

ОПТИМАЛЬНАЯ ФИЛЬТРАЦИЯ И ПРОГНОЗИРОВАНИЕ ЭКОНОМИКО-ПРОИЗВОДСТВЕННЫХ ПРОЦЕССОВ В РЕАЛЬНОМ МАСШТАБЕ ВРЕМЕНИ

Н.Э. Хомякова

(представил д.т.н., проф. Ю.В. Стасев)

Рассмотрена задача фильтрации и прогнозирования нестационарных последовательностей, состоящих из полиномиальной и авторегрессионной компонент, наблюдаемых в сумме с дискретным белым шумом. Сформулирована математическая постановка задачи, даны основные соотношения для алгоритма фильтра Калмана. Синтезированный алгоритм использован для обработки балансовой прибыли предприятия.

Введение. Инженерная постановка проблемы анализа динамики экономико-производственных процессов. Экономико-производственное состояние предприятия является одним из главных факторов, который необходимо учитывать при управлении проектами, связанными с данным предприятием, и при управлении производством [1]. В условиях переходной экономики, например, финансово-экономическое состояние предприятия характеризуется значительной нестабильностью.

Анализ литературы. Анализ публикаций по вопросам динамики экономико-производственных процессов [2, 3], предварительная обработка данных о финансово-экономических показателях предприятия позволяют сделать вывод о целесообразности представления динамики таких показателей в виде составной последовательности.

Основные тенденции представляются полиномиальной компонентой. Квазисезонные изменения хорошо аппроксимируются авторегрессионной гауссовской стационарной последовательностью. В реальных условиях к этим двум компонентам следует добавить некоррелированную компоненту, характеризующую погрешности наблюдений.

В [4] рассмотрена задача фильтрации и прогнозирования составных последовательностей по полной выборке данных. Одним из этапов решения этой задачи был этап оценки статистических характеристик авторегрессионной и некоррелированной компонент входных данных. Стационарность этих компонент позволяет использовать данные статистические характеристики на последующих интервалах времени наблюдения.

В процессе управления реализацией проектов появляется потребность в разработке алгоритма обработки данных (фильтрации и прогнозирования), работающего по выборке данных нарастающего объема, то есть в реальном масштабе времени.

Ниже дано решение этой задачи с использованием метода динамической фильтрации Калмана, позволяющего найти оптимальные оценки процессов по критерию минимума среднеквадратической погрешности.

Метод обработки данных в реальном масштабе времени широко используется в радиоэлектронных комплексах [5] при оценке траекторий и управлении движением аэрокосмических аппаратов, однако практически не применяется при обработке экономико-производственных процессов.

Цель статьи. Целью данной статьи является разработка методики калмановской фильтрации составных последовательностей, а также ее применение для анализа текущего состояния и прогнозирования балансовой прибыли предприятия.

1. Математическая постановка задачи. Рекуррентная модель оцениваемых последовательностей. Наблюдаемые данные представим в виде

$$z_k = \mu_k + \overset{\circ}{\lambda}_k + v_k, \quad (1)$$

где μ_k – полиномиальная компонента; $\overset{\circ}{\lambda}_k$ – авторегрессионная компонента; v_k – некоррелированная компонента; $k=1, 2, \dots, M$ – дискретное время наблюдений.

Ограничимся рассмотрением полинома второй степени и авторегрессии второго порядка. В последующем эти ограничения можно снять, однако предварительная обработка различных финансово-экономических процессов показывает, что именно такая модель представляется наиболее целесообразной. Предварительная обработка свидетельствует также о гауссовском распределении плотности вероятности последовательностей $\overset{\circ}{\lambda}_k, v_k$.

Для полиномиальной компоненты на интервале времени $t \in (t_{k-1}, t_k)$ запишем отрезок ряда Тейлора

$$\mu(t) = C_{0,k-1} + C_{1,k-1}(t - t_{k-1}) + \frac{1}{2}C_{2,k-1}(t - t_{k-1})^2.$$

При $t = t_k$ имеем

$$\mu_k = C_{0,k} = C_{0,k-1} + C_{1,k-1}\Delta t + \frac{1}{2}C_{2,k-1}\Delta t^2,$$

где

$$\Delta t = t_k - t_{k-1} = \text{Const}(k).$$

Введем вектор

$$\bar{C}_k^T = \{C_{2k}, C_{1k}, C_{0k}\}$$

и переходную матрицу

$$\underline{\Phi}_1 = \begin{bmatrix} 1 & 0 & 0 \\ \Delta t & 1 & 0 \\ \Delta t^2/2 & \Delta t & 1 \end{bmatrix}.$$

При этом можно записать рекуррентное соотношение

$$\bar{C}_k = \underline{\Phi}_1 \cdot \bar{C}_{k-1}.$$

В скалярном виде имеем:

$$\begin{cases} C_{2,k} = C_{2,k-1}; \\ C_{1,k} = C_{1,k-1} + C_{2,k-1} \cdot \Delta t; \\ C_{0,k} = C_{0,k-1} + C_{1,k-1} \cdot \Delta t + 0.5 C_{2,k-1} \cdot \Delta t^2. \end{cases}$$

Для последовательности авторегрессии второго порядка имеем

$$\overset{\circ}{\lambda}_k = \gamma_1 \cdot \overset{\circ}{\lambda}_{k-1} + \gamma_2 \cdot \overset{\circ}{\lambda}_{k-2} + \varepsilon_k,$$

где ε_k – некоррелированная гауссовская последовательность с нулевым математическим ожиданием $\langle \varepsilon_k \rangle = 0$ и дисперсией $\langle \varepsilon_k^2 \rangle = \sigma_\varepsilon^2$; γ_1, γ_2 – параметры авторегрессии.

Следуя [5], введем вектор $\overset{\circ}{\lambda}_k^{(k)}$ и переходную матрицу $\underline{\Phi}_2$:

$$\overset{\circ}{\lambda}_k^{(k)} = \begin{pmatrix} \overset{\circ}{\lambda}_k^{(k)} \\ \lambda_k \\ \overset{\circ}{\lambda}_{k-1}^{(k)} \end{pmatrix}; \quad \underline{\Phi}_2 = \begin{bmatrix} \gamma_1 & \gamma_2 \\ 1 & 0 \end{bmatrix}.$$

Верхний индекс показывает текущий момент времени. Нижние индексы показывают позиции компонент вектора.

Теперь можно записать рекуррентное соотношение

$$\overset{\circ}{\lambda}_k^{(k)} = \underline{\Phi}_2 \cdot \overset{\circ}{\lambda}_{k-1}^{(k-1)} + \bar{\varepsilon}_k,$$

где

$$\bar{\varepsilon}_k = \begin{pmatrix} \varepsilon_k \\ 0 \end{pmatrix}.$$

В скалярной форме получим:

$$\begin{cases} \overset{\circ}{\lambda}_k = \gamma_1 \cdot \overset{\circ}{\lambda}_{k-1} + \gamma_2 \cdot \overset{\circ}{\lambda}_{k-2} + \varepsilon_k; \\ \overset{\circ}{\lambda}_{k-1} = \overset{\circ}{\lambda}_{k-1} . \end{cases}$$

Теперь введем составной вектор $\bar{\lambda}_k$ и переходную матрицу $\underline{\Phi}$:

$$\bar{\lambda}_k = \begin{pmatrix} \bar{C}_k \\ \overset{\circ}{\lambda}_k \\ \lambda_k \\ \overset{\circ}{\lambda}_{k-1} \end{pmatrix} = \begin{pmatrix} C_{2,k} \\ C_{1,k} \\ C_{0,k} \\ \overset{\circ}{\lambda}_k \\ \lambda_k \\ \overset{\circ}{\lambda}_{k-1} \end{pmatrix} ; \quad \underline{\Phi} = \begin{bmatrix} 1 & 0 & 0 & 0 & 0 \\ \Delta t & 1 & 0 & 0 & 0 \\ \Delta t^2/2 & \Delta t & 1 & 0 & 0 \\ 0 & 0 & 0 & \gamma_1 & \gamma_2 \\ 0 & 0 & 0 & 1 & 0 \end{bmatrix} .$$

Если ввести вектор $\bar{\varepsilon}_k^T = (0, 0, 0, \varepsilon_k, 0)$, то можно записать

$$\bar{\lambda}_k = \underline{\Phi} \cdot \bar{\lambda}_{k-1} + \bar{\varepsilon}_k, \quad (2)$$

где $\langle \bar{\varepsilon}_k \cdot \bar{\varepsilon}_k^T \rangle = \underline{\Theta}_\varepsilon$; в матрице $\underline{\Theta}_\varepsilon$ все элементы равны нулю, кроме $\Theta_{\varepsilon_{44}} = \sigma_\varepsilon^2$.

Уравнение наблюдения (1) можно записать в виде

$$u_k = \bar{d}^T \cdot \bar{\lambda}_k + v_k, \quad (3)$$

где $\bar{d}^T = (0, 0, 1, 1, 0)$; v_k – последовательность некоррелированных гауссовских случайных величин с нулевым матожиданием и дисперсией σ_v^2 .

2. Основные соотношения калмановской фильтрации составной последовательности. Оценка $\tilde{\lambda}_k$ находится с использованием выражения

$$\tilde{\lambda}_k = \bar{\lambda}_k^* + \bar{h}_k \left(u_k - \bar{d}^T \cdot \bar{\lambda}_k^* \right), \quad (4)$$

где $\bar{\lambda}_k^* = \underline{\Phi} \cdot \tilde{\lambda}_k$ –

вектор прогнозируемых на один шаг оценок; \bar{h}_k – вектор усиления фильтра Калмана, определяемый соотношением

$$\bar{h}_k = \tilde{\Sigma}_k \frac{\bar{d}}{\sigma_v^2}. \quad (6)$$

Матрица $\tilde{\Sigma}_k$ характеризует точность оценки вектора $\bar{\lambda}_k$. На ее главной диагонали находятся дисперсии погрешностей оценок элемен-

тов вектора $\bar{\lambda}_k$, а вне диагонали – коэффициенты взаимной корреляции погрешностей. Эта матрица определяется выражением

$$\tilde{\Sigma}_k = \left[\begin{array}{c} *^{-1} \\ \Sigma_k + \frac{\bar{d} \cdot \bar{d}^T}{\sigma_v^2} \end{array} \right]^{-1}, \quad (7)$$

где
$$\Sigma_k^* = \Phi \cdot \tilde{\Sigma}_{k-1} \cdot \Phi^T + \Theta_\varepsilon - \quad (8)$$

матрица дисперсий и ковариаций погрешностей прогнозируемых оценок.

Для соотношений (5) и (8) необходимо задать начальные условия.

Специфика рассмотренной задачи (полином второй степени и авто-регрессия второго порядка) требует начать вычислительный процесс при $k = 4$, а первые значения исходных данных u_1, u_2, u_3 использовать для

получения вектора $\bar{\lambda}_3^T = \left\{ \tilde{C}_{2,3}, \tilde{C}_{1,3}, \tilde{C}_{0,3}, \lambda_3^{(3)}, \lambda_2^{(3)} \right\}$ и матрицы $\tilde{\Sigma}_3$.

Для оценок начальных значений коэффициентов полинома при $k = 3$ используем соотношения:

$$\tilde{C}_{0,3} = u_3; \quad \tilde{C}_{1,3} = \frac{u_3 - u_2}{\Delta t}; \quad \tilde{C}_{2,3} = \frac{\frac{u_3 - u_2}{\Delta t} - \frac{u_2 - u_1}{\Delta t}}{\Delta t} = \frac{u_3 - 2u_2 + u_1}{\Delta t^2}.$$

В первом приближении положим $\lambda_3^{(3)} = 0; \lambda_2^{(3)} = 0$.

В качестве $\tilde{\Sigma}_3$ положим
$$\tilde{\Sigma}_3 = \sigma_v^2 \begin{bmatrix} 6/\Delta t^4 & 3/\Delta t^2 & 1/\Delta t^2 & 0 & 0 \\ 3/\Delta t^2 & 2/\Delta t^2 & 1/\Delta t & 0 & 0 \\ 1/\Delta t^2 & 1/\Delta t & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 1 \end{bmatrix}.$$

Структура алгоритма обработки данных приведена на рис. 1.

3. Прогнозирование составной последовательности. Для прогнозирования составной последовательности используются соотношения (5) и (8):

$$\bar{\lambda}_{M+1}^* = \Phi \cdot \tilde{\lambda}_M^*; \quad \bar{\lambda}_{M+2}^* = \Phi \cdot \bar{\lambda}_{M+1}^* \quad \text{и т.д.}$$

Точность прогноза определяется выражениями:

$$\Sigma_{M+1}^* = \Phi \cdot \tilde{\Sigma}_M \cdot \Phi^T + \Theta_\varepsilon; \quad \Sigma_{M+2}^* = \Phi \cdot \Sigma_{M+1}^* \cdot \Phi^T + \Theta_\varepsilon \quad \text{и т.д.}$$

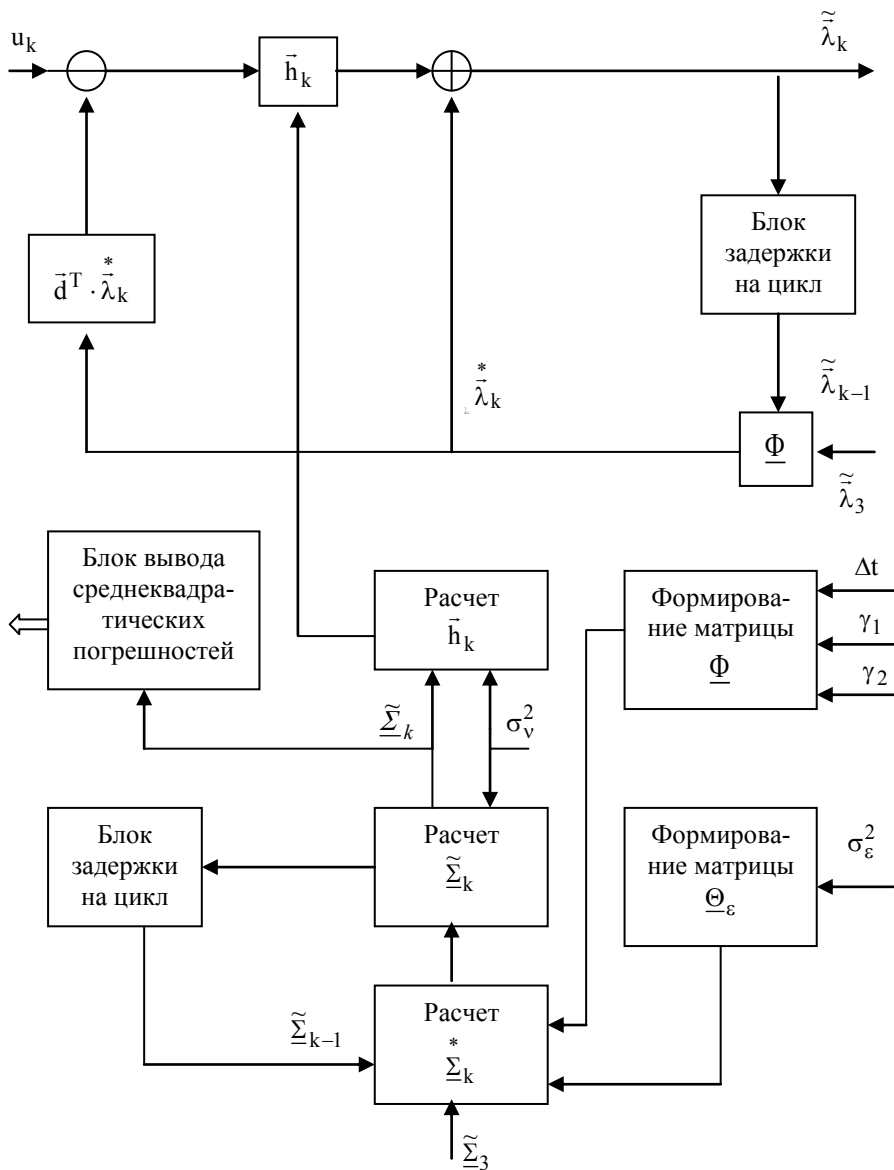


Рис. 1. Структура алгоритма обработки данных при фильтрации составной последовательности

4. Пример обработки данных о балансовой прибыли (БП) предприятия. Исходные данные приведены в табл. 1. Табл. 2 содержит текущие оценки $\tilde{\mu}_k$, $\tilde{\lambda}_k^{(k)}$ ($k=4, 5, \dots, 20$), а также их среднеквадратические погрешности.

Таблица 1

Исходные данные о БП предприятия и статистические характеристики

№ квартала	БП, млн.грн	№ квартала	БП, млн.грн
1	2,33	11	1,0
2	2,27	12	0,667
3	2,13	13	0,361
4	1,87	14	0,267
5	1,125	15	0,3
6	0,533	16	0,533
7	0,467	17	0,946
8	0,533	18	1,33
9	0,636	19	1,47
10	0,8	20	1,53
$\sigma_{\varepsilon}^2 = 0,00314$	$\sigma_v^2 = 0,00764$	$\gamma_1 = 1,21$	$\gamma_2 = -0,8614$

Таблица 2

Результаты фильтрации компонент составной последовательности

k	$\tilde{\mu}_k$	σ_{μ_k}	$\tilde{\lambda}_k^{(k)}$	$\sigma_{\lambda_k}^o$
4	1,88E+00	4,04E-01	-1,40E-02	3,98E-01
5	1,05E+00	2,94E-01	7,65E-02	2,75E-01
6	6,53E-01	2,01E-01	-1,47E-01	1,70E-01
7	6,92E-01	1,94E-01	-2,81E-01	1,62E-01
8	5,92E-01	2,03E-01	-7,25E-02	1,68E-01
9	4,21E-01	1,76E-01	2,15E-01	1,40E-01
10	4,87E-01	1,50E-01	2,91E-01	1,16E-01
11	7,42E-01	1,34E-01	2,17E-01	1,04E-01
12	6,50E-01	1,24E-01	7,48E-02	9,77E-02
13	5,95E-01	1,17E-01	-1,84E-01	9,38E-02
14	6,35E-01	1,12E-01	-3,45E-01	9,09E-02
15	6,45E-01	1,08E-01	-3,01E-01	8,86E-02
16	6,88E-01	1,04E-01	-1,18E-01	8,66E-02
17	8,38E-01	1,00E-01	1,02E-01	8,49E-02
18	1,06E+00	9,71E-02	2,46E-01	8,34E-02
19	1,26E+00	9,42E-02	2,14E-01	8,21E-02
20	1,50E+00	9,15E-02	3,87E-02	8,09E-02

В табл. 3 приведены результаты прогноза на 5 кварталов вперед. Здесь же приведены среднеквадратические погрешности прогнозирования обеих компонент последовательности. Табл. 4 содержит для моментов времени $k = 16 \dots 20$ значения элементов векторов \bar{h}_k , $\bar{\lambda}_k^*$, $\bar{\lambda}_k^{\sim}$.

Таблица 3

Результаты прогнозирования компонент составной последовательности

k	μ_k^*	σ_{μ_k}	λ_k°	$\sigma_{\lambda_k^{\circ}}$
21	1,92	0,08	-0,18	0,1
22	2,26	0,09	-1,77	0,1
23	2,64	0,11	-0,06	0,1
24	3,05	0,13	0,09	0,11
25	3,5	0,15	0,15	0,12

Таблица 4

Фрагмент файла выходных данных фильтра Калмана

16	\bar{h}_k^T	0,019772	0,165611	0,588076	0,163446	-0,149159
	$\bar{\lambda}_k^{*T}$	0,034426	0,145893	0,773894	-0,094456	-0,301338
	$\bar{\lambda}_k^{\sim T}$	0,031531	0,121642	0,687777	-0,118391	-0,279496
17	\bar{h}_k^T	0,016417	0,146549	0,554900	-0,184713	-0,132783
	$\bar{\lambda}_k^{*T}$	0,031531	0,153173	0,825184	0,097517	-0,118391
	$\bar{\lambda}_k^{\sim T}$	0,031914	0,156587	0,838112	0,101821	-0,121484
18	\bar{h}_k^T	0,013816	0,130808	0,525544	0,203458	-0,117766
	$\bar{\lambda}_k^{*T}$	0,031914	0,188501	1,010656	0,227845	0,101821
	$\bar{\lambda}_k^{\sim T}$	0,033178	0,233647	1,058743	0,246461	0,091046
19	\bar{h}_k^T	0,011743	0,117502	0,499095	0,220251	-0,103757
	$\bar{\lambda}_k^{*T}$	0,033178	0,233647	1,275801	0,219774	0,246441
	$\bar{\lambda}_k^{\sim T}$	0,032877	0,230642	1,263037	0,214141	0,249115

20	\bar{h}_k^T	0,010050	0,106034	0,474942	0,235448	-0,90598
	$\bar{\lambda}_k^{*T}$	0,032877	0,263519	1,510117	0,044505	0,214141
	$\tilde{\lambda}_k^T$	0,032630	0,260909	1,498423	0,038707	0,216372

Выводы. В статье приведено решение задачи фильтрации и прогнозирования составной полиномиально-авторегрессионной последовательности, наблюдаемой на фоне некоррелированной последовательности. Фактически решена задача разрешения двух последовательностей с различной динамикой, наблюдаемых в сумме с дискретным белым шумом. Для применения данного вида обработки необходимо знание параметров авторегрессии и дисперсии некоррелированной компоненты. Результаты обработки реальных данных о балансовой прибыли предприятия свидетельствует о работоспособности предложенного алгоритма, а также о достаточно высоком качестве результатов обработки даже при сравнительно малом объеме выборки ($M = 20$). В дальнейшем целесообразно рассмотреть данную задачу для произвольной степени полинома и произвольного порядка авторегрессии.

ЛИТЕРАТУРА

1. Мазур И.И., Шапиро В.Д., Ольдерогге Н.Г. Управление проектами. – М.: Экономика, 2001. – 574 с.
2. Бокс Дж., Дженкинс Г. Анализ временных рядов. – М.: Мир, 1974. – 401 с.
3. Бережная Е.В., Бережной В.И. Математические методы моделирования экономических процессов. – М.: Финансы и статистика, 2001. – 363 с.
4. Хомякова Н.Э. Статистическая идентификация, фильтрация и экстраполяция результатов финансовой деятельности предприятия // *Авіаційно-космічна техніка і технологія*. – 1999. – Вип. 14. – С. 55 – 58.
5. Фалькович С.Е., Хомяков Э.Н. Статистическая теория измерительных радиосистем. – М.: Радио и связь, 1981. – 288 с.

Поступила 14.03.2003

ХОМЯКОВА Надежда Эдуардовна, ассистент кафедры Национального аэрокосмического университета “ХАИ”. Окончила ГАКУ “ХАИ” в 1999 году. Область научных интересов – моделирование динамических процессов.