«Обнаружение разладки с помощью метода SSA»
Презентация ВКР

Кононыхин Иван Александрович, группа 20.М03-мм

Санкт-Петербургский государственный университет
Математико-механический факультет
Кафедра статистического моделирования

Научный руководитель: к.ф.-м.н., доцент Голяндина Н.Э.
Рецензент: Лектор, Университет Кардиффа (Великобритания), Пепелышев А.Н.

Санкт-Петербург
2022г.

Введение: постановка задачи

Временной ряд **однороден**, если его структура постоянна. При внешнем воздействии ряд терпит возмущение, появляется разладка в его структуре и возникает задача найти момент возмущения.

Задача обнаружения разладки: Определить момент изменения структуры ряда. Структура — подпространство сигнала.

Метод: Превышение порога функцией обнаружения неоднородности, основанной на разнице структур скользящих отрезков ряда.

Временной ряд:

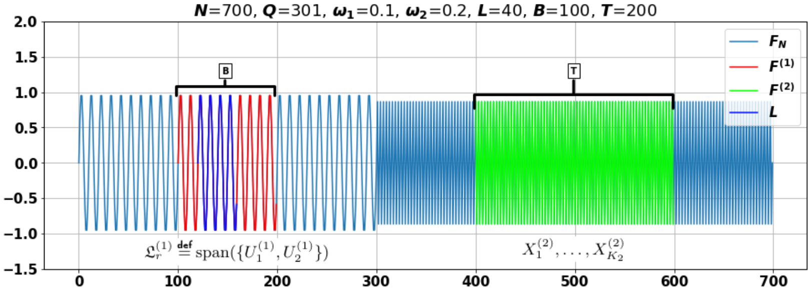
$F_N = (f_1, \ldots, f_N)$, где $f_n = \begin{cases} C_1 \sin(2\pi\omega_1 n + \phi_1), & n < Q, \\ C_2 \sin(2\pi\omega_2 n + \phi_2), & n \geq Q, \end{cases}$

Q — неизвестный момент возмущения.

Цель работы: Создание системы, которая:

1. Определяет разладку, заданную изменением частоты.
2. Автоматически выбирает порог срабатывания.
3. Сообщает о моменте возмущения с заданным значением максимально допустимого запаздывания.

Параметры: L, B, T, $r = 2$.



Индекс неоднородности:

$$g(F^{(1)}; F^{(2)}) = \frac{\sum\limits_{l=1}^{K_2} \mathrm{dist}^2(X_l^{(2)}, \mathfrak{L}_r^{(1)})}{\sum\limits_{l=1}^{K_2} \|X_l^{(2)}\|^2} = 1 - \frac{\sum\limits_{l=1}^{K_2} \sum\limits_{i=1}^{r} \langle X_l^{(2)}, U_i^{(1)} \rangle^2}{\sum\limits_{l=1}^{K_2} \|X_l^{(2)}\|^2}.$$

Матрица неоднородности:

Матрица $\mathbf{G} = \mathbf{G}_{B,T}$, состоящая из элементов g_{ij}:

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

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

Функции обнаружения неоднородности:

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

Постановка задачи: Численно сравнить имеющиеся четыре функций обнаружения неоднородности и выбрать лучшую в смысле момента преодоления заданного порога γ^* и скорости роста для ряда $F_N^{'} = F_N + \epsilon, \epsilon \sim N(0, \sigma^2)$.

Задание неоднородности ряда F_N:

1. Фазовый сдвиг: $\phi_1 \neq \phi_2$.
2. Выброс: $f_n = \begin{cases} C_1 \sin(2\pi\omega_1 n + \phi_1), & n \neq Q, \\ 10 \cdot C_1, & n = Q. \end{cases}$
3. Изменение амплитуды: $C_1 \neq C_2$.
4. Изменение частоты: $\omega_1 \neq \omega_2$.

Сравнение функций обнаружения

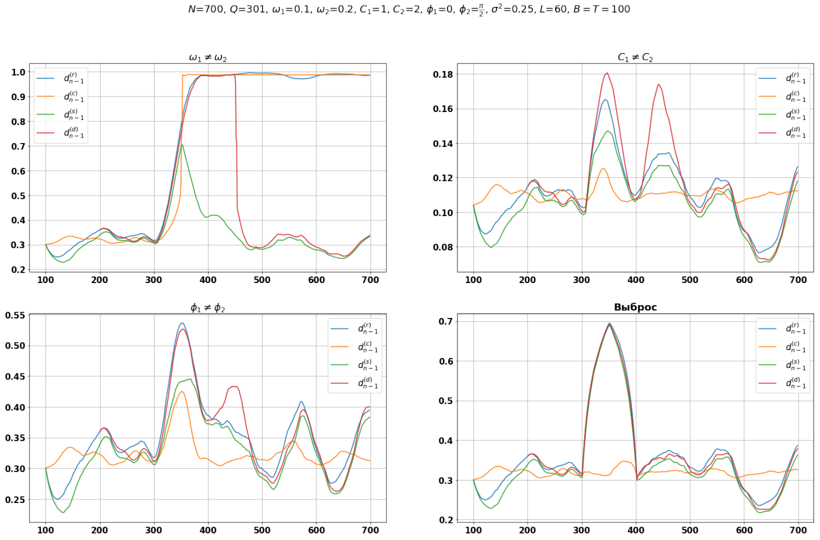


Рис.: Функции обнаружения. Пример на одной реализации шума.

Аппроксимация индекса неоднородности

Рассмотрим ряд F_N с неоднородностью, заданной изменением частоты периодики — $\omega_1 \neq \omega_2$. Пусть $C_1 = C_2 = 1$.

Постановка задачи: Аналитически упростить индекс неоднородности $g(F^{(1)}; F^{(2)})$ для строковой функции обнаружения $d_{n-1}^{(r)}$, чтобы получить в явном виде его зависимость от частот до и после разладки.

Результат:

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

$$\approx 1 - \frac{\left[\left(\frac{\sin(2\pi Lb)}{4\pi b} - \frac{\sin(2\pi La)}{4\pi a} \right)^2 + \left(\frac{\cos(2\pi Lb) - 1}{4\pi b} - \frac{\cos(2\pi La) - 1}{4\pi a} \right)^2 \right]}{\frac{L^2}{4}} = g_a(\omega_1, \omega_2),$$

где $a = \omega_1 + \omega_2$, $b = \omega_1 - \omega_2$.
При $\varepsilon \to 0$ и $L \to \infty, K_2 \to \infty$, $|g(F^{(1)}; F^{(2)}) - g_a(\omega_1, \omega_2)| < \varepsilon$.

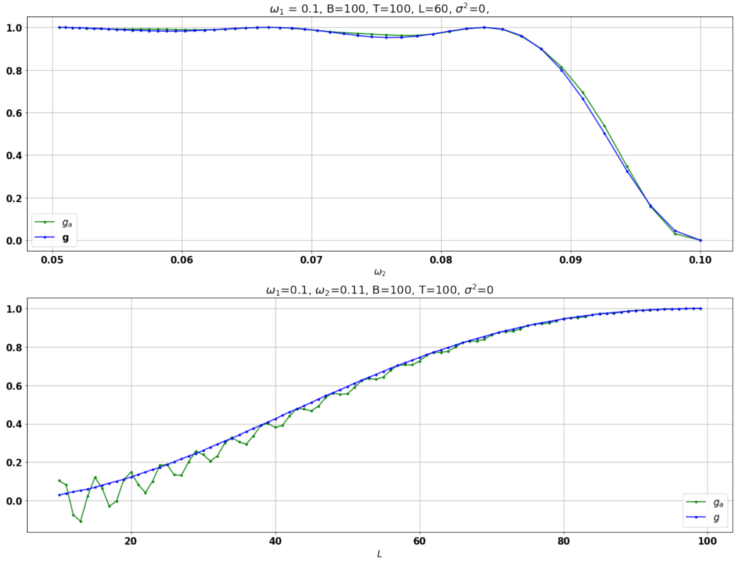


Рис.: Корректность аппроксимации $g(\omega_1, \omega_2)$.

При достаточно маленьком значении L по отношению к T переходный интервал становится линейным.

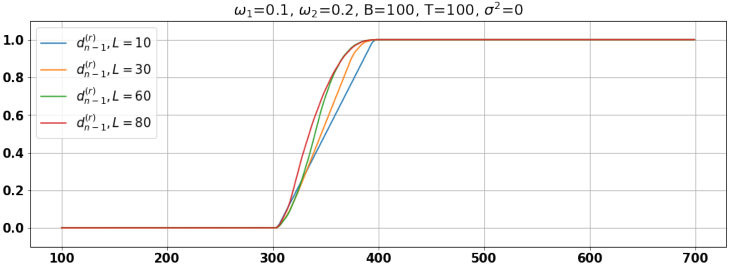


Рис.: Линейность переходного интервала при большом значении $T - L$.

Постановка задачи: Создать систему, способную обнаружить разладку за заданный промежуток времени в синусоидальных временных рядах, порожденную изменением частоты. Сигнал о моменте возмущения — превышение $d_{n-1}^{(r)}$ порога γ^*.

Описание системы:

1. Входные данные: F_N, k, Δ_{min}.
2. Результат: $\hat{Q}$.

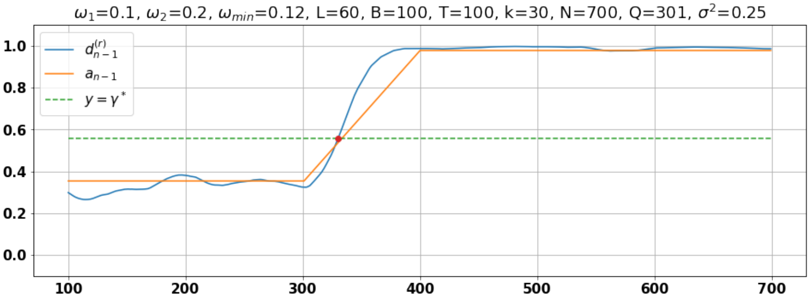


Рис.: Система. Пример работы.

Оценка качества системы

Характеристики системы:

- $\mathrm{FP}(\gamma^*)$ при $\hat{Q} < Q$.
- $\mathrm{TP}(\gamma^*)$ при $\hat{Q} \in [Q, Q+k]$.
- $\mathrm{FN}(\gamma^*)$ при $\hat{Q} > Q+k$.

Промоделируем $n_{iter} = 200$ раз реализацию шума ϵ и на каждой итерации посчитаем характеристики системы.

Вероятности обнаружения:

- $\mathrm{FPR}(\gamma^*) = \dfrac{\sum\limits_{i=1}^{n_{iter}} \mathrm{FP}_i(\gamma^*)}{n_{iter}}$.
- $\mathrm{TPR}(\gamma^*) = \dfrac{\sum\limits_{i=1}^{n_{iter}} \mathrm{TP}_i(\gamma^*)}{n_{iter}}$.
- $\mathrm{FNR}(\gamma^*) = \dfrac{\sum\limits_{i=1}^{n_{iter}} \mathrm{FN}_i(\gamma^*)}{n_{iter}}$.

Оценка системы: параметр T

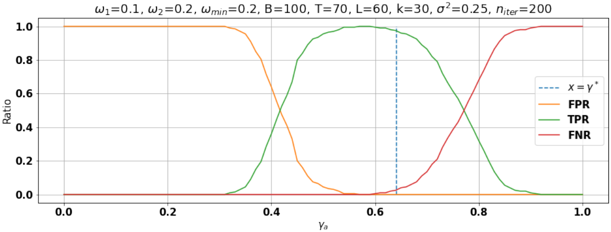


Рис.: Работы системы. Оценка, $T = 70$.

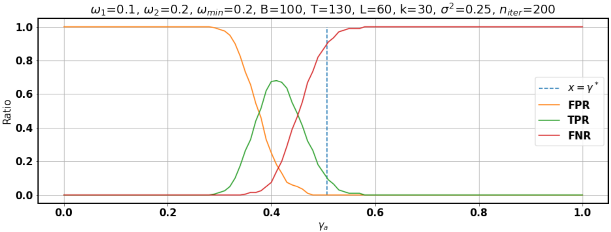


Рис.: Работы системы. Оценка, $T = 130$.

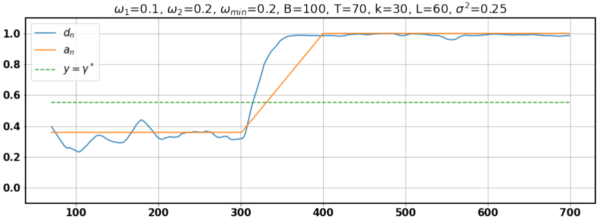


Рис.: Работы системы. Одна итерация, $T - L = 10$.

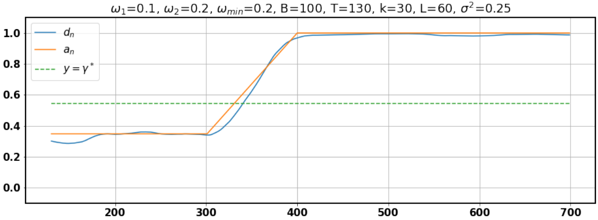


Рис.: Работы системы. Одна итерация, $T - L = 70$.

Оценка системы: параметр B

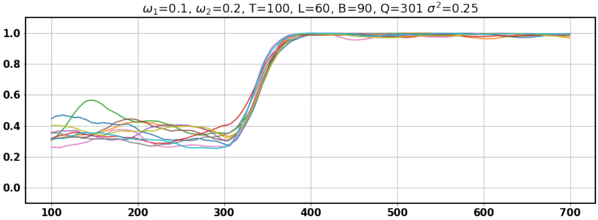


Рис.: Функция d_n. Реализации шума, $B = 90$.

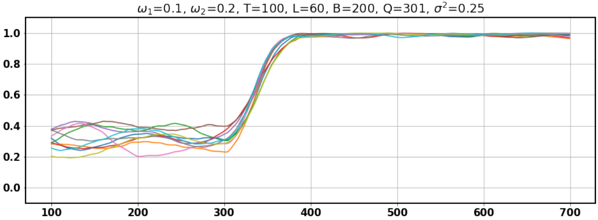


Рис.: Функция d_n. Реализации шума, $B = 200$.

Проблемы

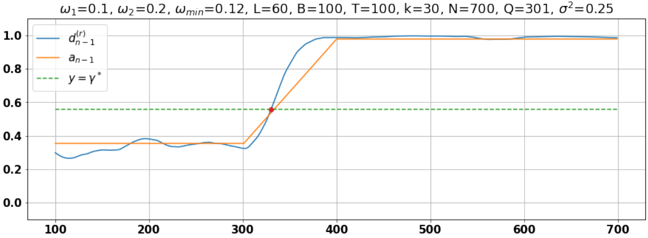


Рис.: Система. Пример работы.

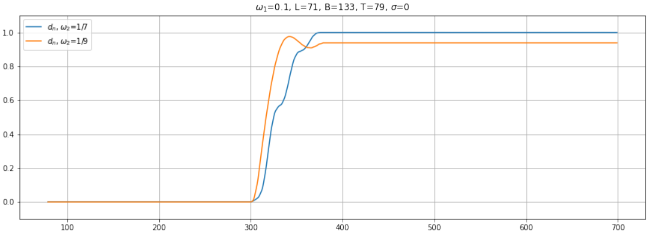


Рис.: Функция d_n. Поведение функции при $\omega_2 = \frac{1}{7}$ и $\omega_2 = \frac{1}{9}$.