Ф.И. Александров

ОЦЕНКА КОЭФФИЦИЕНТОВ ЛИНЕЙНОЙ РЕКУРРЕНТНОЙ ФОРМУЛЫ ПОРЯДКА 2, УПРАВЛЯЮЩЕЙ СИГНАЛОМ ¹

1 Введение

При исследовании временных рядов часто возникают задачи, которые решаются в рамках определенной модели ряда. К таким задачам относится выделение аддитивной составляющей ряда, которая удовлетворяет определенной модели, или аппроксимация ряда функцией известного вида. При таком подходе саму модель предполагают заданной и задача сводится к оценке ее параметров. При этом могут ставиться разные цели – максимально точное определение значений параметров, аппроксимация исходного ряда рядом, удовлетворяющим модели, поиск "модельного" ряда, дающего наилучший прогноз и т.д.

Рассмотрим одномерный вещественнозначный временной ряд $F_N = (f_0, \ldots, f_{N-1})$ длины N, N > 2. Под моделью ряда будем понимать способ задания его элементов f_n . Можно выделить два типа моделей: авторегрессионные (AR) и модели "сигнал плюс шум". В моделях типа авторегрессии элементы ряда $F_N = (f_0, \ldots, f_{N-1})$ представляются следующим образом:

$$f_n = \mathcal{F}(f_0, \dots, f_{n-1}) + \varepsilon_n,$$

где \mathcal{F} – какая-то функция, а ε_n – случайная шумовая составляющая. Видно, что ε_n воздействует не только на f_n , но и на все последующие элементы f_{n+1}, f_{n+2}, \ldots В модели "сигнал плюс шум" элемент представляется следующим образом:

$$f_n = s_n + \varepsilon_n, \quad s_n = \mathcal{S}(s_0, \dots, s_{n-1}),$$

то есть можно представить такой ряд, как детерминированный сигнал s_0, \ldots, s_{N-1} , наблюдения которого проводятся с ошибкой ε_n . Сигнал в общем случае может быть представлен с помощью рекуррентной формулы S.

Для процессов авторегрессии (порядка p) полагают, что функция \mathcal{F} имеет линейный вид и f_n зависит от p предыдущих значений, случайная

¹©Ф.И. Александров, 2004

шумовая составляющая имеет нулевое среднее и постоянную ковариацию. Таким образом, параметрами такого процесса, кроме ковариации, будут коэффициенты a_1, \ldots, a_p функции \mathcal{F} :

$$f_n = a_1 f_{n-1} + \ldots + a_p f_{n-p} + \varepsilon_n$$

Подход, подобный авторегрессионному, используется и в моделях "сигнал плюс шум" следующим образом: полагают, что сигнал ряда задается линейной рекуррентной формулой (ЛРФ):

$$f_n = s_n + \varepsilon_n, \quad s_n = a_1 s_{n-1} + \ldots + a_p s_{n-p}.$$
 (1)

Такая модель тоже имеет p неизвестных параметров a_1, \ldots, a_p вдобавок к параметрам распределения шума. Наложим следующие стандартные ограничения на шум: $\varepsilon_0, \ldots, \varepsilon_{N-1}$ независимы, одинаково распределены, $\mathbf{E}\varepsilon_n = 0$ и $\mathbf{D}\varepsilon_n = \sigma^2$.

Данная работа посвящена разработке и исследованию методов решения задачи оценки параметров a_1 и a_2 для модели (1) "сигнал плюс шум", где сигнал задается ЛРФ второго порядка (p = 2). Рассматриваются два метода, в основании одного лежит идея оценки коэффициентов линейной регрессии с помощью метода наименьших квадратов (МНК), а другой использует общий подход к исследованию временных рядов "Гусеница"-SSA [1, 2, 3]. Мы будем использовать работу [2] в качестве основного источника о нем и придерживаться веденных там обозначений.

Качество метода будем оценивать через близость оценки параметров к их реальным значениям. Сравнивать методы будем с помощью статистического моделирования.

2 Методы оценки параметров

Рассмотрим ряд $F_N = (f_0, \ldots, f_{N-1}), N > 2$:

$$f_n = s_n + \varepsilon_n, \quad s_n = a_1 s_{n-1} + a_2 s_{n-2}, \ \mathbf{E}\varepsilon_n = 0, \ \mathbf{D}\varepsilon_n = \sigma^2,$$
 (2)

 $S_N = (s_0, \ldots, s_{N-1})$ – детерминированный сигнал. Сформулируем методы оценки параметров a_1 и a_2 . Первый метод, назовем его SSALRF, основан на подходе "Гусеница"-SSA, с помощью него оцениваются коэффициенты некоторой ЛРФ, управляющей сигналом ряда, рассматривается ее характеристический полином и, используя значения "главных" корней этого полинома, строится минимальная ЛРФ (порядка 2). Во

втором методе, REGR, рассмотрим стандартную модель линейной регрессии, подставим вместо зависимых и независимых переменных значения ряда F_N таким образом, что они будут связаны соотношением (2), и посчитаем \hat{a}_1 и \hat{a}_2 с помощью формулы для оценки MHK.

2.1 Metod SSALRF

Метод "Гусеница"-SSA может быть использован для различных общих задач исследования временных рядов, в частности – для выделения сигнала и нахождения его ЛРФ. При его использовании по ряду F_N строится траекторная матрица **X** заданного размера $L \times K$, 1 < L < N, K = N - L + 1 (L называется длиной окна), вычисляются собственные числа $\{\lambda_i\}_{i=1}^L$, собственные $\{U_i\}_{i=1}^L$ и факторные $\{V_i\}_{i=1}^L$ вектора матрицы **XX**^T, формируя сингулярное разложение **X** = $\sum_i \sqrt{\lambda_i}U_iV_i^T$ (см. [2, Section 1.2.2],[1, Раздел 1.3.2]). Набор ($\sqrt{\lambda_i}, U_i, V_i$) будем называть называется сингулярной тройкой. Ряду сопоставляется траекторное пространство, аддитивной составляющей ряда при выполнении условий разделимости соответствует собственное траекторное подпространство в этом пространстве (см [2, Section 1.2.3],[1, Раздел 1.3.3]). В условиях приближенной разделимости метод позволяет найти подпространство близкое к траекторному пространству рассматриваемой аддитивной составляющей.

Опишем алгоритм метода SSALRF, в нем можно выделить следующую последовательность шагов.

Алгоритм 2.1.

- 1. Выбор длины окна L и построение траекторной матрицы $\mathbf{X} \in \mathbb{R}^{L \times K}$ по ряду F_N ;
- 2. Сингулярное разложение траекторной матрицы $\mathbf{X} = \sum_{i} \sqrt{\lambda_i} U_i V_i^T$;
- 3. Выбор сингулярных троек, соответствующих сигналу S_N ;
- Построение по собственным векторам выбранных сингулярных троек приближенной ЛРФ сигнала порядка L − 1;
- 5. Нахождение корней характеристического полинома этой ЛРФ;
- 6. Поиск главных корней среди всего множества корней;
- Получение приближенной минимальной ЛРФ (порядка 2) сигнала по главным корням.

Поясним и распишем шаги алгоритма метода SSALRF. Ключевыми стадиями являются: построение $\Pi P \Phi$ сигнала порядка L - 1 и ее минимизация.

2.1.1 Получение ЛРФ порядка L - 1 для сигнала

Рассмотрим сигнал S_N , ряд ранга 2, описываемый ЛРФ порядка 2. То, что ранг ряда равен d по определению означает, что размерность Lтраекторного пространства dim $\mathcal{L}^{(L)} = d$ для любого допустимого L, см. [2, chapter 5],[1, раздел 3].

Воспользуемся теоремой [2, Theorem 5.2] для поиска коэффициентов ЛРФ сигнала порядка L-1. Обозначим для вектора $P = (p_1, \ldots, p_L)^T$: $P^{\nabla} = (p_1, \ldots, p_{L-1})^T, \ \pi = p_L$. Рассмотрим P_1, P_2 – некоторый базис L-траекторного пространства $\mathcal{L}_2(S)$ сигнала S_N .

Если $L > 2, e_L \notin \mathcal{L}_2(S)$, то существует ЛРФ сигнала порядка L - 1 с коэффициентами b_1, \ldots, b_{L-1} , которые задаются равенством:

$$(b_{L-1},\ldots,b_1)^T = \frac{\pi_1 P_1^{\nabla} + \pi_2 P_2^{\nabla}}{1 - \pi_1^2 - \pi_2^2}.$$
(3)

Таким образом, чтобы построить ЛРФ сигнала S_N порядка L - 1, необходимо сначала провести сингулярное разложение траекторной матрицы ряда F_N (шаги 1-2 алгоритма 2.1). Далее, найдя две сингулярные тройки, которые соответствуют сигналу, возьмем их собственные вектора U_i, U_j в качестве базиса $\{P_1, P_2\}$ собственного подпространства сигнала (см. [2, Section 1.2.2]), а на шаге 4 построим по ним ЛРФ порядка L - 1, посчитав ее коэффициенты по формуле (3).

Так как нашей задачей является получение коэффициентов ЛРФ для сигнала порядка 2, то необходимо "минимизировать" полученную ЛРФ. Возможность такой минимизации опирается на тот факт, что для бесконечного ряда конечного ранга его ранг совпадает с порядком минимальной ЛРФ (см. [2, Corollary 5.1],[1, Предложение 3.4]).

2.1.2 Построение минимальной ЛРФ

Рассмотрим характеристический полином, соответствующий ЛРФ порядка L-1 с коэффициентами b_1, \ldots, b_{L-1} :

$$P_{L-1}(\lambda) = \lambda^{L-1} - \sum_{k=1}^{L-1} b_k \lambda^{L-k-1}$$

и пусть $b_{L-1} \neq 0$. Обозначим за k_m $(1 \leq m \leq p)$ кратности несовпадающих корней $\lambda_1, \ldots, \lambda_p$ полинома и пусть $h_n(m, j) = n^j \lambda_m^n, 1 \leq m \leq p, 0 \leq j \leq k_m - 1.$

Известно, что рядFудовлетворяет заданной ЛР
Ф тогда и только тогда, когда

$$f_n = \sum_{m=1}^p \sum_{j=0}^{k_m - 1} c_{mj} h_n(m, j),$$
(4)

где комплексные коэффициенты c_{mj} определяются по первым значениям ряда f_0, \ldots, f_{L-2} (см. [2, Theorem 5.3]).

Таким образом, мы установили связь между коэффициентами ЛРФ и видом ряда: коэффициенты ЛРФ задают характеристический полином, корни которого определяют вид ряда. Для ряда существуют ЛРФ разных порядков, минимальная ЛРФ порождает корни, которые называются главными. Набор корней неминимальной ЛРФ состоит из этих корней, к которым добавляются несущественные корни. Поэтому для того чтобы минимизировать ЛРФ произвольного порядка, необходимо на шаге 6 алгоритма 2.1 среди всех корней ее характеристического полинома отобрать главные и по ним на следующем шаге 7 восстановить коэффициенты минимальной ЛРФ (которые равны коэффициентам полинома с такими корнями).

Замечание 2.1. Поиск главных корней строится на представлении (4). Нетрудно увидеть, что линейному ряду соответствует два главных корня $\lambda_1 = \lambda_2 = 1$, а гармоническому ряду с частотой ω , $1 < \omega < 1/2$, соответствуют два комплексно-сопряженных главных корня $\lambda_1 = e^{i2\pi\omega}$, $\lambda_2 = e^{-i2\pi\omega}$, таких, что $|\lambda_1| = |\lambda_2| = 1$ (см. [2, Example 5.10]).

Рассмотрим также модификацию метода SSALRF, которая использует идею, лежащую в основании итеративной процедуры Cadzow применения "Гусеницы"-SSA (см. [4]) и состоит из двух частей. Сначала с помощью "Гусеницы"-SSA выделим сигнал ряда F_N . После этого применим описанный выше метод SSALRF, но уже к выделенному ряду. Способ выделения сигнала зависит от его вида, решение этой задачи излагается в работах [1, 2].

2.2 Metod REGR

Рассмотрим стандартную модель линейной регрессии: $Y = \mathbf{Z}^T \boldsymbol{\beta} + \boldsymbol{\xi}$, где **Z** – матрица независимых переменных, Y – вектор зависимых переменных, $\boldsymbol{\xi}$ – случайная ошибка, а $\boldsymbol{\beta}$ – вектор коэффициентов регрессии. Наложим стандартные ограничения: $\mathbf{A} = \mathbf{Z}\mathbf{Z}^T$ – неособенная, $\mathbf{E}\boldsymbol{\xi} = \mathbf{0}$, $\mathbf{D}\boldsymbol{\xi} = \sigma^2 \mathbf{E}$. Обозначим через

$$\hat{\boldsymbol{\beta}} = \mathbf{A}^{-1} \mathbf{Z} Y. \tag{5}$$

оценку метода наименьших квадратов для коэффициентов линейной регрессии.

Определим Y таким образом, чтобы его элементами были значения f_2, \ldots, f_{N-1} ряда F_N , обозначим коэффициенты регрессии как a_1, a_2 , а элементы матрицы **Z** определим таким образом, чтобы (n + 1)-ый элемент Y (т. е. f_n) задавался следующим уравнением:

$$f_n = \begin{pmatrix} f_{n-1} & f_{n-2} \end{pmatrix} \begin{pmatrix} a_1 \\ a_2 \end{pmatrix} + \varepsilon_n = a_1 f_{n-1} + a_2 f_{n-2} + \varepsilon_n.$$

Заметим, что при таких подстановках данная модель совпадает с авторегрессионной моделью для членов ряда F_N . Возьмем и формально по формуле для оценки МНК коэффициентов линейной регрессии оценим a_1, a_2 .

Метод REGR состоит в вычислении оценок \hat{a}_1 и \hat{a}_2 по формуле (5): $(\hat{a}_1, \hat{a}_2)^T = \mathbf{A}^{-1} \mathbf{Z} \mathbf{Y}$, где $\mathbf{Z} = [Z_1 : \ldots : Z_{N-1}], Z_i = (f_{i-1}, f_i)^T, Y = (f_2, \ldots, f_{N-1})^T$.

Замечание 2.2. Метод REGR больше подходит для AR-модели, тогда как подход "Гусеница"-SSA, а, следовательно, и метод SSALRF, предназначен для модели "сигнал плюс шум".

3 Сравнение методов на конкретном примере: гармонический ряд

3.1 Методы в применении к гармоническому ряду

Сравним описанные методы, применяя их для оценки коэффициентов a_1, a_2 ряда, сигнал которого задается ЛРФ 2-го порядка. В качестве такого ряда возьмем зашумленный гармонический ряд, $F_N = (f_0, \ldots, f_{N-1})$:

$$f_n = A\sin(2\pi\omega n + \phi) + \varepsilon_n, \quad \varepsilon_n \sim N(0, \sigma^2), \tag{6}$$

или в рекуррентном представлении:

$$f_n = s_n + \varepsilon_n$$
, где $s_n = a_1 s_{n-1} + a_2 s_{n-2}$, $\varepsilon_n \sim N(0, \sigma^2)$.

Как уже было сказано выше, характеристический полином минимальной ЛРФ гармонического ряда имеет два комплексных сопряженных корня с единичным модулем. Поэтому в методе SSALRF будем минимизировать ЛРФ, выделяя такие корни и по ним пересчитывая коэффициенты минимальной ЛРФ, как было описано выше.

В модификации метода SSALRF, когда сначала проводится выделение гармонического сигнала, будем проводить его следующим стандартным способом (см. [1, Раздел 3.2.1],[2, Section 1.6.1]): проведем разложение ряда F_N с длиной окна L, возьмем первые две сингулярные тройки ($\sqrt{\lambda_i}, U_i, V_i$), построим по каждой из них элементарную матрицу $\mathbf{X}_i = \sqrt{\lambda_i}U_iV_i^T$ и диагональным усреднением матрицы $\mathbf{X}_1 + \mathbf{X}_2$ выделим искомый сигнал.

Замечание 3.1. Для гармонического ряда $a_2 = -1$. В приведенных выше методах вычисление \hat{a}_1 и \hat{a}_2 происходит одновременно, но мы будем приводить результаты только для коэффициента a_1 , ведь качество оценки a_2 не важно. Если для ряда, кроме того, что его ранг равен 2, известно, что его параметрическая модель – гармоника, то для получения более точных результатов стоит принимать $a_2 = -1$, невзирая на посчитанное значение оценки \hat{a}_2 .

Замечание 3.2. Значение a_2 , отличное от -1, задает ЛРФ, управляющую экспоненциально-модулированным гармоническим рядом.

Замечание 3.3. В качестве длины окна в методе SSALRF берется $L = \lfloor N/2 \rfloor$. Рассматривались также другие значения для длины окна. Например, при L = 3 траекторное пространство исходного ряда является подпространством \mathbb{R}^3 , т. е. его размерность не может превышать 3 (что соответствует ЛРФ порядка 2), а, так как мы ищем ЛРФ именно порядка 2, значит характеристический полином не имеет лишних корней. Но при L = N/2 методы показали наилучшие результаты, видимо из-за асимптотического по min(L, K) характера разделимости синуса и шума (см. [1, раздел 2.4],[2, section 6.1.3]).

Замечание 3.4. Фаза ϕ в нашей модели зашумленного гармонического ряда (6) не влияет на результаты оценки a_1 и a_1 , а также не играет роли при вычислениях оценок, поэтому при моделировании можно считать $\phi = 0$.

Заметим, что проблема оценки частоты ω зашумленного гармонического ряда встречается чаще, чем задача оценки коэффициентов ЛРФ. Так как из замечания 2.1 следует, что ω выражается через a_1, a_2 следующим образом:

$$\omega = \frac{1}{2\pi} \arccos\left(\frac{a_1}{2\sqrt{-a_2}}\right),\tag{7}$$

то приведем в сравнении методов также результаты для оценки ω , которая расчитывалась через \hat{a}_1 и \hat{a}_2 согласно (7). Результаты для $\hat{\omega}$ удобны также тем, что они проще интерпретируются, в отличии от результатов для \hat{a}_1, \hat{a}_2 .

3.2 Схема сравнения и результаты

Сравнение методов проводилось как при работе с рядами различной длины N, так и для разных стандартов шума σ . Для получения более полной картины при одном фиксированном показателе (например, при фиксированном N) другой изменялся с некоторым шагом в установленных пределах. При этом мы могли видеть, как качество методов (абсолютное и в сравнении с другим методом) изменяется при увеличении N и σ .

Четыре показателя: период $T = 1/\omega$, амплитуда A, длина ряда N и стандарт шума σ определяют в нашей модели соотношение сигналшум. Зафиксируем T и A с тем, чтобы интерпретация результатов стала возможной. Пусть T = 6, A = 5.

Качество модифицированного SSALRF метода Исследование метода, являющегося модификацией метода SSALRF (в нем проводится сначала выделение сигнала и уже к нему применяется SSALRF) показало, что его качество при оценке a_1, ω очень незначительно отличается от качества собственно SSALRF метода. Смещение оценки \hat{a}_1 и дисперсии оценок \hat{a}_1 и $\hat{\omega}$ у модифицированного метода чуть меньше, чем у SSALRF, но эта разница очень мала. Стоит учесть, что в первой части модифицированного метода проводятся трудоемкие вычисления, поэтому он работает медленнее, чем SSALRF. Поэтому при дальнейшем сравнении с методом REGR результаты для модифицированного метода не приводятся.

SSALRF лучше REGR при не слишком больших σ или при достаточно большом N Исследование, проведенное при различных Nи σ , показало, что нет однозначного ответа на вопрос, какой из методов лучше. Метод SSALRF лучше метода REGR как по смещению, так и

по дисперсии при достаточно малом шуме или при достаточно большой длине ряда.

Проиллюстрируем это на примере исследования ряда с фиксированными T = 6, A = 5. Сначала проведем ряд оценок a_1 и ω при фиксированном N = 60 или, иначе, при N = 10T. Будем изменять σ в интервале [0.25, 5] с шагом 0.25 и считать смещение и дисперсию для \hat{a}_1 и $\hat{\omega}$ на 10000 повторах. Результаты приведены на рис. 1-4.

На этих графиках видно, что и смещение, и дисперсия оценок до какого-то уровня σ меньше для SSALRF, а, значит, при таких σ метод SSALRF лучше, чем REGR. Заметим, что область, в которой SSALRF лучше, довольно велика: $\sigma \leq 4.5$ для всех показателей кроме смещения \hat{a}_1 и $\sigma \lesssim 3.25$ для него. Для того чтобы продемонстрировать, что такой порядок шума сравнительно велик, приведем на рис. 10 график зашумленного гармонического ряда с $\sigma = 4$ (все остальные параметры – такие же, как ранее: T = 6, A = 5, N = 60).

Приведем также графики (рис. 6-9), отражающие результаты, полученные при фиксированном $\sigma = 3.5$ для N = 12...62, N увеличивалось с шагом 5 (количество повторов равно 10000). Поскольку увеличение N соответствует уменьшению σ (в смысле влияния на результаты), то графики выглядят подобно приведенным выше графикам зависимости от σ , с точностью до замены направления оси абсцисс.

Рассмотрим теперь график дисперсии $\hat{\omega}$, которая была вычислена при большем N, N = 90 (рис. 5). Видно, что при таком N метод SSALRF лучше REGR в большей области, при $\sigma \leq 5.5$. Таким образом, увеличение длины ряда существенно улучшает качество метода SSALRF, практически не влияя на качество REGR. Это дает методу SSALRF преимущество в случае, когда возможно увеличить количество наблюдений.

Замечание 3.5. Во время оценок параметров методом SSALRF и его модифицированным вариантом ожидалось, что качество будет лучше при L = N/2, делящемся на T, но исследование показало, что если такая зависимость и есть, то очень незначительная. В общем случае заботиться о том, чтобы L делилось на T, необязательно.

Метод SSALRF, основанный на подходе "Гусеница"-SSA, показал при сравнении результаты лучшие, чем метод REGR, в достаточно большой области параметров. Этого следовало ожидать, так как метод REGR больше подходит для AR-моделей. Резкое же ухудшение качества метода SSALRF при достижении некоторого порога отношения сигнал-шум связано с тем, что при таком уровне шума подход "Гусеница"-SSA перестает работать. Это не удивительно, ведь он не настроен на работу с рядами определенной модели, единственная информация о ряде, которая используется в методе "Тусеница"-SSA на этапе разложения – это его ранг, в то время как метод REGR представляет собой оценку параметров заданной модели. Зависимость качества метода SSALRF от изменения длины ряда, описанная выше, также объясняется свойствами "Гусеница"-SSA.

Список литературы

- [1] Голяндина, Н.Э., (2003), Memod "Гусеница"-SSA: анализ временных рядов: Учеб. пособие, Санкт-Петербург: Изд-во СПбГУ, 87 с.
- [2] Golyandina, N.E., Nekrutkin, V.V., Zhigljavsky, A.A., (2001), Analysis of Time Series Structure. SSA and Related Techniques, Chapmap & Hall/CRC.
- [3] Данилов, Д.Л., Жиглявский, А.А., (1997), Главные компоненты временных рядов: метод "Гусеница", С.-Петербург: Пресском, 308 с.
- [4] Cadzow, J.A., (1988), Signal Enhancement A Composite Property Mapping Algorithm // IEEE Transactions on Acoustics, Speech and Signal Processing. - v. 36. - p. 49-62.



Рис. 1: Модуль смещения \hat{a}_1 (σ по Рис. 2: Дисперсия \hat{a}_1 (σ по оси абсоси абсцисс) цисс)



Рис. 3: Модуль смещения $\hat{\omega}$ (σ по Рис. 4: Дисперсия $\hat{\omega}$ (σ по оси абсоси абсцисс) цисс)



Рис. 5: Дисперсия $\hat{\omega}$ (σ по оси абсцисс), N=90



Рис. 6: Модуль смещения \hat{a}_1 (*N* по Рис. 7: Дисперсия \hat{a}_1 (*N* по оси абсоси абсцисс) цисс)



Рис. 8: Модуль смещения $\hat{\omega}$ (*N* по Рис. 9: Дисперсия $\hat{\omega}$ (*N* по оси абсоси абсцисс) цисс)



Рис. 10: Гармонический ряд (жирная черная линия) и он же зашумленный с $\sigma=4$ (тонкая линия с пунктами)