Санкт-Петербургский государственный университет
Математико-механический факультет
Кафедра Статистического Моделирования

«Научно-исследовательская работа» (семестр 6)

Поддерживающие временные ряды в MSSA

Выполнил:
Ткаченко Егор Андреевич
группа 19.Б04-мм

Научный руководитель:
к. ф.-м. н., доцент
Голяндина Нина Эдуардовна

Санкт-Петербург
2022

Оглавление

Введение

Полезность умения строить прогнозы не нуждается в доказательстве. Прогноз временных рядов может использоваться в прогнозе погоды, приливов, спроса на товары и многом другом.

С помощью книги [1] был изучен базовый SSA, разложение рядов, заполнение пробелов в данных, прогноз и базовый MSSA. Для работы с временными рядами и их прогнозом использовался пакет Rssa. Проведены эксперименты с простейшими моделями сигналов для изучения связи между согласованностью сигналов и поддерживающими рядами.

1. Определения

Вещественным временным рядом длины N называется вектор

$$\mathsf{F} = (f_0, \ldots, f_{N-1}),\ f_j \in \mathbb{R}.$$

Многомерным временным рядом $\vec{\mathsf{F}}$ называется набор s временных рядов $\mathsf{F}^{(p)}$ длин N_p:

$$\vec{\mathsf{F}} = \{\mathsf{F}^{(p)} = (f_0^{(p)}, \ldots, f_{N_p-1}^{(p)}),\ p = 1, \ldots, s\}.$$

L-траекторная матрица (или просто траекторная матрица) ряда F имеет структуру ганкелевой матрицы, а ее столбцами являются отрезки длины L ряда F

$$\mathcal{T}_{\mathsf{SSA}}(\mathsf{F}) = \begin{pmatrix} f_0 & f_1 & \ldots & f_{K-1} \\ f_1 & f_2 & \ldots & f_K \\ \vdots & \vdots & \ddots & \vdots \\ f_{L-1} & f_L & \ldots & f_{N-1} \end{pmatrix}.$$

L-траекторная матрица многомерного ряда $\vec{\mathsf{F}}$ состоит из горизонтально склеенных траекторных матриц рядов $\mathsf{F}^{(p)} \in \vec{\mathsf{F}}$:

$$\mathcal{T}_{\mathsf{MSSA}}(\vec{\mathsf{F}}) = [\mathcal{T}_{\mathsf{SSA}}(\mathsf{F}^{(1)}) : \ldots : \mathcal{T}_{\mathsf{SSA}}(\mathsf{F}^{(s)})].$$

Из траекторной матрицы можно восстановить ряд. Из любой матрицы подходящего размера можно получить траекторную матрицу проектированием на пространство ганкелевых матриц (или склеенных горизонтально ганкелевых матриц для многомерного случая).

Ранг ряда равен рангу его траекторной матрицы:

$$\mathsf{rank}\ \mathsf{F} = \mathsf{rank}\ \mathcal{T}_{\mathsf{SSA}}(\mathsf{F}), \qquad \mathsf{rank}\ \vec{\mathsf{F}} = \mathsf{rank}\ \mathcal{T}_{\mathsf{MSSA}}(\vec{\mathsf{F}}).$$

2. Применение SSA и MSSA

Алгоритмы SSA и MSSA могут быть применены для аппроксимации временного ряда рядом конечного ранга.

Алгоритм.

Вход: *Ряд* F *для* SSA *или многомерный ряд* $\vec{\mathsf{F}}$ *для* MSSA*, длина окна* $L \leq N$ *для* SSA *или* $L \leq N_p, \forall N_p$ *для* MSSA*, ранг аппроксимирующего ряда* r*.*

1 *Вложение. Временной ряд переводится в* L*-траекторную матрицу* $\mathbf{X}$

$$\mathbf{X} = \mathcal{T}_{\mathsf{SSA}}(\mathsf{F}) \text{ для SSA}, \qquad \mathbf{X} = \mathcal{T}_{\mathsf{MSSA}}(\vec{\mathsf{F}}) \text{ для MSSA}.$$

2 *Сингулярное разложение. Методом SVD матрица* $\mathbf{X}$ *раскладывается на сумму* d *матриц* $\mathbf{X}_i = \sqrt{\lambda_i} U_i V_i^T$*, где* $d = \mathsf{rank}\ \mathbf{X} = \mathsf{rank}\ \mathbf{X}\mathbf{X}^T \leq L$*,* λ_i *— собственные числа матрицы* $\mathbf{X}\mathbf{X}^T$ *(*$\lambda_1 \geq \ldots \geq \lambda_L \geq 0$*),* U_i *— собственные вектора матрицы* $\mathbf{X}\mathbf{X}^T$*,* $V_i = \mathbf{X}^T U_i / \sqrt{\lambda_i}$ *— факторные вектора матрицы* $\mathbf{X}$*.*

3 *Группировка. Множество индексов* $\{1, \ldots, d\}$ *делится на* m *непересекающихся множеств* $I_1, \ldots, I_m$*. Далее для каждого множества индексов (пусть* $I = i_1, \ldots, i_t$*) получается матрица* $\mathbf{X}_I = \mathbf{X}_{i_1} + \cdots + \mathbf{X}_{i_t}$*.*

Для аппроксимации рядом конечного ранга r*, понадобится множество из первых* r *индексов* $\{1, \ldots, r\}$*, соответствующую ему матрицу обозначу* $\overset{\wedge}{\mathbf{X}}_r = \mathbf{X}_1 + \cdots + \mathbf{X}_r$*.*

4 *Сгруппированные матрицы* $\mathbf{X}_{I_j}$ *восстанавливаются в ряды (*SSA*) или многомерные ряды (*MSSA*). Для получения аппроксимирующего ряда нужно восстановить его из матрицы* $\overset{\wedge}{\mathbf{X}}_r$*.*

Выход: *Аппроксимирующий ряд* $\overset{\wedge}{\mathsf{F}}_r$ *конечного ранга* r*.*

3. ЛРФ

Линейная рекуррентная формула (ЛРФ) выражает каждый член последовательности через линейную комбинацию предыдущих членов.

Ряд F длины N — управляемый ЛРФ, если существуют такие $a_1, \ldots, a_d$, что:

$$f_{i+d} = \sum_{k=1}^{d} a_k f_{i+d-k},\ 0 \leq i \leq N-1-d,\ a_d \neq 0,\ t < N-1.$$

Важно отметить, что ряд конечного ранга является управляемым ЛРФ [1, 2.1.2.2, стр. 35].

Вещественный временной ряд F, управляемый ЛРФ, естественным образом прогнозируется на одну точку:

$$\widetilde{f}_N = \sum_{k=1}^{L-1} a_k f_{N-k}.$$

Но тогда его можно прогнозировать и на любое количество точек.

Глава 1

Постановка задачи

Пусть имеется временной ряд $\mathsf{F}^{(1)} = \mathsf{S}^{(1)} + \mathsf{R}^{(1)}$, где сигнал $\mathsf{S}^{(1)}$ — ряд управляемый ЛРФ, шум $\mathsf{R}^{(1)}$ — ряд без структуры. Рассмотрим задачу прогнозирования $\mathsf{S}^{(1)}$. Эта задача уже решается методом SSA, но как можно улучшить прогноз?

Пусть помимо ряда $\mathsf{F}^{(1)}$ имеется временной ряд $\mathsf{F}^{(2)} = \mathsf{S}^{(2)} + \mathsf{R}^{(2)}$. Если структура сигналов $\mathsf{S}^{(1)}$ и $\mathsf{S}^{(2)}$ похожа, то использование ряда $\mathsf{F}^{(2)}$ может улучшить прогноз сигнала $\mathsf{S}^{(1)}$, потому что второй ряд дает алгоритму больше данных, которые могут улучшить ЛРФ. Возможность такого улучшения прогноза подтверждена [1, 4.3.3.3, стр. 216]. Но второй ряд может сделать прогноз хуже, если структура его сигнала отличается от первого.

Простейший пример похожих по структуре сигналов — гармонические колебания с одинаковыми периодом. Они даже могут быть смещены по фазе или иметь разную амплитуду.

Для объективной оценки качества прогноза буду использовать среднюю квадратичную ошибку.

$$\mathsf{MSE}(\widetilde{\mathsf{S}},\ \mathsf{S}) = \frac{1}{N_f} \sum_{i=N}^{N+N_f-1} (\widetilde{s}_i - s_i)^2,$$

где $\widetilde{\mathsf{S}}$ — прогноз сигнала S на N_f точек, N — длина прогнозируемого сигнала S.

Ряд $F^{(2)}$ называется поддерживающим для прогноза, если прогноз с его использованием лучше чем без него, т.е.

$$\mathsf{MSE}(\widetilde{\mathsf{S}}_{\mathsf{MSSA}},\ \mathsf{S}^{(1)}) < \mathsf{MSE}(\widetilde{\mathsf{S}}_{\mathsf{SSA}},\ \mathsf{S}^{(1)}).$$

Такое определение можно применить только в экспериментах с известным продолжением ряда. На практике не с чем сравнить прогноз, поэтому появляется вопрос. Как понять, что ряд поддерживающий не зная продолжения прогнозируемого ряда? Помочь ответить на этот вопрос может понятие согласованности.

Сигналы $\mathsf{S}^{(1)}$, $\mathsf{S}^{(2)}$ называются полностью согласованными, если ранг $r_{1,2} = r_1 = r_2$ и полностью несогласованными, если $r_{1,2} = r_1 + r_2$, где $r_1 = \mathsf{rank}\ \mathsf{S}^{(1)}$, $r_2 = \mathsf{rank}\ \mathsf{S}^{(2)}$, $r_{1,2} = \mathsf{rank}\ \vec{\mathsf{S}} = \mathsf{rank}\ \{\mathsf{S}^{(1)}, \mathsf{S}^{(2)}\}$.

Глава 2

Численные эксперименты

2.1. Первый эксперимент: выбор компонент для MSSA

Гипотеза: Когда сигналы похожи, их можно считать согласованными и лучше использовать (при прогнозе или восстановлении сигнала) ранг равный рангу одного сигнала. Когда сигналы отличаются, их следует считать не согласованными и использовать ранг равный сумме рангов сигналов.

Выберем в качестве первого ряда простой сигнал, зависящий от параметра (например, у косинуса параметр — период) с аддитивным гауссовым шумом с дисперсией $\sigma_1^2 = 0.2^2$. Второй ряд будет простым сигналом того же вида, с несколько отличающимся параметром и без шума. Восстановим и спрогнозируем первый ряд с помощью SSA, MSSA считая ряды согласованными и MSSA считая ряды несогласованными.

Для устойчивости результатов, повторим это 50 раз и усредним ошибки.

Назовем относительной ошибкой прогноза (восстановления) значение

$$\frac{error_{\mathsf{SSA}} - error_{\mathsf{MSSA}}}{error_{\mathsf{SSA}} + error_{\mathsf{MSSA}}},$$

где $error_{\mathsf{SSA}}, error_{\mathsf{MSSA}}$ — ошибки прогноза (восстановления) методами SSA и MSSA соответственно.

Значения относительной ошибки легко расположить на графике (она принимает значения от -1 до 1). По значению относительной ошибки легче понять, какой метод лучше (не надо сравнивать два значения ошибок, которые просто положительны и могут быть любых порядков).

Как интерпретировать значения относительной ошибки?

- Значения больше 0 значат, что что MSSA лучше SSA;
- Значения меньше 0 значат, что что MSSA хуже SSA;
- значения около 0 значат что ошибки примерно равны.
- значения далеко от 0 значат, что ошибки сильно отличаются.

Все сигналы будут нормироваться, чтобы амплитуда сигнала не влияла на ошибки прогноза и восстановления. Например, для косинуса: $s_j^{(i)} = A\cos(\frac{2\pi j}{T_i})$, где $A = \mathsf{mean}(|s_j^{(i)}|)^{-1}$.

2.1.1. Сигнал косинус

Сигнал $\mathsf{S}^{(1)}$ — косинус с периодом $T_1 = 8$. Сигналы $\mathsf{S}^{(2)}$ — косинусы с периодами $T_2 \in \{8, 8.02, 8.04, 8.06, 8.08, 8.1, 8.15\}$.

Ранг косинуса равен 2, поэтому для MSSA используются первые 2 или первые 4 компоненты разложения, а для SSA только 2.

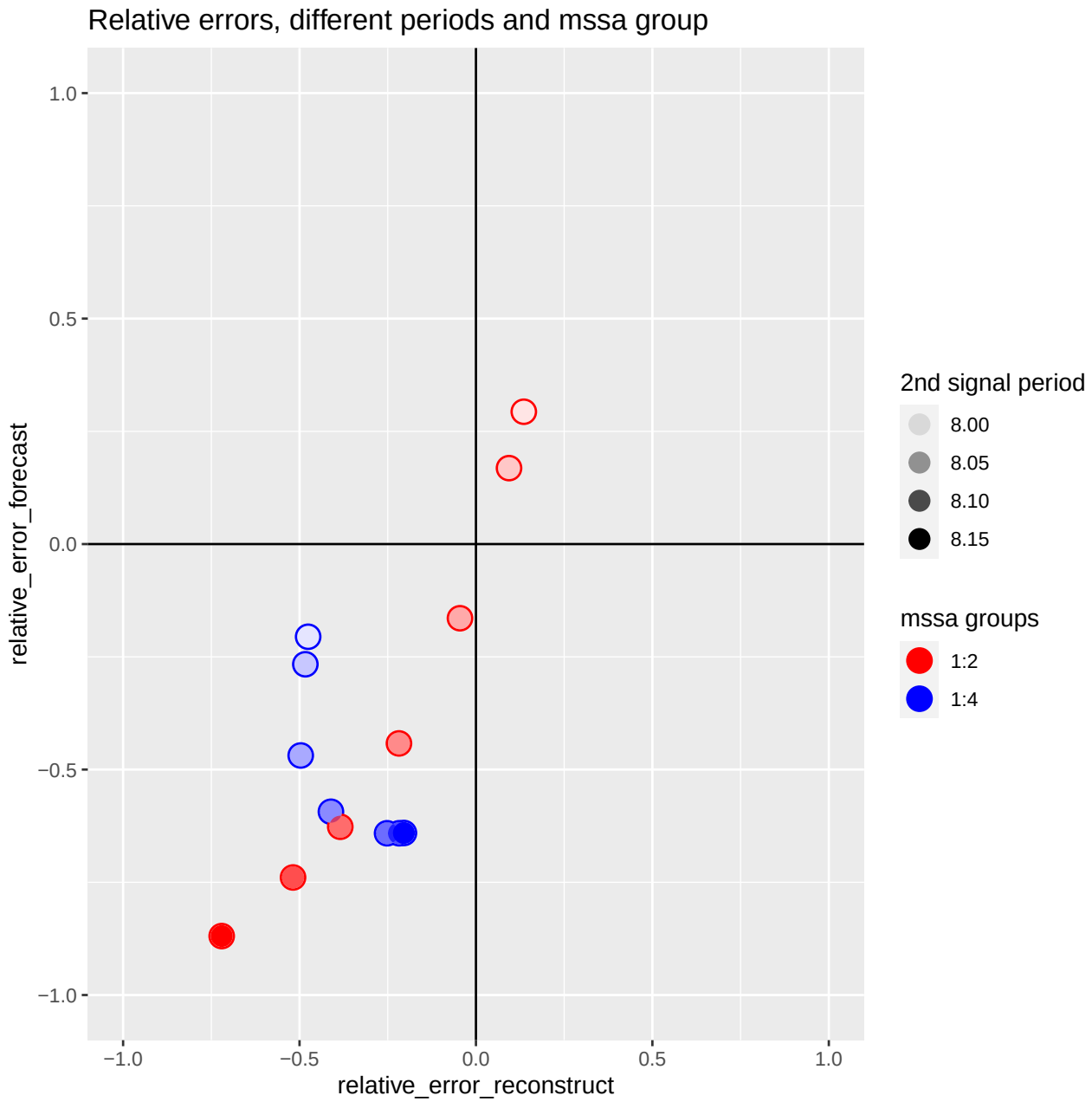

Рис. 2.1. Зависимость относительных ошибок от разницы сигналов и выбора ранга для MSSA.

На рис. 2.1 видим подтверждение гипотезы для косинусов: с увеличением разницы рядов использование четырех компонент становится лучше и для прогноза и для восстановления сигнала.

2.1.2. Сигнал экспонента (показательная функция)

Функция для сигналов — $s_j^{(i)} = A\exp(j\lambda_i)$. Сигнал $\mathsf{S}^{(1)}$ — нормированная показательная функция с $\lambda_1 = 0.005$. Сигналы $\mathsf{S}^{(2)}$ — нормированная показательная функция с $\lambda_2 \in \{0.005, 0.0075, 0.01, 0.0125, 0.015, 0.02, 0.025, 0.03\}$.

Ранг показательной функции равен 1, поэтому для MSSA используется первая или первые 2 компоненты разложения, а для SSA только первая.

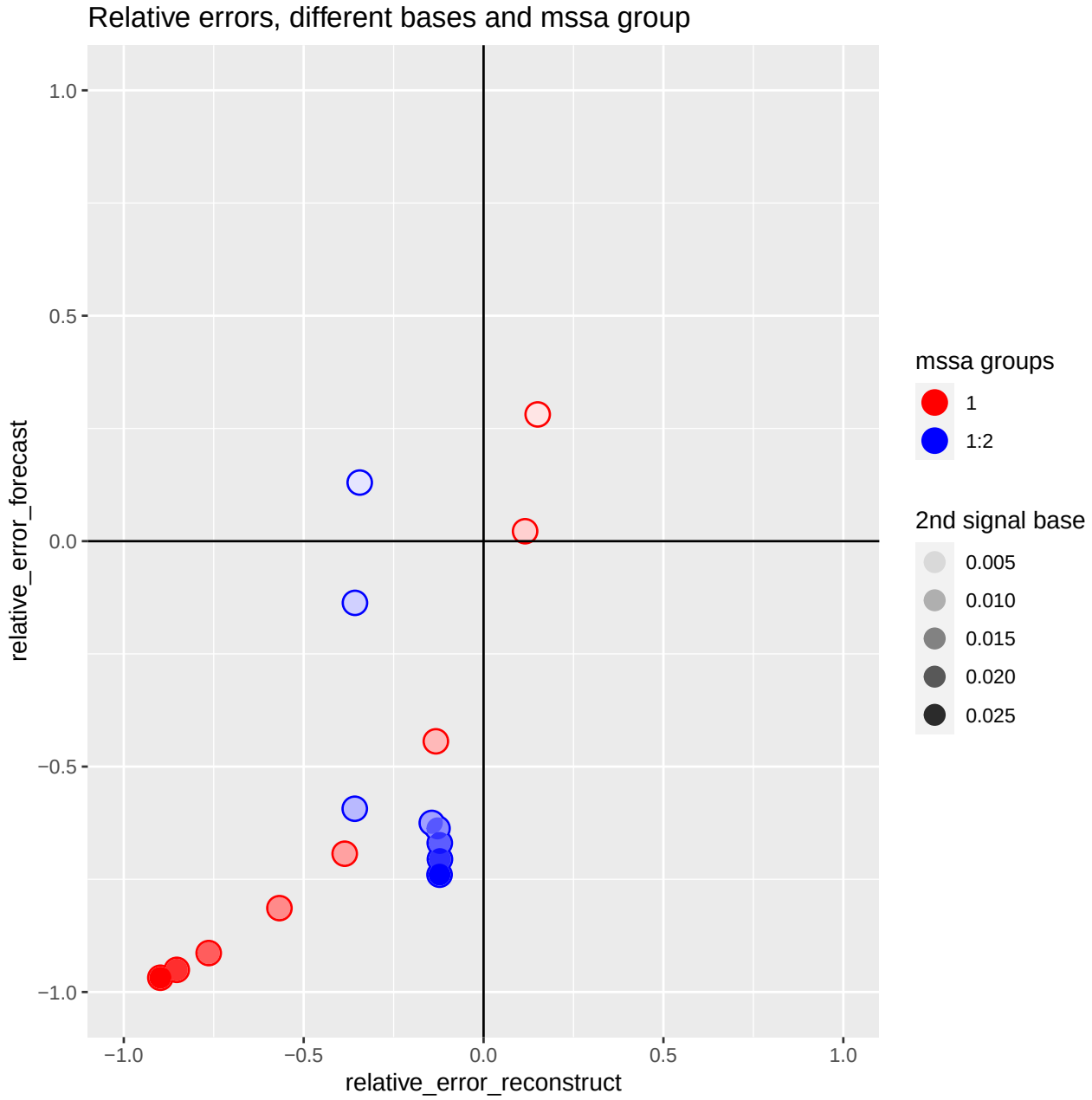

Рис. 2.2. Зависимость относительных ошибок от разницы сигналов и выбора ранга для MSSA.

На рис. 2.2 снова видим подтверждение гипотезы для показательных функций: с увеличением разницы рядов использование двух компонент становится лучше.

2.1.3. Сигнал косинус с показательной модуляцией (общий период, меняющаяся модуляция)

Функция для сигналов — $s_j^{(i)} = A\exp(j\lambda_i)\cos(\frac{2\pi j}{8})$. Сигнал $\mathsf{S}^{(1)}$ — функция с $\lambda_1 = 0.005$. Сигналы $\mathsf{S}^{(2)}$ — функция с $\lambda_2 \in \{0.005, 0.0075, 0.01, 0.0125, 0.015, 0.02, 0.025, 0.03\}$.

Ранг косинуса с модуляцией равен 2, поэтому для MSSA используются первые 2 или первые 4 компоненты разложения, а для SSA только 2.

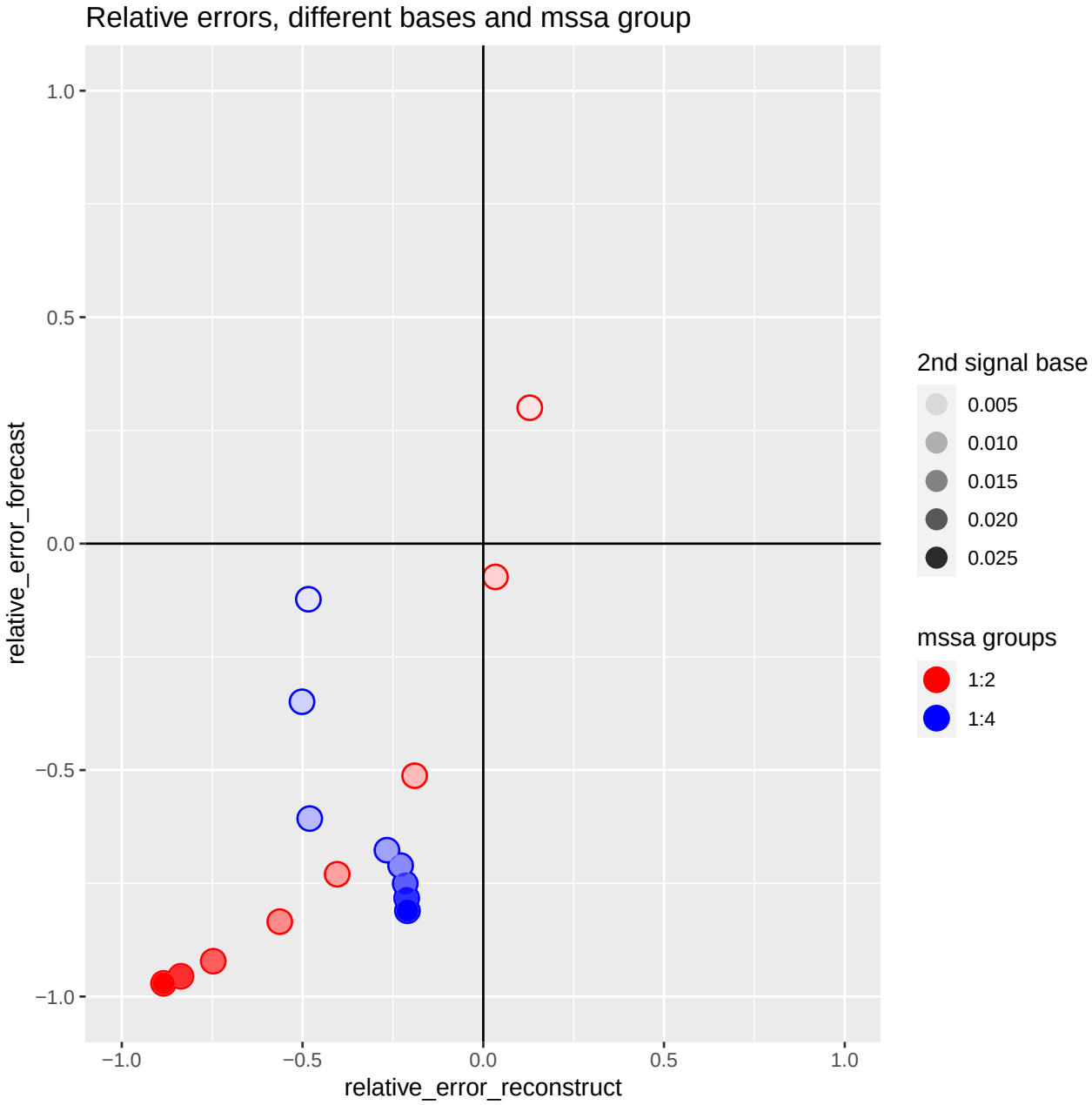

Рис. 2.3. Зависимость относительных ошибок от разницы сигналов и выбора ранга для MSSA.

На рис. 2.3 видим подтверждение гипотезы для косинусов с меняющейся модуляцией.

2.1.4. Сигнал косинус с показательной модуляцией (меняющийся период, общая модуляция)

Функция для сигналов — $s_j^{(i)} = A\exp(0.02j)\cos(\frac{2\pi j}{T_i})$. Сигнал $\mathsf{S}^{(1)}$ — функция с $T_1 = 8$. Сигналы $\mathsf{S}^{(2)}$ — функция с $T_2 \in \{8, 8.02, 8.04, 8.06, 8.08, 8.1, 8.15\}$.

Ранг косинуса с модуляцией равен 2, поэтому для MSSA используются первые 2 или первые 4 компоненты разложения, а для SSA только 2.

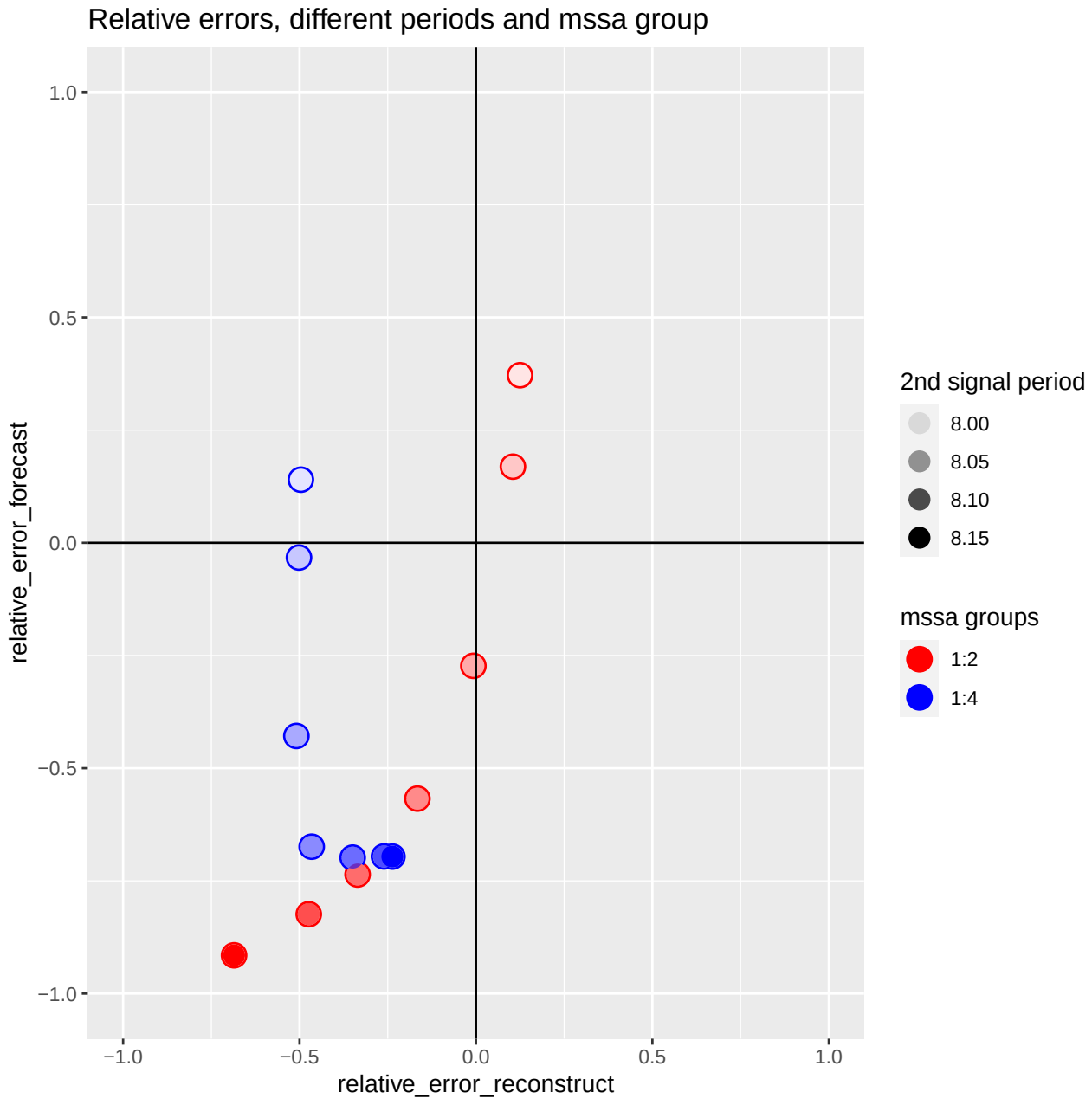

Рис. 2.4. Зависимость относительных ошибок от разницы сигналов и выбора ранга для MSSA.

На рис. 2.4 видим подтверждение гипотезы для модулируемых косинусов с меняющимся периодом.

2.1.5. Результат первого эксперимента

Для всех видов сигналов при отклонении второго сигнала от первого всегда наступал момент, когда использование удвоенного ранга лучше.

Так как большая часть наблюдений оказалась нижней левой четверти графика относительных ошибок, это значит что во всех случаях, кроме MSSA с маленькой разницей сигналов и рангом равным рангу сигнала, использование SSA дает лучший результат.

2.2. Второй эксперимент: сравнение ошибки прогноза методами SSA и MSSA при разных величинах шума первого ряда и отклонениях второго ряда

Никита Федоров в свей выпускной квалификационной работе изучал влияние величины второго шума на результаты работы SSA, MSSA, ProjSSA. Рассмотрим влияние величины первого шума на прогноз SSA и MSSA.

Гипотеза: При увеличении шума первого ряда, MSSA станет лучше для любого отклонения второго ряда. Если это так, то можно найти зависимость граничного значения дисперсии шума от изменения параметра второго сигнала.

2.2.1. Сигнал косинус

Тут не хватает этого эксперимента с остальными видами рядов. и описания того что на рисунках 2.5, 2.6.

2.3. Третий эксперимент: линейные сигналы

Как видно на рис. 2.7 иногда линейный сигнал можно хорошо аппроксимировать показательной функцией.

Ранг линейного сигнала равен 2, а показательного — 1.

2.3.1. Является ли первая компонента разложения линейного ряда показательной функцией?

Из рис. 2.8 видно что первая компонента разложения линейного ряда не является показательной функцией, так как показательные функции не могут дважды пересекаться.

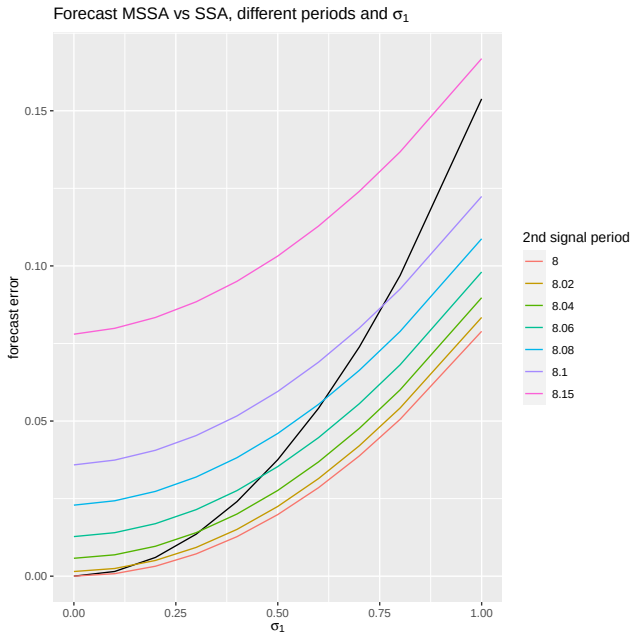

Рис. 2.5. Ошибка прогноза для SSA и MSSA.

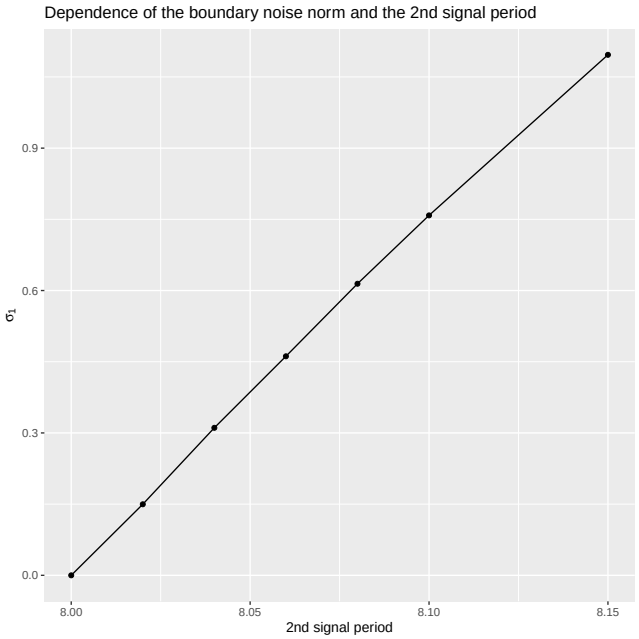

Рис. 2.6. Зависимость значения шума при котором SSA становится хуже MSSA от второго сигнала

Заметим, что при меньших углах наклона, аппроксимация получается лучше. Есть ли зависимость доли второй компоненты в линейном сигнале от угла наклона?

На рис. 2.9 видна линейная зависимость логарифма из доли второй компоненты и угла наклона сигнала, поэтому доля второй компоненты зависит экспоненциально от угла наклона.

Тут не хватает части экспериментов с линейными функциями и описания графиков

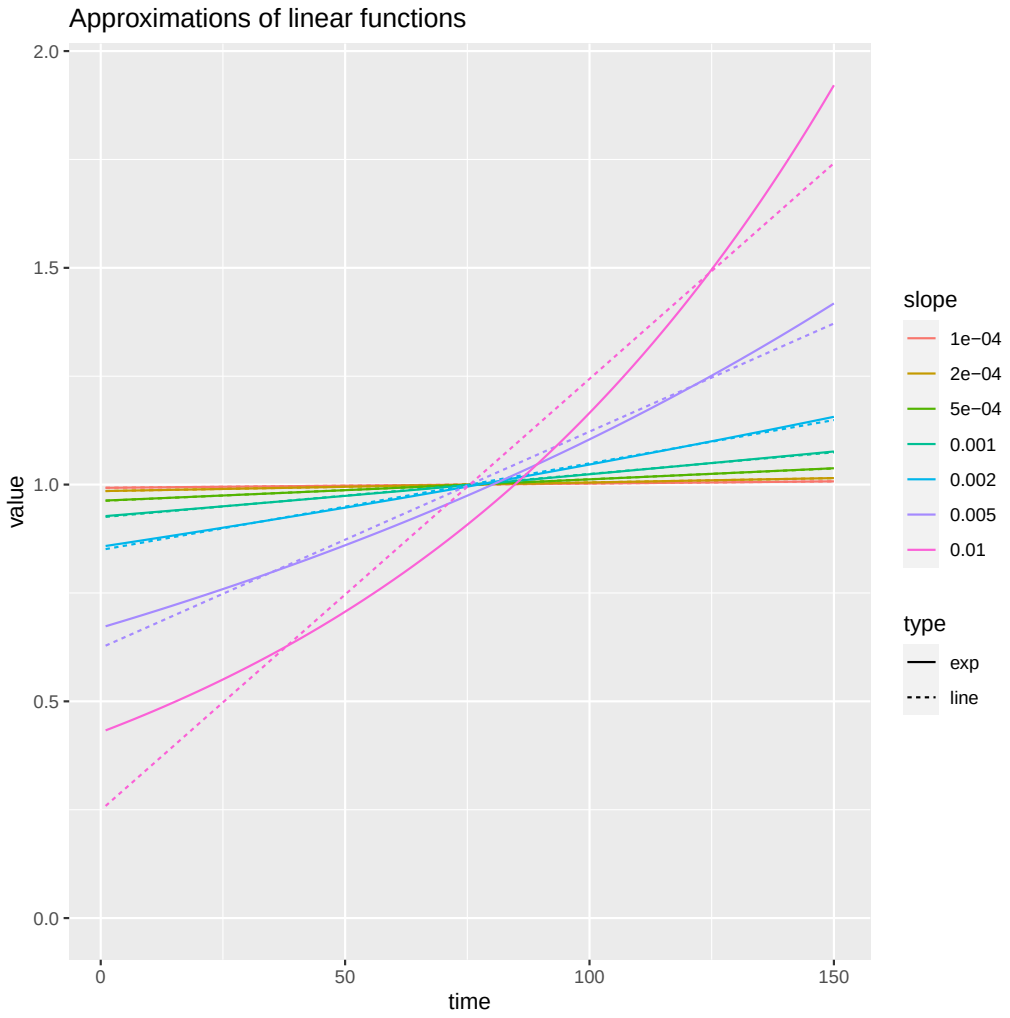

Рис. 2.7. Пример аппроксимации линейных функций показательными.

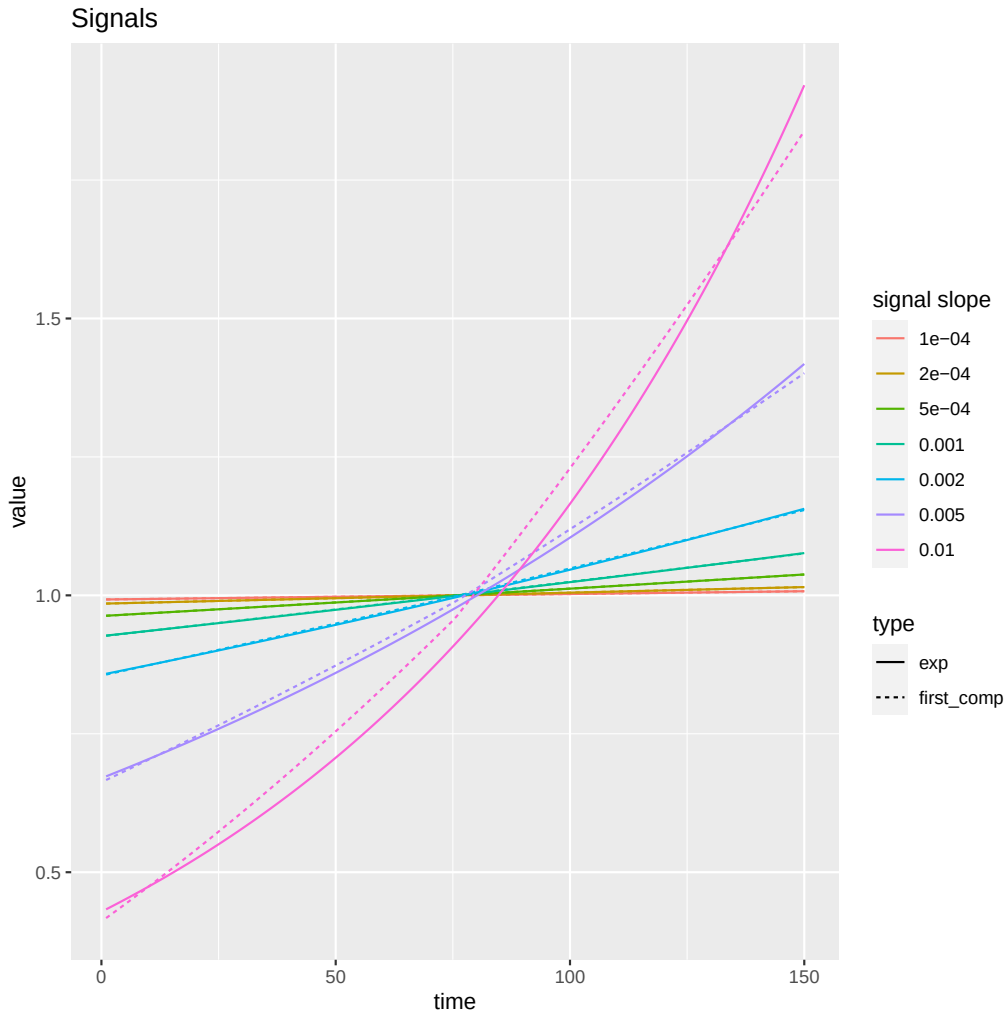

Рис. 2.8. Сравнение первых компонент сигнала и аппроксимирующих экспонент.

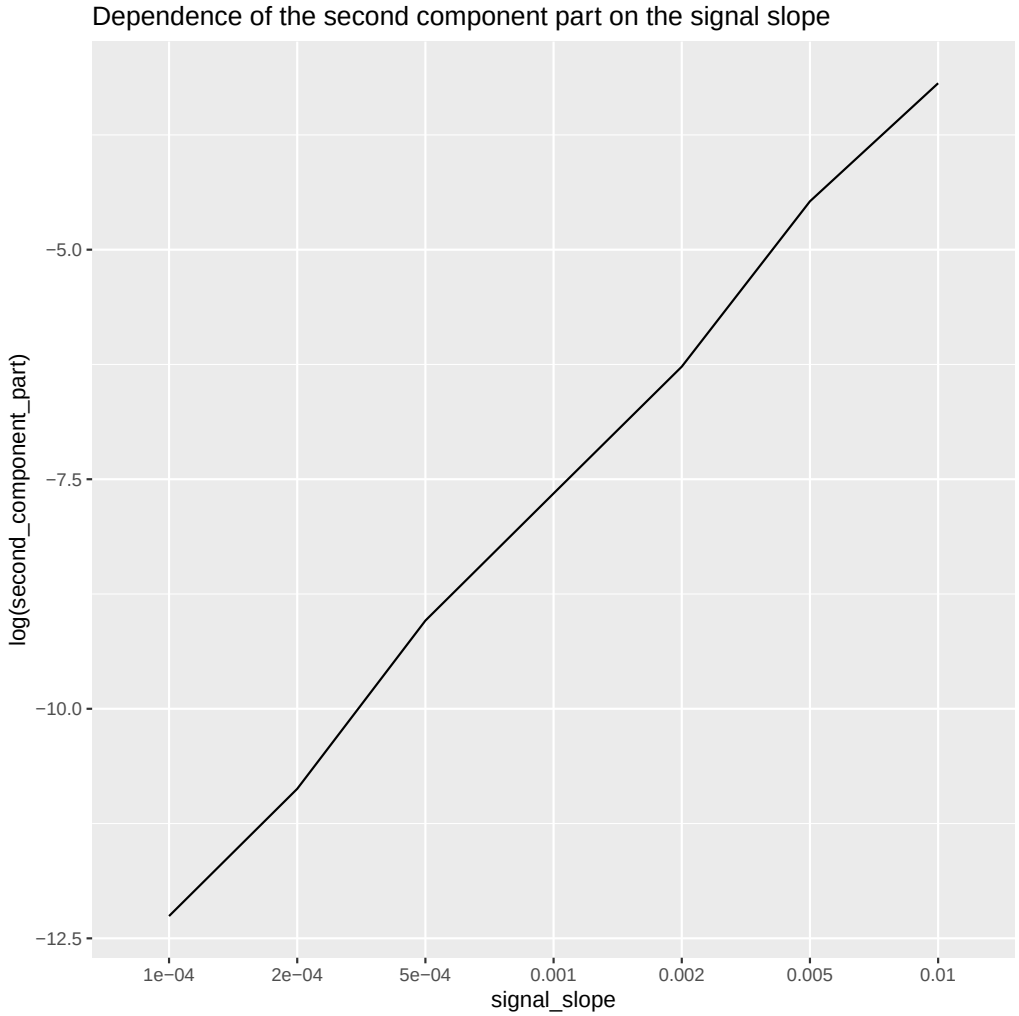

Рис. 2.9. Зависимость доли второй компоненты от угла наклона линейного сигнала.

Заключение

Найдено много интересных зависимостей, характер которых еще предстоит объяснить.

Список литературы

1. Golyandina N, Korobeynikov A, Zhigljavsky A. Singular Spectrum Analysis with R. — Springer, 2018. — P. 272.