

# «Обнаружение разладки с помощью метода SSA»

## Презентация ВКР

Кононыхин Иван Александрович, группа 20.M03-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.

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 заданного порога функцией обнаружения, основанной на скользящих отрезках ряда.

**Временной ряд:**

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

$Q$  — неизвестный момент возмущения.

**Цель работы:** Создать систему, автоматически выбирающую порог срабатывания и способную обнаруживать разладку, заданную изменением частоты в синусоидальном временном ряде  $F_N$  за указанный промежуток времени.

## Параметры:

- $L : 2 \leq L \leq \min(B - 1, T)$  — длина окна,
- $B > L$  — длина базовых подрядов,  $T \geq L$  — длина тестовых подрядов,
- $F^{(1)} = (f_i, \dots, f_{i+B-1}), 1 \leq i \leq N - B + 1$  — базовые подряды,
- $F^{(2)} = (f_i, \dots, f_{i+T-1}), 1 \leq i \leq N - T + 1$  — тестовые подряды,
- $U_l^{(1)}, l = 1, \dots, L$  — собственные векторы траекторной матрицы ряда  $F^{(1)}$ ,
- $I = \{i_1, \dots, i_r\}$  — подмножество  $\{1, \dots, L\}$ ,  $r$  — ранг ряда,
- $\mathfrak{L}_r^{(1)} \stackrel{\text{def}}{=} \text{span}(U_l^{(1)}, l \in I)$ ,
- $X_1^{(2)}, \dots, X_{K_2}^{(2)}$  — векторы вложений длины  $L$  ряда  $F^{(2)}$ .

Индекс неоднородности:

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

Матрица неоднородности:

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

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

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

есть **матрица неоднородности** временного ряда  $F_N$ .

## Строковая:

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

## Столбцовая:

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

## Диагональная:

Ряд  $D_{T+B,N}^{(d)}$  с элементами

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

## Симметричная:

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

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

**Задание неоднородности ряда  $F_N$ :**

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

$N=700$ ,  $Q=301$ ,  $\omega_1=0.1$ ,  $\omega_2=0.2$ ,  $C_1=1$ ,  $C_2=2$ ,  $\phi_1=0$ ,  $\phi_2=\frac{\pi}{2}$ ,  $\sigma^2=0.25$ ,  $L=60$ ,  $B=T=100$

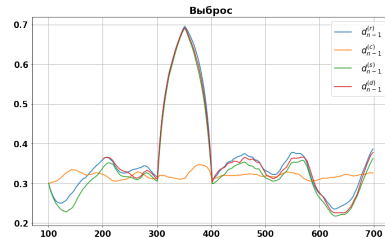
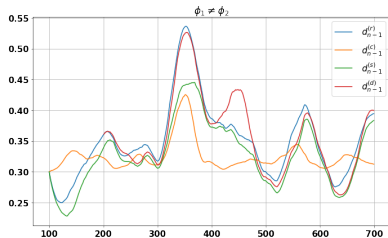
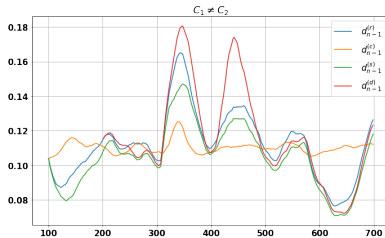
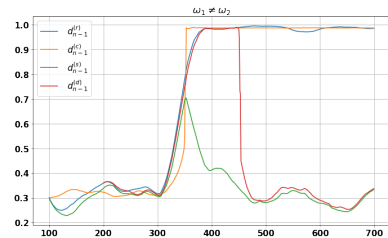


Рис.: Функции обнаружения. Пример на одной реализации шума.

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 явно увидеть, как частоты ряда до и после момента  $Q$  влияют на его значения.

**Результат:**

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

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

При  $\varepsilon \rightarrow 0$  и  $L \rightarrow \infty$ ,  $K_2 \rightarrow \infty$ ,  $|g(F^{(1)}; F^{(2)}) - g_a(\omega_1, \omega_2)| < \varepsilon$ .



# Аппроксимация индекса однородности: точность аппроксимации

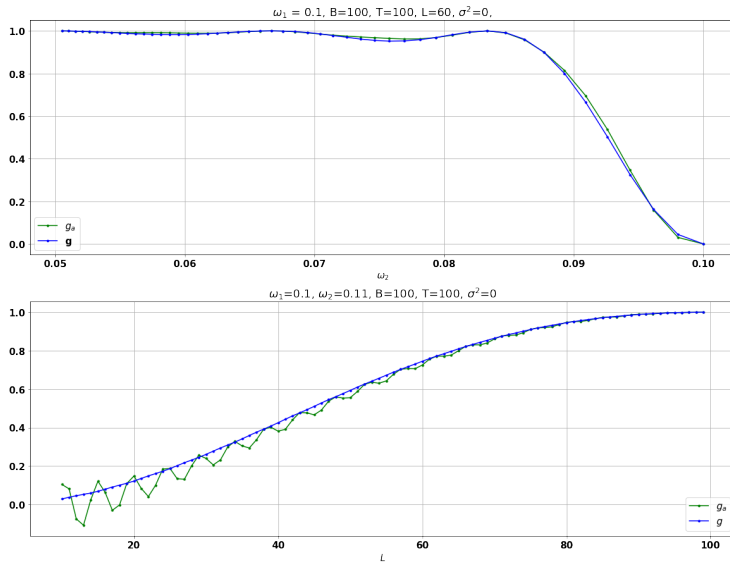


Рис.: Корректность аппроксимации  $g(\omega_1, \omega_2)$ .

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

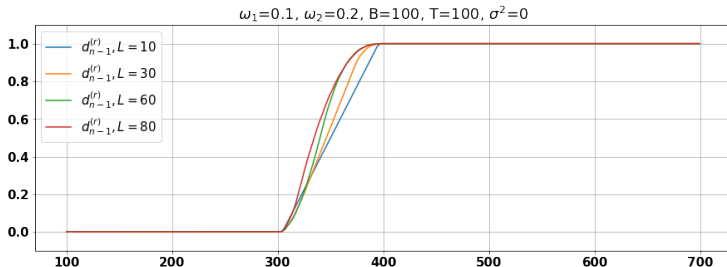


Рис.: Линейность переходного интервала при большом значении  $T - L$ .

**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  $\gamma^*$ .

## Описание системы:

- ❶ Входные данные:  $F_N, k, \Delta_{min}$ .
- ❷ Результат:  $\hat{Q}$ .
- ❸ Алгоритм:
  - ❶ Фиксируем  $B, T, L$ .
  - ❷ Оцениваем  $\omega_1$ .
  - ❸ Оцениваем  $\gamma_{min}$  как 3-й квартиль функции  $d_{n-1}^{(r)}$  на промежутке  $[0; \frac{N}{4} - T]$ .
  - ❹ Вычисляем  $g_a(\omega_1, \omega_{min})$ ,  $\omega_{min} = \omega_1 + \Delta_{min}$ .
  - ❺ Строим прямую  $\gamma$ , соединяющую  $\gamma_{min}$  и  $g_a(\omega_1, \omega_{min})$ ;
  - ❻ Фиксируем  $\gamma^* = \gamma(k)$ ;
  - ❼ Определяем  $\hat{Q}$  как момент преодоления  $d_n$  значения  $\gamma^*$ .

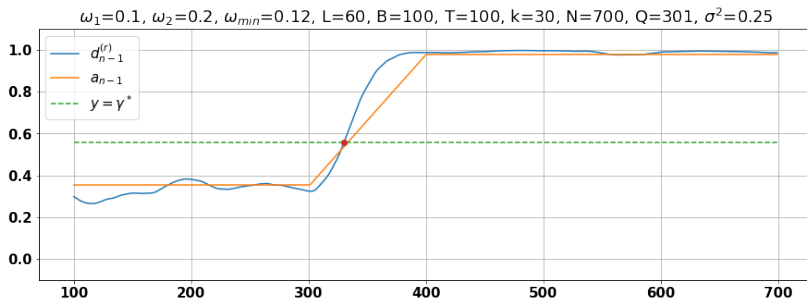


Рис.: Система. Пример работы.

## Характеристики системы:

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

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

## Вероятности обнаружения:

- $FPR(\gamma^*) = \frac{\sum_{i=1}^{n_{iter}} FP_i(\gamma^*)}{n_{iter}}.$
- $TPR(\gamma^*) = \frac{\sum_{i=1}^{n_{iter}} TP_i(\gamma^*)}{n_{iter}}.$
- $FNR(\gamma^*) = \frac{\sum_{i=1}^{n_{iter}} FN_i(\gamma^*)}{n_{iter}}.$

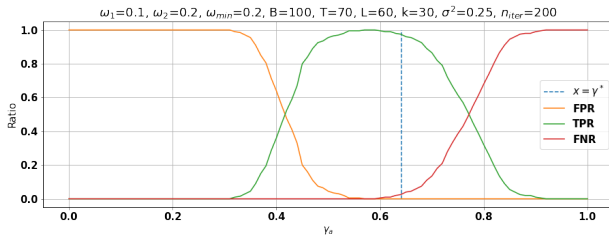


Рис.: Работы системы. Оценка,  $T = 70$ .

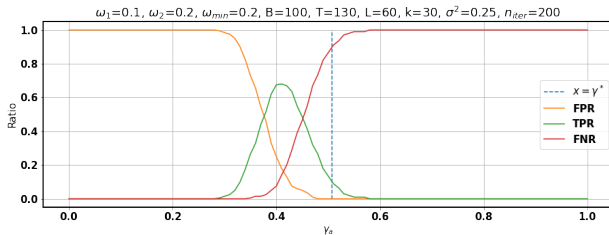


Рис.: Работы системы. Оценка,  $T = 130$ .

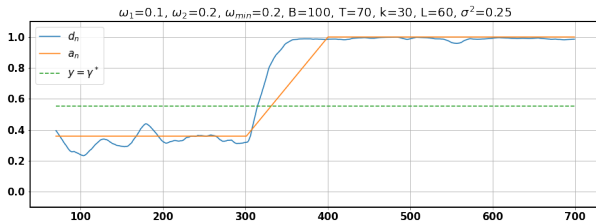


Рис.: Работы системы. Одна итерация,  $T - L = 10$ .

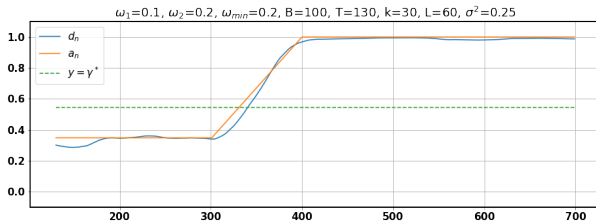


Рис.: Работы системы. Одна итерация,  $T - L = 70$ .

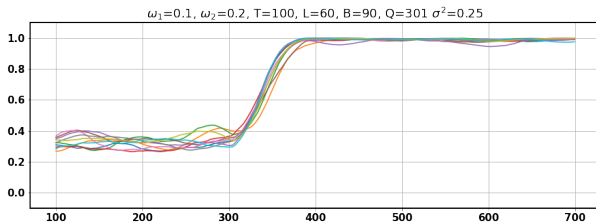


Рис.: Функция  $d_n$ . Реализации шума,  $B = 90$ .

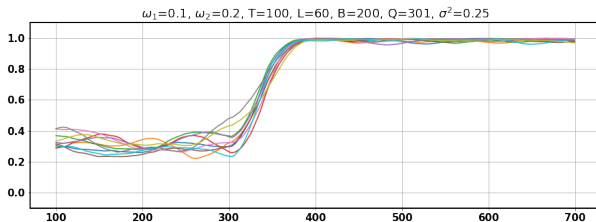


Рис.: Функция  $d_n$ . Реализации шума,  $B = 200$ .



Таблица: Результаты тестирования модели,  $N = 800, Q = 301, k = 15$ .

$\omega_1$	$\omega_2$	$\Delta_{min}$	$\sigma$	$B$	$T$	$L$	$FPR$	$TPR$	$FNR$
1/10	1/3	0.02	0.2	133	79	71	0.000	1.0	0.000
1/10	1/3	0.02	0.3	133	79	71	0.000	0.995	0.005
1/10	1/3	0.02	0.4	133	79	71	0.000	0.895	0.105
1/10	1/3	0.02	0.5	133	79	71	0.040	0.745	0.215
1/10	1/4	0.02	0.2	133	79	71	0.000	1.0	0.000
1/10	1/4	0.02	0.3	133	79	71	0.000	0.980	0.020
1/10	1/4	0.02	0.4	133	79	71	0.000	0.870	0.130
1/10	1/4	0.02	0.5	133	79	71	0.040	0.745	0.215
1/10	1/5	0.02	0.2	133	79	71	0.000	1.0	0.000
1/10	1/5	0.02	0.3	133	79	71	0.000	0.980	0.020
1/10	1/5	0.02	0.4	133	79	71	0.000	0.855	0.145
1/10	1/5	0.02	0.5	133	79	71	0.040	0.720	0.240
1/10	1/6	0.02	0.2	133	79	71	0.000	1.0	0.000
1/10	1/6	0.02	0.3	133	79	71	0.000	0.995	0.005
1/10	1/6	0.02	0.4	133	79	71	0.000	0.925	0.075
1/10	1/6	0.02	0.5	133	79	71	0.040	0.820	0.140

Таблица: Результаты тестирования модели,  $N = 800, Q = 301, k = 15$ .

$\omega_1$	$\omega_2$	$\Delta_{min}$	$\sigma$	$B$	$T$	$L$	$FPR$	$TPR$	$FNR$
1/10	1/7	0.02	0.0	133	79	71	0.000	0.000	1.000
1/10	1/7	0.02	0.1	133	79	71	0.000	0.250	0.750
1/10	1/7	0.02	0.2	133	79	71	0.000	0.335	0.665
1/10	1/7	0.02	0.3	133	79	71	0.000	0.325	0.675
1/10	1/7	0.02	0.4	133	79	71	0.000	0.380	0.620
1/10	1/7	0.02	0.5	133	79	71	0.040	0.340	0.660
1/10	1/8	0.02	0.2	133	79	71	0.000	1.0	0.000
1/10	1/8	0.02	0.3	133	79	71	0.000	1.0	0.000
1/10	1/8	0.02	0.4	133	79	71	0.000	0.995	0.005
1/10	1/8	0.02	0.5	133	79	71	0.040	0.920	0.040
1/10	1/9	0.02	0.4	133	79	71	0.000	1.0	0.000
1/10	1/9	0.02	0.5	133	79	71	0.050	0.950	0.000

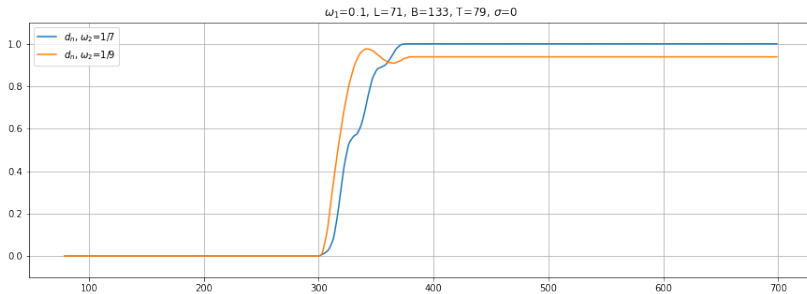


Рис.: Функция  $d_n$ . Поведение функции при  $\omega_2 = \frac{1}{7}$  и  $\omega_2 = \frac{1}{9}$ .