

МЕТОД РЕКУРСИВНОГО ПРОДОЛЖЕНИЯ МАССИВОВ ДАНЫХ СОСТАВА СЫПУЧЕГО МАТЕРИАЛА

Ариненков Ю.Д.

Кафедра ОПИ, факультет ЭМА, ДонНТУ

opi@fizmet.dgtu.donetsk.ua

Abstract

Arinenkov Y.D. Method of recursive continuing the arrays of data of composition of loose material. Put problem to install a correspondence to recursive methods a continuing the arrays of composition of loose material to positions of theory of interpolating in classical production with the instruction of area of such correspondence to the motivation of description technologies and is given its decision.

Постановка проблемы

Развитие научных основ сепарации и создание систем технологической информации с целью повышения эффективности переработки угля, особенно в условиях кризисного состояния горнодобывающей промышленности, является важнейшей научно-технической проблемой, решение которой, согласно Закону Украины об энергосбережении от 1.07.1994 г. № 74/94-ВР, направлено на обеспечение экономической и энергетической безопасности страны. Важной частью этой общей задачи является выбор и обоснование методов описания характеристик сыпучих материалов.

Анализ исследований по проблеме

Гранулометрические и фракционные характеристики формируются и изменяются на протяжении всего цикла операций добычи, транспортировки и переработки угля. Результаты измерения этих характеристик представляют таблицами из 3 - 7 строк, т.е. короткими массивами данных, причём, сами данные, полученные в соответствии с методиками отбора и обработки представительных проб, рассматривают как точные. Попытки авторов соответствующих исследований, например [1], найти подходящие функции для описания таких массивов не привели к какой-либо универсальной рекомендации по решению задачи. Поэтому автор в работе [2] предложил принцип "as is" (AI) как отказ от поиска несуществующих аналитических закономерностей для характеристик указанного типа в пользу применения к ним интерполяции на основе методов *knot-* и *area-* продолжения массива.

Сущность метода *knot*-продолжения массива данных распределения крупности, в отличие от интерактивного задания кривой [3] или сепарабельного программирования [4], состоит в том, что на каждом интервале строится доверительный треугольник, образуемый стягивающей парой точек хордой и продолжениями хорд смежных интервалов, и если вершина этого треугольника не выходит за границы интерполируемого интервала, то с некоторой степенью правдоподобия новая опорная точка назначается в центре тяжести треугольника, иначе - посередине стягивающей хорды как точка перегиба характеристики. Метод *area*-продолжения массива усреднённых по интервалам данных фракционного состава [5] является развитием метода W. MacLarten [6], устанавливающего начало и конец интерполирующей 2-кусочно-линейной функции посередине ступенек фракций с сохранением площади каждой при делении её на две части.

На основе толерантного применения предложенных методов разработаны вопросы форматирования массивов [7] и оценки ошибок интерполяции изучаемых "гипотетических характеристик" [8]. Назрела необходимость проверить соответствие получаемых таким путём результатов основам теории интерполирования. В свою очередь, позаимствовав из неё методологические приёмы, представленные в символах и определениях этой теории, метод рекурсивного продолжения массивов может заполнить пустующую в ней нишу интерполяции неизвестной "гипотетической функции" с переменными интервалами данных.

Постановка задачи

В теории интерполирования основная задача формулируется как требование построить функцию $F(x)$, принадлежащую к известному классу и принимающую в узлах интерполяции x_0, x_1, \dots, x_n те же значения y_0, y_1, \dots, y_n , что и заменяемая интерполяцией функция $f(x)$, т.е. такую, что $F(x_0) = y_0, F(x_1) = y_1, \dots, F(x_n) = y_n$ [9]. Аппарат решения задачи использует конечные разности различных порядков для равных интервалов представления значений функции $f(x)$ с постоянным шагом и разделённые разности или разностные отношения для таблиц с переменным шагом. В качестве функции известного класса $F(x)$ принимают полином n -ого порядка, для нахождения которого фундаментально разработаны интерполяционные формулы Ньютона, Гаусса, Стирлинга, Бесселя для равноотстоящих узлов и Лагранжа для произвольно заданных узлов [10]. Сплайн-интерполяция ориентирована уже только на кусочно-заданный полином. Здесь ставится задача установить соответствие методов рекурсивного продолжения массивов основным положениям теории интерполирования в классической постановке с указанием области такого

соответствия. Решение этой задачи позволит дать обоснование корректности сепарабельного описания сепарационной технологии.

Метод решения задачи

При решении интерполяционных задач с помощью известной (по определению) функции $f(x)$ вычисляют опорные точки $y_0 = f(x_0)$, $y_1 = f(x_1)$, ..., $y_n = f(x_n)$ с постоянным или с переменным шагом. В последующем, однако, эта функция никак не используется. Следовательно, отсутствие функционального описания конкретной зависимости не является препятствием для интерполяции такой гипотетической зависимости $f_r(x)$, если имеется сам массив элементов $y_k(x_k)$, $k = 1, 2, \dots, n$, принадлежащий этой зависимости.

В теории интерполяции используется аппарат конечных разностей и разностных отношений (или разделенных разностей) [9]. Здесь сразу нужно отметить, что нисходящее разностное отношение первого порядка в точности равно коэффициенту наклона хорды на участке между точками (x_k, y_k) и (x_{k+1}, y_{k+1}) , поэтому пронумеруем его по правой, конечной в порядке счёта, точке интервала как

$$A_{k+1} = \frac{y_{k+1} - y_k}{x_{k+1} - x_k} = \frac{\Delta y_k}{\Delta x_k}, \quad (1)$$

а восходящее разностное отношение первого порядка соответствует коэффициенту наклона хорды на аппроксимируемом участке между точками (x_{k-1}, y_{k-1}) и (x_k, y_k) , пронумеруем его также по правой, конечной в порядке счёта, точке интервала как

$$A_k = \frac{y_k - y_{k-1}}{x_k - x_{k-1}} = \frac{\Delta y_{k-1}}{\Delta x_{k-1}}. \quad (2)$$

Очевидно, по аналогии с уже рассмотренными понятиями восходящего и нисходящего разностных отношений, можно ввести понятие открывающего разностного отношения первого порядка, которому соответствует коэффициент наклона хорды на участке между точками (x_{k-2}, y_{k-2}) и (x_{k-1}, y_{k-1}) , пронумеровав его также по правой точке интервала как

$$A_{k-1} = \frac{y_{k-1} - y_{k-2}}{x_{k-1} - x_{k-2}} = \frac{\Delta y_{k-2}}{\Delta x_{k-2}}. \quad (3)$$

Формулы (1) - (3) отождествляют разностные отношения первого порядка с коэффициентами наклона хорд. Теперь, согласно [2],

рассматривая интервал между опорными точками (x_{k-1}, y_{k-1}) и (x_k, y_k) как интерполируемый, а хорду на этом интервале с наклоном по (2) - как основание доверительного треугольника, через указанные опорные точки можем провести прямые с наклонами соответственно по (3) и (1), точка пересечения которых определит искомую вершину этого треугольника.

Для размещения новых опорных точек исходный массив предварительно вдвое продолжается умножением всех индексов на множитель 2. Тогда старые опорные точки получают чётные значения индексов, а для новых опорных точек будут зарезервированы нечётные значения индексов.

Таким образом, если исходный массив будет продолжен вдвое и получит индексы при $k = 0, 2, 4, \dots$ - чётном, то формулы (3), (2), (1) нужно будет записать с удвоенными смещениями следующим образом:

$$A_{k-2} = \frac{y_{k-2} - y_{k-4}}{x_{k-2} - x_{k-4}} = \frac{\Delta y_{k-4}}{\Delta x_{k-4}}, \quad (4)$$

$$A_k = \frac{y_k - y_{k-2}}{x_k - x_{k-2}} = \frac{\Delta y_{k-2}}{\Delta x_{k-2}}, \quad (5)$$

$$A_{k+2} = \frac{y_{k+2} - y_k}{x_{k+2} - x_k} = \frac{\Delta y_k}{\Delta x_k}. \quad (6)$$

В работе [5] даны формулы вычисления координат вершины доверительного треугольника в выражении только через четыре опорные точки. Здесь даём тот же вывод, но в обозначениях открывающего и нисходящего разностных отношений первого порядка по (4) - (6). Для этого составляем систему двух уравнений описания смежных с интерполируемым участком хорд, продолженных до точки их пересечения (x_c, y_c) ,

$$\left. \begin{aligned} y_c &= y_{k-2} + A_{k-2} \cdot (x_c - x_{k-2}) \\ y_c &= y_k + A_{k+2} \cdot (x_c - x_k) \end{aligned} \right\}, \quad (7)$$

и, исключая вначале y_c , с учётом подстановки числителя правой части из определения восходящего разностного отношения первого порядка (5), получаем величину аргумента в искомой точке

$$x_c = \frac{A_k - A_{k-2}}{A_{k+2} - A_{k-2}} \cdot x_{k-2} + \frac{A_{k+2} - A_k}{A_{k+2} - A_{k-2}} \cdot x_k. \quad (8)$$

Теперь, вместо громоздких преобразований подробного вывода формулы для вычисления функции y_c , воспользуемся свойством функциональной обратимости линейного описания, по которому всегда будет верно и $y(x)$, и $x(y)$, и по аналогии с предыдущей формулой запишем сразу окончательное выражение для искомой величины функции в точке пересечения

$$y_c = \frac{A_k^{-1} - A_{k-2}^{-1}}{A_{k+2}^{-1} - A_{k-2}^{-1}} \cdot y_{k-2} + \frac{A_{k+2}^{-1} - A_k^{-1}}{A_{k+2}^{-1} - A_{k-2}^{-1}} \cdot y_k. \quad (9)$$

Обратим внимание на то, что обе зависимости (8) и (9) следуют из одной системы уравнений (7), поэтому будет верно и $y_c(x_c)$, и $x_c(y_c)$, т.е. (7) нужно рассматривать как рекурсивную функцию [11], обеспечивающую указанное соответствие.

Искомая новая опорная точка лежит на медиане, проведенной из полученной вершины доверительного треугольника (x_c, y_c) к середине противоположной ей стороны этого треугольника в точке (x_a, y_a) , имеющей следующие координаты:

$$x_a = \frac{x_k + x_{k-2}}{2}, \quad (10)$$

$$y_a = \frac{y_k + y_{k-2}}{2}. \quad (11)$$

В качестве коэффициента компромисса при назначении новой опорной точки между указанными точками c и a используется коэффициент гладкости функции U , $0 \leq U \leq 1$, с учётом которого вычисление выполняется по формулам:

$$x_{k-1} = x_a \cdot (1 - U) + x_c \cdot U, \quad (12)$$

$$y_{k-1} = y_a \cdot (1 - U) + y_c \cdot U. \quad (13)$$

При $U=0$ новые опорные точки будут назначаться всегда точно посередине стягивающей хорды каждого доверительного треугольника (точка a) во всех рекурсиях, следовательно, сглаживания функции не произойдёт (функция будет представлена вписанными многоугольниками с вершинами точно в точках первичного массива). В работе [5] при продолжении выборочного массива функции Стьюдента (по целым табличным значениям t_n) интерполяция обеспечивала точность в 3 значащих цифры при коэффициенте гладкости в пределах $0,25 \leq U \leq 0,275$.

Там же дано подробное изложение метода *area*-продолжения массива более общего вида - колеблющейся характеристики, и выполнено его сравнение с методом *knot*-продолжения массива. Сравнение основано на том, что последовательность восходящих разностных отношений первого порядка из массива для *knot*-продолжения (его производная) являются исходным массивом (первичным ступенчатым графиком) для осуществления *area*-продолжения массива. Полученный результат интегрируется и сопоставляется с результатом *knot*-продолжения массива. Исследование показало, что существенные расхождения могут проявляться только на нулевом и конечном интервалах, но заданием соответствующих производных в начале и конце массива эти расхождения эффективно устраняются. В то же время, установлено, что вычисленные в первом *area*-продолжении массива значения функции изменяются во втором и т.д., но таким образом, что колебательно стремятся к некоторым, очевидно, правдоподобным, пределам на своих интервалах. Проведенный анализ методов позволяет систематизировать их по следующим признакам: по заданной функции (полиномиальная интерполяция), по кусочно-заданному полиному (сплайн-интерполяция) и по степени правдоподобия получаемых решений (рекурсивное продолжение массивов данных).

Основные результаты и применения

Если в (8) ввести обозначения

$$\psi_{k-2} = \frac{A_k - A_{k-2}}{A_{k+2} - A_{k-2}}, \quad (14)$$

$$\psi_k = \frac{A_{k+2} - A_k}{A_{k+2} - A_{k-2}}, \quad (15)$$

где ψ_k и ψ_{k-2} - коэффициенты веса аргументов x_k и x_{k-2} соответственно, то сразу станет очевидным, что должно соблюдаться равенство $\psi_k + \psi_{k-2} = 1$, благодаря которому упрощается контроль отсутствия точки перегиба в продолжаемом интервале проверкой соблюдения условия $0 < \psi_{k-2} < 1$, и тогда новая опорная точка назначается по (10), (11), (12), (13), иначе - только по (10) и (11) как точка перегиба. Это условие дано следующей формулой:

$$U = \begin{cases} U_{\text{зад.}} & \text{если } 0 < \psi_{k-2} \text{ и } \psi_{k-2} < 1, \\ 0 & \text{если } \psi_{k-2} \leq 0 \text{ или } 1 \leq \psi_{k-2}, \end{cases} \quad (16)$$

где $U_{\text{зад.}}$ - заданная величина коэффициента плавности продолжения характеристики.

При этом вычисления выполняются в следующем порядке перечисления номеров формул: (14), (8) - (11), (16), (12), (13).

Вычислением в цикле для всех k - чётных завершается первая рекурсия. Для следующей рекурсии все номера элементов массива удваиваются и вычислительный процесс повторяется.

Вместе с тем, для продолжения массива на первом/последнем интервале требуется дополнительная информация о наклоне смежной хорды слева/справа. В этом случае имеем следующие формулы описания прогнозируемых начальных условий из [5]:

$$A_0 = \left(\frac{y_2 - y_0}{x_2 - x_0} \right)^2 \cdot \frac{x_4 - x_2}{y_4 - y_2}, \quad (17)$$

$$A_{n+2} = \left(\frac{y_n - y_{n-2}}{x_n - x_{n-2}} \right)^2 \cdot \frac{x_{n-2} - x_{n-4}}{y_{n-2} - y_{n-4}}, \quad (18)$$

где A_0 и A_{n+2} - прогнозируемые производные гипотетической функции $y_i(x)$ по аргументу x в начале и конце интервала этой функции, т.е. в точках 0 и n соответственно.

Ориентация на степень правдоподобия при назначении новых опорных точек и применение экстраполяции концевых производных по (17) и (18) дают основания говорить о распознавании гипотетической функции (здесь заданной $y=10 \cdot x^{1/2}$, см. рис.1).

Это исследование выполнено на массиве $x=(0; 25; 50; 100\text{мм})$ и $y=(0; 50; 70,71; 100\%)$. Здесь видно, что на участке 0 - 25мм аппроксимация завершилась неудовлетворительно и по *knot*-ному, и по методу Лагранжа. Ошибки, особенно большие, наблюдаются на концах интервала. Улучшим интерполяцию уменьшением концевых интервалов массива (рис.2).

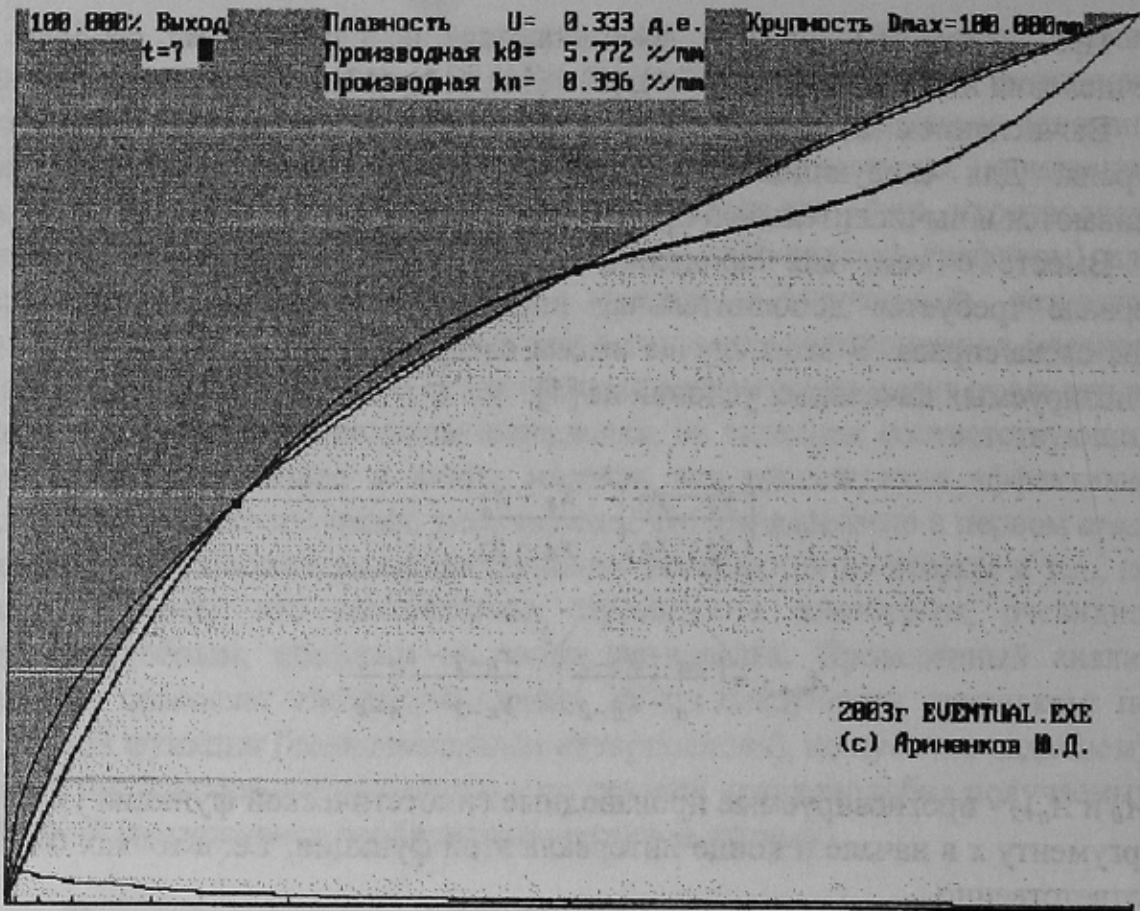


Рисунок 1 - Продолжение короткого массива $n=3$ от известной функции $y=10 \cdot x^{1/2}$

Теперь исследование выполнено при улучшенном выборе узлов массива $x=(0; 3; 50; 95; 100\text{мм})$ и $y=(0; 17,32; 70,71; 97,47; 100\%)$. Аппроксимация завершилась вполне удовлетворительно по *knot*-ному, по методу Лагранжа ошибки существенно возросли.

По этим кривым видно, что *knot*-ном “распознаёт” гипотетическую для него зависимость по производным на смежных интервалах, подстраивается под неё, в то время как полином по методу Лагранжа на коротком массиве работает неудовлетворительно. Так, для *knot*-нома наибольшая наблюдаемая ошибка, имеющая место на начальном интервале (где производная равна $+\infty$), не превышает 1,5%, а полином, полученный по методу Лагранжа, на втором интервале даёт недопустимую ошибку +23%.

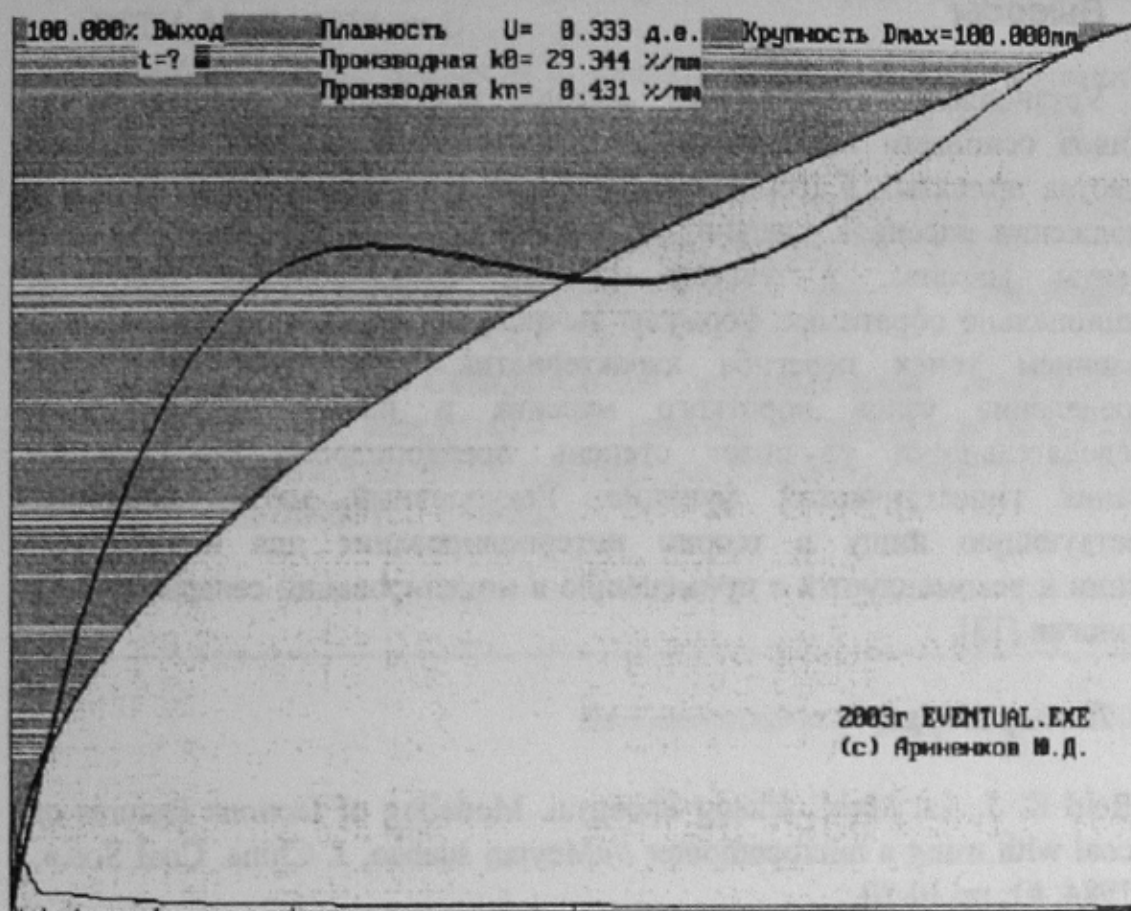


Рисунок 2 - Продолжение короткого массива $n=4$ от известной функции $y=10 \cdot x^{1/2}$

Из данного примера следует, что для повышения точности экстраполяции при определении наклонов характеристики в начале и конце массива необходимо первый и последний интервалы принимать малыми (см. также рис.1 и 2, графики плотности распределения крупности по *knot*-ному). В теории наилучшего выбора узлов интерполяции обнаруживаем, что сгущать узлы у концов отрезка интерполяции - это решение Чебышева П.Л. [12], которое, как показано здесь, оказывается эффективным в применении именно к *knot*-ному, с наибольшей степенью правдоподобия отображающему неизвестную (гипотетическую) функцию интерполируемого короткого массива данных granulометрической характеристики сыпучего материала.

Выводы

Установлено соответствие метода рекурсивного продолжения массивов основным положениям теории интерполирования. На основе тождества принятых в теории интерполяции разностных отношений и в продолжении массивов - величины наклона хорд, стягивающих смежные элементы массива, с учётом степени правдоподобия получены функционально обратимые формулы вычисления новых опорных точек с выявлением точек перегиба характеристик. Повышение плотности распределения узлов короткого массива в начале и конце их последовательности улучшает степень правдоподобия рекурсивного описания гипотетической функции. Рекурсивный метод заполняет существующую нишу в теории интерполирования для неизвестной функции и рекомендуется к применению в моделировании сепарационной технологии [13].

Литература

1. Reid K. J., Lu Maixi, Zhang Shengsui. Modeling of factions features of coal with using a microcomputer // «Meitan suabao, J. China. Coal Soc.», 1984, #1, pp.10-20.
2. Arinenkov Y. Universal model for research both optimization of technological processes of coal preparation factories on the COMPUTER / 2nd Regional APCOM'97 Symposium. Published by: The Moscow State Mining University Publishing Center. Moscow, Russia - 1997. -P. 209-214.
3. Techniques Graphiques Interactives et C.A.O. our Michel Lucas of et Yvon Gardan. Hermes Publishing (France). 1983.
4. Miller C.E. (1963). The simplex method for local separable programming, in R.L. Graves and P. Wolfe (eds), Recent Advances in Mathematical Programming. McGraw-Hill, New York, pp. 89 - 100.
5. Ариненков Ю.Д. Начальные условия в сепарабельном описании характеристик сыпучих материалов // Вісті Донецького гірничого інституту. -№1, 2003. -С. 41-46.
6. Maclarren W. A guarantee of performance for coal washing plant - Coll. Guard. - 1931. Vol. CXLIII, pp. 198 - 201; 375.
7. Ариненков Ю.Д. Метод форматирования массивов данных обогатимости полезных ископаемых // Наукові праці Донецького державного технічного університету. Серія: Гірничо-електромеханічна.- Донецьк, 2000.- Випуск 16.- С. 3-10.
8. Ариненков Ю.Д. Метод минимизации массива классов крупности сыпучих материалов // Наукові праці Донецького державного технічного університету. Серія: Гірничо-електромеханічна.- Донецьк, 2001.- Випуск 27.- С. 17-22.

9. Гончаров В.Л. Теория интерполирования и приближения функций. ГТТИ, М.-Л., 1934, гл.1.
10. Демидович Б.П., Марон И.А. Основы вычислительной математики. Издание третье, исправленное. «Наука», Гл. ред. физ.-мат. лит. -М.: - 1966. -664 с.
11. Shoenfield J.R. Degrees of unsolvability. North-Holland Publishing Company, American Elsevier Publishing Company, 1971, 192p.
12. Ремез Е.Я. Общие вычислительные методы чебышевского приближения, Изд. АН УССР, 1957, ч. 1, гл. I.
13. Ариненков Ю.Д. Сепарабельное моделирование сепарационной технологии // 4 междунар.н.-тех.семинар «Практика и перспективы развития институционального партнёрства» / Известия ДонНТУ-ТГРТУ. - №3, апрель 2003. Т.1, с.75-82.

Дата надходження до редколегії: 20.12.2003 р.