

ГЕОФИЗИЧЕСКАЯ СТРУКТУРОСКОПИЯ НА ОСНОВЕ АЛГОРИТМА САМООБУЧЕНИЯ

Алёшин Ю.Г.

ИГиОН НАН КР

Неполнота априорной информации, состоящая в том, что заранее, как правило, неизвестно ни отношение правдоподобия $f_i(\bar{x})/f_j(\bar{x})$, ни априорные вероятности $P_i \forall i, j \in \overline{1, m}; j \neq i$, является серьёзным препятствием для применения классического байесова подхода [1], да и всех других алгоритмов распознавания, в которых так или иначе используется соотношение вероятностей: P_i/P_j . В этих случаях возникает необходимость в использовании обучающихся алгоритмов, способных после некоторого периода обучения на реальном горном массиве выработать решающее правило, для того чтобы на последующих этапах произвести достоверное распознавание его структуры и свойств. Недостаток априорной информации преодолевается обучением. Чем менее детализирована априорная ФГМ, тем большее время и больше точек опробования необходимо для обучения.

Воспользуемся тем, что априорная модель так или иначе содержит информацию о наличии в массиве горных пород нескольких структурных классов и покажем, что новые решающие правила, способные более достоверно произвести распознавание, могут быть получены единообразным путём с использованием одной и той же обучающей выборки. Специфика распознавания структурных классов в реальных горных массивах состоит в том, что априорная вероятность присут-

ствия того или иного класса, вследствие геологических закономерностей, существенно зависит от глубины его опробования, т.е. фактически следует рассматривать все процедуры распознавания при условии $P_i(h)$.

Формальная постановка и строгое математическое решение задачи обучения дано В.Я. Цыпкиным [2,3]. Воспользуемся этим аппаратом для решения задачи геофизической структуроскопии.

В ситуации, возникшей при исследовании оползневого склона в Кок-Жангаке, число структурных классов заранее было неизвестно, поэтому информации о принадлежности каких-либо геофизических показателей к тем или иным классам нет и единственный путь формирования системы распознавания – применение методов самообучения. К самообучению приходится прибегать и тогда, когда хотя заранее и известно число классов, однако обучающая выборка не дана, имеются лишь некоторая совокупность геолого-геофизических объектов и значения признаков, которыми они характеризуются, т.е. даны объекты $\omega_1, \dots, \omega_m$ и величины $x_1^{(1)}, \dots, x_n^{(1)}, x_1^{(2)}, \dots, x_n^{(2)}, x_1^{(m)}, \dots, x_n^{(m)}$, но не указано, к каким классам относятся эти объекты. При этом необходимо обратить внимание на то, что построение самого обучающего алгоритма распознавания во всех случаях базируется на известном, заранее выбранном словаре признаков.

Рассмотрим в общем виде, как может быть выбрана процедура самообучения, когда неизвестно ни число структурных классов, ни априорная вероятности их залегания в исследуемом горном массиве. Априорная информация сводится лишь к известному множеству объектов $\omega_1, \dots, \omega_l$, которые потенциально могут слагать исследуемый массив горных пород, а также значению векторов $\{\bar{x}_j\}_i$, которыми описывается каждый объект (хотя бы некоторые точечные или интервальные оценки).

Обозначим через $P(A_i)$ вероятность появления геологического объекта класса A_i , а через $f_i(\bar{x})$ условную плотность распределения значений признаков внутри A_i -го класса. Хотя функция $P(A_i)$ и $f_i(x)$ неизвестны, однако относительно совместной плотности распределения вероятностей:

$$f(\bar{x}) = \sum_{i=1}^m P(A_i) f_i(\bar{x}) \quad (1)$$

естественно предположить, что при унимодальных функциях $f_i(\bar{x})$ центрам классов в признаковом пространстве соответствуют существенные максимумы функции $f(\bar{x})$. Поэтому цель самообучения на основе предъявления векторов $\{x\}_t$ и заданных условий объединения объектов в классы состоит в том, чтобы определить совместную плотность распределения, найти её существенные максимумы, а следовательно, и центры классов и тем самым их число m .

Для решения задачи воспользуемся алгоритмом обучения, основанным на методе стохастической аппроксимации [2,3]. Аппроксимируем искомую совместную плотность распределения конечным набором ортонормированных функций $\varphi_j(x)$:

$$\bar{f}(\bar{x}, c) = \sum_{j=1}^N c_j \varphi_j(\bar{x}) = \bar{c}^T \bar{\varphi}(\bar{x}) \quad (2)$$

Введём квадратичную меру уклонения искомой плотности распределения от истинной:

$$J(c) = \int_x^0 \left[f(\bar{x}) - \bar{c}^T \bar{\varphi}(\bar{x}) \right]^2 d\bar{x} \quad (3)$$

Задача состоит в том, чтобы найти такое оптимальное значение вектора $\bar{C} = \bar{C}^0$, которое минимизирует функционал (3), т.е.

$$\bar{C}^0 = \arg \min J(\bar{C}) \quad (4)$$

В силу того, что компоненты вектора функции $\overline{\varphi(\bar{x})}$ ортонормированы, оптимальное значение вектора C определяется следующим образом:

$$\bar{C}^0 = \int \overline{\varphi(\bar{x})} f(\bar{x}) dx = M \{ \overline{\varphi(\bar{x})} \} \quad (5)$$

т.е. равно математическому ожиданию вектора функции $\overline{\varphi(\bar{x})}$, значения реализаций которого могут быть получены на основании априорной информации – величин компонентов векторов информативных параметров $\{\bar{x}_j\}_i$, Я.И. Цыпкиным предложена рекуррентная процедура оценки

$$\bar{C}[n] = \bar{C}[n-1] + \gamma[n] \left[\overline{\varphi(\bar{x}[n])} \right] - \bar{C}[n-1] \quad (6)$$

где: $\gamma[*]$ – элемент диагональной матрицы, обеспечивающий условие сходимости алгоритма обучения.

Зачастую принимается $\gamma[n] = n^{-1}$ для всех компонентов векторов $\bar{C}[n]$, что обеспечивает минимум дисперсии оценки $\bar{C}$ при любом фиксированном значении n .

Определив по (5) значение C^0 и подставив в (6), можно получить функцию: $\bar{f}(\bar{x}C_0^0)$. Её анализ и выявление существенных максимумов позволяет определить число классов, к которым могут быть отнесены объекты изучаемой совокупности, что обеспечивает возможность построения байесовского алгоритма распознавания; эта процедура продемонстрирована нами в [1].

Рассмотрим конкретный пример использования процедуры обучения в геофизической структуроскопии в условиях обследования того же оползневого склона в пос. Кок-Жангак. Перед началом инженерно-геологических работ эксперты сформировали две альтернативные ФГМа (априорные), которые основывались на результатах обследования соседних участков, где произошли оползневые процессы:

– ФГМа-1 – горный массив сложен лёссовидными суглинками мощностью от 7 до 10 м и во всяком случае не превышающими 12 м, кровля коренных пород – меловые отложения (по аналогии с оползневым обнажением на участке, расположенном в 300 м от изучаемого);

– ФГМа-2 – та же двухслойная модель, однако мощность покровного чехла от 30 до 50 м по аналогии с участком Сары-Булак, находящимся в 1,5 км от изучаемого.

Распределение или функция плотности вероятности положения кровли коренных пород, исходя из сформулированных предположений, были приняты гауссовским:

$$f_{am}(H) = \frac{1}{\delta_1 \sqrt{2\pi}} \cdot \exp\left\{-\frac{(H - a_1)^2}{2\delta_1^2}\right\} \quad (7)$$

где: для модели ФГМа-1 $a_1 = 8$ м, $\delta_1 = 1$ м, $i = 1$;

для модели ФГМа-2 $a_2 = 40$ м, $\delta_2 = 4$ м, $i = 2$.

Относительно положения кровли коренных пород по падению горного склона не было выдвинуто никаких предположений. По умолчанию будем полагать, что она параллельна дневной поверхности. По результатам обучения на малой выборке, объёмом N , как это показано на рис. 1, требуется вынести более определённое суждение о мощности лёссовидных суглинков, перекрывающих коренные породы.

Искомая последовательная оценка априорной вероятности $f_{an}(A_2/H_j)$ на любом n -ом шаге обучения в соответствии с алгоритмом стохастической аппроксимации определяется на основе (6):

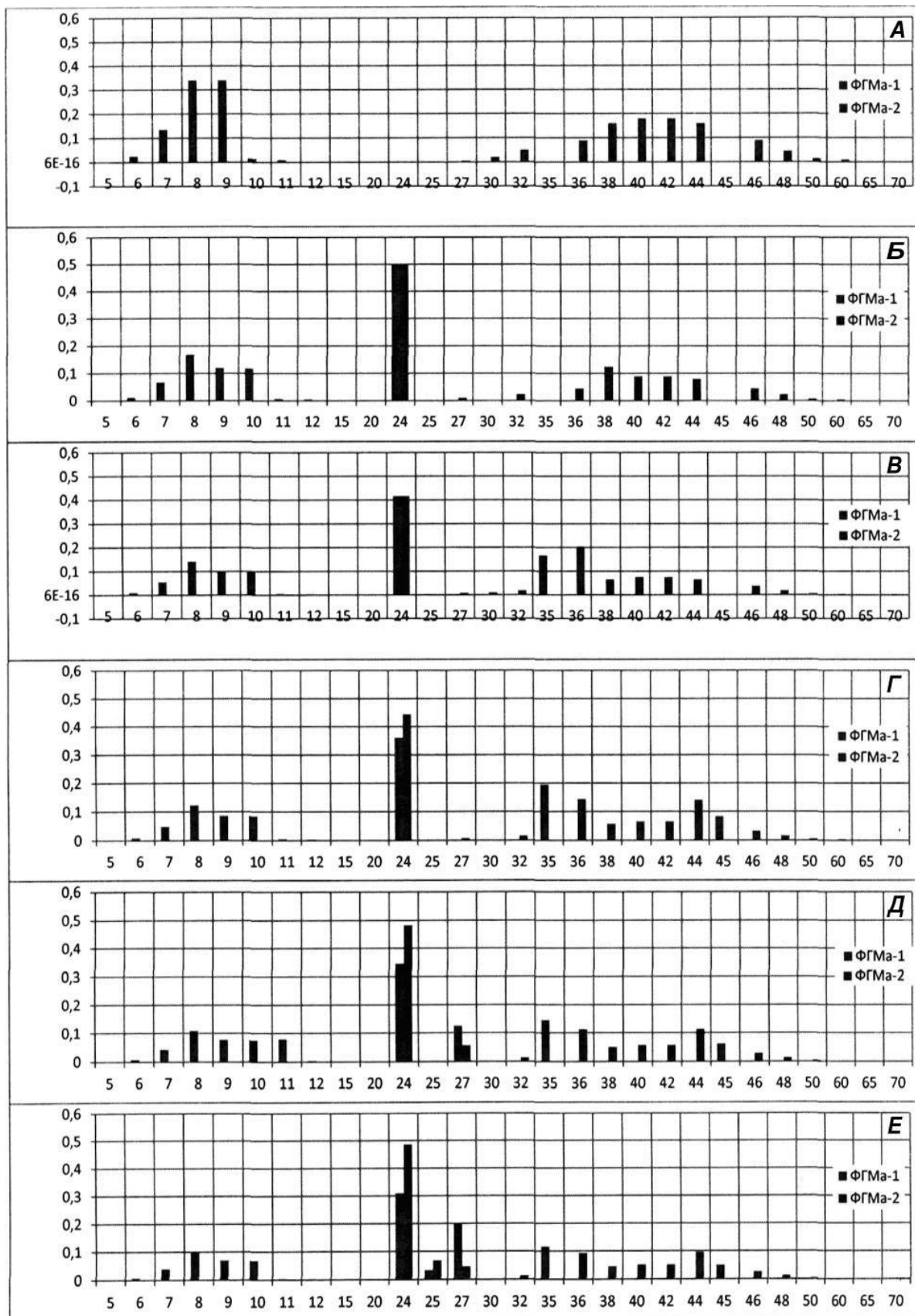


Рис. 1. Трансформация априорных функций плотности вероятности (А) ФГМа-1 и ФГМа-2 по мере получения информации о положении кровли коренных пород по экспериментальным данным: Б- $\eta=1$; В- $\eta=2$;...Е- $\eta=5$.

$$\bar{f}_{an}(A_2/H_j) = \bar{f}_{a(n-1)}(A_2/H_j) + \frac{1}{n} \left[\frac{f_{a0}(A_2/H_j) + K(H_j)}{1+n} - \bar{f}_{a(n-1)}(A_2/H_j) \right] \quad (8)$$

где: $\bar{f}_{a(n-1)}$ – оценка плотности вероятности, полученная на предыдущем шаге обучения после распознавания;
 $K(H_j)$ – количество экспериментов, в которых определён класс A_2 на глубине H_j .

Для удобства вычислений распределение (8) следует дискретизовать и представить в виде типичной гистограммы. На рис. 1, представлены результаты оценки априорного распределения положения кровли коренных пород в период обучения системы распознавания при объёме обучающей выборки N' от 0 до 5 ($n = 0$ соответствует двум априорным моделям ФГМа-1 и ФГМа-2). Хорошо видно, что ни одна из априорных ФГМ, сформированная на основании мнения экспертов, не отражает реальную ситуацию на объекте. В результате проведения обучения средняя мощность четвертичного покрова должна быть принята равной от 23 до 27 м. К этому интервалу значений асимптотически устремляются средние двух исходных ФГМа по мере увеличения объёма обучающей выборки. Важно отметить также и то, что, несмотря на существенное различие в исходных ФГМа- i ($i = 1,2$), в процессе обучения модели довольно быстро сходятся (см. рис. 2) и формируется новая априорная модель ФГМа-3, в которой мощность четвертичного покрова уже определяется распределением, показанным на рис. 1е, с параметрами:

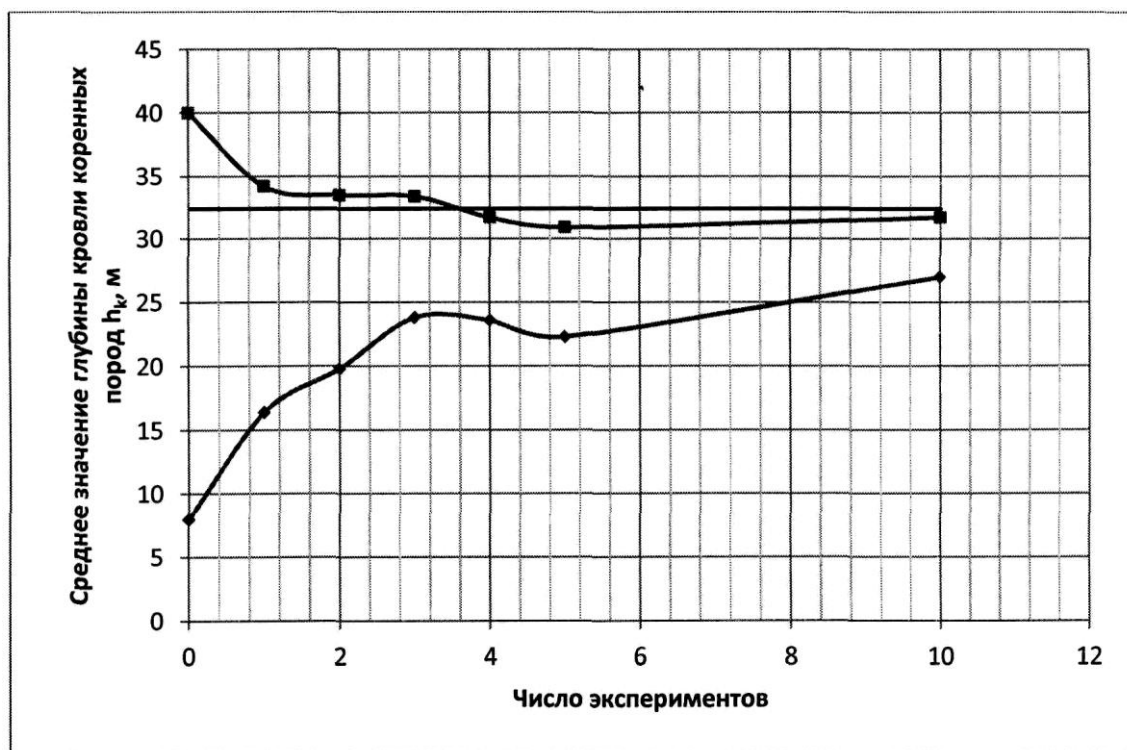


Рис. 2. Конвергенция среднего значения положения кровли коренных пород на оползневом склоне "Интернациональный" по мере получения информации и использовании метода стохастической аппроксимации.

$M\{H\}=26,1$ и дисперсией $\delta^2\{H\} = 106,5 \text{ м}^2$. На оси глубин H в интервале от 6 до 12 м ещё заметны следы распределения "остаточной" вероятности, сформированной ФГМа-1, но не "подавленные" окончательно обучающей выборкой малого объёма. Общую картину сходимости к среднему демонстрирует рис. 2.

По результатам наиболее полного обследования оползневого склона (48 точек опробывания по двум геофизическим профилям) окончательно установлено: среднее значение мощности покровных отложений, включая кору выветривания коренных пород, на оползневом склоне "Интернациональный" в поселке Кок-Жангак составляет 32,7 м при среднеквадратичном отклонений 13,8 м. Вариация результата связана не столько с точностью метода электроразведки, сколько с геологическим фактором – наличием участка разрывного нарушения

в коренных породах, что предопределило чрезвычайно высокую изрезанность кровли коренных пород.

ЛИТЕРАТУРА

1. Алёшин Ю.Г. Исследование универсальных статистических алгоритмов распознавания в геофизической структуроскопии: байесовский классификатор // Современные проблемы механики сплошных сред, вып. 9. 2009. С. 26-34.
2. Цыпкин Я.З. Адаптация и обучение в автоматических системах. М.: Наука, 1958 – 240 с.
3. Цыпкин Я.З. Основы теории обучающихся систем. М.: Наука, 1970 – 252с.