СПИСОК ЛИТЕРАТУРЫ

1. Биологические процессы и самоочищение на загрязненном участке реки (на примере Верхнего Днепра). – Минск.: Изд-во БГУ, 1987.

2. *Белолипетцкий В.М. , Шокин Ю.И.* Математические модели в задачах охраны окружающей среды. – Новосибирск: Издательство «ИНФОЛИО - пресс», 1997.

3. *Самарский А.А.* Введение в теорию разностных схем.- М.: Наука,1971.

4. *Самарский А.А., Вабищевич П.Н.* Численные методы решения задач конвекции-диффузии. – М.: Эдиторал УРСС, 1999.

5. *Вержбицкий В.М.* Основы численных методов: Учебник для вузов. -М.:Высш. шк., 2002.

6. Экологический мониторинг. Состояние окружающей природной среды Томской области в 1998 году. – Томск, 1999.

 Авторы статьи:

Михайлов Михаил Дмитриевич - ст. преп. каф. вычислительной математики и компьютерного моделирования Томского Государственного университета

Трапп Виктория Викторовна - дипломант каф. вычислительной математики и компьютерного моделирования Томского Государственного университета

УДК 622.1:528 (043.2)

В. И. Снетков

КРИТЕРИЙ ОПТИМАЛЬНОСТИ ПРИ АППРОКСИМАЦИИ ГЕОЛОГИЧЕСКИХ ПОКАЗАТЕЛЕЙ С ПОМОЩЬЮ РЯДОВ ФУРЬЕ

Для оценки тенденций в изменении геологического показателя по линии или в пространстве обычно применяют фильтрующие или аппроксимирующие процедуры, позволяющие уменьшить влияние случайной изменчивости. К ним относятся различного рода скользящие средние, полиномиальные функции, сплайны, гармонический анализ Фурье и другие [1-3, 5-7, 9]. Наибольшее распространение в геологии и горном деле получил метод Фурье, однако при его практическом применении существуют трудности, связанные с выбором числа гармоник, необходимых для правильного описания закономерного изменения изучаемого показателя.

Как известно, анализ Фурье применяется при исследовании временных процессов и представляет собой метод разложения сигнала или случайной функции на отдельные гармоники. В геологии и горном деле принято считать сигнал состоящим из трех частей: линейного *тренда*, различных *периодических* или циклических компонент и *случайной* компоненты. Тренд данных обнаруживается подбором соответствующей прямой регрессии. Считается, что тренд существует, если угол наклона прямой является значимым. Отметим, что дисперсионный анализ не всегда способен правильно решить эту задачу, и поэтому практика показала, что тренд необходимо всегда устранять вне зависимости от значимости коэффициентов регрессии. Устранение тренда обязательно для применения метода Фурье для того, чтобы осуществить преобразование исходного временного ряда в последовательность отклонений от подобранной прямой линии. Оставшаяся часть временного ряда состоит из *сигнала* (периодической компоненты) и *шума* (случайной компоненты).

Итак, одной из основных задач анализа Фурье является выделение главной периодической компоненты, что эквивалентно выделению закономерного изменения показателя из композиции в случайном процессе.

Основные положения анализа Фурье и принципы его применения следующие. Любая периодическая функцию может быть разложена в так называемый ряд Фурье [2,3], который для простых последовательностей имеет вид:

$$Y_i = \sum_{n=1}^{\infty}\left(\alpha_n Cos\frac{2n\pi x_i}{\lambda} + \beta_n Sin\frac{2n\pi x_i}{\lambda}\right),$$

где n – гармоническое число; α_n, β_n - коэффициенты гармоник; λ - длина волны основной гармоники.

В действительности, обычно величина λ неизвестна и в качестве основной длины волны выбирают произвольное значение L. Если выбрать L бóльшим или равным длины изучаемого ряда данных и вычислить достаточное число гармоник, то получим оценку периодичности, присутствующей в наших данных. Произвольный выбор значения L скорее всего будет неправильной оценкой длины волны λ, поэтому коэффициенты произвольной основной гармоники могут оказаться близкими к нулю. Коэффициенты всех последовательных гармоник, которые на самом деле не присутствуют в данных, будут также близки к нулю. Поэтому только действительно имеющиеся длины волн будут входить с достаточно большими коэффициентами.

Неизвестные коэффициенты α_n, β_n определяют с помощью решения системы нормальных уравнений, аналогичных используемым в регрессионном анализе [3]. Если же наблюдения Y_i сделаны через равные интервалы значений переменной X_i, то их поиск существенно упрощается:

$$\alpha_n = \frac{2}{N}\sum_{i=1}^{N} Y_i Cos\frac{2n\pi x_i}{\lambda},$$

$$\beta_n = \frac{2}{N}\sum_{i=1}^{N} Y_i Sin\frac{2n\pi x_i}{\lambda}.$$

Для конечных последовательностей, состоящих из N равномерно расположенных наблюдений, можно вычислить N/2 гармоник. Величина $N/2$ носит название частота Найквиста. Каждая волна, имеющая частоту Найквиста, определяется на основании только трех выборочных значений. Известно, что синусоидальную волну невозможно определить по числу точек, меньшему трех, и $N/2$ – максимальная частота, которую можно определить таким образом. Длина волны, имеющей частоту Найквиста, равна 2Δ, где Δ — расстояние между последовательными наблюдениями.

Амплитуда $n-$ой гармоники и её фазовый угол определяются через найденные коэффициенты по формулам:

$$A_n = \frac{1}{2}\sqrt{\alpha_n^2 + \beta_n^2}, \quad \varphi_n = arc\left[tg\left(\frac{-\beta_n}{\alpha_n}\right)\right].$$

В общем случае коэффициенты связаны с *энергией* или *дисперсией* каждой гармоники, определяемой равенством:

$$\sigma_n^2 = \alpha_n^2 + \beta_n^2.$$

Если σ_0^2 - общая дисперсия Y в ряде, то вклад n-й гармоники равен

$$\frac{\sigma_n^2}{\sigma_0^2}.100\%.$$

Таким образом, это дает возможность проверки значимости различных гармоник критериями типа ANOVA [1,3]. На практике обычно так и поступают. Из разложения с помощью дисперсионного анализа выбираются значимые гармоники, которые включаются в аппроксимирующую функцию. Отметим, что значимость и не значимость n-ой гармоники в значительной мере определяется из субъективных соображений. Это связано, прежде всего, с тем, что величина критерия Фишера зависит от принятого уровня значимости α, выбор которого всегда субъективен. При разных уровнях значимости одна и та же гармоника может оказаться значимой и не значимой. В итоге два исследователя с разными взглядами на природу изучаемого процесса и разным уровнем подготовки в области статистических критериев могут получить существенно отличающиеся результаты. В геологии это особенно опасно, так как при таком подходе обычно не уделяется должного внимания высокочастотным гармоникам, играющим важную роль при решении конкретных задач горного дела и геологии.

Дальнейшие действия заключаются в последовательном наращивании числа гармоник и оценке значимости вклада каждой при помощи F-критерия Фишера. Расширение первоначального уравнения с помощью добавления следующих гармоник приводит к тому, что график начинает искривляться и постепенно приближаться к исходной последовательности. Одна дополнительная гармоника заставляет основную волну изменить наклон, вторая приводит к возникновению нескольких точек перегиба и т.д. Увеличивающаяся искривленность позволяет линии более точно подходить к исходным данным.

Как видим, все действия сводятся к тому, чтобы подобрать кривую, наилучшим образом описывающую фактическое изменение изучаемого показателя, причём это стремление в пределе имеет абсурдный результат. Действительно, если число гармоник достигнет $N/2$, то линия регрессии пройдет точно через каждую точку исходной случайной функции. В построении такой линии мало смысла, так как она не является более эффективной, чем сами исходные данные.

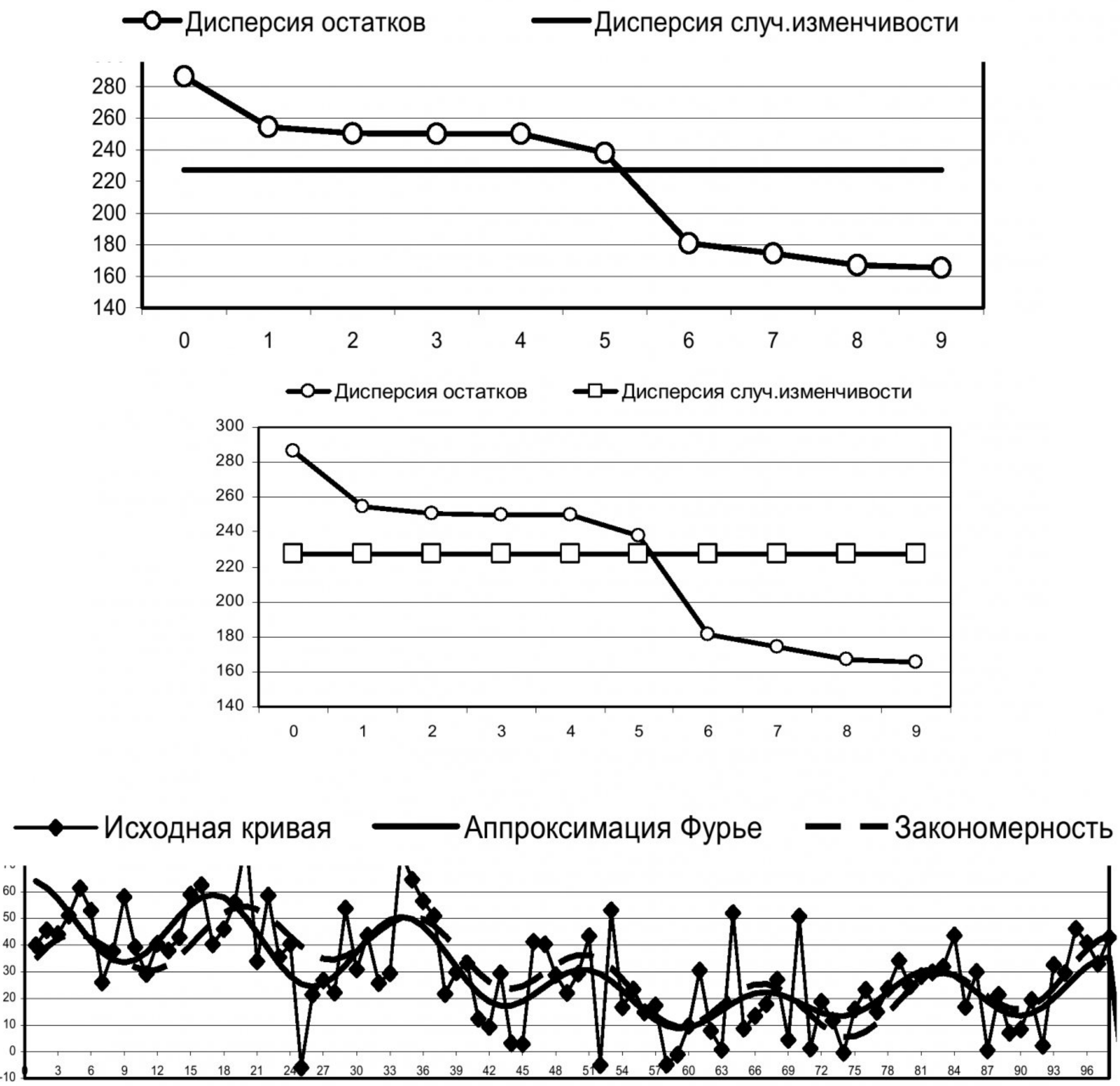
Цель геометризации геологического показателя принципиально отлична от цели полиномиальной аппроксимации и заключается в выявлении и последующей оценке закономерного изменения изучаемого показателя Y_i.

В этом случае надо стремиться не к минимальной дисперсии остатков, а к тому, чтобы дисперсия остатков равнялась бы дисперсии случайной изменчивости σ_C^2. Последняя может быть найдена по одному из нижеприведённых выражений [2,4]:

$$\sigma_{\Delta'}^2 = \frac{\sum_{i=1}^{k}(\Delta'_i)^2}{2k}, \quad \sigma_{\Delta''}^2 = \frac{\sum_{i=1}^{m}(\Delta''_i)^2}{6p},$$

$$\sigma_r^2 = \sigma^2\left[1 - r(1)\right], \quad \sigma_c^2 = \frac{m\left(\sigma^2 - \sigma_{\overline{x}(m)}^2\right)}{m-1},$$

где Δ' - первые разности, $\Delta'_i = x_{i+1} - x_i$; k – количество первых разностей, $k=n-1$; $p=n-2$ – ко-



Энергетический спектр, зависимость дисперсии остатков от числа аппроксимирующих гармоник и модель с результатами реконструкции закономерного изменения (на основе критерия равенства дисперсии остатков и дисперсии случайной изменчивости)

личество вторых разностей; *r(1)* – коэффициент автокорреляции при единичном лаге; $\sigma^2_{\bar{x}(m)}$ - дисперсия сглаженной последовательности по *m* пробам.

Из этого следует, что <u>критерием при выборе необходимого числа гармоник должен быть паритет дисперсии остатков и дисперсии случайной изменчивости</u>. Достижение этого условия автоматически приводит к выполнению основного равенства (1) в математической статистике, а значит и к адекватному отражению закономерной компоненты в изучаемом показателе Y_i.

$$\sigma^2_o = \sigma^2_з + \sigma^2_c \quad , \qquad (1)$$

где $\sigma^2_з$ - дисперсия закономерного изменения показателя; σ^2_c - дисперсия случайной изменчивости показателя; σ^2_o - общая дисперсия.

Применение в качестве критерия дисперсии или стандарта случайной изменчивости, полностью решает проблему неоднозначности принимаемых решений, адекватности аппроксимированной модели закономерному изменению изучаемого показателя, а также и проблему непрерывности и автоматизации процесса вычислений.

Таким образом, методическая основа аппроксимации гармониками Фурье может быть сформулирована в виде следующего алгоритма:

- по изучаемой последовательности Y_i определяют параметры линейного тренда, который исключают, получая новую последовательность Z_i ;

- по вновь полученной последовательности определяют общую дисперсию σ^2 и дисперсию случайной изменчивости σ^2_c ;

- вычисляют все *N/2* гармоник и определяют их коэффициенты;

- выполняют аппроксимацию последовательности Z_i, используя для этого параметры найденных гармоник. Результатом данной операции будет последовательность сглаженных значений $\hat{Z}_i$.

Аппроксимация в свою очередь может быть осуществлена следующими способами:

- последовательным наращиванием числа найденных гармоник до достижения определённого условия;
- сортировкой полученных гармоник в порядке возрастания и далее последовательным их наращиванием до достижения определённого условия;
- условием достаточности аппроксимации должно быть равенство дисперсии отклонений $\left(Z_i - \hat{Z}_I\right)$ дисперсии случайной изменчивости σ^2_c ;
- после аппроксимации к последовательности $\hat{Z}_i$ добавляется исключённый линейный тренд, в результате получают реконструированный закономерный динамический ряд $\hat{Y}$.

Покажем эффективность данного подхода на статистических моделях. В качестве примера возьмём сложную гармоническую модель, приведённую на рисунке. Длина последовательности в рассматриваемом примере равна 98 значений. Линейный тренд для этой модели, найденный по способу наименьших квадратов, имеет вид:

$$Y'_i=44.6796-0.2985\cdot X_i \text{ ..}$$

После исключения линейного тренда дисперсия общей изменчивости по новой последовательности Z_i равна σ^2 =266.480 , дисперсия случайной изменчивости σ^2_c =227.312 , квадрат корреляционного отношения $\eta^2_з$ =0.21, то есть доля гармонической изменчивости составляет 21%. На том же рисунке показан энергетический спектр последовательности Z_i , на котором видны 2 выдающиеся гармоники: первая гармоника с периодом 98 значений и шестая с периодом $\lambda=98/6\sim16.3$ значений (фактические параметры $\lambda_1=94$ и $\lambda_2=16.7$). Гармоники 2-5 являются по всем классическим критериям незначимыми, однако они содержат часть дисперсии из обеих гармоник, поскольку длина основной волны была выбрана произвольно и не соответствовала фактической. Поэтому последовательное наращивание гармоник с первой по шестую даст сложение их дисперсий и приблизит к дисперсии закономерной изменчивости. Критерием, как уже отмечалось, должно быть равенство дисперсии отклонений $\left(Z_i - \hat{Z}_I\right)$ и дисперсии случайной изменчивости.

На рисунке показано соотношение стандарта случайной изменчивости и стандарта остатков в зависимости от числа гармоник. Между 5 и 6 гармониками наблюдается их пересечение, то есть равенство остаточной дисперсии и дисперсии случайной изменчивости. Исследования показали, что в таких случаях наибольшая точность достигается если остановиться на следующей гармонике после пересечения графиков, то есть на шестой. При этом наблюдается минимум дисперсии отклонений аппроксимированных значений от исходного закономерного изменения и, таким образом, обеспечивается наибольшая степень приближения аппроксимирующей функции и математическое ожидание отклонений, близкое к нулю.

Разработанная методика Фурье анализа в совокупности с критерием оптимизации подбора числа гармоник для сглаживания последовательностей были протестированы на многих моделях и фактическом материале [8] и показали эффективность данного подхода, высокую вероятность выявления характера закономерного изменения даже в условиях существенного преобладания случайной изменчивости над закономерной.

Выводы

Разработанный критерий аппроксимации позволяет объективно производить разделение закономерного и случайного изменения показателя и рекомендуется к использованию наряду с традиционными способами, основанными на дисперсионном анализе с применением статистических критериев.

СПИСОК ЛИТЕРАТУРЫ

1. *Андерсон Т.* Статистический анализ временных рядов /Перевод с английского И.Г. Журбенко, В.П. Носко. Под. ред. Ю.К. Беляева. – М.: Мир, 1976. – 755 с.
2. *Базанов Г.А., Снетков В.И.* Анализ распределения показателей методом статистических испытаний. – В кн.: Исследования по проблемам геодезии и маркшейдерского дела на горнодобывающих предприятиях Восточной Сибири, Межвузовский сборник, №1. – Иркутск, 1976, с.14-29.
3. *Дэвис Дж.* Статистика и анализ геологических данных. / Под ред. Д.А. Родионова. – М.: Мир, 1977. – 572 с.
4. *Попов Е.И.* К оценке точности изображения залежи полезного ископаемого по данным разведки. – В кн.: Зап. ЛГИ. – Л. 1959, т.36, вып.2, с.178-189.
5. *Снетков В.И.* Теория сглаживания показателей месторождения независимым окном // Вопросы совершенствования маркшейдерско-геодезических работ. Межвуз. сб. - Л.: ЛГИ,1982, с. 104 – 107.

6. *Снетков В.И., Снеткова А.В.* Обоснование размера фильтра при сглаживании горногеологических показателей способом скользящего среднего // Проблемы развития минеральной базы Восточной Сибири. Сб. научн. тр. – Иркутск, 2003, с.90-94.

7. *Снетков В.И., Снеткова А.В.* Теоретические основы сглаживания статистических рядов способом скользящей средней // Проблемы развития минеральной базы Восточной Сибири. Сб. научн. тр. – Иркутск, 2003, с. 111-120.

8. *Снетков В.И.* Динамическое прогнозирование параметров подсчёта запасов по слоям на глубину для жильных пегматитовых месторождений. - Маркшейдерский вестник. - 2005. -№2. – с. 36 - 41.

9. *Четыркин Е.М.* Статистические методы прогнозирования. - М.: Статистика, 1975. - 184 с.

Автор статьи:

Снетков
Вячеслав Иванович
- доц. каф. маркшейдерского дела
Иркутского Государственного технического университета

УДК 519. 21

А.В. Бирюков, П.А. Бирюков

ДИНАМИЧЕСКИЕ МОНОПОЛИИ В РЕГУЛЯРНЫХ ГРАФАХ

Пусть G – связной граф с вершинами $x_1, \ldots, x_n$. Предположим, что в любой из моментов времени t=0,1,2,…каждая вершина может находиться в одном из двух состояний 0,1. Обозначим состояние вершины x_i в момент t через $x_i(t)$. Вектор $x_i(t)$. =[$x_1(t),\ldots, x_n(t)$] назовем состоянием графа G в момент t.

Вершина x_i и все смежные с ней вершины образуют окрестность $B(x_i)$. Зададим правило перехода от состояния $x(t)$ к состоянию $x(t+1)$:

(**T**): Если в момент t не менее половины вершин из окрестности $B(x_i)$ находились в состоянии 1, то $x_i(t+1)=1$; в противном случае $x_i(t+1)=0$.

Это правило определяет на графе G дискретную динамическую систему с пространством состояний $\{0,1\}^n$. Произвольно заданное начальное состояние $x(0)$ порождает траекторию, т. е. последовательность состояний, в которой каждое состояние $x(t+1)$ получается из предыдущего состояния $x(t)$ по правилу перехода (T). Поскольку число возможных состояний конечно, любая траектория становится μ - периодической, начиная с некоторого t. В [1] показано, что период любой траектории равен 1 или 2. Это означает, что с некоторого момента t граф либо переходит в стационарное состояние $x(t)$ $=x(t+1) =\ldots$, либо принимает два состояния поочередно.

Множество тех вершин, начальное состояние которых равно 1, называется динамической монополией (динамо) в графе G [2], если соответствующая траектория выходит на стационарное состояние (1,1, … ,1). Пусть $m(G)$ - наименьшее число вершин, образующих динамо в графе G, и $\mu(G)=m(G)/n$, где n – порядок G. В [2] для любого k построен граф G_k порядка $2k+3$, для которого $m(G_n)=3$. Таким образом, $\mu(G_k)\to 0$ при $k\to\infty$. Отметим, что граф G_k содержит вершины степени 1, 2, 3, k. Можно предположить, что в классе всех регулярных графов степени r значение $\mu(G)$ ограничено снизу некоторой положительной константой (зависящей от r).Мы докажем эту гипотезу для $r \le 4$. Случай $r=1$ тривиален. При $r=2$ граф G является простым циклом и легко проверить, что $\mu(G)>1/2$.

Пусть M – динамо в регулярном графе G степени $r \le 4$. Тогда любой цикл C графа G должен содержать хотя бы одну вершину из множества M. Действительно, если начальное состояние всех вершин цикла равно 0, то для любой вершины $x_i \in C$ по крайней мере три вершины из окрестности $B(x_i)$ находятся в состоянии 0 при t=0. Поскольку окрестность содержит не более пяти вершин, по правилу перехода вершина x_i остается при t=1 в состоянии 0. Аналогично, все вершины цикла остаются в состоянии 0 в любой момент t, что противоречит условию. Таким образом, хотя бы одна вершина цикла должна иметь начальное состояние 1. Тогда граф G-M , полученный удалением из G всех вершин множества M, не содержит циклов, т. е. представляет собой лес.

Пусть n – порядок графа G, m – число вершин множества M и q – число ребер графа G-M. Так как число ребер леса меньше числа его вершин [3], то $q<n-m$. С другой стороны, граф G имеет $r\cdot n/2$ ребер и удаление m вершин приводит к удалению не более, чем $r\cdot m/2$