МОДЕЛИРОВАНИЕ И ПРОГНОЗИРОВАНИЕ БИЗНЕС- ПРОЦЕССОВ С ПОМОЩЬЮ АЛГОРИТМОВ САМООРГАНИЗАЦИИ ФОРМАЛЬНЫХ ОПИСАНИЙ

В.В. Белов,
д.т.н., профессор кафедры вычислительной и прикладной математики Рязанского государственного радиотехнического университета,
e-mail vvbeloff@yandex.ru
В.И. Чистякова,
к.т.н., доцент кафедры вычислительной и прикладной математики Рязанского государственного радиотехнического университета,
e-mail compvv@mail.ryazan.ru

Предлагаются направления развития метода группового учёта аргументов, разрабатываются новые алгоритмы, отличающиеся применением процедуры оптимизации частных полиномов в последовательных рядах приближений. Разработанные алгоритмы могут применяться для моделирования и прогнозирования показателей производственно-экономической деятельности предприятий. Приводятся примеры описаний, полученных в процессе разработки прогноза социально-экономического статуса тепловой электростанции.

Введение

Предметом рассмотрения являются производственно-экономические процессы, представляющие собой изменение во времени показателей производственно-экономической деятельности предприятия. Стимулом к рассмотрению явились исследование текущей ситуации и прогноз социально-экономического статуса крупной тепловой электростанции, выполненные в рамках хоздоговорной НИР.

В настоящей статье приводятся результаты разработки формализованных описаний, позволяющих прогнозировать значения проблемных показателей и моделировать зависимости результативных показателей от факторных. Проблемными названы показатели, представляющие интерес для исследователя, результативными – те показатели, непосредственное управление которыми практически невозможно, факторными – показатели, на значения которых возможны прямые или косвенные управляющие воздействия.

Предлагаемые алгоритмы осуществляют, по сути дела, поиск наилучшей в некотором смысле нелинейной множественной регрессии. Они могут применяться во всех случаях, когда возникает необходимость синтеза модели, описывающей зависимость одних количественных дискретных величин от других. Наибольший эффект достигается при достаточно большом количестве потенциальных аргументов синтезируемого формального описания и отсутствии сколь-нибудь надёжных предположений о статистических свойствах изучаемых процессов. Естественно, существуют и значительные ограничения:

1) объём статистического материала должен быть адекватным размерности решаемой задачи;

2) зависимость результативных показателей от факторных должна быть близка к детерминированной.

Указанные ограничения весьма нечётки, проверить их выполнение практически невозможно. Выручает следующее: в процессе работы алгоритмов качество построенных прогнозирующих описаний тестируется самым надёжным способом — тестовой последовательностью.

В качестве базовой нелинейной модели используется полином Колмогорова — Габора (дискретный вариант функционального ряда Вольтерра)

$$\hat{y} = a_0 + \sum_{i=1}^{p} a_i x_i + \sum_{i=1}^{p} \sum_{j=1}^{i} a_{ij} x_i x_j + \sum_{i=}^{p} \sum_{j=i}^{p} \sum_{k=j}^{p} a_{ijk} x_i x_j x_k + \ldots.$$

Конкретная структура полинома не задаётся — она является предметом поиска в процессе самоорганизации модели и, в частности, определяет состав эффективных аргументов описания, которые автоматически выбираются из множества учитываемых факторных показателей. Способность алгоритма выделять эффективные аргументы, — оказывающие в своей совокупности наиболее существенное влияние на результативный показатель, — называется его селектирующей способностью. Она, конечно же, легко проверяется на искусственных, специальным образом синтезированных данных.

Постановка задачи

Формально рассматриваемая задача сводится к задаче построения аппроксимации (модели) зависимости скалярной функции от векторного аргумента со следующим требованием: получаемая аппроксимация должна обеспечивать приемлемые значения погрешности экстраполяции, то есть должна обладать достаточно существенной прогностической силой. Функция представляет собой некоторый результативный показатель, а элементы векторного аргумента являются факторными показателями. Исходными данными являются одновременные наблюдения скалярной функции и векторного аргумента $\vec{x} = (x_1, x_2,\ldots,x_p)$:

$$(y_i, \vec{x}_i),\ i = 1,\ 2,\ \ldots,\ m;\quad \vec{x}_i = (x_{i1},\ x_{i2},\ \ldots,\ x_{in}).$$

Известны многочисленные подходы и методы решения указанной формальной задачи. Наиболее часто используются регрессионный анализ [1], алгоритмы поиска «наилучшей» регрессии [2] и искусственные нейронные сети [3]. В излагаемых исследованиях в качестве классической платформы использован метод эвристической самоорганизации [4] в конкретной его форме, называемой методом группового учёта аргументов [5]. Обоснование целесообразности такого выбора можно найти, например, в [6] и [7].

Разработка прогноза осуществляется методом сценарных условий: прогнозные значения факторных показателей находятся по экспертным оценкам и/или методом выявления и экстраполяции тенденций; будущие значений результативного признака вычисляются путём подстановки найденных значений в полученную ранее модель. В ряде приложений хорошо срабатывает метод лагированных (задержанных) переменных: в качестве аргументов модели используются не текущие, а предыдущие значения факторных показателей. При этом глубина прогноза равна величине задержки аргументов. Указанный приём применим в тех случаях, когда имеет место определенная инерция в изменении результативного показателя под влиянием изменений факторных показателей.

Новые алгоритмы эвристической самоорганизации, отличающиеся структурой интегрирующего ядра

Анализ известных методов построения математических моделей по результатам эксперимента показал, что метод группового учёта аргументов (МГУА) остается наиболее эффективным средством создания математического описания процессов. Однако классические варианты МГУА базируются на фиксированных структурах опорных полиномов. Обычно это квадратичные полиномы. При переходе на новый ряд в качестве аргументов используются лучшие приближения предыдущего ряда. Нетрудно убедиться, что такой способ организации селекции изначально имеет методическое ограничение точности восстановления функций. Действительно, пусть неизвестная функция $y = f(x)$ представляет собой просто x^3. Если организовать восстановление этой функции по совокупности дискретных отсчетов (yi, xi), $i = 1, 2, \ldots, m$ в соответствии с классическим МГУА, то первое приближение будет иметь вид

$$y^{(1)} = a_1 x^2 + b_1 x + c_1,$$

а второе

$$\begin{aligned} y^{(2)} &= a_2 (y^{(1)})^2 + b_2 y^{(1)} + c_2 = \\ &= a_2 a_1^2 x^4 + 2a_2 a_1 b_1 x^3 + (a_2 b_1^2 + 2a_2 a_1 c_1 + b_2 a_1) x^2 + \\ &+ (2a_2 b_1 c_1 + b_2 b_1) x + a_2 c_1^2 + b_2 c_1 + c_2. \end{aligned}$$

Очевидно следующее:

1) желательно, чтобы выполнились соотношения:

2) $2a_2 a_1 b_1 = 1;$

$$a_2 a_1^2 = a_2 b_1^2 + 2a_2 a_1 c_1 + b_2 a_1 = 2a_2 b_1 c_1 + b_2 b_1 = = a_2 c_1^2 + b_2 c_1 + c_2 = 0;$$

3) указанные соотношения несовместны;

4) дальнейшие итерации в построении модели не приведут к улучшению точности приближения.

Таким образом, даже простейшая функция не восстанавливается абсолютно точно классическим вариантом МГУА. Причиной этому является фиксированность структуры опорных полиномов, которая обусловливает следующий факт: в процессе самоорганизации описания принципиально отсутствует возможность синтеза произвольного полинома Колмогорова — Габора, в процессе эволюционного отбора порождаются и участвуют в конкуренции только некоторые из всех возможных полиномов.

Для устранения указанного недостатка и существенного улучшения свойств восстановления функций предлагается модифицировать метод группового учета аргументов следующим образом. Для обеспечения возможности быстрого восстановления зависимостей от одного аргумента необходимо ввести дополнительный этап формирования приближений от одиночных аргументов. Для обеспечения возможности точного восстановления простейших зависимостей от одиночных и парных аргументов, а также для повышения точности приближения функции необходимо осуществлять подбор оптимальной степени полинома частного приближения, а не использовать «опорный» (например, квадратичный или иной) полином.

Тестирование классического МГУА путем решения контрольных задач с искусственно формируемыми исходными данными показывает, что его селектирующие способности не достаточно высоки: в некоторых примерах аргументы, не входящие в формулу, определяющую процесс, оказывались в списке аргументов модели процесса.

Для варианта алгоритма МГУА, описанного ниже, таких явлений на тех же тестовых данных не наблюдается, что свидетельствует о его более высоких селектирующих способностях. Кроме того, при обработке больших объемов исходных данных заметно различие во времени работы программы: классический вариант алгоритма обрабатывает одинаковые данные почти вдвое дольше предлагаемого варианта.

В общем случае различные варианты алгоритмов, реализующих рассматриваемый метод построения модели, могут отличаться следующими элементами:

- ✧ структурой интегрирующего ядра; при этом количество аргументов в модифицированном векторе факторов может быть произвольным; на начальном шаге алгоритма может использоваться и один аргумент, алгоритм при этом становится более унифицированным; степень первичного полинома также может быть произвольной; кроме того, в качестве частичных приближений могут использоваться не только полиномы, но и другие функции;
- ✧ видом коэффициентов первичного полинома и аргументов процесса — они могут быть либо скалярными, либо интервальными величинами;
- ✧ методом вычисления значений коэффициентов первичного полинома — это может быть метод наименьших квадратов или линейное программирование (в зависимости от объема исходных данных и вида аргументов);
- ✧ способом деления выборки на рабочую и контрольную группы; контрольная группа может предшествовать рабочей, следовать за ней, или данные групп могут чередоваться.

Критерий выбора наилучшего приближения на каждом шаге алгоритма определяется методом вычисления коэффициентов первичного полинома. В методе наименьших квадратов — это сумма квадратов отклонений расчетных значений от экспериментальных. В задаче линейного программирования с интервальной линейной регрессией — это сумма длин интервалов и отклонения средних значений от экспериментальных. При использовании интервальных величин каждый элемент вектора коэффициентов A представляет собой интервал $A_i = (a_i, \alpha_i)$, где a_i — среднее значение интервала, α_i — половина длины интервала.

В тех случаях, когда предполагается использование синтезируемой модели для решения задачи прогнозирования, оценку точности или соответствующей погрешности формального описания целесообразно осуществлять по тестовой, или проверочной последовательности, — части исходных данных, не использованных в процессе оценки параметров.

Мы будем использовать запись $y^{*(l)} = \mathrm{best}(y^{*(l,j)})$ для отражения факта выбора наилучшего приближения на конкретном этапе серии последовательных приближений, которая эквивалентна записи $y^{*(l)} = y^{*(l,k)}$, где $k = \arg(\min Q^{(l,j)}$; $Q^{(l,j)}$ — критерий, описывающий погрешность приближения $y^{*(l,k)}$ к y.

В общем случае для определения наилучшего приближения возможно использование человеческого фактора в интерактивном режиме, при этом процесс выбора наилучшего приближения не может быть строго формализован.

Алгоритм автопостроения модели без использования селекции в последовательных рядах приближения

На начальном (нулевом) этапе построения приближения каждый из аргументов функции используется для формирования частного аппроксимирующего полинома. Причем степень полинома постепенно увеличивается до тех пор, пока показатель погрешности приближения проверочной последовательности не перестанет уменьшаться. Все полученные таким образом частные приближения включаются в список аргументов. Расширенный набор аргументов используется для реализации следующего шага построения модели.

На этапе №1 алгоритма формируются полиномиальные приближения для всех возможных пар аргументов. При этом степень полинома постепенно увеличивается по той же схеме, что и на этапе №0. Все полученные новые частные приближения также включаются в список аргументов последующего этапа.

На этапе №r алгоритма ($r = 2, 3, \ldots$) формируются полиномиальные приближения для следующих пар аргументов – первый элемент пары берется из списка аргументов этапа №r–2 , а второй – из списка результатов (приближений) этапа №r–1; после исчерпания пар указанного типа формируются все возможные пары из списка результатов этапа №r–1. Степень полинома подбирается по схеме этапа №0. Все полученные новые частные приближения включаются в список аргументов последующего этапа.

На каждом этапе алгоритма оценивается показатель погрешности приближения ε_r, равный наименьшей погрешности среди погрешностей всех частных приближений этапа:

$$\varepsilon_r = \min_{\forall k} (\varepsilon_r^{(k)})$$

Последовательные этапы выполняются до тех пор, пока наименьшая ошибка приближения тестовой последовательности не прекратит уменьшаться, т.е. $\varepsilon r \le \varepsilon_{r+1}$. Конечным результатом алгоритма является лучшее приближение предпоследнего этапа.

Формирование набора аргументов описанного алгоритма схематично можно изобразить в следующем виде:

этап №0 $\quad \vec{x} \to \vec{P}^{(0)}$

этап №1 $\quad \vec{P}^{(0)} * \vec{P}^{(0)} \to \vec{P}^{(1)}$

этап №2 $\quad (\vec{P}^{(0)} * \vec{P}^{(1)}) \& (\vec{P}^{(1)} * \vec{P}^{(1)}) \to \vec{P}^{(2)}$

...

этап №r $\quad (\vec{P}^{(r-2)} * \vec{P}^{(r-1)}) \& (\vec{P}^{(r-1)} * \vec{P}^{(r-1)}) \to \vec{P}^{(r)}$

финальная модель: $\hat{y} = \underset{\varepsilon_r}{\mathrm{best}} (\vec{P}^{(r)})$,

где $\vec{x}$ – начальный вектор аргументов синтезируемой модели;

$\vec{P}^{(i)}$, – вектор частных приближений, порождаемый на i-м шаге построения описания, являющийся вектором аргументов i+1-го шага;

$\underset{\varepsilon_r}{\mathrm{best}} (\vec{P}^{(r)})$– наилучшее по ошибке восстановления тестовой последовательности из частных приближений, полученных на последнем шаге работы алгоритма;

символ $\to$ означает отображение вектора аргументов в вектор частных приближений функции;

$v_1 * v_2$ – операция прямого умножения векторов, результатом которой является упорядоченное множество упорядоченных пар, первый элемент которых принадлежит вектору v_1, а второй – вектору v_2;

$A \& B$ – операция объединения упорядоченных множеств (множество A дополняется справа множеством B).

Алгоритм автопостроения модели с использованием механизмов эвристической селекции в последовательных рядах приближения

Как показывают результаты практических расчётов, рассмотренный в предыдущем подпункте алгоритм автопостроения модели обеспечивает достаточно точное восстановление законов и приемлемую для многих приложений аппроксимацию процессов. Однако он имеет следующие существенные недостатки. Во-первых, он требует большого объема памяти для хранения нарастающего объема частных приближений, используемых в качестве аргументов последующих шагов алгоритма. Во-вторых, он не отсеивает несущественные факторы. Случайные (за счет шумов в исходных данных) проявления их влияния на изучаемый процесс могут негативно влиять на последующие приближения, понижая точность аппроксимации. Избавиться от указанных недостатков можно с помощью механизмов отбрасывания плохих промежуточных приближений [8].

На i-м этапе построения модели в качестве аргументов используются результаты всех предыдущих этапов алгоритма, однако не все, а только лучшие из них, – удовлетворяющие некоторому селектирующему требованию. В частности, можно использовать следующее требование: в состав аргументов i-го этапа включаются частные приближения предыдущих этапов, погрешность которых не превышает удвоенной погрешности (i–1)-го этапа.

Операцию выбора лучших приближений обозначим как $\underset{\varepsilon}{\mathrm{best}}(\vec{P})$, где $\vec{P}$ – вектор приближений;

ε – отбраковывающее значение погрешности приближения.

Как в алгоритме без селекции последовательные этапы выполняются до тех пор, пока наименьшая ошибка приближения тестовой последовательности не прекратит уменьшаться, т.е. $\varepsilon_r \le \varepsilon_{r+1}$.

С учетом изложенного схему формирования набора аргументов улучшенного алгоритма можно изобразить в следующем виде:

этап № 0: $\vec{x} \to \vec{P}_{k_0}^{(0)}$;

этап № 1: $\vec{P}^{(0)} * \vec{P}^{(0)} \to \vec{P}_{k_1}^{(1)}$;

этап № 2: $\vec{P}_{sel}^{(0)} = \underset{2\varepsilon_1}{\mathrm{best}}(\vec{P}_{k_0}^{(0)})$; $\vec{P}_{sel}^{(1)} = \underset{2\varepsilon_1}{\mathrm{best}}(\vec{P}_{k_1}^{(1)})$;

$$(\vec{P}_{sel}^{(0)} * \vec{P}_{sel}^{(1)}) \,\&\, (\vec{P}_{sel}^{(1)} * \vec{P}_{sel}^{(1)}) \to \vec{P}_{k_2}^{(2)};$$

…

этап № r : $\vec{P}_{sel}^{(0)} = \underset{2\varepsilon_{r-1}}{\mathrm{best}}(\vec{P}_{k_0}^{(0)})$; ... $\vec{P}_{sel}^{(r-2)} = \underset{2\varepsilon_{r-1}}{\mathrm{best}}(\vec{P}_{k_{r-2}}^{(r-2)})$;

$$\vec{P}_{sel}^{(r-1)} = \underset{2\varepsilon_{r-1}}{\mathrm{best}}(\vec{P}_{k_{r-1}}^{(r-1)});$$

$$(\vec{P}_{sel}^{(0)} * \vec{P}_{sel}^{(r-1)}) \,\&\, (\vec{P}_{sel}^{(1)} * \vec{P}_{sel}^{(r-1)}) \,\&\, ...$$

$$...(\vec{P}_{sel}^{(r-2)} * \vec{P}_{sel}^{(r-1)}) \,\&\, (\vec{P}_{sel}^{(r-1)} * \vec{P}_{sel}^{(r-1)}) \to \vec{P}_{k_r}^{(r)};$$

финальная модель: $\hat{y} = \underset{\varepsilon_r}{\mathrm{best}}(\vec{P}_{k_r}^{(r)})$,

где $\vec{P}_{ki}^{(i)}$, $i = \overline{0;r}$ – вектор частных приближений, порождаемый на i-м шаге построения описания, состоящий из элементов;

$\vec{P}_{sel}^{(i)}$, $i = \overline{0;r}-1$ – вектор лучших по ошибке восстановления тестовой последовательности частных приближений, полученных на i-м шаге, – результат выполнения процедуры селекции.

Заметим, что с ростом номера этапа i количество аргументов предыдущих этапов, вовлекаемых в построение модели, имеет тенденцию к уменьшению, поскольку погрешность этапов ε_i уменьшается и селектирующее требование становится более жёстким.

В качестве иллюстрации построения формального описания по предложенному модифицированному алгоритму МГУА ниже приведены модели зависимостей *нагрузки на собственные нужды* и *условно-постоянных затрат* от технико-экономических показателей, описывающих условия производственно-экономической деятельности тепловой электростанции.

Представление формального описания в виде псевдоформул

Результат поиска формального описания методом самоорганизации с использованием полиномов в качестве интегрирующего ядра представляет собой некоторый алгоритм, позволяющий вычислять модельные значения результативного показателя для конкретных значений факторных показателей. Естественно, это формальное описание может быть представлено и в виде конкретного полинома Колмогорова – Габора. Конкретика полинома определяется набором членов, значениями коэффициентов и составом аргументов. Однако, получение классического полиномиального представления найденной модели особого практического смысла не имеет по следующим причинам:

1) необходимы не совсем тривиальные символьные преобразования;

2) переход от алгоритмического описания к полиномиальному сопряжён с дополнительными ошибками округления в значениях коэффициентов;

3) сам получаемый полином Колмогорова – Габора выполняет чисто декоративную роль, поскольку его применение для вычислений нерационально из-за указанных выше дополнительных ошибок округления;

4) даже традиционные коэффициенты эластичности могут быть более точно вычислены по алгоритмической модели как отношение приращения моделируемой величины к приращению одного из аргументов в окрестности интересующего набора значений факторных показателей.

Для получения общего аналитического представления о финальной алгоритмической модели более целесообразным является использование предлагаемых ниже двух вариантов псевдоформул. Псевдоформула состоит из композиционной формы записи полинома и двух ассоциированных таблиц значений коэффициентов.

Первый вариант композиционной формы представляет собой символическое обозначение полинома вида $P_n(u, v)$, аргументами которого являются другие символические обозначения полиномов того же вида или символическое обозначение $P_n(x_i)$ полинома с одним аргументом – результативным показателем x_i. Полиномы с одним аргументом являются терминальными элементами композиции. Они вычисляются все одновременно на нулевом этапе алгоритма. Очерёдности вычисления остальных полиномов явно не указываются и определяются косвенно по расположению скобок с аргументами.

Второй вариант композиционной формы содержит символические обозначения только терминальных полиномов вида $P_n(x_i)$. Вместо полиномов с двумя аргументами используется символ конструктора $\bullet_n^k$, в котором n – это порядок полинома, а k – порядок (очередь) срабатывания конструктора.

Очерёдности вычисления можно рассматривать как номера полиномов в композиции и как номера пар их аргументов. Формально конструктор определяется так:

$$u \bullet_k^n v = \sum_{i=0}^{n} \sum_{j=0}^{n-i} a_{(2n-i+3)i/2+j} u^i v^j,$$

где u, v — аргументы полинома.

Заметим, что номер полинома не влияет на формулу его вычисления.

Индекс коэффициента a определяется выбранным способом нумерации пар индексов i и j: пары при их естественном рассмотрении (j изменяется быстрее i), нумеруются последовательно друг за другом, начиная с нуля. При этом, поскольку верхнее значение j является переменным и равно $n-i$, пара (i, j) получает линейный номер, вычисляемый по формуле:

$$index(i, j) = \begin{cases} j, \text{ если } i = 0; \\ \sum_{l=1}^{i} (n - l + 2) + j, \text{ если } i \geq 1. \end{cases}$$

Далее после применения формулы членов арифметической прогрессии получаем:

$$index(i, j) = (2n - i + 3)i / 2 + j\,.$$

Нетрудно убедиться в том, что указанный индекс последовательно пробегает значения от 0 до $N = (n+3)n/2$, где N — общее число всех возможных слагаемых полинома n-й степени с двумя аргументами.

Композиционная форма записи полинома наглядно представляет финальный набор эффективных аргументов и очерёдности вычисления полиномов с двумя аргументами. Необходимая числовая информация (значения коэффициентов и погрешности восстановления тестовой последовательности) размещается в ассоциированных таблицах. Образуемая композиционной формой и таблицами псевдоформула полностью описывает алгоритмическую модель, полученную в процессе самоорганизации, и легко воспроизводится программно.

Модель зависимости мощности, необходимой для собственных нужд, от технико-экономических показателей ГРЭС

Часть электроэнергии, производимой на электростанции, используется самой электростанцией для подачи питательной и циркуляционной воды, работы мельниц, дымососов, дутьевых вентиляторов и т.д. Расход энергии на эти потребности электростанции, называемые собственными нуждами, составляет 3–10 % производимой электроэнергии. Параметры, определяющие изменение расхода мощности на собственные нужды ($N_{\text{собст}}$):

- $N_{\text{макс1}}$, *Nмакс2*, $N_{\text{макс}}$, $N_{\text{ср1}}$, $N_{\text{ср2}}$, $N_{\text{ср3}}$, $N_{\text{ср4}}$, $N_{\text{ср5}}$, $N_{\text{ср6}}$ — электрические характеристики вспомогательного оборудования; с увеличением нагрузки энергоблоков увеличивают потребляемую мощность, при этом наиболее существенно увеличение потребляемой мощности при максимуме нагрузки;
- $N_{\text{соб1}}$, $N_{\text{соб2}}$, $N_{\text{соб3}}$, $N_{\text{соб4}}$, $N_{\text{соб5}}$, $N_{\text{соб6}}$ — составляющие мощности на собственные нужды станции, отражают потребность в мощности на собственные нужды каждого энергоблока в отдельности;
- W_t — затраты электроэнергии оборудования, участвующего только в процессе производства теплоэнергии;
- $W_{\text{хн}}$, $W_{\text{пр}}$ — затраты электроэнергии на хозяйственные нужды и собственных потребителей;
- $W_{\text{пот}}$, $W_{\text{небал}}$ — потери при передаче и учете энергии;
- $T_{\text{раб}}$ — параметр нормативного метода учета электроэнергии;
- $B_{\text{пуск1}}$, $K_{\text{пуск1}}$, $B_{\text{пуск2}}$, $K_{\text{пуск2}}$ — увеличение затрат мощности на пуски энергоблоков по разному типу оборудования;
- $t_{\text{взд}}$ — изменение потребления мощности как от температуры наружного воздуха в котлотурбинном цехе, так и от периода года.

Исходные данные для анализа представляют собой статистические материалы по Рязанской ГРЭС. Результаты предварительного корреляционного анализа позволяют сделать следующие выводы:

- не все факторные признаки коррелированы с результативным признаком;
- коррелированность $N_{\text{собст}}$ с $N_{\text{макс}}$ и $N_{\text{макс1}}$ подтверждает сильную зависимость затрат на собственные нужды станции от пиковых нагрузок; некоррелированность $N_{\text{собст}}$ с $N_{\text{макс2}}$ связана с тем, что блоки второй очереди экономичней блоков первой очереди и работают на оптимальных режимах;
- коррелированность $N_{\text{собст}}$ с W_t и $t_{\text{взд}}$ отражает тот факт, что потребление мощности на собственные нужды носит сезонный характер;
- коррелированность $N_{\text{собст}}$ с $T_{\text{раб}}$ указывает на преобладание в фактическом учете электроэнергии нормативного метода, при котором отнесение затрат электроэнергии на объект в связи с отсутствием приборов учета производится по нормативу, а не по фактическим затратам;

- ✧ коррелированность $N_{\text{собст}}$ с $W_{\text{пот}}$ отражает физический процесс – с увеличением нагрузки увеличиваются потери в трансформаторе;
- ✧ факторные признаки коррелированы между собой.

Зависимость мощности, необходимой для собственных нужд, от факторов, перечисленных выше и пронумерованных от x_1 до x_{26} в порядке их перечисления, в виде композиционной формулы первого вида записывается следующим образом:

$$y = P_1(P_1(P_1(P_2(P_4(x_{16}),\ P_3(x_{26})),\ P_1(P_2(x_{16}),\ P_3(x_1))),$$
$$P_1(x_{21})),\ P_1(P_2(P_3(x_{19}),P_3(x_2)),P_2(P_4(x_3),P_2(x_{14})))).$$

В виде композиционной формулы второго вида эта зависимость имеет вид:

$$y = ((((P_4(x_{16})\bullet_1^2 P_3(x_{26}))\bullet_3^1 (P_2(x_{16})\bullet_2^1 P_3(x_1)))\bullet_4^1 P_1(x_{21}))\bullet$$
$$\bullet_8^1 ((P_3(x_{19})\bullet_5^2 P_3(x_2))\bullet_7^1 (P_4(x_3)\bullet_6^2 P_2(x_{14})))).$$

Параметры модели представлены в *табл. 1* и *2*, где ε – погрешность аппроксимации тестовой последовательности; *k* – номер объединения полиномов в пары операций. В первой таблице представлены параметры полиномов с одним аргументом, полученные на нулевом шаге алгоритма синтеза описания. Во второй таблице приведены параметры полиномов с двумя аргументами, полученные на последующих шагах. Заметим, что процедура селекции привела к тому, что из 26 исходных факторов в финальную модель вошло только 8 из них (№№1, 2, 3, 14, 16, 19, 21, 26).

Таблица 1

Параметры терминальных полиномов с одним аргументом

Номер аргумента	Порядок полинома	ε, %	Коэффициенты a_i при x^i
16	4	15	a_0=3,5; a_1= -233,5; a_2=0,345; a_3=0,089; a_4=96,5
26	3	16	a_0= -31,78; a_1=0,67; a_2= -0,96; a_3=32,57
1	3	19	a_0=562,76; a_1=0,02; a_2=27,74; a_3= -0,06
3	4	22	a_0=3,59; a_1= -454,98; a_2=90,76; a_3=925,62; a_4= -0,35
21	1	26	a_0=13,61; a_1=32,25
19	3	27	a_0= -835,51; a_1=0,73; a_2= -43,09; a_3=178,54
2	3	27	a_0= -0,36; a_1=0,89; a_2=8,14; a_3=38,29
14	2	29	a_0= -36,95; a_1=3298,53; a_2= -27,68

Таблица 2

Параметры полиномов с двумя аргументами

k	Порядок полинома n	ε, %	Коэффициенты $a_{(2n-i+3)i/2+j}$ полинома, *i* - степень 1-го аргумента, *i*=0, …, *n*; *j* - степень 2-го аргумента, *j*=0, …, *n-i*
1	2	10	a_0=0,01; a_1= -233,5; a_2=0,345; a_3=0,089; a_4=96,5; a_5=96,5
2	2	9	a_0=43,5; a_1= -4,23; a_2= -78,35; a_3=12,08; a_4=0,005; a_5=9,56
3	1	8	a_0= -0,0013; a_1= 963,85; a_2= -1233,72
4	2	7	a_0=38,45; a_1= -835,5; a_2=0,31; a_3=1,078; a_4=166,5; a_5= -26,35
5	2	6	a_0=788,5; a_1= -2,05; a_2=0,0003; a_3= -0,189; a_4=45,3; a_5=966,85
6	2	6	a_0=31,25; a_1= -23,5; a_2=1,345; a_3=7,059; a_4=823,43; a_5=16,83
7	1	5	a_0=5,15; a_1= -3373,35; a_2= -247,12
8	1	5	a_0=3347,5; a_1= -0,0028; a_2= -2534,5

Модель зависимости условно-постоянных затрат от технико-экономических показателей работы ГРЭС

К условно-постоянным затратам относятся затраты, которые не изменяются в зависимости от роста или сокращения объема производства. В соответствии с положением о калькулировании себестоимости электрической и тепловой энергии к условно-постоянным затратам относятся все затраты, связанные с производством, за исключением затрат на топливо. Факторы, влияющие на уменьшение или увеличение условно-постоянных затрат:

- ✧ W_1, W_2, W_t, – объемы производства электроэнергии на оборудования первой и второй очередей электростанции и объем производства тепловой энергии;
- ✧ *Мпер* – объем перевозок железнодорожным транспортом;
- ✧ З*, ТП – налоги, уплачиваемые от себестоимости продукции и объема товарной продукции;
- ✧ $S_{\text{фонд}}$ – амортизационные отчисления и налоги, уплачиваемые от стоимости основных фондов;
- ✧ Ч, S_t – затраты на оплату труда и налоги, уплачиваемые от заработной платы;
- ✧ $K_{\text{тл}}$ – общая обеспеченность предприятия оборотными средствами для ведения хозяйственной деятельности;
- ✧ $K_{\text{осс}}$ – наличие собственных оборотных средств у предприятия, необходимых для финансовой устойчивости;

- $K_{\text{впс}}$ – отражает возможность предприятия восстановить свою платежеспособность; коэффициенты $K_{\text{тл}}$, $K_{\text{осс}}$, $K_{\text{впс}}$ в своей совокупности отражают внешние и внутренние факторы, влияющие на производство;
- $t_{\text{нв}}$ – сезонность затрат;
- Д – уровень инфляции в рассматриваемом периоде; за базовый период принят 1995 г.; коэффициенты приведены к базовому периоду;
- $К_{\$}$ – влияние курса рубля на стоимость материальных ресурсов и услуг производственного характера.

Исходные данные для анализа представляют отчетные материалы по Рязанской ГРЭС. Результаты предварительного корреляционного анализа показывают:
- не все факторные признаки коррелированы с результативным признаком;
- коррелированность УП с $З^*$, ТП и $S_{\text{фонд}}$ указывает на наличие влияния налогов на УП;
- коррелированность УП с $S_{\text{фонд}}$ подтверждает зависимость УП от стоимости основных фондов;
- коррелированность УП с Ч и S_t отражает влияние трудозатрат на УП;
- коррелированность УП с Ч и S_t подтверждает существование инфляции и зависимость производства от курса рубля к доллару;
- отсутствие коррелированности УП с $K_{\text{тл}}$, $K_{\text{осс}}$, $K_{\text{впс}}$ означает, что внешние факторы и финансовое состояние предприятия не влияют на УП; это может быть следствием того, что электростанции являются предприятиями-монополистами, кроме того, производство электроэнергии и теплоэнергии имеет исключительную особенность – его продукция всегда реализуется на 100 %;
- факторные признаки коррелированы между собой.

Зависимость условно-постоянных затрат от факторов, перечисленных выше и пронумерованных от x_1 до x_{18} в порядке их перечисления, в виде композиционной формулы первого вида записывается следующим образом:

$$y = P_1(P_1(P_1(P_2(P_3(x_5),\ P_3(x_6)), P_1(P_2(x_9),\ P_2(x_2))),$$
$$P_2(P_3(x_{14}),\ P_3(x_{15}))),\ P_2(P_2(x_2),\ P_3(x_5))).$$

В виде композиционной формулы второго вида эта зависимость имеет вид:

$$y = ((((P_3(x_5) \bullet_1^2 P_3(x_6)) \bullet_3^1 (P_2(x_9) \bullet_2^1 P_2(x_2))) \bullet_5^1 (P_3(x_{14}) \bullet$$
$$\bullet_4^2 P_3(x_{15}))) \bullet_7^1 (P_2(x_2) \bullet_6^2 P_3(x_5))).$$

Параметры модели представлены в *табл. 3* и *4*. Заметим, что процедура селекции привела к тому, что из 18 исходных факторов в финальную модель вошло только 6 из них (№№2, 5, 6, 9, 14, 15).

Таблица 3

Параметры терминальных полиномов с одним аргументом

Номер аргумента	Степень полинома	ε, %	Коэффициенты ai при xi
5	3	8	a_0=17,35; a_1= -3,15; a_2=41,35; a_3=8,09; a_4=31,69
6	3	9	a_0=12,17; a_1= 29,83; a_2= -0,08; a_3= -7,50; a_4=0,91
9	2	11	a_0=32,96; a_1= -456,2; a_2=442,74; a_3= -0,006
2	2	17	a_0=3212,56; a_1= -43,76; a_2=0,077; a_3=57,02
14	3	18	a_0=0,043; a_1= 4444,55; a_2=0,001; a_3=40,09; a_4= -906,03
15	3	19	a_0= -11,03; a_1= -185,61; a_2=0,21; a_3=20,94; a_4=91,31

Таблица 4

Параметры полиномов с двумя аргументами

k	Степень полинома n	ε, %	Коэффициенты $a_{(2n-i+3)i/2+j)}$ полинома, i - степень 1-го аргумента, i=0, …, n; j - степень 2-го аргумента, j=0, …, $n-i$
1	2	14	a_0=44,14; a_1= -17,24; a_2=33,31; a_3=77,019; a_4=61,89; a_5= -262,03
2	1	13	a_0= -1,101; a_1= 3,07; a_2= -187,92
3	2	10	a_0=88,01; a_1= -204,79; a_2=412,59; a_3= -29,48; a_4=76,15; a_5= -9,03
4	2	10	a_0=88,81; a_1= -2,19; a_2= -42,19; a_3= -1799,08; a_4=0,004; a_5=6,013
5	1	7	a_0= -94,33; a_1= 376,94; a_2= -244,24
6	1	6	a_0= -10,03; a_1= 8,012; a_2= -13,30
7	1	6	a_0=0,093; a_1= -3,26; a_2= 771,25

Заключение

Получены следующие основные теоретические результаты:

1) определены направления развития метода группового учёта аргументов, обеспечивающие повышение его селектирующих способностей и точности аппроксимации функции векторного аргумента;

2) разработаны модифицированные алгоритмы метода группового учёта аргументов, повышающие его эффективность посредством оптимизации частных полиномов; высокое качество предложенных алгоритмов (способность надёжно выделять эффективные аргументы, приемлемая точность интерполяции и экстраполяции процессов и высокая производительность) подтверждено методом моделирования; хорошие интерполяционные и экстраполяционные свойства предложенных алгоритмов позволяют использовать их не только для прогноза методом сценарных условий и лагированных переменных, но и для анализа по принципу «что — если»;

3) предложена оригинальная легко реализуемая программно нотация записи алгоритмических описаний, получаемых в результате реализации разработанных алгоритмов, в виде псевдоформул с композиционной формой записи полиномов двух видов, позволяющих наглядно представлять состав актуальных аргументов, порядки частных полиномов, вид частных приближений и последовательность их формирования.

Указанные результаты пополняют арсенал эффективных средств специалистов, занятых разработкой и анализом прогнозов социально-экономических процессов.

Основным практическим результатом является следующее:

1) построены алгоритмические модели, позволяющие исследовать взаимосвязь между результативными и факторными показателями социально-экономической деятельности крупного энергогенерирующего предприятия ОАО «Рязанская ГРЭС» (две модели приведены в данной статье в качестве примеров);

2) построенные модели использованы для разработки прогноза социально-экономического статуса ОАО «Рязанская ГРЭС»;

3) результаты анализа полученных моделей и прогнозных значений проблемных показателей использованы для коррекции и планирования социально-экономической политики предприятия ОАО «Рязанская ГРЭС». ■

Литература

1. Айвазян С.А., МхитарянВ.С. Прикладная статистика и основы эконометрики. М.:ЮНИТИ, 1998. 1024 с.
2. Себер Дж. Линейный регрессионный анализ: Пер с англ. М.: Мир, 1980. 456 с.
3. Уоссерман Ф. Нейрокомпьютерная техника: Теория и практика: Пер. с англ. М.: Мир, 1992. 240 с.
4. Габор Д. Перспективы планирования // Автоматика. 1972. № 2.
5. Ивахненко А. Г. и др. Принятие решений на основе самоорганизации. М.: Сов. радио, 1976. 280 с.
6. Тамура, Кондо. Современная методология групповой обработки данных и ее приложения // Оперсэндзу рисати. 1987. № 2. С. 104 – 111.
7. Прикладные нечеткие системы: Пер. с япон. / К. Асаи, Д. Ватада, С. Иваи и др. Под редакцией Т. Тэрано, К. Асаи, М. Сугано. М.: Мир, 1993. 368 с.
8. Белов В.В., Васильев С.В., Наумкина С.Г. Модифицированный метод группового учета аргументов на основе процедуры оптимизации частных полиномов // Вычислительные машины, комплексы и сети: Межвуз. сб. науч. трудов. Рязань: РГРТА, 1999. С. 95–99.