

АЛГОРИТМ СОВМЕСТНОЙ ИНВЕРСИИ ДАННЫХ ПРИЕМНЫХ ФУНКЦИЙ И ДИСПЕРСИОННЫХ КРИВЫХ ПОВЕРХНОСТНЫХ ВОЛН МЕТОДОМ МОНТЕ-КАРЛО

Ступина Т.А., Бушенкова Н.А.

*Институт нефтегазовой геологии и геофизики им. А.А. Трофимука СО РАН,
г. Новосибирск
StupinaTA@ipgg.sbras.ru*

В настоящей работе приводятся результаты апробации на синтетических моделях разработанной методики (и программного кода) глобальной оптимизации целевого функционала на основе метода Монте-Карло для решения задачи инверсии геофизических данных, представленных функциями приемника и дисперсионными кривыми групповых скоростей поверхностных волн.

В предлагаемом алгоритме целевой функционал определяется через апостериорную функцию распределения $p(m|d)$ с использованием байесовского подхода [1], то есть вычисляются наиболее вероятные параметры модели m , удовлетворяющие наблюдаемым данным d . Выборочное оценивание этого функционала осуществляется через функцию правдоподобия направленным перебором параметров модели с применением цепей Маркова (Markov Chain Monte Carlo, МСМС) [2]. В методах МСМС ключевым моментом является определение состояний, в которые переходит модель, исходя из её текущего состояния. Как отмечают некоторые авторы (например, [3]), неправильно подобранная схема (слишком «сильные» или «слабые» возмущения) может привести к неэффективности алгоритма, существенно замедлив перебор в пространстве моделей. В данной работе мы рассматриваем модель Земли: $m = (V_p, V_s, h)_{i=1, n}$, состоящей из n однородных, изотропных слоев, разделенных $n - 1$ плоскими горизонтальными границами раздела. Нижний n -й слой модели представляет собой полупространство. В такой модели «слабые» возмущения в реализации алгоритма задаются изменением параметров в отдельном слое, «сильные» – изменением количества слоёв. Мощность слоя определяем через координаты границ, упругие свойства каждого слоя - значениями

скорости поперечной волны V_s в слое и отношением скоростей V_p/V_s . Плотность среды ρ в каждом слое определяется согласно линейному закону Бёрча: $\rho = 0,32 V_p + 0,77$. Интервальные значения параметров, удовлетворяющие физическим законам распространения объемных и поверхностных волн в упругой изотропной среде, для всех моделей представлены в таблице 1.

Таблица 1. Интервальные значения параметров

Число слоев, n	V_p , км/сек.	V_p/V_s	ρ , кг/м ³
4 – 13	2 – 5	1,45 – 2,1	1,69 – 4,13

Синтетическое моделирование осуществляется в соответствии со следующей схемой. На первом этапе для некоторой «реальной» модели m^* , состоящей из 6 слоев, в числовых диапазонах близких к реальным ($V_p/V_s = 1.75$, $V_s = \{2.3; 3.1; 5; 2.5; 3; 4.5\}$, $h = \{0 - 10; 10 - 20; 20 - 28; 28 - 38; 38 - 40; 40 - 70\}$) вычисляется вектор наблюдаемых данных $d_{obs} = d(m^*) = (d^{RF}, d^{SWD})$. На рисунке 1 по скоростному параметру (график 3) эта модель и наблюдаемые данные (графики 1 и 2) выделены зеленым цветом. Прямая задача решается с помощью высокоскоростных, что немаловажно при реализации метода Монте-Карло, программных приложений hrftn96 и surf96 из пакета CPS330 [4], написанных на языке Фортран. Вычисление приемных функций d^{RF} в программе hrftn96 основано на матричном методе Томсона-Хаскелла [5], вычисление дисперсионных кривых d^{SWD} групповых скоростей волн Рэля в программе surf96 - на модификации матричного метода Дункина [6]. Далее, аналогичными алгоритмами вычисляются синтетические дисперсионные кривые и приемные функции (на графиках рис. 1 показаны серым цветом), формирующие вектор данных $d(m)$ для случайной модели m с параметрами из равномерного распределения, диапазон которых указан в таблице 1.

Пошаговая схема алгоритма выглядит следующим образом.

1. Генерация стартовой модели m_0 с равномерно распределенными параметрами, которая далее становится текущей моделью.
2. Создание модели кандидата m из текущей модели m_0 при помощи одной из следующих модификаций:
 - удаление случайно выбранного слоя;

- добавление нового слоя в случайное место;
 - изменение значений V_p , V_s или h в случайно выбранном слое;
 - перемена порядка двух слоев (обмен местами).
3. Расчёт вектора модельных данных $d(m_o)$ и $d(m)$.
 4. Вычисление функций правдоподобия $L(m_o)$ и $L(m)$ для стартовой модели и модели кандидата в соответствии с нормальным законом распределения и функцией потерь: $\Phi(m) = (d(m) - d(m^*))^T C^{-1} (d(m) - d(m^*))$;

$$L(m) = \frac{1}{\sqrt{(2\pi)^n |C|}} \times \exp\left\{\frac{-\Phi(m)}{2}\right\}, \text{ где } C - \text{ матрица ковариации}$$

с параметризацией по параметрам (уровень шума) и r из интервала $[0 \div 1]$.

5. Проверка условия: если $L(m) > L(m_o)$, то модель m принимается в качестве текущей модели и $m_o = m$, в противном случае, если отношение функций правдоподобия больше некоторого случайного числа s ($L(m)/L(m_o) > s$) из диапазона $[0 \div r]$, модель m также остается текущей.
6. Повтор шагов 2 – 5 до тех пор, пока число итераций не будет превосходить наперёд заданное число N_{iter} .

С каждой итерацией работы алгоритма плотность распределения выборки моделей приближается к распределению апостериорной плотности вероятности с учетом того, что оптимальное соотношение между количеством принятых моделей и моделей кандидатов мы определили в диапазоне от 0,25 до 0,50 в соответствии с результатами работ [7, 8].

В результате моделирования, выполненного при $N_{\text{iter}} = 6,4 \times 10^6$, мы получили наилучшее байесовское решение, оцененное как среднее из 64 апостериорных моделей с максимальной функцией правдоподобия относительно остальных. На рисунке 1 это решение выделено желтым цветом, его минимальное отклонение от исходной модели указывает на работоспособность применяемой методики к данному классу задач на «идеальных» входных данных. Решение, вычисленное как среднее из 100 случайно выбранных из апостериорного распределения решений, на рисунке 1 показано черным цветом с очевидным его отклонением от исходной модели.

Далее, на модельные приемные функции накладывался гауссовский шум, на дисперсионные кривые - экспоненциальный шум с дисперсиями $\sigma = 0,0; 0,05; 0,1; 0,15; 0,2; 0,25$ и величиной корреляции 0,85 во всех случаях.

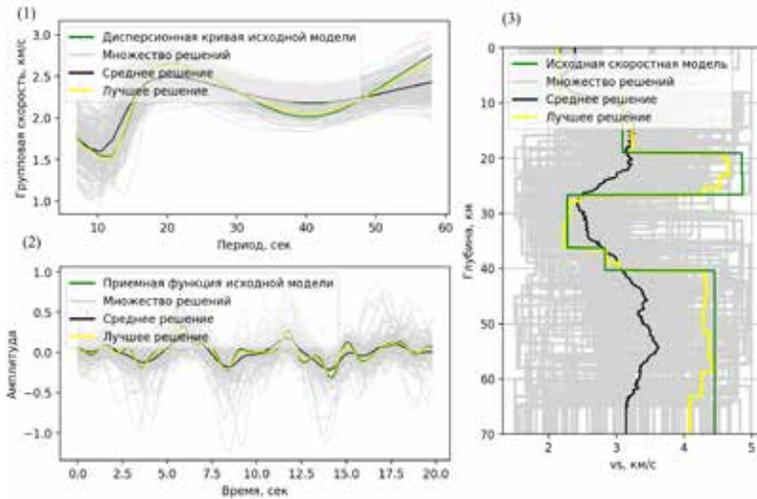


Рис. 1. Результаты совместной инверсии синтетических данных без шума. Зеленая линия: на графике (1) – дисперсионная кривая, на графике (2) – приемная функция, на графике (3) – распределение скорости по глубине для исходной модели m_0 . Серые линии – множество соответствующих графикам решений, смоделированных по методу Монте – Карло. Черная линия – среднее решение. Желтая линия - наилучшее байесовское решение

В случае модельных данных без шума и с шумом $\sigma = 0,05$, результаты показали хорошее восстановление исходного профиля скорости S-волны. Для данных с шумом $\sigma = 0,25$ восстановить заданную модель практически не удастся, различить возможно лишь самые крупные структуры, такие как слой толщиной 8 км на глубине 20 км и слой мощностью 10 км на глубине 28 км, мелкие детали гарантированно теряются.

Работа выполнена при финансовой поддержке ФНИ FWZZ-2022-0017.

Литература

1. Tarantola A. Inverse problem theory: methods for data fitting and model parameter estimation // Elsevier. 1987.
2. Mosegaard K., Tarantola A. Monte Carlo sampling of solutions to inverse problems // Journal of Geophysical Research. 1995. V. 100. № B 7. P. 12431-12447.
3. Malinverno A. 2002. Parsimonious Bayesian Markov chain Monte Carlo inversion in a nonlinear geophysical problem // Geophys. J. Int. 2002. V. 151. № 3. P. 675-688.
4. Herrmann R. B. Computer programs in seismology: An evolving tool for instruction and research // Seismological Research Letters. 2013. V. 84. P. 1081-1088.
5. Haskell N.A. The dispersion of surface waves on multilayered media // Bull. seism. Soc. Am. 1953. V. 43. P. 17-34.
6. Dunkin, J.W. Computation of modal solutions in layered, elastic media at high frequencies // Bull. seism. Soc. Am. 1965. V. 12. P. 335-358.
7. Mosegaard K. Monte Carlo analysis of inverse problem // PhD thesis, Copenhagen University. 2006.
8. Rosenthal J.S. Optimal proposal distribution and adaptive MCMC // Chapter for MCMC Handbook. Chapman and Hall/CRC Press. 2009.