УДК 519.216.3 EDN ZRCKRE
https://www.doi.org/10.47813/nto.5.2024.1001

Исследование гессиана объединения прогнозов моделей временных рядов, минимизирующего оценку дисперсии прогноза

Д.А. Петрусевич*, М.Е. Савельев

МИРЭА – Российский Технологический Университет, пр. Вернадского, 78, Москва, 119454, Россия

*E-mail: petrdenis@mail.ru

Аннотация. В работе исследуется структура гессиана функции дисперсии оценки прогноза для моделей ARIMA и ETS. Показано, что в некоторых случаях гессиан равен 0 и охарактеризовать поведение экстремума стандартными способами не получается. Построен гессиан функции дисперсии оценки прогноза при оптимальном взвешенном объединении двух прогнозов (с условием минимизации оценки дисперсии прогноза). Для построенных комбинаций прогнозов гессианы уже не обращаются в ноль в экстремумах за исключением вырожденных случаев. Результаты проверены на подмножестве моделей ARIMA, ETS. Показано, что в окрестности точек экстремумов гессиан при объединении прогнозов принимает диагональный вид, а коэффициенты в блоках, соответствующих побочной диагонали, имеют второй порядок малости по параметрам объединяемых моделей.

Ключевые слова: взвешенное объединение моделей, минимизация оценки дисперсии прогноза, гессиан функции оценки дисперсии прогноза.

Analysis of the Hessian of time series models’ combination minimizing estimate of forecast variance

D. Petrusevich*, M. Savelev

MIREA – Russian Technological University, Prospekt Vernadskogo, 78, Moscow, 119454, Russia

*E-mail: petrdenis@mail.ru

Abstract. In the paper structure of the Hessian of the forecast variance function is investigated in case of ARIMA and ETS models. It is shown that in case of certian models the Hessian is equal to 0 and it is impossible to characterize the behavior of the extremum using standard methods. The Hessian of the forecast variance function has been constructed for the optimal weighted combination of two forecasts (aiming to minimize estimate of forecast variance). For the constructed combinations of forecasts, the Hessians aren't equal to zero at the extrema, except in degenerate cases. The results are tested on a subset of ARIMA, ETS models. It is shown that in the vicinity of the extremum points, the Hessian takes a diagonal form, and the coefficients in the blocks corresponding to the side diagonal are of the second order of smallness in terms of the parameters of the combined models.

Keywords: optimal weighted combination, forecast variance estimate minimization, Hessian of forecast variance estimate.

1. Введение

В представленной работе рассматривается построение объединения пары моделей временных рядов в соответствии с принципом минимизации оценки дисперсии прогноза. При работе с моделями ARIMA и ETS [1] дисперсию прогноза можно выразить через параметры самой модели, что делает возможным аналитически связать ширину доверительного интервала и параметры модели. В большинстве других моделей для оценки ширины доверительного интервала используется критерий Стьюдента, подобную связь обнаружить тяжело. В рамках работ [2, 3] было произведено построение взвешенной комбинации двух моделей, минимизирующей оценку дисперсии прогноза. Вместе с тем, было показано, что для простых моделей ARIMA [1] гессиан оценки функции дисперсии прогноза обращается в ноль, что заставляет прибегать к нестандартным методам анализа экстремумов. В представленной работе приведён анализ гессиана функции оценки дисперсии прогноза при взвешенном объединении прогнозов моделей. Показано, что он не обращается в ноль за исключением вырожденных случаев.

2. Постановка задачи

Для того, чтобы построить объединение прогнозов двух моделей, минимизирующее оценку дисперсии прогноза, надо задать аналитически эту оценку. По теореме Вальда [1, 4] модель ARIMA можно представить в виде бесконечной суммы (в общем случае) элементов скользящего среднего MA(q) [1]. Коэффициенты этого разложения (ψ -веса) используются для оценки дисперсии прогноза.

Для оценки дисперсии прогноза модели авторегрессии p-го порядка AR(p) [1] имеет следующие рекуррентные соотношения:

$$
\begin{aligned}
&\psi_0 = 1,\\
&\psi_1 = \varphi_1,\\
&\psi_2 = \varphi_1\psi_1 + \varphi_2,\\
&\psi_3 = \varphi_1\psi_2 + \varphi_2\psi_1 + \varphi_3, \ldots\\
&\psi_k = \sum_{i=1}^{k} \varphi_i \psi_{k-i}.
\end{aligned}
$$

В случае процесса авторегрессии AR(1) $X_t = c + \varphi_1 X_{t-1}$, как показано в [2, 3], оценка дисперсии прогноза на n шагов вперёд имеет особенно простой и красивый вид геометрической прогрессии:

$$Var_{AR(1)}(\hat{x}_n - x_n) = \sigma^2 \sum_{j=0}^{n-1} \psi_j{}^2 = \sigma^2 (1 + \sum_{j=1}^{n-1} \varphi_1{}^{2j}) = \sigma^2 \frac{1-\varphi_1{}^{2n}}{1-\varphi_1}.$$

При этом сама функция оценки дисперсии прогноза является выпуклой вниз.

Для модели ARIMA(p, d, q) в выражении пси-весов появляются параметры скользящего среднего:

$$\begin{aligned}
&\psi_0 = 1,\\
&\psi_1 = \varphi_1 + \theta_1,\\
&\psi_2 = \varphi_1\psi_1 + \theta_2 = \varphi_1(\varphi_1 + \theta_1) + \theta_2,\\
&\psi_{q+1} = \varphi_1\psi_q + \theta_{q+1}.
\end{aligned}$$

В случае модели ARIMA(1, d, q) геометрическая прогрессия начинается только с q-го элемента (следует учитывать, что количество периодов, на который делается прогноз n, может быть меньше порядка q, - в таком случае геометрическая прогрессия не наблюдается):

$$Var_{ARMA(1,q)}(\hat{x}_n - x_n) = \sigma^2 \left[1 + (\varphi_1 + \theta_1)^2 + (\varphi_1(\varphi_1 + \theta_1) + \theta_2)^2 + \ldots + \psi_{q-1}{}^2 + \psi_q{}^2 \frac{1-\varphi_1{}^{2n}}{1-\varphi_1} \right].$$

Интересно, что сама функция оценки дисперсии прогноза обладает в случае моделей ARMA(1, 1), ARMA(1, 2) нулевым гессианом (т.к. от порядка интегрирования d наши рассуждения не зависят, иногда опускаем этот параметр и получаем модель ARMA(p, q)). Подробности вычислений представлены в работах [2, 3], приведём пример вычислений для простейшей модели ARMA(1, 1) при прогнозировании на два шага вперёд. Исследуемая функция имеет вид:

$$\begin{aligned}
&Var_{ARMA(1,1)}(\hat{x}_1 - x_1) = \sigma^2,\\
&Var_{ARMA(1,1)}(\hat{x}_2 - x_2) = \sigma^2 \left[1 + (\varphi_1 + \theta_1)^2 \right].
\end{aligned}$$

Вторые производные этой функции выглядят так:

$$(Var_{ARMA(1,1)}(\hat{x}_2 - x_2))''|_{\varphi_1^2} = (Var_{ARMA(1,1)}(\hat{x}_2 - x_2))''|_{\theta_1^2} = (Var_{ARMA(1,1)}(\hat{x}_2 - x_2))''|_{\varphi_1\theta_1} = 2\sigma^2.$$

Из-за наличия одинаковых строк гессиан функции оценки дисперсии прогноза нулевой:

$$H_{ARMA(1,1)}(\hat{x}_2 - x_2) = \begin{vmatrix} 2\sigma^2 & 2\sigma^2 \\ 2\sigma^2 & 2\sigma^2 \end{vmatrix} = 0.$$

Одинаковые или нулевые строки наблюдаются везде или в точке экстремума и для гессианов оценки дисперсии прогноза моделей ARIMA(1, d, 2) на 2-3 шага вперёд [2, 3].

2.1. Гессиан взвешенного объединения прогнозов

Далее рассмотрим объединение нескольких моделей (или их прогнозов) в комбинацию по какому-либо принципу [5]. Дальнейшее изложение следует работе [6]. Рассмотрим линейную комбинацию прогнозов f с весами w: $wf_1 + (1-w)f_2$. Дисперсия оценки комбинации прогнозов имеет вид:

$$Var(\sigma_1, \sigma_2, \rho, w) = (w\sigma_1)^2 + ((1-w)\sigma_2)^2 + 2w(1-w)\rho\sigma_1\sigma_2.$$

Здесь ρ - корреляция объединяемых прогнозов. В работе [6] продемонстрировано, что оптимальные веса имеют вид:

$$w = \frac{\sigma_2^2 - \rho\sigma_1\sigma_2}{\sigma_1^2 + \sigma_2^2 - 2\rho\sigma_1\sigma_2},$$

$$1 - w = \frac{\sigma_1^2 - \rho\sigma_1\sigma_2}{\sigma_1^2 + \sigma_2^2 - 2\rho\sigma_1\sigma_2}.$$

Этот результат не зависит от типов моделей, чьи прогнозы объединяются. Анализ полученной функции представлен в [6].

3. Исследование гессиана взвешенного объединения прогнозов моделей, минимизирующего оценку дисперсии прогноза

Рассмотрим гессиан функции оценки дисперсии прогноза при оптимальном объединении двух моделей. В работе [6] показано, что вторые производные функции имеют вид (здесь x отмечает производные по параметрам первой модели в объединении, y – второй):

$$Var(\sigma_1, \sigma_2, \rho)''_{x_1x_2} = \frac{\sigma_2^3(1-\rho^2)}{(\sigma_1^2 + \sigma_2^2 - 2\rho\sigma_1\sigma_2)^2}\left[(\sigma_2 - \rho\sigma_1)(\sigma_1^2)''_{x_1x_2} - (\sigma_1^2)'_{x_1}(\sigma_1)'_{x_2}\frac{-3\rho(\sigma_1^2+\sigma_2^2) + 2(\rho^2+2)\sigma_1\sigma_2}{\sigma_1^2 + \sigma_2^2 - 2\rho\sigma_1\sigma_2}\right],$$

$$Var(\sigma_1, \sigma_2, \rho)''_{xy} = \sigma_2^2(1-\rho^2)(\sigma_1^2)'_x(\sigma_2)'_y\frac{-3\rho(\sigma_1^2+\sigma_2^2) + 2(\rho^2+2)\sigma_1\sigma_2}{(\sigma_1^2 + \sigma_2^2 - 2\rho\sigma_1\sigma_2)^3}.$$

Непосредственно в точке экстремума $(\sigma_1^2)'_{x_1} = (\sigma_1^2)'_{x_2} = (\sigma_2^2)'_y = 0$, поэтому второе слагаемое производной по параметрам первой модели $Var(\sigma_1, \sigma_2, \rho)''_{x_1x_2}$ и производная $Var(\sigma_1, \sigma_2, \rho)''_{xy}$ обращаются в ноль. Но в отличие от функции оценки дисперсии прогноза отдельных моделей, целиком гессиан в ноль не обращается, приобретая диагональный вид, т.к. в смешанной производной $Var(\sigma_1,\sigma_2,\rho)''_{x_1x_2}$ в точке экстремума остаётся ненулевым первое слагаемое:

$$Var(\sigma_1,\sigma_2,\rho)''_{x_1x_2}=\sigma_2^{\,3}(1-\rho^2)\frac{\sigma_2-\rho\sigma_1}{[\sigma_1^{\,2}+\sigma_2^{\,2}-2\rho\sigma_1\sigma_2]^2}(\sigma_1^{\,2})''_{x_1x_2},$$
$$Var(\sigma_1,\sigma_2,\rho)''_{xy}=0.$$

Таким образом, например, для двух моделей AR(1), объединяемых с минимизацией оценки дисперсии прогноза гессиан функции оценки дисперсии комбинации моделей имеет вид:

$$\sigma_1^{\,2}=se_1^{\,2}(1+\varphi_1^{\,2}),(\sigma_1^{\,2})''_{\varphi_{11}}=2se_1^{\,2},$$
$$\sigma_2^{\,2}=se_2^{\,2}(1+\varphi_2^{\,2}),(\sigma_2^{\,2})''_{\varphi_{21}}=2se_2^{\,2},$$
$$H=\begin{vmatrix}2se_1^{\,2}\sigma_2^{\,3}(1-\rho^2)\dfrac{\sigma_2-\rho\sigma_1}{[\sigma_1^{\,2}+\sigma_2^{\,2}-2\rho\sigma_1\sigma_2]^2} & 0\\ 0 & 2se_2^{\,2}\sigma_1^{\,3}(1-\rho^2)\dfrac{\sigma_1-\rho\sigma_2}{[\sigma_1^{\,2}+\sigma_2^{\,2}-2\rho\sigma_1\sigma_2]^2}\end{vmatrix}.$$

Здесь se – стандартные ошибки объединяемых моделей, которые оцениваются во время обучения, φ - коэффициенты моделей.

Для моделей ETS(A, N, N) [1, 7] результат выглядит аналогичным образом (с условием замены обозначений параметров моделей $\varphi\leftrightarrow\alpha$).

В [6] приведён гессиан для моделей ETS(A, A, N), обладающих бОльшим количеством параметров. В целом, он сохраняет блочный диагональный вид. Интересно оценить порядок элементов, которые находятся в блоках на побочной диагонали, а также, в целом, добавок, которые обращаются в ноль в точке экстремума (экстремум достигается при $(\sigma_1^{\,2})'_x=(\sigma_2^{\,2})'_y=0$. Например, в случае модели ETS(A, N, N) $(\sigma_1^{\,2})'_{\alpha_1}=(\sigma_2^{\,2})'_{\alpha_2}=0,\alpha_1=\alpha_2=0$ и можно оценить множители в выражениях производных второго порядка с точностью до бесконечно малых второго порядка:

$$\sigma_1^{\,2}=se_1^{\,2}(1+\alpha_1^{\,2}),\sigma_1=se_1\sqrt{1+\alpha_1^{\,2}}\sim se_1(1+\alpha_1^{\,2}/2),(\sigma_1^{\,2})'_{\alpha_1}\sim\alpha_1,(\sigma_1)'_{\alpha_1}\sim\alpha_1,$$
$$\sigma_2^{\,2}=se_2^{\,2}(1+\alpha_2^{\,2}),\sigma_2=se_2\sqrt{1+\alpha_2^{\,2}}\sim se_2(1+\alpha_2^{\,2}/2),(\sigma_2^{\,2})'_{\alpha_2}\sim\alpha_2,(\sigma_2)'_{\alpha_2}\sim\alpha_2$$
$$(\sigma_1^{\,2}+\sigma_2^{\,2}-2\rho\sigma_1\sigma_2)^{-n}\sim k_0+k_1\alpha_1^{\,2}+k_2\alpha_2^{\,2},$$
$$\sigma_1\sigma_2=se_1\sqrt{1+\alpha_1^{\,2}}se_2\sqrt{1+\alpha_2^{\,2}}\sim se_1se_2(1+\alpha_1^{\,2}/2+\alpha_2^{\,2}/2).$$

Здесь k – некоторые коэффициенты, зависящие от параметров моделей. Первое слагаемое производной $Var(\sigma_1, \sigma_2, \rho)''_{\alpha_1\alpha_2}$ имеет нулевой порядок малости, второе – второй порядок малости за счёт произведения $(\sigma_1^2)'_{\alpha_1}(\sigma_1)'_{\alpha_1} \sim \alpha_1^2$. Таким образом, в блоках, расположенных на главной диагонали, добавляется величина второго порядка малости по коэффициентам моделей. Что же элементов в блоках побочной диагонали, зависящих от смешанных производных, то $(\sigma_1^2)'_{\alpha_1}(\sigma_2)'_{\alpha_2} \sim \alpha_1\alpha_2$. На побочной диагонали находятся величины, которые в окрестности экстремумов имеют второй порядок малости по коэффициентам моделей. Хотя этот вывод для наглядности сделан в терминах параметров моделей ETS, но, т.к. той же структурой оценки дисперсии прогноза модели обладает большинство часто применяющихся инструментов, его можно распространить и на них.

4. Полученные результаты

Приведём пример объединения моделей ETS [1, 7] при обработке данных индекса потребительских цен в период 1991 – 2022 гг. [8]. В качестве моделей, которые можно объединить используются различные версии моделей ETS (см. таблицу 1). Для сравнения даны характеристики модели ARIMA на том же периоде обучения и тестовой выборке.

Таблица 1. Модели временных рядов ETS, ARIMA для процесса месячных показателей индекса потребительских цен РФ [8].

Модель	RMSE	MAE	Оценка дисперсии модели
ETS(A,A,A)	0.35	0.27	0.59
ETS(A,A_d,A)	0.35	0.26	0.62
ETS(A,A,N)	0.50	0.36	1.60
ETS(A,A_d,N)	0.50	0.36	1.51
ARIMA	0.34	0.28	-

Пары моделей, для которых с определёнными весами объединяются прогнозы, указаны в таблице 2. Дана корреляция между прогнозами на тестовый период, от которой, вкупе с дисперсиями моделей, и зависит вес в объединении.

Таблица 2. Объединения прогнозов моделей временных рядов ET S для процесса месячных показателей индекса потребительских цен РФ [8].

Модель	Корреляция прогнозов	RMSE	MAE	Оценка дисперсии объединённого прогноза
ETS(A,A,A) + ETS(A,A_d,A)	0.77	0.45	0.40	0.54
ETS(A,A_d,N) + ETS(A,A_d,A)	-0.10	0.49	0.43	0.40
ETS(A,A_d,N) + ETS(A,A,A)	**0.20**	**0.28**	**0.22**	**0.50**
ETS(A,A,N) + ETS(A,A_d,A)	-0.11	0.51	0.45	0.40
ETS(A,A,N) + ETS(A,A,A)	**0.19**	**0.29**	**0.23**	**0.50**

Лучшие объединённые прогнозы в таблице 2 отмечены жирным шрифтом. Отметим, что они превосходят по качеству прогноза как модель ARIMA, так и модели ETS, которые участвовали в объединении.

5. Выводы

В работах [2, 3] построены и исследуются взвешенные линейные комбинации моделей, в которых веса подбираются на основе минимизации теоретической оценки дисперсии прогноза. В представленной работе рассмотрен гессиан функции оценки дисперсии прогноза в стационарных точках при таком объединении. Показано, что, в отличие от гессианов функций прогнозов AR(1), ARMA(1, q) и ETS(A, N, N), которые обращались в 0, рассмотренная функция для объединённого прогноза в общем случае в 0 не обращается, за исключением некоторых вырожденных случаев.

В дополнение гессиан функции оценки дисперсии объединённого прогноза рассмотрен в окрестности экстремума. Показано, что он имеет блочную структуру. При этом, в экстремуме блоки, расположенные на главной диагонали, ненулевые. Блоки на побочной диагонали – нулевые, в малой окрестности экстремума блоки на главной диагонали получают добавки порядка x^2 (x – коэффициент одной из объединяемых моделей), на побочной диагонали – порядка произведения xy (x – коэффициент первой из пары объединяемых моделей, y – коэффициент из второй модели).

Список литературы

1. Hyndman R. J. Forecasting: principles and practice / R. J. Hyndman, G. Athanasopoulos. – Melbourne: OTexts, 2021.
2. Beletskaya N. V. Linear combinations of time series models with minimal forecast variance / N. V. Beletskaya, D. A. Petrusevich // Journal of communications technology and electronics. – 2022. – № 67(1). – C. 144-158.
3. Musatov D. Modeling of forecasts variance reduction at multiple time series prediction averaging with ARMA (1, q) functions / D. Musatov, D. Petrusevich // V International Workshop on Modeling, Information Processing and Computing (MIP: Computing-V 2022) / Eugene Semenkin, Igor Kovalev (eds.). – Krasnoyarsk, Russia, 25 January 2022. – 2022.
4. Wald H. A study in the analysis of stationary time series / H. Wald. – Uppsala: Almqvist and Wiksell Book Co., 1954.
5. Hyndman R. J. Optimal combination forecasts for hierarchical time series / R. J. Hyndman, R. A. Ahmed, G. Athanasopoulos, H. L. Shang // Computational Statistics & Data Analysis. – 2011. - №55.9. – c. 2579-2589.
6. Белецкая Н. В. Минимизация оценки дисперсии прогноза на примере моделей ETS / Н. В. Белецкая, Д. А. Петрусевич // Информационные процессы. – 2024. – № 24(1). – С. 16-29.
7. Stock J. H. Introduction to Econometrics / J. H. Stock, M. W. Watson. – Pearson, 2021.
8. Единый архив экономических и социологических данных. Динамические ряды макроэкономической статистики РФ. Индекс потребительских цен. – URL: http://sophist.hse.ru/hse/nindex.shtml (дата обращения: 02.09.2024)