

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.М03-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	5
1.1. Алгоритм базового метода SSA	5
1.1.1. Вложение	5
1.1.2. Сингулярное разложение	5
1.1.3. Группировка	6
1.1.4. Реконструкция	6
1.2. Ранги ряда	6
1.2.1. Пример	7
Глава 2. Поиск разладки	8
2.1. Матрица неоднородности и функции неоднородности	8
2.1.1. Матрица неоднородности	8
2.1.2. Функции обнаружения	10
2.2. Однородность и неоднородность	11
2.2.1. Типы неоднородности	11
Глава 3. Сравнение функций разладки	13
3.1. Постановка задачи	13
3.2. Экспериментальные установки	14
3.3. Моделирование	15
3.3.1. Ряды без шума	16
3.3.2. Ряды с шумом	22
3.3.3. Выводы	29
Глава 4. Аналитическая оценка индекса неоднородности при изменении частоты гармоник	30
4.1. Вычисление индекса однородности	31
4.1.1. Знаменатель	31
4.1.2. Числитель	31
4.2. Индекс неоднородности	32

4.3.	Проверка точности аппроксимации	33
4.3.1.	Одинаковые частоты	33
4.3.2.	$L\omega_1$ и $L\omega_2$ целые, $\omega_1 \neq \omega_2$	33
4.3.3.	Предположения об L	34
4.3.4.	Разность ω_1 и ω_2	34
Глава 5. Система обнаружения структурной неоднородности ряда с ав-		
томатически выстраиваемым порогом срабатывания на основе выве-		
денной аналитической формулы		36
5.1.	Оценка γ	36
5.1.1.	Нижняя граница	36
5.1.2.	Верхняя граница	37
5.2.	Аппроксимация переходного интервала	38
5.3.	Алгоритм работы	40
5.4.	Оценка системы	41
5.5.	Оценка параметров	43
5.5.1.	Оценка параметров: T	43
5.5.2.	Оценка параметров: B	47
5.5.3.	Оценка параметров: L	47
Заключение		49
Список литературы		50

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, анализу поведения переходного интервала функций обнаружения и создания системы обнаружения неоднородности.

В «Главе 1» приведена теория базового алгоритма SSA.

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

«Глава 3» посвящена 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 тесты подтверждающие эквивалентность новой формулы и старой.

«Глава 5» посвящена созданию системы, которая обнаруживает момент возникновения неоднородности не позже, чем через заданный момент 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

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

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 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 $\alpha_1, \dots, \alpha_d \forall n : f_{n+r} = \sum_{i=1}^r \alpha_i \cdot f_{n+r-i}, \alpha_i \neq 0$.

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 (Гл. 1.1.1) (это значение должно быть больше размерности минимального LRR) вложенные векторы покрывают одно и то же линейное пространство $\mathfrak{L}^{(L)}$ независимо от N (если N достаточно велико). Таким образом, нарушения однородности ряда могут быть описаны в терминах соответствующих сдвинутых векторов: возмущения заставляют эти векторы покидать пространство $\mathfrak{L}^{(L)}$. Соответствующие расхождения определяются в терминах расстояний между сдвинутыми векторами 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 L -сдвинутые векторы 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 $X_1^{(2)}, \dots, X_{K_2}^{(2)} (K_2 = N_2 - L + 1)$ L -сдвинутые векторы временного ряда $F^{(2)}$.

Введем меру, называемую **индексом неоднородности**, которая характеризует несоответствие между рядом $F^{(2)}$ и структурой ряда $F^{(1)}$ (описываемого 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 $F_N : F_N = (f_0, \dots, f_{N-1})$, $N > 2$;
2. Подряды (интервалы) $F_{i,j}$ временного ряда $F_N : F_{i,j} = (f_{i-1}, \dots, f_{j-1})$, $1 \leq i < j \leq N$;
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 L -сдвинутые векторы 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 L -сдвинутые векторы 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.

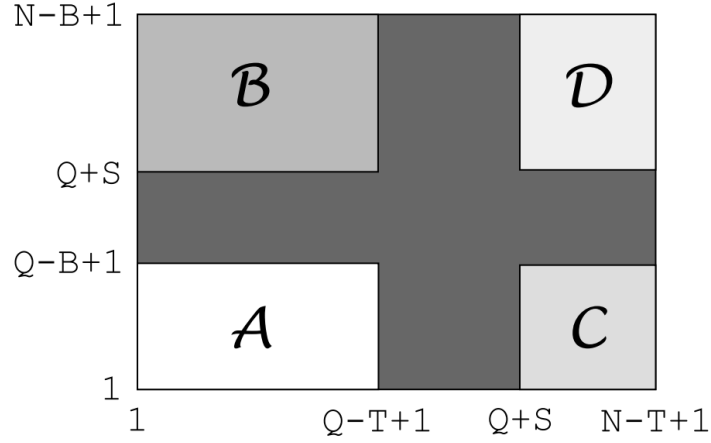


Рис. 2.1. Общая форма матрицы неоднородности.

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

Сравнение функций разладки

Так как у нас есть четыре функции обнаружения неоднородности:

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

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: $\omega_1 \neq \omega_2$.

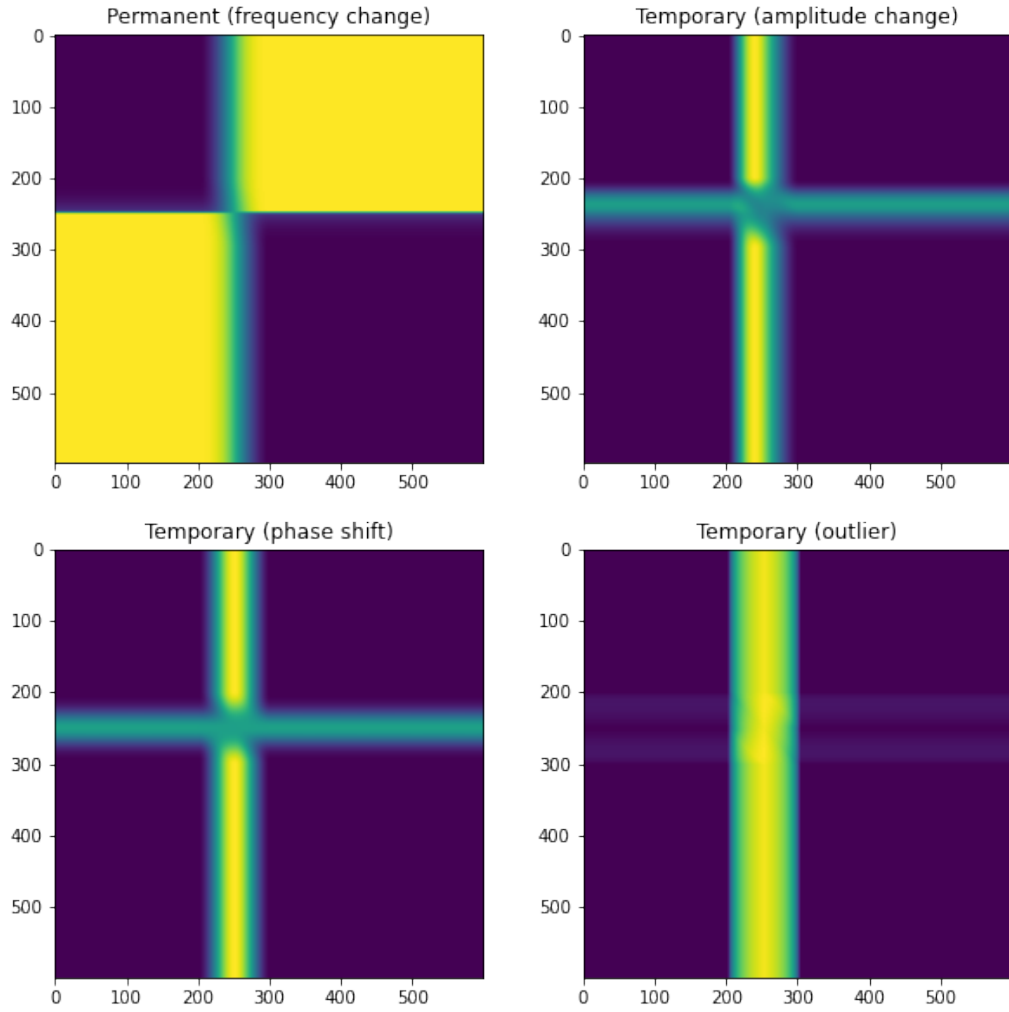


Рис. 3.1. Временные ряды без шума.

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. Экспериментальные установки

Для реализации тестов был выбран язык *Python3*. Взаимодействие с пакетом *RSSA* осуществлялось через библиотеку *rpy2* (стадии вложения и сингулярного раз-

ложения). Подсчет индексов неоднородности велось 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 ниже.

3.3. Моделирование

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}_r .

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 ($\omega_1 \neq \omega_2$)	row	col	sym	diag
meanMax	0.2391	0.2062	0.2339	0.2185
95 procentile	0.2352	0.2054	0.2297	0.2167
Temporary ($C_1 \neq C_2$)	row	col	sym	diag
meanMax	0.0685	0.0593	0.0669	0.0624
95 procentile	0.0673	0.0591	0.0657	0.0616
Temporary ($\phi_1 \neq \phi_2$)	row	col	sym	diag
meanMax	0.2372	0.2044	0.2319	0.2166
95 procentile	0.2336	0.2036	0.2279	0.2143
Temporary (Outlier)	row	col	sym	diag
meanMax	0.2336	0.2069	0.2287	0.2161
95 procentile	0.2307	0.2061	0.2254	0.2139

3.3.1. Ряды без шума

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

Полученные результаты тестирования функций разладки приведены в таблицах 3.2, 3.3, 3.4, 3.5.

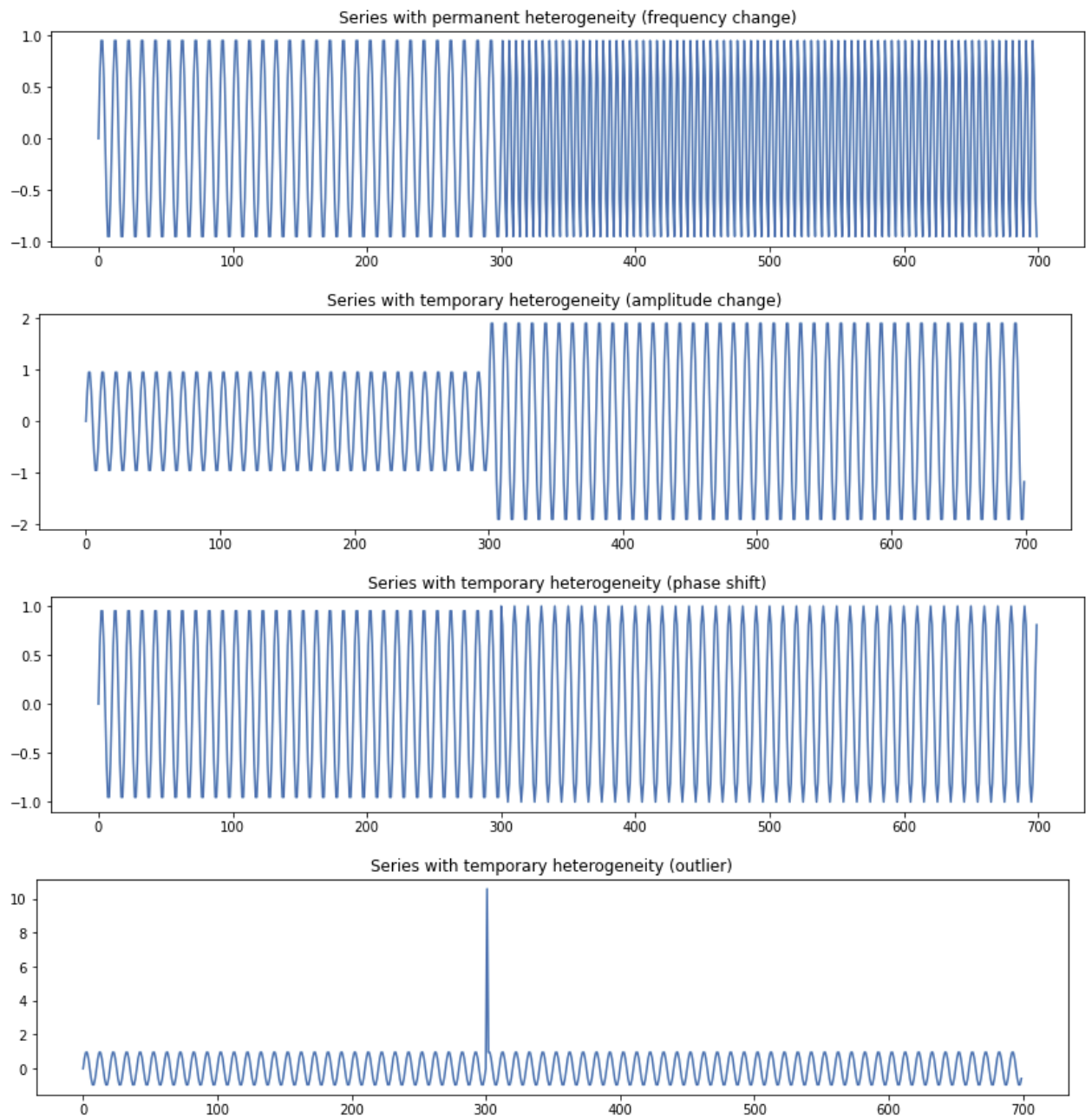


Рис. 3.2. Временные ряды без шума.

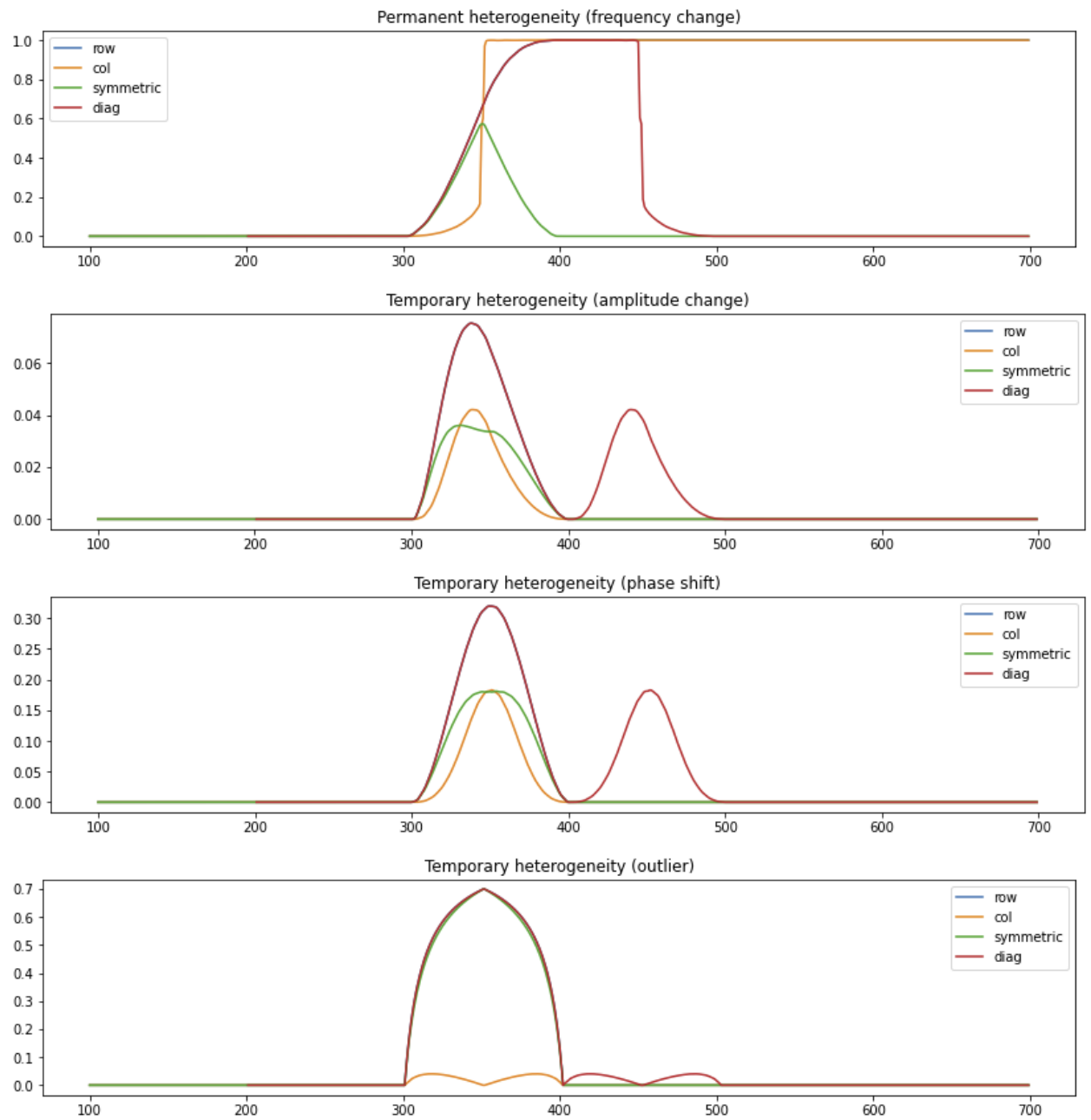


Рис. 3.3. Функции неоднородности рядов без шума.

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

Row	meanMax	95 procentile	Col	meanMax	95 procentile
$n_{overcome}$	328	328	$n_{overcome}$	350	350
$D_Q^{(r)}$	0.0000	0.0000	$D_Q^{(c)}$	0.0000	0.0000
$D_{Q+10}^{(r)}$	0.0428	0.0428	$D_{Q+10}^{(c)}$	0.0028	0.0028
$D_{Q+20}^{(r)}$	0.1468	0.1468	$D_{Q+20}^{(c)}$	0.0140	0.0140
$D_{Q+30}^{(r)}$	0.2962	0.2962	$D_{Q+30}^{(c)}$	0.0385	0.0385
Sym	meanMax	95 procentile	Diag	meanMax	95 procentile
$n_{overcome}$	329	329	$n_{overcome}$	327	327
$D_Q^{(s)}$	0.0000	0.0000	$D_Q^{(d)}$	0.0000	0.0000
$D_{Q+10}^{(s)}$	0.0402	0.0402	$D_{Q+10}^{(d)}$	0.0428	0.0428
$D_{Q+20}^{(s)}$	0.1354	0.1354	$D_{Q+20}^{(d)}$	0.1468	0.1468
$D_{Q+30}^{(s)}$	0.2706	0.2706	$D_{Q+30}^{(d)}$	0.2962	0.2962

Таблица 3.3. Характеристики функций неоднородности для временной разладки ($C_1 \neq C_2$).

Row	meanMax	95 procentile	Col	meanMax	95 procentile
$n_{overcome}$	330	330	$n_{overcome}$		
$D_Q^{(r)}$	0.0000	0.0000	$D_Q^{(c)}$		
$D_{Q+10}^{(r)}$	0.0186	0.0186	$D_{Q+10}^{(c)}$		
$D_{Q+20}^{(r)}$	0.0491	0.0491	$D_{Q+20}^{(c)}$		
$D_{Q+30}^{(r)}$	0.0703	0.0703	$D_{Q+30}^{(c)}$		
Sym	meanMax	95 procentile	Diag	meanMax	95 procentile
$n_{overcome}$			$n_{overcome}$	327	326
$D_Q^{(s)}$			$D_Q^{(d)}$	0.0000	0.0000
$D_{Q+10}^{(s)}$			$D_{Q+10}^{(d)}$	0.0186	0.0186
$D_{Q+20}^{(s)}$			$D_{Q+20}^{(d)}$	0.0491	0.0491
$D_{Q+30}^{(s)}$			$D_{Q+30}^{(d)}$	0.0703	0.0703

Таблица 3.4. Характеристики функций неоднородности для временной разладки ($\phi_1 \neq \phi_2$).

Row	meanMax	95 procentile	Col	meanMax	95 procentile
$n_{overcome}$	334	333	$n_{overcome}$		
$D_Q^{(r)}$	0.0008	0.0008	$D_Q^{(c)}$		
$D_{Q+10}^{(r)}$	0.0392	0.0392	$D_{Q+10}^{(c)}$		
$D_{Q+20}^{(r)}$	0.1215	0.1215	$D_{Q+20}^{(c)}$		
$D_{Q+30}^{(r)}$	0.2161	0.2161	$D_{Q+30}^{(c)}$		
Sym	meanMax	95 procentile	Diag	meanMax	95 procentile
$n_{overcome}$			$n_{overcome}$	332	331
$D_Q^{(s)}$			$D_Q^{(d)}$	0.0008	0.0008
$D_{Q+10}^{(s)}$			$D_{Q+10}^{(d)}$	0.0392	0.0392
$D_{Q+20}^{(s)}$			$D_{Q+20}^{(d)}$	0.1215	0.1215
$D_{Q+30}^{(s)}$			$D_{Q+30}^{(d)}$	0.2161	0.2161

Таблица 3.5. Характеристики функций неоднородности для временной разладки (выброс).

Row	meanMax	95 procentile	Col	meanMax	95 procentile
$n_{overcome}$	306	306	$n_{overcome}$		
$D_Q^{(r)}$	0.0000	0.0000	$D_Q^{(c)}$		
$D_{Q+10}^{(r)}$	0.4012	0.4012	$D_{Q+10}^{(c)}$		
$D_{Q+20}^{(r)}$	0.5470	0.5470	$D_{Q+20}^{(c)}$		
$D_{Q+30}^{(r)}$	0.6223	0.6223	$D_{Q+30}^{(c)}$		
Sym	meanMax	95 procentile	Diag	meanMax	95 procentile
$n_{overcome}$	306	306	$n_{overcome}$	305	305
$D_Q^{(s)}$	0.0000	0.0000	$D_Q^{(d)}$	0.0000	0.0000
$D_{Q+10}^{(s)}$	0.3806	0.3806	$D_{Q+10}^{(d)}$	0.4012	0.4012
$D_{Q+20}^{(s)}$	0.5288	0.5288	$D_{Q+20}^{(d)}$	0.5470	0.5470
$D_{Q+30}^{(s)}$	0.6101	0.6101	$D_{Q+30}^{(d)}$	0.6223	0.6223

В таблицах 3.3, 3.4 и 3.5 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, 3.3, 3.4, 3.5 можем сделать вывод, что строковая $d_{n-1}^{(r)}$ и диагональная $d_{n-1}^{(d)}$ функции неоднородности более устойчивые к шуму ϵ , возрастают сильнее, однако $d_{n-1}^{(d)}$ раньше преодолевает промоделированное значение, что значит и более раннее обнаружение разладки.

3.3.2. 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 моделировать реализации шума во временных рядах и подсчитывать количество преодолений $\#n_{overcome}$, на основе которых будем считать средний момент преодоления значений $n_{overcome}$ из таблицы 3.1.

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 3.6, 3.7, 3.8, 3.9.

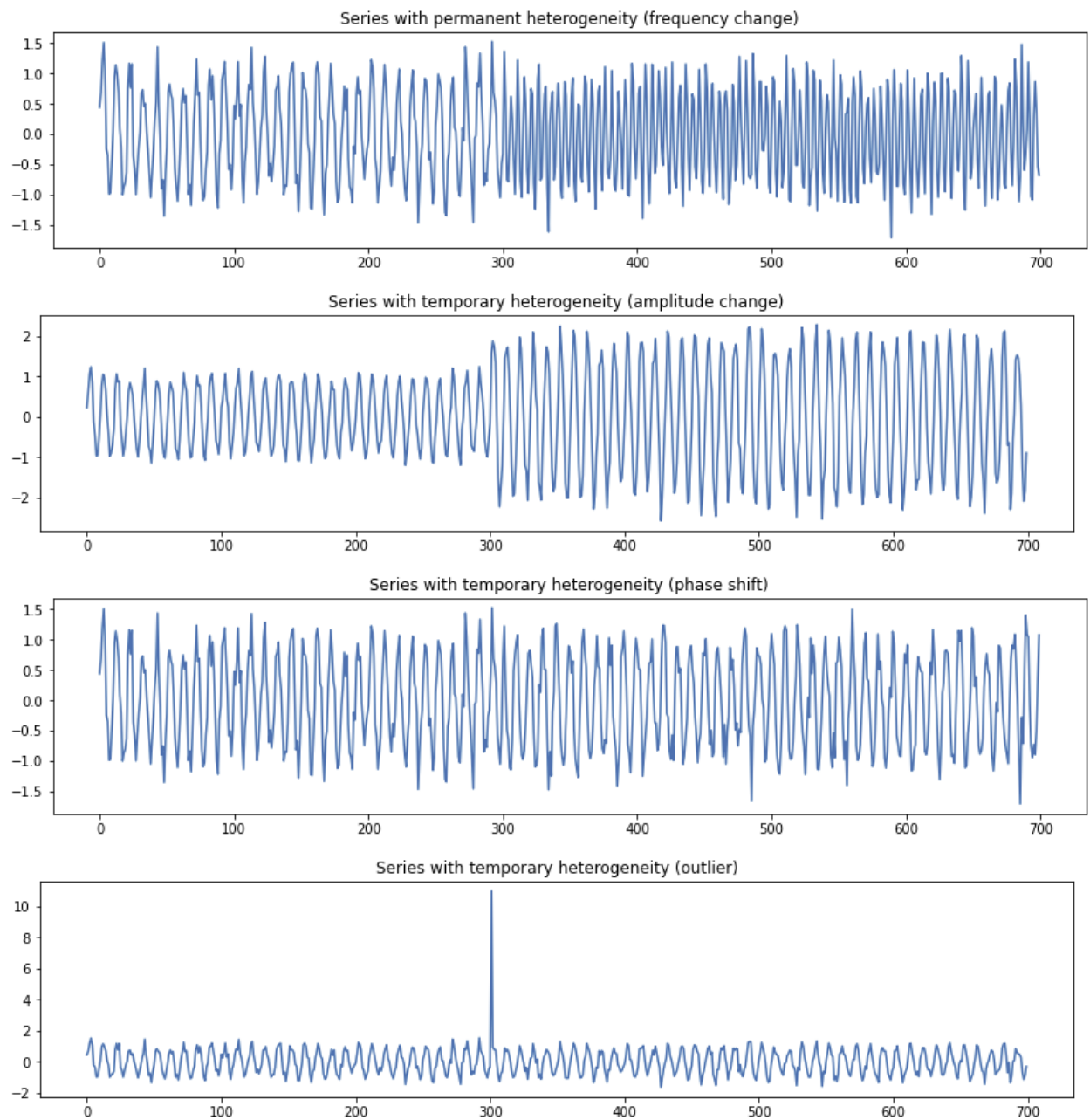


Рис. 3.4. Временные ряды с шумом.

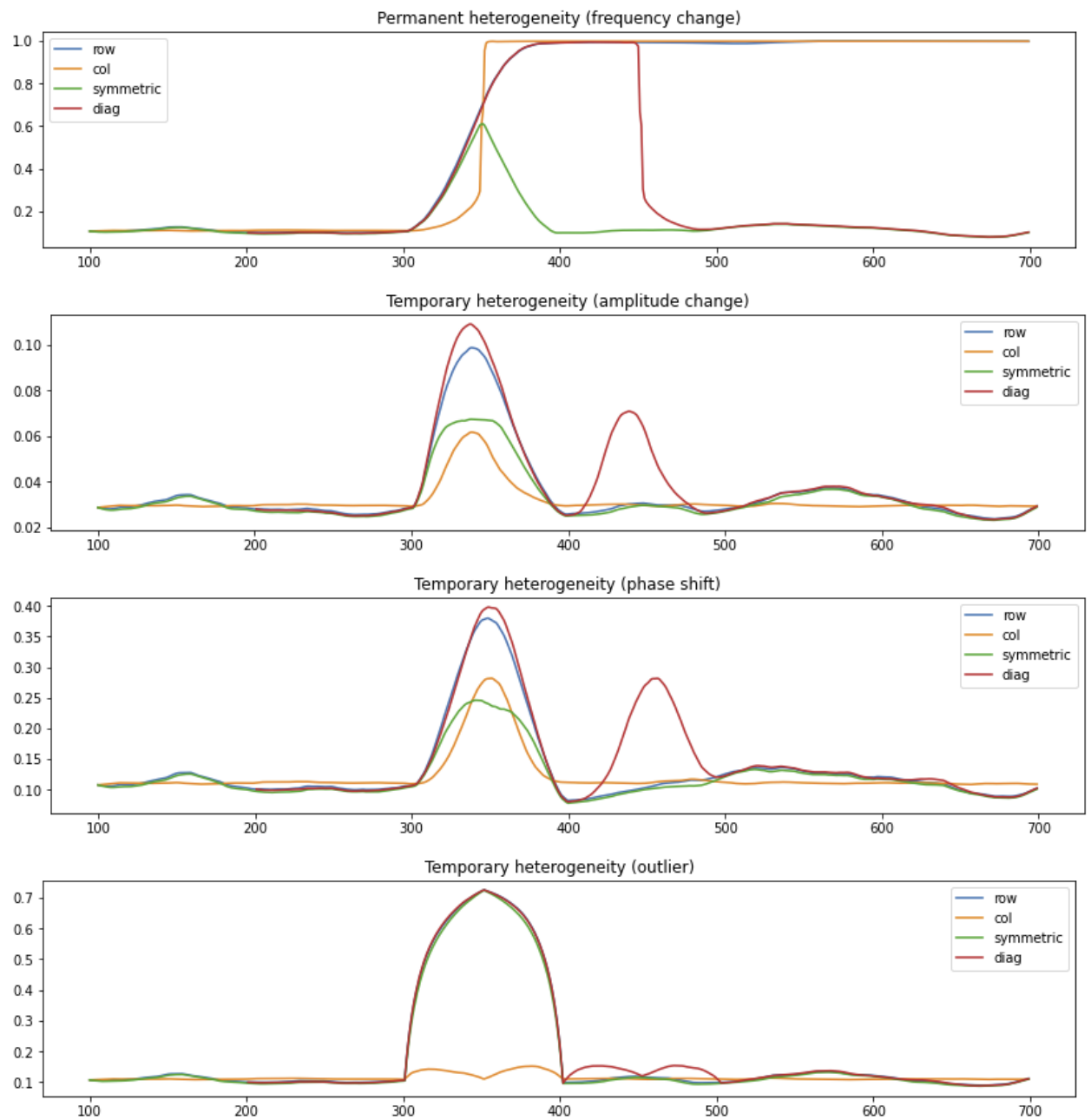


Рис. 3.5. Функции неоднородности рядов с шумом.

Таблица 3.6. Характеристики функций неоднородности для постоянной разладки с шумом ($\omega_1 \neq \omega_2$).

Row	meanMax	95 procentile	Col	meanMax	95 procentile
$\#n_{overcome}$	200	200	$\#n_{overcome}$	200	200
$n_{overcome}$	312.36	311.71	$n_{overcome}$	315.69	315.32
$D_Q^{(r)}$	0.1949	0.1949	$D_Q^{(c)}$	0.1979	0.1979
$D_{Q+10}^{(r)}$	0.2308	0.2308	$D_{Q+10}^{(c)}$	0.2010	0.2010
$D_{Q+20}^{(r)}$	0.3148	0.3148	$D_{Q+20}^{(c)}$	0.2140	0.2140
$D_{Q+30}^{(r)}$	0.4345	0.4345	$D_{Q+30}^{(c)}$	0.2379	0.2379
Sym	meanMax	95 procentile	Diag	meanMax	95 procentile
$\#n_{overcome}$	200	200	$\#n_{overcome}$	200	200
$n_{overcome}$	312.91	312.165	$n_{overcome}$	309.195	308.79
$D_Q^{(s)}$	0.1903	0.1903	$D_Q^{(d)}$	0.1950	0.1950
$D_{Q+10}^{(s)}$	0.2234	0.2234	$D_{Q+10}^{(d)}$	0.2283	0.2283
$D_{Q+20}^{(s)}$	0.2976	0.2976	$D_{Q+20}^{(d)}$	0.3056	0.3056
$D_{Q+30}^{(s)}$	0.4044	0.4044	$D_{Q+30}^{(d)}$	0.4210	0.4210

Таблица 3.7. Характеристики функций неоднородности для временной разладки с шумом ($_1 \neq _2$).

Row	meanMax	95 procentile	Col	meanMax	95 procentile
$\#n_{overcome}$	200	200	$\#n_{overcome}$	200	200
$n_{overcome}$	308.475	308.03	$n_{overcome}$	309.07	308.88
$D_Q^{(r)}$	0.0564	0.0564	$D_Q^{(c)}$	0.0578	0.0578
$D_{Q+10}^{(r)}$	0.0768	0.0768	$D_{Q+10}^{(c)}$	0.0612	0.0612
$D_{Q+20}^{(r)}$	0.1028	0.1028	$D_{Q+20}^{(c)}$	0.0735	0.0735
$D_{Q+30}^{(r)}$	0.1181	0.1181	$D_{Q+30}^{(c)}$	0.0861	0.0861
Sym	meanMax	95 procentile	Diag	meanMax	95 procentile
$\#n_{overcome}$	200	200	$\#n_{overcome}$	200	200
$n_{overcome}$	311.14	309.94	$n_{overcome}$	305.64	305.34
$D_Q^{(s)}$	0.0551	0.0551	$D_Q^{(d)}$	0.0564	0.0564
$D_{Q+10}^{(s)}$	0.0723	0.0723	$D_{Q+10}^{(d)}$	0.0789	0.0789
$D_{Q+20}^{(s)}$	0.0868	0.0868	$D_{Q+20}^{(d)}$	0.1095	0.1095
$D_{Q+30}^{(s)}$	0.0900	0.0900	$D_{Q+30}^{(d)}$	0.1284	0.1284

Таблица 3.8. Характеристики функций неоднородности для временной разладки с шумом ($\phi_1 \neq \phi_2$).

Row	meanMax	95 procentile	Col	meanMax	95 procentile
$\#n_{overcome}$	200	200	$\#n_{overcome}$	200	200
$n_{overcome}$	312.08	311.395	$n_{overcome}$	310.81	310.52
$D_Q^{(r)}$	0.1994	0.1994	$D_Q^{(c)}$	0.1997	0.1997
$D_{Q+10}^{(r)}$	0.2313	0.2313	$D_{Q+10}^{(c)}$	0.2053	0.2053
$D_{Q+20}^{(r)}$	0.2978	0.2978	$D_{Q+20}^{(c)}$	0.2312	0.2312
$D_{Q+30}^{(r)}$	0.3672	0.3672	$D_{Q+30}^{(c)}$	0.2790	0.2790
Sym	meanMax	95 procentile	Diag	meanMax	95 procentile
$\#n_{overcome}$	200	200	$\#n_{overcome}$	200	200
$n_{overcome}$	313.8	312.815	$n_{overcome}$	308.79	308.355
$D_Q^{(s)}$	0.1948	0.1948	$D_Q^{(d)}$	0.1993	0.1993
$D_{Q+10}^{(s)}$	0.2217	0.2217	$D_{Q+10}^{(d)}$	0.2277	0.2277
$D_{Q+20}^{(s)}$	0.2671	0.2671	$D_{Q+20}^{(d)}$	0.2874	0.2874
$D_{Q+30}^{(s)}$	0.3009	0.3009	$D_{Q+30}^{(d)}$	0.3599	0.3599

Таблица 3.9. Характеристики функций неоднородности для временной разладки с шумом (выброс).

Row	meanMax	95 procentile	Col	meanMax	95 procentile
$\#n_{overcome}$	200	200	$\#n_{overcome}$	180	181
$n_{overcome}$	302.25	302.185	$n_{overcome}$	311.2389	311.0663
$D_Q^{(r)}$	0.1982	0.1982	$D_Q^{(c)}$	0.2050	0.2049
$D_{Q+10}^{(r)}$	0.4814	0.4814	$D_{Q+10}^{(c)}$	0.2318	0.2316
$D_{Q+20}^{(r)}$	0.5929	0.5929	$D_{Q+20}^{(c)}$	0.2334	0.2333
$D_{Q+30}^{(r)}$	0.6555	0.6555	$D_{Q+30}^{(c)}$	0.2268	0.2266
Sym	meanMax	95 procentile	Diag	meanMax	95 procentile
$\#n_{overcome}$	200	200	$\#n_{overcome}$	200	200
$n_{overcome}$	302.31	302.275	$n_{overcome}$	301.845	301.795
$D_Q^{(s)}$	0.1935	0.1935	$D_Q^{(d)}$	0.1980	0.1980
$D_{Q+10}^{(s)}$	0.4619	0.4619	$D_{Q+10}^{(d)}$	0.4855	0.4855
$D_{Q+20}^{(s)}$	0.5760	0.5760	$D_{Q+20}^{(d)}$	0.5971	0.5971
$D_{Q+30}^{(s)}$	0.6440	0.6440	$D_{Q+30}^{(d)}$	0.6598	0.6598

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.

3.3.3. Выводы

Таблица 3.10. Средние значения характеристик функций обнаружения.

Row	meanMax	95 percentile
$Mean(\#n_{overcome})$	200	200
$Mean(n_{overcome})$	308.791	308.330
$Mean(\ [D_Q^{(r)}, D_{Q+10}^{(r)}, D_{Q+20}^{(r)}, D_{Q+30}^{(r)}] \ _{l_2})$	0.597	0.597
Col	meanMax	95 percentile
$Mean(\#n_{overcome})$	195	195.25
$Mean(n_{overcome})$	311.702	311.447
$Mean(\ [D_Q^{(c)}, D_{Q+10}^{(c)}, D_{Q+20}^{(c)}, D_{Q+30}^{(c)}] \ _{l_2})$	0.370	0.370
Sym	meanMax	95 percentile
$Mean(\#n_{overcome})$	200	200
$Mean(n_{overcome})$	310.040	309.299
$Mean(\ [D_Q^{(s)}, D_{Q+10}^{(s)}, D_{Q+20}^{(s)}, D_{Q+30}^{(s)}] \ _{l_2})$	0.558	0.558
Diag	meanMax	95 percentile
$Mean(\#n_{overcome})$	200	200
$Mean(n_{overcome})$	306.368	306.070
$Mean(\ [D_Q^{(d)}, D_{Q+10}^{(d)}, D_{Q+20}^{(d)}, D_{Q+30}^{(d)}] \ _{l_2})$	0.595	0.595

Явными фаворитами (таблица 3.10) 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 гармоник

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 $F_N = (f_0, \dots, f_{N-1})$, причем

$$f_n = \begin{cases} C_1 \sin(2\pi\omega_1 n + \phi_1), & n \in [0, Q-1] \\ C_2 \sin(2\pi\omega_2 n + \phi_2), & n \in [Q, N-1] \end{cases}$$

Обозначим

$$F^{(1)} = f_n^{(1)}|_{n=0}^{B-1},$$

$$F^{(2)} = f_n^{(2)}|_{n=0}^{T-1},$$

$$X_l^{(2)} = (f_l^{(2)}, \dots, f_{l+L-1}^{(2)})^T, \quad 0 \leq l < K_2.$$

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, 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 некие нормированные $\sin(2\pi\omega_1 n + \psi)$ и $\cos(2\pi\omega_1 n + \psi)$.

Пусть $p_1 = \sin(2\pi\omega_1 n + \psi)$, $p_2 = \cos(2\pi\omega_1 n + \psi)$. 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 $U_1^{(1)} = \frac{\sin(2\pi\omega_1 n + \psi)}{\sqrt{L/2}}$, $U_2^{(1)} = \frac{\cos(2\pi\omega_1 n + \psi)}{\sqrt{L/2}}$.

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

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

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

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

Так как $\frac{\sum_{l=1}^{K_2} \sum_{i=1}^r \langle X_l^{(2)}, U_i^{(1)} \rangle^2}{\sum_{l=1}^{K_2} \|X_l^{(2)}\|^2}$ - линейная проекция на пространство $\mathfrak{L}_r^{(1)}$, эта проекция не должна зависеть от базиса, поэтому $\psi = 0$.

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:

$$\begin{aligned}
g(F^{(1)}; F^{(2)}) &= 1 - \frac{\sum_{l=0}^{K_2-1} \sum_{i=0}^{r-1} \langle X_l^{(2)}, U_i^{(1)} \rangle^2}{\sum_{l=0}^{K_2-1} \|X_l^{(2)}\|^2} \approx \\
&\approx 1 - \frac{\frac{K_2 \cdot 2}{L} \cdot \left[\left(\frac{\sin(2\pi Lb)}{4\pi b} - \frac{\sin(2\pi La)}{4\pi a} \right)^2 + \left(\frac{\cos(2\pi Lb)-1}{4\pi b} - \frac{\cos(2\pi La)-1}{4\pi a} \right)^2 \right]}{K_2 \cdot \frac{L}{2}} = \\
&= 1 - \frac{\left[\left(\frac{\sin(2\pi Lb)}{4\pi b} - \frac{\sin(2\pi La)}{4\pi a} \right)^2 + \left(\frac{\cos(2\pi Lb)-1}{4\pi b} - \frac{\cos(2\pi La)-1}{4\pi a} \right)^2 \right]}{\frac{L^2}{4}} = g_a(\omega_1, \omega_2). \tag{4.1}
\end{aligned}$$

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**¹ репозитории в файле **Analytical.ipynb**².

4.3.1. Одинаковые частоты

Пусть $N = 700$, $Q = 301$, $B = 100$, $T = 100$. Зададим $w_1 = \frac{1}{10}$, $w_2 = \frac{1}{10}$, $L = 60$. 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.

¹ <https://github.com/Loulaan/researchWork>

² <https://github.com/Loulaan/researchWork/blob/main/Analytical.ipynb>

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.

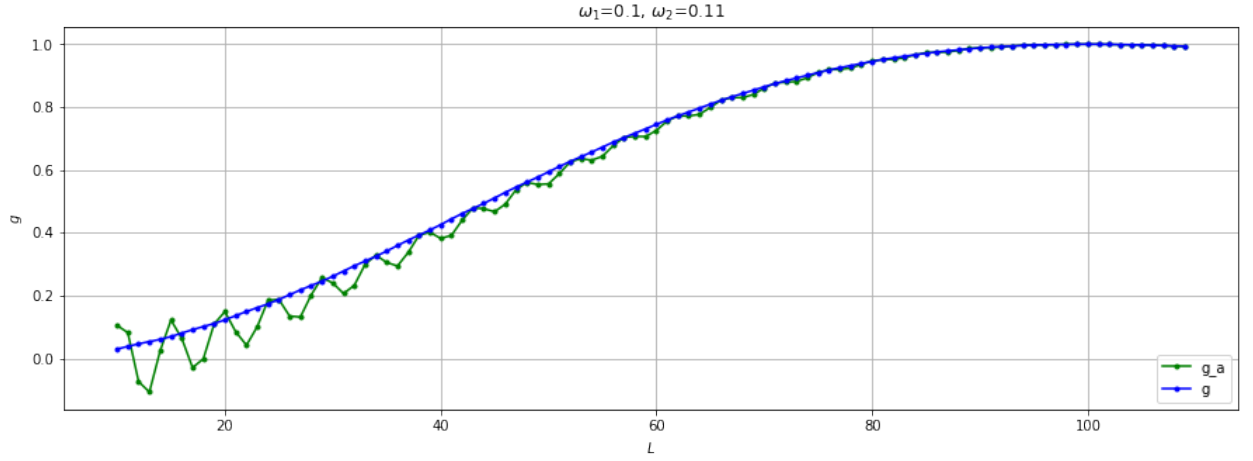


Рис. 4.1. Зависимость индексов g и g_a от L при $\omega_2 = 0.11$.

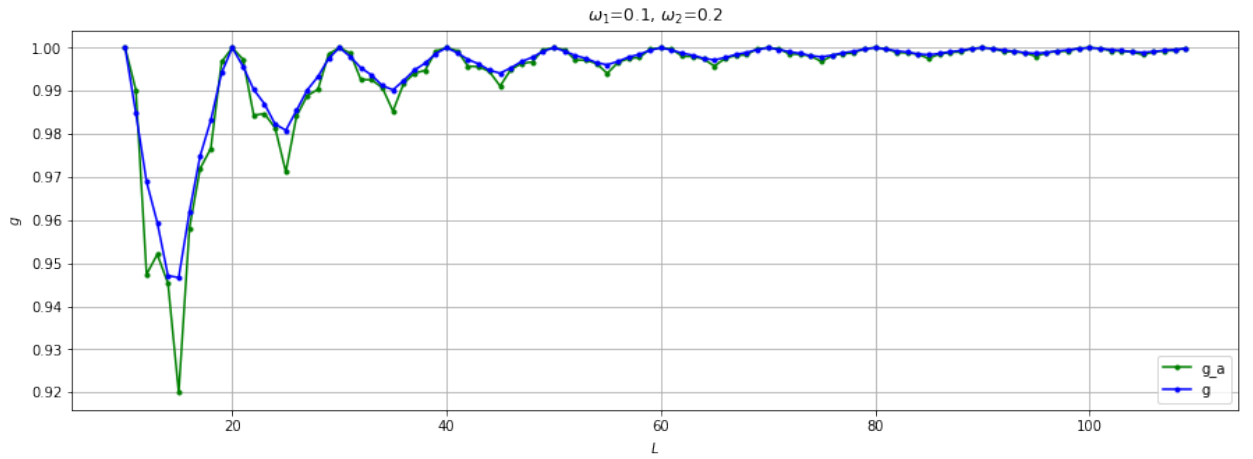


Рис. 4.2. Зависимость индексов g и g_a от L при $\omega_2 = 0.2$.

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

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

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

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

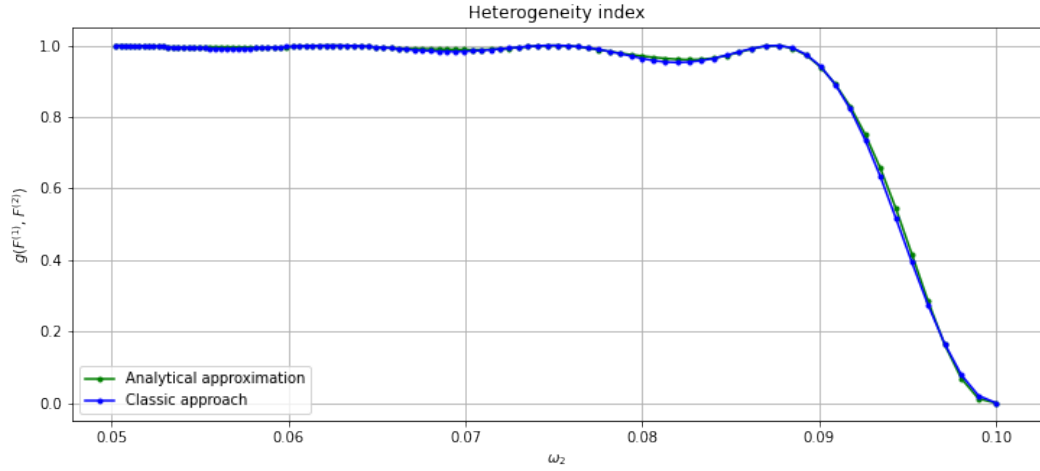


Рис. 4.3. Зависимость индексов g и g_a от ω_2 .

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 на основе выведенной аналитической формулы

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

Рассмотрим ряд $F_N = (f_0, \dots, f_{N-1})$, причем

$$f_n = \begin{cases} C \sin(2\pi\omega_1 n + \phi_1), & n \in [0, Q-1] \\ C \sin(2\pi\omega_2 n + \phi_2), & n \in [Q, N-1], \end{cases}$$

где Q — момент возмущения.

Обозначим

$$F^{(1)} = F_{i,i+B}|_{i=0}^{N-B-1},$$

$$F^{(2)} = F_{j,j+T}|_{j=0}^{N-T-1},$$

$$X_l = (f_l, \dots, f_{l+L-1})^T, \quad 0 \leq l < K_2.$$

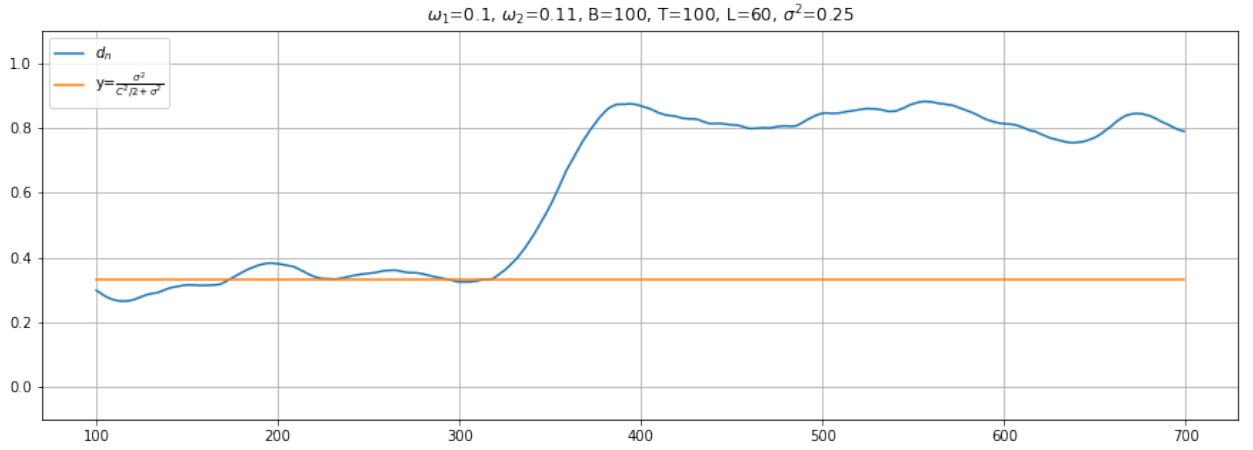
Рассмотрим систему, которая получает на вход ряд F_N и длину интервала $k \in \mathbb{N}$, за которую нужно определить момент возмущения \hat{Q} , причем $\hat{Q} \in [Q, Q+k]$. Параметр k — максимальное запаздывание для обнаружения неоднородности

Будем определять \hat{Q} как момент преодоления порога γ строковой функции обнаружения d_n

5.1. Оценка γ

5.1.1. Нижняя граница

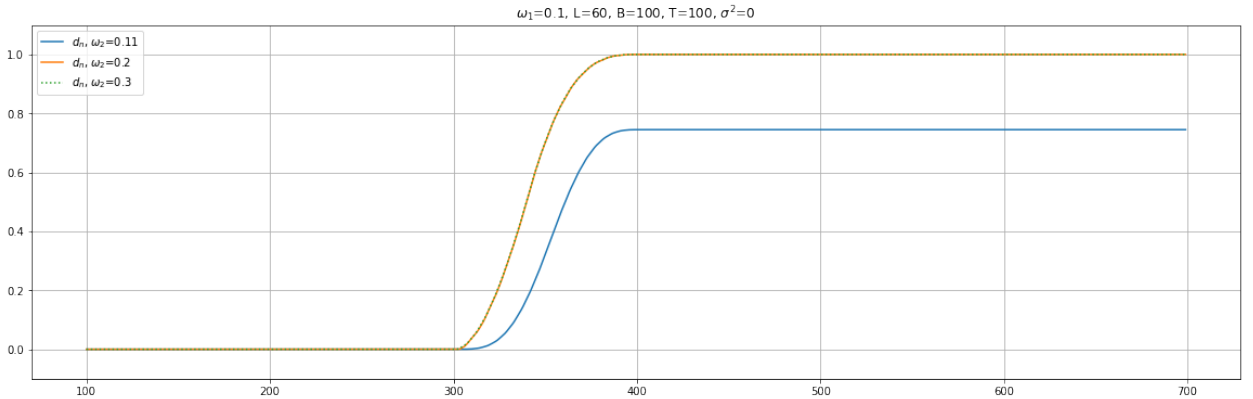
При наличии шума ϵ с дисперсией σ^2 , значения d_n до момента Q смещаются от 0 вверх примерно на $\frac{\sigma^2}{C^2/2+\sigma^2}$ (Рис. 5.1). Если взять значение γ слишком маленьким, $\hat{Q} < Q$. Пусть нижняя граница $\gamma_{min} = \frac{\sigma^2}{C^2/2+\sigma^2}$.

Рис. 5.1. Оценка нижней границы γ .

5.1.2. Верхняя граница

Для определения верхней границы γ мы можем воспользоваться аналитической аппроксимацией $g_a(\omega_1, \omega_2)$ по формуле (4.1), однако для этого 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 $g_a(\omega_1, \omega_{min})$, в качестве верхней границы γ мы можем взять значение на переходном интервале кривой d_n в точке k .

Так как значения на переходном интервале нам неизвестны, мы можем попытаться аппроксимировать его. Обозначим эту аппроксимацию a_T .

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.

$$\begin{aligned} g(F^{(1)}; F^{(2)}) &= 1 - \frac{\sum_{l=1}^{K_{test}} \sum_{i=1}^r \langle X_l, U_i^{(1)} \rangle^2}{\sum_{l=1}^{K_{test}} \|X_l\|^2} = \\ &= 1 - \frac{\sum_{l=1}^{n_B} \sum_{i=1}^r \langle X_l, U_i^{(1)} \rangle^2 + \sum_{l=n_B}^{n_B+n_Q} \sum_{i=1}^r \langle X_l, U_i^{(1)} \rangle^2 + \sum_{l=n_B+n_Q}^{K_{test}} \sum_{i=1}^r \langle X_l, U_i^{(1)} \rangle^2}{\sum_{l=1}^{n_B} \|X_l\|^2 + \sum_{l=n_B}^{n_B+n_Q} \|X_l\|^2 + \sum_{l=n_B+n_Q}^{K_{test}} \|X_l\|^2} = \\ &= 1 - \frac{0 + \sum_{l=n_B}^{n_B+n_Q} \sum_{i=1}^r \langle X_l, U_i^{(1)} \rangle^2 + n_A(j) \cdot c_H}{c_1 + \sum_{l=n_B}^{n_B+n_Q} \|X_l\|^2 + c_2} = 1 - \frac{n_A(j) \cdot c_H}{\tilde{c}}, \end{aligned} \quad (5.1)$$

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 (5.1) 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).

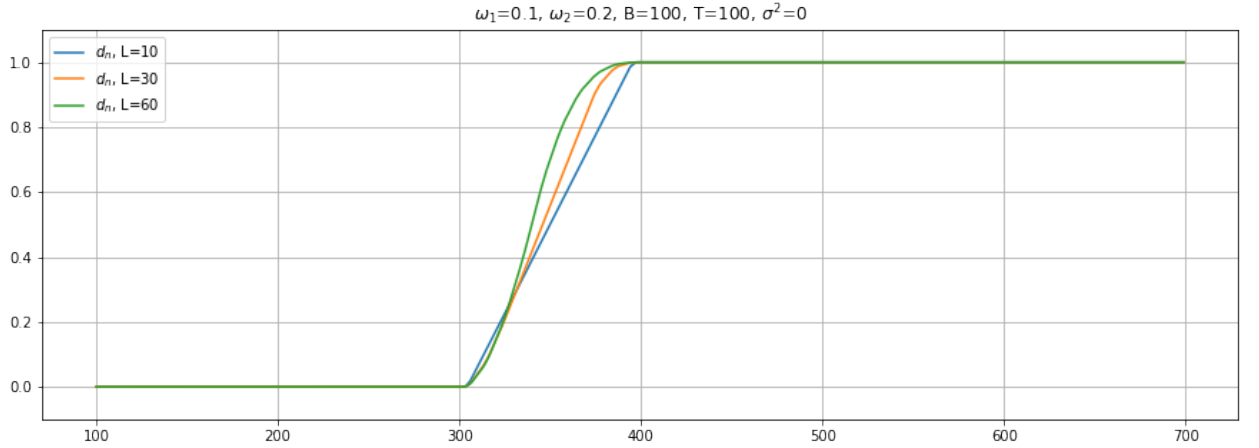


Рис. 5.3. Линейность переходного интервала при разностях $T - L$.

Таким образом, мы можем аппроксимировать переходный интервал кривой d_n от γ_{min} до $g_a(\omega_1, \omega_{min})$ прямой a_T (Рис. 5.4, 5.5).

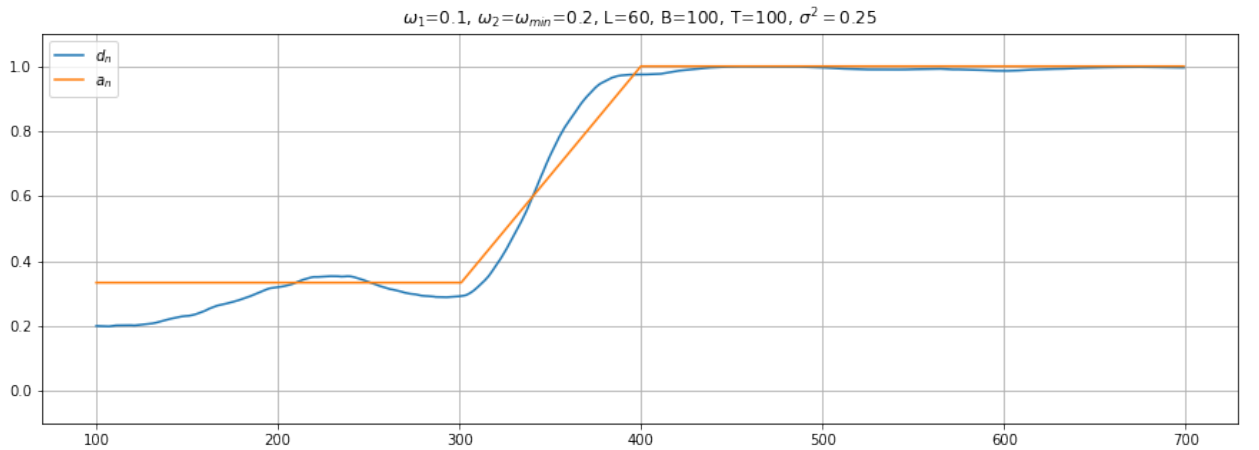


Рис. 5.4. Аппроксимация при отсутствии шума. Кривая a_n с переходным интервалом a_T синхронизирована номерами значений с d_n

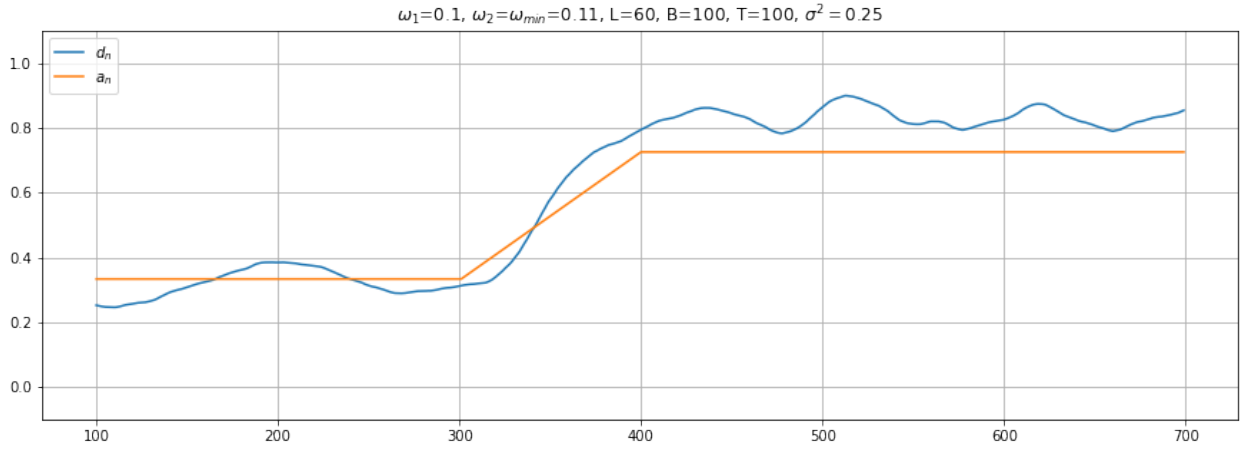


Рис. 5.5. Аппроксимация при наличии шума. Кривая a_n с переходным интервалом a_T синхронизирована номерами значений с d_n .

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 a_T , нам важно чтобы первые k значений a_T были меньше d_n , ведь в противном случае $\hat{Q} > Q + k$. Таким образом на первый план выходит вопрос не о корректности аппроксимации прямой a_T переходного интервала d_n , а 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, описываемая система имеет следующий вид:

- Входные данные: $F_N, \omega_1, \omega_{min}, k, \sigma^2$;
- Результат: \hat{Q} ;
- Алгоритм:
 1. Оцениваем γ_{min} ;
 2. Вычисляем $g_a(\omega_1, \omega_{min})$;
 3. Строим a_T ;
 4. Фиксируем γ .
 5. Определяем \hat{Q} как момент преодоления d_n значения γ .

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 $FNR(\gamma_a)$.

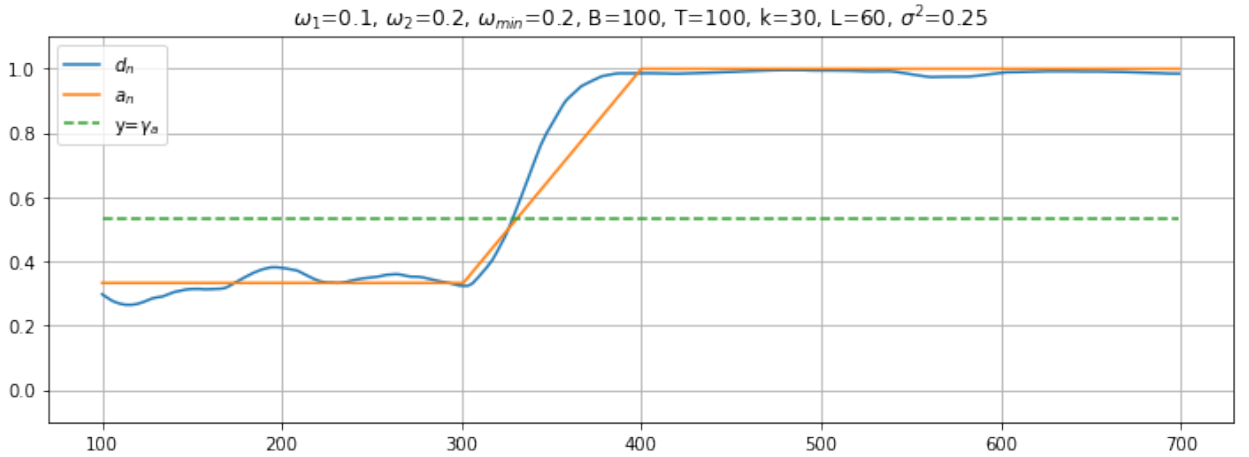


Рис. 5.6. Пример работы системы на одной итерации при $\sigma^2 = 0.25$.

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 $\hat{Q} \in [Q, Q + k]$.

Уменьшая дисперсию 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 a_T раньше, чем в момент k . Об этом нам говорит интервал значений γ на Рис. 5.9, где вероятность точного обнаружения $TPR(\gamma) = 1$.

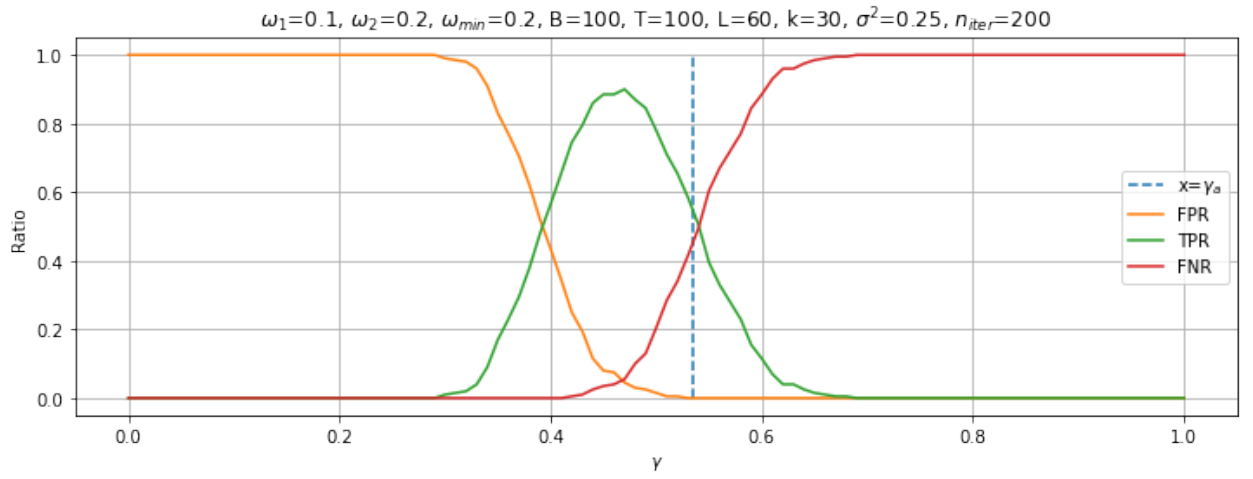


Рис. 5.7. Характеристики системы при $\sigma^2 = 0.25$.

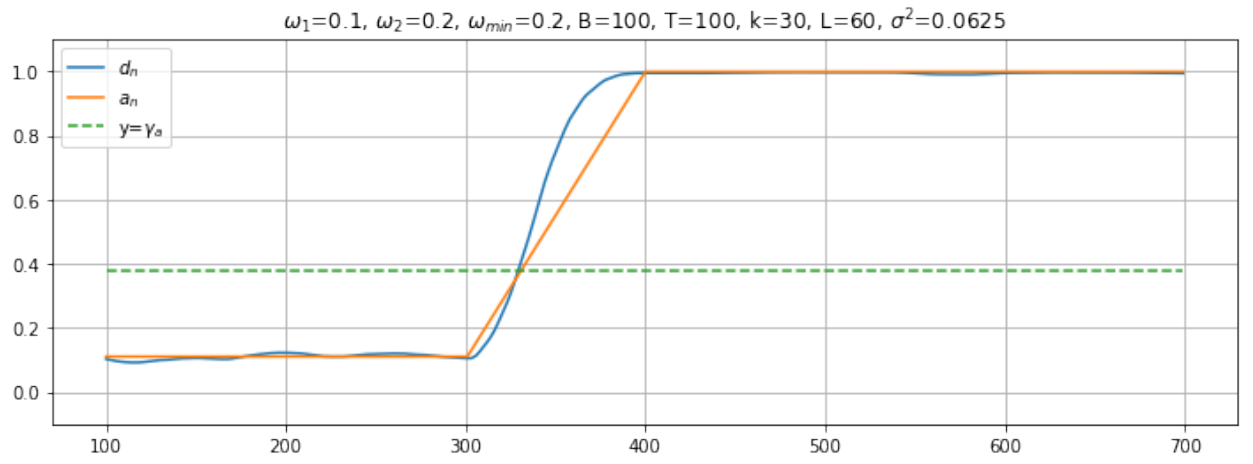


Рис. 5.8. Пример работы системы на одной итерации при $\sigma^2 = 0.0625$.

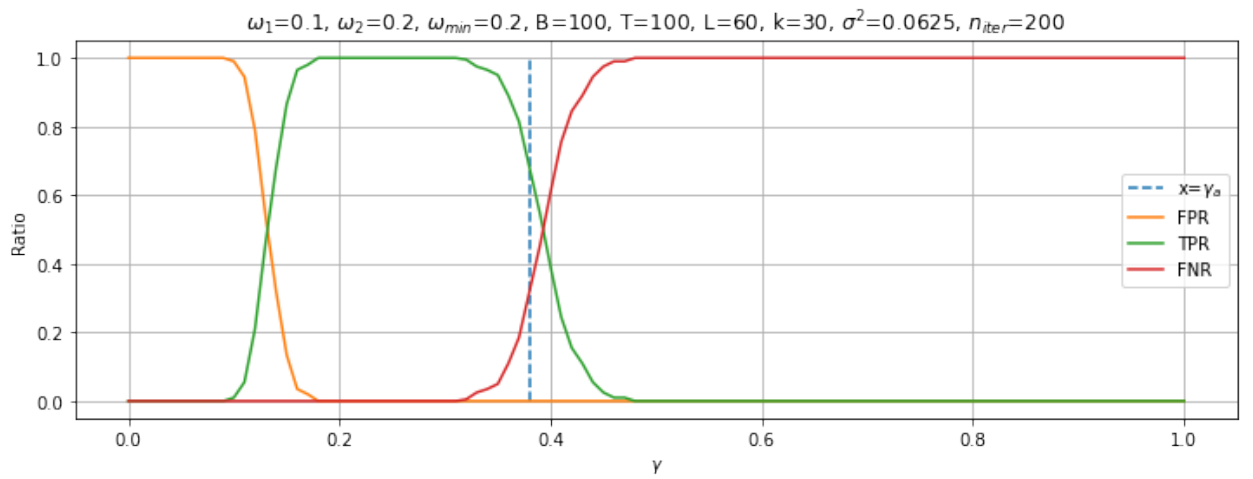


Рис. 5.9. Характеристики системы при $\sigma^2 = 0.0625$.

5.5. Оценка параметров

Все параметры в описываемой системе можно разделить на 4 группы:

1. Входные: $F_N, \omega_1, \omega_{min}, k, \sigma^2$;
2. Зависимые от входных параметров: γ_{min}, γ_a .
3. Неизвестные: ω_2, Q ;
4. Свободные, выбираемые системой: L, B, T .

Первую категорию параметров можно интерпретировать как заданные пользователем системы. Они фиксированы и не могут меняться для определения более хорошего порога. Исключением из этого является параметр k в силу его смысла — обнаружить разладку не позже момента k .

Вторая категория зависит от входных параметров и определяется алгоритмом работы.

Третья категория зависит от ряда, подаваемого системе и вообще говоря не известны.

Четвертая категория параметров - те, которые система может подстраивать под тот или иной ряд. Попробуем оценить их влияние на оценку системы.

5.5.1. Оценка параметров: T

При уменьшении параметра T длина переходного интервала сокращается, что увеличивает наклон прямой a_T (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: $FPR(\gamma_a) = 0$, $TPR(\gamma_a) = 0.99$, $FNR(\gamma_a) = 0.01$.

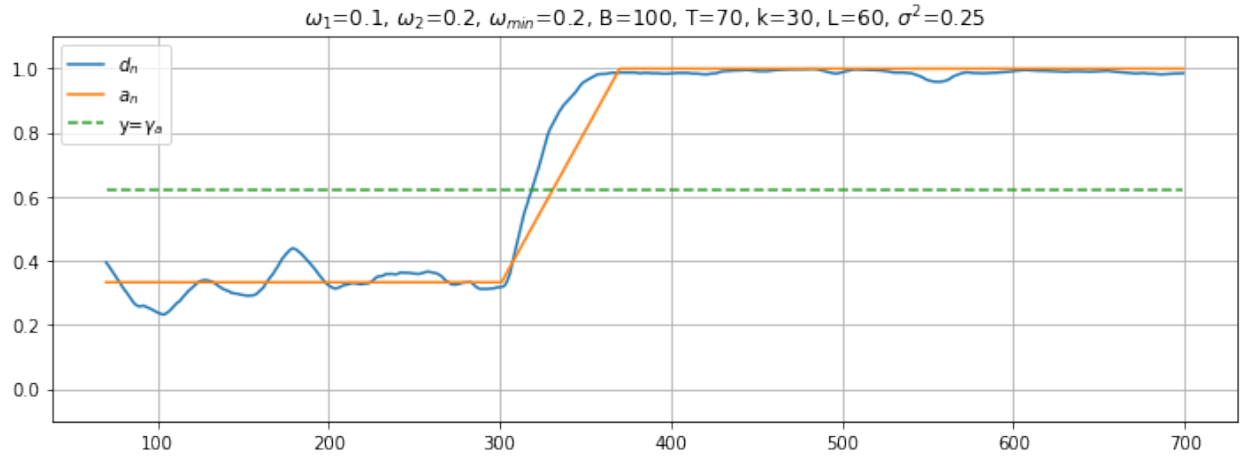


Рис. 5.10. Пример работы системы на одной итерации при $T = 70$.

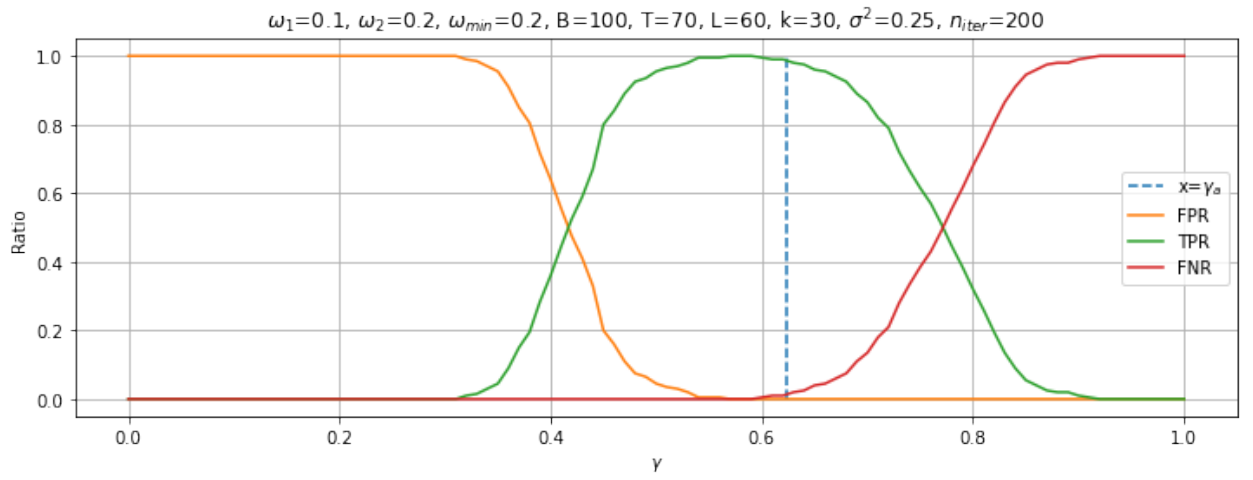


Рис. 5.11. Характеристики системы при $T = 70$.

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.

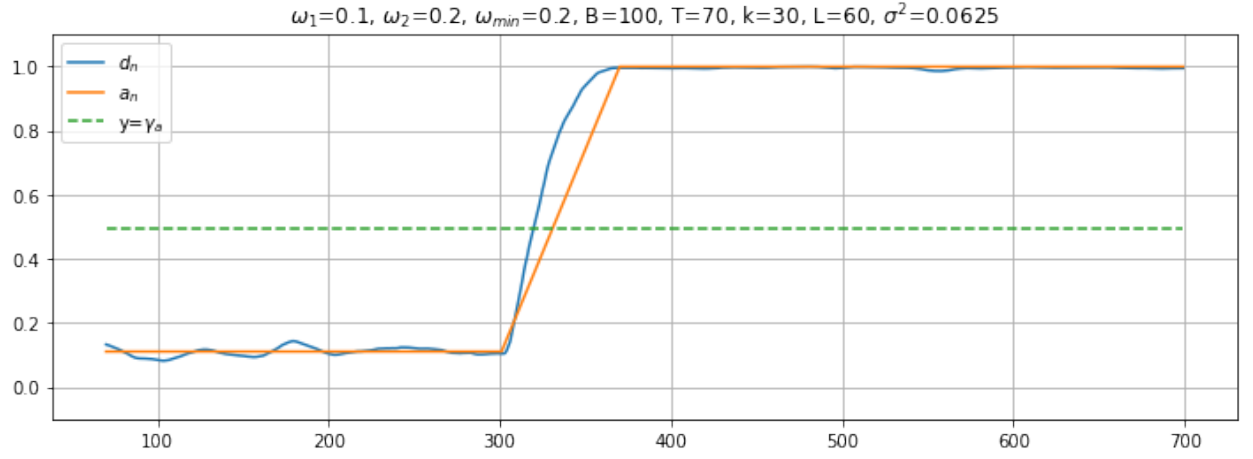


Рис. 5.12. Пример работы системы на одной итерации при $T = 70$, $\sigma^2 = 0.0625$.

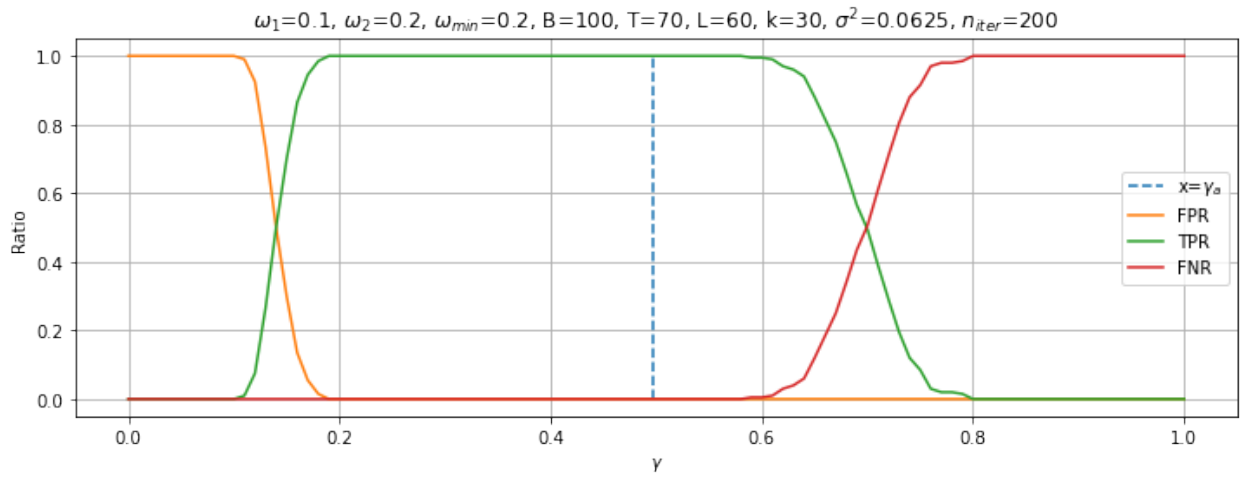


Рис. 5.13. Характеристики системы при $T = 70$, $\sigma^2 = 0.0625$.

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 a_T некорректна и первые значения a_T 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.

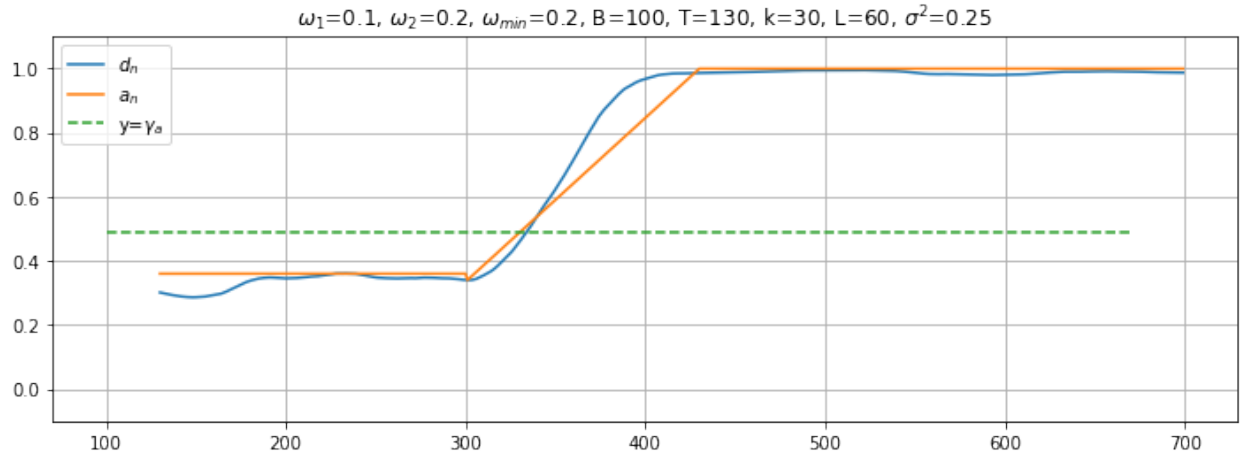


Рис. 5.14. Пример работы системы на одной итерации при $T = 130$.

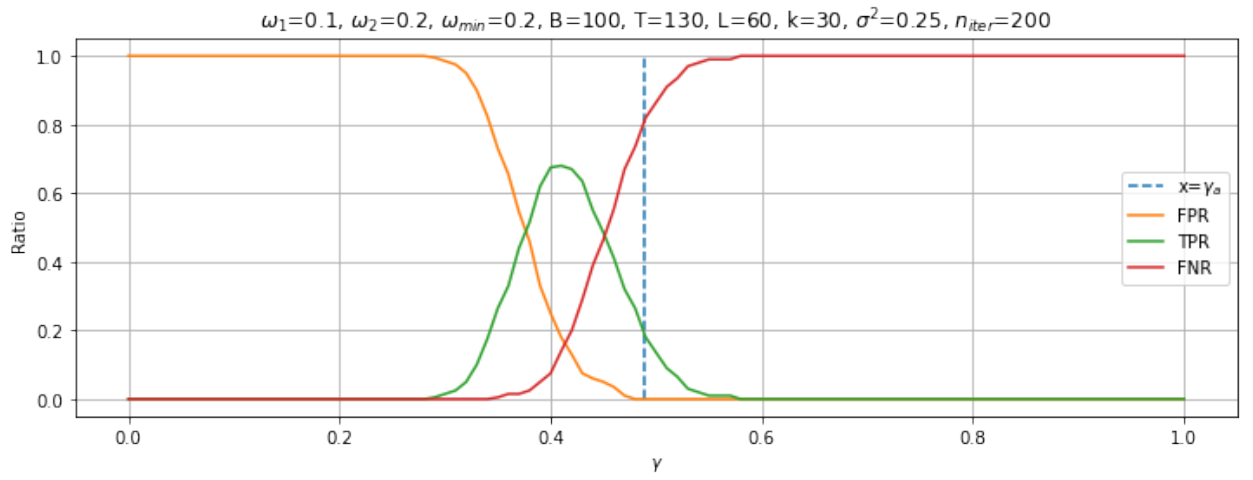
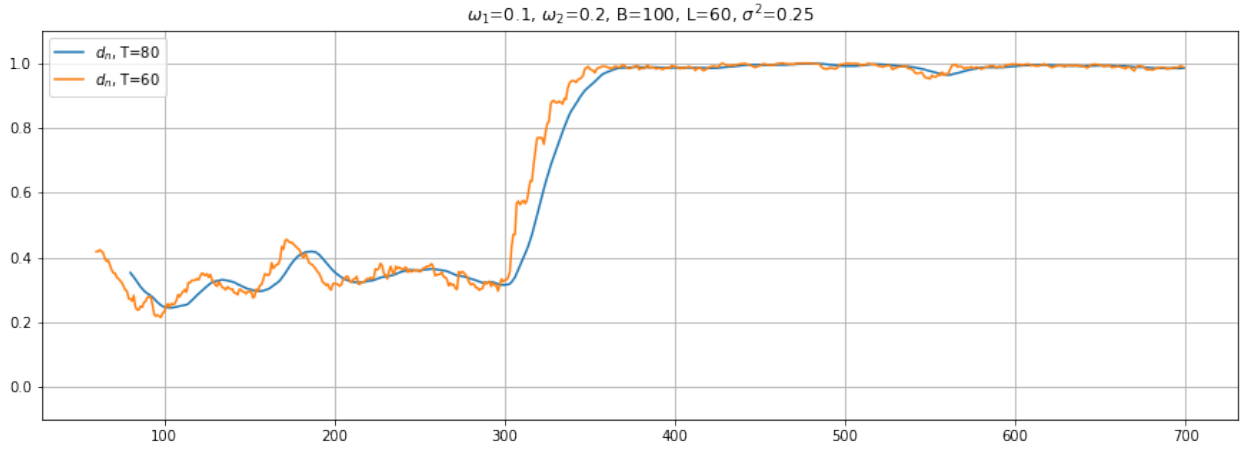


Рис. 5.15. Характеристики системы при $T = 130$.

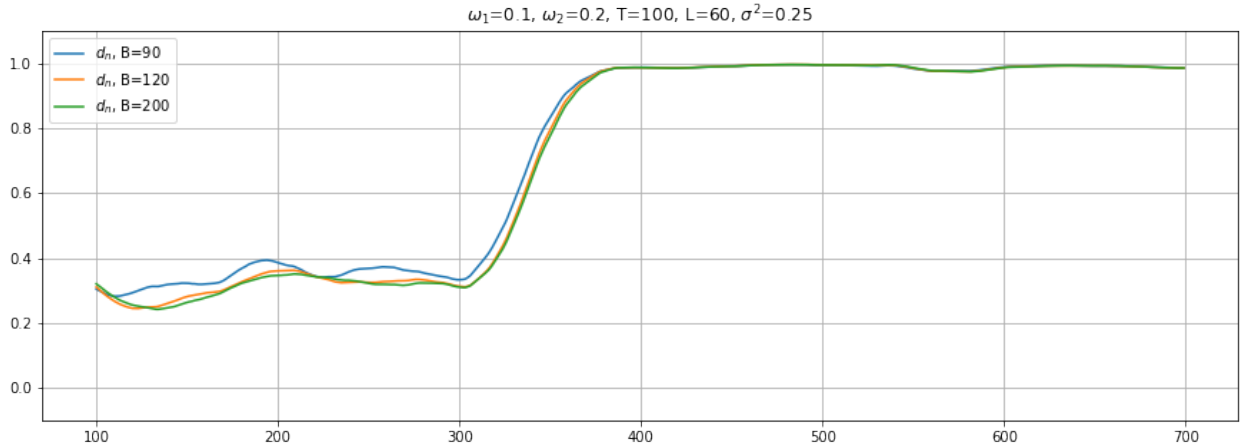
Таким образом, уменьшая T , мы увеличиваем устойчивость системы, а также интервал, меньший размера k , на котором $\text{TPR}(\gamma) = 1$, следствием чего является $\hat{Q} \rightarrow Q$ — более точное обнаружение момента нарушения однородности.

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).

Рис. 5.16. $T \rightarrow L$.

5.5.2. Оценка параметров: B

Параметр B не оказывает значимого влияния на устойчивость системы (Рис. 5.17), так как значения d_n до момента Q мы оцениваем по входному параметру σ^2 .

Рис. 5.17. d_n при разных значениях B .

5.5.3. Оценка параметров: 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, однако, при увеличении L , количество векторов вложений n_Q , содержащих момент возмущения, также возрастет, и при фиксированном T равенство в формуле (5.1) достигаться не будет в силу $n_Q = O(L)$. Линейность переходного интервала нарушается, однако скорость роста, при увеличении L становится

быстрее линейной (Рис. 5.18).

Действительно, чем больше L , тем меньше K_{test} и больше размерность самих векторов вложений. При преодолении момента возмущения, начиная с $j = Q - T$, количество векторов $n_B : \rho(X_l, \mathfrak{L}_\tau) \neq 0, X_l \in \mathbb{X}_{test}^{(j)}$ увеличивается. Отсюда увеличивается вклад и в числитель, и в знаменатель индекса неоднородности g каждого X_l . Поскольку растет норма векторов X_l , увеличивается и их скалярное произведение. С ростом L размер элементов базового пространства увеличивается, что позволяет раньше установить принадлежность вектора вложения базовому пространству \mathfrak{L}_τ . Отсюда увеличивается скорость роста g .

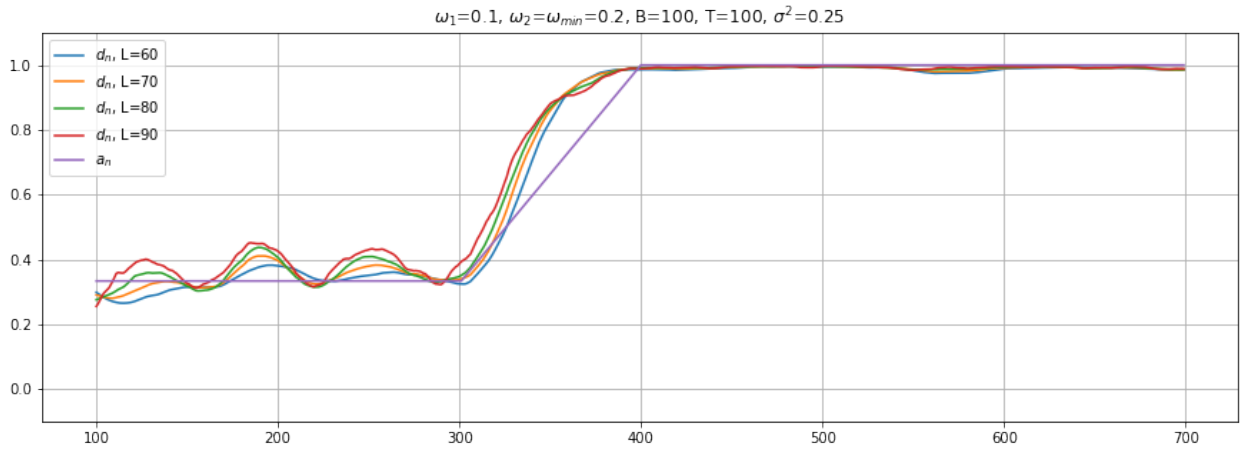


Рис. 5.18. d_n при разных значениях L .

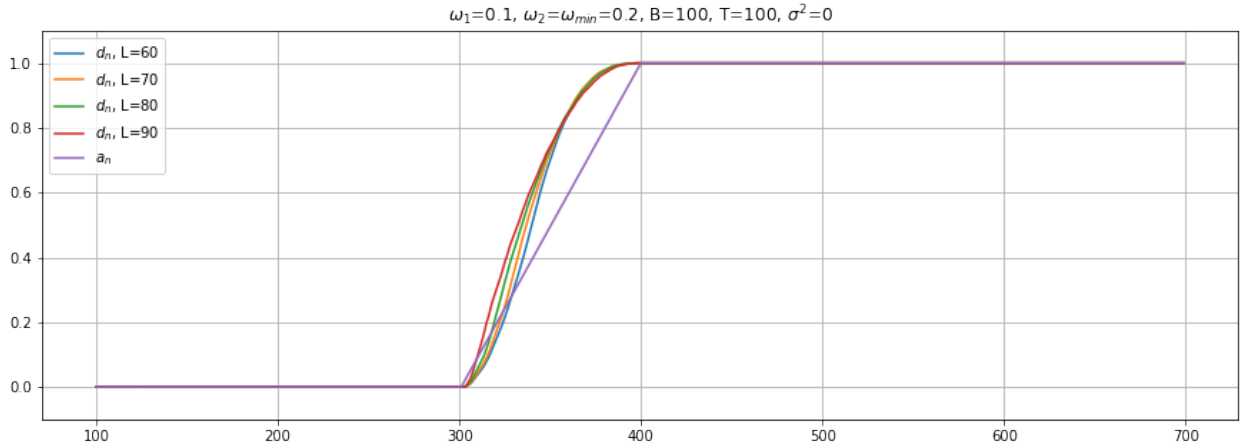


Рис. 5.19. d_n при разных значениях $L, \sigma^2 = 0$.

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 тесты, доказывающие эквивалентность классической формуле.

Была разработана система автоматического обнаружения неоднородности за указанный временной интервал.

Также был проведен анализ поведения строковой функции неоднородности на переходном интервале.

Список литературы

1. Nina Golyandina, Vladimir Nekrutkin, and Anatoly Zhigljavsky. Analysis of time series structure: SSA and related techniques. *Monographs on Statistics and Applied Probability*, 90, January 2001.