

ПРЕДСКАЗАНИЕ ВРЕМЕННЫХ РЯДОВ НА ОСНОВЕ МОДЕЛЕЙ В ПРОСТРАНСТВЕ СОСТОЯНИЙ И ФИЛЬТРА КАЛМАНА

М. И. Малафеева, С.В. Лобач

malafeevamaria02@gmail.com, fpm.lobach@gmail.com;

Научный руководитель – С. В. Лобач, старший преподаватель

В статье исследуется применимость моделей в пространстве состояний и фильтра Калмана для предсказания и восстановления пропущенных значений временных рядов. Для решения используются классические формулы фильтра Калмана, а также рекуррентные формулы вычисления моментов.

Ключевые слова: модели в пространстве состояний; прогнозирование временных рядов; фильтр калмана, моделирование; временные ряды с пропусками.

ВВЕДЕНИЕ

Многие математические модели временных рядов могут быть сведены к моделям в пространстве состояний. В свою очередь модели в пространстве состояний позволяют применить к исходной модели временного ряда широкий спектр процедур, включая столь актуальные, как оценивание параметров и прогнозирование [1].

ФИЛЬТР КАЛМАНА НА ПРАКТИЧЕСКОМ ПРИМЕРЕ

Рассмотрим практический пример.

Модель ARIMA(4, 1, 1):

$$y_t = y_{t-4} + \varepsilon_t + \varepsilon_{t-1}, t = 0, 1, \dots, T$$

$$\varepsilon_t = N(0, 1), y_0 = 1$$

Данную модель можно представить в форме модели в пространстве состояний:

$$\begin{cases} y_t = (1 \ 0 \ 0 \ 0)x_t \\ x_{t+1} = \begin{pmatrix} 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \\ 1 & 0 & 0 & 0 \end{pmatrix} x_t + \begin{pmatrix} 1 \\ 1 \\ 0 \\ 0 \end{pmatrix} \varepsilon_t \end{cases}, t = 0, 1, \dots, T, T = 100$$

$$\varepsilon_t = N(0, 1), y_0 = 1.$$

Требуется найти оценки $\hat{x}_t, \hat{y}_t$, используя фильтр Калмана.

Для нахождения оценок будем пользоваться следующими формулами фильтра Калмана [2]:

$$x_{t|t-1} = Tx_{t-1},$$

$$P_{t|t-1} = TP_{t-1}T' + RQR',$$

$$x_t = x_{t|t-1} + K_tv_t,$$

$$P_t = [I - K_tZ]P_{t|t-1},$$

$$v_t = y_t - Zx_{t|t-1},$$

$$K_t = P_{t|t-1}Z'(ZP_{t|t-1}Z' + H)^{-1},$$

Где

$$\begin{cases} y_t = Z_tx_t, \\ x_{t+1} = T_tx_t + R_t\varepsilon_t \end{cases}, \quad t = 0, 1, \dots, T, \quad T = 100,$$

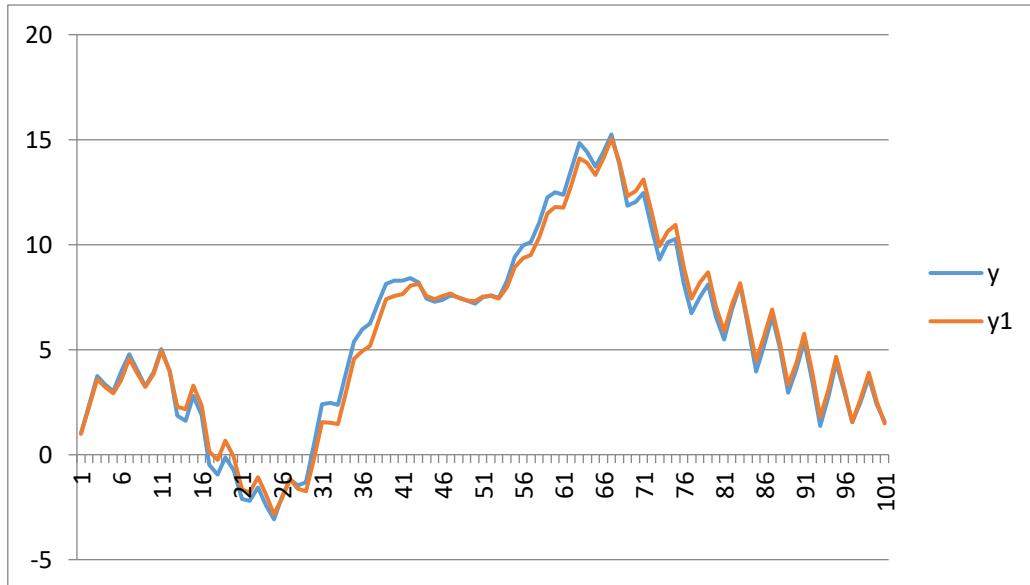
$$E(\varepsilon_{t-1}\varepsilon_{\tau-1}) = \begin{cases} Q, & t = \tau \\ 0, & t \neq \tau \end{cases},$$

$$E(\varepsilon_t\varepsilon_\tau) = \begin{cases} H, & t = \tau \\ 0, & t \neq \tau \end{cases}.$$

В нашем случае: $Z = (1 \ 0 \ 0 \ 0)$, $T = \begin{pmatrix} 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \\ 1 & 0 & 0 & 0 \end{pmatrix}$, $R = \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix}$.

Зададим: $H = 1$, $Q = 1$, $x_0 = \begin{pmatrix} 1 \\ 1 \\ 1 \\ 1 \end{pmatrix}$, $P_0 = \begin{pmatrix} 16 & 0 & 0 & 0 \\ 0 & 16 & 0 & 0 \\ 0 & 0 & 16 & 0 \\ 0 & 0 & 0 & 16 \end{pmatrix}$.

На рисунке представлены графики исходного временного ряда y и полученного в результате фильтрации ряда оценок y_1 .



Графики исходного временного ряда y и ряда оценок y_1

ПРОГНОЗИРОВАНИЕ НА ОСНОВЕ МОДЕЛЕЙ В ПРОСТРАНСТВЕ СОСТОЯНИЙ

Рассмотрим стационарный в широком смысле процесс со спектральной плотностью [3]:

$$\hat{f}(\lambda) = \frac{|e^{i\lambda} + 1|^2}{|e^{2i\lambda} + \frac{1}{2}e^{i\lambda} + \frac{1}{2}|^2}.$$

Построим гауссовский процесс со спектральной плотностью. Такой процесс может быть получен как решение уравнения:

$$\xi_{t+2} + \frac{1}{2}(\xi_{t+1} + \xi_t) = \varepsilon(t+2) + \varepsilon(t+1).$$

Введём вспомогательную переменную θ и составим систему:

$$\begin{cases} \theta_{t+1} = -\frac{1}{2}\theta_t - \frac{1}{2}\xi_t + \frac{1}{2}\varepsilon(t+1) \\ \xi_{t+1} = \theta_t + \varepsilon(t+1). \end{cases}$$

Далее смоделируем два вектора θ и ξ при заданных начальных условиях $\theta_0 = 1, \xi_0 = 1$.

Для задачи прогнозирования строим дополнительные вектора n_1 и n_2 (n_1 - для θ , n_2 - для ξ) по формулам :

$$\begin{cases} n_1(t+1, s) = -\frac{1}{2}n_1(t, s) - \frac{1}{2}n_2(t, s) \\ n_2(t+1, s) = n_1(t, s) \\ n_1(s, s) = m_s, n_2(s, s) = \xi_s \end{cases}.$$

Для того, чтобы начать строить прогноз с момента времени s , нужно знать начальное значение $n_1(s, s)$. Для этого построим ещё одну последовательность m :

$$\begin{cases} m_{s+1} = -\frac{1}{2}m_s - \frac{1}{2}\xi_s + \frac{1-\gamma_s}{2(1+\gamma_s)}(\xi_{s+1} - m_s) \\ \gamma_{s+1} = \frac{\gamma_s}{1+\gamma_s} \\ m_0 = 0, \gamma_0 = 1. \end{cases}$$

Указанные выше формулы можно также применить для восстановления пропущенных значений временного ряда. Рассматриваются три варианта пропусков: расположенные в начале, в середине и в конце временного ряда, а также детерминированный и случайный шаблоны пропусков.

Результаты представлены в таблицах 1, 2, 3.

Таблица 1

Ретроспективный прогноз с момента $s = 90$

№	θ	n_1	ξ	n_2
91	0.467	0.425	0.233	0.156
92	-0.247	-0.290	0.673	0.425
93	-0.331	-0.067	-0.483	-0.290

Таблица 2

Восстановление пропущенных значений ($s = 70$)

№	θ	n_1	ξ	n_2
71	0.243	-0.284	0.969	-0.088
72	-0.331	0.186	0.793	-0.284
73	-0.003	0.049	0.125	0.186

Таблица 3

Восстановление пропущенных значений ($s = 50$)

№	θ	n_1	ξ	n_2
51	-0.509	1.053	-3.069	0.123
52	0.382	-0.588	-3.322	1.053
53	0.205	-0.232	-2.148	-0.588

Результаты, приведенные в таблице, показывают, что наименее точными оценками пропущенных наблюдений, а также прогноз получается в случае, когда пропуски находятся в начале временного ряда, а более точные, когда пропуски находятся в конце временного ряда.

Библиографические ссылки

1. *В.И. Лобач, Р. И. Меркулов.* Прогнозирование временных рядов с пропусками на основе моделей в пространстве состояний.. Режим доступа: https://libeldoc.bsuir.by/bitstream/123456789/27683/1/Merkulov_Prognozirovaniye.PDF.
2. James Durbin. State Space and Unobserved Component Models: Theory and Applications. 2004. 394 с.
3. *Липцер Р.Ш., Ширяев А.Н.,* Статистика случайных процессов, – Москва: Главная редакция физико-математической литературы изд-ва «Наука», 1974 – 696 с.