

Санкт-Петербургский государственный университет
Прикладная математика и информатика

Производственная практика 2 (научно-исследовательская работа) (семестр 4)

ОБНАРУЖЕНИЕ РАЗЛАДКИ С ПОМОЩЬЮ МЕТОДА SSA

Выполнил:

Кононыхин Иван Александрович

группа 20.M03-мм

Научный руководитель:

к. ф.-м. н., доцент

Голяндина Нина Эдуардовна

Кафедра статистического моделирования

Санкт-Петербург

2022

Оглавление

Введение	4
Глава 1. Singular spectrum analysis	6
1.1. Алгоритм базового метода SSA	6
1.1.1. Вложение	6
1.1.2. Сингулярное разложение	6
1.1.3. Группировка	7
1.1.4. Реконструкция	7
1.2. Ранги ряда	7
1.2.1. Пример	8
Глава 2. Поиск разладки	9
2.1. Матрица неоднородности и функции неоднородности	9
2.1.1. Матрица неоднородности	9
2.1.2. Функции обнаружения	11
2.2. Однородность и неоднородность	12
2.2.1. Типы неоднородности	12
Глава 3. Численное сравнение функций разладки	14
3.1. Постановка задачи	14
3.2. Организация численного эксперимента	15
3.2.1. Ряды без шума	17
3.2.2. Ряды с шумом	23
3.3. Выводы	30
Глава 4. Аналитическая оценка индекса неоднородности при изменении частоты гармоник	31
4.1. Вычисление индекса однородности	32
4.1.1. Знаменатель	32
4.1.2. Числитель	32
4.2. Индекс неоднородности	33
4.3. Проверка точности аппроксимации	34

4.3.1.	Одинаковые частоты	34
4.3.2.	$L\omega_1$ и $L\omega_2$ целые, $\omega_1 \neq \omega_2$	34
4.3.3.	Предположения об L	34
4.3.4.	Разность ω_1 и ω_2	35
Глава 5. Система обнаружения структурной неоднородности ряда с ав-		
томатически выстраиваемым порогом срабатывания		37
5.1.	Оценка γ	38
5.1.1.	Нижняя граница	38
5.1.2.	Верхняя граница	38
5.2.	Аппроксимация переходного интервала	39
5.3.	Алгоритм работы	42
5.4.	Оценка системы	42
5.5.	Оценка параметров	44
5.5.1.	Оценка параметров: T	45
5.5.2.	Оценка параметров: V	49
Заключение		50
Список литературы		51

Введение

Будем называть временной ряд $F_N = (f_1, \dots, f_N)$ **однородным**, если он управляется некоторым линейным рекуррентным соотношением (LRR) $f_{n+r} = \sum_{i=1}^r \alpha_i \cdot f_{n+r-i}$, $\alpha_i \neq 0$, размерность которого (r) мала по отношению к N .

Предположим, что из-за внешнего воздействия (или по другой причине) однородный временной ряд подвергается мгновенному возмущению, то есть он перестает следовать исходному LRR. Однако по прошествии определенного периода времени он снова становится управляемым неким LRR, которое может отличаться от исходного. В результате, ряд в целом перестает быть однородным и возникает проблема изучения этой неоднородности.

Метод, основанный на алгоритме SSA (Singular Spectrum Analysis) [1], позволяет изучить эту неоднородность путем построения матрицы разладки, где для обнаружения неоднородности используются строковые, столбцовые и диагональные подмножества этой матрицы.

Целью данной работы является сравнение методов обнаружения неоднородности в синусоидальных рядах, аналитическая аппроксимация индекса неоднородности, анализ поведения переходного интервала функций обнаружения и создание системы обнаружения неоднородности.

Главы 1 и 2 носят реферативный характер.

В «Главе 1» приведена теория базового алгоритма SSA.

В «Главе 2» введены индекс неоднородности, матрица разладки, функции обнаружения и обзореваются типы неоднородности.

«Глава 3» посвящена численному сравнению функций обнаружения неоднородностей на синусоидальных временных рядах с шумом и без. Неоднородность рядов задавалась изменением частоты, амплитуды, фазовым сдвигом и выбросом.

«Глава 4» посвящена аналитическому анализу индекса неоднородности, его упрощению и аналитической аппроксимацией. Приведены численные эксперименты, показывающие условие сходимости аналитической аппроксимации и точной формулы.

«Глава 5» посвящена созданию системы, которая обнаруживает момент возникновения неоднородности не позже, чем через заданный интервал времени. Также анализируется поведение строковой функции обнаружения неоднородности на переходном

интервале.

Глава 1

Singular spectrum analysis

1.1. Алгоритм базового метода SSA

Определим метод SSA как любой метод, состоящий из четырех этапов, описанных ниже. Обозначим входной объект как F_N — упорядоченный набор из N действительных чисел (временной ряд).

1.1.1. Вложение

Процедура вложения переводит исходный временной ряд $F_N = (f_1, \dots, f_N)$ в последовательность многомерных векторов вложения

$$X_i = (f_i, \dots, f_{i+L-1})^T, \quad 1 \leq i \leq K$$

размерности L (длина окна), где $K = N - L + 1$.

Создается траекторная матрица $\mathbf{X} = [X_1, \dots, X_K]$ размерности $L \times K$, имеющую Ганкелеву структуру с одинаковыми значениями на анти-диагоналях

$$\mathbf{X} = \begin{pmatrix} f_1 & f_2 & f_3 & \dots & f_K \\ f_2 & f_3 & f_4 & \dots & f_{K+1} \\ f_3 & f_4 & f_5 & \dots & f_{K+2} \\ \vdots & \vdots & \ddots & \vdots & \\ f_L & f_{L+1} & f_{L+2} & \dots & f_N \end{pmatrix}. \quad (1.1)$$

1.1.2. Сингулярное разложение

Результатом этого шага является сингулярное разложение траекторной матрицы \mathbf{X} .

Пусть $\mathbf{S} = \mathbf{X}\mathbf{X}^T$, $\lambda_1 \geq \dots \geq \lambda_L \geq 0$ — собственные числа матрицы \mathbf{S} , $d = \text{rank } \mathbf{X} = \max\{j : \lambda_j > 0\}$, U_1, \dots, U_d — ортонормированный набор собственных векторов матрицы \mathbf{S} , соответствующий собственным числам, и $V_j = \mathbf{X}^T U_j / \sqrt{\lambda_j}$, $j = 1, \dots, d$ — факторные векторы. Тогда разложение будет иметь вид

$$\mathbf{X} = \sum_{i=1}^d \mathbf{X}_i, \quad \mathbf{X}_i = \sqrt{\lambda_i} U_i V_i^T. \quad (1.2)$$

1.1.3. Группировка

На основе разложения шага 2 процедура группировки делит множество индексов $\{1, \dots, d\}$ на m непересекающихся подмножеств I_1, \dots, I_m .

Пусть $I = \{i_1, \dots, i_p\} \subset \{1, \dots, d\}$ - набор индексов. Тогда результирующая матрица \mathbf{X}_I , соответствующая группе $I = \{i_1, \dots, i_p\}$ определяется как

$$\mathbf{X}_I = \mathbf{X}_{i_1} + \dots + \mathbf{X}_{i_p}.$$

Такие матрицы вычисляются для $I = \{I_1, \dots, I_m\}$, тем самым, разложение (1.2) может быть записано в сгруппированном виде

$$\mathbf{X} = \mathbf{X}_{I_1} + \dots + \mathbf{X}_{I_m}. \quad (1.3)$$

1.1.4. Реконструкция

На последнем шаге базового алгоритма каждая матрица сгруппированного разложения (1.3) преобразуется в новый ряд длины N диагональным усреднением элементов.

Пусть Y — матрица $L \times K$ с элементами y_{ij} , где $1 \leq i \leq L$, $1 \leq j \leq K$. Обозначим $L^* = \min(L, K)$, $K^* = \max(L, K)$. Пусть $y_{ij}^* = y_{ij}$ если $L < K$ и $y_{ij}^* = y_{ji}$ в противном случае. Диагональное усреднение преобразует матрицу Y в ряд g_0, \dots, g_{N-1} по формуле

$$g_k = \begin{cases} \frac{1}{k+1} \sum_{m=1}^{k+1} y_{m,k-m+2}^* & 0 \leq k < L^* - 1, \\ \frac{1}{L^*} \sum_{m=1}^{L^*} y_{m,k-m+2}^* & L^* - 1 \leq k < K^*, \\ \frac{1}{N-k} \sum_{m=k-K^*+2}^{N-K^*+1} y_{m,k-m+2}^* & K^* \leq k < N. \end{cases} \quad (1.4)$$

Применяя диагональное усреднение (1.4) к результирующим матрицам $\mathbf{X}_{\mathbf{I}_k}$, получаем m рядов $\tilde{F}_k = (\tilde{f}_0^{(k)}, \dots, \tilde{f}_{N-1}^{(k)})$. Тогда исходный ряд F_N раскладывается в сумму рядов:

$$F_N = \sum_{k=1}^m \tilde{\mathbb{X}}_k.$$

1.2. Ранги ряда

Рассмотрим ряд $F_N = (f_1, \dots, f_N)$. После процедуры вложения мы получаем подряды длины L : $X_i^{(L)} = X_i = (f_i, \dots, f_{i+L-1})^T$, $1 \leq i \leq K$, $\mathfrak{L}^{(L)} = \mathfrak{L}^{(L)}(F_N) \stackrel{\text{def}}{=} \text{span}(X_1, \dots, X_K)$ — траекторное пространство ряда F_N .

Будем говорить, что ряд имеет L -ранг d , если $\dim \mathfrak{L}^{(L)} = d$ и записывать это как $\text{rank}_L(F_N) = \text{rank}(\mathbf{X}) = \text{rank}(\mathbf{X}\mathbf{X}^T) = \text{rank}(\mathbf{X}^T\mathbf{X}) = d$. Справедливым равенство будет в том случае, если $d \leq \min(L, K)$. Если же равенство $\text{rank}_L(F_N) = d < \frac{N}{2}$ будет достигаться \forall допустимого L , то говорим, что ряд F_N имеет ранг d ($\text{rank}(F_N) = d$) и при существовании такого d ряд F_N — ряд конечного ранга.

Если же мы рассмотрим бесконечный временной ряд $F = (\dots, f_{-1}, f_0, f_1, \dots)$, то такой ряд будем называть рядом конечного ранга тогда и только тогда, когда он управляется LRR размерности d , то есть существуют такие $\alpha_1, \dots, \alpha_d \forall n : f_{n+r} = \sum_{i=1}^r \alpha_i \cdot f_{n+r-i}$, $\alpha_d \neq 0$.

1.2.1. Пример

Рассмотрим ряд $F_N = C \sin(2\pi\omega n + \phi)$, $\phi \in [0, 2\pi)$, $\omega \in [0, \frac{1}{2}]$. Если $\omega < \frac{1}{2}$, то $\forall L \geq 2$ и $N \geq L + 1$ сингулярное разложение траекторной матрицы имеет 2 члена (то есть ранг равен 2). При $\omega = \frac{1}{2}$ ранг равен 1. Отсюда, на практике при SSA разложении в такого вида рядах (где сигнал задан синусом) даже при наличии шума, в качестве количества рассматриваемых собственных векторов указывают $r = 2$, что соответствует компонентам сигнала.

Глава 2

Поиск разладки

Основная идея метода решения задачи обнаружения структурных изменений может быть описана следующим образом. Временной ряд F_N , управляемый LRR, характеризуется тем, что для достаточно больших значений длины окна L (Раздел 1.1.1) (это значение должно быть больше размерности минимального LRR) векторы вложения принадлежат одному и тому же линейному пространству $\mathfrak{L}^{(L)}$ независимо от N (если N достаточно велико). Таким образом, нарушения однородности ряда могут быть описаны в терминах соответствующих векторов вложения: возмущения заставляют эти векторы покидать пространство $\mathfrak{L}^{(L)}$. Соответствующие расхождения определяются в терминах расстояний между векторами вложения длины L и пространством $\mathfrak{L}^{(L)}$, которые могут быть определены для различных частей ряда (например, до и после возмущения).

2.1. Матрица неоднородности и функции неоднородности

2.1.1. Матрица неоднородности

Рассмотрим два временных ряда $F^{(1)} = F_{N_1}^{(1)}$ и $F^{(2)} = F_{N_2}^{(2)}$ и зададим число $L : 2 \leq L \leq \min(N_1 - 1, N_2)$. Обозначим $\mathfrak{L}^{(L,1)}$ линейное пространство, натянутое на векторы вложения длины L ряда $F^{(1)}$.

Пусть $U_l^{(1)} (l = 1, \dots, L)$ — левые сингулярные векторы траекторной матрицы ряда $F^{(1)}$. Для $l > d \stackrel{\text{def}}{=} \dim \mathfrak{L}^{(L,1)}$ в качестве собственных векторов $U_l^{(1)}$ мы берем векторы из любого ортонормированного базиса пространства, ортогонального $\mathfrak{L}^{(L,1)}$.

Пусть $I = \{i_1, \dots, i_r\}$ — подмножество $\{1, \dots, L\}$ и $\mathfrak{L}_r^{(1)} \stackrel{\text{def}}{=} \text{span}(U_l^{(1)}, l \in I)$. Обозначим через $X_1^{(2)}, \dots, X_{K_2}^{(2)} (K_2 = N_2 - L + 1)$ векторы вложения длины L временного ряда $F^{(2)}$.

Введем меру, называемую **индексом неоднородности**, которая характеризует несоответствие между рядом $F^{(2)}$ и структурой ряда $F^{(1)}$ (описываемой подпространством $\mathfrak{L}_r^{(1)}$):

$$\begin{aligned}
g(F^{(1)}; F^{(2)}) &= \frac{\sum_{l=1}^{K_2} \text{dist}^2(X_l^{(2)}, \mathfrak{L}_r^{(1)})}{\sum_{l=1}^{K_2} \|X_l^{(2)}\|^2} = \\
&= \frac{\sum_{l=1}^{K_2} (\|X_l^{(2)}\|^2 - \sum_{i=1}^r \langle X_l^{(2)}, U_i^{(1)} \rangle^2)}{\sum_{l=1}^{K_2} \|X_l^{(2)}\|^2} = 1 - \frac{\sum_{l=1}^{K_2} \sum_{i=1}^r \langle X_l^{(2)}, U_i^{(1)} \rangle^2}{\sum_{l=1}^{K_2} \|X_l^{(2)}\|^2}.
\end{aligned} \tag{2.1}$$

Значения g принадлежат интервалу $[0, 1]$.

Введем обозначения:

1. Исходный временной ряд $F_N : F_N = (f_1, \dots, f_N), N > 2$;
2. Подряды (интервалы) $F_{i,j}$ временного ряда $F_N : F_{i,j} = (f_i, \dots, f_j), 1 \leq i < j \leq N$;
3. Длина окна $L : 1 < L < N$;
4. Длина B базовых подрядов ряда $F_N : B > L$;
5. Длина T тестовых подрядов ряда $F_N : T \geq L$;
6. Предполагаем набор $I = \{j_1, \dots, j_r\}$ различных натуральных чисел: $j < \min(L, B - L + 1) \forall j \in I$.
7. Базовые пространства ($i = 1, \dots, N - B + 1$) натянуты на собственные векторы с индексами из I , полученные сингулярным разложением траекторных матриц $\mathbf{X}^{(i,B)}$ ряда $F_{i,i+B-1}$ с длиной окна L . Соответствующий набор собственных троек называется **базовым набором собственных троек**.

Учитывая эти обозначения, матрица $\mathbf{G} = \mathbf{G}_{B,T}$, состоящая из элементов g_{ij} :

$$g_{ij} = g(F_{i,i+B-1}; F_{j,j+T-1}), \tag{2.2}$$

$$1 \leq i \leq N - B + 1, \quad 1 \leq j \leq N - T + 1,$$

есть **матрица неоднородности** ($H - matrix$) временного ряда F_N . Отсюда следует, что пространство $\mathfrak{L}_{I,B}^{(L,i)}$ соответствует пространству $\mathfrak{L}_r^{(1)}$. Ряд $F_{i,i+B-1}$ называют **базовым подрядом** (или базовым интервалом), а $F_{j,j+T-1}$ — **тестовым подрядом** (интервалом). По определению, величина g_{ij} является нормированной суммой расстояний между L -сдвинутыми векторами тестового подряда и линейным пространством $\mathfrak{L}_{I,B}^{(L,i)}$.

2.1.2. Функции обнаружения

На основе матрицы неоднородности \mathbf{G} введем различные функции обнаружения.

1. Строковая функция обнаружения

Строковой функцией обнаружения является ряд $D_{T,N}^{(r)}$ с членами, определяемыми

$$d_{n-1}^{(r)} \stackrel{\text{def}}{=} g(F_{1,B}; F_{n-T+1,n}), \quad T \leq n \leq N, \quad (2.3)$$

что соответствует обнаружению изменений по отношению к начальной части ряда (или, точнее, к его первым B членам, которые представлены пространством $\mathfrak{L}_{I,B}^{(L,1)}$).

2. Столбцовая функция обнаружения

Столбцовой функцией обнаружения является ряд $D_{B,N}^{(c)}$ с членами, определяемыми

$$d_{n-1}^{(c)} \stackrel{\text{def}}{=} g(F_{n-B+1,n}; F_{1,T}), \quad B \leq n \leq N. \quad (2.4)$$

3. Функция диагонального обнаружения

Функцией диагонального обнаружения является ряд $D_{T+B,N}^{(d)}$ с членами, определяемыми

$$d_{n-1}^{(d)} \stackrel{\text{def}}{=} g(F_{n-T-B+1,n-T+1}; F_{n-T+1,n}), \quad T+B \leq n \leq N. \quad (2.5)$$

Поскольку промежуток между базовым и тестовым интервалами отсутствует, данная функция обнаружения может использоваться для обнаружения резких структурных изменений на фоне медленных.

4. Функция симметричного обнаружения

Пусть $T = B$. Функцией симметричного обнаружения является ряд $D_{B,N}^{(s)}$ с членами, определяемыми

$$d_{n-1}^{(s)} \stackrel{\text{def}}{=} g(F_{n-B+1,n}; F_{n-B+1,n}), \quad B \leq n \leq N. \quad (2.6)$$

Эта функция обнаружения измеряет качество приближения базового ряда выбранными собственными тройками.

2.2. Однородность и неоднородность

Пусть F_N — однородный временной ряд, управляемый минимальным LRR размерности d . Выберем L и r такие, что $L \geq d$, $d \leq r \leq \min(L, N - L + 1)$.

Если выбрать $I = \{1, \dots, r\}$, то матрица неоднородности (2.1) будет нулевой, поскольку $B \geq L$, $\forall i$ $\mathfrak{L}^{(L)}(F_{i,i+B-1}) = \mathfrak{L}^{(L)}(F_N)$, следовательно, все векторы вложения длины L ряда $F_{j,j+T-1}$ лежат в пространстве $\mathfrak{L}^{(L)}(F_{i,i+B-1}) \forall i, j$. Это означает, что любой однородный ряд F_N порождает нулевую матрицу неоднородности, а наличие ненулевых элементов g_{ij} в этой матрице свидетельствует о нарушении однородности.

Рассмотрим несколько типов нарушений.

2.2.1. Типы неоднородности

Как указывалось ранее, временной ряд F_N задан LRR до определенного времени Q . Затем происходит мгновенное возмущение, хотя через короткое время ряд снова становится однородным и подчиняется LRR, которое может отличаться от исходного.

Если начальное LRR восстанавливается, то мы имеем **временное** нарушение структуры временного ряда. В противном случае нарушение является **постоянным**.

Момент времени Q будем называть **моментом возмущения** или **точкой изменения**. Положим $d = \text{rank}_L(F_{1,Q-1})$.

Предположим, что через некоторое время $S \geq 0$ после возмущения, временной ряд стал опять однородным (ряд $F_{Q+S,N}$). Обозначим $d_1 = \text{rank}_L(F_{Q+S,N})$. Временной интервал $[Q, Q+S]$ называется **переходным интервалом** (поведение ряда на котором нас не интересует).

Пусть $L \geq \max(d, d_1)$. Дополнительно введем ограничения $L \leq Q - 1$ и $L \leq N - Q - S + 1$. Если векторы вложения длины L ряда F_N покрывают исходное подпространство $\mathfrak{L}^{(L)}(F_{1,Q-1})$ после того, как они покинули переходный интервал (то есть $\mathfrak{L}^{(L)}(F_{1,Q-1}) = \mathfrak{L}^{(L)}(F_{Q+S,N})$), тогда обе однородные части временного ряда соответствуют одному минимальному LRR — случай временной неоднородности. Отсюда вытекает случай постоянной неоднородности.

Опишем вид матрицы неоднородности. Пусть длины базового и тестового интервалов удовлетворяют условию $\max(B, T) < Q$. Предположим, что $I = \{1, \dots, r\}$ и $r = d \leq \min(L, B - L + 1)$. Тогда все элементы g_{ij} матрицы $\mathbf{G}_{B,T}$ равны нулю для $i + B \leq Q$ и $j + T \leq Q$. Это связано с тем, что для этих индексов, и базовый, и тестовый

подряды исходного ряда F_N также являются подрядами однородного ряда $F_{1,Q-1}$. Значения остальных элементов матрицы неоднородности зависят от типа неоднородности и значений параметров.

Схематично, общая форма матрицы неоднородности изображена на Рис. 2.1

Регион \mathcal{A} соответствует элементам g_{ij} где ряды $F_{i,i+B-1}$ и $F_{j,j+T-1}$ являются подрядами однородного ряда $F_{1,Q-1}$. Отсюда следует, что данный регион состоит из нулевых элементов.

В регионе \mathcal{D} те же самые ряды являются подрядами ряда $F_{Q+S,N}$. Поэтому, если размерность d_1 ряда $F_{Q+S,N}$ не больше размерности d ряда $F_{1,Q-1}$, то данный регион матрицы также состоит из нулевых элементов.

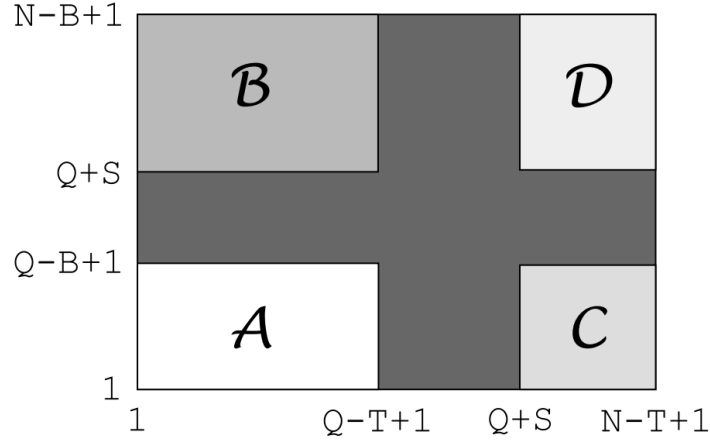


Рис. 2.1. Общая форма матрицы неоднородности.

Крест неоднородности — регион в матрице неоднородности с индексами элементов

$$Q - B + 1 \leq i \leq Q + S - 1, \quad Q - T + 1 \leq j \leq Q + S - 1.$$

Эти элементы соответствуют тому, что либо тестовый интервал, либо базовый имеет пересечение с переходным.

Глава 3

Численное сравнение функций разладки

Так как у нас есть четыре функции обнаружения неоднородности:

1. Строковая: $d_{n-1}^{(r)} \stackrel{\text{def}}{=} h_{b-T}^{(r,1)} = g(F_{1,B}; F_{n-T+1,n}), \quad T \leq n \leq N.$
2. Столбцовая: $d_{n-1}^{(c)} \stackrel{\text{def}}{=} h_{n-B}^{(1,c)} = g(F_{n-B+1,n}; F_{1,T}), \quad B \leq n \leq N.$
3. Диагональная: $d_{n-1}^{(d)} \stackrel{\text{def}}{=} h_{n-T-B}^{(d,B)} = g(F_{n-T-B+1,n-T+1}; F_{n-T+1,n}), T + B \leq n \leq N.$
4. Симметричная: $d_{n-1}^{(s)} \stackrel{\text{def}}{=} h_{n-B}^{(s)} = g(F_{n-B+1,n}; F_{n-B+1,n}), \quad B \leq n \leq N.$

можем экспериментальным путем попытаться определить, какая из них лучше обнаруживает разладку в ряде.

3.1. Постановка задачи

Рассмотрим ряд

$$f_n = \begin{cases} C_1 \sin(2\pi\omega_1 n + \phi_1), & n < Q, \\ C_2 \sin(2\pi\omega_2 n + \phi_2), & n \geq Q, \end{cases}$$

чьи параметры буду задаваться типом разладки и соответствующим изменением параметров. Рассмотрим два типа неоднородности:

1. Временную, заданную

а. Фазовым сдвигом: $\phi_1 \neq \phi_2$;

б. Выбросом:

$$f_n = \begin{cases} C_1 \sin(2\pi\omega_1 n + \phi_1) & n \neq Q, \\ 10 \cdot C_1 & n = Q. \end{cases}$$

в. Изменением амплитуды: $C_1 \neq C_2$.

2. Постоянную, заданную

а. Изменением частоты: $\omega_1 \neq \omega_2$.

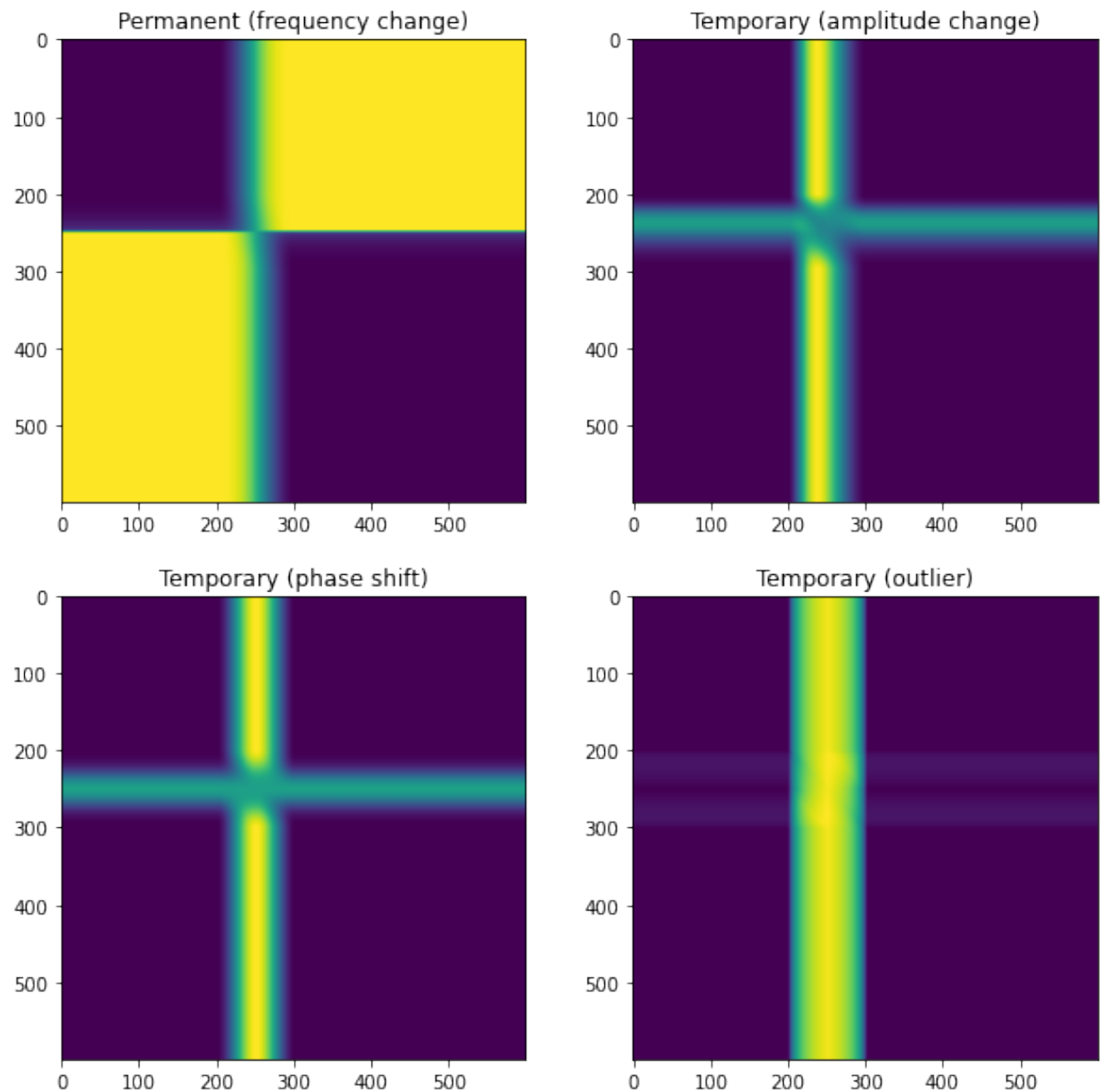


Рис. 3.1. Матрицы неоднородности для рядов без шума.

Ряды, заданные параметрами выше, порождают матрицы неоднородности, изображенные на Рис. 3.1

В качестве оценок функций неоднородности будем учитывать скорость возрастания значений и момент преодоления $n_{overcome}$ заданного порога δ .

3.2. Организация численного эксперимента

Для реализации тестов был выбран язык *Python3*. Взаимодействие с пакетом RSSA [2] [3] осуществлялось через библиотеку *rpy2* (стадии вложения и сингулярного

разложения). Подсчет индексов неоднородности велась функционалом языка *Python3* и библиотеку *numpy*. Работа с графикой велась через библиотеку *matplotlib*.

Параметры ряда были заданы следующим образом:

$N = 700$, $\omega_1 = \frac{1}{10}$, $\omega_2 = \frac{1}{5}$, $C_1 = 1$, $C_2 = 2$, $\phi_1 = 0$, $\phi_2 = \frac{\pi}{2}$, $Q = 301$, $B = T = 100$, $L = 50$, $r = d = \text{rank}(f_n) = 2$.

В тестах предполагаем, что момент разладки Q известен и для оценки скорости возрастания будем выводить значения функций в точках $[Q, Q + 10, Q + 20, Q + 30]$.

Порог, относительно которого будем определять, какая из функций неоднородности раньше обнаруживает разладку, зададим в соответствии с промоделированными значениями, описанными далее.

3.2.1. Ряды без шума

Временные ряды и их функции неоднородности изображены на Рис. 3.2 и Рис. 3.3 соответственно.

Для определения момента преодоления значения порога δ необходимо этот порог задать. Для этого к рассмотренным рядам добавим шум $\epsilon \sim N(0, \sigma^2)$, где $\sigma = 0.5$. Для ряда с временной разладкой, заданной изменением амплитуды ($C_1 \neq C_2$), зададим дисперсию шума до разладки как $\frac{\sigma^2}{2}$, чтобы шум ϵ был пропорционален амплитуде ряда, так как в противном случае, сила шума не позволит корректно определить базис пространства \mathfrak{L}_τ .

Промоделируем реализации шума $n_{mod} = 200$ раз и посчитаем такие характеристики ряда на промежутке $[0, \dots, Q - 1]$, как средний максимум и 95-й процентиль. Эти два значения возьмем в качестве параметра δ .

Результаты моделирования представлены в таблице 3.1.

Таблица 3.1. Значения моделирования.

Permanent ($\omega_1 \neq \omega_2$)	$d_{n-1}^{(r)}$	$d_{n-1}^{(c)}$	$d_{n-1}^{(s)}$	$d_{n-1}^{(d)}$
meanMax	0.133376	0.110945	0.130618	0.126147
95 percentile	0.130877	0.110506	0.127974	0.124446
Temporary ($C_1 \neq C_2$)	$d_{n-1}^{(r)}$	$d_{n-1}^{(c)}$	$d_{n-1}^{(s)}$	$d_{n-1}^{(d)}$
meanMax	0.036142	0.030146	0.035325	0.034518
95 percentile	0.035530	0.030024	0.034683	0.034116
Temporary ($\phi_1 \neq \phi_2$)	$d_{n-1}^{(r)}$	$d_{n-1}^{(c)}$	$d_{n-1}^{(s)}$	$d_{n-1}^{(d)}$
meanMax	0.132096	0.114525	0.129501	0.124849
95 percentile	0.129943	0.114098	0.127145	0.123502
Temporary (Outlier)	$d_{n-1}^{(r)}$	$d_{n-1}^{(c)}$	$d_{n-1}^{(s)}$	$d_{n-1}^{(d)}$
meanMax	0.132471	0.1108	0.129573	0.127397
95 percentile	0.130347	0.110320	0.127320	0.126176

Полученные результаты тестирования функций разладки приведены в таблицах 3.2, 3.4, 3.6, 3.8, а значения функций неоднородности в точках $[Q, Q + 10, Q + 20, Q + 30]$

приведены в таблицах 3.3, 3.5, 3.7, 3.9.

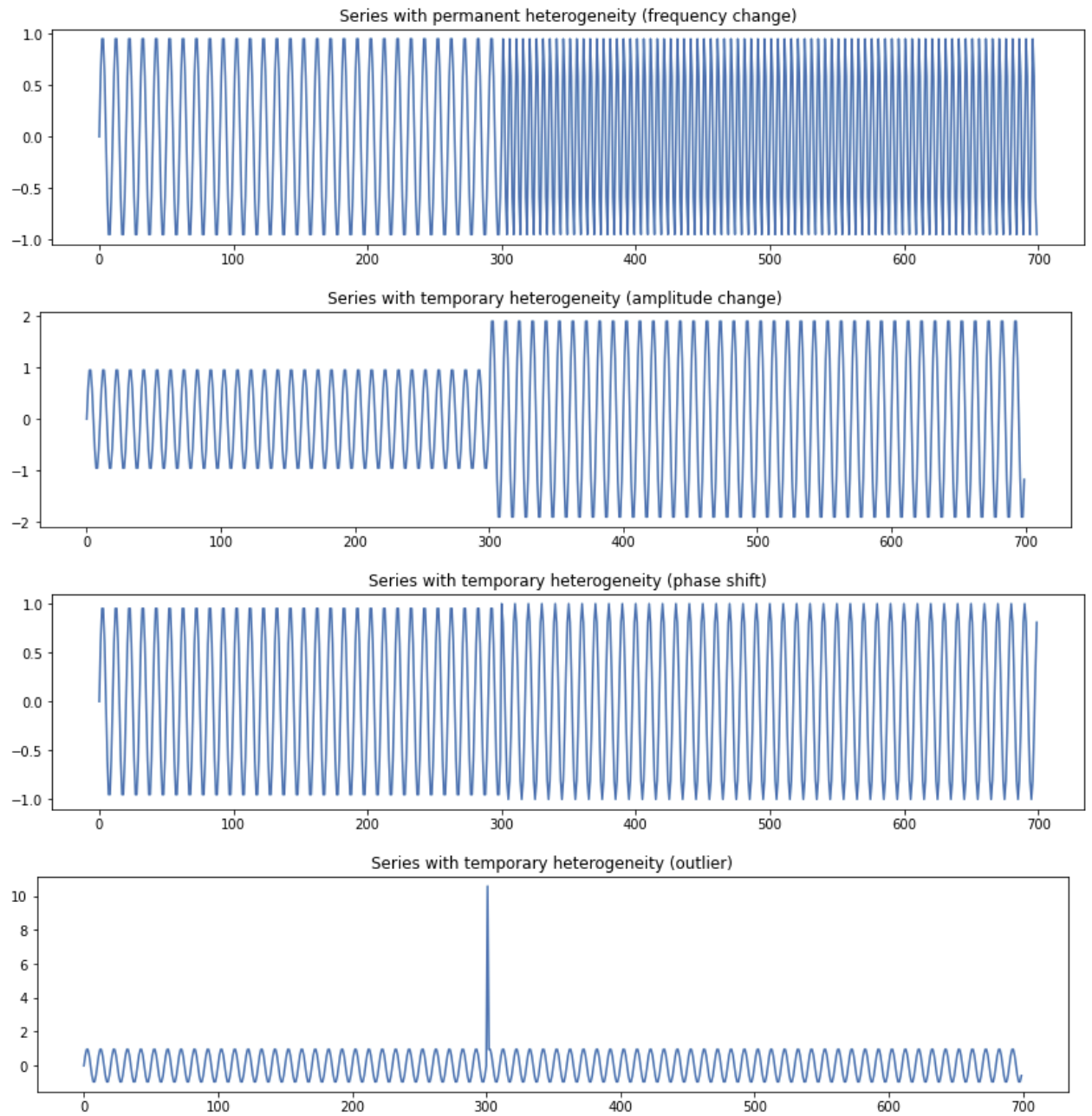


Рис. 3.2. Временные ряды без шума.

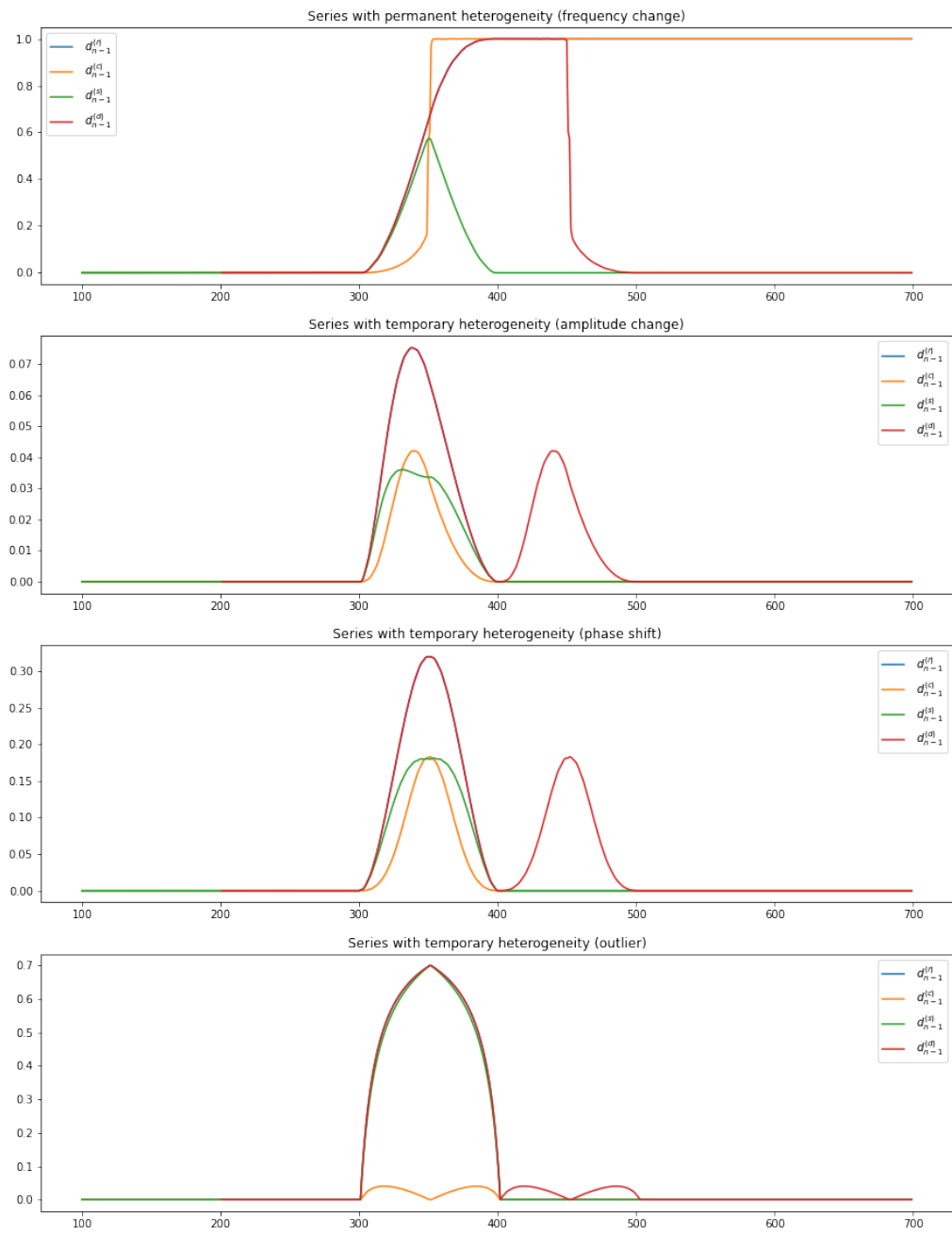


Рис. 3.3. Функции неоднородности рядов без шума.

Таблица 3.2. Характеристики функций неоднородности для постоянной разладки ($\omega_1 \neq \omega_2$).

Row			Col		
Sym	Statistic	Mean Max	Mean 95 proc	Mean Max	Mean 95 proc
	$n_{overcome}$	320	320	345	345
	$D_{n_{overcome}}$	0.134217	0.134217	0.115557	0.115557
	Diag				
Sym	Statistic	Mean Max	Mean 95 proc	Mean Max	Mean 95 proc
	$n_{overcome}$	321	321	320	320
	$D_{n_{overcome}}$	0.135379	0.135379	0.134217	0.134217
	Diag				

Таблица 3.3. Значения функций неоднородности в моменты $[Q, Q + 10, Q + 20, Q + 30]$ при постоянной разладки ($\omega_1 \neq \omega_2$).

	Row	Col	Sym	Diag
D_Q	0	0	0	0
D_{Q+10}	0.042795	0.002815	0.040179	0.042795
D_{Q+20}	0.146766	0.013995	0.135379	0.146766
D_{Q+30}	0.296227	0.038518	0.270609	0.296227

Таблица 3.4. Характеристики функций неоднородности для временной разладки ($C_1 \neq C_2$).

Row			Col		
Sym	Statistic	Mean Max	Mean 95 proc	Mean Max	Mean 95 proc
	$n_{overcome}$	317	317	328	328
	$D_{n_{overcome}}$	0.037079	0.037079	0.031576	0.031576
	Diag				
Sym	Statistic	Mean Max	Mean 95 proc	Mean Max	Mean 95 proc
	$n_{overcome}$	327	326	317	317
	$D_{n_{overcome}}$	0.035488	0.035066	0.037079	0.037079
	Diag				

Таблица 3.5. Значения функций неоднородности в моменты $[Q, Q + 10, Q + 20, Q + 30]$ при временной разладки ($C_1 \neq C_2$).

	Row	Col	Sym	Diag
D_Q	0	0	0	0
D_{Q+10}	0.018616	0.003571	0.015156	0.018616
D_{Q+20}	0.04911	0.018519	0.031535	0.04911
D_{Q+30}	0.070292	0.036105	0.036025	0.070292

Таблица 3.6. Характеристики функций неоднородности для временной разладки ($\phi_1 \neq \phi_2$).

Row	Col				
Statistic	Mean	Max	Mean	95 proc	
$n_{overcome}$	323		322		
$D_{n_{overcome}}$	0.140235		0.130778		
				Mean	Max
				Mean	95 proc
				336	336
				0.120716	0.120716
Sym	Diag				
Statistic	Mean	Max	Mean	95 proc	
$n_{overcome}$	327		327		
$D_{n_{overcome}}$	0.131381		0.131381		
				Mean	Max
				Mean	95 proc
				322	322
				0.130778	0.130778

Таблица 3.7. Значения функций неоднородности в моменты $[Q, Q + 10, Q + 20, Q + 30]$ при временной разладки ($\phi_1 \neq \phi_2$).

	Row	Col	Sym	Diag
D_Q	0.000752	0.000029	0.000723	0.000752
D_{Q+10}	0.03919	0.00507	0.034446	0.03919
D_{Q+20}	0.12146	0.030249	0.096102	0.12146
D_{Q+30}	0.21607	0.085474	0.150779	0.21607

В таблице 3.8 можно заметить пустые элементы, которые соответствуют ситуациям, когда рассматриваемая функция неоднородности не смогла преодолеть соответствующее промоделированное значение из таблицы 3.1.

По таблицам 3.2, 3.4, 3.6, 3.8 можем сделать вывод, что строковая $d_{n-1}^{(r)}$ и диагональная $d_{n-1}^{(d)}$ функции неоднородности более устойчивые к шуму ϵ , возрастают сильнее,

Таблица 3.8. Характеристики функций неоднородности для временной разладки (выброс).

Row				Col	
	Statistic	Mean	Max	Mean	95 proc
	$n_{overcome}$	304		304	
	$D_{n_{overcome}}$	0.178852		0.178852	
Sym				Diag	
	Statistic	Mean	Max	Mean	95 proc
	$n_{overcome}$	304		303	
	$D_{n_{overcome}}$	0.167006		0.128127	

Таблица 3.9. Значения функций неоднородности в моменты $[Q, Q + 10, Q + 20, Q + 30]$ для временной разладки (выброс).

	Row	Col	Sym	Diag
D_Q	0	0	0	0
D_{Q+10}	0.401244	0.00357	0.380619	0.401244
D_{Q+20}	0.546991	0.01851	0.528819	0.546991
D_{Q+30}	0.622343	0.036111	0.610083	0.622343

однако $d_{n-1}^{(d)}$ раньше преодолевает промоделированное значение, что значит и более раннее обнаружение разладки.

3.2.2. Ряды с шумом

Возьмем те же параметры ряда, что и в предыдущем примере, однако к ряду добавим шум $\epsilon \sim N(0, \sigma^2)$, $\sigma = 0.5$.

Для ряда с временной разладкой, заданной изменением амплитуды ($C_1 \neq C_2$), зададим дисперсию шума до разладки как $\frac{\sigma^2}{2}$, чтобы шум ϵ был пропорционален амплитуде ряда, так как в противном случае, сила шума не позволит корректно определить базис пространства \mathfrak{L}_τ .

Для оценки функций неоднородности будем считать порог δ как максимум и 95-й процентиль до момента Q на каждой итерации.

Временные ряды и их функции неоднородности изображены на Рис. 3.4 и Рис. 3.5 соответственно.

Полученные результаты тестирования функций разладки на рядах с шумом приведены в таблицах 3.10, 3.12, 3.14, 3.16, а значения функций неоднородности в точках $[Q, Q + 10, Q + 20, Q + 30]$ приведены в таблицах 3.11, 3.13, 3.14, 3.17. В данных таблицах приведены средние значения по всем итерациям.

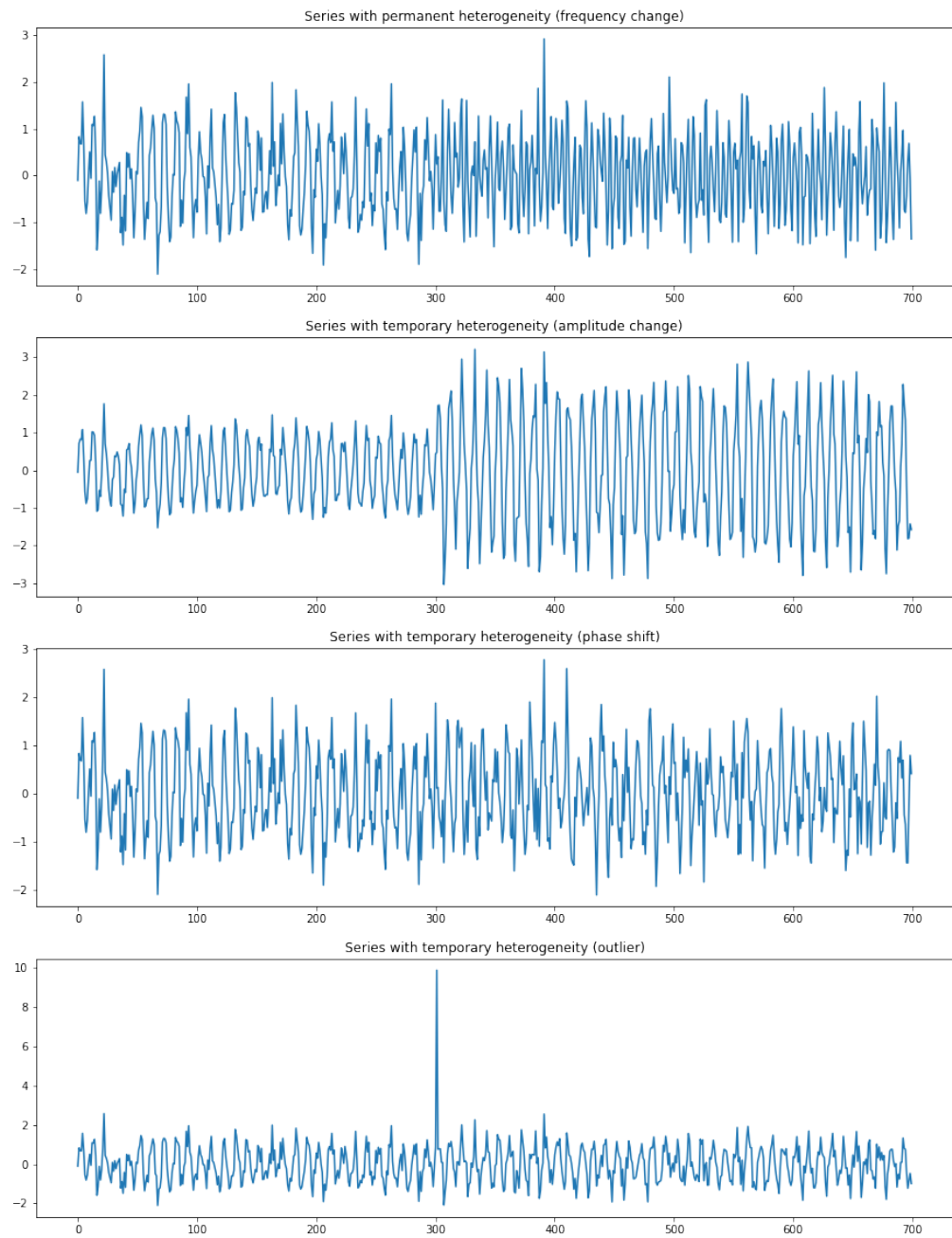


Рис. 3.4. Временные ряды с шумом.

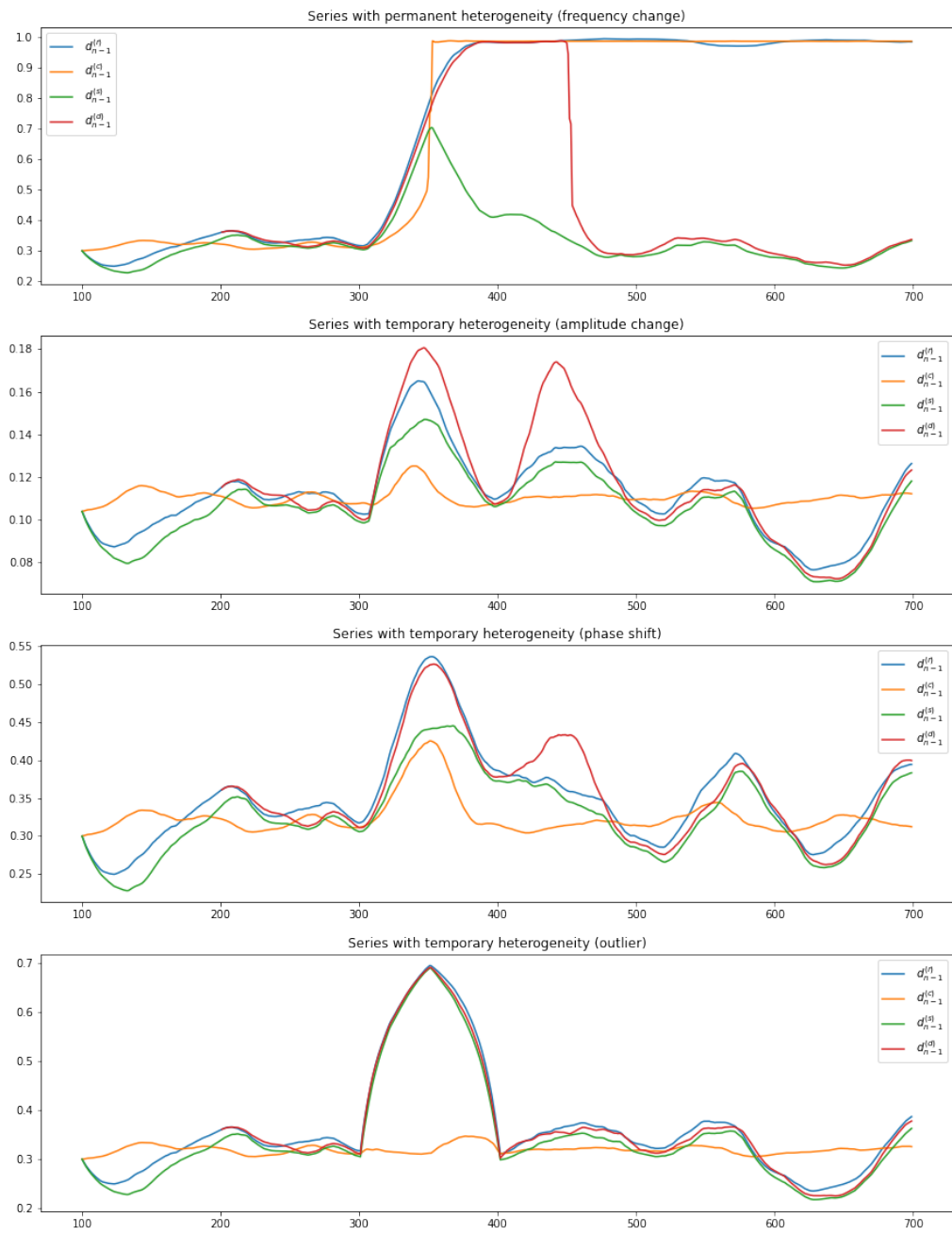


Рис. 3.5. Функции неоднородности рядов с шумом.

Таблица 3.10. Характеристики функций неоднородности для постоянной разладки с шумом ($\omega_1 \neq \omega_2$).

Row			Col		
Statistic	Mean Max	Mean 95 proc	Mean Max	Mean 95 proc	
$\#n_{overcome}$	200	200	200	200	
$n_{overcome}$	309.12	308.65	313,46	313,29	
Confidence	[309.09, 309.15]	[308.61, 308.68]	[313.35, 313.57]	[313.17, 313.39]	
$D_{n_{overcome}}$	0.137	0.135	0.1167	0.1165	
Sym			Diag		
Statistic	Mean Max	Mean 95 proc	Mean Max	Mean 95 proc	
$\#n_{overcome}$	200	200	200	200	
$n_{overcome}$	309.46	308.94	307.78	307.45	
Confidence	[309.42, 309.49]	[308.91, 308.98]	[307.74, 307.81]	[307.41, 307.48]	
$D_{n_{overcome}}$	0.134	0.131	0.13	0,128	

Таблица 3.11. Значения функций неоднородности в моменты $[Q, Q + 10, Q + 20, Q + 30]$ при разладке, вызванной изменением частот с шумом)

	Row	Col	Sym	Diag
D_Q	0,1085	0,1075	0,1055	0,1084
D_{Q+10}	0,1469	0,1102	0,1416	0,1466
D_{Q+20}	0,2405	0,1206	0,2273	0,2400
D_{Q+30}	0,3734	0,1433	0,3476	0,3727

Таблица 3.12. Характеристики функций неоднородности для временной разладки с шумом ($\tau_1 \neq \tau_2$).

Row			Col		
Statistic	Mean	Max	Mean	Max	95 proc
$\#n_{overcome}$	200		200		200
$n_{overcome}$	306,885		306,525		306,465
Confidence	[306.86, 306.91]		[306.50, 306.55]		[306.41, 306.52]
$D_{n_{overcome}}$	0,0373		0,0366		0,032
Sym			Diag		
Statistic	Mean	Max	Mean	Max	95 proc
$\#n_{overcome}$	200		200		200
$n_{overcome}$	307,53		307,07		305,745
Confidence	[307.50, 307.56]		[307.04, 307.10]		[305.72, 305.77]
$D_{n_{overcome}}$	0,0363		0,0357		0,0354

Таблица 3.13. Значения функций неоднородности в моменты $[Q, Q + 10, Q + 20, Q + 30]$ для временной разладки с шумом ($\tau_1 \neq \tau_2$).

	Row	Col	Sym	Diag
D_Q	0,0296	0,0297	0,0289	0,0296
D_{Q+10}	0,0458	0,0331	0,0418	0,0458
D_{Q+20}	0,0713	0,0475	0,0537	0,0713
D_{Q+30}	0,0876	0,0645	0,0536	0,0873

Таблица 3.14. Характеристики функций неоднородности для постоянной разладки ($\phi_1 \neq \phi_2$).

Row			Col		
	Statistic	Mean Max	Mean 95 proc	Mean Max	Mean 95 proc
	$\#n_{overcome}$	200	200	200	200
	$n_{overcome}$	309,37	308,925	311,73	311,485
	Confidence	[309.33, 309.41]	[308.89, 308.96]	[311.64, 311.82]	[311.40, 311.57]
	$D_{n_{overcome}}$	0,1351	0,1331	0,1193	0,1190
Sym			Diag		
	Statistic	Mean Max	Mean 95 proc	Mean Max	Mean 95 proc
	$\#n_{overcome}$	200	200	200	200
	$n_{overcome}$	310,105	309,58	307,84	307,535
	Confidence	[310.06, 310.15]	[309.54, 309.62]	[307.80, 307.88]	[307.50, 307.57]
	$D_{n_{overcome}}$	0,1320	0,1298	0,1282	0,1269

Таблица 3.15. Значения функций неоднородности в моменты $[Q, Q + 10, Q + 20, Q + 30]$ для временной разладки с шумом ($\phi_1 \neq \phi_2$).

	Row	Col	Sym	Diag
D_Q	0,1078	0,1077	0,1050	0,1077
D_{Q+10}	0,1421	0,1123	0,1351	0,1421
D_{Q+20}	0,2158	0,1347	0,1907	0,2158
D_{Q+30}	0,3008	0,1836	0,2407	0,3005

Таблица 3.16. Характеристики функций неоднородности для временной разладки с шумом (выброс).

Row	Statistic	Mean Max	Mean 95 proc	Col	Mean Max	Mean 95 proc
	$\#n_{overcome}$	200	200		200	200
	$n_{overcome}$	301,935	301,92		303,401	303,394
	Confidence	[301.93, 301.94]	[301.916, 301.93]		[303.34, 303.46]	[303.33, 303.45]
	$Dn_{overcome}$	0,1579	0,1571		0,1185	0,1182
Sym	Statistic	Mean Max	Mean 95 proc	Diag	Mean Max	Mean 95 proc
	$\#n_{overcome}$	200	200		200	200
	$n_{overcome}$	301,98	301,95		301,88	301,87
	Confidence	[301.976, 301.984]	[301.946, 301.953]		[301.876, 301.883]	[301.866, 301.873]
	$Dn_{overcome}$	0,1531	0,1517		0,1550	0,1545

Таблица 3.17. Значения функций неоднородности в моменты $[Q, Q + 10, Q + 20, Q + 30]$ для временной разладки с шумом ($\phi_1 \neq \phi_2$).

	Row	Col	Sym	Diag
D_Q	0,1072	0,1100	0,1042	0,1071
D_{Q+10}	0,4369	0,1423	0,5462	0,4366
D_{Q+20}	0,5652	0,1459	0,1907	0,5649
D_{Q+30}	0,6336	0,1387	0,6204	0,6332

По результатам видим что в среднем, диагональная функция неоднородности $d_{n-1}^{(d)}$ раньше обнаруживает разладку среди остальных трех, при этом, уступая в скорости возрастания строковой в примерах с постоянной ($\omega_1 \neq \omega_2$) и временной (выброс) разладками.

3.3. Выводы

Таблица 3.18. Средние значения характеристик функций обнаружения по всем типам неоднородности.

Row	meanMax	95 percentile
$\#n_{overcome}$	200	200
$n_{overcome}$	306.83	306.5
$ [D_Q^{(r)}, D_{Q+10}^{(r)}, D_{Q+20}^{(r)}, D_{Q+30}^{(r)}] _{L_2}$	0.49439	
Col	meanMax	95 percentile
$\#n_{overcome}$	199.25	199.5
$n_{overcome}$	308.79	308.66
$ [D_Q^{(c)}, D_{Q+10}^{(c)}, D_{Q+20}^{(c)}, D_{Q+30}^{(c)}] _{L_2}$	0.2199	
Sym	meanMax	95 percentile
$\#n_{overcome}$	200	200
$n_{overcome}$	307.27	306.89
$ [D_Q^{(s)}, D_{Q+10}^{(s)}, D_{Q+20}^{(s)}, D_{Q+30}^{(s)}] _{L_2}$	0.45628	
Diag	meanMax	95 percentile
$\#n_{overcome}$	200	200
$n_{overcome}$	305.86	305.65
$ [D_Q^{(d)}, D_{Q+10}^{(d)}, D_{Q+20}^{(d)}, D_{Q+30}^{(d)}] _{L_2}$	0.49389	

Явными фаворитами (таблица 3.18) являются строковая $d_{n-1}^{(r)}$ и диагональная $d_{n-1}^{(d)}$ функции неоднородности. Они обе показывают превосходство над столбцовой $d_{n-1}^{(c)}$ и симметричной $d_{n-1}^{(s)}$ в устойчивости к шуму ϵ , моментом обнаружения разладки $n_{overcome}$ и скорости возрастания значений $[D_Q, D_{Q+10}, D_{Q+20}, D_{Q+30}]$ после момента нарушения однородности Q .

Глава 4

Аналитическая оценка индекса неоднородности при изменении частоты гармоника

Исходя из заключений прошлой главы о качестве функций обнаружения разладки, далее будем рассматривать элементы строковой функции неоднородности $d_{n-1}^{(r)}$ в качестве индекса неоднородности g .

Рассмотрим ряд $F_N = (f_1, \dots, f_N)$, причем

$$f_n = \begin{cases} C_1 \sin(2\pi\omega_1 n + \phi_1), & n \in [1, Q-1] \\ C_2 \sin(2\pi\omega_2 n + \phi_2), & n \in [Q, N] \end{cases}$$

Обозначим

$$F^{(1)} = f_n^{(1)}|_{n=1}^B,$$

$$F^{(2)} = f_n^{(2)}|_{n=1}^T,$$

$$X_l^{(2)} = (f_l^{(2)}, \dots, f_{l+L-1}^{(2)})^T, \quad 1 \leq l < K_2.$$

В обозначениях выше, $F^{(1)}$ — некий подряд ряда F_N длины B , целиком лежащий в промежутке от начала ряда до точки разладки Q , а $F^{(2)}$ — некий подряд ряда F_N длины T , целиком лежащий в промежутке от точки разладки Q до конца ряда F_N .

В соответствии с формулой (2.1), индекс неоднородности задается как:

$$g(F^{(1)}; F^{(2)}) = 1 - \frac{\sum_{l=0}^{K_2-1} \sum_{i=0}^{r-1} \langle X_l^{(2)}, U_i^{(1)} \rangle^2}{\sum_{l=0}^{K_2-1} \|X_l^{(2)}\|^2} = 1 - z(F^{(1)}; F^{(2)}),$$

где $z(F^{(1)}; F^{(2)})$ — индекс однородности.

В данной главе будем предполагать $\omega_1 \neq \omega_2$; $C_1 = C_2$. Для простоты зададим амплитуды $C_1 = C_2 = 1$.

Попробуем аналитически упростить данную формулу, чтобы явно увидеть, как разности частот ряда до и после разладки влияют на значения g .

4.1. Вычисление индекса однородности

4.1.1. Знаменатель

Начнем со знаменателя $\sum_{l=1}^{K_2} \|X_l^{(2)}\|^2$, а точнее, с квадрата нормы $\|X_l^{(2)}\|^2$. Оценим его:

$$\|X_l^{(2)}\|^2 = \sum_{i=1}^L (X_l^{(2)})_i^2 \approx \int_0^L \sin^2(2\pi\omega_2 y + \psi_l) dy = \frac{L}{2} - \frac{\sin(4\pi L\omega_2 + \psi_l) - \sin(2\psi_l)}{8\pi\omega_2} \approx \frac{L}{2},$$

где ψ_l формируется из ϕ_2 и сдвига, порождаемого номером вектора вложения.

$$\text{Отсюда } \sum_{l=1}^{K_2} \|X_l^{(2)}\|^2 \approx K_2 \cdot \frac{L}{2}.$$

4.1.2. Числитель

$$\begin{aligned} \sum_{l=1}^{K_2} \sum_{i=1}^r \langle X_l^{(2)}, U_i^{(1)} \rangle^2 &= \sum_{l=1}^{K_2} \left(\langle X_l^{(2)}, U_1^{(1)} \rangle^2 + \langle X_l^{(2)}, U_2^{(1)} \rangle^2 \right) = \\ &= \sum_{l=1}^{K_2} \left[\left(\sum_{j=1}^L (X_l^{(2)})_j \cdot (U_1^{(1)})_j \right)^2 + \left(\sum_{j=1}^L (X_l^{(2)})_j \cdot (U_2^{(1)})_j \right)^2 \right]. \end{aligned}$$

Рассмотрим $U_1^{(1)}$ и $U_2^{(1)}$. В силу задания ряда, ортонормированным базисом $U_1^{(1)}$ и $U_2^{(1)}$ пространства $\mathfrak{L}_r^{(1)}$, порожденного элементами $f_n^{(1)} = \sin(2\pi\omega_1 n + \phi_1)$ являются $\sin(2\pi\omega_1 n)$ и $\cos(2\pi\omega_1 n)$, при условии $L\omega_1$ — целое число.

Пусть $p_1 = \sin(2\pi\omega_1 n)$, $p_2 = \cos(2\pi\omega_1 n)$. Вычислим нормы p_1 и p_2 для поиска $U_1^{(1)}$ и $U_2^{(1)}$. По аналогии со знаменателем индекса однородности (п. 5.1.1.), $\|p_1\| = \|p_2\| \approx \sqrt{\frac{L}{2}}$, откуда $U_1^{(1)} = \frac{\sin(2\pi\omega_1 n)}{\sqrt{L/2}}$, $U_2^{(1)} = \frac{\cos(2\pi\omega_1 n)}{\sqrt{L/2}}$.

Пусть

$$I_l = \left(\sum_{j=1}^L (X_l^{(2)})_j \cdot (U_1^{(1)})_j \right)^2,$$

$$J_l = \left(\sum_{j=1}^L (X_l^{(2)})_j \cdot (U_2^{(1)})_j \right)^2,$$

$$a = \omega_1 + \omega_2, \quad b = \omega_1 - \omega_2.$$

Тогда

$$I_l \approx \left(\int_0^L \sin(2\pi\omega_2 y + \psi_l) \cdot \frac{\sin(2\pi\omega_1 y)}{\sqrt{L/2}} dy \right)^2 =$$

$$\begin{aligned}
&= \frac{2}{L} \left(\int_0^L \sin(2\pi\omega_2 y) \cdot \sin(2\pi\omega_1 y) dy \right)^2 = \\
&= \frac{2}{L} \left(\frac{\sin(2\pi Lb - \psi_l) + \sin(\psi_l)}{4\pi b} - \frac{\sin(2\pi La + \psi_l) - \sin(\psi_l)}{4\pi a} \right)^2. \\
\\
J_l &\approx \left(\int_0^L (\sin(2\pi\omega_2 y + \psi_l) \cdot \frac{\cos(2\pi\omega_1 y)}{\sqrt{L/2}}) dy \right)^2 = \\
&= \frac{2}{L} \left(\int_0^L (\sin(2\pi\omega_2 y + \psi_l) \cdot \cos(2\pi\omega_1 y)) dy \right)^2 = \\
&= \frac{2}{L} \left(\frac{\cos(2\pi Lb - \psi_l) - \cos(\psi_l)}{4\pi b} - \frac{\cos(2\pi La + \psi_l) - \cos(\psi_l)}{4\pi a} \right)^2.
\end{aligned}$$

Так как ψ_l формируются из сдвига, порождаемого номером вектора вложения, а исходный ряд F_N задан синусом, при суммировании и достаточно большом значении K_2 , сумма по периоду обращается в 0, поэтому при переходе к сумме по элементам $X_l^{(2)}, l = 1, \dots, K_2$, зависимость от ψ_l пропадает.

Обозначим

$$\begin{aligned}
I &= \frac{2}{L} \left(\frac{\sin(2\pi Lb)}{4\pi b} - \frac{\sin(2\pi La)}{4\pi a} \right)^2. \\
J &= \frac{2}{L} \left(\frac{\cos(2\pi Lb) - 1}{4\pi b} - \frac{\cos(2\pi La) - 1}{4\pi a} \right)^2.
\end{aligned}$$

С учетом предположения выше, получаем:

$$\begin{aligned}
&\sum_{l=1}^{K_2} \sum_{i=1}^r \langle X_l^{(2)}, U_i^{(1)} \rangle^2 \approx K_2 \cdot [I_l + J_l] = \\
&= \frac{K_2 \cdot 2}{L} \cdot \left[\left(\frac{\sin(2\pi Lb)}{4\pi b} - \frac{\sin(2\pi La)}{4\pi a} \right)^2 + \left(\frac{\cos(2\pi Lb) - 1}{4\pi b} - \frac{\cos(2\pi La) - 1}{4\pi a} \right)^2 \right]
\end{aligned}$$

4.2. Индекс неоднородности

Собирая все вместе, получаем:

$$g(F^{(1)}; F^{(2)}) = 1 - \frac{\sum_{l=0}^{K_2-1} \sum_{i=0}^{r-1} \langle X_l^{(2)}, U_i^{(1)} \rangle^2}{\sum_{l=0}^{K_2-1} \|X_l^{(2)}\|^2} \approx$$

$$\begin{aligned}
& \approx 1 - \frac{\frac{K_2 \cdot 2}{L} \cdot \left[\left(\frac{\sin(2\pi Lb)}{4\pi b} - \frac{\sin(2\pi La)}{4\pi a} \right)^2 + \left(\frac{\cos(2\pi Lb)-1}{4\pi b} - \frac{\cos(2\pi La)-1}{4\pi a} \right)^2 \right]}{K_2 \cdot \frac{L}{2}} = \\
& = 1 - \frac{\left[\left(\frac{\sin(2\pi Lb)}{4\pi b} - \frac{\sin(2\pi La)}{4\pi a} \right)^2 + \left(\frac{\cos(2\pi Lb)-1}{4\pi b} - \frac{\cos(2\pi La)-1}{4\pi a} \right)^2 \right]}{\frac{L^2}{4}} = g_a(\omega_1, \omega_2). \quad (4.1)
\end{aligned}$$

4.3. Проверка точности аппроксимации

При сравнении индекса неоднородности, вычисленного классическим способом и аналитически упрощенным, результаты оказались довольно похожи, причем при $L \rightarrow \infty$ оба значения сходятся друг к другу. Все тесты доступны в гитхаб [4] репозитории в файле **Analytical.ipynb**.

4.3.1. Одинаковые частоты

Пусть $N = 700$, $Q = 301$, $B = 100$, $T = 100$. Зададим $\omega_1 = \frac{1}{10}$, $\omega_2 = \frac{1}{10}$, $L = 60$. При одинаковых частотах значения индексов неоднородности должны быть равны 0. Действительно, по определению g пространство \mathfrak{L}_τ , порожденное рядом F_N является одним и тем же для любых подрядов $F^{(1)}$ и $F^{(2)}$ ряда F_N так как структура ряда не менялась. Проверая этот теоретический факт на практике, получили также 0.

4.3.2. $L\omega_1$ и $L\omega_2$ целые, $\omega_1 \neq \omega_2$

Еще один теоретический факт:

При целых $L\omega_1$ и $L\omega_2$ индекс однородности z обращается в 0. Действительно, в силу построения ряда (гармоника, описываемая синусом с какими-то фиксированными параметрами), скалярное произведение частей исходного ряда по целому периоду на элементы базиса (ортогональные друг другу) обращают числитель в 0, следовательно индекс неоднородности g всегда равен 1, что наблюдается на практике.

4.3.3. Предположения об L

Зафиксируем ω_2 и будем изменять L .

Чтобы наглядно продемонстрировать стремление значений друг к другу, посмотрим на Рис. 4.1 и 4.2.

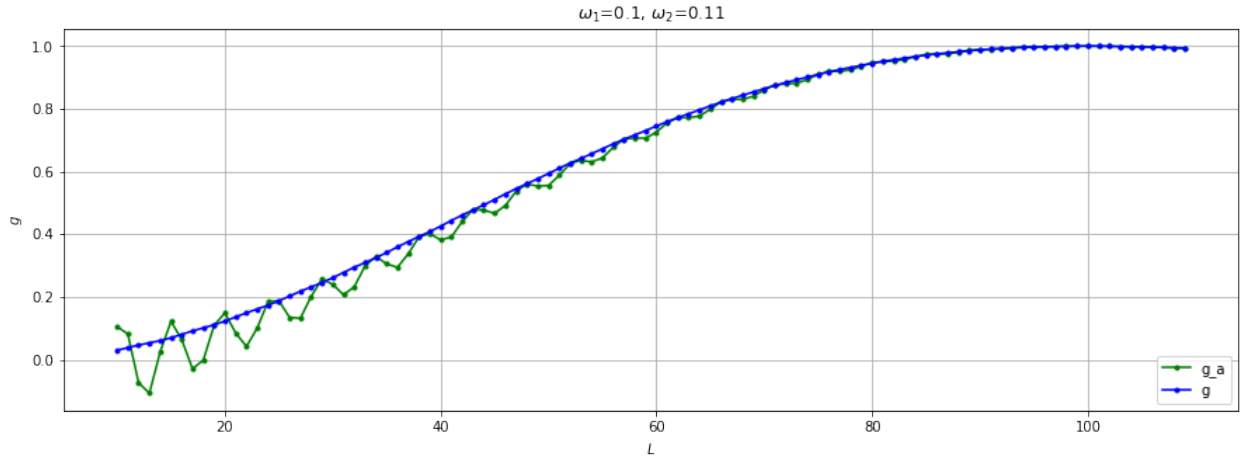


Рис. 4.1. Зависимость индексов g и g_a от L при $\omega_2 = 0.11$.

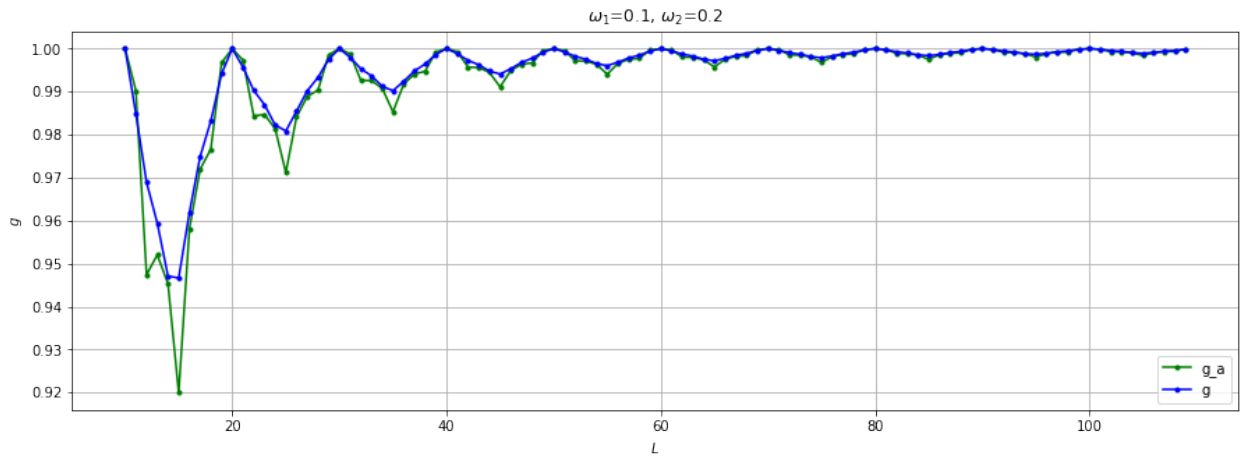


Рис. 4.2. Зависимость индексов g и g_a от L при $\omega_2 = 0.2$.

Исходя из кривых на Рис. 4.1 и 4.2, $g \rightarrow 1$ при $L \rightarrow \infty$ — справедливо для обеих формул (2.1), (4.1).

Данный эксперименты подтверждает что при достаточно больших L аппроксимация индекса неоднородности выведенной аналитической формулой (4.1) точна.

4.3.4. Разность ω_1 и ω_2

Утверждение 1. *Чем сильнее разница ω_1 и ω_2 , тем проще определить разладку.*

Иными словами, чем больше разница ω_1 и ω_2 , тем быстрее индекс неоднородности g переходит в 1.

Пусть $\omega_1 \geq \omega_2$, $\omega_2 \rightarrow \omega_1$. Аналитически, в пределе $a = 2\omega_1$, $b = 0$. Тогда формула

(4.1) примет вид

$$g_a(\omega_1, \omega_2) \approx g_a(\omega_1, \omega_1) = 1 - \frac{\left(\frac{L}{2} - \frac{\sin(4\pi L\omega_1)}{8\pi\omega_1} + \frac{\omega_1}{2\pi\omega_1^2}\right)^2}{\frac{L^2}{4}} \approx 1 - \frac{\left(\frac{L}{2}\right)^2}{\frac{L^2}{4}} = 1 - 1 = 0$$

При проверке данного предположения, получились значения $g = 0.0$, $g_a = 3.330669e-16 \approx 0$.

Посмотрим на график зависимости g от ω_2 при фиксированном $\omega_1 = \frac{1}{10}$.

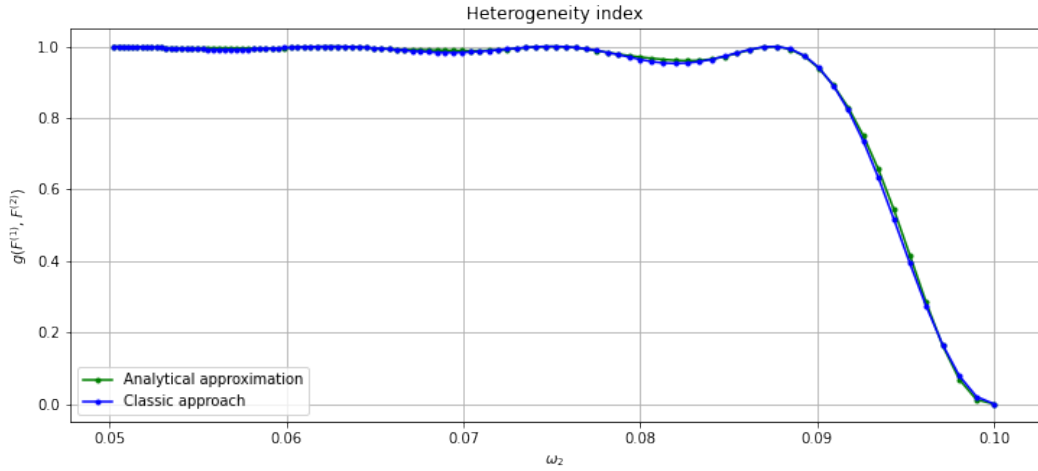


Рис. 4.3. Зависимость индексов g и g_a от ω_2 .

По кривым на Рис. 4.3 видим, что чем ближе частоты друг к другу, тем ближе индекс неоднородности к 0, и, соответственно, чем дальше, тем ближе g к 1, что подтверждает утверждение 1.

Глава 5

Система обнаружения структурной неоднородности ряда с автоматически выстраиваемым порогом срабатывания

В данной главе будет рассматриваться только строковая функция обнаружения неоднородности $d_n^{(r)}$, обозначим ее как d_n

Рассмотрим ряд $F_N = (f_1, \dots, f_N)$, причем

$$f_n = \begin{cases} C \sin(2\pi\omega_1 n + \phi_1), & n \in [1, Q-1] \\ C \sin(2\pi\omega_2 n + \phi_2), & n \in [Q, N], \end{cases}$$

где Q — момент возмущения.

Обозначим

$$F^{(1)} = F_{i,i+B}|_{i=1}^{N-B},$$

$$F^{(2)} = F_{j,j+T}|_{j=1}^{N-T},$$

$$X_l = (f_l, \dots, f_{l+L-1})^T, \quad 1 \leq l < K_2.$$

Рассмотрим систему, которая получает на вход ряд F_N и длину интервала $k \in \mathbb{N}$, за которую нужно определить момент возмущения \hat{Q} .

Алгоритм 1. *Описание системы:*

1. *Входные данные:* $F_N, \gamma^* \in \mathbb{R}_{[0,1]}$;
2. *Результат:* \hat{Q} — момент преодоления d_n значения γ ;

Введем требование к системе — $\hat{Q} \in [Q, Q+k]$, где параметр $k \in \mathbb{N}$ — максимальное запаздывание для обнаружения неоднородности, полученный из входных данных.

Поскольку подача на вход значения γ^* не гарантирует соблюдения данного требования, будем выбирать его алгоритмически. Таким образом, основная задача состоит в нахождении порога.

Алгоритм 2. *Нахождение порога γ^* :*

1. Построить функцию $\gamma : \mathbb{N} \rightarrow [0, 1]$, аппроксимирующую поведение функции d_n после момента Q .
2. Брать γ^* как значение функции γ в точке k .

Для построения γ , сначала рассмотрим его нижнюю и верхнюю границы, то есть значения функции d_n до разладки и после.

5.1. Оценка γ

5.1.1. Нижняя граница

При наличии шума ϵ с дисперсией σ^2 , значения d_n до момента Q смещаются от 0 вверх примерно на $\frac{\sigma^2}{C^2/2 + \sigma^2}$ (Рис. 5.1). Если взять нижнюю границу γ слишком маленьким, $\hat{Q} < Q$, то есть произойдет ложное срабатывания системы. Пусть нижняя граница γ_{min} задается как

$$\gamma_{min}(\sigma^2) = \frac{\sigma^2}{C^2/2 + \sigma^2} \quad (5.1)$$

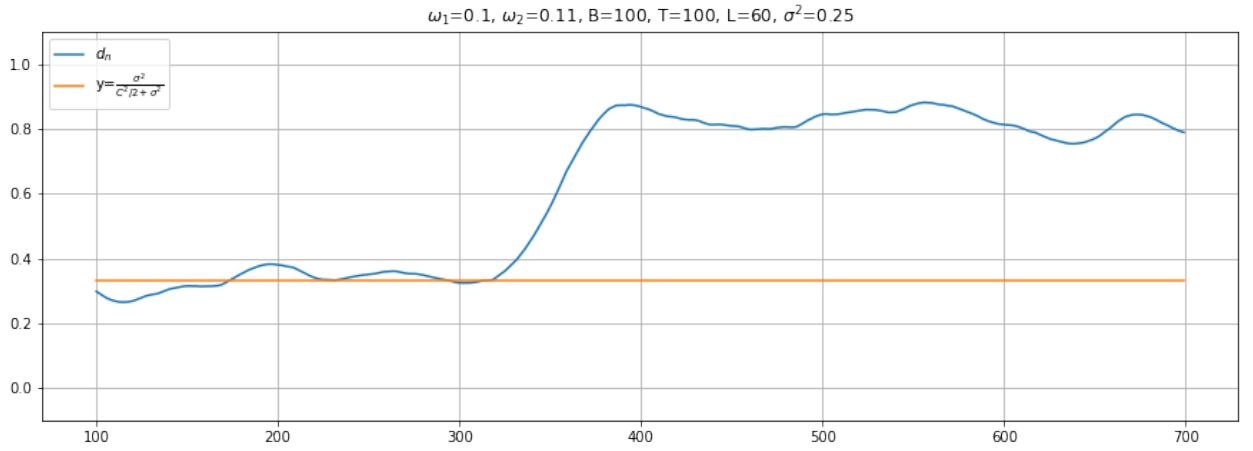


Рис. 5.1. Оценка нижней границы γ .

5.1.2. Верхняя граница

В качестве верхней границы γ будет выступать значение функции d_n после нарушения однородности, для вычисления которой мы можем воспользоваться аналитической аппроксимацией индекса неоднородности $g_a(\omega_1, \omega_2)$ по формуле (4.1), однако нам нужно знать частоты до и после разладки.

По свойству индекса неоднородности, чем больше $|\omega_2 - \omega_1|$, тем ближе g к 1 после переходного интервала, следовательно, кривая d_n на переходном интервале будет иметь более крутой наклон (Рис. 5.2).

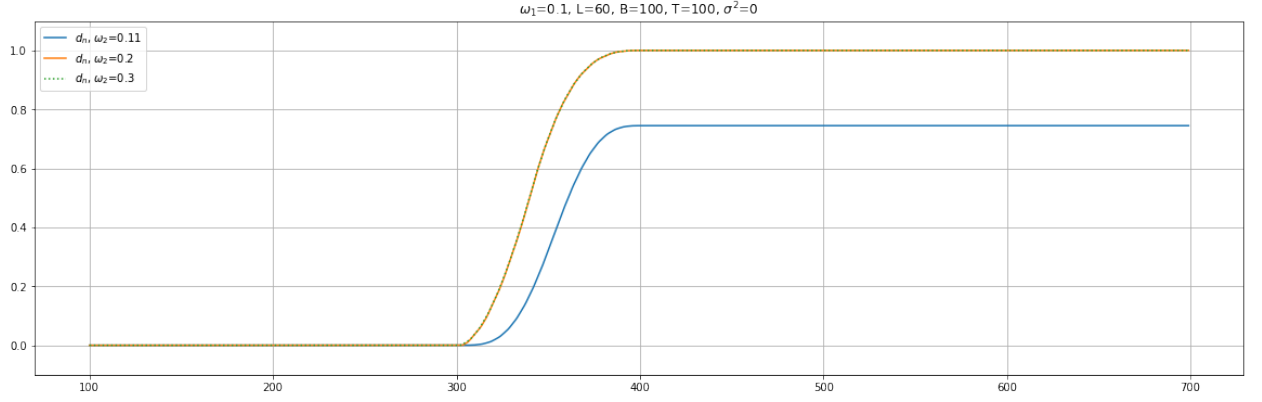


Рис. 5.2. Скорость возрастания d_n на переходном интервале.

Добавим еще 2 параметра, подаваемых на вход системе:

1. ω_1 — начальная частота ряда;
2. $\omega_{min} = \omega_1 + \Delta_{min}$, где Δ_{min} — минимальное для обнаружения неоднородности отклонение частоты ряда от ω_1 ;

Таким образом, имея значение $g_a(\omega_1, \omega_{min})$, мы можем вычислить γ^* как значение аппроксимации переходного интервала $\gamma(i), i = 1, \dots, T$ кривой $d_n, n = Q, \dots, Q + T$, в точке k , которая соединяет γ_{min} и $g_a(\omega_1, \omega_{min})$.

Введем еще одно обозначение для наглядности дальнейших результатов исследования. Так как γ — аппроксимация переходного интервала функции d_n , введем a_n — аппроксимацию всей кривой d_n .

$$a_n = \begin{cases} \gamma_{min}, n < Q, \\ \gamma(n), n \in [Q, Q + T], \\ g_a(\omega_1, \omega_{min}), n > Q. \end{cases} \quad (5.2)$$

5.2. Аппроксимация переходного интервала

Рассмотрим траекторные матрицы $\mathbb{X}_{test}^{(j)}$ размерности $L \times K_{test}$ тестовых рядов $F_N^{(2)} = F_{j,j+T}$, где $K_{test} = T - L + 1, j \in [0, N - T]$.

$$\forall j \in [0, Q - T), \forall n \in [1, K_{test}] : X_n \in \mathbb{X}_{test}^{(j)}, X_n \in \mathfrak{L}_r^{(1)}.$$

$$\forall j \in [Q + T, N - T), \forall n \in [1, K_{test}] : X_n \in \mathbb{X}_{test}^{(j)}, X_n \notin \mathfrak{L}_r^{(1)}.$$

При $T > 2 \cdot L$, $\forall j \in [Q - T + L, Q - L)$, $\mathbb{X}_{test}^{(j)}$ состоит из:

- $n_B = n_B(j)$ векторов вложений, лежащих в $\mathfrak{L}_r^{(1)}$;
- $n_Q = n_Q(j)$ векторов вложений, содержащих момент возмущения;
- $n_A = n_A(j)$ векторов вложений, содержащих только значения ряда после разладки.

Причем $K_{test} = n_B + n_Q + n_A$

Пусть L — фиксировано.

$$\begin{aligned} g(F^{(1)}; F^{(2)}) &= 1 - \frac{\sum_{l=1}^{K_{test}} \sum_{i=1}^r \langle X_l, U_i^{(1)} \rangle^2}{\sum_{l=1}^{K_{test}} \|X_l\|^2} = \\ &= 1 - \frac{\sum_{l=1}^{n_B} \sum_{i=1}^r \langle X_l, U_i^{(1)} \rangle^2 + \sum_{l=n_B}^{n_B+n_Q} \sum_{i=1}^r \langle X_l, U_i^{(1)} \rangle^2 + \sum_{l=n_B+n_Q}^{K_{test}} \sum_{i=1}^r \langle X_l, U_i^{(1)} \rangle^2}{\sum_{l=1}^{n_B} \|X_l\|^2 + \sum_{l=n_B}^{n_B+n_Q} \|X_l\|^2 + \sum_{l=n_B+n_Q}^{K_{test}} \|X_l\|^2} = \\ &= 1 - \frac{0 + \sum_{l=n_B}^{n_B+n_Q} \sum_{i=1}^r \langle X_l, U_i^{(1)} \rangle^2 + n_A(j) \cdot c_H}{c_1 + \sum_{l=n_B}^{n_B+n_Q} \|X_l\|^2 + c_2} = 1 - \frac{n_A(j) \cdot c_H}{\tilde{c}}, \end{aligned} \quad (5.3)$$

при $T \rightarrow \infty$ в силу $n_Q = o(T)$.

Так как для вычисления d_n мы последовательно смещаем тестовые ряды на 1 элемент, $n_A(j)$ возрастает линейно, начиная с $j = Q - T + L$, $n_A(j) = j - Q + T - L$. Получили в правой части формулы (5.3) линейную функцию.

Действительно, при увеличении T , увеличивается количество K_{test} элементов, что приводит к увеличению n_B и n_A , и уменьшению вклада каждого X_l в g . Так как $\forall j n_Q(j) \leq L$ при $T \rightarrow \infty$ вклад соответствующих X_l в g пренебрежительно мал.

Аналогичная ситуация и при уменьшении L . Таким образом, чем больше разность $T - L$, тем линейнее переходный интервал кривой d_n (Рис. 5.3).

Таким образом, мы можем аппроксимировать переходный интервал кривой d_n от γ_{min} до $g_a(\omega_1, \omega_{min})$ функцией $\gamma(i)$, $i = 1, \dots, T$ (Рис. 5.4, 5.5).

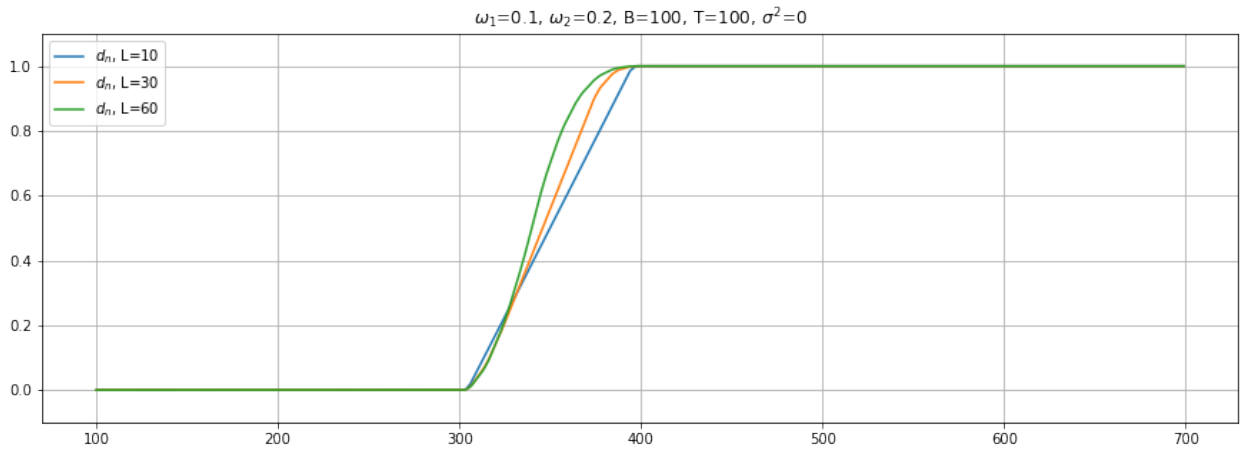


Рис. 5.3. Линейность переходного интервала при увеличении разности $T - L$.

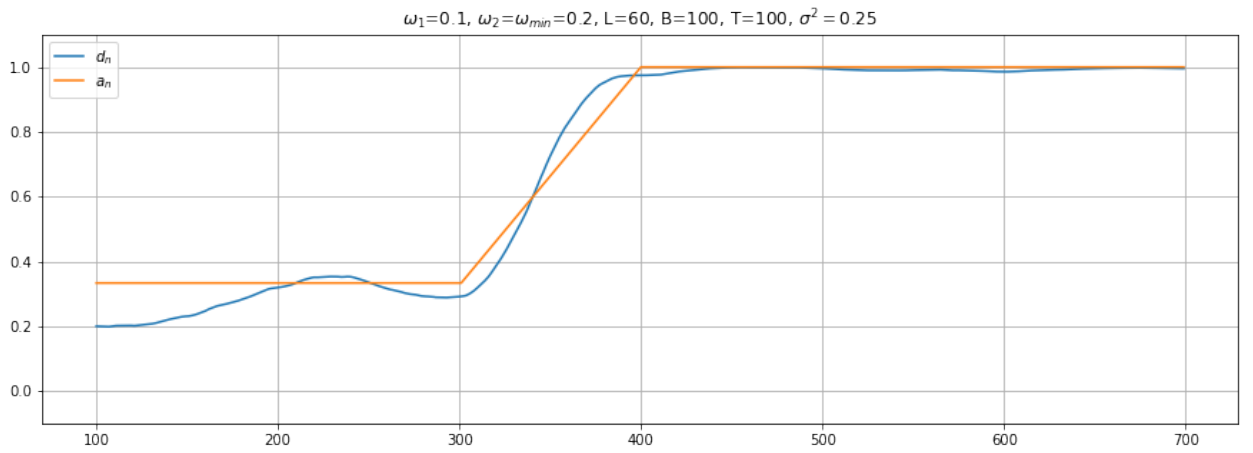


Рис. 5.4. Аппроксимация при отсутствии шума.

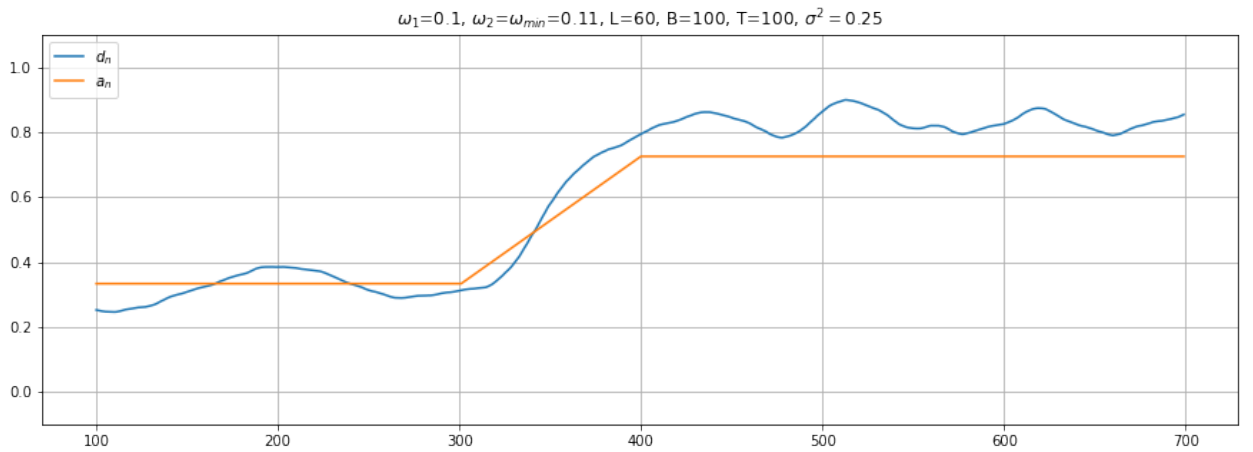


Рис. 5.5. Аппроксимация при наличии шума.

Важно отметить, что такая аппроксимация не всегда корректна (Рис. 5.5). В силу сходимости g_a к g , у нас есть ограничения на L . Однако аппроксимация прямой требует

увеличения $T - L$ и при уменьшении L возникает противоречие.

Поскольку мы хотим брать γ^* как значение в точке k прямой γ , нам важно чтобы первые k значений γ были меньше d_n , ведь в противном случае $\hat{Q} > Q + k$. Таким образом на первый план выходит вопрос не только о корректности аппроксимации γ переходного интервала d_n , но и о выборе доступных для изменения алгоритмом значений параметров B, T и L . Позже попробуем установить их влияние на устойчивость системы.

5.3. Алгоритм работы

Таким образом, с учетом разделов 5.1.1, 5.1.2 и 5.2 уточним алгоритм 1:

Алгоритм 3. *Описание системы с учетом оценки γ и аппроксимации переходного интервала кривой d_n :*

1. *Входные данные:* $F_N, k, \omega_1, \omega_{min}, \sigma^2$;
2. *Результат:* \hat{Q} ;
3. *Алгоритм:*
 - а. *Оцениваем γ_{min} по формуле 5.1;*
 - б. *Вычисляем $g_a(\omega_1, \omega_{min})$ по формуле 4.1;*
 - в. *Строим прямую γ , соединяющую γ_{min} и $g_a(\omega_1, \omega_{min})$;*
 - г. *Фиксируем $\gamma^* = \gamma(k)$;*
 - д. *Определяем \hat{Q} как момент преодоления d_n значения γ^* .*

Дальнейшие исследования системы будут проводиться в соответствии с алгоритмом 3.

5.4. Оценка системы

Зафиксируем дисперсию шума $\sigma^2 = 0.25$ и введем характеристики системы.

Будем считать, что произошло ложноположительное обнаружение неоднородности $FP(\gamma)$ при пороге γ если $\hat{Q} < Q$. Если $\hat{Q} \in [Q, Q + k]$ при пороге γ , то у нас точное обнаружение $TP(\gamma)$. Если же $\hat{Q} > Q + k$ для порога γ , то произошло ложноотрицательное обнаружение неоднородности $FN(\gamma)$

Промоделируем n_{iter} раз реализацию шума ϵ и на каждой итерации посчитаем $FP(\gamma)$, $TP(\gamma)$ и $FN(\gamma)$. Будем характеризовать систему вероятностью ложноположительного обнаружения $FPR(\gamma) = \frac{\sum_{i=1}^{n_{iter}} FP_i(\gamma)}{n_{iter}}$, вероятностью точного обнаружения $TPR(\gamma) = \frac{\sum_{i=1}^{n_{iter}} TP_i(\gamma)}{n_{iter}}$ и вероятностью ложноотрицательного обнаружения $FNR(\gamma) = \frac{\sum_{i=1}^{n_{iter}} FN_i(\gamma)}{n_{iter}}$.

Обозначим порог γ , выбранный по алгоритму выше как γ_a . Посчитаем $\forall \gamma \in [0, 1]$ $FPR(\gamma)$, $TPR(\gamma)$ и $FNR(\gamma)$. Также будем смотреть на $FPR(\gamma_a)$, $TPR(\gamma_a)$ и $FNR(\gamma_a)$.

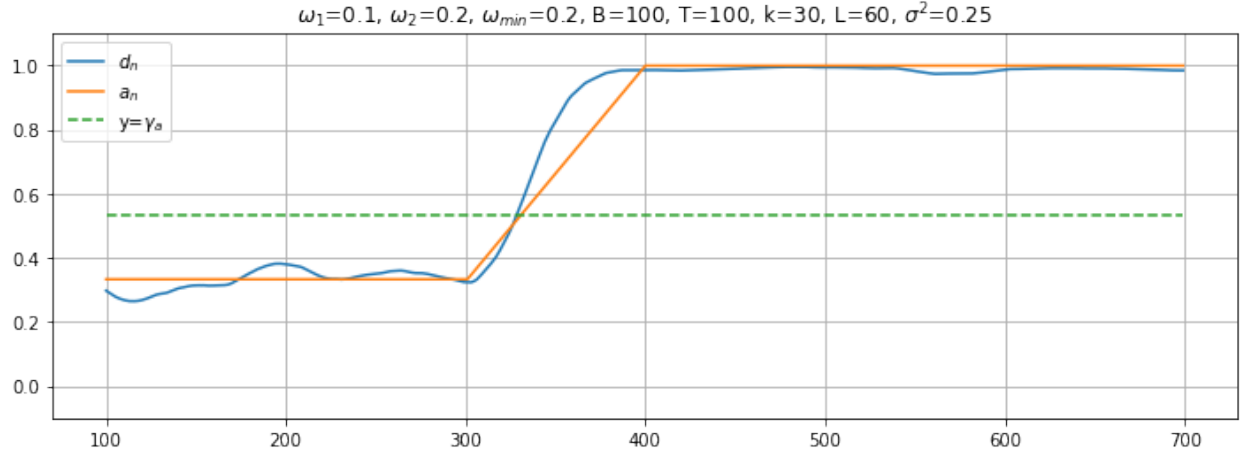


Рис. 5.6. Пример работы системы на одной итерации при $\sigma^2 = 0.25$.

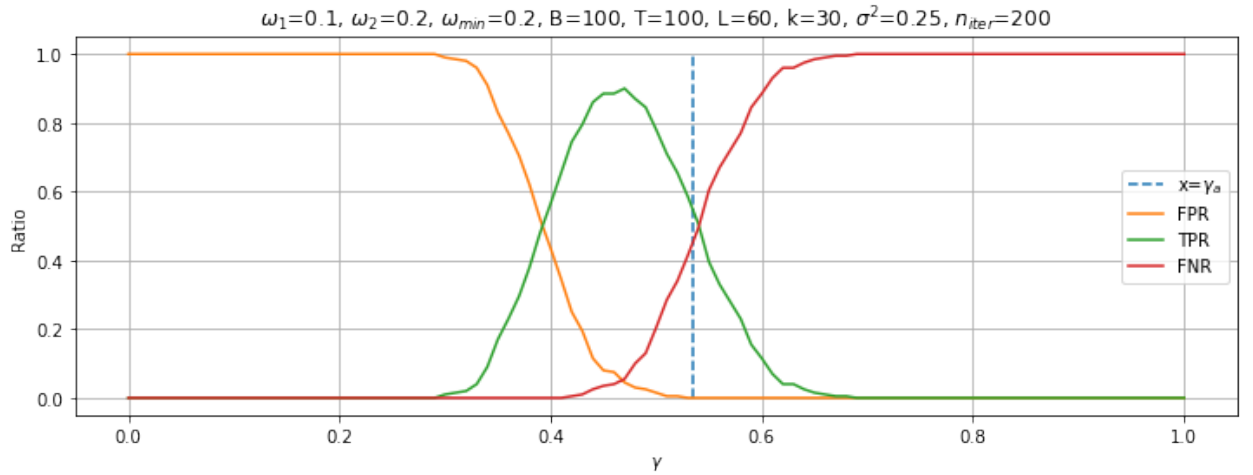


Рис. 5.7. Характеристики системы при $\sigma^2 = 0.25$.

Для параметров, указанных на Рис. 5.7, $FPR(\gamma_a) = 0$, $TPR(\gamma_a) = 0.585$, $FNR(\gamma_a) = 0.415$. Исходя из этих значений, с вероятностью 0.415 предложенный алгоритм не справляется с требованием $\hat{Q} \in [Q, Q + k]$.

При меньшей дисперсии шума σ^2 (Рис. 5.9, 5.8), мы получаем большую устойчивость к ложноположительным обнаружениям, что позволяет уменьшить порог γ_a путем выбора значения на прямой γ раньше, чем в момент k . Об этом нам говорит интервал значений γ на Рис. 5.9, где вероятность точного обнаружения $\text{TPR}(\gamma) = 1$.

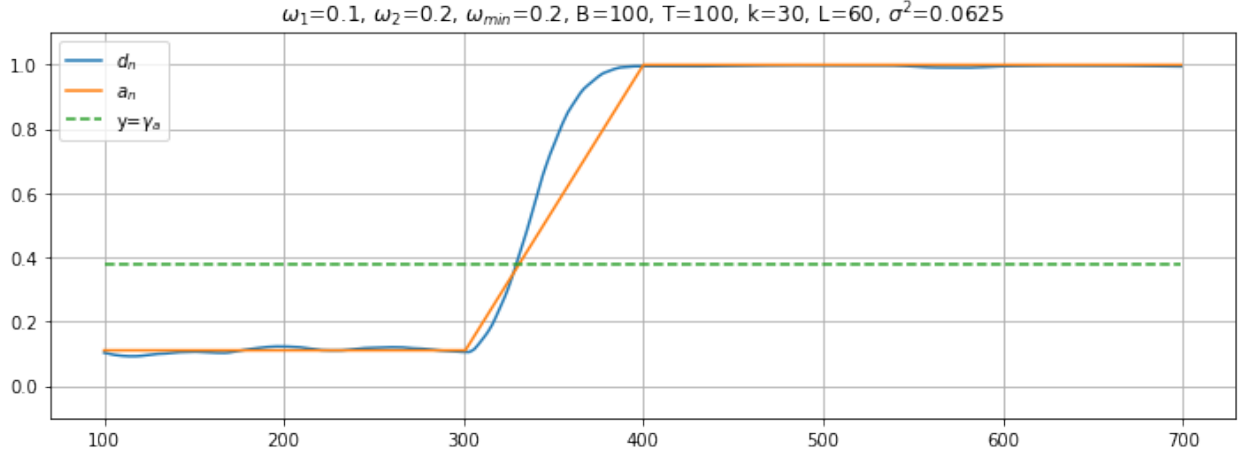


Рис. 5.8. Пример работы системы на одной итерации при $\sigma^2 = 0.0625$.

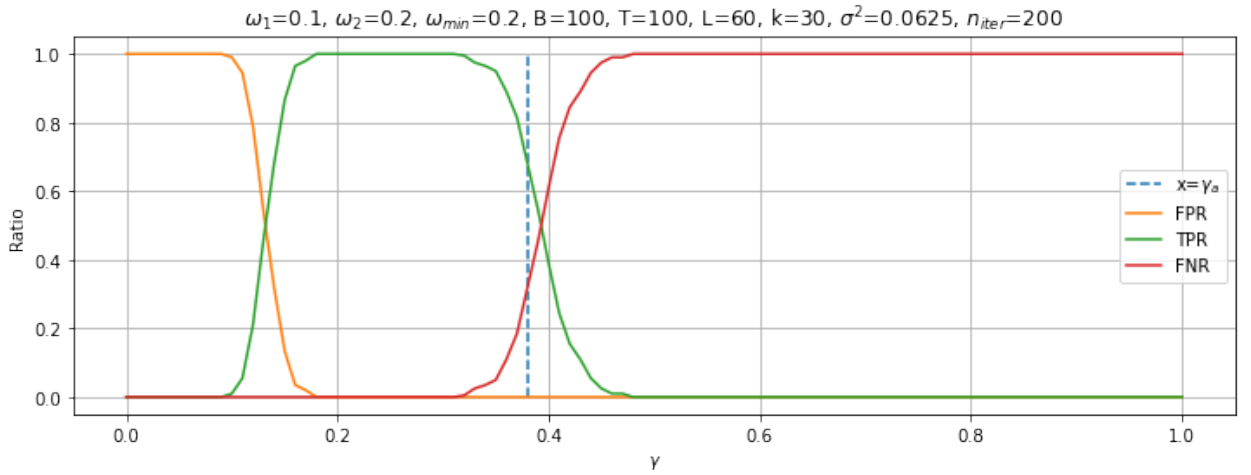


Рис. 5.9. Характеристики системы при $\sigma^2 = 0.0625$.

5.5. Оценка параметров

Все параметры в описываемой системе можно разделить на 4 группы:

1. Входные — ряд и его характеристики до момента нарушения однородности: F_N, ω_1, σ^2 ;
2. Входные — необходимые для работы алгоритма 3: ω_{min}, k ;

3. Неизвестные: ω_2, Q ;
4. Свободные, выбираемые системой: L, B, T . Именно эти параметры мы можем подстраивать под разные ряды. Оценим их.

5.5.1. Оценка параметров: T

При уменьшении параметра T длина переходного интервала сокращается, что увеличивает наклон прямой γ (Рис. 5.10), следовательно, порог γ_a имеет большее значения, что делает систему более устойчивой к шуму и при параметрах, указанных на Рис. 5.11 исключает ложноположительные обнаружения: $\text{FPR}(\gamma_a) = 0$, $\text{TPR}(\gamma_a) = 0.99$, $\text{FNR}(\gamma_a) = 0.01$.

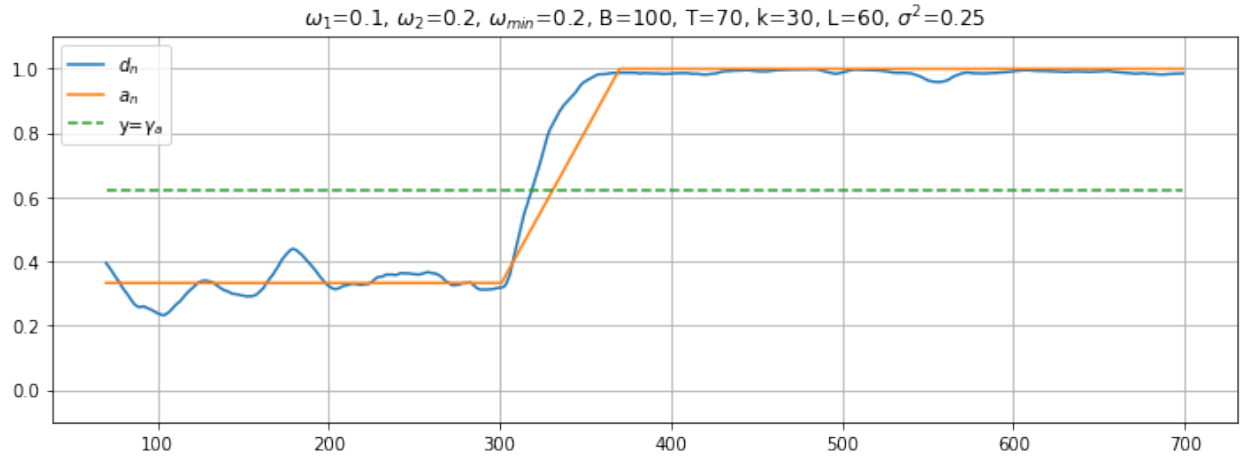


Рис. 5.10. Пример работы системы на одной итерации при $T = 70$.

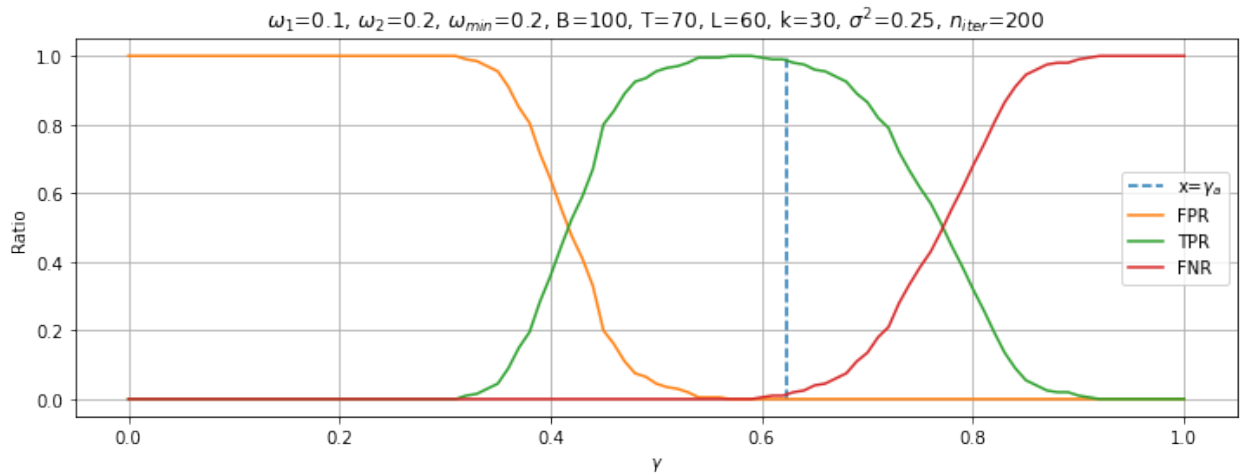


Рис. 5.11. Характеристики системы при $T = 70$.

При маленькой дисперсии шума σ^2 (Рис. 5.13), уменьшая T интервал значений γ , где вероятность точного обнаружения $\text{TPR}(\gamma) = 1$ становится еще больше, по сравнению с ситуацией на Рис. 5.9.

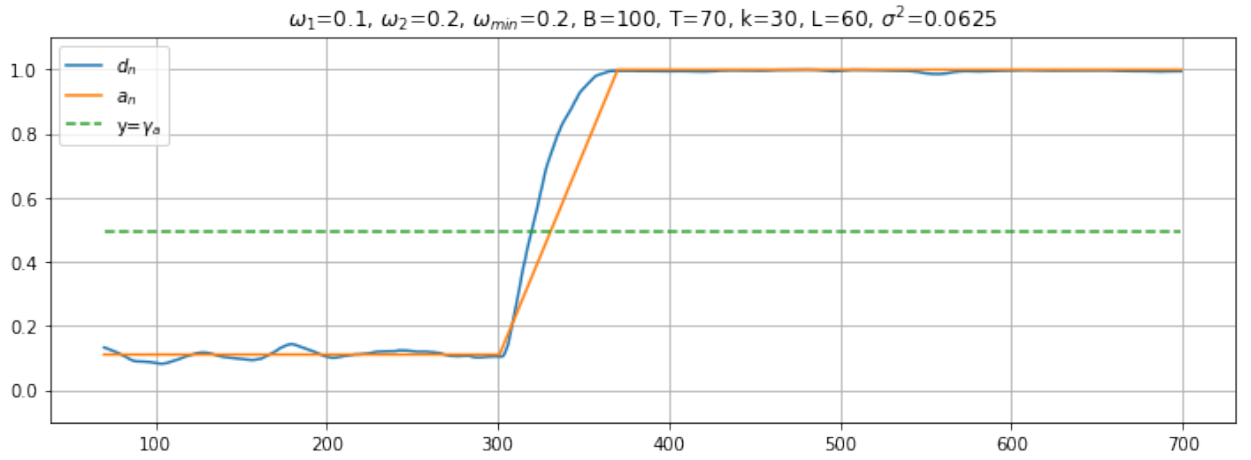


Рис. 5.12. Пример работы системы на одной итерации при $T = 70, \sigma^2 = 0.0625$.

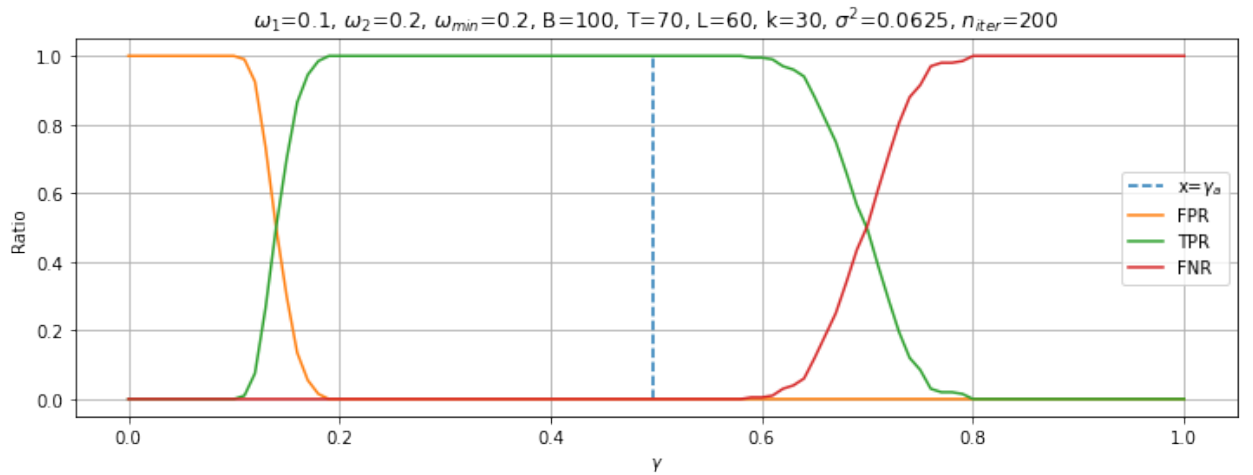


Рис. 5.13. Характеристики системы при $T = 70, \sigma^2 = 0.0625$.

При увеличении параметра T (Рис. 5.14) мы увеличиваем тестовый интервал, занижая параметр γ_a , потенциально, делая систему менее устойчивой при большой дисперсии σ^2 шума ϵ . Также, при недостаточно высокой величине $|T - L|$ линейная аппроксимация γ некорректна и первые значения γ превосходят значения d_n на переходном интервале, что влечет увеличение $\text{FNR}(\gamma_a)$.

При параметрах на Рис. 5.15, $\text{FPR}(\gamma_a) = 0$, $\text{TPR}(\gamma_a) = 0.265$, $\text{FNR}(\gamma_a) = 0.735$. Получили, что с вероятностью в 0.735 требование $\hat{Q} \in [Q, Q + k]$ будет нарушено.

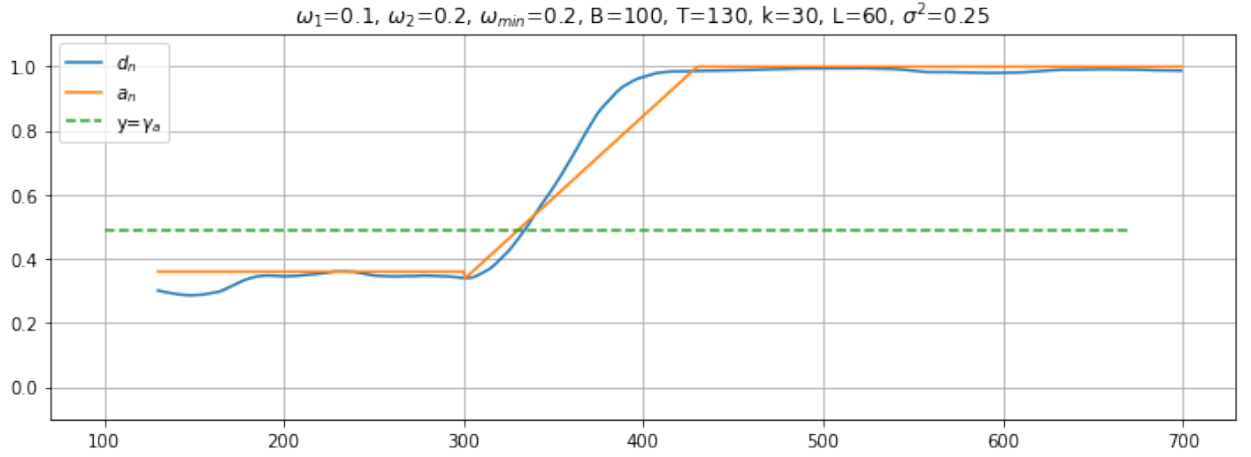


Рис. 5.14. Пример работы системы на одной итерации при $T = 130$.

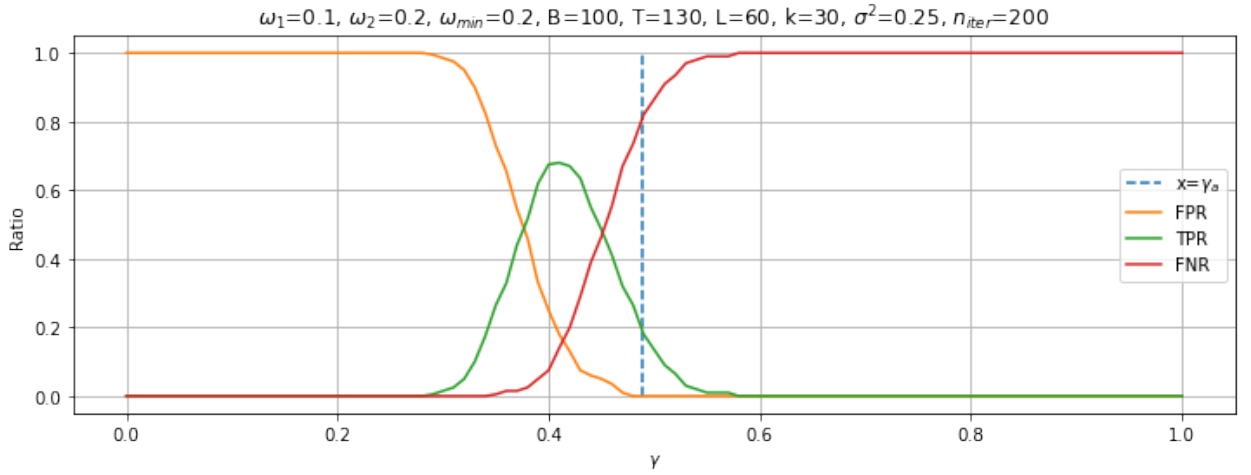


Рис. 5.15. Характеристики системы при $T = 130$.

Таким образом, уменьшая T , мы увеличиваем устойчивость системы, а также интервал, меньший размера k , на котором $\text{TPR}(\gamma) = 1$, следствием чего является $\hat{Q} \rightarrow Q$ — более точное обнаружение момента нарушения однородности.

Однако, при $T \rightarrow L$, $K_{test} \rightarrow 1$, и вклад шума в значения d_n увеличивается, что влечет увеличение ложноположительных обнаружений (Рис. 5.16). Помимо этого, при $T \rightarrow L$ эмпирически установлено, что скорость роста переходного интервала становится быстрее линейной.

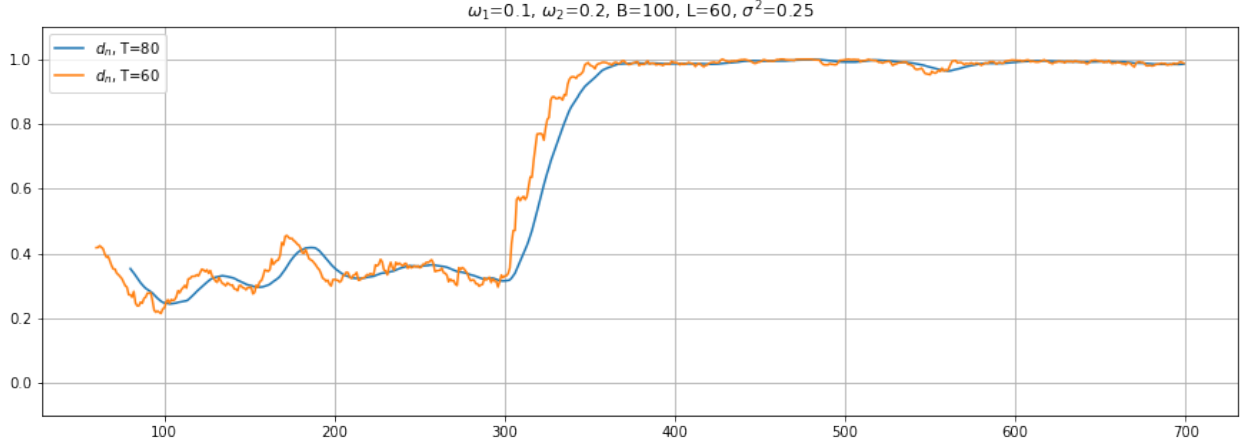


Рис. 5.16. $T \rightarrow L$.

Как было отмечено ранее, сходимость g_a и g достигается при достаточно большом L , однако линейность переходного интервала достигается при противоположном условии. Получается противоречие, так как при увеличении L , количество векторов вложений n_Q , содержащих момент возмущения, также возрастет, и при фиксированном T равенство в формуле (5.3) достигаться не будет в силу $n_Q = O(L)$. Линейность переходного интервала нарушается, однако скорость роста, при уменьшении $T - L$ становится быстрее линейной (Рис. 5.16 и 5.17).

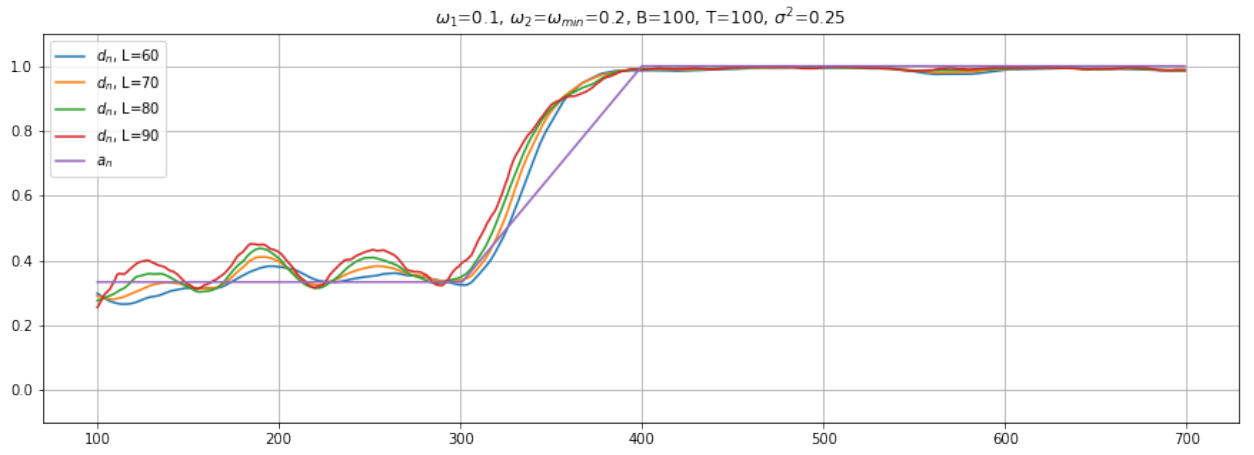


Рис. 5.17. d_n при разных значениях L .

5.5.2. Оценка параметров: B

Параметр B не оказывает значимого влияния на устойчивость системы (Рис. 5.18), так как значения d_n до момента Q мы оцениваем по входному параметру σ^2 .

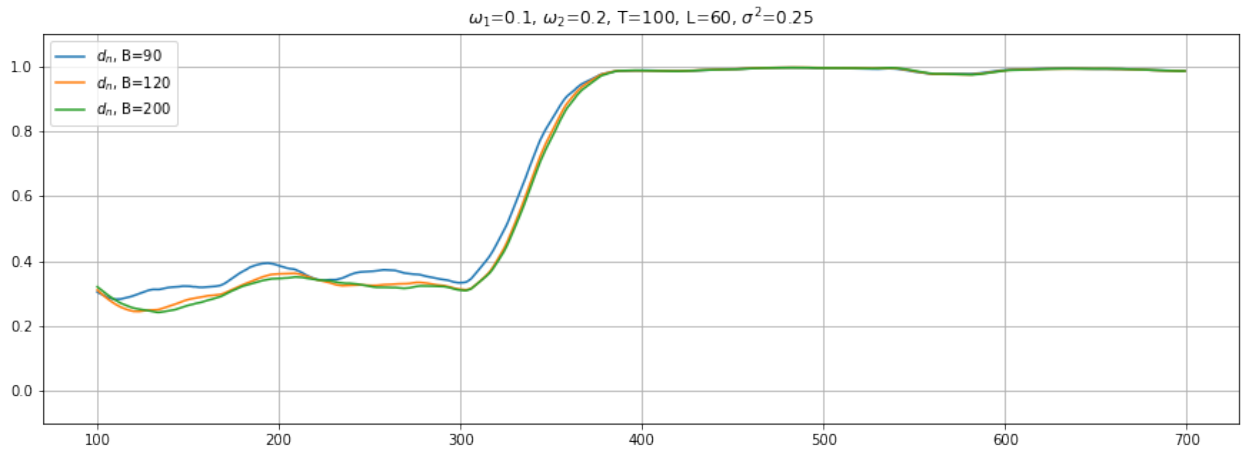


Рис. 5.18. d_n при разных значениях B .

Заключение

В данной работе были рассмотрены и сравнены функции обнаружения неоднородности в синусоидальных временных рядах с неоднородностями, заданными изменением частоты, амплитуды, фазовым сдвигом и выбросом.

Был тщательно рассмотрен и аналитически упрощен индекс неоднородности $g(F^{(1)}, F^{(2)})$. Для аналитической аппроксимации g_a были приведены численные эксперименты, подтверждающие хорошее качество полученной аппроксимации при достаточно большой длине окна L .

Была разработана система автоматического обнаружения неоднородности за указанный временной интервал после неизвестной точки разладки на основе анализа поведения функции неоднородности на переходном интервале.

Анализ показал, что алгоритм построения порога нуждается в доработке с учетом нелинейности функции разладки. Однако, в случае правильного выбора параметров, в частности, и L , алгоритм работает хорошо с точки зрения малого количества ложных срабатываний и запаздываний при обнаружении разладки.

Список литературы

1. Golyandina N., Nekrutkin V., Zhigljavsky A. Analysis of Time Series Structure: SSA and Related Techniques // Monographs on Statistics and Applied Probability. — 2001. — Jan. — Vol. 90.
2. Golyandina N., Korobeynikov A., Zhigljavsky A. Singular spectrum analysis with R. Use R! — Springer-Verlag Berlin Heidelberg, 2018.
3. Korobeynikov A. Computation- and space-efficient implementation of SSA // Statistics and Its Interface. — 2010. — Vol. 3, no. 3. — P. 357–368. — R package version 1.0.4.
4. Access mode: <https://github.com/Loulaan/researchWork>.