

Известия Алтайского государственного университета. 2025. № 1 (141). С. 112–117.
Izvestiya of Altai State University. 2025. No 1 (141). P. 112–117.

Научная статья

УДК: 519.688

DOI: 10.14258/izvasu(2025)1-15

Исследование эффекта фильтрации в задачах интервального анализа экспериментальных данных

Николай Михайлович Оскорбин¹, Ерлан Кананиянович Ергалиев²,
Лариса Ленгардовна Смолякова³

¹Алтайский государственный университет, Барнаул, Россия, osk46@mail.ru

²Восточно-Казахстанский университет им. С. Аманжолова, Усть-Каменогорск,
Казахстан, ergaliev79@mail.ru

³Алтайский государственный университет, Барнаул, Россия, knaus.larisa@gmail.com

Original article

Study of Filtering Effect in Problems of Interval Analysis of Experimental Data

Nikolai M. Oskorbin¹, Erlan K. Ergaliev², Larisa L. Smolyakova³

¹Altai State University, Barnaul, Russia, osk46@mail.ru

²East Kazakhstan University Named after S. Amanzholov, Ust-Kamenogorsk,
Kazakhstan, ergaliev79@mail.ru

³Altai State University, Barnaul, Russia, knaus.larisa@gmail.com

Аннотация. В данной работе рассматривается проблема фильтрации или разделения реальной таблицы наблюдений по возможности на полезную информацию и составляющую погрешностей наблюдения, которая в данной области носит название шума. Исследование включает разработку и обоснование моделей полезной информации и шума в каждом наблюдении; выбор критерия оптимальности; постановку математической задачи оптимизации. Задача фильтрации решается при моделировании линейного процесса, наблюдения переменных которого проводятся с интервальными погрешностями. При фильтрации получаем оценки коэффициентов связи и таблицы истинных значений переменных процесса. Показано, что общее число оценок превышает число связей и независимо от числа наблюдений точного разделения полезной информации от шума получить не удастся. В содержательной постановке задача оптимальной фильтрации состоит в том, чтобы указанное разделение было по возможности максимальным. Разработка алгоритмов фильтрации проводится с использованием систем линейных интервальных уравнений. Компьютерное моделирование процесса фильтрации и реализация многовариантных вычислительных экспериментов выполнены в среде Excel.

Ключевые слова: фильтрация погрешностей измерения, интервальный анализ данных, компьютерное моделирование, вычислительные эксперименты

Abstract. This paper examines the problem of filtering or dividing a real table of observations into possible useful data and observation errors called noise. The research includes the development and justification of useful information models and noise models for each observation, along with selection of optimality criterion and a mathematical optimization model. The filtering problem is solved while modeling the linear process with its variables observed with interval errors. There are estimated coupling coefficients and tables of true values of process variables obtained when filtering. It is shown that the number of estimates is greater than the number of connections. Therefore, it is impossible to accurately separate useful information from noise. Following the problem description, the task of optimal filtering is to ensure that the specified separation is the maximum possible. Filtering algorithms are developed using systems of linear interval equations. Computer simulations of the filtration processes and multi-variant computational experiments are performed using the Microsoft Excel software product.

Keywords: filtering measurement errors, interval data analysis, computer modeling, computational experiments

Для цитирования: Оскорбин Н.М., Ергалиев Е.К., Смолякова Л.Л. Исследование эффекта фильтрации в задачах интервального анализа экспериментальных данных // Известия Алтайского государственного университета. 2025. № 1 (141). С. 112–117. DOI: 10.14258/izvasu(2025)1-15

For citation: Oskorbin N.M., Ergaliev E.K., Smolyakova L.L. Study of Filtering Effect in Problems of Interval Analysis of Experimental Data. *Izvestiya of Altai State University*. 2025. No 1 (141). P. 112–117. (In Russ.). DOI: 10.14258/izvasu(2025)1-15

Введение

В данной работе рассматривается проблема фильтрации или разделения реальной таблицы наблюдений на полезную по возможности информацию и составляющую погрешностей наблюдения, которая в данной области носит название шума. Проектирование оптимальных фильтров в технических приложениях включает следующие этапы [1]: разработку и обоснование моделей полезной информации и шума в каждом наблюдении; выбор критерия оптимальности; постановку математической задачи оптимизации.

Следует отметить, что основы математической теории оптимальных фильтров и практика их применения по времени датируются 1958–1961 гг. В этот период разработана теория фильтрации сигналов динамических систем с внутренним шумом, наблюдение которых производится на фоне белого шума с дополнительными выбросами [2]. С развитием вычислительной техники и микропроцессорных систем эти методы нашли применение для фильтрации ошибок измерения в случаях, когда последовательность чистого сигнала имеет постоянное или медленно меняющееся значение, а модель ошибок измерения совпадает с моделью фильтров динамических систем [3–5].

Менее изучены задачи фильтрации погрешностей измерения при моделировании причинно-следственных связей с использованием базы данных (таблицы наблюдений) и базы знаний. В данном случае последовательности истинных значений входных и выходной переменных в разные моменты наблюдений не коррелированы, а базовым основанием для проектирования фильтров выступает наличие устойчивой связи значений этих переменных в отдельных наблюдениях. Указанный тип фильтров может проектироваться при моделировании процессов методами регрессионного, конъюнктного или прикладного интервального анализа данных. В первых двух статистических методах модель погрешностей наблюдений — белый шум, а при интервальной модели погрешностей измерения в последнем методе даже возможные связи ошибок наблюдения при анализе данных не учитываются [6].

Исследование эффекта фильтрации в интервальном анализе экспериментальных данных проводится при следующих соглашениях:

1. В качестве моделируемого принимаем линейный процесс: $b = x_1 a_1 + \dots + x_n a_n$, где b — значение выходной переменной, $a = (a_1, \dots, a_n)$ — вектор значений входных переменных, $x = (x_1, \dots, x_n)$ — вектор иско-

мых коэффициентов влияния входов на значение выходной переменной.

2. Таблицу N наблюдений ($N \geq n$) истинных значений совокупности переменных $(b_j^M, a_{1j}^M, \dots, a_{nj}^M)$ в каждом наблюдении j получаем как результаты измерений $(b_j, a_{1j}, \dots, a_{nj})$ с погрешностями $(\varepsilon_{bj}, \varepsilon_{1j}, \dots, \varepsilon_{nj})$, интервальные оценки которых известны и заданы выражениями: $|\varepsilon_{bj}| \leq \varepsilon_b^0; |\varepsilon_{ij}| \leq \varepsilon_i^0; i = 1, \dots, n, j = 1, \dots, N$. Здесь без потери общности интервальные погрешности измерений для отдельных переменных одинаковы для всех N наблюдений.

3. В качестве модели состояния процесса принимаем равенство:

$$x_1 a_{1j}^M + \dots + x_n a_{nj}^M = b_j^M, j = 1, \dots, N. \quad (1)$$

Задача фильтрации в указанных предположениях заключается в оценках вектора $x \in R^n$ и таблицы истинных значений переменных процесса $(b_j^M, a_{1j}^M, \dots, a_{nj}^M)$.

В результате получаем оценки ненаблюдаемых погрешностей измерения переменных моделируемого процесса как разность таблиц наблюдений и их оценки. В нашем случае общее число оценок превышает число связей, и независимо от числа наблюдений точного разделения полезной информации от шума получить не удастся. В содержательной постановке задача оптимальной фильтрации состоит в том, чтобы указанное разделение было по возможности максимальным.

Разработка алгоритмов фильтрации в интервальном анализе данных проводится с использованием интервальных систем линейных алгебраических уравнений (ИСЛАУ) [6, 7], множество решений которых связано с оценками коэффициентов $x \in R^n$ в выражении (1). Числовые оценки эффекта фильтрации получаем для частного случая линейного процесса с n , равным двум, и числом наблюдений N , равным 20. Компьютерное моделирование процесса фильтрации и реализация многовариантных вычислительных экспериментов выполнены в среде Excel.

1. Базовая математическая модель процесса фильтрации и ее модификации

Оценки переменных моделируемого процесса, вектора коэффициентов связи и интервальных погрешностей их измерения как результат фильтрации обозначим соответственно в следующем виде: $(\hat{b}_j, \hat{a}_{1j}, \dots, \hat{a}_{nj})$; $\hat{x} = (\hat{x}_1, \dots, \hat{x}_n)$; $(\hat{\varepsilon}_{bj}, \hat{\varepsilon}_{1j}, \dots, \hat{\varepsilon}_{nj})$. Считаем, что базы данных и знаний удовлетворяют исходным предположе-

ниям интервального анализа данных при моделировании процесса, т.е. являются результатами правильных наблюдений [8].

Оценки вектора $x \in R^n$ находим решением ИСЛАУ, которая в матричной форме для исходной базы данных

$$a_{ij} = [a_{ij}^H, a_{ij}^V]; a_{ij}^H = a_{ij} - \varepsilon_i^0; a_{ij}^V = a_{ij} + \varepsilon_i^0; i = 1, \dots, n; j = 1, \dots, N; \quad (2)$$

$$b_j = [b_j^H, b_j^V]; b_j^H = b_j - \varepsilon_b^0; b_j^V = b_j + \varepsilon_b^0; j = 1, \dots, N. \quad (3)$$

При исследовании процессов фильтрации рассматриваются три множества решений ИСЛАУ, в том числе объединенное $\Xi_{uni}(A, B)$, допусковое $\Xi_{tol}(A, B)$

записывается интервальной $(N \times n)$ матрицей коэффициентов и N -мерным вектором правой части: $Ax = B$. Элементы интервальных матриц A и B задаются в следующем виде [7]: $A^H \leq A \leq A^V$; $B^H \leq B \leq B^V$, где:

и управляемое $\Xi_{ctl}(A, B)$, предикатные формулы последних двух имеют вид:

$$\Xi_{tol}(A, B) = \{x \in R^n: (\forall A \in A)(\exists B \in B)(Ax = B)\}. \quad (4)$$

$$\Xi_{ctl}(A, B) = \{x \in R^n: (\forall B \in B)(\exists A \in A)(Ax = B)\}. \quad (5)$$

Анализ выражений (4) и (5) показывает, что оценки вектора $x \in R^n$, которые удовлетворяют равенству (1), можно искать на этих множествах. Объединенное множество решений ИСЛАУ содержит эти множества, накладывает слабые связи между переменными и используется при оценке итогового эффекта фильтрации реальных данных.

Рассматриваем оценки таблицы наблюдений и таблицы погрешностей. В этой части математическое описание процесса фильтрации аналогично работе [9], в которой коррекция экспериментальных данных имеет вид:

$$\hat{a}_{ij}^M = a_{ij} + e_{ij}\varepsilon_i^0; \hat{b}_j^M = b_j + e_{bj}\varepsilon_b^0; |e_{ij}| \leq 1; |e_{bj}| \leq 1. \quad (6)$$

Здесь $e = (e_{ij}; e_{bj}) \in E$ — матрица коэффициентов корректирования размером $((n+1) \times N)$, а E — множество ее допустимых значений. Для скорректированной базы наблюдений запишем ИСЛАУ $\hat{A}x = \hat{B}$ и множества решений $\Xi_{uni}(\hat{A}, \hat{B})$, $\Xi_{tol}(\hat{A}, \hat{B})$ и $\Xi_{ctl}(\hat{A}, \hat{B})$ аналогично, как для исходной базы данных. Таким об-

разом, для каждого значения $e \in E$ согласно (6) получаем оценки экспериментальных данных и погрешностей их наблюдений.

Рассмотрим критерий оптимальной фильтрации, который для выражений (1), (6) запишем в следующем виде:

$$S(\hat{A}, \hat{B}, E) = \sum_{j=1}^N (\hat{b}_j^M - (x_1 \hat{a}_{1j}^M + \dots + x_n \hat{a}_{nj}^M))^2, x \in \Xi_0(\hat{A}, \hat{B}); e \in E \quad (7)$$

В выражении (7) $\Xi_0(\hat{A}, \hat{B})$ — одно из множеств решений ИСЛАУ $\hat{A}x = \hat{B}$.

Базовая модель фильтрации — это задача выбора x^*, e^* , при которых $S(\hat{A}, \hat{B}, E)$ достигает минимального значения. В реальных постановках задача фильтрации относится к классу задач большой размерности, и требуется использовать соответствующие методы ее решения [7].

Эффект фильтрации экспериментальных данных на практике оценивается точностью оценок параметров моделируемого процесса по исходным и фильтрованным наблюдениям. В модельных вычислительных экспериментах этот эффект предлагается оценивать показателем сокращения евклидова расстояния матрицы исходных ошибок наблюдения и вычислять по формуле:

$$f_\varepsilon = \frac{R(\varepsilon_b, \varepsilon_a) - R(\hat{\varepsilon}_b, \hat{\varepsilon}_a)}{R(\varepsilon_b, \varepsilon_a)}; R(\varepsilon_b, \varepsilon_a) = \sqrt{\sum_{j=1}^N ((\varepsilon_{bj})^2 + \sum_{i=1}^n (\varepsilon_{ij})^2)}. \quad (8)$$

Показатель $R(\hat{\varepsilon}_b, \hat{\varepsilon}_a)$ рассчитывается аналогично для скорректированных погрешностей измерения согласно выражению (6).

Компьютерное моделирование процесса фильтрации проведено при следующих исходных данных: $n = 2$; $N = 20$; $a_1, a_2 \in [10, 100]$; $x_1^M = x_1^M = 1$. Генерация значений входных переменных и погрешностей на-

блюдения выполнена с использованием псевдослучайных чисел СЛЧИС(), равномерно распределенных на интервале $[0, 1]$. Оптимальные оценки параметров моделируемого процесса находим инструментом Excel «Поиск решения». Для реализации многовариантных вычислительных экспериментов базы данных и знаний обновляются новыми значе-

ниями. Модификации базовой модели фильтрации нумеруются как Ф1, Ф2 и т.д.

2. Исследование эффекта фильтрации с использованием допускового множества решений ИСЛАУ

В задачах моделирования линейных процессов допусковое множество обеспечивает сильную совмест-

ность зависимости [6, с. 222], а искомые оценки имеют высокую точность. Однако в общем случае разброса интервальных погрешностей это множество оказывается пустым, поэтому в работе [9] предлагается использовать прием уширения правой части ИСЛАУ и определять это множество в следующем виде:

$$\Xi_{tol}(A, B, k) = \{x \in R_+^n: A^V x \leq B^V(k); A^H x \geq B^H(k)\}; k \geq 0. \quad (9)$$

В выражении (9) границы правой части ИСЛАУ определяются по формуле (3), в которой ε_b^0 умножено на k . Минимальное расширение k_{min} , при котором $\Xi_{tol}(A, B, k_{min}) \neq \emptyset$, находим с использованием инструмента «Поиск решения». Фильтр Ф1 отличается от базового фильтра введением в целевую функцию слагаемого $W_E \|E\|$, который при большом числе свободных переменных выполняет роль демпфера.

Опишем фильтр Ф2, который обеспечивает не только корректировку базы данных, но и базы знаний, минимизируя радиусы интервалов погрешностей наблюдения. Радиусы интервалов определяются по формулам: $\varepsilon_1^0(1 - Ll_1)$; $\hat{\varepsilon}_2^0 = \varepsilon_2^0(1 - Ll_2)$; $\hat{\varepsilon}_b^0 = \varepsilon_b^0(1 - Ll_b)$, где $L \geq 0$ — общий параметр, значение которого следует максимизировать, а значения параметров l_1, l_2, l_3 задаются в диалоговом режиме вычислений.

Задача оптимизации фильтра Ф2 сводится к минимизации критерия: $S = W_S S(\hat{A}, \hat{B}, E) + W_E \|E\| - W_L L$

по переменным x, e, L , а ее решение проводится аналогично как для фильтра Ф1.

Точность моделирования оценивается, следуя [6], радиусом интервала прогноза выходной переменной в точке $a_1 = a_2 = 55$, который обозначим $Rad b$ для исследуемых фильтров. Для сравнения вычисляем радиус интервала прогноза $Rad bu$ для исходной таблицы наблюдений с использованием множества $\Xi_{uni}(A, B)$.

Многовариантные исследования описанных фильтров проведены для четырех вариантов погрешностей наблюдения переменных линейного процесса (табл. 1). Все результаты получены настройкой многокритериальных задач оптимизации в диалоговом режиме. Для средних значений погрешностей экспериментальных данных реально можно ожидать эффект фильтрации порядка 20 %, и эта оценка выше для фильтра Ф2.

Таблица 1

Результаты исследования фильтров Ф1 и Ф2

№ п/п	ε_1^0	ε_2^0	ε_b^0	$Rad bu$	k_{min}	Ф1		Ф2	
						f_ε	$Rad b$	f_ε	$Rad b$
1	4	4	2	5,02	2,52	19,7%	1,33	24,2%	0,23
2	4	4	4	6,32	4,51	16,9%	1,67	17,5%	1,64
3	2	2	4	2,96	1,3	31,6%	0,82	32,6%	2,13
4	0	4	0,01	0,62	39,4	61,9%	0,70	92,6%	1,48

3. Исследование эффекта фильтрации с использованием управляемого множества решений ИСЛАУ

В работе [7] показано, что управляемое множество решений ИСЛАУ $\Xi_{cl}(A, B)$ позволяет обеспечить точность оценок линейных процессов, сравнимую с до-

пусковым множеством $\Xi_{tol}(A, B)$. Как и допусковое множество, при анализе экспериментальных данных оно может оказаться пустым. В этом случае используем прием уширения левой части ИСЛАУ, который при m_{min} обеспечивает непустоту множества

$$\Xi_{tol}(A, B, m) = \{x \in R_+^n: A^V(m) x \geq B^V; A^H(m) x \leq B^H\}; m \geq 0. \quad (10)$$

В выражении (10) границы левой части ИСЛАУ определяются по формуле (2), в которой $\varepsilon_1^0, \varepsilon_2^0$ умножено на m .

Исследуемые фильтры Ф3, Ф4 аналогичны фильтрам Ф1, Ф2 и отличаются использованием в них множества $\Xi_{cl}(A, B, m_{min})$. Результаты исследования пред-

ставлены в таблице 2, в которой первые пять столбцов совпадают с таблицей 1. Анализ результатов вычислительных экспериментов показывает, что фильтр Ф4 доминирует над Ф3 по уровню коррекции погрешностей измерения.

Таблица 2

Результаты исследования фильтров Ф3 и Ф4

№ п/п	ε_1^0	ε_2^0	ε_b^0	Rad bu	k_{min}	Ф3		Ф4	
						f_ε	Rad b	f_ε	Rad b
1	4	4	2	5,02	0,88	20,98%	1,15	23,18%	1,02
2	4	4	4	6,32	1,25	18,06%	1,35	20,17%	1,97
3	2	2	4	2,96	2,21	30,80%	1,47	31,62%	1,27
4	0	4	0	0,62	0,18	74,34%	1,28	87,87%	1,17

Данные таблиц 1 и 2 позволяют утверждать, что использование множеств $\Xi_{tol}(A, B)$ и $\Xi_{crit}(A, B)$ в задачах фильтрации на группе исследованных вычислительных экспериментов дают сравнительные итоговые результаты и обеспечивают условие гарантированного оценивания [10].

Заключение

В данной статье задачи фильтрации исследуются при моделировании линейного процесса, наблюдение

переменных которого проводятся с интервальными погрешностями. При фильтрации получаем уточненные оценки коэффициентов связи и таблицы истинных значений переменных процесса в сравнении с оценками исходных баз данных и знаний. Результаты имеют теоретическое и прикладное значение при проведении экспериментальных исследований реальных процессов.

Библиографический список

1. Тисленко В.И. Статистические методы обработки сигналов в радиотехнических системах. Томск: ТУСУР, 2007. 245 с.
2. Юнусова Л.Р., Магсумова А.Р. Фильтрация шумов // Проблемы науки. 2020. № 2 (50). С. 39–43.
3. Kalman R.E., Bucy R.S. New Results in Linear Filtering and Prediction Theory // Journal of Basic Engineering. 1961. Vol. 83. P. 35–45.
4. Bucy R.S., Joseph P.D. Filtering for Stochastic Process with Application to Guidance. N.Y.: Interscience Publishers, 1968. 195 p.
5. Wiesenfeld K. Virtual Hopf Phenomenon: A New Precursor of Period-Doubling Bifurcations // Physical Review A, 1985. Vol. 32. No 3. P. 1744–1751.
6. Баженов А.Н., Жилин С.И., Кумков С.И., Шарый С.П. Обработка и анализ интервальных данных. М., Ижевск: Институт компьютерных исследований, 2024. 356 с.
7. Оскорбин Н.М., Алгазин Г.И., Суханов С.И. Методы и модели интервального анализа экспериментальных данных : учебное пособие. Барнаул: Изд-во Алт. ун-та, 2023. 95 с.
8. Zhilin S.I. Simple Method for Outlier Detection in Ztting Experimental Data under Interval Error // Chemometrics and Intellectual Laboratory Systems. 2007. Vol. 88. No 1. P. 60–68.
9. Ергалиев Е.К., Мадияров М.Н., Оскорбин Н.М., Смолякова Л.Л. Согласование базы данных в прикладном интервальном анализе // Известия Алтайского государственного университета. 2022. № 1 (123). С. 90–94. DOI: 10.14258/izvasu(2022)1-141
10. Шелудько А.С. Гарантированное оценивание параметров дискретных моделей хаотических процессов // Вестник ЮУрГУ. Серия: Вычислительная математика и информатика. 2018. Т. 7. № 1. С. 25–39. DOI: 10.14529/cmse180103

References

1. Tislenko V.I. *Statistical Methods of Signal Processing in Radio Engineering Systems*. Tomsk: TUSUR, 2007. 245 p. (In Russ.).
2. Yunusova L.R., Magsumova A.R. Noise Filtering. *Problems of Science*. 2020. No 2 (50). P. 39–43. (In Russ.).
3. Kalman R.E., Bucy R.S. New Results in Linear Filtering and Prediction Theory. *Journal of Basic Engineering*. 1961. Vol. 83. P. 35–45.
4. Bucy R.S., Joseph P.D. *Filtering for Stochastic Process with Application to Guidance*. N.Y.: Interscience Publishers, 1968. 195 p.
5. Wiesenfeld K. Virtual Hopf Phenomenon: A New Precursor of Period-Doubling Bifurcations. *Physical Review A*. 1985. Vol. 32. No 3. P. 1744–1751.
6. Bazhenov A.N., Zhilin S.I., Kumkov S.I., Shary S.P. *Processing and Analysis of Interval Data*. M.: Izhevsk: Institute of Computer Research, 2024. 356 p. (In Russ.).

7. Oskorbin N.M., Algazin G.I., Sukhanov S.I. *Methods and Models of Interval Analysis of Experimental Data*. Barnaul: ASU, 2023. 95 p. (In Russ.)
8. Zhilin S.I. Simple Method for Outlier Detection in Ztting Experimental Data under Interval Error. *Chemometrics and Intellectual Laboratory Systems*. 2007. Vol. 88. No 1. P. 60–68.
9. Ergaliev E.K., Madiyarov M.N., Oskorbin N.M., Smolyakova L.L. Database Reconciliation in Applied Interval Analysis. *Izvestiya of Altai State University*. 2022. No 1 (123). P. 90–94. DOI: 10.14258/izvasu(2022)1-141 (In Russ.)
10. Sheludko A.S. Guaranteed Estimation of Parameters of Discrete Models of Chaotic Processes. *Vestnik YUUrGU: Series: Computational Mathematics and Computer Science*. 2018. T. 7. No 1. P. 25-39. DOI: 10.14529/cmse180103 (In Russ.)

Информация об авторах

Н.М. Оскорбин, доктор технических наук, профессор кафедры теоретической кибернетики и прикладной математики, Алтайский государственный университет, Барнаул, Россия;

Е.К. Ергалиев, кандидат физико-математических наук, заведующий кафедрой математики, Восточно-Казахстанский университет им. С. Аманжолова, Усть-Каменогорск, Казахстан;

Л.Л. Смолякова, старший преподаватель кафедры информатики, Алтайский государственный университет, Барнаул, Россия.

Information about the authors

N.M. Oskorbin, Doctor of Sciences in Technology, Professor of the Department of Theoretical Cybernetics and Applied Mathematics, Altai State University, Barnaul, Russia;

E.K. Ergaliev, Candidate of Sciences in Physics and Mathematics, Head of the Department of Mathematics, East Kazakhstan University named after S. Amanzholov, Ust-Kamenogorsk, Kazakhstan;

L.L. Smolyakova, Senior Lecturer of the Department of Computer Science, Altai State University, Barnaul, Russia.