Если векторы $\vec{\pi}_{ij}, \vec{\pi}_{ji}$ из (3) удовлетворяют условиям согласования с точностью до слагаемых, длины которых достаточно малы, то имеет место σ-фокусировка.

Полученные результаты справедливы и в случае, когда множество всех состояний исследуемого процесса Π является счетным или континуальным. Для каждого из этих случаев число возможных фрагментов бесконечно. Условия согласования здесь такие же, как и в случае конечного числа состояний. Доказательство сходимости к финальному распределению проводится сначала для процессов Π_n ($n = 1, 2, \dots$), в каждом из которых число возможных фрагментов конечно и неограниченно возрастает с ростом n. Каждый процесс Π_n при $t \uparrow t_0$ фокусирует на распределение $\vec{\pi}_n^*$ и аппроксимирует с заданной точностью процесс Π. С ростом n точность аппроксимации возрастает. Далее устанавливается, что $\vec{\pi}_n^* \to \vec{\pi}^*$ ($n \to \infty$).

Пусть локальные возмущения процесса Π таковы, что векторы распределений, на которые фокусируют фрагменты, являются случайными. Мощность множества всех состояний процесса Π по-прежнему не более чем континуальна. Предполагается, что любой фрагмент при каждом очередном его возмущении фокусирует на один и тот же случайный вектор. Условия согласования теперь состоят в том, чтобы были параллельны векторы математических ожиданий $M\vec{\pi}_{ij}$, $M\vec{\pi}_{ji}$. Остальные предположения об исследуемом процессе остаются прежними (см. условия а), б)). Тогда процесс Π с вероятностью 1 фокусирует на случайный вектор $\vec{\pi}^*$. Равенство (4) теперь имеет вид $\lim\limits_{t \uparrow t_0} M\vec{\pi}_{ij}(t) = \lim\limits_{t \uparrow t_0} M\vec{\pi}_{ji}(t)$.

Литература: 1. *Дикарев В. А.* Точки фокусировки и теоремы о существовании предельных вероятностей. Харьков, 1995. 9 с. Деп. в ГНТБ Украины 28.02.95, № 526 – Ук95. 2. *Dikarev V. A.* Modeling of disintegrating economy by Markov process // II Intern. Conf. "Computing in economics and finance". Geneva, 1996. P. 170-171. 3. *Дикарев В. А.* Фокусировка распределений марковских процессов // Доп. НАН України. 1999. №11. С. 100 -103.

Поступила в редколлегию 07.12.99

Рецензент: д-р физ.-мат. наук Руткас А. Г.

Дикарев Вадим Анатольевич, д-р физ.-мат. наук, профессор кафедры прикладной математики ХТУРЭ. Научные интересы: функциональный анализ, стохастический анализ и его приложения. Адрес: Украина, 61164, Харьков, пр. Ленина 66, кв. 21, тел. (0572) 33-57-03, (0572) 40-94-36.

УДК 658.012

ДВУХСТУПЕНЧАТЫЙ АЛГОРИТМ ОЦЕНИВАНИЯ ПАРАМЕТРОВ, ОСНОВАННЫЙ НА МЕТОДЕ ИНСТРУМЕНТАЛЬНЫХ ПЕРЕМЕННЫХ

ТИМОФЕЕВ В.А.

Рассматривается задача оценивания параметров линейной регрессионной модели в случае коррелированных помех. Проводится анализ используемых при этом алгоритмов, в частности, МНК, двойной МНК. В данной ситуации несмещенная оценка может быть получена с использованием МИП, однако целесообразным представляется разработка такой модификации алгоритма, которая позволяла бы получить оценку с минимальной дисперсией. Предлагается двухступенчатая процедура алгоритма МИП. Приводится вычислительная процедура алгоритма. Доказывается его сходимость. Показывается, что предлагаемый алгоритм имеет меньшую СКО по сравнению с аналогами.

Многие задачи распознавания образов, обработки экспериментальных данных, идентификации и прогнозирования связаны с построением множественной регрессии вида

$$\vec{Y} = X\vec{a} + \vec{\varepsilon},$$

где $\vec{Y} \in R^n$ – вектор выходной переменной; $X \in R^{n \times m}$ – матрица наблюдений; $\vec{a} \in R^n$ – вектор оцениваемых параметров; $\varepsilon \in R^n$ – помеха измерений; n – число наблюдений.

На практике наиболее часто для оценивания параметров $\vec{a}$ используются алгоритмы метода наименьших квадратов (МНК). Если помехи измерений распределены по нормальному закону, МНК-оценки становятся оценками максимального правдоподобия, а также несмещенными, состоятельными, эффективными и достаточными [1]. Наличие коррелированности входных переменных, приводящей к плохой обусловленности информационной матрицы, резко ухудшает свойства МНК-оценок.

Уменьшение среднеквадратичной ошибки (СКО) оценивания может быть достигнуто переходом от МНК к двухступенчатому МНК. Двухступенчатый метод, основанный на МНК, был предложен в [2] для анализа остатков линейного регрессионного уравнения. Задача оценивания параметров с помощью такого метода рассмотрена в [3,4]. Обобщение двухступенчатого метода на многомерный случай регрессионной модели осуществлено в [5]. Кроме того, в этой работе показано, что двухступенчатый МНК (ДМНК), обладая тем же смещением, что и МНК, позволяет получить меньшую по сравнению с МНК среднеквадратичную ошибку оценивания. Эффективность ДМНК существенно возрастает при наличии помех и сильной коррелированности входных переменных, т.е. при плохой обусловленности матрицы наблюдений.

В ряде случаев, в частности, при наличии корреляции полезного сигнала и помехи предпочтительным оказывается метод инструментальных переменных (МИП) [4]. Однако известный МИП имеет ряд недостатков, поэтому представляется целесообразным получение такой его модификации, которая помимо несмещенной оценки давала бы еще и оценку с меньшей СКО по сравнению с МИП.

Для получения модифицированного МИП поступим следующим образом. Разобьём матрицу X и вектор $\vec{a}$ на блоки

$$X = (X_1, X_2)\ ,\ \vec{a}^T = \left(\vec{a}_1^{\,T}, \vec{a}_2^{\,T}\right). \tag{1}$$

Здесь $X_1 \in R^{n\times p}$, $X_2 \in R^{n\times q}$, $\vec{a}_1 \in R^p$, $\vec{a}_2 \in R^q$, $p+q=m$, T – символ транспонирования.

С учетом (1) уравнение регрессии можно записать в виде

$$\vec{Y} = x_1\vec{a}_1 + x_2\vec{a}_2 + \vec{\varepsilon}. \qquad (2)$$

Обозначим $x_2\vec{a}_2 + \vec{\varepsilon}_t = \vec{\chi}_t$ и оценим из полученного уравнения $\vec{a}_1$. В этом случае оценка МИП будет иметь вид

$$\vec{\alpha}_1 = (W_1^T X_1)^{-1} W_1^T Y. \qquad (3)$$

Здесь матрица ИП W разбита также на блоки $W = (W_1, W_2)$, $W_1 \in R^{n\times p}$, $W_2 \in R^{n\times q}$.

Обозначим $\vec{\tilde{Y}} = \vec{Y} - x_1\vec{\alpha}_1$. Тогда оценка МИП вектора $\vec{a}$ запишется следующим образом:

$$\alpha_2 = (W_2{}^T X_2)^{-1} W_2 \vec{\tilde{Y}}$$

или с учетом выражений для $\vec{\alpha}_1$ и $\vec{\tilde{Y}}$

$$\alpha_2 = (W_2{}^T X_2)^{-1} W_2{}^T (\vec{Y} - X_1\vec{\alpha}_1) =$$

$$= (W_2{}^T X_2)^{-1} W_2{}^T P\vec{Y}, \qquad (4)$$

где $P = I - X_1 (W_1{}^T X_1)^{-1} W_2{}^T$; I – единичная матрица размерности $n \times n$.

Таким образом, оценки ДМИП имеют вид

$$\vec{\alpha}_1 = (W_1{}^T X_1)^{-1} W_1{}^T \vec{Y}; \qquad (5)$$

$$\vec{\alpha}_2 = (W_2{}^T X_2)^{-1} W_2{}^T P\vec{Y}. \qquad (6)$$

При изучении свойств ДМИП предположим, что:

1) $\varepsilon_t \sim N(0,\sigma^2 I)$, σ^2 – дисперсия помехи;

2) переменные X, W – неслучайны;

3) $X_2{}^T X_1 \neq 0$, $W_1^T X_2 \neq 0$, $W_2^T X_1 \neq 0$, $W_2^T W_1 \neq 0$.

Определим смещение оценок ДМИП. С учетом (5), (6) и свойств помехи ε получаем

$$\Delta\vec{\alpha}_1 = M\{\vec{\alpha}_1 - \vec{a}_1\} = A\vec{a}_2;$$

$$\Delta\vec{\alpha}_2 = M\{\vec{\alpha}_2 - \vec{a}_2\} = -A_2 A_1 \vec{a}_2,$$

где

$$A_1 = (W_1{}^T X_1)^{-1} W_1{}^T X_2;$$

$$A_2 = (W_2{}^T X_2)^{-1} W_1{}^T X_1.$$

Ковариационная матрица оценок ДМИП определится следующим образом:

$$D\{\vec{\alpha}_1\} = M\left\{(\vec{\alpha}_1 - M\{\vec{\alpha}_1\})(\vec{\alpha}_1 - M\{\vec{\alpha}_1\})^T\right\} =$$

$$= (W_1{}^T X_1)^{-1} W_1{}^T W_1 (X_1{}^T W_1)^{-1}\sigma^2; \qquad (7)$$

$$D\{\vec{\alpha}_2\} = M\left\{(\vec{\alpha}_2 - M\{\vec{\alpha}_2\})(\vec{\alpha}_2 - M\{\vec{\alpha}_2\})^T\right\} =$$

$$= A_2 P A_2^T \sigma^2; \qquad (8)$$

$$\mathrm{cov}\{\vec{\alpha}_1, \vec{\alpha}_2\} = M\left\{(\vec{\alpha}_1 - M\{\vec{\alpha}_1\})(\vec{\alpha}_2 - M\{\vec{\alpha}_2\})\right\} = 0. \qquad (9)$$

Использование формул (7), (8) позволяет получить выражение для СКО оценивания ДМИП:

$$M\left\{\|\vec{\alpha}_1 - \vec{a}_1\|^2\right\} = \mathrm{tr}D\{\vec{\alpha}_1\} + \|\Delta\vec{a}_1\|^2 =$$

$$= \sigma^2 \mathrm{tr}(W_1{}^T X_1)^{-1} W_1{}^T W_1 (X_1{}^T W_1)^{-1} + \|A_1 a_2\|^2; \qquad (10)$$

$$M\left\{\|\vec{\alpha}_2 - \vec{a}_2\|^2\right\} = \mathrm{tr}D\{\vec{\alpha}_2\} + \|\Delta\vec{\alpha}_2\|^2 =$$

$$= \sigma^2 \mathrm{tr}A_2 P A_2^T + \left\|A_2 \tilde{A}_1 a_2\right\|^2. \qquad (11)$$

Здесь trA – след матрицы A. Так как ковариационная матрица оценок МИП имеет вид

$$D_{МИП}\{\vec{\alpha}\} = (W^T X)^{-1} W^T W (X^T W)^{-1}\sigma^2,$$

то разбиение матриц X и W на блоки $X = (X_1, X_2)$, $W = (W_1, W_2)$ позволяет записать эту матрицу следующим образом:

$$D_{МИП}\{\vec{\alpha}\} = \begin{pmatrix} D\{\vec{\alpha}_1\} & \mathrm{cov}\{\vec{\alpha}_1, \vec{\alpha}_2\} \\ \mathrm{cov}\{\vec{\alpha}_1, \vec{\alpha}_2\} & D\{\vec{\alpha}_2\} \end{pmatrix}.$$

Для определения элементов матрицы воспользуемся правилами обращения блочных матриц [6]. Тогда

$$(W^T X)^{-1} = \begin{pmatrix} A & B \\ C & D \end{pmatrix},$$

где

$$A = (W_1 X_1)^{-1} + (W_1^T X_1)^{-1} W_1^T X_2 N^{-1} W_2 X_1 (W_1^T X_1)^{-1};$$

$$B = -(W_1{}^T X_1)^{-1} W_1{}^T X_2 N^{-1};$$

$$C = -N^{-1} W_2{}^T X_1 (W_1{}^T X_1)^{-1};$$

$$D = N^{-1} = (W_2^T X_2 - W_2^T X_1 (W_1^T X_1)^{-1} W_1^T X_2)^{-1} =$$

$$= (W_2^T P X_2)^{-1}.$$

По аналогии с $(W^T X)^{-1}$ определяем элементы матрицы $(X^T W)^{-1}$:

$$(X^T W)^{-1} = \begin{pmatrix} \tilde{A} & \tilde{B} \\ \tilde{C} & \tilde{D} \end{pmatrix},$$

здесь

$$\tilde{A} = (X_1^T W_1) + (X_1^T W_1)^{-1} \tilde{X}_1^T W_2 N^{-1} X_2^T W_1 (X_1^T W_1)^{-1};$$

$$\tilde{B} = -(X_1{}^T W_1)^{-1} X_1{}^T \tilde{W}_2 N^{-1};$$

$$\tilde{C} = -\tilde{N}^{-1} X_2{}^T W_1 (X_1{}^T W_1)^{-1};$$

$$\tilde{D} = \tilde{N}^{-1} = (X_2{}^T W_2 - X_2{}^T W_1 (X_1{}^T W_1)^{-1} X_1 W_2)^{-1} =$$

$$= (X_2^T P^T W_2)^{-1}.$$

Ковариационную матрицу $D_{МИП}\{\vec{\alpha}\}$ можно представить в виде

$$D_{МИП} = \begin{pmatrix} D_{11} & D_{12} \\ D_{21} & D_{22} \end{pmatrix}.$$

Подстановка в формулу для $D_{МИП}\{\vec{\alpha}\}$ выражений для обратных матриц $(W^T X)^{-1}$ и $(X^T W)^{-1}$ и перемножение дают

$$
\begin{aligned}
D_{11} &= (W_1{}^T X_1)^{-1} W_1{}^T W_1 (X_1{}^T W_1)^{-1} + \\
&+ (W_1{}^T X_1)^{-1} W_1{}^T W_1 (X_1{}^T W_1)^{-1} X_1{}^T W_2 \tilde{N}^{-1} X_2{}^T W_1 (X_1{}^T W_1)^{-1} + \\
&+ (W_1{}^T X_1)^{-1} W_1{}^T X_2 N^{-1} W_2{}^T X_1 (W_1{}^T X_1)^{-1} W_1{}^T W_1 (X_1{}^T W_1) + \\
&+ (W_1{}^T X_1)^{-1} W_1{}^T X_2 N^{-1} W_2{}^T X_1 (W_1{}^T X_1)^{-1} \times \\
&\times W_1{}^T W_1 (X_1{}^T W_1)^{-1} X_1{}^T W_2 \tilde{N}^{-1} X_2{}^T W_1 (X_1{}^T W_1)^{-1} - \\
&- (W_1{}^T X_1)^{-1} W_1{}^T X_2 N^{-1} X_2{}^T W_1 (X_1{}^T W_1)^{-1} - \\
&- (W_1{}^T X_1)^{-1} W_1{}^T X_2 N^{-1} X_2{}^T W_1 (X_1{}^T W_1)^{-1} X_1 W_2 \tilde{N}^{-1} X_2 W_1 (X_1{}^T W_1)^{-1} - \\
&- (W_1{}^T X_1)^{-1} W_1{}^T W_2 \tilde{N}^{-1} X_2{}^T W_1 \times (X_1{}^T W_1)^{-1} - \\
&- (W_1{}^T X_1)^{-1} W_1{}^T X_2 N^{-1} W_2{}^T X_1 (W_1{}^T X_1)^{-1} W_1^T W_2 \tilde{N}^{-1} X_2{}^T W_1 (X_1{}^T W_1)^{-1} + \\
&+ (W_1{}^T X_1)^{-1} W_1{}^T \times X_2 N^{-1} W_2{}^T W_2 \tilde{N}^{-1} X_2{}^T W_1 (X_1{}^T W_1)^{-1}\,;
\end{aligned} \tag{12}
$$

$$
\begin{aligned}
D_{22} &= N^{-1} W_2{}^T X_1 (W_1{}^T X_1)^{-1} W_1{}^T W_1 (X_1{}^T W_1)^{-1} X_1{}^T W_2 \tilde{N}^{-1} - \\
&- N^{-1} W_2{}^T W_1 (X_1{}^T W_1)^{-1} X_1 W_2 \tilde{N}^{-1} + N^{-1} W_2{}^T \times \\
&\times W_2 \tilde{N}^{-1} N^{-1} W_2{}^T X_1 (W_1{}^T X_1)^{-1} W_2{}^T W_2 \tilde{N}^{-1}.
\end{aligned} \tag{13}
$$

По аналогии можно выписать выражения и для матриц D_{12} и D_{21}. Однако поскольку нас интересует СКО-оценки ДМИП, в выражения для которой входят только D_{11} и D_{22}, ограничимся приведенными формулами.

Используя (12) и (13), получаем следующие выражения для СКО оценивания МИП:

$$M\{\|\hat{\vec{\alpha}}_1 - \vec{a}_1\|^2\} = M\{\|\hat{\vec{\alpha}}_1 - M\{\hat{\vec{\alpha}}_1\}\|^2\} = \sigma^2 \mathrm{tr} D_{11}; \tag{14}$$

$$M\{\|\hat{\vec{\alpha}}_2 - \vec{a}_2\|^2\} = M\{\|\hat{\vec{\alpha}}_2 - M\{\hat{\vec{\alpha}}_2\}\|^2\} = \sigma^2 \mathrm{tr} D_{22}, \tag{15}$$

где D_{11} и D_{22} – определяются формулами (12) и (13) соответственно.

Сравним СКО-оценки ДМИП и МИП:

$$
\begin{aligned}
&M\{\|\hat{\vec{\alpha}}_1 - \vec{a}_1\|^2\} - M\{\|\vec{\alpha}_1 - \vec{a}_1\|^2\} = M\{\|\hat{\vec{\alpha}}_1 - M\{\hat{\vec{\alpha}}_1\}\|^2\} = \\
&= \sigma^2 \mathrm{tr} D_{11} - \sigma^2 \mathrm{tr} (W_1{}^T X_1)^{-1} W_1{}^T W_1 (X_1{}^T W_1)^{-1} - \|A_1 \vec{a}_2\|^2; \\
&M\{\|\hat{\vec{\alpha}}_2 - \vec{a}_2\|^2\} - M\{\|\vec{\alpha}_2 - \vec{a}_2\|^2\} = \\
&= M\{\|\hat{\vec{\alpha}}_1 - M\{\hat{\vec{\alpha}}_1\}\|^2\} = \sigma^2 \mathrm{tr} D_{22} - \sigma^2 \mathrm{tr} A_2 P A_2{}^T - \|A_2 A_1 \vec{a}_2\|^2.
\end{aligned}
$$

Подставив в данные формулы выражения для D_{11}, D_{22}, A_1, A_2 и P, после несложных преобразований получим, что оценки ДМИП имеют меньшую СКО по сравнению с МИП при выполнении условий

$$
\begin{aligned}
&\mathrm{tr} \tilde{W}_1{}^T [\sigma^2 (X_1 N^{-1} W_2{}^T P - I) P^T W_2 \tilde{N}^{-1} X_2{}^T - \\
&- X_2 (\sigma^2 N^{-1} W_2{}^T P + \vec{\alpha}_2 \vec{\alpha}_2{}^T X_2{}^T)] \tilde{W}_1 > 0\,;
\end{aligned} \tag{16}
$$

$$
\begin{aligned}
&\mathrm{tr}[\sigma^2 N^{-1} W_2{}^T P (I - P^T) W_2 \tilde{N}^{-1} - \\
&- \sigma^2 \tilde{W}_2{}^T P \tilde{W}_2 - A_2 A_1 \vec{\alpha}_2 \vec{\alpha}_2{}^T A_2{}^T A_2] > 0\,.
\end{aligned} \tag{17}
$$

Здесь

$$\tilde{W}_1^T = (W_1{}^T X_1)^{-1} W_1{}^T; \ \tilde{W}_2^T = (W_2{}^T X_2)^{-1} W_2{}^T.$$

Так как след положительно-определённой матрицы всегда положителен, то условия (16), (17) выполняются, если

$$
\begin{aligned}
&\tilde{W}_1{}^T [\sigma^2 (X_2 N^{-1} W_2{}^T P - I) P^T W_2 \tilde{N}^{-1} X_2{}^T - \\
&- X_2 (\sigma^2 N^{-1} W_2{}^T P + \vec{\alpha}_1 \vec{\alpha}_1{}^T X_2{}^T)] \tilde{W}_1 > 0\,;
\end{aligned} \tag{18}
$$

$$
\begin{aligned}
&\sigma^2 N^{-1} W_2{}^T P (I - P^T) W_2 \tilde{N}^{-1} - \sigma^2 \tilde{W}_2{}^T P \tilde{W}_2 - \\
&- A_2 A_1 \vec{\alpha}_2 \vec{\alpha}_2{}^T A_2{}^T A_2{}^T > 0
\end{aligned} \tag{19}
$$

в смысле положительной определённости матрицы разностей.

Если матрица $\tilde{W}_1$ невырождена, т.е. $\mathrm{rank} \tilde{W} = p$, то для выполнения первого неравенства требуется положительная определённость матрицы

$$
\begin{aligned}
&\sigma^2 (X_2 N^{-1} W_2{}^T P - I) P^T W_2 \tilde{N}^{-1} X_2{}^T - \\
&- X_2 (\sigma^2 N^{-1} W_2{}^T P + \vec{\alpha}_2 \vec{\alpha}_2{}^T X_2{}^T).
\end{aligned}
$$

Следует отметить, что полученные результаты являются обобщением результатов на случай ДМНК. Действительно, при построении ДМНК следует принять $W = X$, т.е. в качестве матрицы ИП используется сама матрица наблюдений X. В этом случае

$$W^T X = X^T W = X^T X,$$

$$(W^T X)^{-1} = (X^T W)^{-1} = (X^T X)^{-1}$$

и полученные формулы существенно упрощаются. Так,

$$D\{\vec{\alpha}_1\} = \sigma^2 (X_1{}^T X_1)^{-1};$$

$$D\{\vec{\alpha}_2\} = \sigma^2 (I - A_2 A_2)(X_2{}^T X_2)^{-1};$$

$$D_{МНК}\{\alpha\} = \sigma^2 (X^{\ T} X\ )^{-1}.$$

Кроме того, упрощаются элементы ковариационной матрицы $D\{\alpha\}$, так как $N^{-1} = N^{-1}$, $P = P^T$, $PP^T = P$. Учет этих соотношений и несложные преобразования позволяют получить результаты, приведенные в [5].

Литература: 1. *Себер Дж.* Линейный регрессионный анализ: Пер. с англ. М. Мир, 1980. 456с. 2. *Freud R.J., Vail R.W., Clumies-Ross C.W.* Residnal analisis. J.Amer. Stat. Association, 1961. Vol. 56. P.98-104. 3. *Wallance T.D.* Efficiencies for stepwise regressions. J. Amer. Stat. Association, 1964. Vol. 59. P.1179-1182. 4. *Джонстон Дж.* Эконометрические методы: Пер. с англ. М.: Статистика, 1980. 444 с. 5. *Абдуллаев Ф.М., Гейдаров Э.Х.* Рекуррентный двухступенчатый метод наименьших квадратов // Автоматика и телемеханика, 1981. N1. С.77-83. 6. *Фадеев Д.К., Фадеева В.Н.* Вычислительные методы линейной алгебры. М-Л: ГИФМЛ, 1963. 734 с.

Поступила в редколлегию 09.09.99

Рецензент: д-р техн. наук Нефедов Л.И.

Тимофеев Владимир Александрович, канд. техн. наук, доцент, ведущий научный сотрудник кафедры ЭВМ ХТУРЭ. Научные интересы: контроль динамических объектов. Адрес: Украина, 61166, Харьков, пр. Ленина, 14, тел. 40-93-54.