

СПЕКТРАЛЬНЫЙ АНАЛИЗ ДАННЫХ РАДИОЛОКАЦИОННЫХ НАБЛЮДЕНИЙ АСТЕРОИДОВ С ПОМОЩЬЮ МЕТОДА НАЛОЖЕНИЯ ЭПОХ

Ю.В. ЧЕРКАС, Ю.И. ВОЛОЩУК

Метод наложения эпох относится к непараметрическим методам выделения скрытой периодичности и основывается на построении фазовых диаграмм ряда для пробных значений периода. Его достоинством является отсутствие априорного предположения о форме периодической составляющей, которая в большинстве случаев классического спектрального анализа предполагается гармоникой, и, следовательно, его применение позволяет определить ее истинную форму. Отсутствует требование равномерной дискретизации сигнала, что облегчает анализ временных рядов в некоторых задачах радиотехники и других наук. В рамках данной работы с помощью метода наложения эпох был подтвержден факт наличия периодической составляющей в распределениях астероидов по большой полуоси, найденный ранее классическими методами спектрального анализа [1]. Сделана оценка формы присутствующей периодической компоненты.

Ключевые слова: спектральный анализ, непараметрические методы спектрального анализа, метод наложения эпох, гистограмма численности, околоземные астероиды.

ВВЕДЕНИЕ

Классические методы спектрального анализа будь то непараметрические (периодограмма, коррелограмма) или параметрические (модели авторегрессии, скользящего среднего, авторегрессии и скользящего среднего и др.) обладают хорошо разработанной теоретической базой и широким спектром применений. Однако если целью исследования является не получение спектра исходного сигнала как такового (т.е. оценка распределения мощности или энергии в частотной области), а нахождение скрытой периодической составляющей, эти методы обладают одним ограничением — поиск ведется в пространстве гармоник. Когда периодическая составляющая обладает формой отличной от синусоидальной, а сам сигнал является зашумленным, то классические методы спектрального анализа могут давать существенно искаженные параметры периодической компоненты. Преимуществом метода наложения эпох является то, что он может использоваться для анализа как равномерных рядов, так и рядов с меняющимся шагом дискретизации, что значительно расширяет область его потенциального применения.

Объект исследования — выборка орбит околоземных астероидов NEA (Near Earth Asteroids) [2].

Предмет исследования — частотный состав дискретного равномерного ряда, формируемого на базе гистограммы численности распределения имеющейся выборки орбит астероидов по большой полуоси. Указанное преобразование приводит данные к форме, позволяющей использовать весь спектр существующих методов анализа рядов и дискретных сигналов в радиотехнике.

Целями работы являются:

- 1) Рассмотрение и адаптация метода наложения эпох применительно к объекту исследований.
- 2) Подтверждение наличия периодической составляющей в ряде распределения околозем-

ных астероидов, присутствие которой было выявлено в работе [1];

- 3) Оценка формы найденной периодической составляющей.

Актуальность работы заключается в получении новых знаний об объекте исследования и тонкой структуре распределения орбит околоземных астероидов NEA в пространстве, что особенно важно ввиду стремительного развития направлений науки и техники, направленных на изучение и предотвращение кометно-астероидной опасности для Земли. Предложенная схема анализа и сам алгоритм метода наложения эпох могут быть использованы в ряде задач спектрального анализа рядов как радиотехники, так и других наук, если не в качестве основного метода, то как дополнительный способ более всестороннего изучения данных.

1. АЛГОРИТМ МЕТОДА НАЛОЖЕНИЯ ЭПОХ

При изучении рядов иногда можно найти приближенное значение \hat{P} истинного периода P_0 , а затем построить фазовую диаграмму, т.е. последовательные отрезки ряда $\{y_k\}$ длиной \hat{P}_0 . Эти отрезки накладывают друг на друга и рассматривают зависимость y_k от фазы [3]:

$$x_k = \text{FR} \left(\frac{t_k - t_0}{\hat{P}_0} \right), k = 0, 1, \dots, N-1, \quad (1)$$

где $\text{FR}(z)$ дробная часть числа z , t_0 — произвольное значение аргумента ряда (которым в нашем случае будет значение большой полуоси), N — количество отсчетов ряда.

Остается рассмотреть зависимость y_k от x_k . Эта зависимость и является фазовой диаграммой ряда. Цель построения фазовой диаграммы заключается в данном случае в том, чтобы при-

ближенно определить форму истинной периодической зависимости $y(x)$.

Очевидно, что при P_0 близком к истинному периоду P_0 точки на фазовой диаграмме располагаются наиболее упорядоченным образом. В противном случае разброс точек будет обусловлен не только неизбежным шумом, но и причинами систематического характера, т.е. будет превышать разброс, соответствующий сворачиванию ряда с периодом P_0 . Таким образом, разброс точек на фазовой диаграмме может служить показателем близости какого-либо значения периода к истинному периоду P_0 .

Используя метод наложения эпох, можно вообще отказаться от нахождения приближенного значения \hat{P}_0 с помощью классических методов спектрального анализа и предложить следующую процедуру выделения скрытой периодической зависимости [3]:

1) определить некоторую числовую характеристику Γ степени упорядоченности расположения точек на фазовой диаграмме;

2) построить фазовые диаграммы изучаемого ряда для всех интересующих нас значений пробного периода P (или частоты $f=1/P$) и для каждой фазовой диаграммы вычислить соответствующую характеристику $\Gamma(f)$;

3) выбрать то значение Γ_0 , которое отвечает наиболее упорядоченному расположению точек (обычно это минимальное или максимальное – в зависимости от определения – значение Γ). Соответствующая Γ_0 частота f_0 и укажет на искомое значение этого параметра, а расстояние $\Gamma(f_0) \equiv \Gamma_0$ от типичных для других частот значений $\Gamma(f)$ – на степень надежности найденной периодической зависимости.

2. СТАТИСТИКА АББЕ-ЛАФЛЕРА-КИНМАНА (ALK) [3]

Одним из исторически первых и удачных критериев оценки степени упорядоченности точек на фазовой диаграмме является статистика Аббе-Лафлера-Кинмана, которая базируется на вычислении суммы квадратов разностей для последовательных точек на ней, нормированную к общей дисперсии ряда.

Рассмотрим ряд $\{t_k, y_k\}$, $k=0, 1, \dots, N-1$, включающий N не обязательно равноотстоящих отсчетов процесса $y(t)$. Для каждого момента t_k при заданном пробном периоде $P=f^{-1}$ вычислим фазу (1).

Рассмотрим далее все фазовые диаграммы, т.е. зависимости y_k от $x_k(f)$, для набора пробных частот с равномерным шагом по частоте, обеспечивающим должное разрешение.

Описываемая характеристика степени разброса точек на фазовой диаграмме строится следующим образом.

Расположим фазы x_k в порядке возрастания и обозначим упорядоченные значения $x_{[k]}$, так что

$$0 \leq x_{[0]} \leq x_{[1]} \leq \dots \leq x_{[N-1]} < 1. \quad (2)$$

Пусть

$$\bar{y}_k \equiv y(x_{[k]}) \quad (3)$$

– соответствующие (2) отсчеты ряда, последовательные не на оси аргумента ряда, а на фазовой диаграмме (т.е. близкие в смысле пробного периода P). Ясно, что величина

$$s_{in}^2 = \frac{1}{2(N-1)} \sum_{k=0}^{N-2} (\bar{y}_{k+1} - \bar{y}_k)^2 \quad (4)$$

характеризует упорядоченность точек на фазовой диаграмме: если последовательные точки в среднем близки друг к другу, то s_{in}^2 мала, а если они расположены совершенно хаотично, то s_{in}^2 будет величиной порядка общей дисперсии ряда:

$$s^2 = \frac{1}{N-1} \sum_{k=0}^{N-2} (y_k - \bar{y})^2. \quad (5)$$

Таким образом, следует ожидать, что отношение

$$\Gamma_{ALK}(f) = \frac{\sum_{k=0}^{N-2} (\bar{y}_{k+1} - \bar{y}_k)^2}{2 \sum_{k=0}^{N-1} (y_k - \bar{y})^2} \quad (6)$$

будет существенно меньше единицы при значении пробного периода, близкого к истинному периоду P_0 , и $\Gamma_{ALK} \approx 1$ на других частотах. В первом случае точки на фазовой диаграмме будут располагаться тесно вокруг некоторой средней кривой, которая и отражает форму искомой периодичности.

3. ФАЗОВАЯ ГРУППИРОВКА ДАННЫХ

Для вычисления статистики ALK необходимо при каждом значении пробного периода произвести поточечное упорядочение отсчетов в соответствии с их фазами. Способы адресной сортировки позволяют выполнить эту операцию достаточно быстро, однако можно и вообще избежать полной сортировки, разбивая предварительно весь круговой интервал изменения фазы $[0, 1]$ на подинтервалы (бины) и формируя при каждом значении пробного периода гистограммы отсчетов. Далее на основании таких сгруппированных по бинам данных нужно оценить, насколько хорошо выделяется средняя кривая.

Достоинства методов, опирающихся на фазовую группировку данных, связаны с относительно высокой скоростью расчетов и более гладкой, чем для статистики ALK, периодограммой. Вместе с тем указанные методы пригодны лишь при относительно большом количестве измерений и вносят в окончательные результаты субъективность, обусловленную выбором биновой структуры [3]. Кроме того, при современном быстродействии компьютеров выгоды, доставляемые формированием гистограммы вместо сортировки отсчетов, не имеют решающего зна-

чения. Приведенные соображения обусловили преимущественное использование метода фазовой группировки на предварительном, поисковом этапе исследования.

4. ДИСПЕРСИОННЫЙ АНАЛИЗ Р. ФИШЕРА [3]

Рассмотрим произвольный временной ряд $\{t_k, y_k\}$, $k=0, \dots, N-1$ и вычислим для каждого отсчета фазу $x_k(v)$ (1). Разобьем интервал изменения фазы $[0, 1)$ на M неперекрывающихся бинов, содержащих последовательно N_m ($m=1, 2, \dots, M$) отсчетов каждый, так что

$$\sum_{m=1}^M N_m = N. \quad (7)$$

Пусть y_{mj} обозначает j -й отсчет в m -том бине ($j=1, 2, \dots, N_m$). Введем оценки групповых среднего и дисперсии отсчетов:

$$\bar{y}_m = \frac{1}{N_m} \sum_{j=1}^{N_m} y_{mj}, \quad (8)$$

$$s_m^2 = \frac{1}{N_m - 1} \sum_{j=1}^{N_m} (y_{mj} - \bar{y}_m)^2,$$

а также общего среднего значения выборки и полной дисперсии:

$$\bar{y} = \frac{1}{N} \sum_{k=1}^N y_k = \frac{1}{N} \sum_{m=1}^M N_m \bar{y}_m = \frac{1}{N} \sum_{m=1}^M \sum_{j=1}^{N_m} y_{mj}, \quad (9)$$

$$s^2 = \frac{1}{N-1} \sum_{k=1}^N (y_k - \bar{y})^2 = \frac{1}{N-1} \sum_{m=1}^M \sum_{j=1}^{N_m} (y_{mj} - \bar{y})^2.$$

Обозначив вошедшие сюда двойные суммы просто как $\sum_{m,j}$, запишем упомянутое тождество в виде:

$$\sum_{m,j} (\bar{y}_{m,j} - \bar{y})^2 = \sum_{m,j} (\bar{y}_m - \bar{y})^2 + \sum_{m,j} (y_{m,j} - \bar{y}_m)^2. \quad (10)$$

Для доказательства следует в скобках исходной суммы добавить и вычесть \bar{y}_m , возвести во вторую степень и учесть, что сумма перекрестных членов равна нулю ввиду (8).

Таким образом, сумма квадратов всех отклонений от общего среднего разбивается на сумму квадратов отклонений локальных средних от общего среднего и на сумму квадратов отклонений отсчетов от локальных средних внутри бинов. Каждая из этих частичных сумм может рассматриваться как мера «качества» средней кривой блеска на фазовой диаграмме; отличие фазовой картины от хаотической проявляется тем более четко, чем больше первая сумма или меньше вторая сумма. Но учитывая, что сумма в левой части тождества, т. е. общий разброс данных, остается фиксированной, оба приведенных критерия эквивалентны.

Учитывая определения (8) и (9), перепишем тождество (10) в следующем виде:

$$(N-1)s^2 = \sum_{m=1}^M N_m (\bar{y}_m - \bar{y})^2 + \sum_{m=1}^M (N_m - 1) s_m^2. \quad (11)$$

Пусть далее

$$v_1^2 = \frac{1}{M-1} \sum_{m=1}^M N_m (\bar{y}_m - \bar{y})^2 \quad (12)$$

и

$$v_2^2 = \frac{1}{N-M} \sum_{m=1}^M (N_m - 1) s_m^2 \quad (13)$$

– соответственно, характеристики разброса локальных средних и внутренней дисперсии совокупности точек на фазовой диаграмме. Тогда тождество Фишера принимает окончательный вид:

$$\frac{M-1}{N-1} \frac{v_1^2}{s^2} + \frac{N-M}{N-1} \frac{v_2^2}{s^2} = 1 \quad (14)$$

Характерная особенность непараметрических методов состоит в том, что они не отдают предпочтения «простым» кривым, как это сделал бы человек при сравнении фазовых диаграмм. Предположим, например, что мы исходим из критерия максимума дисперсии средних (12). Тогда и последовательность $\{\bar{y}_m\}$, очерчивающая простую кривую с двумя экстремумами, и произвольно переставленная последовательность тех же групповых средних, очерчивающая, например, кривую с четырьмя экстремумами, дадут одно и то же значение v_1^2 . Формально задача решается путем введения дополнительного критерия «простоты» типа ограничения количества экстремумов.

5. АНАЛИЗ ДАННЫХ

Подготовка данных выборки астероидов NEA [2] для спектрального анализа их распределения по большой полуоси на предмет наличия периодической составляющей включала в себя следующие этапы [1, 4]:

1) Построение гистограммы распределения астероидов NEA по большой полуоси (на рис. 1 приведена гистограмма для $N=100$ интервалов разбиения диапазона изменения аргумента).

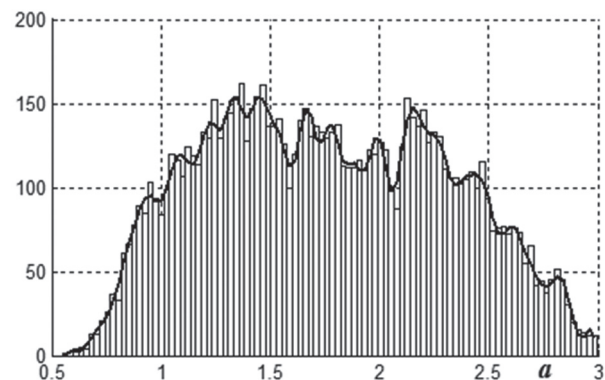


Рис. 1. Гистограмма распределения астероидов NEA по большой полуоси

2) Нахождение тренда построенного распределения. Для его получения было использовано сглаживание ряда гистограммы сплайном по методу наименьших квадратов. Возможны также другие методы выделения тренда: аппроксимация полиномом, скользящее среднее, цифровая фильтрация, параметрические методы – каждый из них имеет свои преимущества и недостатки [5].

3) Нахождения ряда колебания численности (рис. 2) путем вычитания из ряда гистограммы выделенного тренда.

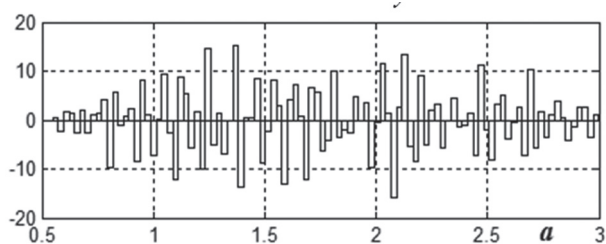


Рис. 2. Ряд колебания численности астероидов NEA по большой полуоси

В качестве начальной точки спектрального анализа будем использовать классическую немодифицированную периодограмму (рис. 3) ряда [6, 7, 8]. Стоит отметить, что, несмотря на такой недостаток, как неустойчивость спектра, периодограмма является одной из наиболее робастных оценок в том смысле, что может быть использована для анализа практически любых рядов, обладающих свойством стационарности. А в случае линейного спектра дает оценки, не уступающие по спектральному разрешению и обнаружимости сигнала другим методам спектрального анализа [6].

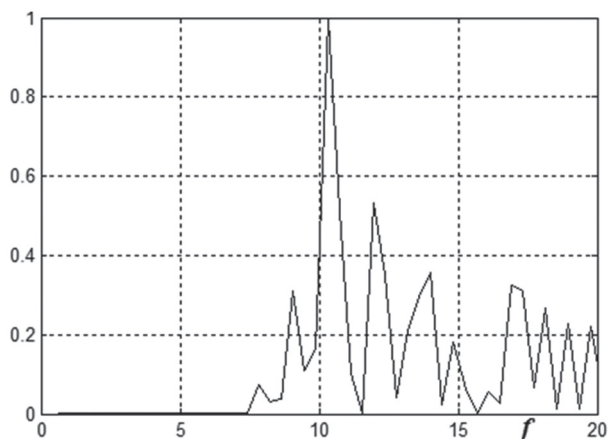


Рис. 3. Немодифицированная периодограмма ряда колебания численности

Периодограмма рис. 3 содержит отчетливо проявляющуюся спектральную составляющую на частоте $f \approx 10.3 a.e.^{-1}$. Полное отсутствие составляющих на нижних частотах объясняется удалением тренда, который в данном случае был выделен путем сглаживания исходного ряда сплайном. Данная процедура по своему характеру равноценна пропуску ряда через цифровой фильтр верхних частот.

Найденная оценка частоты периодической составляющей $f \approx 10.3 a.e.^{-1}$ соответствует гармоническому колебанию, поскольку при вычислении значений классической периодограммы за основу берется корреляция ряда с набором гармоник в наперед заданной сетке частот.

Проведем спектральный анализ ряда колебания численности методом наложения эпох. Для

этого были рассчитаны и построены значения статистики Аббе-Лафлера-Кинмана для сетки частот с шагом $f = 10^{-3} a.e.^{-1}$. Минимум выбранной статистики, соответствующий наиболее упорядоченному расположению точек на фазовой диаграмме, находится на частоте $f \approx 10.2 a.e.^{-1}$ рис. 4. Соответствующая ему фазовая диаграмма приведена на рис. 5. Сплошной линией показано сглаженное полиномом среднее значение. Для большей наглядности область изменения фазы повторится дважды.

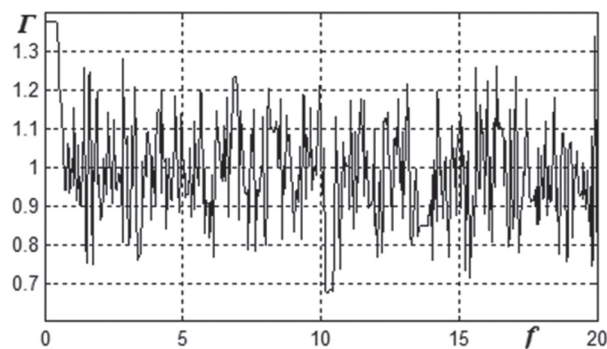


Рис. 4. Зависимость статистики Аббе-Лафлера-Кинмана от частоты

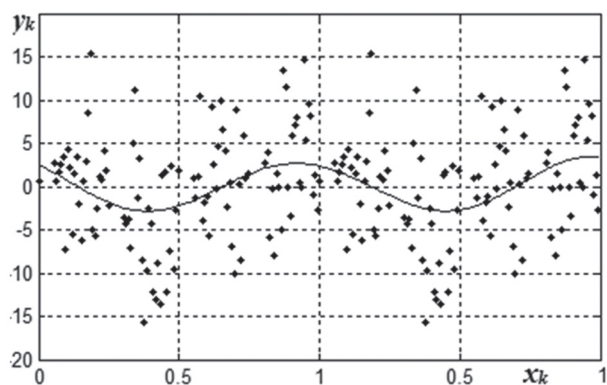


Рис. 5. Фазовая диаграмма распределения астероидов NEA для частоты $f \approx 10.2 a.e.^{-1}$

ВЫВОДЫ

Полученное значение частоты периодической составляющей $f \approx 10.2 a.e.^{-1}$ хорошо согласуется со значениями периодической составляющей, найденными с помощью как периодограммного метода, так и параметрического метода авторегрессии (АР) для астероидов NEA [4, 5]. По фазовой диаграмме (рис. 5) видим, что даже наиболее упорядоченное расположение точек является достаточно зашумленным; среднее значение, полученное сглаживанием полиномом, по форме мало отличается от гармонического колебания. К достоинствам метода наложения эпох следует отнести его чувствительность при обнаружении колебаний, форма которых сильно отличается от синусоиды. В этом случае спектральная плотность содержит высокие гармоники (обертоны), выделение которых из шумов нередко затруднительно. Однако для случая, когда спектр сигнала является линейным, т.е.

содержит только гармонические составляющие, использование классических (периодограмный, коррелограмный) или же параметрических методов спектрального анализа является предпочтительным.

Литература

- [1] Черкас Ю.В. Применение спектрального анализа для поиска периодических составляющих в распределениях астероидов по большой полуоси / Ю.В.Черкас, Ю.И. Волощук // Прикладная радиоэлектроника. – 2012. – Т. 11, № 3. – С. 335-341.
- [2] Программа мониторинга объектов, орбиты которых сближаются с орбитой Земли [Электронный ресурс]. – Режим доступа: http://neo.jpl.nasa.gov/cgi-bin/neo_elem.
- [3] Тербиж В.Ю. Анализ временных рядов в астрофизике. – М.: Наука, 1992. – 392 с.
- [4] Черкас Ю.В. Анализ тонкой структуры распределенной орбит малых тел Солнечной системы. Околоземные астероиды. // Ю.В. Черкас, Ю.И. Волощук, С.В. Коломиец // Материалы 22-й Международной Крымской конференции КрыМиКо, 10-14 сентября 2012. – Том 2. – С. 1003-1004.
- [5] Андерсон Т. Статистический анализ временных рядов: Пер. с англ. – М.: Мир, 1976. – 755 с.
- [6] Марпл.-мл. С.Л. Цифровой спектральный анализ и его приложения: Пер. с англ. – М.: Мир, 1990. – 584 с.
- [7] Бриллинджер Д. Временные ряды. Обработка данных и теория: Пер. с англ. – М.: Мир, 1980. – 536 с.
- [8] Stoica P. Introduction to Spectral Analysis. – Prentice Hall, New Jersey, 1997 – 319 p.

Поступила в редколлегия 04.06.2014



Волощук Юрий Иванович, доктор технических наук, профессор, ХНУРЭ. Научные интересы: радиолокация метеоров, обработка радиосигналов, спектральный анализ временных рядов



Черкас Юрий Васильевич, аспирант, ХНУРЭ. Научные интересы: математическое моделирование, обработка радиосигналов, спектральный анализ временных рядов.

УДК 523.24:521.1; 523.6

Спектральный анализ данных радиолокационных спонсеров астероидов за допомогою методу накладання епох / Ю.В. Черкас, Ю.И. Волощук // Прикладна радіоелектроніка: наук.-техн. журнал. – 2014. – Том 13. – № 2. – С. 164–168.

Метод накладання епох належить до непараметричних методів виділення прихованої періодичності та базується на побудові фазових діаграм ряду для пробних значень періоду. Його перевагою є відсутність апіорного припущення про форму періодичної складової, яка в більшості випадків вважається гармонікою, а отже, дозволяє визначити істинну форму періодичної складової. Також відсутня вимога рівномірної дискретизації сигналу, що полегшує аналіз часових рядів у деяких задачах радіотехніки та інших наук. У рамках даної роботи з допомогою методу накладання епох був підтверджений факт наявності періодичної складової в розподілах астероїдів по великій півосі, знайдений раніше класичними методами спектрального аналізу [1]. Зроблена оцінка форми наявної періодичної складової.

Ключові слова: спектральний аналіз, непараметричні методи спектрального аналізу, періодограма, гістограма чисельності, навколосемні астероїди.

Л.: 5. Бібліогр.: 8 найм.

UDC 523.24:521.1; 523.6

Spectral data analysis of radar observations of asteroids using the epoch folding method / Yu.V. Cherkas, Yu.I. Voloshchuk // Applied Radio Electronics: Sci. Journ. – 2014. – Vol. 13. – № 2. – P. 164–168.

The method of epoch folding is a non-parametric method of finding hidden periodicity and is based on building of series phase diagrams for different trial values of periods. Its advantage is the absence of a priori assumptions about the shape of a periodic component, which in most cases of classic spectral analysis is assumed harmonic and therefore it allows to determine the true shape of the periodic component. Besides, there is no requirement of uniform sampling of a signal, which facilitates the analysis of time series in some tasks of radio engineering and other sciences. In this work using the epoch folding method has confirmed the presence of a periodic component in the semimajor axis distributions of asteroids, which has been previously found by classical methods of spectral analysis [1]. The estimation of the periodic component shape is made.

Keywords: spectral analysis, nonparametric methods of spectral analysis, epoch folding method, number histogram, near Earth asteroids.

Fig.: 5. Ref: 8 items.