

УДК 681.5(075.8)

В.В. Белов, А.А. Рыбаков

ВЫЧИСЛЕНИЕ ЗНАЧЕНИЙ ЛИНЕЙНОЙ МОДЕЛИ БЕЗ ЕЁ ПАРАМЕТРИЧЕСКОЙ ИДЕНТИФИКАЦИИ

Предлагается способ вычисления значений линейной модели, в частности MLR (Multiple Linear Regression), без оценивания её параметров, эквивалентной по значениям описанию, параметры которого оцениваются методом наименьших квадратов. Предлагаемый способ может использоваться в алгоритмах поиска наилучшего в некотором смысле линейного описания процесса, представленного дискретными значениями, например временным рядом.

Ключевые слова: линейная модель, проекционная матрица, рекуррентные вычисления, поиск регрессоров

Предварительные замечания

Результаты, приведённые в данной статье, получены в процессе выполнения работ по заказам Министерства труда и социального развития России, а также Федеральной службы государственной статистики. Главные задачи работ состояли в получении прогнозных значений для заданных групп показателей производственного травматизма в РФ и занятости населения в экономике страны. Необходимые статистические данные были предоставлены Федеральной службой государственной статистики. Обусловливалось обязательное построение варианта модели в виде линейной множественной регрессии (MLR-описания) с регрессорами, входящими в сценарные условия прогнозирования. Допускалось использование альтернативных методов моделирования для дополнительной оценки надежности полученного прогноза. Поскольку крайне желательно находить MLR-описание с наименьшим количеством параметров (это, в частности, способствует минимизации дисперсии коэффициентов модели), в практике реальных приложений регрессионного анализа часто приходится решать задачу поиска «наилучшей» в некотором смысле регрессии. В процессе решения этой задачи происходит последовательное добавление регрессоров в модель в разных сочетаниях до достижения требуемой погрешности аппроксимации.

Длительные упражнения в решении задачи последовательного синтеза наилучшей линейной модели и оценки качества частичных описаний привели к возникновению идеи вычисления векторов значений MLR-описания без предварительного оценивания его параметров, т.е. без осуществления операции параметрической идентификации. Такой приём позволяет ускорить процесс поиска: новый регрессор добавляется без пересчёта обратной матрицы, связанной с решением системы нормальных уравнений Гаус-

са. Кроме того, происходит, хотя и несущественное, повышение точности вычислений. Платой за эти преимущества является ухудшение пространственных характеристик программного приложения – увеличиваются затраты на память.

История вопроса

Задача расширения линейной модели, параметры которой оцениваются методом наименьших квадратов, впервые рассмотрена в [1]. Однако, В.Г. Кохран ограничился рассмотрением случая добавления в модель одной переменной. На случай нескольких переменных подход В.Г. Кохрана распространил М.Г. Квинауил [2]. Указанные результаты описаны в [3] и доступны на русском языке в переводе [4]. Общим началом подходов В.Г. Кохрана и М.Г. Квинауила является то, что они предлагают схему добавления новых регрессоров, *меняющую* коэффициенты предшествующих описаний: добавление нового слагаемого ($a_p \cdot x_p$) в MLR-описание сопровождается пересчетом всех коэффициентов a_1, a_2, \dots, a_{p-1} предыдущего варианта модели.

Последовательное расширение описания, предлагаемое в настоящей статье, отличается тем, что параметры модели не оцениваются вовсе: 1) коэффициент нового добавляемого в модель регрессора не вычисляется; 2) значения коэффициентов «старых» регрессоров не пересчитываются. Предлагаемый способ целесообразен в тех случаях, когда конкретика значений коэффициентов не важна. Это, прежде всего, как уже указывалось, – задачи поиска наилучших линейных моделей с минимальным количеством параметров.

Классическое решение задачи последовательного введения регрессоров

При введении дополнительных регрессоров по Кохрану и Квинауилу предполагается, что подобрана модель регрессии

$$\varepsilon(\mathbf{Y}) = \mathbf{X}\boldsymbol{\beta}, \quad \mathcal{D}(\mathbf{Y}) = \sigma^2 \mathbf{I}_n,$$

где $\mathbf{Y} = (y_1, y_2, \dots, y_n)^T$ – временной ряд, т.е. вектор последовательных равноотстоящих по времени значений некоторого процесса (результативного признака); \mathbf{X} – матрица регрессоров (аргументов или факторных признаков) размером $n \times p$ ранга p ; $\boldsymbol{\beta}$ – вектор коэффициентов регрессии, состоящий из p элементов; \mathbf{I}_n – единичная матрица порядка n ; σ – дисперсия значений временного ряда \mathbf{Y} . Оценка $\hat{\boldsymbol{\beta}}$ вектора $\boldsymbol{\beta}$ находится методом наименьших квадратов: $\hat{\boldsymbol{\beta}} = (\mathbf{X}^T \mathbf{X})^{-1} \mathbf{X}^T \mathbf{Y}$.

Рассматривается задача добавления новых регрессоров таким образом, чтобы новая модель имела вид

$$G: \varepsilon(\mathbf{Y}) = \mathbf{X}\boldsymbol{\beta}_G + \mathbf{Z}\boldsymbol{\gamma}_G = (\mathbf{X} \quad \mathbf{Z}) \begin{pmatrix} \boldsymbol{\beta}_G \\ \boldsymbol{\gamma}_G \end{pmatrix} = \mathbf{W}\boldsymbol{\delta}_G,$$

где буква G идентифицирует новую модель; \mathbf{Z} – матрица дополнительных регрессоров размером $n \times t$ ранга t ; $\boldsymbol{\beta}_G$ – вектор новых коэффициентов при «старых» регрессорах, объединённых в матрицу \mathbf{X} ; $\boldsymbol{\gamma}_G$ – вектор коэффициентов при новых регрессорах, объединённых в матрицу \mathbf{Z} , состоящий из t элементов; $\mathbf{W} = (\mathbf{X} \quad \mathbf{Z})$ – матрица объединённых регрессоров, т.е. матрица всех регрессоров модели G , имеющая размер $n \times (p+t)$ и ранг $p+t$; $\boldsymbol{\delta}_G = (\boldsymbol{\beta}_G \quad \boldsymbol{\gamma}_G)^T$ – вектор всех коэффициентов модели G , состоящий из $p+t$ элементов.

Указанные предположения означают, что имеет место стандартная ситуация: длина n временного ряда \mathbf{Y} больше числа параметров модели, т.е. $n > p+t$ и все столбцы матрицы \mathbf{W} (все регрессоры) линейно независимы.

Естественно, вычислить оценку $\hat{\boldsymbol{\delta}}_G$ вектора $\boldsymbol{\delta}_G$ и её дисперсионную матрицу $\mathcal{D}(\hat{\boldsymbol{\delta}})$ можно следующим образом:

$$\hat{\boldsymbol{\delta}}_G = (\mathbf{W}^T \mathbf{W})^{-1} \mathbf{W}^T \mathbf{Y} \quad \mathcal{D}(\hat{\boldsymbol{\delta}}) = \sigma^2 (\mathbf{W}^T \mathbf{W})^{-1}.$$

Однако можно сократить объём вычислений, используя результат обращения матрицы $(\mathbf{X}^T \mathbf{X})$, полученный ранее при вычислении оценки $\hat{\boldsymbol{\beta}}$ вектора коэффициентов $\boldsymbol{\beta}$.

Суть метода Кохрана-Квинаула выражается следующей теоремой [4, с. 69].

□ **Теорема Себера**

Пусть $\mathbf{R} = \mathbf{I}_n - \mathbf{X}(\mathbf{X}^T \mathbf{X})^{-1} \mathbf{X}^T$,

$$\mathbf{R}_G = \mathbf{I}_n - \mathbf{W}(\mathbf{W}^T \mathbf{W})^{-1} \mathbf{W}^T, \quad \mathbf{L} = (\mathbf{X}^T \mathbf{X})^{-1} \mathbf{X}^T \mathbf{Z},$$

$$\mathbf{M} = [\mathbf{Z}^T \mathbf{R} \mathbf{Z}]^{-1} \text{ и } \hat{\boldsymbol{\delta}}_G = \begin{pmatrix} \hat{\boldsymbol{\beta}}_G \\ \hat{\boldsymbol{\gamma}}_G \end{pmatrix}.$$

Тогда

$$(i) \hat{\boldsymbol{\beta}}_G = (\mathbf{X}^T \mathbf{X})^{-1} \mathbf{X}^T (\mathbf{Y} - \mathbf{Z}\hat{\boldsymbol{\gamma}}_G) = \hat{\boldsymbol{\beta}} - \mathbf{L}\hat{\boldsymbol{\gamma}}_G.$$

$$(ii) \hat{\boldsymbol{\gamma}}_G = (\mathbf{Z}^T \mathbf{R} \mathbf{Z})^{-1} \mathbf{Z}^T \mathbf{R} \mathbf{Y}.$$

$$(iii) \mathbf{Y}^T \mathbf{R}_G \mathbf{Y} = (\mathbf{Y} - \mathbf{Z}\hat{\boldsymbol{\gamma}}_G)^T \mathbf{R} (\mathbf{Y} - \mathbf{Z}\hat{\boldsymbol{\gamma}}_G).$$

$$(iv) \mathbf{Y}^T \mathbf{R}_G \mathbf{Y} = \mathbf{Y}^T \mathbf{R} \mathbf{Y} - \hat{\boldsymbol{\gamma}}_G^T \mathbf{Z}^T \mathbf{R} \mathbf{Y}.$$

$$(v) \mathcal{D}(\hat{\boldsymbol{\delta}}_G) = \sigma^2 \begin{pmatrix} (\mathbf{X}^T \mathbf{X})^{-1} + \mathbf{L} \mathbf{M} \mathbf{L}^T & -\mathbf{L} \mathbf{M} \\ -\mathbf{M} \mathbf{L}^T & \mathbf{M} \end{pmatrix} \blacksquare$$

Замечания:

1) пункт (ii) теоремы определяет алгоритм вычисления коэффициентов при новых переменных, добавляемых в модель; если вместо нескольких переменных добавить одну переменную, то матрица \mathbf{Z} трансформируется в вектор \mathbf{z} , одновременно матрицы $\mathbf{Z}^T \mathbf{R} \mathbf{Z}$ и $\mathbf{Z}^T \mathbf{R} \mathbf{Y}$ трансформируются в скаляры, формула коэффициента при добавляемой переменной принимает

$$\text{вид } \hat{\boldsymbol{\gamma}}_G = \frac{\mathbf{z}^T \mathbf{R} \mathbf{Y}}{\mathbf{z}^T \mathbf{R} \mathbf{z}}, \text{ и это уже не вектор, а скаляр;}$$

2) пункт (i) определяет алгоритм коррекции вектора $\hat{\boldsymbol{\beta}}$ коэффициентов при переменных, уже входивших в модель; заметим, что он предполагает использование вектора $\hat{\boldsymbol{\gamma}}_G$ коэффициентов при добавляемых переменных, определяемого в следующем пункте, и, главное, – формула для матрицы \mathbf{L} предполагает вычисление обратной матрицы $(\mathbf{X}^T \mathbf{X})^{-1}$;

3) семантически матрица $\mathbf{L} = (\mathbf{X}^T \mathbf{X})^{-1} \mathbf{X}^T \mathbf{Z}$ представляет собой совокупность векторов модельных значений – j -й столбец этой матрицы представляет собой результат объяснения j -й добавляемой переменной всеми предыдущими переменными;

4) пункты (iii) и (iv) определяют две эквивалентные по значениям формулы вычисления остаточной суммы квадратов $\mathbf{Y}^T \mathbf{R}_G \mathbf{Y}$ для итоговой модели через матрицы \mathbf{R} и \mathbf{Z} , формируемые старыми и новыми переменными.

Предлагаемое решение задачи последовательного введения регрессоров без параметрической идентификации

□ **Теорема 1**

Пусть

$\mathbf{u} = (y_1, y_2, \dots, y_n)^T$ – вектор вещественных чисел, условно называемый результативным признаком;

$\mathbf{x}_1 = (x_{1,1}, x_{2,1}, \dots, x_{n,1})^T$, $\mathbf{x}_2 = (x_{1,2}, x_{2,2}, \dots, x_{n,2})^T$,
 $\mathbf{x}_p = (x_{1,p}, x_{2,p}, \dots, x_{n,p})^T$ – векторы вещественных чисел, условно называемые объясняющими переменными или факторными признаками;
 $\mathbf{J}_n = [1]_{n \times 1}$ – вектор единиц, состоящий из n элементов; \mathbf{I}_n – единичная матрица размером $n \times n$,
 $\mathbf{X} = [\mathbf{x}_0 \ \mathbf{x}_1 \ \mathbf{x}_2 \ \dots \ \mathbf{x}_p]$ – матрица объясняющих переменных, причём $\mathbf{x}_0 = \mathbf{J}_n$ и $\text{rank}(\mathbf{X}) = p + 1$;
 $\hat{\mathbf{y}} = (\hat{y}_1, \hat{y}_2, \dots, \hat{y}_n)^T$ – вектор значений линейной модели зависимости результирующего признака от объясняющих переменных, т.е. $\hat{\mathbf{y}} = \mathbf{X} \cdot \mathbf{a}$, где $\mathbf{a} = (a_0, a_1, \dots, a_p)$ – вектор параметров (коэффициентов) модели, вычисляемых методом наименьших квадратов: $\mathbf{a} = (\mathbf{X}^T \cdot \mathbf{X})^{-1} \cdot \mathbf{X}^T \cdot \mathbf{y}$;
 $\mathbf{P} = \mathbf{X} \cdot (\mathbf{X}^T \cdot \mathbf{X}) \cdot \mathbf{X}^T$ – проекционная матрица, порождающая значения модели $\hat{\mathbf{y}} = \mathbf{P} \cdot \mathbf{y}$.

Тогда

- 1) проекционная матрица $\mathbf{P} = \mathbf{X} \cdot (\mathbf{X}^T \cdot \mathbf{X})^{-1} \cdot \mathbf{X}^T$, порождающая значения $\hat{\mathbf{y}}$ линейной модели, построенной с использованием факторных признаков x_1, x_2, \dots, x_p , эквивалентной по значениям описанию, параметры которого оцениваются методом наименьших квадратов, может быть получена по рекуррентной схеме: $\mathbf{P} = \mathbf{P}_p$, где \mathbf{P}_p – финальное значение в последовательности рекуррентных вычислений

$$\mathbf{P}_j = \mathbf{P}_{j-1} + \frac{\mathbf{R}_{j-1} \cdot \mathbf{x}_j \cdot \mathbf{x}_j^T \cdot \mathbf{R}_{j-1}}{\mathbf{x}_j^T \cdot \mathbf{R}_{j-1} \cdot \mathbf{x}_j}, \quad j = \overline{1, p};$$

$$\mathbf{R}_{j-1} = \mathbf{I}_n - \mathbf{P}_{j-1};$$

$$\mathbf{P}_0 = \frac{\mathbf{J}_n \cdot \mathbf{J}_n^T}{\mathbf{J}_n^T \cdot \mathbf{J}_n} = \frac{\mathbf{J}_n \cdot \mathbf{J}_n^T}{n};$$

- 2) остаточная сумма квадратов $RSS = \mathbf{y}^T \cdot (\mathbf{I}_n - \mathbf{P}_p) \cdot \mathbf{y}$.

Доказательство.

Запишем формулу (3.37) [4, с. 72], описывающую значения модели с ортогональной структурой, для частного случая, когда матрица \mathbf{Z} вырождается в вектор \mathbf{x}_j в новых обозначениях:

$$\hat{\mathbf{y}}^{[j]} = \mathbf{X}_{j-1} \cdot \mathbf{a}_{j-1}^{[j-1]} + \mathbf{R}_{j-1} \cdot \mathbf{x}_j \cdot a_j^{[j]}. \quad (1)$$

В соответствии с пунктом (ii) теоремы Себера (см. выше) для рассматриваемого частного случая и используемых обозначений имеем:

$$a_j^{[j]} = \frac{\mathbf{x}_j^T \cdot \mathbf{R}_{j-1} \cdot \mathbf{y}}{\mathbf{x}_j^T \cdot \mathbf{R}_{j-1} \cdot \mathbf{x}_j}. \quad \text{Кроме того, учтём, что}$$

$\mathbf{X}_{j-1} \cdot \mathbf{a}_{j-1}^{[j-1]} = \hat{\mathbf{y}}^{[j-1]}$. Подставив указанные выражения в (1), получим

$$\hat{\mathbf{y}}^{[j]} = \hat{\mathbf{y}}^{[j-1]} + \mathbf{R}_{j-1} \cdot \mathbf{x}_j \cdot \frac{\mathbf{x}_j^T \cdot \mathbf{R}_{j-1} \cdot \mathbf{y}}{\mathbf{x}_j^T \cdot \mathbf{R}_{j-1} \cdot \mathbf{x}_j}.$$

Далее с учётом того, что $\hat{\mathbf{y}}^{[j]} = \mathbf{P}_j \cdot \mathbf{y}$ и

$$\hat{\mathbf{y}}^{[j-1]} = \mathbf{P}_{j-1} \cdot \mathbf{y}, \text{ имеем}$$

$$\mathbf{P}_j \cdot \mathbf{y} = \mathbf{P}_{j-1} \cdot \mathbf{y} + \frac{\mathbf{R}_{j-1} \cdot \mathbf{x}_j \cdot \mathbf{x}_j^T \cdot \mathbf{R}_{j-1}}{\mathbf{x}_j^T \cdot \mathbf{R}_{j-1} \cdot \mathbf{x}_j} \cdot \mathbf{y}.$$

Откуда и следует, что матрица \mathbf{P}_j может быть вычислена рекуррентно:

$$\mathbf{P}_j = \mathbf{P}_{j-1} + \frac{\mathbf{R}_{j-1} \cdot \mathbf{x}_j \cdot \mathbf{x}_j^T \cdot \mathbf{R}_{j-1}}{\mathbf{x}_j^T \cdot \mathbf{R}_{j-1} \cdot \mathbf{x}_j}.$$

Терминальная ветвь рекурсии для случая $j=1$ определяется значением \mathbf{P}_0 . По определению $\mathbf{P}_0 = \mathbf{x}_0 \cdot (\mathbf{x}_0^T \cdot \mathbf{x}_0)^{-1} \cdot \mathbf{x}_0^T$. Поскольку $\mathbf{x}_0 = \mathbf{J}_n$, а произведение $\mathbf{J}_n^T \cdot \mathbf{J}_n$ скалярно и равно n , имеем

$$\mathbf{P}_0 = \frac{\mathbf{J}_n \cdot \mathbf{J}_n^T}{n}.$$

Остаточная сумма квадратов RSS по определению равна $\mathbf{y}^T \cdot (\mathbf{I}_n - \mathbf{P}) \cdot \mathbf{y}$. Равенство $\mathbf{P} = \mathbf{P}_p$ определяет возможность вычисления остаточной суммы по формуле $RSS = \mathbf{y}^T \cdot (\mathbf{I}_n - \mathbf{P}_p) \cdot \mathbf{y}$. ■

Алгоритм вычислений

В соответствии с указанной теоремой алгоритм вычисления вектора значений $\hat{\mathbf{y}}$ линейной модели без операции её параметрической идентификации имеет вид:

- 3) вспомогательные величины:

- (1) $\mathbf{J}_n = [1]_{n \times 1}$ – вектор единиц, состоящий из n элементов;

- (2) $\mathbf{I}_n = \text{identity}(n)$ – единичная матрица размером $n \times n$;

- 4) начальное значение:

$$\mathbf{P}_0 = \frac{\mathbf{J}_n \cdot \mathbf{J}_n^T}{\mathbf{J}_n^T \cdot \mathbf{J}_n} = \frac{\mathbf{J}_n \cdot \mathbf{J}_n^T}{n}$$

– квадратная матрица размером $n \times n$, все элементы которой равны $\frac{1}{n}$, – начальное значение матрицы,

порождающей линейную модель, построенную по первому столбцу $\mathbf{x}_0 = \mathbf{J}_n$ матрицы объясняющих переменных \mathbf{X}_p ;

- 5) рекуррентные вычисления для $j = \overline{1, p}$:

- (1) $\mathbf{R}_{j-1} = \mathbf{I}_n - \mathbf{P}_{j-1}$ – вспомогательная матрица для упрощения записи формул;
 (2) если $\mathbf{R}_{j-1} \cdot \mathbf{x}_j \neq [0]_{n \times 1}$, то вычисляется

$$\mathbf{P}_j = \mathbf{P}_{j-1} + \frac{\mathbf{R}_{j-1} \cdot \mathbf{x}_j \cdot \mathbf{x}_j^T \cdot \mathbf{R}_{j-1}}{\mathbf{x}_j^T \cdot \mathbf{R}_{j-1} \cdot \mathbf{x}_j} - \text{очередное}$$

значение порождающей матрицы \mathbf{P} ;

- (3) иначе, если $\mathbf{R}_{j-1} \cdot \mathbf{x}_j = [0]_{n \times 1}$, где $[0]_{n \times 1}$ – вектор, состоящий из одних нулей, то переменная x_j исключается из списка потенциальных аргументов модели и осуществляется переход к следующему значению j ;
- б) вычисляется вектор $\hat{\mathbf{y}}(\mathbf{X}) = \mathbf{P}_p \cdot \mathbf{y}$ значений модели, построенной по матрице объясняющих переменных $\mathbf{X} = [x_0 \ x_1 \ x_2 \ \dots \ x_p]$.

Замечания:

1) вычисление значений модели по указанному алгоритму не требует выполнения операции обращения матрицы;

2) на каждом шаге рекуррентных вычислений могут быть получены значения частных описаний, построенных по части объясняющих переменных: $\mathbf{y}^{[j]} = \hat{\mathbf{y}}(x_0, x_1, \dots, x_j) = \mathbf{P}_j \cdot \mathbf{y}$;

3) условие $\mathbf{R}_{j-1} \cdot \mathbf{x}_j \neq [0]_{n \times 1}$ исключает переменные, векторы значений которых коллинеарны с векторами, уже включёнными в модель; произведение $\mathbf{R}_{j-1} \cdot \mathbf{x}_j$ представляет собой вектор ошибок объяснения вектора \mathbf{x}_j векторами $\mathbf{x}_0, \mathbf{x}_1, \mathbf{x}_2, \dots, \mathbf{x}_{j-1}$.

Вычислительный эксперимент

Проверка правильности алгоритма была осуществлена путём генерации нескольких векторов случайных чисел, равномерно распределённых от нуля до десяти. Один из векторов интерпретировался как значения результативного признака $\mathbf{y} = (y_1, y_2, \dots, y_n)^T$, а остальные – как значения факторных признаков $\mathbf{x}_1 = (x_{1,1}, x_{2,1}, \dots, x_{n,1})^T$, $\mathbf{x}_2 = (x_{1,2}, x_{2,2}, \dots, x_{n,2})^T$, ..., $\mathbf{x}_p = (x_{1,p}, x_{2,p}, \dots, x_{n,p})^T$. Матрица, порождающая значения модели, находилась двумя способами – рекуррентно \mathbf{P}_p и путём прямых вычислений $\mathbf{P} = \mathbf{X} \cdot (\mathbf{X}^T \cdot \mathbf{X})^{-1} \cdot \mathbf{X}^T$. Затем находились векторы значений модели $\hat{\mathbf{y}}^{[p]} = \mathbf{P}_p \cdot \mathbf{y}$ и $\hat{\mathbf{y}} = \mathbf{P}_p \cdot \mathbf{y}$.

Длины векторов, участвовавших в вычислениях, выbralись равными: 10; 30; 100 и 500. Ко-

личество факторных признаков: 4; 7 и 15.

Расчёты выполнялись с использованием восьми байтов для представления вещественных чисел. Результаты показали, что во всех случаях абсолютная величина разности элементов матриц \mathbf{P} и \mathbf{P}_p , а также векторов $\hat{\mathbf{y}}^{[p]}$ и $\hat{\mathbf{y}}$ мала:

$$\max(\overline{|\mathbf{P} - \mathbf{P}_p|}) \approx 10^{-15}, \quad \max(\overline{|\hat{\mathbf{y}}^{[p]} - \hat{\mathbf{y}}|}) \approx 10^{-14},$$

где стрелка над выражением символизирует операцию векторизации, в данном случае – формирования вектора абсолютных значений разностей соответствующих элементов матриц и векторов. Таким образом, различие результатов вычисления проекционных матриц и значений модели классическим и предлагаемым способами имеет порядок погрешности внутримашинного представления вещественных данных.

Выводы

Научным результатом, представленным в настоящей статье, является рекуррентный метод вычисления проекционной матрицы \mathbf{P} , порождающей вектор значений линейной модели без оценки её параметров.

Прикладное значение указанного метода состоит в том, что он положен в основу предложенного алгоритма вычисления значений модели. Этот алгоритм предназначен, в первую очередь, для использования в процедурах поиска наилучшего состава регрессоров MLR-модели в заданном множестве потенциальных аргументов x_1, x_2, \dots, x_p , ориентированной на решение задачи прогнозирования. На первом этапе этих процедур отыскиваются регрессии, лучшие по RSS – остаточной сумме квадратов с одним, двумя, тремя и так далее p регрессорами. При этом не требуется параметрическая идентификация модели. Затем среди найденных моделей находится наиболее устойчивое описание, прогнозирующие способности которой наиболее существенны.

Использование предложенного алгоритма на первом этапе процедуры поиска наилучшей модели в некоторых случаях позволяет сократить время и повысить точность вычислений. Указанные эффекты достигаются за счёт увеличения используемых объёмов оперативной памяти.

Можно утверждать, что изложенные в настоящей статье материалы позволяют исследователям глубже проникнуть в тонкости процедур оценки параметров линейной модели и расширяют арсенал средств аналитиков-практиков.

Библиографический список

1. Cochran W.G. The omission or addition an independent variable in multiple linear regression // J. R.

Stat. Soc. Suppl., 1938, Vol. 5, pp.171 – 176..

2. *Quenouille M.H.* An application of least squares to family diet surveys // *Econometrica*, January 1950, Vol. 18, Issue 1, pp. 27 – 44.

3. *Seber G.A.F.* Linear regression analysis. Wiley: New York, London, Sydney and Toronto. 1977.

4. *Себер Дж.* Линейный регрессионный анализ: пер с англ. – М.: Мир, 1980. – 456 с.