

ВОЗМОЖНОСТИ МЕТОДА ГРУППОВОГО УЧЕТА АРГУМЕНТОВ В ЗАДАЧАХ ПРОГНОЗИРОВАНИЯ БЕЗОПАСНОСТИ ХИМИЧЕСКИХ ПРОИЗВОДСТВ

**Е.С. Калинина, кандидат педагогических наук, доцент.
Санкт-Петербургский университет ГПС МЧС России**

Рассмотрен алгоритм применения комбинаторного метода группового учета аргументов при математическом моделировании технологических производств. Исследованы вопросы оценки точности и прогнозирующих характеристик, полученных методом группового учета аргументов математических моделей различных стадий химического производства.

Ключевые слова: безопасность химических производств, математическая модель, метод группового учета аргументов, прогнозирование

POSSIBILITIES OF THE GROUP METHOD OF DATA HANDLING IN PROBLEMS OF FORECASTING OF SAFETY OF CHEMICAL PRODUCTIONS

E.S. Kalinina. Saint-Petersburg university of State fire service of EMERCOM of Russia

The algorithm of application of a group method of data handling at mathematical modeling of technological productions is considered. Questions of an assessment of accuracy and the predicting characteristics received by the group method of data handling of mathematical models of various stages of chemical production are investigated.

Keywords: safety of chemical productions, mathematical model, group method of data handling, forecasting

В настоящее время химические производства являются одними из наиболее опасных техногенных источников воздействия на человека и окружающую среду. Химическая опасность производств подразделяется на токсическую, пожаро- и взрывоопасную. К химически опасным объектам относятся не только предприятия химической, нефтехимической, металлургической и других отраслей промышленности, где токсические химические вещества содержатся в сырье, вспомогательных материалах, технологических смесях, продуктах и отходах. Значительные массы сильнодействующих токсических веществ сосредоточены на объектах пищевой, мясомолочной промышленности, в жилищно-коммунальном хозяйстве и т.д. [1].

Таким образом, задача предотвращения чрезвычайных ситуаций на химических производствах, анализ, оценка риска и управление их безопасностью на сегодняшний день является очень актуальной. Эффективное решение поставленной задачи может быть получено на основе системного подхода к изучаемой проблеме с использованием математического моделирования.

Эффективным аппаратом построения математических моделей прогнозирования, дающим возможность автоматически находить скрытые взаимозависимости в исследуемых процессах и выбирать оптимальную структуру модели, является метод группового учета аргументов [2].

Метод группового учета аргументов (МГУА) применяется в самых различных областях для анализа данных, прогнозирования, моделирования систем, оптимизации и распознавания образов. Метод базируется на индуктивных принципах – нахождение лучшего решения основано на переборе всевозможных вариантов [3].

Математические модели, полученные МГУА, отвечают всем требованиям, которые предъявляются к моделям, используемым для промышленных систем оптимизации: информативностью, адекватностью и точностью в диапазоне возможных режимов, малым временем счета и сравнительно небольшим объемом оперативной памяти на ЭВМ [4].

В регрессионных моделях диапазон корректного использования модели устанавливается с помощью интервала доверительной вероятности, который оценивает близость математической модели объекта к математическим моделям процессов с нормальным распределением. В отличие от регрессионных уравнений, которые являются уравнениями момента первого порядка, модели, полученные МГУА, можно рассматривать как модели закономерностей моделируемого объекта, построенные по экспериментальным данным [5, 6]. Поэтому математические модели, полученные МГУА, необходимо дополнить уравнениями, которые установят условия их применения. Особенно это необходимо при расчетах, когда изменение переменных выходит за диапазон значений, на которых были построены МГУА, используемые в модели.

На сегодняшний день разработано много различных МГУА [7]. В настоящей работе рассмотрим применение однорядного или, так называемого, комбинаторного алгоритма МГУА, который наиболее эффективен для решения технических задач. В работе [6] установлено, что комбинаторный алгоритм не имеет «ошибки многорядности» и не может потерять оптимальную модель, потому что обеспечивает перебор всех возможных, при заданной опорной функции, моделей. Практически его можно применять в случаях, когда число слагаемых полного полинома меньше 20.

Пусть дана совокупность экспериментальных значений входных переменных:

$$X = (X_1; X_2; \dots; X_m)$$

и выходной величины Y , где X_i и Y – n -мерные векторы; n – число точек массива исходных данных; m – число переменных.

Математические модели объектов, описывающие статические режимы их функционирования, ищутся чаще всего в виде полиномов:

$$Y = f(X) = A_0 + \sum_{i=1}^m A_i X_i + \sum_{i=1}^m A_{ii} X_i^2 + \sum_{i,j=1}^l A_{ij} X_i X_j + \sum_{i=1}^m A_{iii} X_i^3 + \dots, \quad (1)$$

где $l = C_m^2, i \neq j$, A – вектор неизвестных коэффициентов размерности k :

$$A = \{A_0, A_1, \dots, A_m, \dots\} \in E^k.$$

Если порядок h -полинома по всем переменным X_i равен числу m входных координат, то выражение (1) является полным полиномом Колмогорова-Габбора [3]. Для определения вектора $A \in E^k$ применяют, как правило, метод наименьших квадратов, при этом сталкиваются с двумя следующими затруднениями:

– вследствие отсутствия априорной информации о моделируемых объектах, выходные переменные математической модели могут быть коррелированными, что приводит к плохой обусловленности системы нормальных уравнений, определяющих параметры A . Высокая размерность вектора A делает невозможным предварительный анализ степени обусловленности матрицы системы уравнений, поэтому исследователь вынужден убеждаться в неустойчивости решения задачи определения A лишь в результате многократных вычислений на ЭВМ;

– задача определения вектора A заранее не заданной размерности k полинома (1) является неустойчивой в евклидовой метрике. Обоснованный выбор размерности k и структуры полинома (1) предполагает наличие априорной достоверной информации о моделируемом объекте, которая часто на практике отсутствует. Кроме того, задание структуры полинома (1) является субъективным фактором, что является потенциальным источником ошибочных решений.

Перечисленные трудности можно устранить, если для построения математической модели использовать МГУА. Согласно МГУА решение задачи моделирования большой размерности заменяется многостадийным процессом решения большого числа простых задач аппроксимации данных функциями-полиномами заданной структуры с невысокой размерностью вектора A . Тогда полином типа (1) примет вид:

$$Y_{ij} = A_{0ij} + A_{1i}X_i + A_{1j}X_j + A_{ij}X_iX_j + A_{ii}X_i^2 + A_{jj}X_j^2. \quad (2)$$

Возможны и другие структуры, например, при коэффициентах $A_{ij} = 0$ или $A_{jj} = A_{ii} = 0$ и т.д. Такие полиномы называются частными описаниями. Массив экспериментальных точек разбивается, по крайней мере, на два подмножества: обучающее N_a и проверочное N_b . На множестве N_a методом наименьших квадратов определяются векторы коэффициентов:

$$A_{ij} = \{A_{0ij}, A_{1ij}, \dots, A_j\} \in E^k$$

каждого частного описания. На множестве N_b вычисляется значение критерия селекции CR для всех частных описаний данного r ряда селекции. На следующий $r + 1$ ряд селекции пропускается F частных описаний с наименьшими значениями критерия селекции, где F – свобода выбора.

На $r + 1$ ряду селекции процесс повторяется, причем аргументами в частном описании (2) принимаются ординаты, полученные на предыдущем r ряду селекции, то есть частные описания примут вид:

$$Z_{ij} = A_{0ij} + A_{1i}Y_i + A_{1j}Y_j + A_{ij}Y_iY_j + A_{ii}Y_i^2 + A_{jj}Y_j^2.$$

Процесс селекции продолжается до тех пор, пока значение критерия селекции на некотором $r^* + 1$ ряду не будет удовлетворять условию:

$$\min_k CR_k^{r^*+1} \geq \min_k CR_k^{r^*},$$

где $k = 1, 2, \dots, F$.

Для оценки возможности метода МГУА в задачах прогнозирования безопасности химических производств исследуем математические модели трех последовательных стадий производства фенола [2, 4].

При построении моделей в качестве аппроксимирующего полинома вида (2) был выбран полином второй степени Колмогорова-Габора от переменных \bar{X} и \bar{U} :

$$\bar{Y} = P^2(\bar{X}, \bar{U}),$$

где \bar{Y} – выходная величина.

Для исследования влияния нестационарности химических производств на корректность математических моделей, полученных МГУА, определим интервал стационарности T_c , соответствующий времени, по истечении которого погрешность расчета значений выходной величины достигнет и превысит заданную величину δ .

Учитывая, что метод группового учета аргументов может применяться в ситуациях, когда выборка содержит небольшое количество элементов, при котором использование статистических гипотез о плотности распределения невозможно [5], определим оптимальный объем $N_{\text{опт}}$ выборки экспериментальных значений, взятых из имеющейся последовательности данных. На выборках 5, 7, 9 и 11 точек последовательно осуществлялся синтез модели МГУА соответствующего технологического процесса по одному из каналов [4]. Для полученных моделей рассчитаем оценку точности – среднюю относительную погрешность δ по формуле:

$$\delta = \frac{\sum_{i=1}^N |Y_{ip} - Y_{cp}|}{N \cdot Y_{cp}} \cdot 100\%,$$

где Y_{ip}, Y_{cp} – расчетное по модели и экспериментальное значение выходной величины.

Минимум относительной погрешности δ будет соответствовать оптимальному объему выборки экспериментальных данных $N_{\text{опт}}$, и, следовательно, интервалу стационарности T_c .

Результаты эксперимента представлены в табл. 1. Оптимальный объем выборки составляет:

- для первой стадии производства $N_{\text{опт}} = 5$ ($T_c = 40$ ч);
- для второй стадии производства $N_{\text{опт}} = 7$ ($T_c = 48$ ч);
- для третьей стадии производства $N_{\text{опт}} = 5$ ($T_c = 40$ ч).

Таблица 1. Зависимость средней относительной погрешности математических моделей различных стадий химического производства фенола от объема выборки экспериментальных данных

Объем выборки, N	Стадии производства		
	первая	вторая	третья
	$\delta, \%$	$\delta, \%$	$\delta, \%$
5	0,9	1,35	5,9
7	1,7	1,135	7,3
9	1,7	2,13	17,5
11	1,9	1,876	–

Экспериментальные зависимости для различных стадий производства были получены в виде возрастающих функций $\delta = f(N)$, что можно объяснить нестационарностью условий работы. Условия функционирования исследованных стадий производства были обычные – специальных мероприятий по стабилизации режимов не проводилось.

Количественные результаты экспериментов позволяют утверждать, что на интервалах времени вне T_c величина ошибки δ для первой и третьей стадий производства не превышает 2,4 %. Для второй стадии погрешность модели значительно выше.

Для оценки прогнозирующих характеристик моделей химических производств определим интервал времени T_{Π} , на котором величина дисперсии σ^2 не станет больше значения σ_c^2 , полученного на выборке, соответствующей интервалу стационарности. Дисперсию рассчитаем по формуле:

$$\sigma^2 = \frac{\sum_{j=1}^N (Y_{эj} - Y_{рj})^2}{N},$$

где $Y_{эj}, Y_{рj}$ – экспериментальное и расчетное значения выходной величины.

Математические модели синтезировались на выборках N_c , соответствующих интервалу T_c . Затем на новых выборках, хронологически последующих за T_c , рассчитывались оценки дисперсий σ^2 при увеличении объема выборки [2, 4]. Результаты расчетов представлены в табл. 2.

Величина интервала T_{Π} , определяющего прогнозирующие способности, составляет: для первой стадии производства – 32 ч, для второй стадии – 64 ч, для третьей стадии – 56 ч, что свидетельствует о высоких прогнозирующих возможностях математических моделей, полученных МГУА.

Таблица 2. Прогнозирующие характеристики математических моделей различных стадий химического производства фенола

Стадии производства	Объем выборки, N	Оценка дисперсии, σ^2	Интервал T_{Π} прогноза, ч
Первая	5	3,88	32
	7	3,69	
	9	4,11	
	11	4,13	
Вторая	7	0,613	64
	9	0,637	
	11	0,525	
	13	0,502	
Третья	7	0,0085	56
	9	0,027	
	11	0,065	
	13	0,104	

Таким образом, анализ МГУА показывает, что по сравнению со статистическими методами прогнозирования МГУА обладает определенным разнообразием возможностей на всех этапах процесса моделирования и следующими преимуществами, позволяющими улучшить прогнозирование моделей сложных объектов:

- в методе МГУА не требуется предварительно задавать структуру модели $f(X, A)$. Она формируется в процессе многостадийного отбора «наилучших» частных описаний;
- находится оптимальная сложность структуры модели, адекватная уровню помех в выборке данных (для решения реальных проблем с зашумленными или короткими данными, упрощенные прогнозирующие модели оказываются более точными);

– гарантируется нахождение наиболее точной или несмещенной модели – метод не пропускает наилучшего решения во время перебора всех вариантов (в заданном классе функций);

– главным достоинством алгоритмов МГУА, отличающим их от других методов, является разделение имеющихся наблюдений на две выборки: обучающую, по которой производится идентификация моделей заданного класса структур, и проверочную, обеспечивающую селекцию моделей оптимальной сложности с точки зрения некоторого внешнего критерия;

– любые нелинейные функции или воздействия, которые могут иметь влияние на выходную переменную, используются как входные параметры;

– МГУА дает возможность корректировки прогноза при получении новых факторов;

– автоматически находит интерпретируемые взаимосвязи в данных и выбирает эффективные входные переменные;

– метод использует информацию непосредственно из выборки данных и сводит к минимуму влияние субъективных факторов при построении модели.

Применение МГУА может оказать значительную помощь в решении задач стабилизации статических режимов химических процессов при синтезе безопасных технологических систем, инвариантных к внешним возмущениям.

Математические модели, построенные на основе МГУА, могут эффективно применяться при оптимизации действующих химико-технологических процессов с целью повышения их надежности и безопасности, а также для поиска оптимальных решений при разработке новых высоконадежных химико-технологических систем.

Метод группового учета аргументов обладает высокими прогнозирующими характеристиками и может быть использован в комплексном решении задач оценки риска и прогнозирования безопасности современных химических производств.

Литература

1. Ивахненко А.Г. Индуктивный метод самоорганизации моделей сложных систем. Киев: Наукова думка, 1982.

2. Ивахненко А.Г., Юрачковский Ю.П. Моделирование сложных систем по экспериментальным данным. М.: Радио и связь, 1987.

3. Ивахненко А.Г., Степашко В.С. Помехоустойчивость моделирования. Киев: Наукова думка, 1985.

4. Калинина Е.С., Сметанин Ю.В. Метод группового учета аргументов в задачах прогнозирования пожароопасности // Проблемы обеспечения безопасности при чрезвычайных ситуациях: Материалы междунар. науч.-практ. конф. СПб.: С.-Петербург. ин-т ГПС МЧС России, 2003.

5. Меньшиков В.В., Швыряев А.А. Опасные химические объекты и техногенный риск: учеб. пособие. М.: Изд-во «Химия», 2003.

6. Сметанин Ю.В. Математическое моделирование и оперативная оптимизация химико-технологических процессов и систем. Л.: ЛТИ.-Л., 1988. 158 с.

7. Официальный сайт метода МГУА. URL: <http://www.gmdh.net> (дата обращения: 14.09.2016).