МИНИСТЕРСТВО ОБРАЗОВАНИЯ И НАУКИ,
МОЛОДЁЖИ И СПОРТА УКРАИНЫ
ГОСУДАРСТВЕННОЕ ВЫСШЕЕ УЧЕБНОЕ
ЗАВЕДЕНИЕ
«НАЦИОНАЛЬНЫЙ ГОРНЫЙ УНИВЕРСИТЕТ»

Кафедра геофизических методов разведки

МЕТОДИЧЕСКИЕ УКАЗАНИЯ

к лабораторной работе
«ОПРЕДЕЛЕНИЕ ВИДА НЕЛИНЕЙНОЙ АНАЛИТИЧЕСКОЙ ЗАВИСИМОСТИ»

по дисциплине
“Статистическая обработка геологической информации ”

для студентов направления 6.040103 – ГЕОЛОГИЯ

Днепропетровск
2012

Методические указания к лабораторной работе «Определение вида нелинейной аналитической зависимости» по курсу «Статистическая обработка геологической информации» для студентов специальности 6.040103 – ГЕОЛОГИЯ / Сост. В.П. Солдатенко. – Днепропетровск: НГУ, 2010. – 9 с.

Составитель В.П. Солдатенко, кандидат геол.-мин. наук, доцент

Ответственный за выпуск заведующий кафедрой геофизических методов разведки доцент, докотор геологических наук Довбнич М.М.

Цель работы: изучение методов и приемов определения вида аналитической зависимости при нахождении связи между эмпирическими данными.

Введение

Анализ различных процессов приводит к необходимости выявления существенных факторов, влияющих на исследуемый процесс, а также к выбору формы связи между факторами и к оценке параметров полученных уравнений.

Будем считать, что некоторое явление характеризуется двумя варьируемыми величинами x и y, из которых x выбирается как независимая, а y – как зависимая переменная величина. Предположим, что между переменными x и y существует однозначное соответствие, т.е. Каждому значению независимой переменной величины x соответствует с заданной степенью точности одно значение зависимой переменной. В этом случае возникает задача – выявление формы связи и определение формульной зависимости, задающей y как функцию $f(x)$.

В общем виде задачу можно сформулировать следующим образом. Пусть в результате исследований некоторой величины x значениям x_1, x_2,..., x_n поставлены в соответствие значения y_1, y_2, ..., y_n некоторой величины y. Требуется подобрать вид аналитической зависимости $y=f(x)$, связывающей переменные x и y.

Будем называть аналитические зависимости, полученные в результате наблюдений *эмпирическими*. Выявление эмпирических зависимостей делится на два основных этапа – ***выбор эмпирической формулы*** и ***уточнение коэффициентов*** выбранной формулы.

1. Выбор эмпирических формул для нелинейных зависимостей

На практике обычно пытаются представить аналитическую зависимость в виде прямой общего положения, которая задается формулой $y=ax+b$. Однако, это не всегда удается. Графическое же построение нелинейной зависимости не дает ответа на вопрос о том, какой аналитический вид имеет эта функция, т.е. Будет ли эта зависимость степенной, дробно-рациональной, логарифмической и т.д.

Пусть y есть функция одной переменной с двумя параметрами a и b. В качестве набора функций, из которых будем выбирать эмпирическую зависимость, рассмотрим:

1) линейную функцию $y=ax+b$;

2) показательную функцию $y=ab^x$;

3) дробно-рациональную функцию $y=1/(ax+b)$;

4) логарифмическую функцию $y=a\ln x+b$;

5) степенную функцию $y=ax^b$ (она определяет параболическую зависимость, если параметр $b>0$, и гиперболическую зависимость, если $b<0$; если же параметр $b=0$, то зависимость вылождается в линейную);

6) гиперболическую зависимость вида $y=a+b/x$;

7) дробно-рациональную функцию вида $y=x/(ax+b)$.

Для наилучшего выбора вида аналитической зависимости $y=f(x,a,b)$, соответствующей построенному графику, выполним следующие промежуточные вычисления. На заданном отрезке изменения независимой переменной выберем точки, достаточно надежные и по возможности далеко отстоящие друг от друга. Для простоты будем считать, что это точки x_1 и x_n. Вычислим среднее арифметическое, среднее геометрическое и среднее гармоническое

$$x_{ар} = \frac{x_1 + x_n}{2}, \qquad x_{геом} = \sqrt{x_1 x_n}\,, \qquad x_{гарм} = \frac{2x_1 x_n}{x_1 + x_n}.$$

По вычисленным значениям независимой переменной найдем из построенного графика соответствующие значения зависимой переменной

$$x_{ар} \to y_1^*, \qquad x_{геом} \to y_2^*, \qquad x_{гарм} \to y_3^*.$$

для пока еще неизвестной аналитической зависимости $y=f(x,a,b)$. Выполним вспомогательные вычисления для зависимой переменной. Вычислим среднее арифметическое, среднее геометрическое и среднее гармоническое крайних значений

$$y_{ар} = \frac{y_1 + y_n}{2}, \qquad y_{геом} = \sqrt{y_1 y_n}, \qquad y_{гарм} = \frac{2y_1 y_n}{y_1 + y_n}.$$

Сравним найденные из графика y_1^*, y_2^*, y_3^* с вычисленными $y_{ар}$, $y_{геом}$, $y_{гарм}$ и оценим следующие погрешности:

$$|y_1^* - y_{ар}| = \varepsilon_1, \qquad |y_1^* - y_{геом}| = \varepsilon_2, \qquad |y_1^* - y_{гарм}| = \varepsilon_3,$$

$$|y_2^* - y_{ар}| = \varepsilon_4, \qquad |y_2^* - y_{геом}| = \varepsilon_5,$$

$$|y_3^* - y_{ар}| = \varepsilon_6, \qquad |y_3^* - y_{гарм}| = \varepsilon_7.$$

Найдем из этих ошибок минимальную: $\varepsilon = \min(\varepsilon_1, \varepsilon_2, ..., \varepsilon_7)$.

1 Если наименьшей среди всех абсолютных ошибок окажется ε_1, то в качестве аналитической зависимости для данного графика хорошим приближением служит линейная функция $y=ax+b$.

2 Если наименьшей абсолютной ошибкой является ε_2, то в качестве эмпирической зависимости следует выбрать показательную функцию $y=ab^x$.

3 В том случае, когда наименьшая из абсолютных ошибок есть ε_3, искомая эмпирическая зависимость определяется дробно-рациональной функцией вида $y=1/(ax+b)$.

4 Если наименьшая из абсолютных ошибок есть ε_4, то хорошим приближением служит логарифмическая функция $y=a\ln x+b$.

5 Для случая, когда наименьшей абсолютной ошибкой является ε_5, в качестве эмпирической зависимости выбирается степенная функция $y=ax^b$.

6 Если наименьшей из абсолютных ошибок окажется ε_6, то за искомую зависимость следует выбрать гиперболическую $y=a+b/x$.

7 Наконец, в том случае, когда наименьшая из абсолютных ошибок есть ε_7, в качестве аналитической зависимости вибирается дробно-рациональная функция вида $y=x/(ax+b)$.

Пример 1. Подобрать эмпирическую зависимость для функции заданной таблично:

x	1	2	3	4	5	6	7	8	9
y	521	308	240,5	204	183	171	159	152	147

Решение. 1) Предположим, что в данном примере крайние табличные значения достаточно надежны. Проведем вспомогательные вычисления и найдем для крайних значений среднее арифметическое $x_{ар}=5$, среднее геометрическое $x_{геом}=3$ и среднее гармоническое $x_{гарм}=1,8$.

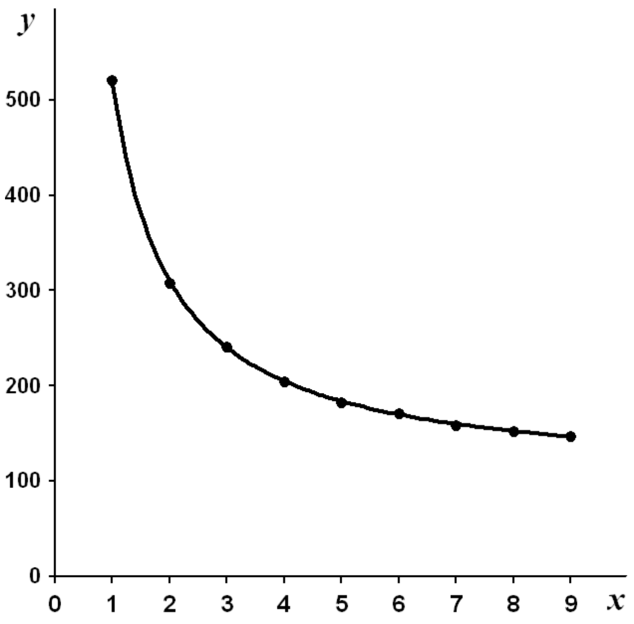

Рис.1

2) Из графика (рис. 1) найдем значение функции, соответствующие вычисленным значениям аргумента: для $x_{ар}=5$ имеем ≈180; для $x_{геом}=3$ имеем ≈240; для $x_{гарм}=1,8$ имеем ≈341.

3) Выполним дополнительные расчеты для зависимой переменной. Найдем для крайних значений среднее арифметическое $y_{ар}=334$, среднее геометрическое $y_{геом}=274$, среднее гармоническое $y_{гарм}=228$.

4) Сравним найденные графически значения зависимой переменной с $y_{ар}$, $y_{геом}$ и $y_{гарм}$.

$\varepsilon_1,=|180–334|=154$, $\varepsilon_2,=|180–274|=106$, $\varepsilon_3,=|180–228|=48$,

$\varepsilon_4,=|240–334|=94$, $\varepsilon_5,=|240–274|=34$, $\varepsilon_6,=|341–334|=7$,

$\varepsilon_7,=|341–228|=113$.

Поскольку наименьшая из абсолютных ошибок есть ε_6, в качестве аналитической зависимости следует выбрать гиперболическую зависимость $y=a+b/x$.

2. Уточнение коэффициентов

Для уточнения коэффициентов выбранной аналитической зависимости $y=f(x,a,b)$ можно воспользоваться несколькими методами.

2.1. Метод выбранных точек. На построенной кривой возьмем две произвольные точки M(x_1^* ; y_1^*) и N(x_2^*; y_2^*). Составим систему

$$\begin{cases} y_1^* = f\left(x_1^*, a, b\right) \\ y_2^* = f\left(x_2^*, a, b\right) \end{cases}$$

и решим ее относительно искомых параметров a и b.

2.2. Метод средних. В эмпирическую формулу $y=f(x,a,b)$ последовательно подставляем табличные значения $x_i(i=1, 2, ..., n)$. Полученные значения функции $y_i=f(x_i,a,b)$, вообще говоря, будут отклоняться от табличных значений: $y_i-f(x_i,a,b)=\varepsilon_i$.

Согласно методу средних, за наилучшее положение кривой принимается то, для которого алгебраическая сумма отклонений равна нулю, т.е.

$$\sum_{i=1}^{n} \varepsilon_i = \sum_{i=1}^{n} \left[y_i - f\left(x_i, a, b\right)\right] = 0 .$$

Для определения параметров a и b по методу средних вся совокупность $\varepsilon_i(i=1, 2, ..., n)$ разбивается на две группы так, чтобы алгебраическая сумма уклонений каждой группы равнялась нулю. Таким образом, для определения параметров a и b получаем следующую систему двух уравнений:

$$\begin{cases} \sum_{j=1}^{l} \left[y_j^I - f\left(x_j^I, a, b\right) \right] = 0, \\ \sum_{k=1}^{n-l} \left[y_k^{II} - f\left(x_k^{II}, a, b\right) \right] = 0, \end{cases}$$

где l и $n–l$ – числа табличных данных соответственно для первой и второй группы.

Заменив сумму разностей разностью сумм в каждом уравнении системы, получим

$$\begin{cases} \sum_{j=1}^{l} y_j^I = \sum_{j=1}^{l} f\left(x_j^I, a, b\right) \\ \sum_{k=1}^{n-l} y_k^{II} = \sum_{k=1}^{n-l} f\left(x_k^{II}, a, b\right) \end{cases}$$

Совместное решение этой системы определяет численное значение двух параметров a и b, подставив которые в $y=f(x,a,b)$, получим искомое эмпирическое соотношение.

2.3. Метод наименьших квадратов. Согласно методу наименьших квадратов, наилучшими параметрами a и b считаются те, для которых сумма квадратов уклонений минимальна:

$$F(a,b) = \sum_{i=1}^{n} \left[y_i - f(x_i, a, b) \right]^2 = \min .$$

Найдем частные производные функции $F(a,b)$ по варьируемым параметрам a и b:

$$\frac{\partial F(a,b)}{\partial a} = -2 \sum_{i=1}^{n} \left[y_i - f(x_i, a, b) \right] f_a'(x_i, a, b),$$
$$\frac{\partial F(a,b)}{\partial b} = -2 \sum_{i=1}^{n} \left[y_i - f(x_i, a, b) \right] f_b'(x_i, a, b).$$

В силу необходимого условия экстремума функции многих переменных, наилучшими значениями параметров a и b служат те, при которых частные производные этой функции по варьируемым параметрам обращаются в нуль:

$$\begin{cases} \dfrac{\partial F(a,b)}{\partial a} = 0, \\ \dfrac{\partial F(a,b)}{\partial b} = 0, \end{cases}$$

т.е.

$$\begin{cases} \sum_{i=1}^{n} [y_i - f(x_i, a, b)] f_a'(x_i, a, b) = 0, \\ \sum_{i=1}^{n} [y_i - f(x_i, a, b)] f_b'(x_i, a, b) = 0. \end{cases}$$

Решение этой системы относительно a и b и дает искомое наилучшее значение числовых параметров.

2.4. Метод преобразования координат. Рассмотрим в системе координат xOy некоторую нелинейную зависимость $y=f(x,a,b)$, непрерывную и монотонную на замкнутом отрезке $[x_1,x_n]$. Перейдем к новым переменным $q=\varphi(x)$ и $z=\psi(y)$ так, чтобы в новой системе координат qOz эмпирическая зависимость стала линейной: $z=aq+b$.

Точки N_i с координатами $[\varphi(x_i);\psi(y_i)]$ в плоскости qOz практически лежат на прямой линии. Обратно, если при построении на плоскости qOz окажется, что точки лежат на прямой линии, то между переменными q и z имеет место зависимость $\psi(y)=a\varphi(x)+b$.

Покажем, как некоторые нелинейные зависимости преобразованием координат можно свести к линейной.

1) Для показательной зависимости вида $y=ab^x$, логарифмируя, имеем $\lg y=x\lg b+\lg a$. Полагаем $\lg b=B$, $\lg a=A$, $\lg y=z$, $x=q$ и в плоскости qOz получим уравнение прямой $z=Bq+A$.

2) Дробно-линейную зависимость $y=1/(ax+b)$ также можно преобразовать в линейную.Действительно, найдем обратную зависимость $1/y=ax+b$ и введем

новые переменные $q=x$, $z=1/y$; тогда получим зависимость вида $z=aq+b$.

3) Для логарифмической зависимости $y=a\lg x+b$ введем новые переменные $q=\lg x$, $z=y$. Тогда снова получим линейную зависимость $z=aq+b$.

4) Рассмотрим степенную зависимость $y=bx^a$, где $a>0$, $b>0$. Логарифмируя, находим $\lg y=a\lg x+\lg b$, откуда, полагая $z=\lg y$, $q=\lg x$, $B=\lg b$, имеем $z=aq+B$. Прямую полученного вида удобно строить, если оси q и z в плоскости qOz взять в логарифмическом масштабе. Началом координат логарифмической сетки служит точка $q=0$, $z=B$.

5) В выражении $y=a+b/x$ произведем замену переменных $1/x=q$, $y=z$. Тем самым заданная гиперболическая зависимость преобразуется в линейную $z=a+bq$.

6) Рассмотрим дробно-рациональную функцию $y=x/(ax+b)$. Функция, обратная данной, имеет вид $1/y=a+b/x$. Введя новые координаты $z=1/y$, $q=1/x$, опять получим линейную зависимость $z=a+bq$.

Пример 2. Опытные данные занесены в следующую таблицу:

x	1	2	3	4	5	6	7	8	9
y	0,33	0,49	0,59	0,65	0,71	0,755	0,77	0,81	0,82

Составить эмпирическую формулу.

Решение. Построим график (рис. 2) и определим тип кривой так, как это было указано в пункте 1. Аналитическая функция, достаточно хорошо соответствующая таблице, есть дробно-рациональная функция вида $y=x/(ax+b)$.

Рассмотрим функцию $1/y=a+b/x$, обратную данной, и введем новые переменные $z=1/y$, $q=1/x$. Тогда в плоскости qOz получим линейную зависимость $z=a+bq$. Для того, чтобы найти коэффициенты q и z , составим таблицу:

q	1	0,5	0,333	0,25	0,2	0,166	0,143	0,125	0,111
z	3	2	1,66	1,54	1,41	1,32	1,3	1,24	1,22

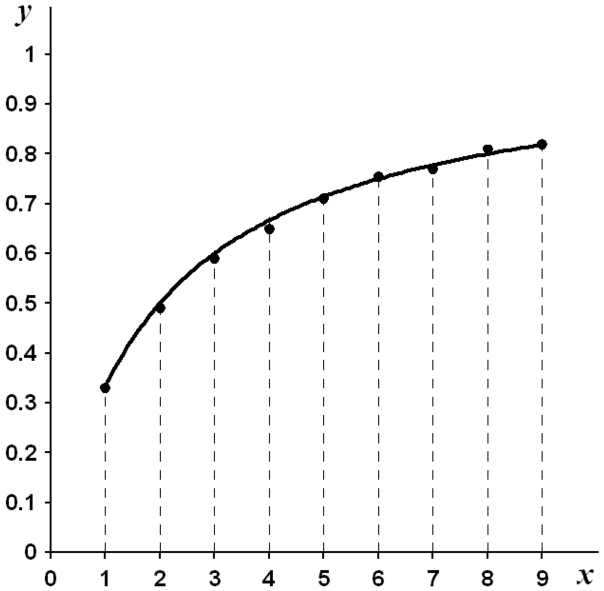

Рис. 2

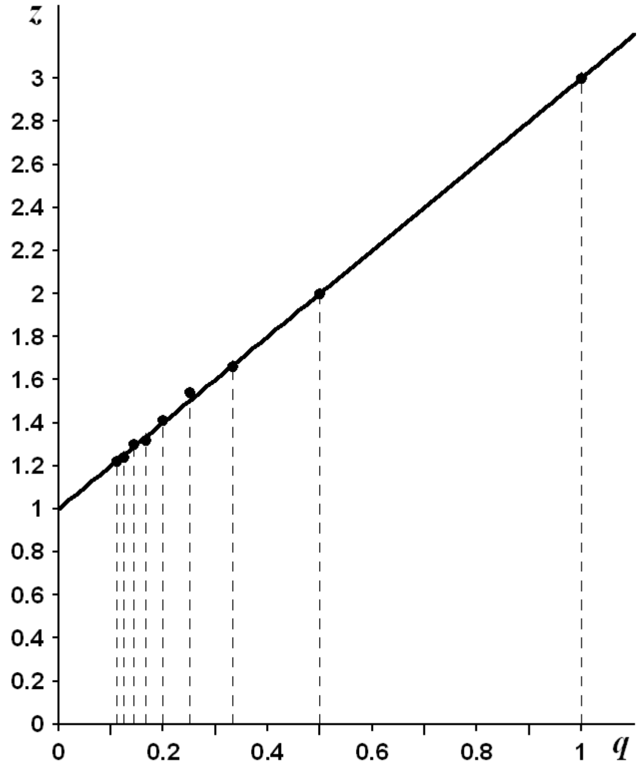

Рис. 3

Построим в плоскости *qOz* точки, соответствующие табличным значениям (рис. 3). Мы видим, что эти точки лежат на одной прямой. Следовательно, наши предположения о типе кривой оказались правильными. Остается определить параметры a и b: a=1; b=(3-1)/(1-0)=2. Таким образом, эмпирическая формула имеет вид $y=x/(x+2)$.

Порядок выполнения работы

1 Вынести на график дискретные значения y=f(x), согласно полученному варианту задания.

2 Построить график функции y=f(x), которая должна быть монотонной и однозначной.

3 Определить тип зависимости (см. п. 2.1 и пример 1)

4 Определить коэффициенты полученной зависимости по способу, указанному преподавателем (см. п.2.2 и пример 2).

Требования к отчету.

В отчете необходимо привести исходные данные, выполнить необходимые графические построения и расчеты (с подробными пояснениями !). В качестве основы следует ориентироваться на пример 1.