Преобразование Гильберта-Хуанга для анализа временных рядов

Сенов Александр Алексеевич, гр. 522

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

Научный руководитель: к.ф.-м.н., доц. Голяндина Н.Э.
Рецензент: к.ф.-м.н., асс . Коробейников А.И.

Санкт-Петербург
2012г.

- "Простая" модель

$$x(t) = a\cos(\omega t)$$

$$x(t) = \sum_{k=1}^{K} a_k \cos(\omega_k t)$$

- "Сложная" модель

$$x(t) = a(t)\cos(\phi(t))$$

$$x(t) = \sum_{k=1}^{K} a_k(t)\cos(\phi_k(t))$$

Введение. Постановка задачи

- “Простая” модель

$$x(t) = a\cos(\omega t)$$

$$x(t) = \sum_{k=1}^{K} a_k \cos(\omega_k t)$$

- “Сложная” модель

$$x(t) = a(t)\cos(\phi(t))$$

$$x(t) = \sum_{k=1}^{K} a_k(t) \cos(\phi_k(t))$$

Введение. Постановка задачи

- "Простая" модель

$$x(t) = a\cos(\omega t)$$

$$x(t) = \sum_{k=1}^{K} a_k \cos(\omega_k t)$$

- "Сложная" модель

$$x(t) = a(t)\cos(\phi(t))$$

$$x(t) = \sum_{k=1}^{K} a_k(t) \cos(\phi_k(t))$$

Введение. Постановка задачи

- “Простая” модель

$$x(t) = a\cos(\omega t)$$

$$x(t) = \sum_{k=1}^{K} a_k \cos(\omega_k t)$$

- “Сложная” модель

$$x(t) = a(t)\cos(\phi(t))$$

$$x(t) = \sum_{k=1}^{K} a_k(t) \cos(\phi_k(t))$$

Введение. Две подзадачи

1. Разделение

$$x(t) \to \{x_k(t)\}_{k=1}^{K} \quad : \quad \sum x_k(t) = x(t),$$

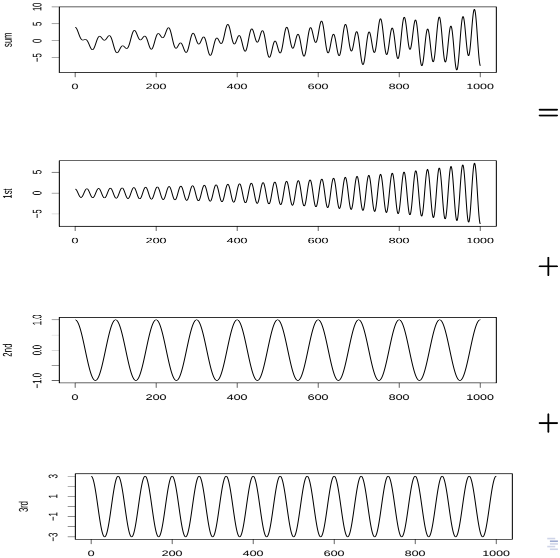
2. Оценка параметров

$$x_k(t) \to a_k(t) \cos(\phi_k(t)), \quad 1 \le k \le K.$$

HHT — Hilbert-Huang Transform (Huang, et. al., 1996) — это

1. Разделение: EMD — Emperical Mode Decomposition
2. Оценка параметров: HT — Hilbert Transform

EMD. Исходные компоненты



Отсеивание первой компоненты

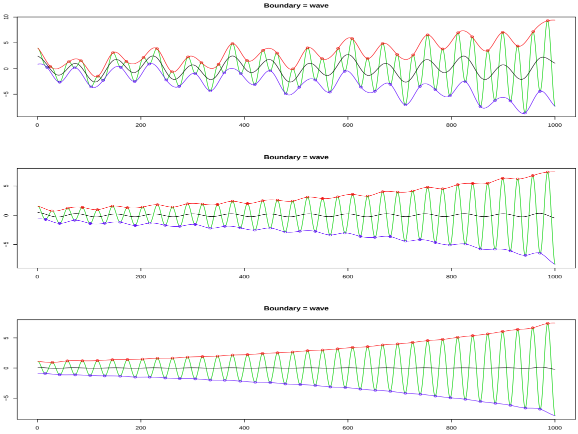


Иллюстрация алгоритма EMD

Отсеивание второй компоненты

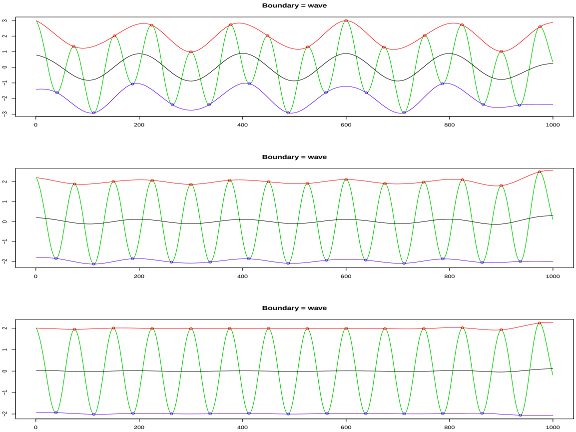
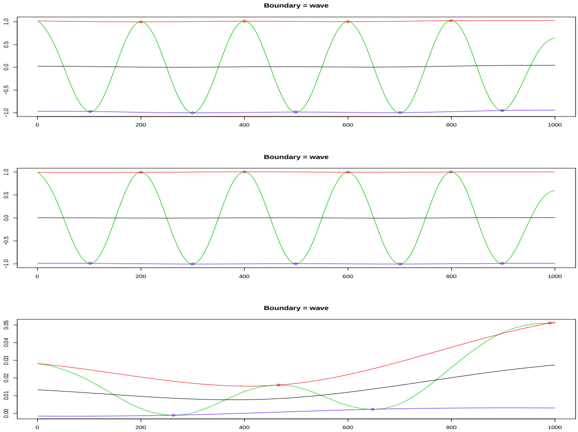
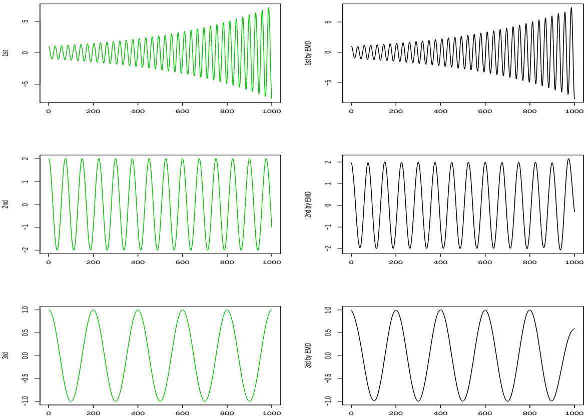


Иллюстрация алгоритма EMD

Отсеивание третьей компоненты



Результаты работы алгоритма

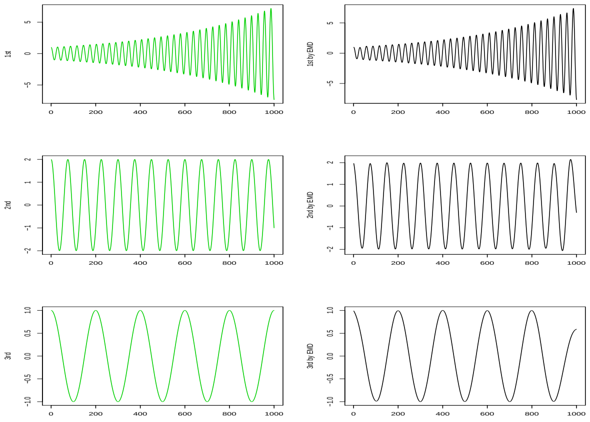


Алгоритм EMD

1. $i = 0,\ h_i(t) = x(t)$
2. $i = i + 1,\ j = 0,\ h_{i,j}(t) = h_{i-1}(t),\ r_i(t) = 0 \forall t$
3. T_{min}, T_{max} — локальные минимумы и максимумы $h_{i,j}(t)$. Если $\min(|T_{min}|, |T_{max}|) < 2$, то алгоритм останавливается.
$$\min(t) = S(T_{min}), \quad \max(t) = S(T_{max})$$
$$m(t) = \frac{\min(t) + \max(t)}{2} \qquad h_{i,j+1}(t) = h_{i,j}(t) - m(t).$$
4. Если
$$\mathrm{SD} = \frac{||h_{i,j+1}(t) - h_{i,j}(t)||_{L_2}}{||h_{i,j}(t)||_{L_2}} < \alpha, \tag{1}$$
то $h_i(t) = h_{i,j+1}(t)$ и возвращаемся к пункту 2 .
Иначе, $j = j + 1$ и возвращаемся к пункту 3 .

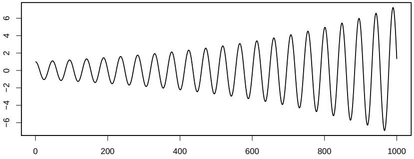
Особенности EMD

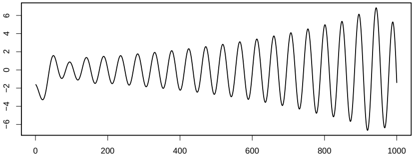
Сперва высокочастотные, затем низкочастотные



Особенности EMD

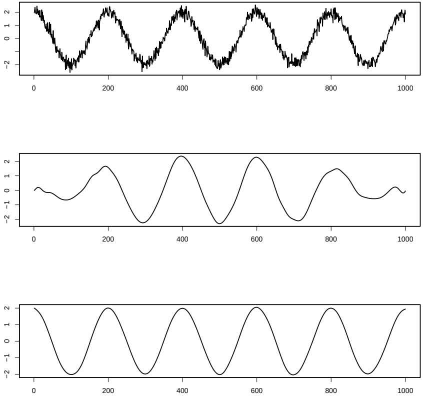
Большие ошибки на концах





Особенности EMD

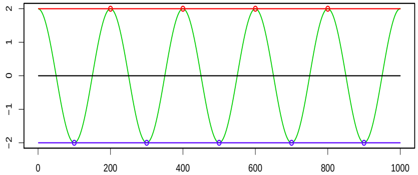
Неустойчивость к шуму. Пример разложения



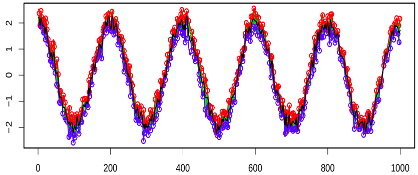
Особенности EMD

Неустойчивость к шуму. Пример построения огибающих

Boundary = periodic



Boundary = periodic



Определение мгновенных характеристик

Определение

Преобразование Гильберта функции $x(t) \in L_2(\mathbb{R})$ *— это*

$$H[x](t) = \frac{1}{\pi} v.p. \int_{\mathbb{R}} \frac{x(\eta)}{t-\eta} d\eta.$$

Определение

Мгновенаая амплитуда и фаза функции $x(t) \in L_2(\mathbb{R})$

$$a(t) = \sqrt{x^2(t) + H[x]^2(t)} \qquad \theta(t) = \arctan\left(\frac{H[x](t)}{x(t)}\right).$$

А почему именно Гильберт?

- Пусть $x(t) = \rho(t)\cos(\phi(t))$, где $\rho(t) \geq 0$.
- Когда $a(t) = \rho(t)$, $\theta(t) = \phi(t)$?
- Если $H[x](t) = \rho(t)\sin(\phi(t))$, то

$$a(t) = \rho(t) \qquad \theta(t) = \phi(t) \mod 2\pi.$$

- А когда $H[a(\cdot)\cos(\phi(\cdot))](t) = \rho(t)\sin(\phi(t))$?

Теоремы. Точное восстановление

Теорема (Nuttall, A., 1966)

Пусть $f(t) = \rho(t)\exp(i\varphi(t)) \in L_2(\mathbb{R})$, $\omega_0 > 0$ *и*

$$\mathcal{F}[f](\omega) = 0 \qquad \forall \omega < -\omega_0.$$

Тогда выполняется

$$||H[\rho(\cdot)\cos(\omega_0 \cdot + \varphi(\cdot))](t) - \rho(t)\sin(\omega_0 t + \varphi(t))||_{L_2}.$$

Теорема (Bedrosian, E., 1962)

Пусть функции $f, g \in L_2(\mathbb{R})$ *и существует такое* $a > 0$*:*

$$\mathcal{F}[f](\omega) = 0 \quad \forall |\omega| > a, \qquad \mathcal{F}[g](\omega) = 0 \quad \forall |\omega| < a.$$

Тогда выполняется $H[fg](t) = f(t)H(g)(t)$.

Следствие. Приближение.

Следствие

Пусть $x(t) = \rho(t)\cos(\omega_0 t + M\sin(\omega t))$, *где* $\rho(t) \in L_2(\mathbb{R})$,

$$\mu = \inf\{v : \mathcal{F}[\rho](\omega) = 0 \quad \forall \omega > v\}$$

и $r_0 = \lceil \frac{\omega_0}{\omega+\mu} \rceil$. *Обозначим*

$$x_Q(t) = \rho(t)\sin(\omega_0 t + M\sin(\omega t))$$

Тогда

$$\int\limits_{\mathbb{R}} |H[x](t) - x_Q(t)|^2 dt \quad \leq \quad 4\pi||\rho||^2_{L_2} \sum_{k=r_0}^{k=\infty} J^2_{2k}(M) \to_{r_0\to\infty} 0.$$

Алгоритм SSA

SSA (Singular Spectrum Analysis, "Гусеница") — метод анализа и прогноза временных рядов. Шаги алгоритма для исходного ряда $f = (f_0, \dots, f_{N-1})$

1. *Вложение*. Построение траекторной матрицы $\mathbf{X} = (x_{i,j})_{i=1,j=1}^{L,K} : x_{i,j} = f_{i-1+(j-1)L}$ (ганкелизация).
2. *Сингулярное разложение*. SVD-разложение траекторной матрицы $\mathbf{X} = \sum_{i=1}^{\min(L,K)} \sqrt{\lambda_i} U_i V_i^T = \sum_{i=1}^{\min(L,K)} \mathbf{X_i}$, где $\lambda_{\min(L,K)} \geq \lambda_1 \leq 0$.
3. *Группировка*. $\mathbf{X} = \sum_{m=1}^{M} \mathbf{X_{I_m}}$, где $I_m{}_{m=1}^{M}$ — разбиение $\{1, \dots, \min(L, K)\}$.
4. *Диагональное усреднение*. "Деганкелизация": проекция $\mathbf{X_{I_m}}$ на пространство ганкелевых матриц и составление соответствующего ряда.

Важный параметр алгоритма $L = \frac{N}{2}$ как наиболее благоприятный в отсутствии информации о ряде f.

Как сравнивать

- Берем $x^{(k)} = (x_0^{(k)}, \dots, x_{N-1}^{(k)})$ $k \in \{1, \dots, K\}$

$$x = \sum_{k=1}^{K} x^{(k)}.$$

- Применяем EMD и SSA

$$\mathrm{EMD} : x \to \{\mathrm{EMD}(x)^k\}_{k=1}^{K_1} \qquad \mathrm{SSA} : x \to \{\mathrm{SSA}(x)^k\}_{k=1}^{K_2}$$

- $\mathrm{Diff}(x^{(k)}, \mathrm{EMD}(x)^{(k)}) \quad vs \quad \mathrm{Diff}(x^{(k)}, \mathrm{SSA}(x)^{(k)})$

$$\mathrm{Diff}(x, y) = \frac{\sum_{n=N_1}^{N_2} (x_n - y_n)^2}{\sum_{n=N_1}^{N_2} x_n^2}.$$

$$N_1 = \lceil 0.1N \rceil \qquad \lfloor N_2 = 0.9N \rfloor$$

Как сравнивать

- Берем $x^{(k)} = (x_0^{(k)}, \ldots, x_{N-1}^{(k)})$ $k \in \{1, \ldots, K\}$

$$x = \sum_{k=1}^{K} x^{(k)}.$$

- Применяем EMD и SSA

$$\mathrm{EMD} : x \to \{\mathrm{EMD}(x)^k\}_{k=1}^{K_1} \qquad \mathrm{SSA} : x \to \{\mathrm{SSA}(x)^k\}_{k=1}^{K_2}$$

- $\mathrm{Diff}(x^{(k)}, \mathrm{EMD}(x)^{(k)}) \quad vs \quad \mathrm{Diff}(x^{(k)}, \mathrm{SSA}(x)^{(k)})$

$$\mathrm{Diff}(x, y) = \frac{\sum_{n=N_1}^{N_2} (x_n - y_n)^2}{\sum_{n=N_1}^{N_2} x_n^2}.$$

$$N_1 = \lceil 0.1N \rceil \qquad \lfloor N_2 = 0.9N \rfloor$$

Как сравнивать

- Берем $x^{(k)} = (x_0^{(k)}, \ldots, x_{N-1}^{(k)})\ k \in \{1, \ldots, K\}$

$$x = \sum_{k=1}^{K} x^{(k)}.$$

- Применяем EMD и SSA

$$\mathrm{EMD} : x \to \{\mathrm{EMD}(x)^k\}_{k=1}^{K_1} \qquad \mathrm{SSA} : x \to \{\mathrm{SSA}(x)^k\}_{k=1}^{K_2}$$

- $\mathrm{Diff}(x^{(k)}, \mathrm{EMD}(x)^{(k)}) \quad vs \quad \mathrm{Diff}(x^{(k)}, \mathrm{SSA}(x)^{(k)})$

$$\mathrm{Diff}(x, y) = \frac{\sum_{n=N_1}^{N_2} (x_n - y_n)^2}{\sum_{n=N_1}^{N_2} x_n^2}.$$

$$N_1 = \lceil 0.1N \rceil \qquad \lfloor N_2 = 0.9N \rfloor$$

Сумма косинусов с постоянной амплитудой

$$N = 1000, \quad M = 50, \quad \alpha = 0.05 \quad \beta = 0.15$$

$$\omega_1, \omega_2 \in \Omega = \{\alpha + k\frac{\beta - \alpha}{M}\}_{k=0}^{M-1}$$

Рассмотрим ряд

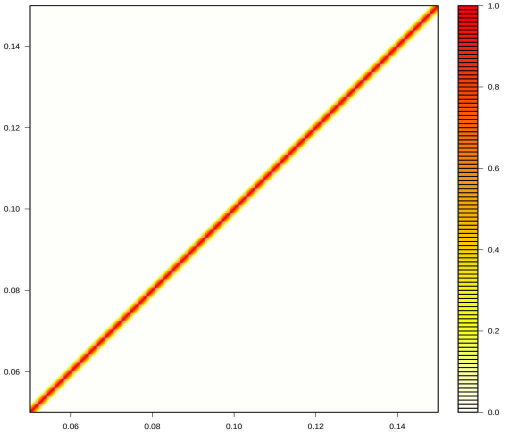
$$x_n^{(1)} = 4\cos(2\pi\omega_1 n), \qquad x_n^{(2)} = \cos(2\pi\omega_2 n).$$

Сравним

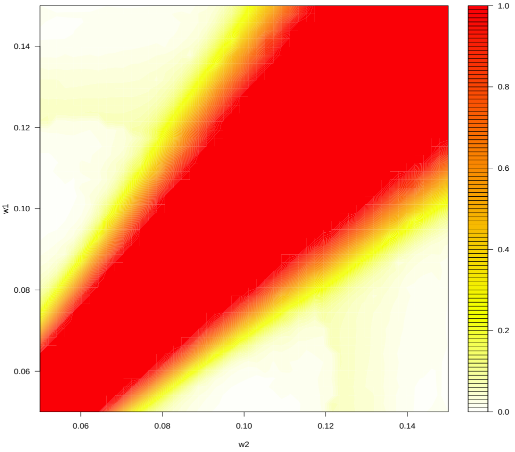
$$\mathrm{Diff}(x^{(k)}, \mathrm{EMD}[x]^{(k)})(\omega_1, \omega_2) \qquad \mathrm{Diff}(x^{(k)}, \mathrm{SSA}[x]^{(k)})(\omega_1, \omega_2)$$

при $k = 1, 2$.

Восстановление двух синусов. Ошибка SSA



Восстановление двух синусов. Ошибка EMD



Сумма синусов — одна компонента?

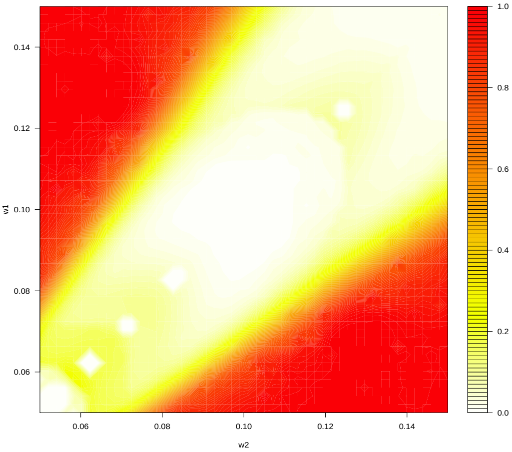


Figure: График $\mathrm{Diff}(x, \mathrm{EMD}[x]^{(1)})$.

EMD не виноват

Оказывается,

$$a_1 \cos(2\pi\omega_1 t) + a_2 \cos(2\pi\omega_2 t) = a(t)\cos\left(2\pi\omega_0 t + \varphi(t)\right),$$

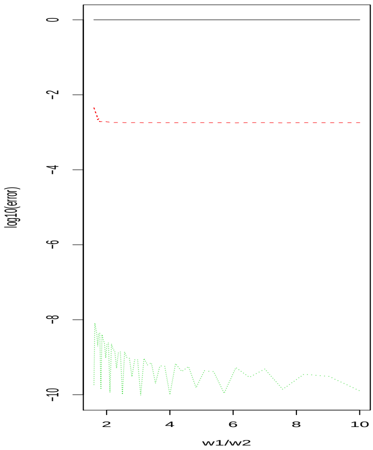
где

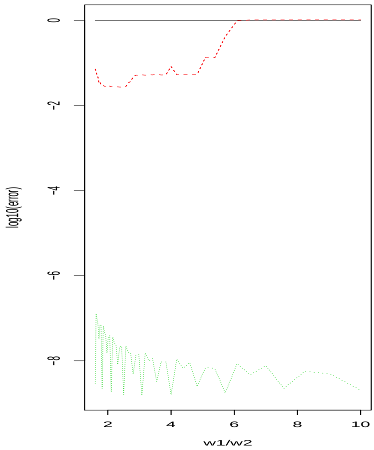
$$a(t) = \sqrt{a_1^2 + a_2^2 + 2a_1 a_2 \cos(2\pi(\omega_1 - \omega_2)t)}$$

$$\varphi(t) = -\arctan\left(\frac{a_1 - a_2}{a_1 + a_2}\tan\left(\frac{\omega_1 - \omega_2}{2}t\right)\right) \qquad \omega_0 = \frac{\omega_1 + \omega_2}{2}.$$

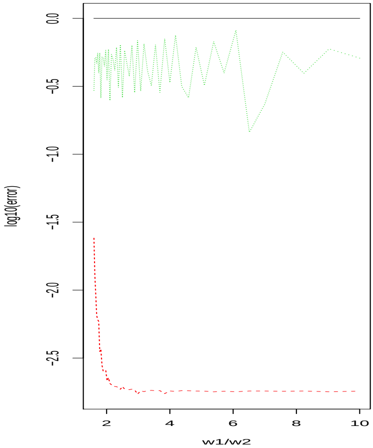
- Чем ближе ω_1 и ω_2, тем больше похоже модулированный косинус.
- Возьмем ω_1 и ω_2 подальше.

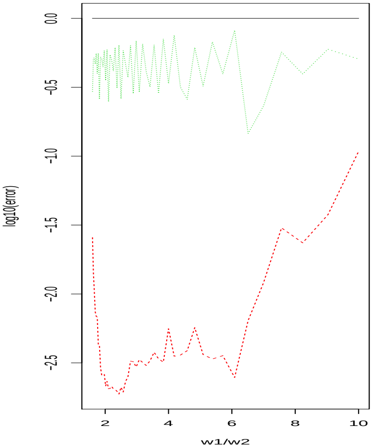
Сумма косинусов. Сравнение EMD и SSA





Сумма косинусов с одинаковой амплитудой. Сравнение EMD и SSA





Сравнение EMD и SSA. Добавим шуму

$$N = 1000, \quad M = 50, \quad \alpha = 0.05 \quad \beta = 0.15$$

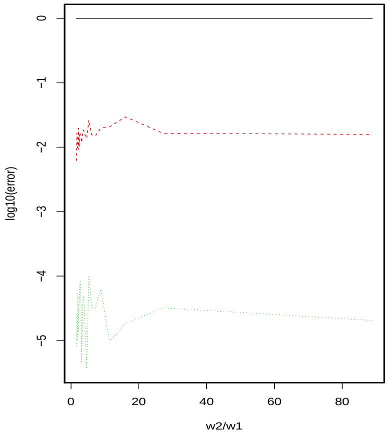
$$\omega_1, \omega_2 \in \Omega = \{\alpha + k\frac{\beta - \alpha}{M}\}_{k=0}^{M-1}$$

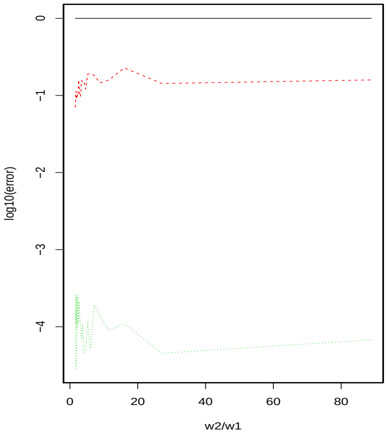
Рассмотрим ряд

$$x_n^{(1)} = 4\cos(2\pi\omega_1 n), \qquad x_n^{(2)} = \cos(2\pi\omega_2 n)$$
$$x_n = x_n^{(1)} + x_n^{(2)} + r_n, \qquad r_n \in N(0, \sigma^2).$$

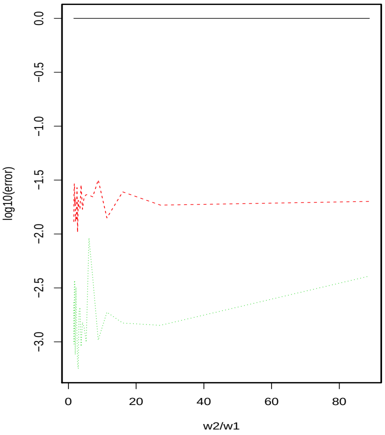
Появился новый параметр — σ, интересна зависимость.

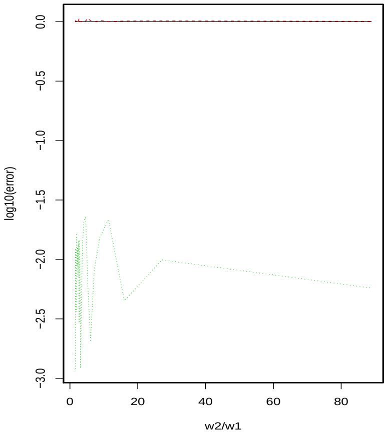
Зашумленная сумма косинусов. Сравнение EMD и SSA





Зашумленная сумма косинусов. Сравнение EMD и SSA





Заключение

- Рассмотрено преобразование Гильберта-Хуанга как метода анализа временных рядов.
- Получены условия для приближенного нахождения мгновенных характеристик.
- Проведено сравнение метода EMD с SSA как методов декомпозиции временных рядов.
- Изучены отличительные особенности метода EMD.