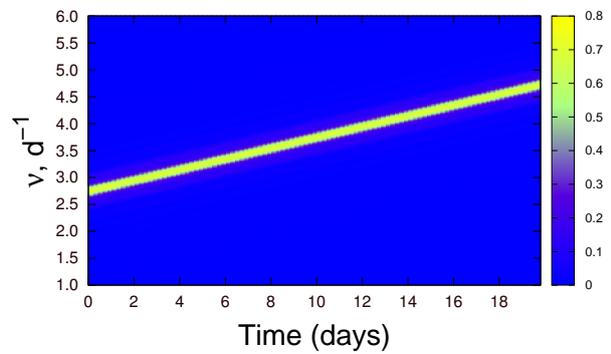
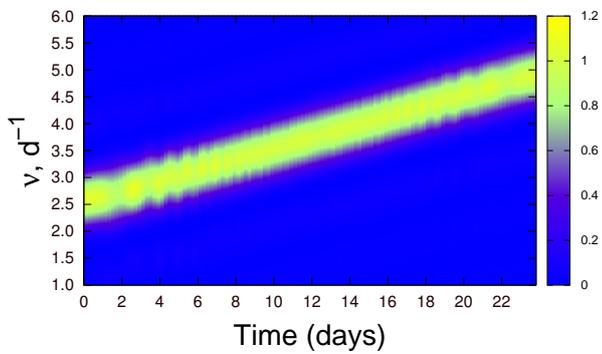


А.Ф. Холтыгин

ПРЕОБРАЗОВАНИЕ ФУРЬЕ: АЛГОРИТМЫ И СТАТИСТИКА



Санкт-Петербург 2022

Аннотация

В пособии даны основные сведения о Фурье-преобразовании. Исследованы статистические свойства периодограммы Шустера белого шума для различных временных сеток. Рассмотрено оконное преобразование Фурье. Представлены оконные преобразования Фурье модельных функций. Описано использование оконного преобразования Фурье для анализа вариаций профилей линий в спектре ВП звезды ρ Leo.

Содержание

1	Преобразование Фурье	4
1.1	Спектр мощности и корреляционная функция	4
1.1.1	Периодограмма Шустера	5
2	Обнаружение гармонических компонентов сигналов: равномерные временные ряды	7
2.1	Чисто гармоническая функция	7
2.2	Периодограмма белого шума	9
2.3	Выделение гармонических компонентов из периодограммы	9
3	Выделение гармонических компонентов для неравномерных временных рядов	13
3.1	Статистика отсчетов периодограммы белого шума	13
3.2	Статистика максимальных отсчетов периодограммы белого шума	15
4	Дискретное преобразование Фурье	15
4.1	Независимость от сдвига начала отсчета временного ряда	16
5	Оконное преобразование Фурье	18
5.1	Периодограмма Шустера для оконного преобразования Фурье	21
5.1.1	Выбор интервала частот при расчете периодограммы Шустера для оконного преобразования Фурье	22
5.2	Оконное преобразования Фурье модельных функций	23
5.2.1	Модельные функции с постоянным периодом	23
5.2.2	Модельные функции с переменным периодом	24
6	Анализ вариаций профилей в спектрах звезд с использованием оконного преобразования Фурье	27
6.1	Вариации профилей в спектре ρ Leo	27

1 Преобразование Фурье

Пусть задана функция $f(t)$ вещественной переменной t . Ее преобразование Фурье

$$\hat{f}(\omega) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} f(t)e^{-i\omega t} dt, \quad (1)$$

Обратное преобразование Фурье:

$$f(t) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} \hat{f}(\omega)e^{i\omega t} d\omega. \quad (2)$$

То есть функция $f(t)$ может быть представлена в виде разложения по гармоникам угловой частоты ω .

Перейдем от угловой частоты ω к частоте ν ($\omega = 2\pi\nu$), тогда преобразование Фурье примет следующий вид:

$$\hat{f}(\nu) = \sqrt{2\pi} \int_{-\infty}^{\infty} f(t)e^{-i2\pi\nu t} dt, \quad (3)$$

Обратное преобразование Фурье в пространстве частот ν будет равно:

$$f(t) = \sqrt{2\pi} \int_{-\infty}^{\infty} \hat{f}(\nu)e^{i2\pi\nu t} d\nu, \quad (4)$$

1.1 Спектр мощности и корреляционная функция

Случайным процессом в теории вероятностей называется совокупность случайных величин, индексированных какой-то упорядоченной величиной (обычно значениями времени или координаты). Будем означать через $F = F(t)$ какую-то реализацию случайного процесса. К случайному процессу можно отнести результаты измерения какой-либо физической величины в зависимости от времени измерения.

Случайный процесс называется стационарным, если функции распределения случайных величин, определяющих случайный процесс, не зависят от времени. Большинство процессов измерения являются стационарными. Важным подмножеством стационарных случайных процессов являются эргодические случайные процессы.

Стационарный случайный процесс называется эргодическим, если любая его вероятностная характеристика, полученная усреднением по множеству возможных реализаций, с вероятностью, сколь угодно близкой к единице, равна временному среднему, полученному усреднением за достаточно большой промежуток времени по одной единственной реализации случайного процесса. То есть для эргодического процесса среднее по времени равно среднему по множеству возможных реализаций.

Обозначим через $E[F] = E[F(t)]$ математическое ожидание случайной величины $F(t)$, а $D[F] = D[F(t)]$ – его дисперсию. Случайные процессы с $E[F] = 0$ называются центрированными. Очевидно, что любой случайный процесс можно представить в виде суммы постоянной величины и центрированного случайного процесса:

$$F(t) = E + F_c(t),$$

где $E = E[F(t)]$, а $F_c(t) = F(t) - E$.

Для случайных процессов вводится понятие *плотность спектра мощности* (в дальнейшем просто *спектр мощности*):

$$S(\omega) = E \left(\lim_{T \rightarrow \infty} \frac{1}{2\pi T} \left| \int_0^T f_i(t) e^{-i\omega t} dt \right|^2 \right) \quad (5)$$

и корреляционная функция

$$K(t, t') = E[f_i(t) \cdot f_i(t')]. \quad (6)$$

Корреляционная функция является функцией двух аргументов. В случае стационарного случайного процесса корреляционная функция зависит только от разности аргументов:

$$K(t, t') = C(\tau), \quad \tau = t - t'. \quad (7)$$

Для центрированных случайных процессов справедлива теорема *Винера-Хинчина* (см., например, [13]). Если центрированный случайный процесс задан на промежутке $t \in [0, T]$ множеством реализаций

$$F(t) = \{f_i(t)\}_{i=1}^N \quad (8)$$

То его спектр мощности и корреляционная функция взаимно обратимы по Фурье:

$$S(\omega) = \frac{1}{2\pi} \int_{-\infty}^{\infty} C(\tau) e^{-i\omega\tau} d\tau, \quad (9)$$

а

$$C(\tau) = \frac{1}{N^2} \sum_{k=0}^{N-1} S(\omega) e^{i\omega\tau} d\omega. \quad (10)$$

1.1.1 Периодограмма Шустера

В практических задачах мы располагаем не всем множеством реализаций случайного процесса (случайной функции) $F(t)$, а одной реализацией, причем не для всего набора возможных значений аргумента t , а для дискретного множества значений $\{t\} = \{t_0, t_1, \dots, t_{N-1}\}$, где N — число значений аргумента. В дальнейшем случайные функции, зависящие от времени t будем по традиции называть *временными рядами*.

Временные ряды разделяются на *равномерные*, когда разности соседних отсчетов постоянны для всего временного ряда и *неравномерные*, когда это условие не выполняется. Для равномерного временного ряда

$$t_k = k * \Delta t, k = 0, 1, \dots, N - 1 \quad (11)$$

где Δt — постоянный шаг временного ряда, а N — число временных отсчетов (длина временного ряда).

Для временных рядов в силу их конечности и единственности реализации вместо спектра мощности и корреляционной функции случайного процесса мы можем получить только их оценки.

В качестве оценки спектра мощности используют *периодограмму Шустера*, определенную следующим образом [11]:

$$D(\nu) = \frac{1}{N^2} \left| \sum_{k=0}^{N-1} f_k e^{-i2\pi\nu t_k} \right|^2. \quad (12)$$

Здесь f_k — значения (отсчеты) случайного процесса
Из формулы Коши-Буняковского следует:

$$D(\nu) \leq \frac{1}{N^2} \sum_{k=0}^{N-1} f_k^2 \sum_{k=0}^{N-1} |e^{-i2\pi\nu t_k}|^2 \leq \frac{1}{N} \sum_{k=0}^{N-1} f_k^2 \leq (f^{\max})^2,$$

где f^{\max} — максимальное по модулю значение отсчетов f_k . То есть значения периодограммы Шустера конечны при любом значении N .

Важной величиной при анализе временных рядов является так называемая частота Найквиста, определяемая следующим выражением:

$$\nu_{\text{Nq}} = \frac{1}{2\Delta t} = \frac{N}{2T}. \quad (13)$$

В случае равномерного ряда Δt — это шаг временных отсчетов, для неравномерных рядов $\Delta t = \overline{\Delta t} = T/N$ — среднее расстояние между временными отсчетами, где T — полная длина временного ряда.

Частота Найквиста соответствует максимально возможной частоте гармонического компонента, который еще может быть выделен из временного ряда.

Для равномерного ряда важное значение имеет *Фундаментальная система частот*, определяемая следующим образом (см. формулу (10) в [13]):

$$\nu_j = \frac{j}{N\Delta t} = \left(\frac{2\nu_{\text{Nq}}}{N} \right) \cdot j, \quad j = 0, 1, \dots, N_{\max}. \quad (14)$$

При N четном $N_{\max} = N/2$, при нечетном — $N_{\max} = (N + 1)/2$. Витязев [13] предлагает полагать N всегда четным, а при N нечетном добавлять в систему фундаментальных частот значение $j = (N + 1)/2$.

При этом последний отсчет системы фундаментальных частот $\nu_{N_{\max}} = \nu_{\text{Nq}}(1 + 1/N)$, что слегка выше частоты Найквиста, но при $N \gg 1$ это не существенно. Величина $\Delta\nu_{\text{Nq}} = 2\nu_{\text{Nq}}/N = 1/T$ представляет собой шаг по частоте в системе фундаментальных частот. То есть $\Delta\nu_{\text{Nq}}$ — разрешение Фурье-спектра для данного временного ряда.

2 Обнаружение гармонических компонентов сигналов: равномерные временные ряды

Пусть в результате какого-то физического эксперимента (астрономических наблюдений и т.д.) получено N значений (отсчетов) случайного процесса $\{f(t_k)\}, k = 0, 1, \dots, N - 1$. В наиболее часто встречающемся случае значения $\{f(t_k)\}$ являются функцией времени.

Задачей гармонического (Фурье) анализа является выяснение вопроса о том, если в этом процессе периодический (гармонический) компонент (компоненты) и определение его (их) параметров.

В общем случае случайный процесс $f(t)$ можно представить как сумму трех слагаемых

$$f(t) = f^0(t) + \sum_{l=1}^M A_l \cos(2\pi\nu_l t + \phi_l) + N(t), \quad t = t_0, t_1, \dots, t_{N-1}. \quad (15)$$

Первое слагаемое обычно называется *трендом*. Тренд может быть постоянным $f^0(t) = const$, линейным $f^0(t) = A + Bt$, квадратичным $f^0(t) = A + Bt + Ct^2$ и т.д. Для выделения гармонических компонентов требуется удалить тренд из значений случайного процесса. В первом случае достаточно из всех значений временного ряда вычесть среднее значение отсчетов $f_k = f(t_k)$. Это процесс называется центрированием временного ряда. Во всех следующих случаях для определения параметров тренда $f^0(t)$ обычно используется метод наименьших квадратов.

В дальнейшем будем предполагать, что тренд уже удален из значений временного ряда. Тогда

$$f(t) = \sum_{l=1}^M A_l \cos(2\pi\nu_l t + \phi_l) + N(t), \quad t = t_0, t_1, \dots, t_{N-1}. \quad (16)$$

Здесь A_l, ν_l и ϕ_l — соответственно амплитуда, частота и фаза гармонического компонента с номером l , $N(t)$ — вклад шумового компонента.

Использование косинусоиды вместо синусоиды — дань традиции. Замена в формуле (16) косинуса на синус эквивалентно сдвигу фазы на $\pi/2$. Рассмотрим частные случаи соотношения (16).

2.1 Чисто гармоническая функция

Пусть

$$f(t) = A_0 \cos(2\pi\nu_0 t + \phi_0), \quad t = t_0, t_1, \dots, t_{N-1}. \quad (17)$$

Спектр мощности этой функции

$$S(\nu) = \frac{A_0^2}{4} [\delta(\nu - \nu_0) + \delta(\nu + \nu_0)]. \quad (18)$$

где $\delta(x)$ — дельта функция аргумента x .

В общем случае периодограмма Шустера $D(\nu)$ для функции (18) не имеет аналитического выражения. Однако для равномерного временного ряда с шагом Δt согласно [13]

$$D(\nu) = \frac{A_0^2}{4} [W(\nu - \nu_0) + W(\nu + \nu_0)], \quad (19)$$

где *спектральное окно*

$$W(\nu) = \frac{\sin^2(N\pi\nu\Delta t)}{N^2 \sin^2(\pi\nu\Delta t)}. \quad (20)$$

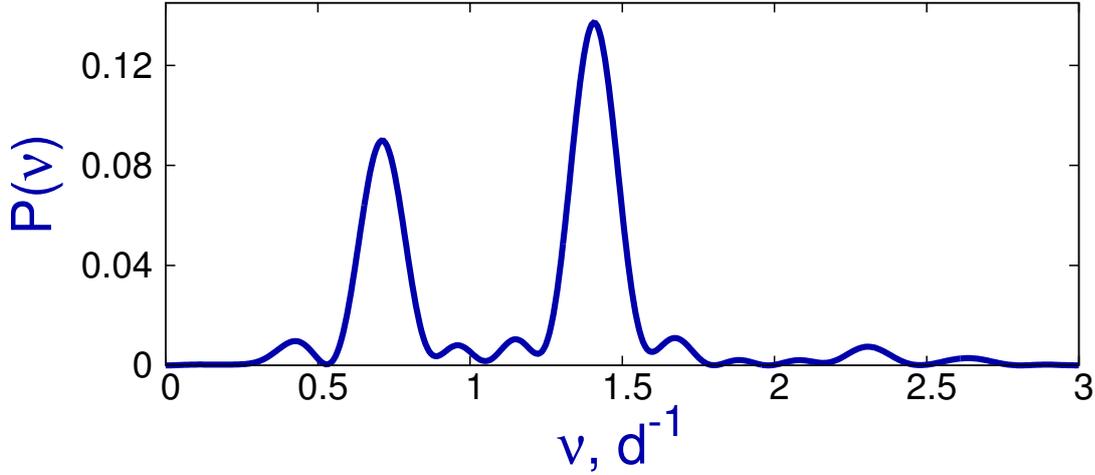


Рис. 1: Периодограмма Шустера (12) равномерного ряда с гармоническими компонентами $\nu_1 = 0.7 d^{-1}$ и $\nu_2 = 1.4 d^{-1}$.

Из выражения (20) следует, что спектральное окно является периодической функцией частоты с периодом $P_\nu = 2\nu_{Nq}^{-1}$. Из этого следует, что периодограмма $D(\nu)$ также является периодической функцией с тем же периодом.

По этой причине значения периодограммы $D(\nu)$ имеет смысл рассчитывать только на основном периоде $[-$

$nu_{Nq}, nu_{Nq}]$. Для вещественного временного ряда периодограмма является четной функцией, поэтому ее достаточно рассматривать только на промежутке $[0, \nu_{Nq}]$.

На Рис. 1 показана периодограмма равномерного ряда с гармоническими компонентами $\nu_1 = 0.7 d^{-1}$, $A = 0.5$ и $\nu_2 = 1.4 d^{-1}$, $A = 0.7$. Из анализа рисунка видно, что периодограмма гармонической функции представляет собой основной и боковые лепестки с быстро убывающей амплитудой.

Ширина основного (главного) лепестка периодограммы

$$\Delta\nu = \frac{2}{N\Delta T} = \frac{2}{T}. \quad (21)$$

Величина $\Delta\nu/2 = 1/T$ — полуширина главного лепестка периодограммы определяет точность, с которой может быть определена частота гармонического компонента временного ряда из анализа его периодограммы. То есть, если максимальный отсчет периодограммы соответствует некоторому значению ν , то реальная частота находится в интервале $\nu \pm 1/T$.

¹Учтено, что период $\sin^2(x) = \pi$

2.2 Периодограмма белого шума

Для шумовой компоненты временного ряда $N(t)$ возможны разные представления, однако наиболее часто используется модель *белого шума*, в которой предполагается, что все значения временного ряда η_k^σ являются независимыми и представляют собой выборку значений случайной величины η^σ с постоянной дисперсией σ^2 и нулевым математическим ожиданием (дискретный белый шум).

Проанализируем статистические свойства периодограммы белого шума для случая единичной дисперсии ($\sigma^2 = 1$) и равномерного временного ряда с шагом Δ , нулевым начальным отсчетом $t_0 = 0$ и длиной ряда N . В этом случае отсчеты временного ряда η_k представляют собой значения случайной величины $\eta = \eta^1$ с единичной дисперсией и нулевым математическим ожиданием.

Вычислим периодограмму Шустера (12) для фундаментальной системы частот (14). С учетом того, что временные отсчеты $t_k = k\Delta t$, получим:

$$D_j = D_j(\nu_j) = \frac{1}{N^2} \left| \sum_{k=0}^{N-1} \eta_k e^{-i\frac{2\pi}{N}kj} \right|^2. \quad (22)$$

где $\nu_j = j/(N\Delta t)$, а $\nu_j t_k = kj/N$.

В дальнейшем кроме периодограммы D_j будем использовать также нормированную периодограмму $d_j = ND_j$. Для отсчетов нормированной периодограммы справедлива теорема, полученная Шустером в 1898 (см., например, [13], стр. 10)²:

теорема Шустера. Отсчеты нормированной периодограммы d_j представляют собой случайную величину со следующей плотностью распределения:

$$p(x) = \begin{cases} e^{-x} & , \quad j = 1, 2, \dots, N_{\max} - 1, \\ \frac{e^{-x/2}}{\sqrt{2\pi x}} & , \quad j = 0, N_{\max}, \end{cases} \quad (23)$$

то есть при $j = 0$, N_{\max} величина d_j имеет распределение χ^2 с одной степенью свободы, а при $j = 1, 2, \dots, N_{\max} - 1$ — величина d_j подчиняются распределению χ^2 с двумя степенями свободы.

В дальнейшем будем рассматривать только внутренние отсчеты периодограммы $j = 1, 2, \dots, N_{\max} - 1$, так как отсчет при $j = 0$ соответствует нулевой частоте, а при $j = N_{\max}$ — максимально возможному значению частоты.

Вероятность того, что один из внутренних отсчетов превысит заданное число $x = X_1$ равна

$$P(x > X_1) = \int_{X_1}^{\infty} e^{-x} dx = e^{-X_1}. \quad (24)$$

2.3 Выделение гармонических компонентов из периодограммы

В общем случае временной ряд с гармоническим компонентом (компонентами) представляет собой сумму (16)³. Анализ периодограммы данного временного ряда во многих случаях позволяет определять характеристики гармонических компонентов (гармоник). Присутствие шумового компонента временного ряда значительно усложняет эту задачу, поэтому часто говорят об *выделении сигнала из шума*, имея в виду задачу определения

²см. доказательство в [12], с. 93

³напомним, что тренд уже удален из значений временного ряда

параметров гармоник из анализа отсчетов временного ряда, в которых присутствует шумовой компонент.

Сначала несколько общих соображений. Как мы видим из Рис. 1, значения периодограммы на частотах гармонических компонентов сигнала достигают локальных максимумов. Логично предположить, что если сигнал достаточно сильный (амплитуда гармонического компонента велика), то на частоте этого компонента периодограмма будет иметь максимум (по крайней мере локальный). То есть, если мы видим какой-то максимум периодограммы, то возможно, что положение этого максимума дает частоту одной из гармоник анализируемого временного ряда.

Но возникает вопрос, насколько велико должно быть значение этого максимума, чтобы мы могли бы уверенно утверждать о наличии гармоник в анализируемом временном ряде. Для получения ответа на этот вопрос обычно используют методы проверки статистических гипотез, описанные в любом стандартном учебнике по теории вероятностей и математической статистике (см., например [7]).

Выглядит это так. Сначала выдвигается основная гипотеза, которую мы хотим проверить. Назовем ее гипотезой H_0 . В нашем случае гипотеза H_0 заключается в том, что обнаруженный нами максимум периодограммы соответствует частоте гармоники, реально присутствующей в анализируемом нами временном ряде. Альтернативная гипотеза H_1 состоит в том, что гипотеза H_0 не верна.

Далее выбирается какой-то критерий \mathcal{K} , позволяющий судить о достоверности гипотезы H_0 . Выполнение этого критерия позволяет говорить, что гипотеза H_0 с высокой вероятностью верна. Но что в данном случае значит "с высокой вероятностью"? Обычно поступают так. Выбирается малая величина $\alpha \ll 1$, традиционная называемая *уровнем значимости*⁴ гипотезы. Критерий \mathcal{K} формулируется таким образом, что если он выполняется, то вероятность того, что гипотеза H_0 неверна составляет всего лишь малую величину α , тогда как вероятность того, что гипотеза H_0 верна $1 - \alpha \approx 1$.

Критерий \mathcal{K} для периодограмм традиционно составляется так: для каждого уровня значимости α определяется величина $D_{\max} = D_{\max}(\alpha)$ такая, что если какой-то отсчет периодограммы $D_j > D_{\max}(\alpha)$, то с вероятностью $1 - \alpha$ этот отсчет соответствует реальному сигналу (гармоническому компоненту) в анализируемом временном ряде

Пусть в рассматриваемом временном ряде присутствует шумовой компонент, представляющий собой дискретный белый шум с дисперсией σ^2 (стандартным отклонением σ). Отсчеты временного ряда η_k^σ в момент времени t_k представляют собой значения случайной величины η^σ с нулевым математическим ожиданием и дисперсией σ^2 (см. пункт 2.2).

Нормированная периодограмма для данного временного ряда (случайной величины η^σ) равна

$$d_j^\sigma = ND_j^\sigma = \frac{1}{N} \left| \sum_{k=0}^{N-1} \eta_k^\sigma e^{-i\frac{2\pi}{N}kj} \right|^2 = \frac{\sigma^2}{N} \left| \sum_{k=0}^{N-1} \left(\frac{\eta_k^\sigma}{\sigma} \right) e^{-i\frac{2\pi}{N}kj} \right|^2. \quad (25)$$

Введем случайную величину $\eta^{(1)} = \eta^\sigma / \sigma$. Дисперсия случайной величины $k\eta$, $D(k\eta) = k^2 D(\eta)$, поэтому дисперсия величины $\eta^{(1)}$ равна 1. То есть в сумме (25) значения η_k^σ / σ^2 можно заменить на значения η_k случайной величины η с нулевым математическим ожиданием и единичной дисперсией:

$$d_j^\sigma = \frac{\sigma^2}{N} \left| \sum_{k=0}^{N-1} \left(\eta_k^{(1)} \right) e^{-i\frac{2\pi}{N}kj} \right|^2 = \sigma^2 d_j. \quad (26)$$

⁴В англоязычной литературе используется для уровня значимости термин FAP (False Alarm Probability, вероятность ложной тревоги)

Определим вероятность $P(D_j^\sigma > A)$ того, что отсчеты периодограммы $D_j^\sigma = (1/N)d_j^\sigma$ превысят заданное число A . Очевидно, что

$$P(D_j^\sigma > A) = P(d_j^\sigma > AN) = P\left(d_j > \frac{AN}{\sigma^2}\right). \quad (27)$$

Теорема Шустера является основой для построения критериев выделения сигнала из шума. Возможны 2 случая применения критерия выделения гармонического компонента к реальному ряду. В первом (чаще всего встречающемся) случае частота периодического компонента заранее не известна.

Зададим малое значение уровня значимости α и найдем такое значение D_{\max} отсчета периодограммы (22), что вероятность P_0 того, что какой-то произвольный отсчет периодограммы с номером k (исключая крайние отсчеты $k = 0$ и $k = N - 1$) превысит значение D_{\max} будет равна α .

Представим значение D_{\max} в следующем виде:

$$D_{\max} = \frac{\sigma^2}{N}X_1 = \sigma^2X_q, \quad \text{где} \quad X_q = \frac{X_1}{N}. \quad (28)$$

Тогда, как видно из анализа формулы (27), P_0 — это вероятность того, что хотя бы один отсчет периодограммы d_j будет больше величины X_1 . Вычислим величину $P_1 = 1 - P_0$ — вероятность того, что ни один из отсчетов периодограммы d_j не превысит X_1 . В силу независимости отсчетов временного ряда из соотношения (24) получим⁵:

$$P_1 = P(d_1 < X_1) \cdot P(d_2 < X_1) \cdot \dots \cdot P(d_{N-2} < X_1) = (1 - e^{-X_1})^{N-2}.$$

Вероятность $P_0 = 1 - P_1 = 1 - (1 - e^{-X_1})^{N-2}$. Из равенства $P_0 = \alpha$ легко найти, что

$$X_1 = -\ln\left(1 - (1 - \alpha)^{\frac{1}{N-2}}\right), \quad X_q = -\frac{\ln\left(1 - (1 - \alpha)^{\frac{1}{N-2}}\right)}{N}. \quad (29)$$

Отметим, что выражение (29) близко к формуле (24) в [13], однако в цитированной работе вместо степени $1/(N - 2)$ используется степень $2/(N - 2)$.

При числе временных отсчетов $N \gg 1$ также и $N - 2 \gg 1$. Пользуясь приближенным соотношением $\ln(1 - \alpha)^K \approx 1 - \alpha K$ при $\alpha \ll 1$ получим:

$$X_1 \approx \ln\left(\frac{N - 2}{\alpha}\right), \quad X_q \approx \frac{1}{N} \ln\left(\frac{N - 2}{\alpha}\right). \quad (30)$$

Зависимость параметра X_1 в зависимости от числа N отсчетов временного ряда в функции уровня значимости α приведена на Рис. 2 (слева). Слева на этом рисунке сравниваются точные значения параметра X_1 , вычисленные по формуле (29) с приближенными значениями, рассчитанными согласно (30) для значения $\alpha = 0.1$. Даже для этого значения разность точной и приближенной величин X_1 незначительна. Для меньших значений α эта разность пренебрежимо мала.

Основываясь на приведенных выше формулах (28 - 29) можно сформулировать критерий \mathcal{K} присутствия гармонического компонента в анализируемом временном ряду.

Если выполняется неравенство

$$D_{\max} \geq \frac{\sigma^2}{N}X_1 \quad \text{или} \quad D_{\max} \geq \sigma^2X_q, \quad (31)$$

⁵Исключая крайние отсчеты $k = 0$ и $k = N - 1$ временного ряда

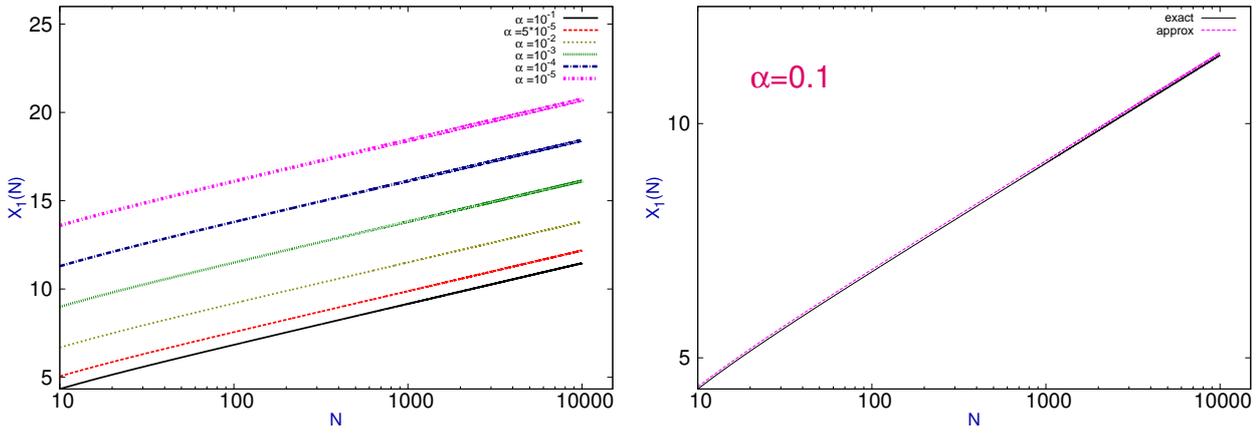


Рис. 2: Параметр X_1 в зависимости от числа N отсчетов временного ряда (слева). Сравнение точной и приближенной зависимостей параметра X_1 от N для $\alpha = 0.1$ (справа).

где величины X_1 и X_q определяются формулой (29), где D_{\max} — наибольший отсчет периодограммы, то с вероятностью $1 - \alpha$ принимается утверждение, что отсчет D_{\max} соответствует сигналу, а не шуму.

Если же частота искомого сигнала уже известна и соответствует отсчеты периодограммы D_l , то, используя вышеописанные рассуждения, критерий \mathcal{K} можно сформулировать согласно [13]) таким образом: если выполняется неравенство

$$D_l \geq \frac{\sigma^2}{N} X_2, \quad (32)$$

где

$$X_2 = \ln\left(\frac{1}{\alpha}\right), \quad (33)$$

то с вероятностью $1 - \alpha$ можно утверждать, что отсчет D_l соответствует сигналу, а не шуму.

Следует отметить, что теорема Шустера об экспоненциальном распределении отсчетов периодограммы центрированного белого шума справедлива только в случае, когда известна дисперсия ряда. На практике ее приходится оценивать по формуле

$$\sigma^2 = \frac{1}{N-1} \sum_{k=0}^{N-1} (f_k)^2, \quad (34)$$

где f_k — отсчеты анализируемого временного ряда, то выражение (29) для предельных отсчетов следует заменить на величины из Табл. 1 в [13].

3 Выделение гармонических компонентов для неравномерных временных рядов

3.1 Статистика отсчетов периодограммы белого шума

Статистика отсчетов периодограммы для равномерных временных рядов, отсчеты которых представляют собой белый шум, определяется теоремой Шустера и формулой (23). В то же время для неравномерных временных рядов данная теорема, вообще говоря, не верна.

Тем не менее, многие исследователи продолжают пользоваться формулами (29) и (31) для получения границы отсчетов периодограммы, такой, что отсчеты, ее превышающие, считаются принадлежащими гармоническим компонентам временного ряда, что может для сильно неравномерных рядов привести к значительным ошибкам.

По этой причине целесообразно построить эмпирическую функцию распределения отсчетов периодограммы белого шума для заданной временной сетки и использовать ее при выделении сигнала из шума. Опишем простой алгоритм построения такой эмпирической функции.

Пусть временная сетка описывается набором N значений временных отсчетов $\{T\} = \{t_k\}$, $k = 0, 1, \dots, N - 1$.

1) Зададим для каждого из временных отсчетов значение η_k случайной величины η — нормально распределенной случайной величины с нулевым математическим ожиданием и единичной дисперсией (см. пункт 2.2). Для генерации значений случайной величины η можно воспользоваться алгоритмом Бокса-Мюллера⁶.

2) Определим равномерную сетку по частоте в интервале $[0, \nu_{\text{Nq}}, \nu_j = j * \Delta\nu, j = 1, 2, \dots, N_{\text{max}}$, где частота Найквиста для неравномерного ряда определяется формулой (13). При N четном $N_{\text{max}} = N/2$, при нечетном — $N_{\text{max}} = (N + 1)/2$, а $\Delta\nu = \nu_{\text{Nq}}/N_{\text{max}}$.

3) Для каждой точки сетки по частоте рассчитаем значения d_j — нормированной периодограммы для частоты ν_j , то есть всего N_{max} значений.

4) Выполним $M \gg 1$ генераций временного ряда значений η_k случайной величины η и рассчитаем $N_{\text{max}} \times M$ значений случайной величины d_j .

5) Определим границы изменения величины d_j и разобьем интервал изменений этой величины $[d_j^{\text{min}}, d_j^{\text{max}}]$ на N_d промежутков длиной $\Delta d = (d_j^{\text{max}} - d_j^{\text{min}})/N_d$. Так как все значения $d_j \geq 0$, то удобно положить $d_j^{\text{min}} = 0$.

6) Рассчитаем эмпирическую дифференциальную функцию распределения величины d_j по следующей формуле:

$$f(d_j^k + \Delta d/2) = \frac{N(d_j^k, d_j^k + \Delta d)}{M\Delta d} = \frac{N(d_j^k, d_j^{k+1})}{M\Delta d}, \quad (35)$$

отнеся значение эмпирической функции распределения к середине промежутка $[d_j^k, d_j^{k+1}]$, где $d_j^k = d_j^{\text{min}} + k\Delta d$, $k = 0, 1, \dots, N_d - 1$. Здесь $N(d_j^k, d_j^{k+1})$ — число значений случайной величины d_j , попадающих в интервал $[d_j^k, d_j^{k+1}]$.

На Рис. 3 показано распределение отсчетов периодограммы белого шума d_j .

Для вычисления значения $X_1 = X_1(\alpha) = d_j^\alpha$, соответствующего уровню значимости α можно воспользоваться следующей процедурой. На первом шаге следует упорядочить все полученные значения случайной величины d_j по возрастанию. Далее следует найти такое

⁶см, например, <https://habrahabr.ru/post/208684/>

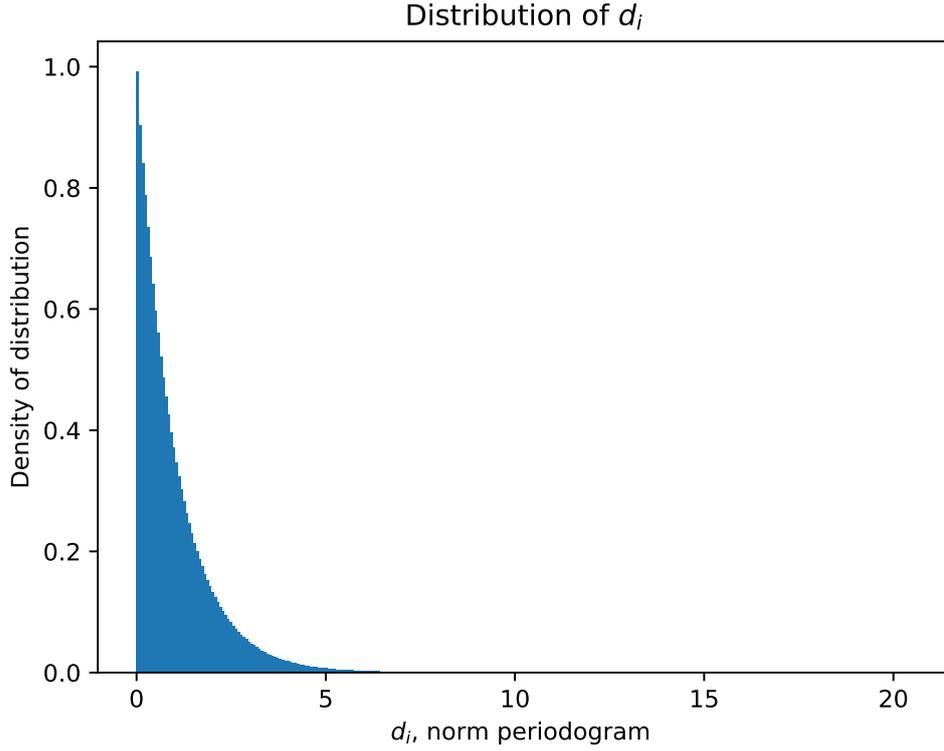


Рис. 3: Функция распределения отсчетов d_j периодограммы белого шума.

значение $d_j^*(\alpha)$, что число $N^*(\alpha)$ значений отсчетов d_j , превышающих $d_j^*(\alpha)$ будет равно $[\alpha N_{\max} M]$, где прямыми скобками обозначена целая часть числа.

Полученное значение $d_j^*(\alpha)$ как раз и будет оценкой величины $X_1 = d_j^\alpha$, тем более точной, чем больше было выполнено число M генераций временного ряда.

Далее следует для значений $\alpha = 10^{-2}, 10^{-3}, 10^{-4}, \dots$ найти значения d_j^α , определяющие для каких отсчетов периодограммы при заданном значении α можно утверждать о наличии гармонического сигнала в анализируемом временном ряде. Естественно, что значения d_j^α зависят от конкретной временной сетки.

Для тестирования вышеописанной методики целесообразно рассчитать данным методом значения $X_1(\alpha) = d_j^\alpha$ для равномерной временной сетки и сравнить с величинами, определяемыми по формуле (29).

Такие расчеты были выполнены нами для временной сетки наблюдений звезды ρ Leo [9]. Полученная зависимость $X_1(\alpha)$ представлена на Рис. 4. Из анализа рисунка видно, что для равномерного ряда зависимость $X_1(\alpha)$ близка к получаемой по формуле Шустера (29).

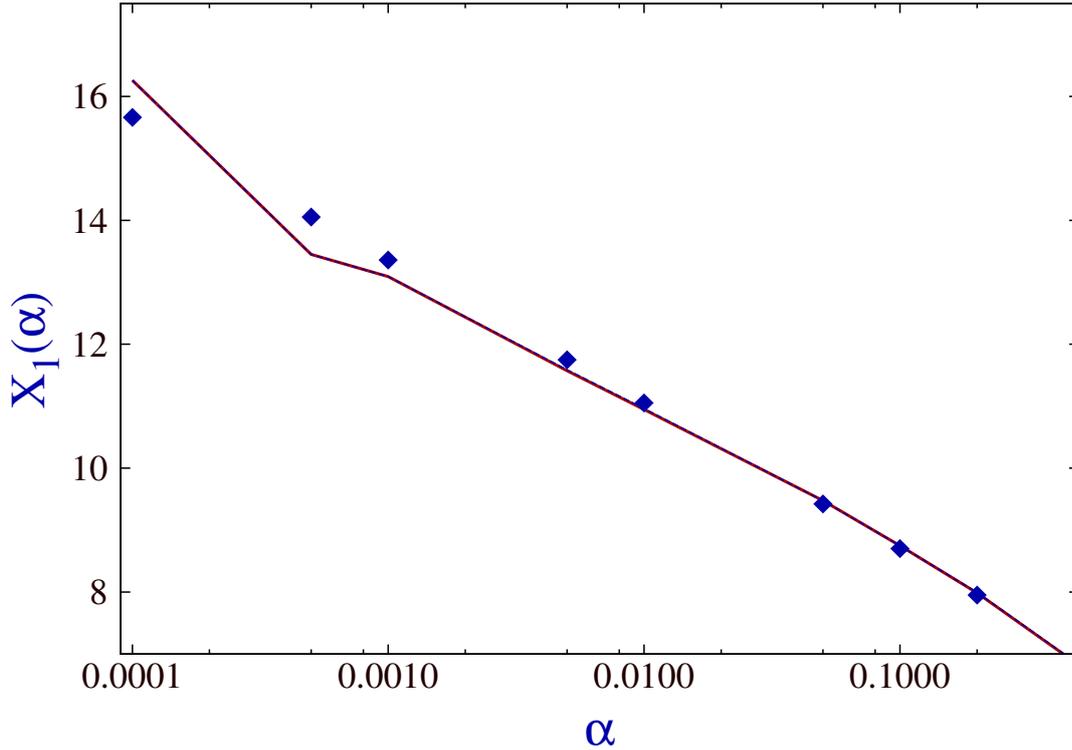


Рис. 4: Зависимость граничной величины X_1 от уровня значимости α . Сплошная линия — численные расчеты для $N = 14000$ модельных рядов белого шума. Прямоугольники — расчеты по формуле Шустера (29).

3.2 Статистика максимальных отсчетов периодограммы белого шума

Описанную выше процедуру построения функции распределения отсчетов периодограммы d_j можно также использовать для получения распределения максимальных отсчетов периодограммы d_j^{\max} .

Полученная функция распределения представлена на Рис. 5. Данная функция распределения оказалась достаточно узкой, что подтверждает корректность используемой нами методики определения величины $X_1(\alpha)$.

4 Дискретное преобразование Фурье

При вычислении преобразования Фурье (3) временного ряда $f(t)$ следует использовать какие-либо интегральные формулы. Наиболее часто используется так называемое *дискретное преобразование Фурье (DFT)*:

$$\tilde{F}(\nu) = \sum_{k=0}^{N-1} f_k e^{-i2\pi\nu t_k} = \sum_{k=0}^{N-1} f_k [\cos(2\pi\nu t_k) - \sin(2\pi\nu t_k)i], \quad (36)$$

где $f_k = f(t_k)$ — значение анализируемой функции при $t = t_k$. Здесь $k = 0, 1, \dots, N - 1$, где N — длина временного ряда.

Обычно DFT используют для равномерной сетки по времени и по частоте с равным числом отсчетов по времени и по частоте (см., например, [2, 13]). Выражение же (36) более общее и применимо и для неравномерных сеток.

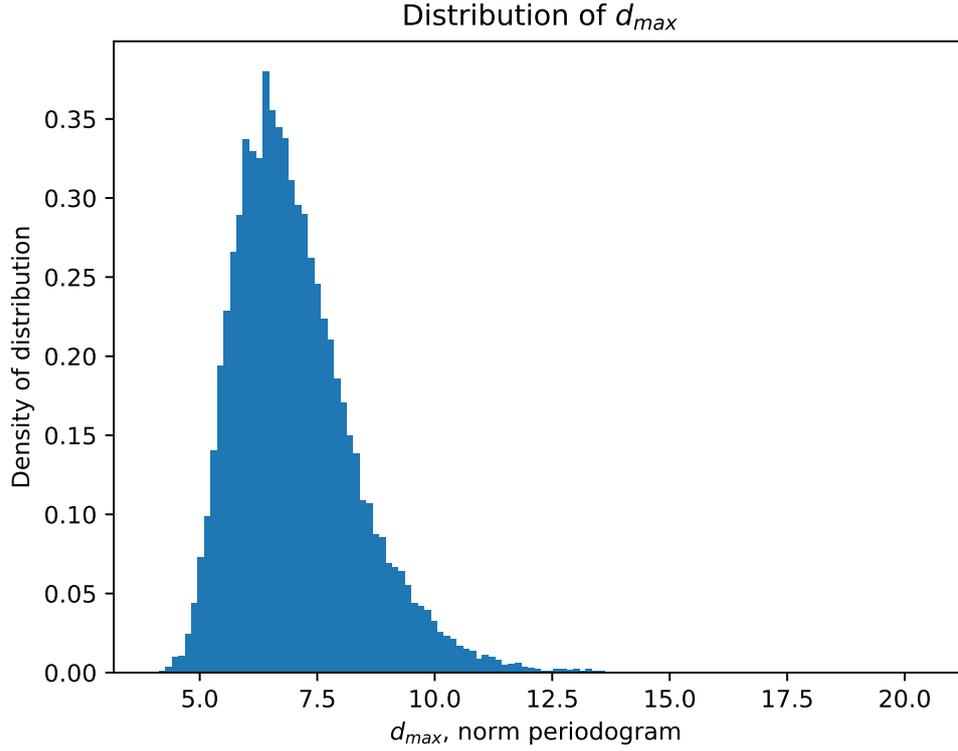


Рис. 5: Функция распределения максимальных отсчетов периодограммы белого шума.

Амплитуда дискретного преобразования Фурье

$$|\tilde{F}(\nu)|^2 = \sum_{k=0}^{N-1} |f_k e^{-i2\pi\nu t_k}|^2, \quad (37)$$

Сравнивая это выражение с формулой для периодограммы Шустера $D(\nu)$ (12), получим:

$$D(\nu) = \frac{1}{N^2} |\tilde{F}(\nu)|^2. \quad (38)$$

Нормированная периодограмма Шустера

$$d(\nu) = D(\nu) \cdot N = \frac{1}{N} |\tilde{F}(\nu)|^2, \quad (39)$$

Таким образом можно сделать вывод, что статистические свойства амплитуды дискретного преобразования Фурье (37) с точностью до множителя такие же, как у периодограммы Шустера.

4.1 Независимость от сдвига начала отсчета временного ряда

Традиционно временные ряды рассматриваются с начального отсчета $T_0 = 0$. В то же время реальные временные ряды зачастую имеют ненулевые начальные отсчеты. Для изучения того, существует ли зависимость дискретного преобразования Фурье (периодограмм Шустера) от начального отсчета рассмотрим модельный равномерный временной ряд и его копии, сдвинутые по оси времени.

Пусть модельный временной ряд, представляет собой сумму двух косинусоид:

$$F(t) = A_1 \cos(2\pi\nu_1 t + \phi_1) + A_2 \cos(2\pi\nu_2 t + \phi_2), \quad (40)$$

где $A_1 = 1.5, \nu_1 = 2.5, \phi_1 = 0.2$ и $A_2 = 2.0, \nu_2 = 4.0, \phi_2 = 0.5$ с начальными отсчетами $T_0 = 0.0, T_0 = 10.0$ и $T_0 = 50.0$ дней. Длина начального и сдвинутых модельных рядов составляет 5 дней, а разность между временными отсчетами составляет 0.02 дня. При таких параметрах каждый временной ряд состоит из 251 отсчета.

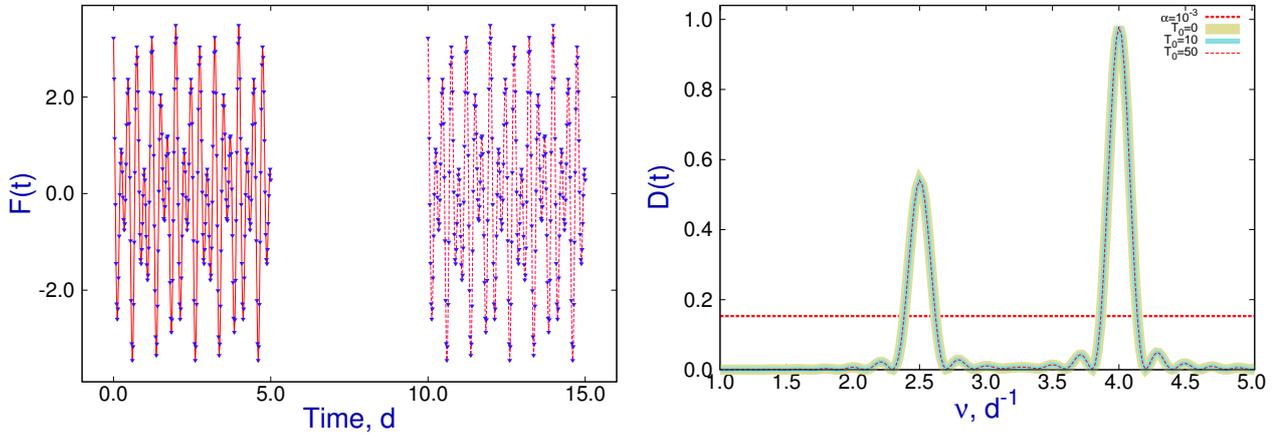


Рис. 6: Временные ряды с начальными отсчетами $T_0 = 0.0$ и $T_0 = 10.0$ (слева), периодограммы Шустера для временных рядов с начальными отсчетами $T_0 = 0.0, T_0 = 10.0$ и $T_0 = 50.0$ (справа).

На рисунке 6 (слева) показаны исходный и сдвинутый к началу отсчету $T_0 = 10.0$ дня временные ряды. Рассчитанные по начальному и сдвинутым на $T_0 = 10.0$ и $T_0 = 50.0$ дней периодограммы показаны на Рис. 6 (справа). Анализ рисунка показывает, что сдвиг начального отсчета временного ряда не меняет периодограммы (и связанное с ними) дискретное преобразование Фурье.

5 Оконное преобразование Фурье

Для длинных временных рядов, в которых свойства гармонических компонентов меняются со временем, бывает целесообразно использовать оконное Фурье-преобразование, например, [1]:

$$\tilde{F}_w(t, \nu) = \int_{-\infty}^{\infty} f(\tau)W(\tau - t)e^{-i2\pi\nu\tau} d\tau, \quad (41)$$

где $W(t)$ — оконная функция.

Для численно заданной функции удобно использовать дискретное оконное Фурье-преобразование (wDFT):

$$F_w(t_m, \nu) = \sum_{k=-\infty}^{\infty} f(t_k)W(k - m)e^{-i2\pi\nu t_k}, \quad (42)$$

где $W(n)$ — дискретная оконная функция. Дискретная оконная функция $W(n)$ зависит от разности $k - m$. В формуле (42) формально суммирование производится по интервалу значений $k \in [-\infty, \infty]$. Реально временной ряд задан на массиве временных отсчетов $\mathcal{T} = \{t_0, t_1, \dots, t_{N-1}\}$. Здесь N — полное число временных отсчетов. За границами промежутка $[t_0, t_{N-1}]$ значения ряда задаются равными нулю. В стандартном представлении [3, 13] wDFT вычисляется для значений $t = t_m$, где $0 \leq m \leq N - 1$.

Простейший тип оконной функции — прямоугольное окно ширины M :

$$W_{\text{rect}}(n) = \begin{cases} 1 & , \quad n \in [0, M - 1], \\ 0 & , \quad n \notin [0, M - 1]. \end{cases} \quad (43)$$

Для прямоугольного окна дискретное оконное Фурье-преобразование будет иметь следующий вид:

$$F_w(t_m, \nu) = \sum_{k=m}^{m+M-1} f(t_k)e^{-i2\pi\nu t_k}, \quad (44)$$

Такое окно получается автоматически при ограничении выборки M отсчетами: $k \in [m, m + M - 1]$.

Для произвольного окна wDFT записывается так:

$$F_w(t_m, \nu) = \sum_{k=m}^{m+M-1} W(k - m)f(t_k)e^{-i2\pi\nu t_k}, \quad (45)$$

Если индекс k выходит за пределы промежутка $[0, N - 1]$, то значения t_k полагаются равными нулю.

Часто используются следующие окна высокого разрешения.

А) Окно Ханна (Hann window):

$$W(n) = 0.5 \left(1 - \cos \left[\frac{2\pi n}{N - 1} \right] \right), \quad n \in [0, N - 1]. \quad (46)$$

Б) Окно Хемминга (Hamming window):

$$W(n) = 0.53836 - 0.46164 \cos \left[\frac{2\pi n}{N - 1} \right], \quad n \in [0, N - 1]. \quad (47)$$

Полный список оконных функций представлен на WWW странице [5].

Использование дискретных оконных функций (43)–(47) наиболее целесообразно использовать для равномерных рядов с достаточно близко расположенными временными отсчетами. В случае неравномерных рядов с существенно различающимися временными интервалами между соседними отсчетами использование дискретных оконных функций может быть неудобно.

Модифицируем оконные функции следующим образом. Определим ширину окна

$$T_w < T_{\max}, \quad (48)$$

где $T_{\max} = T_{N-1} - T_0$ — полная длина временного ряда.

Для прямоугольного окна непрерывная оконная функция

$$W_{T_w}(t_0, t) = W_{\text{rect}}(t_0, t) = \begin{cases} 1 & , t \in [t_0, t_0 + T_w], \\ 0 & , t \notin [t_0, t_0 + T_w]. \end{cases} \quad (49)$$

Здесь t_0 — начало окна.

При таком выборе оконной функции использование оконного преобразования Фурье эквивалентно применению Фурье-преобразования к отрезку ряда со значениями в промежутке $[t_0, t_0 + T_w]$.

Удобнее вместо размерного времени t использовать безразмерную временную переменную $\tau = (t - t_0)/T_w$. Переменная $\tau \in [0, 1]$. Сделав замену $(t - t_0)/T_w \rightarrow \tau$ в формуле (49), получим:

$$W_{\text{rect}}(t_0, t) = W_{\text{rect}}(t_0, \tau) = \begin{cases} 1 & , \tau \in [0, 1], \\ 0 & , \tau \notin [0, 1]. \end{cases} \quad (50)$$

Анализируя дискретные оконные функции, (46) и (47), получим, что переменная τ соответствует отношению $n/(N - 1)$. С учетом этого соответствия более сложные непрерывные оконные функции будут иметь следующий вид:

А) Окно Ханна (Hann window):

$$W_{\text{Hann}}(t_0, \tau) = \begin{cases} 0.5 - 0.5 \cos(2\pi\tau) & , \tau \in [0, 1], \\ 0 & , \tau \notin [0, 1]. \end{cases} \quad (51)$$

В) Окно Хемминга (Hamming window):

$$W_{\text{Hammm}}(t_0, \tau) = \begin{cases} 0.53836 - 0.46164 \cos(2\pi\tau) & , \tau \in [0, 1], \\ 0 & , \tau \notin [0, 1]. \end{cases} \quad (52)$$

С) Окно Блэкмана (Blackman window):

$$W_{\text{Blackm}}(t_0, \tau) = \begin{cases} a_0 - a_1 \cos(2\pi\tau) + a_2 \cos(4\pi\tau) & , \tau \in [0, 1], \\ 0 & , \tau \notin [0, 1], \end{cases} \quad (53)$$

где $a_0 = 0.42659$, $a_1 = 0.49656$, $a_2 = 0.076849$ (exact Blackman, [4]).

Д) Окно Наталла (Nuttall window):

$$W_{\text{Nutt}}(t_0, \tau) = \begin{cases} a_0 - a_1 \cos(2\pi\tau) + a_2 \cos(4\pi\tau) - a_3 \cos(6\pi\tau) & , \tau \in [0, 1], \\ 0 & , \tau \notin [0, 1], \end{cases} \quad (54)$$

где $a_0 = 0.355768$, $a_1 = 0.487396$, $a_2 = 0.144232$, $a_3 = 0.012604$ ([4]).

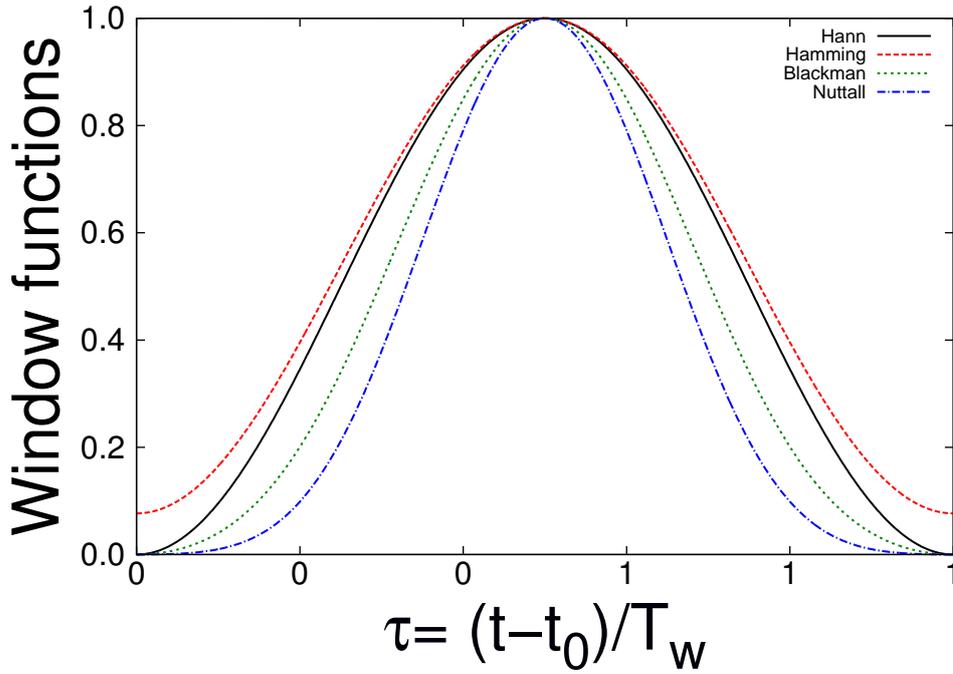


Рис. 7: Оконные функции Ханна, Хемминга, Блэкмана и Наталла.

Для иллюстрации на Рис. 7 представлены описанные выше оконные функции.

На Рис. 8 показаны результаты Фурье-преобразования оконных функций Хемминга и Наталла. Эти графики иллюстрируют общее поведение Фурье-спектров оконных функций. Чем шире сама оконная функция, тем уже центральный пик Фурье-спектра, определяющий разрешение окна, и тем выше интенсивность его боковых лепестков. Напротив, узкие оконные функции, такие как функция Наталла имеют широкий центральный пик (низкое разрешение) и слабые боковые лепестки.

Прямоугольное окно является окном самого высокого разрешения, но с максимальным уровнем боковых лепестков, окна А и В являются окнами высокого разрешения с низким уровнем боковых лепестков, окна С и D являются окнами низкого разрешения с очень низким уровнем боковых лепестков (см. подробнее в [5]).

Используя непрерывные оконные функции, запишем дискретное оконное преобразование Фурье следующим образом:

$$F_w(t_0, T_w, \nu) = \sum_{k, t_k \in [t_0, t_0 + T_w]} f(t_k) W\left(t_0, \frac{t_k - t_0}{T_w}\right) e^{-i2\pi\nu t_k}, \quad (55)$$

Из определения оконных функций (49)–(54) следует, что суммирование в формуле (55) выполняется только для тех отсчетов t_k временного ряда, которые попадают в промежуток $[t_0, t_0 + T_w]$. Это означает, в частности, что возможные значения $t_k \in [t_0, t_0 + T_w]$.

Отметим, что:

- а) Число отсчетов, попадающих в промежуток $[t_0, t_0 + T_w]$ зависит от t_0 и может быть различным для различных значений t_0 ,
- б) Значение t_0 не обязательно должно совпадать с каким либо из отсчетов временного ряда и может отличаться от них.

Получаемый в результате применения оконного преобразования Фурье-спектр зависит от трех параметров: частоты ν , начального положения окна t и ширины окна T_w .

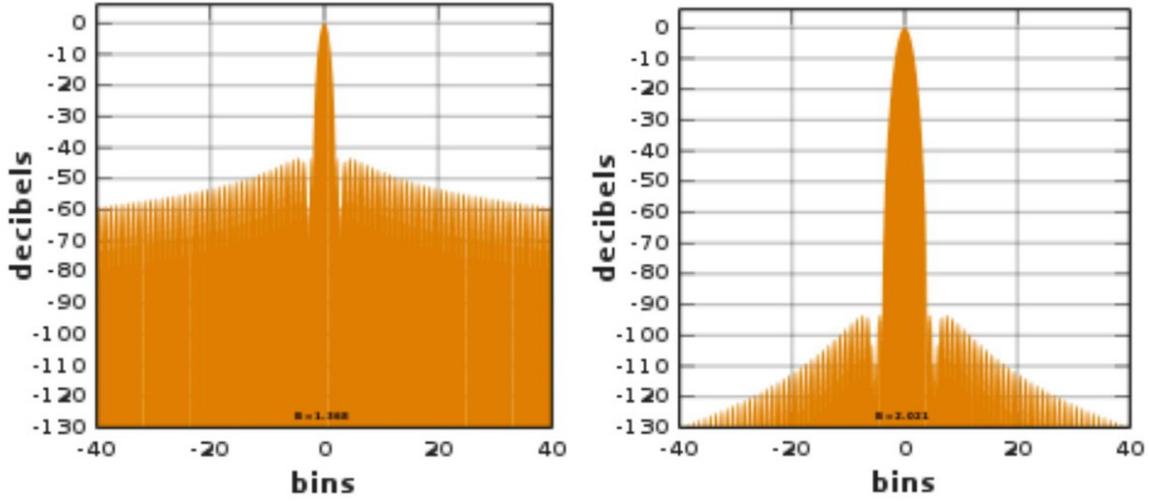


Рис. 8: Фурье-спектры оконных функций Хемминга (слева) и Наталла (справа) согласно [6].

5.1 Периодограмма Шустера для оконного преобразования Фурье

Используя формулу (55) запишем выражение для аналога периодограммы Шустера для оконного преобразования Фурье:

$$D_w(t_0, T_w, \nu) = \frac{1}{N(t_0, T_w)^2} \left| \sum_{k, \tau_k \in [0,1]} f(t_k) W(t_0, \tau_k) e^{-i2\pi\nu t_k} \right|^2. \quad (56)$$

Здесь $N(t_0, T_w)^2$ – нормировочный коэффициент, определяемый как число отсчетов временного ряда в промежутке $[t_0, t_0 + T_w]$, а величина $\tau_k = (t_k - t_0)/T_w$. В случае, если в промежуток $[t_0, t_0 + T_w]$ не попал ни один отсчет временного ряда, то следует положить $D_w(t_0, T_w, \nu) = 0$.

Возможен другой подход к определению нормировочного коэффициента (см, например, формулу (16) в [8]):

$$N_{\text{exact}}^{\text{mod}}(t_0, T_w) = \int_0^1 W(t_0, \tau) d\tau, \quad \text{где, как и ранее} \quad \tau = \frac{t - t_0}{T_w}. \quad (57)$$

Численный аналог данного модифицированного нормировочного коэффициента зависит от величины t_0 и выражается следующим образом:

$$N^{\text{mod}}(t_0, T_w) = \sum_{k, \tau_k \in [0,1]} W(t_0, \tau_k). \quad (58)$$

В этом случае

$$D_w^{\text{mod}}(t_0, T_w, \nu) = \frac{1}{N^{\text{mod}}(t_0, T_w)^2} \left| \sum_{k, \tau_k \in [0,1]} f(t_k) W(t_0, \tau_k) e^{-i2\pi\nu t_k} \right|^2. \quad (59)$$

Определим следующие величины:

$$P = \sum_{k, \tau_k \in [0,1]} f(t_k) W(t_0, \tau_k) \cos(2\pi\nu t_k) \quad (60)$$

и

$$Q = \sum_{k, \tau_k \in [0,1]} f(t_k)W(t_0, \tau_k) \sin(2\pi\nu t_k). \quad (61)$$

Периодограмма Шустера для оконного преобразования Фурье примет следующий вид:

$$D_w(t_0, T_w, \nu) = \frac{1}{N(t_0, T_w)^2} |P^2 + Q^2|, \quad (62)$$

И аналогично для модифицированного нормировочного коэффициента (58).

В силу того, что формулы для вычисления оконного преобразования Фурье по формулам (56) и (59) включают весовые коэффициенты $W(t_0, \tau_k)$, следует изменить выражение (34) для дисперсии временного ряда⁷. Для конкретного окна с началом t_0 и шириной T_w дисперсию ряда следует оценивать таким образом:

$$\sigma_w^2(t_0, T_w) = \frac{1}{N(t_0, T_w) - 1} \sum_{k, \tau_k \in [0,1]} (W(t_0, \tau_k) \times f_k)^2, \quad (63)$$

Значение $\sigma_w^2(t_0, T_w)$ для конкретного окна следует использовать при вычислении величины D_{\max} , определенной формулой (28).

5.1.1 Выбор интервала частот при расчете периодограммы Шустера для оконного преобразования Фурье

В отличие от равномерного ряда, у неравномерного нет фундаментальной системы частот, потому выбор области частот, в которой вычисляется периодограмма Шустера (56) определяется в достаточной степени произвольно, исходя из поставленной задачи. Более подробно выбор частоты можно описать следующим образом:

В отрезке ряда $[T_i, T_i + T_w]$ частота выбирается в интервале $[\nu_{\min}, \nu_{\max}]$, при этом в самом интервале частота может быть выбрана с постоянным шагом $\Delta\nu = (\nu_{\max} - \nu_{\min})/N_\nu$, где N_ν — число промежутков разбиения частоты. Частота $\nu_{\min} = 1/T_w$. При меньшей частоте длина соответствующей синусоиды с периодом $P = 1/\nu$ больше длины окна T_w . Интервал $[\nu_{\min}, \nu_{\max}]$ нужно выбирать единым для для всех окон.

Частота ν_{\max} — это частота Найквиста ν_{Nq} . Для неравномерного ряда частота Найквиста не вполне определена. Можно использовать 2 определения:

- $\nu_{Nq}^1 = 1/(2 * \delta T_{\text{mean}})$, где δT_{mean} — средний промежуток времени между отсчетами.
- $\nu_{Nq}^2 = 1/(2 * \delta T_{\text{min}})$, где δT_{min} — минимальный промежуток времени между отсчетами в отрезке $[T_i, T_i + T_w]$. Очевидно, что $\nu_{Nq}^1 \leq \nu_{Nq}^2$.

Для равномерного ряда эти частоты равны. Обычно выбирают частоту ν_{Nq}^1 . И только когда ряд очень неравномерен, можно выбрать в качестве ν_{\max} какую-то частоту в интервале $[\nu_{Nq}^1, \nu_{Nq}^2]$.

Выбор шага $\Delta\nu$ (или соответствующего числа N_ν) в достаточной степени произволен. Чем меньше $\Delta\nu$, тем лучше можно описать соответствующий Фурье-спектр, однако при этом объем данных существенно возрастает. Наилучший выбор величины $\Delta\nu$ определяется опытным путем.

В случае значительной (на порядки величины) разности величин ν_{\min} и ν_{\max} вместо равномерного шага по частоте можно выбрать логарифмический шаг, то есть равномерный шаг по логарифму частоты $\log(\nu)$.

⁷В данном случае дисперсия вычисляется для рассматриваемого отрезка временного ряда

5.2 Оконное преобразования Фурье модельных функций

5.2.1 Модельные функции с постоянным периодом

Рассмотрим модельный временной ряд, описываемый формулой (40), с такими же параметрами, которые описаны в пункте 4.1, но с полной длительностью ряда $T = 20$ дней. На Рис. 9 представлены результаты применения оконного Фурье-преобразования (59) к данному временному ряду для фиксированного окна $[10, 15]$ и различных типов оконных функций.

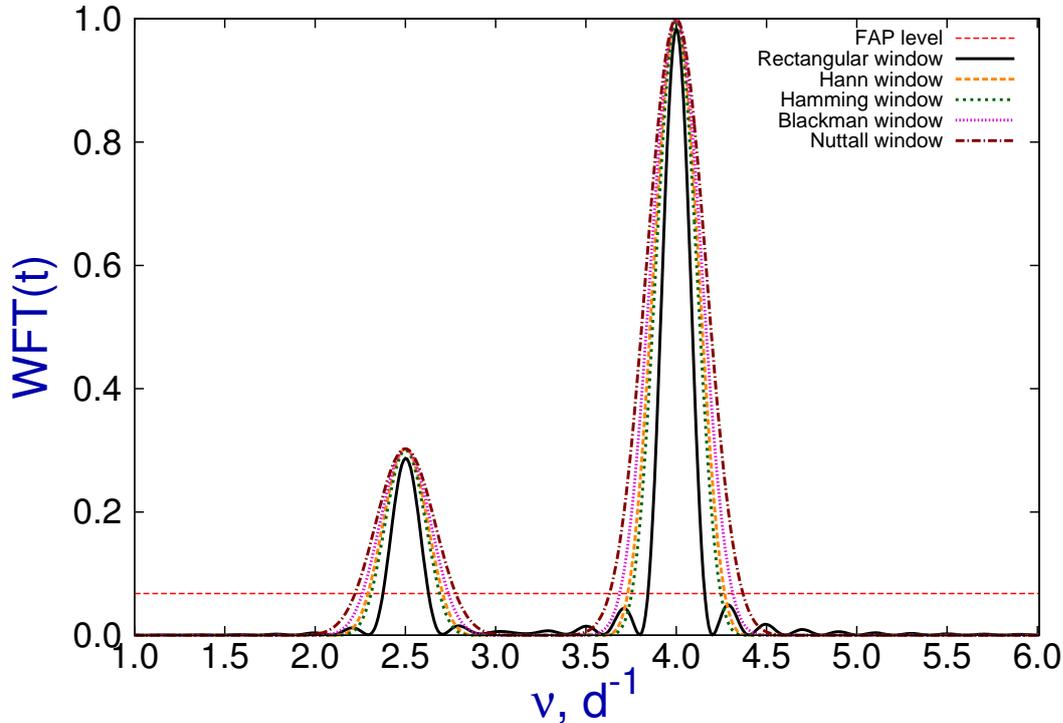


Рис. 9: Оконное Фурье-преобразование модельной функции (40) при $T = 20$ дней и фиксированного окна $t_0 = 10.0$ с шириной $T_w = 5.0$. Пунктиром показано значение D_{\max} для уровня значимости $\alpha = 10^{-3}$ и для прямоугольного окна.

Использование модифицированного нормировочного коэффициента (58) позволяет, как видно на рисунке, получить близкие амплитуды для различных типов оконных Фурье-преобразований.

Таблица 1: Положения максимумов оконной Фурье-периодограммы ν_{\max} и величин D_{\max} для окна $[10, 15]$ при значении $\alpha = 10^{-3}$ и различных оконных функции в сравнении с точными значениями.

Окно	Max ₁	Max ₂	D_{\max}
Rectangular	2.45	4.04	0.0676
Hann	2.53	3.99	0.0342
Hamming	2.50	4.01	0.0368
Blackman	2.50	4.00	0.0288
Nuttall	2.45	4.02	0.0244
exact	2.50	4.00	—

В таблице 1 сравниваются положения максимумов оконной Фурье-периодограммы и

величин D_{\max} для окна $[t_0, t_0 + T_w] = [10, 15]$. Отметим, что наиболее узкие максимумы периодограммы получаются при использовании прямоугольного окна, однако при этом положения максимумов получаются с меньшей точностью. Наиболее точные положения соответствуют окну Блэкмана.

В общем случае оконное преобразование зависит от положения окна. Однако, если окно шире, чем наибольший период самой медленной компоненты временного ряда, то эта зависимость практически отсутствует.

В общем случае периодограмма Шустера (59) является функцией трех переменных t_0 , T_w и ν . При фиксировании ширины окна T_w периодограмма $D_w^{\text{mod}}(t_0, T_w, \nu)$ будет функцией двух переменных t_0 и ν . Отметим, что начала окна $t_0 \in [0, T - T_w]$. При большем значении t_0 окно выходит за границы задания временного ряда.

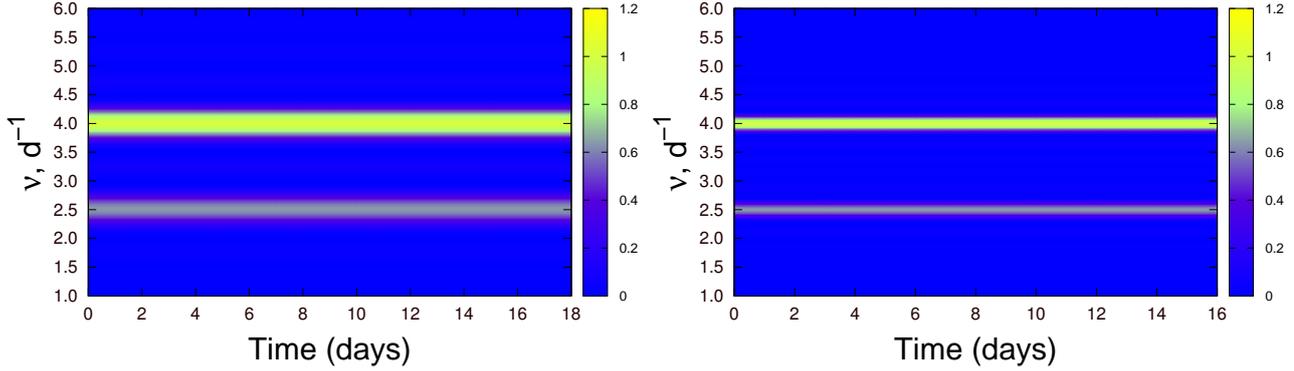


Рис. 10: Оконное Фурье-преобразование модельной функции с постоянными периодами 2.5 и 4.0 дня в интервале $t_0 \in [0, 20]$ дня для окна шириной $t_w = 2$ дня (слева) и $t_w = 4$ дня (справа).

Оконное преобразование модельной функции (40) при $T = 20$ дней и значениях параметров $T_w = 2.0$ представлено на Рис. 10. Хорошо видно, что при увеличении ширины окна точность определения компонент временного ряда для всех возможных значений t_0 увеличивается.

5.2.2 Модельные функции с переменным периодом

Пусть модельный временной ряд, представляет собой косинусоиду с переменным периодом без шумового компонента:

$$F(t) = \begin{cases} 2.0 \cos(2\pi\nu(t)t + \phi_1) & , t \in [0, 25], \\ \nu(t) = \nu_0 + \nu_1 t & , \nu_0 = 2.5. \end{cases} \quad (64)$$

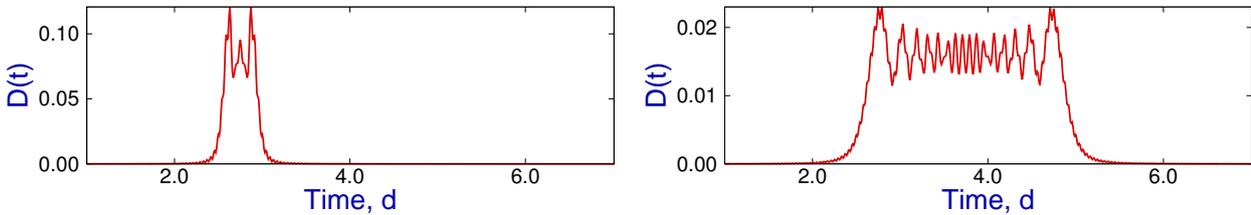


Рис. 11: Периодограмма Шустера модельной функции с переменным периодом, заданной в интервале $[0, 20]$ дня с параметром $\nu_1 = 0.01$ (слева) и $\nu_1 = 0.05$ (слева).

При таком задании временного ряда период изменяется от значения $\nu = 2.5 \text{ d}^{-1}$ при $t = 0$ до $\nu = 2.75 \text{ d}^{-1}$ при значении $\nu_1 = 0.01$ и до $\nu = 3.75 \text{ d}^{-1}$ при значении $\nu_1 = 0.05$. На Рис. 11 представлена периодограмма Шустера для модельного ряда (64) при указанных значениях ν_1 .

Как видно из анализа рисунка, даже небольшие изменения периода временного ряда приводят к значительному расплыванию периодограммы и невозможности точного определения периода гармонических компонент ряда.

Применение Фурье-анализа к коротким отрезкам ряда (64) позволяет точнее определить частоты таких компонент. Рассмотрим два отрезка временного ряда (64) $[0, 2]$ и $[20, 22]$. Периодограмма Шустера для этих отрезков представлена на Рис. 12.

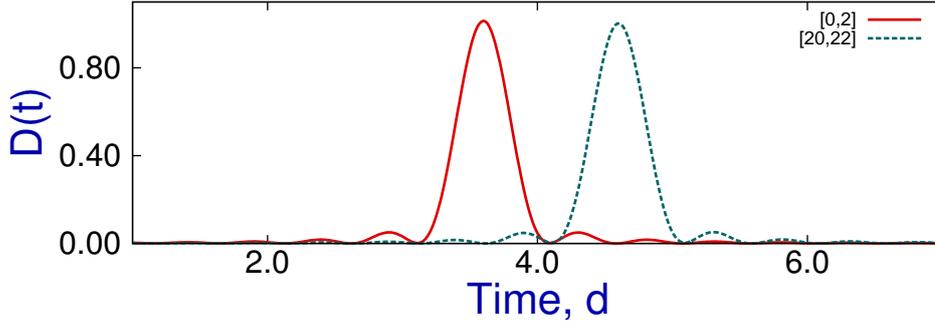


Рис. 12: Периодограмма Шустера отрезков модельной функции (64) в интервалах $[0, 2]$ дня (красная сплошная линия) и $[20, 22]$ дня (пунктир) для параметра $\nu_1 = 0.05$.

Как видно на рисунке, периодограмма для отрезка $[20, 22]$ сдвинута на $\approx 1.0 \text{ d}^{-1}$. Такой сдвиг связан с формой модельного ряда (64) и равен произведению $\nu_1 T$, что для $\nu_1 = 0.05$ и $T = 20$ дней как раз и равно 1 d^{-1} .

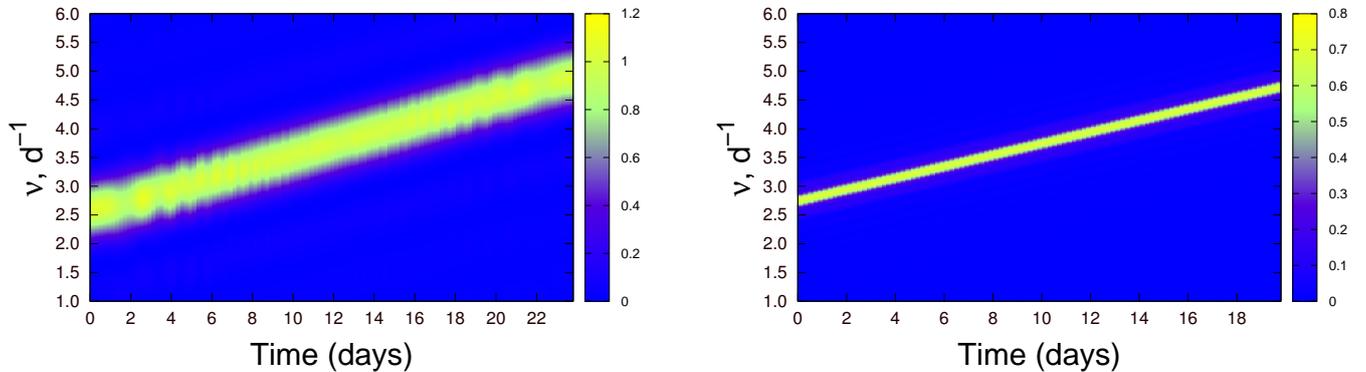


Рис. 13: Оконное преобразование модельной функции (64) для ширины окна $T_w = 1$ день (слева) и $T_w = 5$ дней (справа).

Для детального анализа временных рядов с переменными периодами следует использовать оконное преобразование (59). На Рис. 13 даны результаты применения оконного преобразования к модельному ряду (64) для прямоугольного окна с ширинами $T_w = 1$ и $T_w = 5$ дней. Видно, что увеличение ширины окна приводит к увеличению точности определения частоты, соответствующей конкретному отрезку временного ряда.

В силу конечности ширины окна T_w оконное преобразование может быть вычислено не для всей длительности временного ряда T , а для отрезка $T - T_w$, что хорошо

видно на Рис. 13. Также, как в рассмотренном выше случае, для построения периодограммы Шустера, оконное преобразование дает дополнительный сдвиг частоты порядка $\sim \nu_1(T - T_w) \sim 1 - 1.2$ дня. Этот сдвиг следует учитывать при интерпретации оконного преобразования реальных временных рядов.

6 Анализ вариаций профилей в спектрах звезд с использованием оконного преобразования Фурье

6.1 Вариации профилей в спектре ρ Leo

В статье [9] оконное преобразование Фурье с прямоугольным окном было использовано для анализа быстрых вариаций профилей линий в спектре звезды ρ Leo, полученных с высоким временным разрешением на 6-м телескопе БТА со спектрографом низкого разрешения SCORPIO в 2015 г.

Полученный оконный Фурье-спектр зависит от четырех параметров: частоты ν , t_0 — начального положения окна, T_w — ширины окна и V — доплеровского смещения на профиле линии от центральной длины волны линии.

Для полного профиля линии профиля мы имеем функцию четырех переменных:

$$F_w(t_0, \nu) = F_w(t_0, \nu, T_w, V).$$

Такую функцию весьма сложно анализировать, поэтому значения указанной функции были просуммированы для всех доплеровских смещений V в пределах профиля линии.

При фиксированном значении T_w усредненная по профилю функция $\bar{F}_w(t_0, \nu, T_w)$ становится функцией только двух параметров, t_0 и ν . Анализ результатов применения оконного преобразования Фурье к вариациям профилей линий в спектре ρ Leo показал, что оптимальное значение $\Delta T = 20$ минут.

На Рис. 14 представлен результат применения оконного Фурье-преобразования к вариациям профилей бальмеровских линий в спектре ρ Leo для частот в интервале $[0, 1.2]$ минут⁻¹.

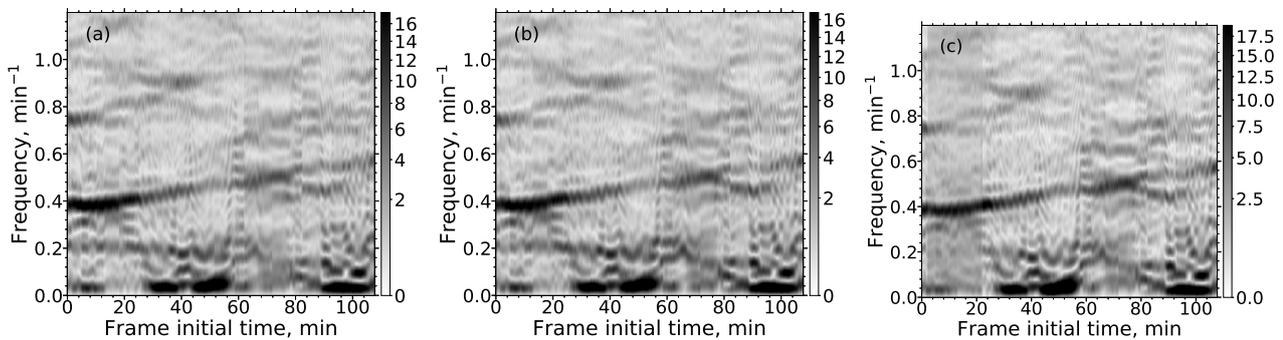


Рис. 14: Оконный Фурье-спектр вариаций профиля линии H_δ (a), H_γ (b) и H_β (c) для значения $\Delta T = 20$ минут в спектре ρ Leo в зависимости от частоты (в минут⁻¹) от положения начала окна t_0 (Frame initial time) в минутах от начала наблюдений.

В нижней части оконных Фурье периодограмм видны регулярные компоненты вариаций профилей, обнаруженные при анализе всех 1271, полученных на БТА спектров ρ Leo. В то же время, кроме этих компонент, обнаруживаются и новые. Наиболее ярко выражен компонент с переменной частотой $\nu = 0.4$ минут⁻¹ (период $P \sim 2.5$ минут) при $t_0 = 0$ и $\nu = 0.5$ минут⁻¹ при $t_0 = 100$ минут ($P \sim 2$ минуты). В области $t_0 \approx 70$ минут данный компонент раздваивается. Нижняя ветвь примерно соответствует частотам вариаций при $t_0 = 0$.

Кроме того, на периодограммах виден менее выраженный компонент с переменной частотой $\nu = 0.2$ минут⁻¹ ($P \sim 5$ минут) при $t_0 = 0$ и $\nu = 0.1$ минут⁻¹ ($P \sim 10$ минут) при

$t_0 = 110$ минут. Возможно также присутствие компонента с $\nu \sim 0.75$ минут⁻¹ с периодом $P \sim 1.3$ минут. Однако этот компонент прослеживается только до значения $t_0 \leq 10$ минут.

Указанные частоты не были обнаружены при применении преобразования Фурье ко всему временному ряду вариаций профилей линий в спектре ρ Leo. Таким образом можно сделать вывод, что использование оконного преобразования Фурье позволяет обнаружить новые гармонические компоненты временных рядов и, в том числе компоненты с переменной частотой.

Список литературы

- [1] И. Добеши, *Десять лекций по вейвлетам*, Ижевск, НИЦ РХД (2001)
- [2] https://ru.wikipedia.org/wiki/Дискретное_преобразование_Фурье
- [3] https://ru.wikipedia.org/wiki/Оконное_преобразование_Фурье
- [4] https://en.wikipedia.org/wiki/Window_function
- [5] <https://web.archive.org/web/20110325050652/http://dspssystem.narod.ru/add/win/win.html>
- [6] https://en.wikipedia.org/wiki/Window_function
- [7] В.Е.Гмурман, теория вероятностей и математическая статистика. Учебн. пособие для вузов, 9-е изд, М., изд. Высшая школа (2003)
- [8] А.Ю. Гришенцев, А.Г. Коробейников, С.А. Арустамов, "Цифровые системы широкополосной связи. Часть 2. оконные и вейвлет-функции и преобразования". Учебное пособие, изд. Университета ИТМО (2019)
- [9] А.Ф.Холтыгин, А.А.Батраков, С.Н.Фабрика и др., *Астрофизический Бюллетень*, **73**, 498 (2018)
- [10] D.H. Roberts, J. Lehar, J.W. Dreher, *AJ* **93**, 968 (1987)
- [11] A. Schuster, *Terrestrial Magnetism* **3**, 13 (1898)
- [12] В.Ю.Теребиж, *Анализ временных рядов в астрофизике*, М, изд. Наука (1992)
- [13] В.В.Витязев, *Спектрально-корреляционный анализ равномерных временных рядов*, СПб, изд. СПбГУ (2001)
- [14] В.В.Витязев, *Анализ неравномерных временных рядов*, СПб, изд. СПбГУ (2001)