L – число рассматриваемых моделей. Подбираем пары близкие по величине $p_s(\alpha)$.

Модели с низкими величинами согласия с большей вероятностью подвержены отбору, чем модели с высокими величинами согласия. Две модели с самыми низкими значениями образуют пару. Из оставшихся моделей пару образуют модели с низкими величинами согласия и так далее. Отбор будет продолжаться до тех пор, пока модели не будут разбиты на пары. Далее при помощи генетических операций скрещивания и мутации каждая пара моделей порождает две вторичные пары. В результате образуется новая популяция. Скрещивание позволяет произвести обмен информацией между парными моделями и как следствие породить новые модели. Предполагается, что скрещивание может быть одноточечным и многоточечным. В случае одноточечного скрещивания простая (однобитовая) позиция (ТС) двоичного набора выбирается случайно в соответствии с равномерным распределением (рис. 1). В случае многоточечного скрещивания эта операция выполняется независимо для каждого параметра модели.

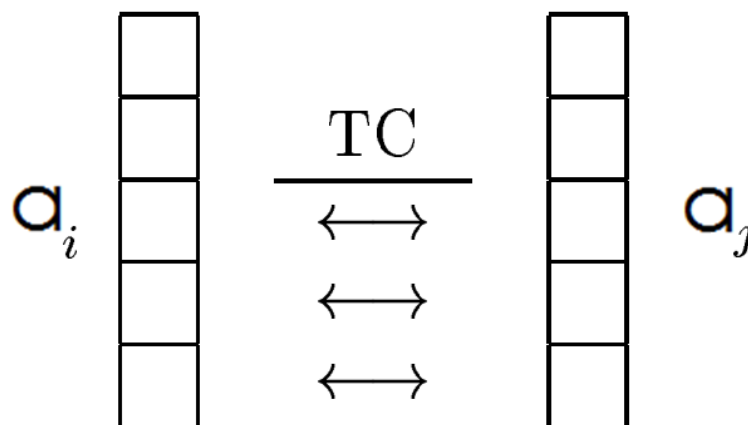


Рис. 1. Работа оператора скрещивания

Целью мутации является случайное изменение двоичного кода, как правило, в последнем разряде. Процедура мутации выполняется, если величина ε , полученная генератором случайных чисел, меньше

порогового значения вероятности выполнения мутации. Пороговое значение в программе задается небольшим числом и обычно не превосходит значения 0.2. Выбор значения зависит от количества экстремальных точек. Если ожидается небольшое количество экстремумов, то значение мутации можно установить на значении 0. Если количество экстремумов большое и необходимо определить глобальное значение из них, то значение мутации можно установить на значении 0.1 или 0.2. Величина ε варьируется от 0 до 1. Заключительным этапом работы алгоритма является выбор. Выбор осуществляется для каждой $L/2$ пар моделей, которые имеют наименьшие функции согласия. В результате получаем набор из $L/2$ моделей. Далее мы снова возвращаемся к этапу скрещивания и мутации. Эти $L/2$ моделей аналогично формируют $L/4$ пары и т. д. Эта процедура повторяется до тех пор, пока не будет получена оптимальная модель, т. е. до тех пор, пока алгоритм не достигнет глобального минимума.

СИНТЕТИЧЕСКОЕ ТЕСТИРОВАНИЕ

Глубина до границы раздела сред принималась равной 400 метрам. После границы скорость волны составляла 4 км/с. Угловой коэффициент меняется от 0.1 до 1.3, истинное значение 0.3. В формуле линейного изменения скорости свободный член также меняется по линейному закону, начиная от 2 км/с. Параметры при реализации генетического алгоритма принимались следующие: пороговое значение для отбора выбиралось равным 0.1, позиция скрещивания подбирается случайно в соответствии с равномерным распределением вероятностей, пороговое значение для мутации выбиралось равным 0.2, число итераций подбиралось интерактивным способом. Мы решаем волновое уравнение для градиентного слоя методом конечных разностей [Калиткин, 1978]. Сетку выбираем прямоугольную и равномерную, а схему неявного типа. В данной работе в качестве источника сигнала мы выбрали импульс Рикера потому, что он соответствует реальному сигналу и может быть описан математически. Таким образом, мы получаем синтетическую сейсмограмму (рис. 2).

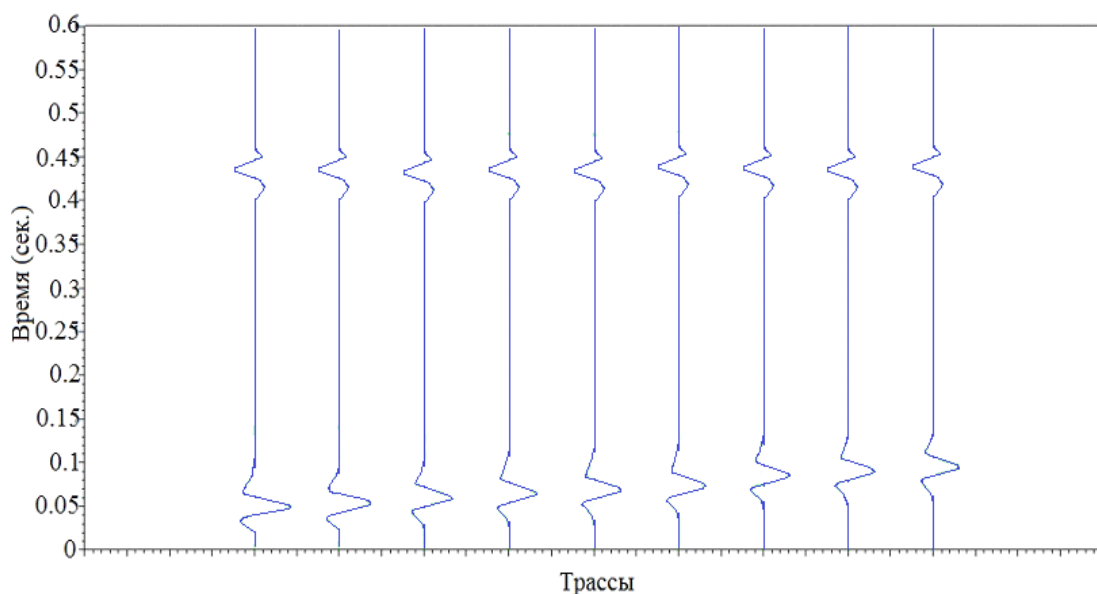


Рис. 2. Синтетическая сейсмограмма без шума

Сейсмограмма состоит из нескольких трасс, каждая из которых соответствует отдельному приемнику. На каждой трассе изображено два сигнала: прямой, идущий от источника, и отраженный, пришедший от границы раздела двух сред.

Сейсмограмма, отвечающая реальной модели, всегда содержит компоненту шума. К шуму можно отнести влияние метеоусловий и микросейсм. Учитывая физическую природу шума, предположим, что он соответствует нормальному распределению. Используем для моделирования компоненты шума генератор случайных нормально распределенных чисел. Математическое ожидание принимается равным нулю, а среднеквадратическое отклонение выбирается из заданного отношения сигнал/помеха. Важно отметить, что амплитуду необходимо брать от отраженного сигнала. Если взять амплитуду от прямого сигнала, то информация от отраженного затеряется среди помех. По этой причине первый сигнал на каждой трассе обнуляется (рис. 3).

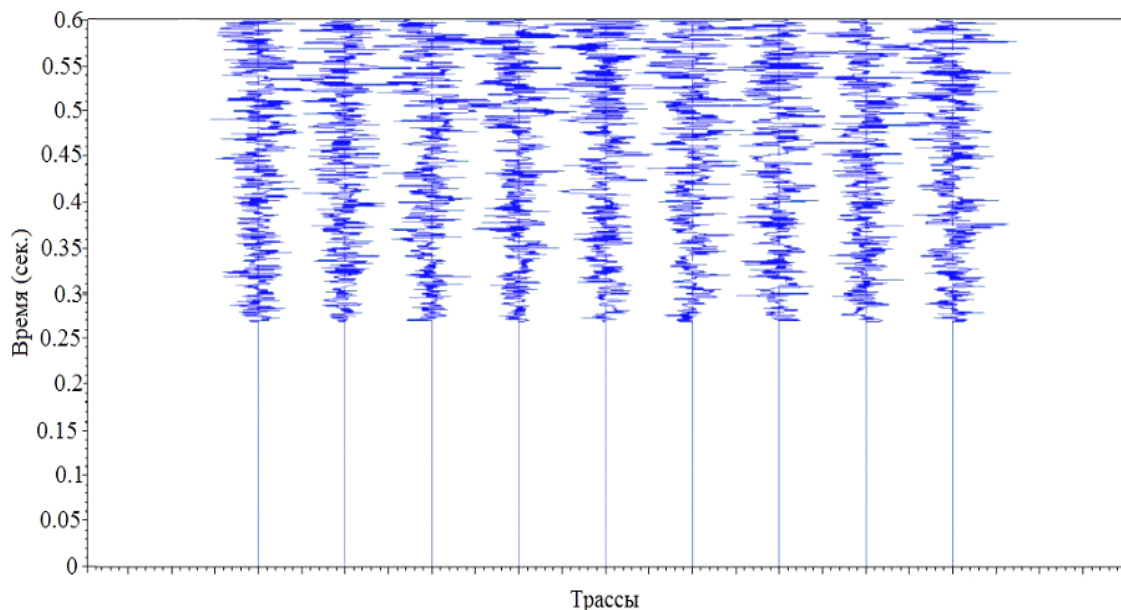


Рис. 3. Сейсмограмма с отношением сигнал/помеха, равный 1

Мы рассматриваем три случая отношения сигнал/помеха: 1; 0.75 и 0.5. Для наглядности целевые функции для всех трех случаев представлены на одном графике (рис. 4).

График функции для случая отношения сигнал/помеха, равного 1, содержит только один минимум, который соответствует значению параметра α , равный 0.3. Для этого случая можно было воспользоваться, например, градиентными методами. Вполне логично может возникнуть вопрос: зачем использовать генетический алгоритм, если существуют методы определения экстремума, такие как метод наискорейшего спуска и др.? Во-первых, ГА находит решение трудоемкой задачи относительно быстро и с достаточной точностью. Во-вторых, что особенно важно, целевая функция может иметь более одного экстремума и применение известных методов может в конечном счете привести не к глобальному, а к локальному экстремуму.

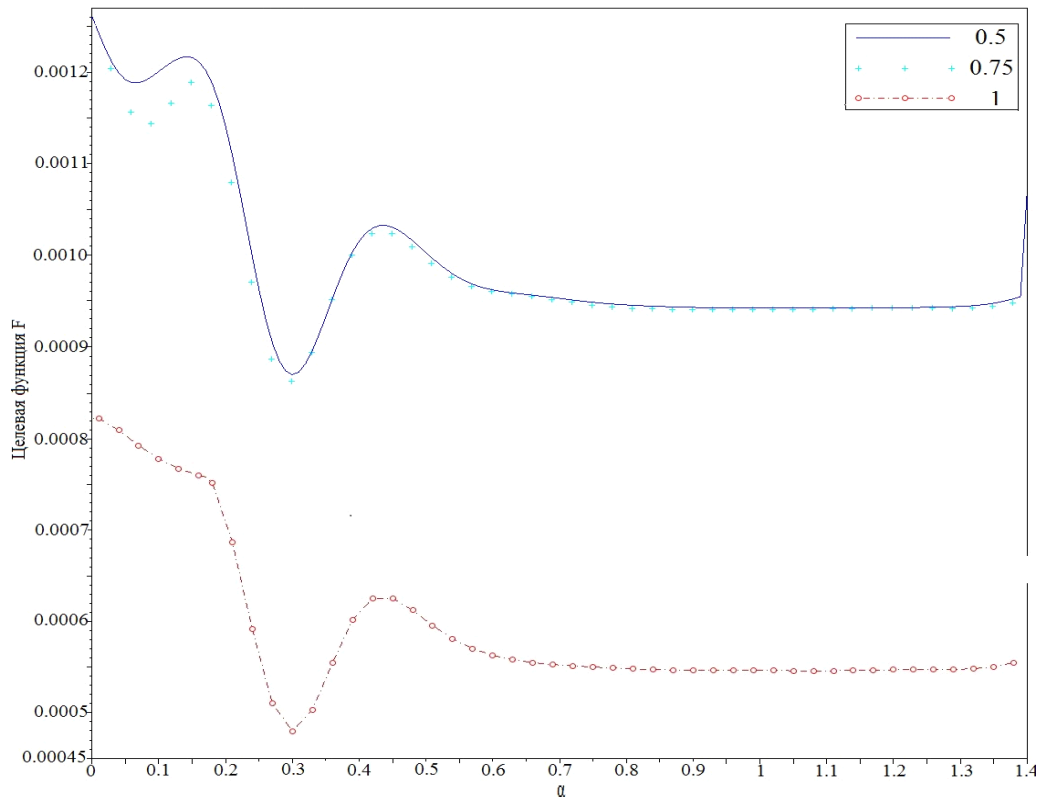


Рис. 4. График целевых функций

ГА позволяет для минимального отношения сигнал/помеха получить точную оценку скорости градиентной среды (изменение среды происходит по вертикальной составляющей), а для более высоких значений отношения сигнал/помеха можно воспользоваться другими методами, например, градиентными методами для поиска глобального экстремума.

$$F(\alpha, V_0) = \sum_{j=1}^K \sum_{i=1}^n (u_{ij0} - u_{ij}(\alpha, V_0))^2. \quad (3)$$

Аналогично предыдущему этапу оценки должны минимизировать целевую функцию F .

Каждый из графиков целевых функций (рис. 5, 6, 7) имеет более десяти минимумов. Точка пересечения горизонтальных линий есть глобальным экстремум, который соответствует значению углового коэффициента α , равного 0.3 и параметру V_0 , равного 2. Данные значения и есть правильное решение задачи.

Для конкретного случая уже сложно представить использование градиентных методов, так как они могут привести к локальному минимуму и, соответственно, к неправильному решению. Данный пример наглядно показывает преимущество ГА в поиске экстремума сложной функции. Алгоритм, несмотря на случайный поиск, находит корректное решение и в условии отношения сигнал/помеха, равного 1, и в случае отношения сигнал/помеха, равного 0.5, т. е. при достаточно высоком уровне шума.

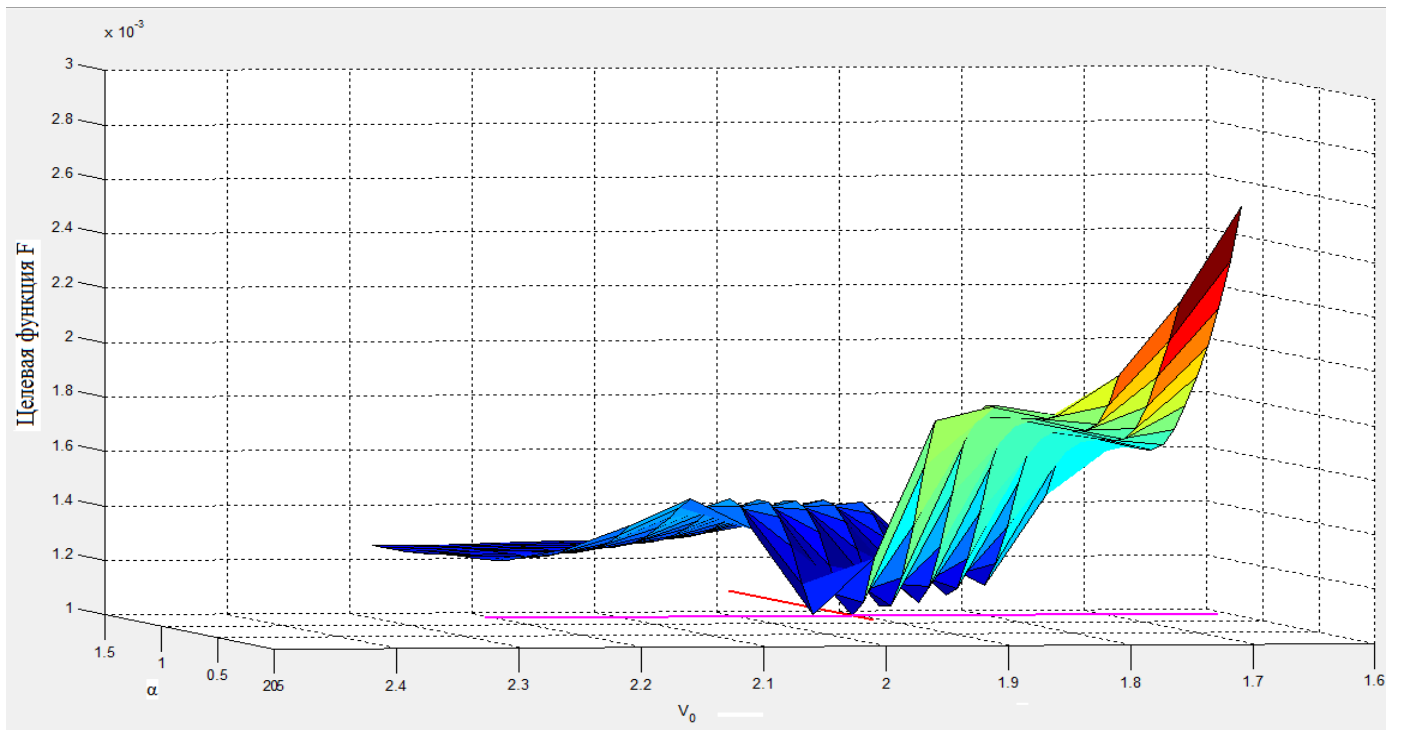


Рис. 5. Целевая функция для отношения сигнал/помеха 1

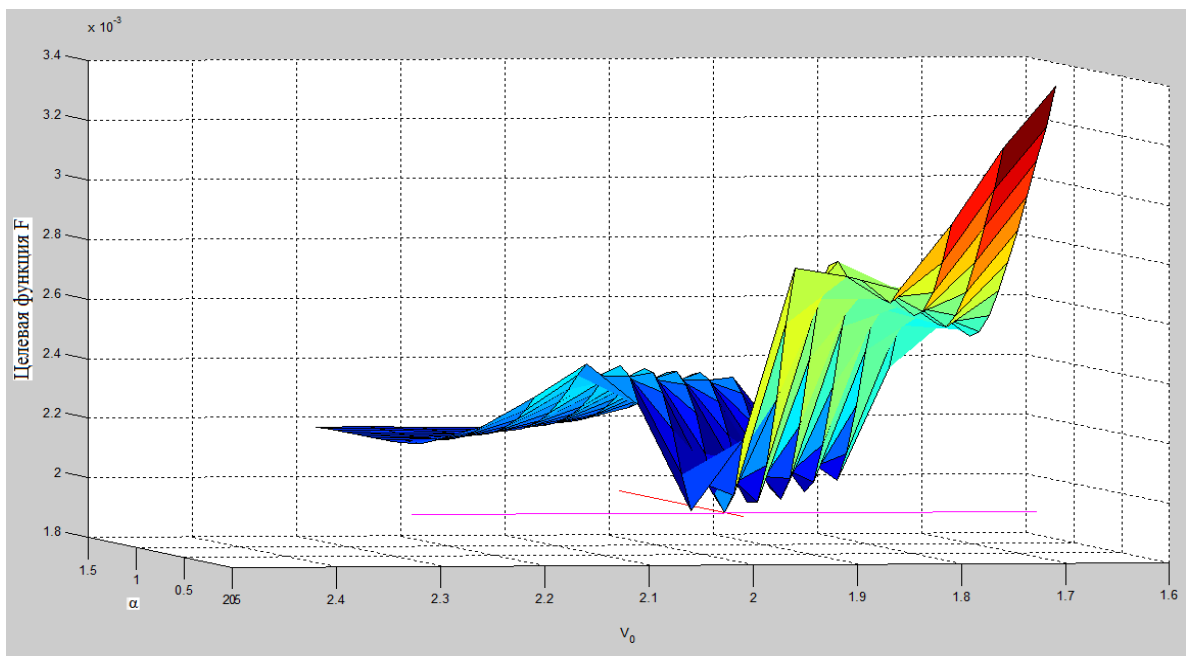


Рис. 6. Целевая функция для отношения сигнал/помеха 0.75

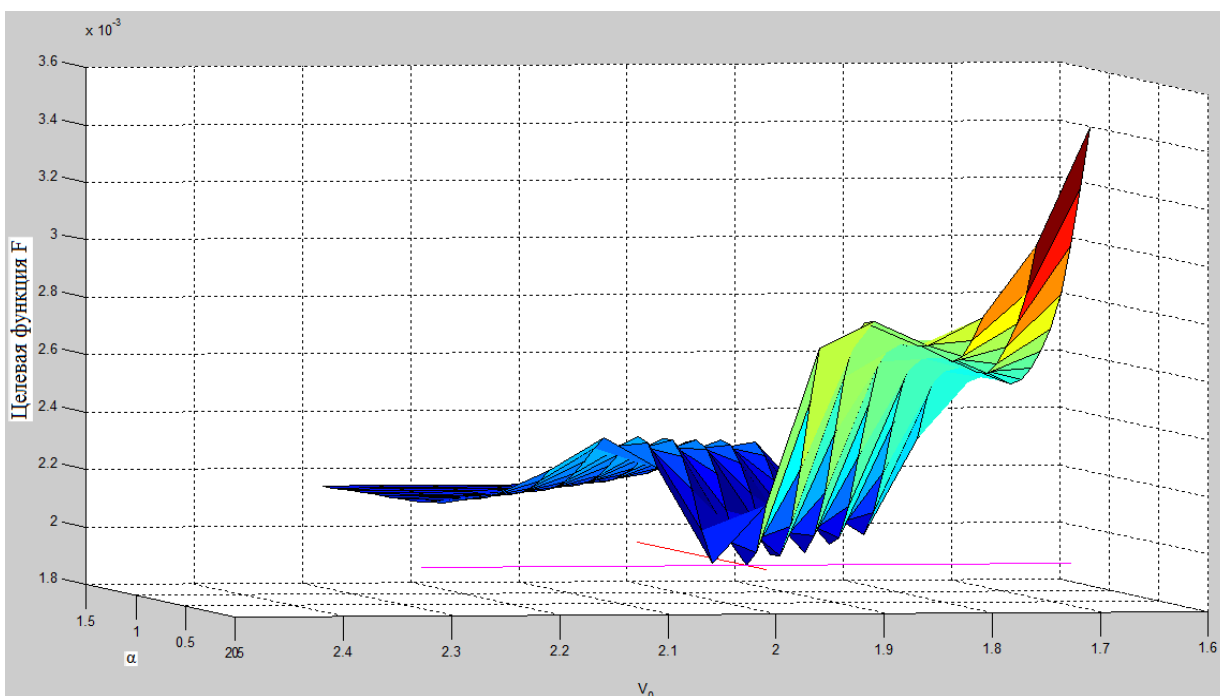


Рис. 7. Целевая функция для отношения сигнал/помеха 0.5

ЗАКЛЮЧЕНИЕ

Рассмотрено линейное изменение скорости в градиентной среде, у которой происходит изменение по горизонтальной и вертикальной составляющей. Результаты проведенных численных экспериментов показали:

- высокую точность генетического алгоритма при определении скорости распространения сейсмических волн в случае градиентного слоя;
- простоту в реализации;
- алгоритм не требует никакой информации о поведении функции;
- возможность использования алгоритма с изменяющейся средой;
- данный метод не требует задания начального значения и всегда приходит к глобальному экстремуму, в отличие от градиентных методов.

Следует отметить, что для сложных задач при поиске глобального экстремума генетический алгоритм показывает хороший результат при высоком уровне шума в экспериментальных данных. Об этом говорят результаты модельных экспериментов, представленных в данной работе.

Результаты, проведенных исследований убедительно доказывают, что ГА можно успешно использовать для оценивания скорости распространения сейсмических волн в градиентных средах.

ЛИТЕРАТУРА

Калиткин Н.Н. Численные методы. – М.: Наука, 1978. – 512 с.

Holland J.H. Adaptation in natural and artificial systems. – Ann Arbor: University of Michigan Press, 1975. – 183 p.

Reeves C.R., Rowe J.E. Genetic algorithms: principles and perspectives: a guide to GA theory. – Kluwer academic publishers, Norwell, MA. 2002. – 332 p.

Troyan V., Kiselev Yu. Statistical methods of geophysical data processing. – Singapore: World Scientific Publishing Co. Pte. Ltd, 2010. – 436 p.

Sambridge M., Drijkoningen G. Genetic algorithms in seismic waveform inversion // Geophysical Journal International. – 1992. – Vol. 109, Iss. 2. – P. 323-342.

Everett M., Schultz A. Two-dimensional nonlinear magnetotelluric inversion using a genetic algorithm // Journal of geomagnetism and geoelectricity. – 1993. – Vol. 45, Iss. 9. – P. 1013-1026.

Wilson W., Vasudevan K. Application of the genetic algorithm to residual statics estimation // Geophysical Research Letters. – 1991. – Vol. 18, Iss. 12. – P. 2181-2184.

КОРОТКО ОБ АВТОРАХ

СУЛТАНГАЛЕЕВ Руслан Робертович – аспирант кафедры физики Земли Санкт-Петербургского государственного университета.

ТРОЯН Владимир Николаевич – доктор физико-математических наук, профессор кафедры физики Земли Санкт-Петербургского государственного университета.