УДК 683.513

ПАВЛОВ А.В.

МНОГОЭТАПНЫЙ АЛГОРИТМ МГУА С ОРТОГОНАЛИЗАЦИЕЙ ПЕРЕМЕННЫХ И ЕГО РЕКУРРЕНТНЫЙ МЕТОД РАССЧЁТА КОЭФФИЦИЕНТОВ И КРИТЕРИЯ СЕЛЕКЦИИ МОДЕЛЕЙ

Найдены условия, при которых обобщённый релаксационный итерационный алгоритм моделирования находит точное решение и показано, что при выполнении этих условий он принадлежит классу переборных алгоритмов МГУА с направленным перебором. Предложен многоэтапный алгоритм с ортогонализацией переменных и рекуррентный метод оценивания параметров и расчёта критерия селекции. Сложность расчёта одной модели на каждом этапе алгоритма является постоянной, в отличие квадратичной в известном алгоритме за счет применения ортогонализации переменных и рекуррентных вычислений.

Conditions under which generalized relaxational iterative algorithm finds exact solution have been found; belonging the algorithm to class of multistage GMDH algorithms has been proved in this case. Multistage algorithm with features orthogonalization and recurrent method for models' parameters and selection criterion estimation has been suggested. Model calculation complexity at each algorithm's stage is constant in contrast to quadratic one for well-known algorithm due to features orthogonalization and recurrent computations.

1. Введение

Прикладные задачи принятия решений часто решаются на основе регрессионных моделей (далее просто моделей) объектов и процессов, построенных по эксперименталь-ным данным. Такими задачами, например, являются прогнозирование, экстраполирова-ние, классификация. Одним из эффективных методов построения моделей является метод группового учета аргументов (МГУА). Различают итерационные и переборные алгоритмы МГУА.

Переборные алгоритмы в зависимости от генератора структур могут находить модель глобального минимума внешнего критерия, либо нет – как в случае направленного (ограниченного) поиска модели способом вложенных структур. Поэтому в настоящее время переборные алгоритмы делятся на комбинаторные, реализующие полный пере-бор структур моделей и, те которые выполняют направленный перебор, называе-мые далее многоэтапными. Если комбинатор-ные алгоритмы, которые всегда находят модель, соответствующую глобальному оптимуму выбранного критерия, имеют «проклятие размерности» по количеству переменных, то многоэтапные и итерацион-ные – могут обрабатывать значительно большее количество аргументов (до 1000), но находят при этом решение, соответствующее локальному минимуму внешнего критерия. Отличие многоэтапных алгоритмов от итерационных заключатся в том, что многоэтапные находят структуру и параметры модели (точное решение) за конечное число этапов.

Среди итерационных алгоритмов выделяется класс релаксационных итерационных алгоритмов, основной особенностью которого является способ усложнения структуры модели от итерации к итерации: ошибка между вектором выхода модели (решением) текущей итерации и исходным вектором выходной переменной уменьшается за счёт явного добавления в модель исходных аргументов на каждой итерации.

Метод усложнения структуры модели, в случае добавления в неё на каждом этапе (или итерации) одного отсутствующего из множества исходных аргументов назван мето-дом вложенных структур. Между моделями, полученными релаксационным и много-этапным и алгоритмами, использующие метод вложенных структур, при условии их (структур) совпадения, имеется различие в величинах оценок параметров. Итеративно рассчитываемые оценки в итерационных алгоритмах, в общем случае не совпадают с оценками для той же структуры в многоэтапных алгоритмах, которые вычис-ляются по методу наименьших квадратов (МНК). Следовательно, в отличие от многоэтапных алгоритмов, итерационные не дают точного решения.

Поэтому актуальной задачей является поиск условий, при которых релаксационные итерационные алгоритмы также находят точное решение. Работа посвящена поиску получения точных решений для обобщённого релаксационного итерационного алгоритма (ОРИА), идентичных многоэтапному алгоритму. Дадим общую для обоих алгоритмов МГУА постановку задачи.

2. Постановка задачи МГУА

Пусть задана матрица данных $\mathbf{X}$, $\dim \mathbf{X} = n_W \times m$, где m – число признаков (аргументов) x_i, $i = \overline{1,m}$, n_W – число наблюдений, и выходной вектор $\mathbf{y}$, $\dim \mathbf{y} = n_W \times 1$. Необходимо найти оптималь-ную по минимуму заданного критерия CR модель $\hat{\mathbf{y}}_{\mathbf{d}_k^*} = \mathbf{f}(\mathbf{Z}_{\mathbf{d}_k^*}, \widehat{\Theta}_{\mathbf{d}_k^*})$:

$$\mathbf{f}(\mathbf{Z}_{\mathbf{d}_k^*}, \widehat{\Theta}_{\mathbf{d}_k^*}) = \arg\min_{\mathbf{d}_k \in D} CR(\mathbf{y}, \mathbf{f}(\mathbf{Z}_{\mathbf{d}_k}, \widehat{\Theta}_{\mathbf{d}_k})), \quad (1)$$

где оценки параметров $\Theta_{\mathbf{d}_k}$, $\forall$ $\mathbf{d}_k \in D$ являются решением задачи:

$$\widehat{\Theta}_{\mathbf{d}_k} = \arg\min_{\Theta_{\mathbf{d}_k} \in \Re^{s_k}} QR(\mathbf{y}, \mathbf{f}(\mathbf{Z}_{\mathbf{d}_k}, \Theta_{\mathbf{d}_k})), \quad (2)$$

$\mathbf{d}_k$ – бинарный вектор, который определяет подмножество аргументов модели; D – множество всех возможных бинарных векторов размерности l; $\Theta_{\mathbf{d}_k}$ – вектор размерности s_k, составленный из ненулевых компонентов вектора $diag(\mathbf{d}_k)\cdot\Theta$; $\mathbf{Z}_{\mathbf{d}_k}$ – матрица, полученная из вектор-столбцов $\mathbf{z}_j$ матрицы $\mathbf{Z}$, где элементы вектора $\mathbf{d}_{k,j} = 1$, а остальные компоненты вектора $\mathbf{d}_k$ – нулевые; $\mathbf{Z}$ – матрица, образованная из вектор-столбцов, соответствующих переменным в полиномиальной модели вида:

$$f(x_1,...,x_m,\Theta) = \theta_0 + \sum_{i=1}^{m} \theta_i x_i + \sum_{i=1}^{m}\sum_{j=i}^{m} \theta_{i,j} x_i x_j + + \sum_{i=1}^{m}\sum_{j=i}^{m}\sum_{k=i+j}^{m} \theta_{i,j,k} x_i x_j x_k + ...,$$

где θ_i, $\theta_{i,j}$, $\theta_{i,j,k}$,… являются коэффициентами при этих переменных; p – заданная максимальная степень полинома; $\dim \mathbf{Z} = n_W \times l$, $l = C_{m+p}^{p} - 1$. Пусть $QR(\cdot)$ – критерий качества решения задачи оценивания параметров, $CR(\cdot)$ – критерий качества решения задачи определения оптимальной модели.

В качестве критерия $CR(\cdot)$ в МГУА используют «внешние» критерии, основанные на разбиении выборки $\mathbf{W} = (\mathbf{X} \vdots \mathbf{y})$, например, критерий регулярности:

$$AR_B \overset{\Delta}{=} AR_{B/A} = \| \mathbf{y}_B - \mathbf{Z}_{B,\mathbf{d}_k} \widehat{\Theta}_{A,\mathbf{d}_k} \|^2,$$

где запись $AR_{B/A}$ указывает на то, что на выборке A при заданной структуре находятся оценки вектора параметров $\Theta_{A,\mathbf{d}_k}$, а на B рассчитывается ошибка модели как значение критерия. Выборка A называется обучающей, а B – проверочной, $W = A \cup B$, $A \cap B = \varnothing$. Задача (2) решается по МНК. $QR(\cdot)$ – критерий остаточной суммы наименьших квадратов (residual sum square (RSS)).

Перейдём к анализу ОРИА описанному в [1].

3. Анализ обобщённого релаксационного алгоритма МГУА

Алгоритм строит модель итерационным способом, уменьшая невязку между решением r-й итерации $\hat{\mathbf{y}}_{A,r}$ и выходом $\mathbf{y}_A$, при добавлении одного из вектор-столбцов $\mathbf{x}_{A,r+1}$ матрицы $\mathbf{X}_A$. Задача (2), решаемая в данном алгоритме на $(r+1)$-й итерации, в матричной форме выглядит следующим образом:

$$\mathbf{y}_A = \omega_{A,r+1}^{r+1} \mathbf{x}_{A,r+1} + \omega_{A,r+1} \hat{\mathbf{y}}_{A,r}, \quad (3)$$

где $\omega_{A,r+1}^{r+1}$, $\omega_{A,r+1}$ – неизвестные параметры, оценки которых определяются по МНК; $\mathbf{x}_{A,r+1}$ – вектор-столбец матрицы $\mathbf{X}_A$, добавляемы в модель на $(r+1)$-й итерации; $\hat{\mathbf{y}}_{A,r}$ – решение, полученное на предыдущей итерации.

Матрица нормальных уравнений для (3) имеет следующий вид:

$$\begin{pmatrix} \hat{\mathbf{y}}_{A,r}^T \hat{\mathbf{y}}_{A,r} & \hat{\mathbf{y}}_{A,r}^T \mathbf{x}_{A,r+1} \\ \mathbf{x}_{A,r+1}^T \hat{\mathbf{y}}_{A,r} & \mathbf{x}_{A,r+1}^T \mathbf{x}_{A,r+1} \end{pmatrix} \begin{pmatrix} \omega_{A,r+1}^{r+1} \\ \omega_{A,r+1} \end{pmatrix} = \begin{pmatrix} \hat{\mathbf{y}}_{A,r}^T \mathbf{y}_A \\ \mathbf{x}_{A,r+1}^T \mathbf{y}_A \end{pmatrix}.$$

Оценки коэффициентов $\widehat{\omega}_{A,r+1}^{r+1}$, $\widehat{\omega}_{A,r+1}$ рассчитываются по формулам:

$$\widehat{\omega}_{A,r+1} = \frac{c_{A,r}(\mathbf{x}_{A,r+1}^T \mathbf{x}_{A,r+1}) - b_{A,r}(\mathbf{x}_{A,r+1}^T \mathbf{y}_A)}{a_{A,r}(\mathbf{x}_{A,r+1}^T \mathbf{x}_{A,r+1}) - b_{A,r}^2}, \quad (4)$$

$$\widehat{\omega}_{A,r+1}^{r+1} = \frac{a_{A,r}(\mathbf{x}_{A,r+1}^T \mathbf{y}_A) - b_{A,r} c_{A,r}}{a_{A,r}(\mathbf{x}_{A,r+1}^T \mathbf{x}_{A,r+1}) - b_{A,r}^2}, \quad (5)$$

где $a_{A,r} = \hat{\mathbf{y}}_{A,r}^T \hat{\mathbf{y}}_{A,r}$, $b_{A,r} = \hat{\mathbf{y}}_{A,r}^T \mathbf{x}_{A,r+1}$, $c_{A,r} = \hat{\mathbf{y}}_{A,r}^T \mathbf{y}_A$ рассчитаны на r-й итерации.

Решение $(r+1)$-й итерации:

$$\hat{\mathbf{y}}_{A,r+1} = \widehat{\omega}_{A,r+1}^{r+1} \mathbf{x}_{A,r+1} + \widehat{\omega}_{A,r+1} \hat{\mathbf{y}}_{A,r}.$$

Критерий RSS_A рассчитывается по

$$RSS_A = \mathbf{y}_A^T \mathbf{y}_A - 2c_{A,r+1} + a_{A,r+1}, \quad (6)$$

а величины $a_{A,r+1}$ и $c_{A,r+1}$ рекуррентно:

$$a_{A,r+1} = a_{A,r}(\widehat{\omega}_{A,r+1})^2 + 2\widehat{\omega}_{A,r+1}\widehat{\omega}^{r+1}_{A,r+1}b_{A,r} + \\ + (\widehat{\omega}^{r+1}_{A,r+1})^2 \mathbf{x}^T_{A,r+1}\mathbf{x}_{A,r+1}, \quad (7)$$

$$c_{A,r+1} = c_{A,r}\widehat{\omega}_{A,r+1} + \mathbf{x}^T_{A,r+1}\mathbf{y}_A\widehat{\omega}^{r+1}_{A,r+1}, \quad (8)$$

$$b_{A,r} = \widehat{\mathbf{y}}^T_{A,r}\mathbf{x}_{A,r+1} = \mathbf{x}_{A,r+1}\mathbf{G}_{A,r}(\widehat{\Omega}^r_{A,r})^T = \\ = \sum_{j=1}^{r}(\mathbf{x}^T_{A,r+1}\mathbf{g}_{A,j})\cdot\widehat{\omega}^j_{A,r},$$

где $\mathbf{G}_{A,r}$ – матрица, состоящая из вектор-столбцов матрицы $\mathbf{X}_A$, вошедших в модель на первых r итерациях; $\widehat{\Omega}^r_{A,r}$ – вектор коэффициентов, элементы которого соответствуют вектор-столбцам матрицы $\mathbf{G}_{A,r}$. Верхний индекс $\widehat{\Omega}^r_{A,r}$ указывает размерность вектора, а нижний – номер итерации.

После расчёта оценок коэффициентов $\widehat{\omega}^{r+1}_{A,r+1}$, $\widehat{\omega}_{A,r+1}$ корректируются коэффициенты предыдущих r итераций при вектор-столбцах матрицы $\mathbf{G}_{A,r}$:

$$\widehat{\Omega}^r_{A,r+1} = \widehat{\Omega}^r_{A,r}\widehat{\omega}_{A,r+1}.$$

Критерий регулярности в ОРИА рассчитывается по формуле:

$$AR_{B,r+1} = \mathbf{y}^T_B\mathbf{y}_B - 2c_{B,r+1} + a_{B,r+1}, \quad (9)$$

где величины $c_{B,r+1}$ и $a_{B,r+1}$ рассчитываются по формулам (7), (8) на выборке B.

Скорость сходимости ОРИА существенно зависит от степени ортогональности вектор-столбцов матрицы $\mathbf{X}_A$, а именно: чем более ортогональной является система, тем быстрей скорость сходимости алгоритма [2] . Докажем, что: максимальная скорость сходимости ОРИА к точному решению достигается при ортогональной системе вектор-столбцов матрицы $\mathbf{X}_A$; решение, к которому сходится алгоритм совпадает с истинной моделью (точным решением).

4. Доказательство сходимости алгоритма к точному решению

Утверждение 1. Минимальное количество итераций ОРИА, необходимое для внутренней сходимости алгоритма к точному решению

$$\mathbf{y}_A = \breve{f}(\mathbf{X}_A, \breve{\Theta}) = \breve{\theta}_0 + \sum_{i=1}^{m}\mathbf{x}_{A,i}\breve{\theta}_i$$

по критерию RSS_A равно m и это решение достигается при условии, что вектор-столбцы матрицы $\mathbf{X}_A$ являются ортогональной системой.

Доказательство. Пусть $\mathbf{y}_A = \mathbf{X}_A\Theta_A$, $m = 3$ и вектор-столбцы матрицы $\mathbf{X}_A$ взаимно ортогональны (рис. 1).

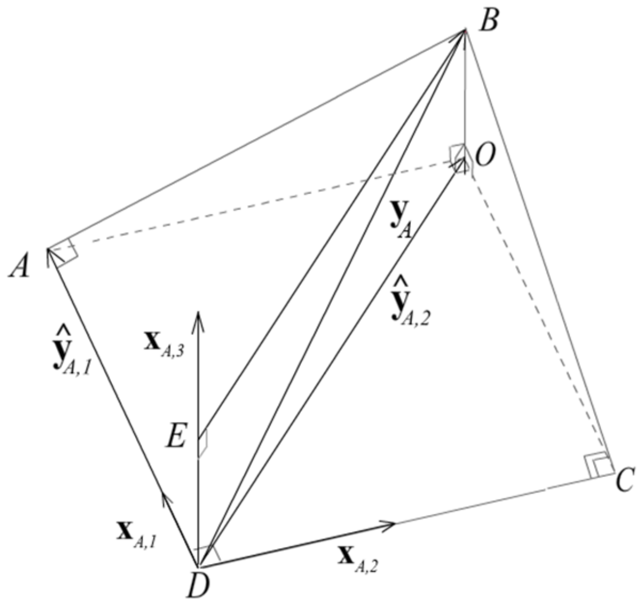

Рис. 1. Процесс построения модели алгоритмом при взаимно ортогональных вектор-столбцах матрицы $\mathbf{X}_A$, $m = 3$

В данном случае скалярное произведение $\widehat{\mathbf{y}}^T_{A,2}\mathbf{x}_{A,3} = 0$, поскольку вектор $\widehat{\mathbf{y}}_{A,2}$ лежит в плоскости векторов $\mathbf{x}_{A,1}$ и $\mathbf{x}_{A,2}$. При этом невязка $|\mathbf{y}_A - \mathbf{G}_{A,3}\widehat{\Omega}^3_{A,3}| = |\mathbf{y}_A - \mathbf{X}_{A,3}\widehat{\Omega}_{A,3}| = 0$. Рассмотрим общий случай, когда $\mathbf{y}_A = \mathbf{X}_A\Theta_A$. Пусть алгоритм выполнил r итераций, тогда на (r+1)-й итерации $\widehat{\mathbf{y}}^T_{A,r}\mathbf{x}_{A,r+1} = 0$ и невязка $|\mathbf{y}_A - \mathbf{G}_{A,r+1}\widehat{\Omega}^{r+1}_{A,r+1}| = |\mathbf{y}_A - \mathbf{X}_{A,r+1}\widehat{\Omega}_{A,r+1}|$ является минимальным расстоянием между вектором выхода $\mathbf{y}_A$ и вектором решения $\widehat{\mathbf{y}}_{A,r+1} = \sum_{i=1}^{r+1}\widehat{\theta}_{A,i}\mathbf{x}_{A,i}$, которое может быть получено в плоскости, определяемой первыми $r+1$ вектор-столбцами матрицы $\mathbf{X}_A$. Если продолжить описанный процесс до $r = m$, алгоритм сойдётся к вектору $\mathbf{y}_A$. Причём полученные оценки коэффициентов $\widehat{\theta}_{A,i}$, $i = \overline{1,m}$ совпадают с оценками МНК: $\widehat{\Theta}_A = \left(\mathbf{X}^T_A\mathbf{X}_A\right)^{-1}\mathbf{X}^T_A\mathbf{y}_A$.

Опишем метод оценивания коэффициентов и расчёта критерия селекции модели для ОРИА при условии взаимно-ортогональных вектор-столбцов входной матрицы $\mathbf{X}^T = (\mathbf{X}^T_A \vdots \mathbf{X}^T_B)$.

5. Рекуррентный метод оценивания параметров и расчёта критериев селекции

Пусть системы вектор-столбцов матриц $\mathbf{X}_A$ и $\mathbf{X}_B$ является ортогональными, тогда $\hat{\mathbf{y}}_{A,r}^T \mathbf{x}_{A,r+1} = 0$, следовательно $b_{A,r} = 0$. Поскольку треугольник ΔDBO – прямоугольный (рис. 1):

$$\mathbf{y}_A^T \mathbf{y}_A = \hat{\mathbf{y}}_{A,r}^T \hat{\mathbf{y}}_{A,r} + (\mathbf{y}_A - \hat{\mathbf{y}}_{A,r})^T (\mathbf{y}_A - \hat{\mathbf{y}}_{A,r}), \qquad (10)$$

Из выполнения равенства (10) следует, что $\hat{\mathbf{y}}_{A,r}^T \hat{\mathbf{y}}_{A,r} = \hat{\mathbf{y}}_{A,r}^T \mathbf{y}_A$, т.е. $a_{A,r} = c_{A,r}$. Тогда формулы (4) и (5) примут вид:

$$\hat{\omega}_{A,r+1} = 1, , \qquad (11)$$

$$\hat{\omega}_{A,r+1}^{r+1} = \frac{(\mathbf{x}_{A,r+1}^T \mathbf{y}_A)}{(\mathbf{x}_{A,r+1}^T \mathbf{x}_{A,r+1})}. , \qquad (12)$$

С учётом $b_{A,r} = 0$ и (11) формулы (7) и (8) примут вид:

$$a_{A,r+1} = a_{A,r} + (\hat{\omega}_{A,r+1}^{r+1})^2 \mathbf{x}_{A,r+1}^T \mathbf{x}_{A,r+1}, \qquad (13)$$

$$c_{A,r+1} = c_{A,r} + \mathbf{x}_{A,r+1}^T \mathbf{y}_A \hat{\omega}_{A,r+1}^{r+1}, \qquad (14)$$

С учётом $a_{A,r} = c_{A,r}$ формула (6) примет вид:

$$RSS_A = \mathbf{y}_A^T \mathbf{y}_A - a_{A,r+1}.$$

Аналогичным образом формулы для расчёта $a_{B,r+1}$, $c_{B,r+1}$ примут вид:

$$a_{B,r+1} = a_{B,r} + (\hat{\omega}_{A,r+1}^{r+1})^2 \mathbf{x}_{B,r+1}^T \mathbf{x}_{B,r+1}, \qquad (15)$$

$$c_{B,r+1} = c_{B,r} + \mathbf{x}_{B,r+1}^T \mathbf{y}_B \hat{\omega}_{A,r+1}^{r+1}, \qquad (16)$$

Критерий регулярности рассчитывается по той же формуле (9).

Запишем метод оценивания параметров и расчёта критерия селекции подсчитаем его вычислительную сложность – количество операций умножения и деления. Обозначим это число через *OpNum*. Метод разобьем на две стадии: 1) $r = 1$; 2) $r > 1$, r – номер этапа.

Стадия 1. $r = 1$.

1.1. Рассчитаем коэффициент и критерий селекции (критерий регулярности):

$$\hat{\omega}_{A,1}^1 = \frac{\mathbf{x}_{A,1}^T \mathbf{y}_A}{\mathbf{x}_{A,1}^T \mathbf{x}_{A,1}},$$

$$AR_{B,1} = \mathbf{y}_B^T \mathbf{y}_B - 2\hat{\omega}_{A,1}^1 \cdot (\mathbf{x}_{B,1}^T \mathbf{y}_B) + (\hat{\omega}_{A,1}^1)^2 (\mathbf{x}_{B,1}^T \mathbf{x}_{B,1}).$$

1.2. Вычисляем начальные значения для величин $a_{A,1}$, $c_{A,1}$, $a_{B,1}$, $c_{B,1}$ по формулам (*OpNum* = 11):

$$a_{A,1} = (\hat{\mathbf{y}}_{A,1}^T \hat{\mathbf{y}}_{A,1}) = (\mathbf{x}_{A,1}^T \mathbf{x}_{A,1})(\hat{\omega}_{A,1}^1)^2,$$

$$c_{A,1} = (\hat{\mathbf{y}}_{A,1}^T, \mathbf{y}_A) = (\mathbf{y}_A^T \mathbf{x}_{A,1})\hat{\omega}_{A,1}^1,$$

$$a_{B,1} = (\hat{\mathbf{y}}_{B,1}^T \hat{\mathbf{y}}_{B,1}) = (\mathbf{x}_{B,1}^T \mathbf{x}_{B,1})(\hat{\omega}_{A,1}^1)^2,$$

$$c_{B,1} = (\hat{\mathbf{y}}_{B,1}^T, \mathbf{y}_B) = (\mathbf{y}_B^T \mathbf{x}_{B,1})\hat{\omega}_{A,1}^1.$$

Стадия 2. $r = \overline{2, m}$. Запишем алгоритм для (r+1)-го этапа.

2.1 Вычисляем оценку коэффициента $\hat{\omega}_{A,r+1}^{r+1}$ по формуле (12). *OpNum* = 1.

2.2 Вычисляем величины $a_{A,r+1}$, $c_{A,r+1}$, $a_{B,r+1}$, $c_{B,r+1}$ по формулам (13)-(16). *OpNum* = 6.

2.3 Рассчитываем критерий регулярности по формуле (9). *OpNum* = 1.

Общее количество операций для построения модели от m аргументов равно $8m + 3$. При этом сложность расчёта модели на текущем этапе является постоянной, и равна *восьми* элементарным операциям. Как видим вычислительная сложность метода не зависит от количества наблюдений n_W и использует лишь $\mathbf{x}_{A,i}^T \mathbf{x}_{A,i}$, $\mathbf{x}_{B,i}^T \mathbf{x}_{B,i}$, $\mathbf{x}_{A,i}^T \mathbf{y}_A$, $\mathbf{x}_{B,i}^T \mathbf{y}_B$, $i = \overline{1, m}$.

Таким образом, ОРИА, принадлежащий классу итерационных алгоритмов, при использовании предложенного рекуррентного метода построения модели имеет такое же решение как и многоэтапный алгоритм МГУА. Поэтому назовём такой алгоритм Многоэтапным Алгоритмом с Ортогонализацией переменных (МАО).

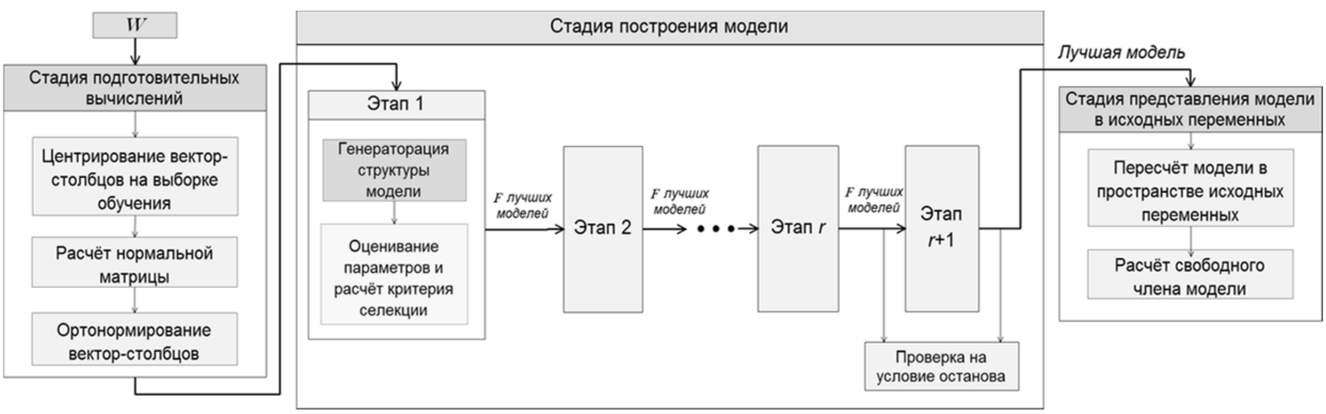

Рис. 2. Процесс построения модели в МАО

6. Многоэтапный алгоритм МГУА с ортогонализацией переменных

Процесс построения функции $f(\mathbf{X},\widehat{\Theta})$ в МАО, такой же, как и в многоэтапном алгоритме MULTI [3]. Общая схема работы алгоритма представлена на рис. 2. Как видим, алгоритм состоит из трёх стадий:

1. Стадия подготовительных вычислений, на которой осуществляется:

1.1 Центрирование вектор-столбцов матрицы $\mathbf{W} = (\mathbf{X} \vdots \mathbf{y})$ на обучающей выборке A для ухода от необходимости оценивания свободного члена моделей по формулам:

$$\widetilde{\mathbf{x}}_{ji} = \mathbf{x}_{ji} - \overline{\mathbf{x}}_{A,i}\,,\; \overline{\mathbf{x}}_{A,i} = \frac{1}{n_A}\sum_{j=1}^{n_A}\mathbf{x}_{ji}\,,\; i = \overline{1,m}\,,\; j = \overline{1,n_W}$$

$$\widetilde{\mathbf{y}}_j = \mathbf{y}_j - \overline{y}_A\,,\; j = \overline{1,n_W}\,,\; \overline{y}_A = \frac{1}{n_A}\sum_{j=1}^{n_A}\mathbf{y}_j\,.$$

1.2 Ортогонализация вектор-столбцов матриц $\mathbf{X}_A$ и $\mathbf{X}_B$ для использования рекуррентного метода построения модели, предложенного в предыдущем разделе.

2. Стадия построения модели в ортогонализированном пространстве перемен-ных. Процесс построения модели следующий. На каждом этапе в текущую модель

$$\widehat{\mathbf{y}}_s(\mathbf{x}_1,...,\mathbf{x}_s,\widehat{\theta}_1,...,\widehat{\theta}_s) = \sum_{i=1}^{s}\mathbf{x}_i\widehat{\theta}_i$$

добавляется два параметра: аргумент $\mathbf{x}_{s+1}$, не содержащийся среди множества аргументов текущей модели и оценка коэффициента $\widehat{\theta}_{s+1}$ при $\mathbf{x}_{s+1}$:

$$\widehat{\mathbf{y}}_{s+1}(\mathbf{x}_1,...,\mathbf{x}_s,\mathbf{x}_{s+1},\widehat{\theta}_1,...,\widehat{\theta}_s) = \sum_{i=1}^{s+1}\mathbf{x}_i\widehat{\theta}_i\,,$$

где s – номер этапа. Оценка коэффициента $\widehat{\theta}_{s+1}$ рассчитывается по формуле (12), а критерий селекции модели по формуле (9). На первом этапе строится множество моделей, зависящих от одного аргумента, на втором – от двух, и т.д. до модели, содержащей все аргументы. На каждом этапе осуществляется выбор F лучших моделей по критерию селекции, которые переходят на следующий этап.

3. Стадия представления модели через исходные признаки, где осуществляется переход в исходное пространство и расчёт свободного члена модели как:

$$\widehat{\theta}_{A,0} = \overline{y}_A - \sum_{i=1}^{m}\widehat{\theta}_{A,i}\overline{\mathbf{x}}_{A,i}\,.$$

Правилом останова алгоритма может быть, например, условие: CR_s и CR_{s+1} – значения критериев селекции лучших из F моделей s-го и (s+1)-го этапа соответственно.

7. Ортогонализация и переход к исходным переменным

Для того чтобы вычислительная сложность метода построения моделей не зависела от количества наблюдений n_W, он должен использовать элементы матриц $\mathbf{W}_A^T\mathbf{W}_A$, $\mathbf{W}_B^T\mathbf{W}_B$, $\mathbf{W}_A = (\mathbf{X}_A \vdots \mathbf{y}_A)$, $\mathbf{W}_B = (\mathbf{X}_B \vdots \mathbf{y}_B)$. Входными данными для стадии построения модели должны быть элементы $\mathbf{z}_{A,i}^T\mathbf{z}_{A,i}$, $\mathbf{z}_{A,i}^T\mathbf{y}_A$, $\mathbf{z}_{B,i}^T\mathbf{z}_{B,i}$, $\mathbf{z}_{B,i}^T\mathbf{y}_B$, $i = \overline{1,m}$ ортогональных матриц $\mathbf{Z}_A$ и $\mathbf{Z}_B$.

Метод ортогонализации. В его основе лежит ортогонализация Грама-Шмидта [4]. Опишем метод пошагово, получая элементы $\mathbf{z}_{A,i}^T\mathbf{z}_{A,i}$, $\mathbf{z}_{A,i}^T\mathbf{y}_A$, $i = \overline{1,m}$ на выборке A, и подсчитаем его вычислительную сложность. Для выборки B он аналогичен.

1) $\mathbf{z}_{A,1} = \mathbf{x}_{A,1}$, $\mathbf{z}_{A,1}^T\mathbf{z}_{A,1} = \mathbf{x}_{A,1}^T\mathbf{x}_{A,1}$, $\mathbf{z}_{A,1}^T\mathbf{y}_A = \mathbf{x}_{A,1}^T\mathbf{y}_A$;

2) $\mathbf{z}_{A,2} = \mathbf{x}_{A,2} - \gamma_{A,1,2}\mathbf{z}_{A,1}$, $\gamma_{A,1,2} = \dfrac{\mathbf{z}_{A,1}^T\mathbf{x}_{A,2}}{\mathbf{z}_{A,1}^T\mathbf{z}_{A,1}}$;

$$\mathbf{z}_{A,2}^T\mathbf{z}_{A,2} = \mathbf{x}_{A,2}^T\mathbf{x}_{A,2} - 2\gamma_{A,1,2}\mathbf{z}_{A,1}^T\mathbf{x}_{A,2} + \gamma_{A,1,2}^2\mathbf{z}_{A,1}^T\mathbf{z}_{A,1}\,.$$

$$\mathbf{z}_{A,2}^T\mathbf{y}_A = \mathbf{x}_{A,2}^T\mathbf{y}_A - \gamma_{A,1,2}\mathbf{z}_{A,1}^T\mathbf{y}_A,$$

3) $\mathbf{z}_{A,3} = \mathbf{x}_{A,3} - \gamma_{A,2,3}\cdot\mathbf{z}_{A,2} - \gamma_{A,1,3}\mathbf{z}_{A,1}$,

$$\gamma_{A,1,3} = \frac{\mathbf{z}_{A,1}^T\mathbf{x}_{A,3}}{\mathbf{z}_{A,1}^T\mathbf{z}_{A,1}}\,,$$

$$\gamma_{A,2,3} = \frac{\mathbf{z}_{A,2}^T\mathbf{x}_{A,3}}{\mathbf{z}_{A,2}^T\mathbf{z}_{A,2}} = \frac{\mathbf{x}_{A,3}^T\mathbf{x}_{A,2} - \gamma_{A,1,2}\mathbf{x}_{A,3}^T\mathbf{x}_{A,1}}{\mathbf{z}_{A,2}^T\mathbf{z}_{A,2}}\,,$$

тогда с учётом $\mathbf{z}_{A,i}^T\mathbf{z}_{A,j} = 0$, $i \neq j$:

$$\mathbf{z}_{A,3}^T\mathbf{z}_{A,3} = \mathbf{x}_{A,3}^T\mathbf{x}_{A,3} + \gamma_{A,2,3}^2\cdot\mathbf{z}_{A,2}^T\mathbf{z}_{A,2} +$$
$$+\gamma_{A,1,3}^2\mathbf{z}_{A,1}^T\mathbf{z}_{A,1} - 2(\gamma_{A,2,3}\mathbf{x}_{A,3}^T\mathbf{z}_{A,2} + \gamma_{A,1,3}\mathbf{x}_{A,3}^T\mathbf{z}_{A,1}),$$
$$\mathbf{z}_{A,3}^T\mathbf{y}_A = \mathbf{x}_{A,3}^T\mathbf{y}_A - \gamma_{A,2,3}\cdot\mathbf{z}_{A,2}^T\mathbf{y}_A - \gamma_{A,1,3}\mathbf{z}_{A,1}^T\mathbf{y}_A.$$

Для вектора $\mathbf{z}_{A,k}$:

$$\mathbf{z}_{A,k} = \mathbf{x}_{A,k} - \sum_{i=1}^{k-1}\gamma_{A,i,k}\mathbf{z}_{A,i}\,,\; \gamma_{A,1,k} = \frac{\mathbf{z}_{A,1}^T\mathbf{x}_{A,k}}{\mathbf{z}_{A,1}^T\mathbf{z}_{A,1}}\,,$$

$$\gamma_{A,2,k} = \frac{\mathbf{z}_{A,2}^T\mathbf{x}_{A,k}}{\mathbf{z}_{A,2}^T\mathbf{z}_{A,2}} = \frac{\mathbf{x}_{A,k}^T\mathbf{x}_{A,2} - \gamma_{A,1,2}\mathbf{x}_{A,k}^T\mathbf{z}_{A,1}}{\mathbf{z}_{A,2}^T\mathbf{z}_{A,2}}\,,\;\ldots$$

$$\gamma_{A,k-1,k} = \frac{\mathbf{z}_{A,k-1}^T\mathbf{x}_{A,k}}{\mathbf{z}_{A,k-1}^T\mathbf{z}_{A,k-1}} = \frac{\mathbf{x}_{A,k}^T\mathbf{x}_{A,k-1} - \sum\limits_{i=1}^{k-2}\gamma_{A,i,k-1}\mathbf{x}_{A,k}^T\mathbf{z}_{A,i}}{\mathbf{z}_{A,k-1}^T\mathbf{z}_{A,k-1}}\,.$$

Как видим при вычислении коэффициентов $\gamma_{A,i,k}$, $i=\overline{1,k-1}$ можно использовать величины $\mathbf{z}_{A,i}^T\mathbf{z}_{A,i}$, $i=\overline{1,k-1}$, $\gamma_{A,i,j}$, $i=\overline{1,k-2}$, $j=\overline{1,k-1}$, полученные при расчете векторов $\mathbf{z}_{A,i}$, $i=\overline{1,k-1}$, а также величины $\mathbf{x}_{A,k}^T\mathbf{z}_{A,i}$, $i=\overline{1,k-2}$.

При вычислении

$$\mathbf{z}_{A,k}^T\mathbf{z}_{A,k}=\mathbf{x}_{A,k}^T\mathbf{x}_{A,k}+\sum_{i=1}^{k-1}\gamma_{A,i,k}^2\mathbf{z}_{A,i}^T\mathbf{z}_{A,i}-2\sum_{i=1}^{k-1}\gamma_{A,i,k}\mathbf{x}_{A,k}^T\mathbf{z}_{A,i}$$

используются числители $\mathbf{x}_{A,k}^T\mathbf{z}_{A,i}$, $i=\overline{1,k-1}$ формул расчёта коэффициентов $\gamma_{A,i,k}$, $i=\overline{1,k-1}$, а так же $\mathbf{z}_{A,i}^T\mathbf{z}_{A,i}$, $i=\overline{1,k-1}$.

В общем случае

$$\mathbf{z}_{A,k}^T\mathbf{y}_A=\mathbf{x}_{A,k}^T\mathbf{y}_A-\sum_{i=1}^{k-1}\gamma_{A,i,k}\mathbf{z}_{A,i}^T\mathbf{y}_A$$

Если в памяти хранить $\gamma_{A,i,j}$, $i=\overline{1,m-1}$, $j=\overline{1,m}$, $\mathbf{z}_{A,i}^T\mathbf{z}_{A,i}$, $i=\overline{1,m}$, $\mathbf{z}_{A,i}^T\mathbf{y}_A$, $i=\overline{1,m}$, $\mathbf{x}_{A,k}^T\mathbf{z}_{A,i}$, $i=\overline{1,k-1}$, и те же самые величины на выборке B, то $OpNum = 1/3\cdot m^3+4m^2-7/3\cdot m$.

Если же в памяти хранить те же величины, что и в первом случае, кроме $\mathbf{x}_{A,k}^T\mathbf{z}_{A,i}$, $\mathbf{x}_{B,k}^T\mathbf{z}_{B,i}$ $i=\overline{1,k-1}$, то $OpNum = 2/3\cdot m^3 + 5m^2 - 5/3\cdot m$.

Метод перехода в исходное пространство. Пусть в ортогонализированном пространстве была найдена модель: $\hat{\mathbf{y}}_z=\sum_{i=1}^{m}\hat{\varphi}_{A,i}\mathbf{z}_{A,i}$. Модель в исходных переменных имеет вид: $\hat{\mathbf{y}}_x=\sum_{i=1}^{m}\hat{\theta}_{A,i}\mathbf{x}_{A,i}$. Необходимо получить коэффициенты $\phi_{A,i,j}$ при $\mathbf{x}_{A,i}$ в $\mathbf{z}_{A,j}$, $i,j=\overline{1,m}$, и тогда $\hat{\theta}_{A,i}$, рассчитать по формуле:

$$\hat{\theta}_{A,i}=\sum_{j=1}^{m}\hat{\varphi}_{A,j}\phi_{A,i,j}\ ,i=\overline{1,m} \qquad (17)$$

Опишем метод пошагово для каждого $\mathbf{z}_{A,i}$, $i=\overline{1,m}$.

1) $\mathbf{z}_{A,1}=\mathbf{x}_{A,1}$;

2) $\mathbf{z}_{A,2}=\mathbf{x}_{A,2}-\gamma_{A,1,2}\mathbf{z}_{A,1}$, $\phi_{A,1,2}=-\gamma_{A,1,2}$;

3) $\mathbf{z}_{A,3}=\mathbf{x}_{A,3}-\gamma_{A,2,3}\cdot\mathbf{z}_{A,2}-\gamma_{A,1,3}\mathbf{z}_{A,1}=$
$=(\gamma_{A,2,3}\gamma_{A,1,2}-\gamma_{A,1,3})\mathbf{x}_{A,1}-\gamma_{A,1,3}\mathbf{x}_{A,2}+\mathbf{x}_{A,3}$

$\phi_{A,1,3}=-\gamma_{A,2,3}\phi_{A,1,2}-\gamma_{A,1,3}\phi_{A,2,2}$, $\phi_{A,2,3}=-\gamma_{A,2,3}$;

Для вектора $\mathbf{z}_{A,k}$:

$$\phi_{A,1,k}=-(\gamma_{A,1,k}+\sum_{i=2}^{k-1}\gamma_{A,i,k}\phi_{A,1,i}),$$

$$\phi_{A,2,k}=-(\gamma_{A,2,k}+\sum_{i=3}^{k-1}\gamma_{A,i,k}\phi_{A,1,i}),\ldots$$

$$\phi_{A,k-1,k}=-\gamma_{A,k-1,k}.$$

Как видим, для вычисления коэффициентов $\phi_{A,i,k}$, $i=\overline{1,k-1}$ используются величины $\phi_{A,i,k-1}$, $i=\overline{1,k-1}$, полученные при расчёте $\mathbf{z}_{A,k-1}$. При расчёте формулы (17) $\phi_{A,i,i}=1$, $i=\overline{1,m}$. В данном алгоритме в памяти необходимо хранить коэффициенты $\phi_{A,i,j}$ $i=\overline{1,m-1}$, $j=\overline{1,m}$, $OpNum = m^3/6-0.5m^2+m/3$.

Таблица 1. Сравнение вычислительной сложности алгоритмов COMBI* и МАО

Стадии	Составляющие стадий	Количество операций *OpNum*	
		МАО	COMBI*
1. Подготовительные вычисления	Центрирование	$m + 1$	0
	Расчёт матриц $\mathbf{W}_A^T\mathbf{W}_A$, $\mathbf{W}_B^T\mathbf{W}_B$	$n_W(0.5m^2 + 1.5m + 1)$	$n_W(0.5m^2 + 2.5m + 3)$
	Ортогонализация	$1/3\cdot m^3+4m^2-7/3\cdot m$	0
2. Построение m моделей с количеством аргументов от 1 до m		$8m + 3$	$\frac{4}{3}m^3+\frac{25}{3}m^2+22m+\frac{45}{6}$
3. Представление модели в исходных переменных	Переход в исходное пространство переменных	$\frac{1}{6}m^3-\frac{1}{2}m^2+\frac{1}{3}m$	0
	Расчёт свободного члена модели	m	0
Общая вычислительная сложность		$n_W(0.5m^2 + 1.5m + 1) + \frac{5}{6}m^3+\frac{7}{2}m^2+8m+4$	$n_W(0.5m + 2.5m + 3) + \frac{4}{3}m^3+\frac{25}{3}m^2+22m+\frac{45}{6}$
Сложность построения одной модели на $(s+1)$-м этапе		8	$4s^2 + 15s + 16$

Наиболее быстрым алгоритмом МГУА, позволяющим получить точные решения для m моделей, содержащих от 1 до m аргументов, является модифицированная версия комбинаторного алгоритма МГУА COMBI* [5]. Алгоритм COMBI* использует метод вложенных структур для формирования структуры модели и метод окаймления матрицы нормальных уравнений (рекуррент-ный способ получения оценок МНК) – для оценивания параметров Θ. Этот алгоритма также как и МАО для расчёта моделей использует элементы матриц $\mathbf{W}_A^T\mathbf{W}_A$, $\mathbf{W}_B^T\mathbf{W}_B$.

Вычислительная сложность каждой стадии МАО и COMBI*, а также общее количество операций алгоритмов приведены в таблице 1. Предполагается, что при расчетах МАО в памяти хранятся величины $\mathbf{x}_{A,k}^T\mathbf{z}_{A,i}$, $\mathbf{x}_{B,k}^T\mathbf{z}_{B,i}$ $i=\overline{1,k-1}$ и свобода выбора МАО F = 1. При расчёте матриц $\mathbf{W}_A^T\mathbf{W}_A$, $\mathbf{W}_B^T\mathbf{W}_B$ алгоритма COMBI* в них добавлен столбец единиц, отвечающий за свободный член. На второй стадии в COMBI* для каждой из m моделей рассчитывается свободный член, в отличие от МАО, где он рассчитан только для полной модели.

8. Выводы

Теоретически доказано, что при ортогональной системе вектор-столбцов матрицы входов ОРИА позволяет получить точное решение идентичное многоэтапному алгортму.

Предложен многоэтапный алгоритм с ортогонализацией входных переменных.

Для МАО разработан рекуррентный метод оценивания параметров и расчёта критерия селекции моделей. Сложность расчёта одной модели метода является постоянной, в отличие от квадратичной в известной модификации комбинаторного алгоритма COMBI*. Это позволяет решать задачи моделирования количество переменных, в которых на два порядка больше.

Вычислительная сложность МАО при решении системы линейных уравнений размерности $m \times m$ меньше, по сравнению с алгоритмом COMBI* при идентичности их решений, несмотря на то, что в МАО дополнительно осуществляется ортогонализа-ция.

Список литературы

1. Павлов А.В. Обобщённый релаксационный итерационный алгоритм МГУА // Індуктивне моделювання складних систем. Збірник наук. праць. – К.: МННЦІТС, 2011. – С. 95-108.
2. Павлов А.В. Методика экспериментальных исследований сходимости итерационных алгоритмов метода группового учёта аргументов / А.В. Павлов, В.А.Павлов // Вісник НТУУ „КПІ”. Інформатика, управління та обчислювальна техніка: Зб. наук. пр. – К.: «Век+», – 2011. – № 54. – С. 3-8.
3. Ивахненко А. Г. Помехоустойчивость моделирования / Ивахненко А. Г., Степашко В. С. – К.: Наукова думка, 1985. – 214 с.
4. Хорн Р. Матричный анализ / Хорн Р., Джонсон Ч. – М.: Мир, 1989. – 655 с.
5. Степашко В.С. О последовательном оценивании параметров регрессионных моделей / Степашко В.С., Ефименко С.Н. // Кибернетика и системный анализ. – 2005. – № 4. – С. 184–187.