УДК 519.25

А.П. Сарычев

# ИТЕРАЦИОННАЯ ПРОЦЕДУРА ОЦЕНИВАНИЯ ПАРАМЕТРОВ СИСТЕМЫ РЕГРЕССИОННЫХ УРАВНЕНИЙ

*Рассмотрена задача моделирования в классе систем регрессионных уравнений в условиях неопределенности по степени статистической зависимости между случайными составляющими выходных переменных объекта. Разработана итерационная процедура оценивания параметров системы регрессионных уравнений, эффективность которой подтверждена методом статистических испытаний.*

*Ключевые слова: задача структурно-параметрической идентификации.*

Статистические методы математического моделирования технических систем, для которых отсутствуют точные априорные гипотезы, являются современным инструментом описания и прогнозирования состояния таких систем. Как объекты математического моделирования технические системы могут быть разделены на два класса. К первому относятся объекты, для которых исследователь может выбрать структуру модели, т.е. априорно задать модель с точностью до неизвестных параметров, тогда задачу построения модели называют задачей параметрической идентификации. Ко второму классу относятся объекты, для которых невозможно однозначно априорно выбрать структуру модели, тогда говорят о задаче структурно-параметрической, или просто структурной идентификации.

Распространенным классом моделей в задачах структурной идентификации является класс систем регрессионных уравнений, линейных по параметрам. Модели этого класса позволяют описывать и прогнозировать состояния объектов, нелинейных по входным переменным: для этого необходимо предварительно расширить множество входных переменных за счет нелинейных функций. Кроме этого, такой класс моделей можно применять и при моделировании динамических систем в установившихся режимах функционирования.

Одним из видов структурной неопределенности при моделировании в классе систем регрессионных уравнений является неопределенность по степени статистической зависимости между случайными составляющими разных выходных переменных объекта. Если такая зависимость существует, а ковариационная матрица случайных составляющих разных выходных переменных неизвестна, задача оценивания сводится к задаче минимизации функционала, который представляет собой логарифм определителя ковариационной матрицы остатков регрессионных уравнений. Известным методом решения этой задачи является так называемый "двухшаговый" метод: на первом шаге параметры регрессионных уравнений оцениваются методом наименьших квадратов независимо для каждой выходной переменной, а на втором шаге они оцениваются совместно с использованием оценки ковариационной матрицы, полученной на первом шаге по остаткам регрессионных уравнений [1]. Этот метод традиционно считался эвристическим, и поэтому его аналитическое обоснование и обобщение представляет собой актуальную задачу.

## 1. Априорные предположения об объекте

Пусть статический объект описывается множеством $m$ входных переменных $X = \{x_1, x_2, \ldots, x_m\}$ и множеством $h$ выходных переменных $Y = \{y(1), y(2), \ldots, y(h)\}$.

Пусть модель функционирования объекта имеет вид

$$y(k) = \overset{o}{y}(k) + \xi\ (k) = \sum_{j=1}^{m(k)} \overset{o}{\theta}_j(k)\, x_j(k) + \xi\ (k), \quad k = 1, 2, \ldots, h\,, \tag{1}$$

где $k$ – номер выходной переменной; $h$ – число выходных переменных; $y(k)$ – измеряемая с ошибкой $k$-я выходная переменная; $\overset{o}{y}(k)$ – незашумленная (ненаблюдаемая) составляющая $k$-й выходной переменной; $\xi\ (k)$ – ненаблюдаемая аддитивная случайная составляющая в $k$-й выходной переменной; $x_j(k)$ – $j$-я входная переменная из множества входных переменных $X(k) \neq \varnothing$ ($\varnothing$ – пустое множество), участвующих в формировании $k$-й выходной переменной; $m(k)$ – число входных переменных, принадлежащих множеству $X(k)$; $\overset{o}{\theta}\ (k) = (\overset{o}{\theta}_1(k), \overset{o}{\theta}_2(k), \ldots, \overset{o}{\theta}_{m(k)}(k))^{\mathrm{T}}$ – вектор неизвестных коэффициентов.

Пусть в результате наблюдения объекта для каждой выходной переменной $y(k)$, $k = 1, 2, \ldots, h$, получены: 1) $\mathrm{X}(k)$ – $(n \times m(k))$-матрица

$n$ наблюдений $m(k)$ входов множества $X(k)$, имеющая полный ранг, равный $m(k)$; 2) $\mathrm{y}(k)$ – $(n\times 1)$-вектор наблюдаемых значений выходной переменной $y(k)$. В соответствии с моделью (1) для наблюдений выполняется

$$\mathrm{y}(k)=\overset{o}{\mathrm{y}}(k)+\xi(k)=\mathrm{X}(k)\overset{o}{\theta}(k)+\xi(k),\quad k=1,2,\ldots,h, \tag{2}$$

где $\overset{o}{\mathrm{y}}(k)$ – $(n\times 1)$-вектор ненаблюдаемых незашумленных значений $k$-й выходной переменной; $\xi(k)$ – $(n\times 1)$-вектор ненаблюдаемых аддитивных случайных составляющих в наблюдениях $k$-й выходной переменной.

Пусть векторная случайная величина $\xi=(\xi(1),\xi(2),\ldots,\xi(h))^{\mathrm{T}}$ распределена по $h$-мерному нормальному закону: $\xi\sim N(0_h,\Sigma)$, и относительно $(n\times 1)$-векторов $\xi(k)$ выполняется

$$E\{\xi(k)\}=0_n,\quad E\{\xi(k)\xi^{\mathrm{T}}(k)\}=\sigma_{kk}\mathrm{I}_n,\quad k=1,2,\ldots,h; \tag{3}$$

$$E\{\xi(k)\xi^{\mathrm{T}}(q)\}=\sigma_{kq}\mathrm{I}_n,\quad k,q=1,2,\ldots,h,\quad k\neq q; \tag{4}$$

$$E\{\xi_{i_1}(k)\xi_{i_2}(k)\}=0,\quad i_1,i_2=1,2,\ldots,n,\quad i_1\neq i_2, \tag{5}$$

где $E\{\cdot\}$ – знак математического ожидания по всем возможным реализациям случайных векторов $\xi(k)$ и $\xi(q)$; $0_h$ – $(h\times 1)$-вектор, состоящий из нулей; $\sigma_{kk}$ – неизвестная конечная величина, дисперсия случайной величины $\xi(k)$; $\sigma_{kq}$ – неизвестная конечная величина, ковариация случайных величин $\xi(k)$ и $\xi(q)$; $\mathrm{I}_n$ – единичная $(n\times n)$-матрица.

**2. Вывод формул для оценивания коэффициентов в системе регрессионных уравнений**

Запишем (2) в объединенном виде. Введем обозначения

$$\mathrm{y}=\begin{pmatrix}\mathrm{y}(1)\\ \hline \mathrm{y}(2)\\ \hline \vdots\\ \hline \mathrm{y}(h)\end{pmatrix},\quad \overset{o}{\mathrm{y}}=\begin{pmatrix}\overset{o}{\mathrm{y}}(1)\\ \hline \overset{o}{\mathrm{y}}(2)\\ \hline \vdots\\ \hline \overset{o}{\mathrm{y}}(h)\end{pmatrix},\quad \overset{o}{\theta}=\begin{pmatrix}\overset{o}{\theta}(1)\\ \hline \overset{o}{\theta}(2)\\ \hline \vdots\\ \hline \overset{o}{\theta}(h)\end{pmatrix},\quad \xi=\begin{pmatrix}\xi(1)\\ \hline \xi(2)\\ \hline \vdots\\ \hline \xi(h)\end{pmatrix}, \tag{6}$$

$$\underline{\underline{\mathrm{X}}} = \left[\begin{array}{c|c|c|c} \mathrm{X}(1) & \mathrm{O}_{(n\times m(2))} & \cdots & \mathrm{O}_{(n\times m(h))} \\ \hline \mathrm{O}_{(n\times m(1))} & \mathrm{X}(2) & \cdots & \mathrm{O}_{(n\times m(h))} \\ \hline \vdots & \vdots & \ddots & \vdots \\ \hline \mathrm{O}_{(n\times m(1))} & \mathrm{O}_{(n\times m(2))} & \cdots & \mathrm{X}(h) \end{array}\right], \tag{7}$$

где $\mathrm{y}$ – объединенный $(N\times 1)$-вектор наблюдаемых зашумленных значений; $\overset{\circ}{\mathrm{y}}$ $-(N\times 1)$-вектор ненаблюдаемых значений; $\overset{\circ}{\theta}$ – $(M\times 1)$-вектор неизвестных коэффициентов; $\xi$ $-(N\times 1)$-вектор ненаблюдаемых случайных аддитивных составляющих; $\underline{\underline{\mathrm{X}}}$ – объединенная $(N\times M)$-матрица регрессоров; $N = n\,h$; $M = m(1)+m(2)+\ldots+m(h)$.

С учетом (6)–(7) систему $h$ регрессионных уравнений (2) можно записать

$$\mathrm{y} = \overset{\circ}{\mathrm{y}} + \xi = \underline{\underline{\mathrm{R}}}\,\overset{\circ}{\theta} + \xi \;. \tag{8}$$

Будем искать оценку неизвестных коэффициентов $\overset{\circ}{\theta}$ в виде

$$\mathrm{d} = \mathrm{C}\;\mathrm{y}, \qquad \mathrm{d} = \begin{pmatrix} \mathrm{d}(1) \\ \hline \mathrm{d}(2) \\ \hline \vdots \\ \hline \mathrm{d}(h) \end{pmatrix}, \qquad \mathrm{d}(k) = \begin{pmatrix} \mathrm{d}(k,1) \\ \hline \mathrm{d}(k,2) \\ \hline \vdots \\ \hline \mathrm{d}(k,h) \end{pmatrix}, \qquad k = 1,2,\ldots,h\,, \tag{9}$$

где матрицу $\mathrm{C}$, которая зависит от $\underline{\underline{\mathrm{X}}}$ и имеет размер $(M\times N)$, требуется определить.

Будем искать такую матрицу $\mathrm{C}$, при которой логарифм определителя ковариационной матрицы оценки коэффициентов (9) принимает минимальное значение и оценки коэффициентов несмещены.

Математическое ожидание и ковариационную матрицу оценки (9) вычислим по всем возможным реализациям случайных величин $\xi\,(k)$, $k = 1,2,\ldots,h$. Для математического ожидания оценки (9) должно выполняться

$$E\{\mathrm{d}\} = E\{\mathrm{C}\,\mathrm{y}\} = E\{\mathrm{C}\,(\overset{\circ}{\mathrm{y}} + \xi\,)\} = E\{\mathrm{C}\underline{\underline{\mathrm{R}}}\,\overset{\circ}{\theta}\,\} + E\{\mathrm{C}\xi\,\} = \overset{\circ}{\theta}\;. \tag{10}$$

Справедливость (10) следует из выполнения условий

$$\mathrm{C}\underline{\underline{\mathrm{R}}} = \mathrm{I}_M,\;\; E\{\mathrm{C}\,\xi\,\} = 0_M\,, \tag{11}$$

 ISSN 1562-9945

где первое следует из требования несмещенности оценок, а второе – из независимости элементов матрицы $\underline{\underline{X}}$ от случайных величин $\xi(k)$, $k = 1, 2, \ldots, h$, с учетом (3).

Пусть $\Sigma_{\xi}$ – ковариационная матрица введенного в (6) объединенного $(N \times 1)$-вектора ненаблюдаемых аддитивных случайных составляющих $\xi$. Тогда для ковариационной матрицы вектора оценок (9) выполняется

$$\mathrm{Cov}(\mathrm{d}) = E\left\{(\mathrm{d} - \overset{o}{\theta})(\mathrm{d} - \overset{o}{\theta})^{\mathrm{T}}\right\} = E\left\{(\mathrm{C}\overset{o}{\mathrm{y}} + \mathrm{C}\xi - \overset{o}{\theta})(\mathrm{C}\overset{o}{\mathrm{y}} + \mathrm{C}\xi - \overset{o}{\theta})^{\mathrm{T}}\right\} =$$

$$= E\left\{\mathrm{C}\xi\,\xi^{\mathrm{T}}\mathrm{C}^{\mathrm{T}}\right\} = \mathrm{C}\Sigma_{\xi}\mathrm{C}^{\mathrm{T}}, \tag{12}$$

где $E\{\cdot\}$ – операция математического ожидания, введенная при вычислении (10).

Запишем функцию Лагранжа в виде

$$L(\mathrm{C}, \Lambda) = \ln(\det[\mathrm{C}\Sigma_{\xi}\mathrm{C}^{\mathrm{T}}]) + \mathrm{tr}\,[\Lambda(\mathrm{C}\underline{\underline{\mathrm{R}}} - \mathrm{I}_M)]. \tag{13}$$

Тогда необходимые условия оптимальности имеют вид

$$\frac{\partial L}{\partial \mathrm{C}} = \frac{\partial}{\partial \mathrm{C}}\Big(\ln(\det[\mathrm{C}\Sigma_{\xi}\mathrm{C}^{\mathrm{T}}])\Big) + \frac{\partial}{\partial \mathrm{C}}\Big(\mathrm{tr}\,[\Lambda(\mathrm{C}\underline{\underline{\mathrm{R}}} - \mathrm{I}_M)]\Big) = \mathrm{O}_{M\times N}, \tag{14}$$

$$\frac{\partial L}{\partial \Lambda} = \frac{\partial}{\partial \Lambda}\Big(\mathrm{tr}\,[\Lambda(\mathrm{C}\underline{\underline{\mathrm{R}}} - \mathrm{I}_M)]\Big) = \mathrm{C}\underline{\underline{\mathrm{X}}} - \mathrm{I}_M = \mathrm{O}_{M\times M}. \tag{15}$$

Применяя правила матричного дифференцирования, из (14)–(15) получаем

$$\mathrm{C} = (\underline{\underline{\mathrm{R}}}^{\mathrm{T}}\Sigma_{\xi}^{-1}\underline{\underline{\mathrm{R}}})^{-1}\underline{\underline{\mathrm{R}}}^{\mathrm{T}}\Sigma_{\xi}^{-1}. \tag{16}$$

Для математического ожидания и ковариационной матрицы оценки выполняется

$$E\{\mathrm{d}\} = E\{\mathrm{Cy}\} = E\{(\underline{\underline{\mathrm{R}}}^{\mathrm{T}}\Sigma_{\xi}^{-1}\underline{\underline{\mathrm{R}}})^{-1}\underline{\underline{\mathrm{R}}}^{\mathrm{T}}\Sigma_{\xi}^{-1}(\underline{\underline{\mathrm{R}}}\overset{o}{\theta} + \xi)\} = \overset{o}{\theta}, \tag{17}$$

$$\mathrm{Cov}\{(\mathrm{d} - \overset{o}{\theta})(\mathrm{d} - \overset{o}{\theta})^{\mathrm{T}}\} = E\{(\underline{\underline{\mathrm{R}}}^{\mathrm{T}}\Sigma_{\xi}^{-1}\underline{\underline{\mathrm{R}}})^{-1}\underline{\underline{\mathrm{R}}}^{\mathrm{T}}\Sigma_{\xi}^{-1}\xi\,\xi^{\mathrm{T}}\Sigma_{\xi}^{-1}\underline{\underline{\mathrm{R}}}(\underline{\underline{\mathrm{R}}}^{\mathrm{T}}\Sigma_{\xi}^{-1}\underline{\underline{\mathrm{R}}})^{-1}\} = (\underline{\underline{\mathrm{R}}}^{\mathrm{T}}\Sigma_{\xi}^{-1}\underline{\underline{\mathrm{R}}})^{-1}. \tag{18}$$

Вычислим ковариационную матрицу $\Sigma_{\xi}$, т. е. дисперсии и ковариации случайных величин $\xi_i(k)$, $i = 1, 2, \ldots, n$, $k = 1, 2, \ldots, h$. Учитывая (3)–(5), получаем:

$$\Sigma_{\xi} = E\{\xi\ \xi^{\mathrm{T}}\} = \left[\begin{array}{c|c|c|c} E\{\xi(1)\xi^{\mathrm{T}}(1)\} & E\{\xi(1)\xi^{\mathrm{T}}(2)\} & \cdots & E\{\xi(1)\xi^{\mathrm{T}}(h)\} \\ \hline E\{\xi(2)\xi^{\mathrm{T}}(1)\} & E\{\xi(2)\xi^{\mathrm{T}}(2)\} & \cdots & E\{\xi(2)\xi^{\mathrm{T}}(h)\} \\ \hline \vdots & \vdots & \ddots & \vdots \\ \hline E\{\xi(h)\xi^{\mathrm{T}}(1)\} & E\{\xi(h)\xi^{\mathrm{T}}(2)\} & \cdots & E\{\xi(h)\xi^{\mathrm{T}}(h)\} \end{array}\right] = \left[\Sigma \otimes \mathrm{I}_n\right] \tag{19}$$

где $\mathrm{I}_n$ – единичная $(n \times n)$-матрица; $\Sigma \otimes \mathrm{I}_n$ – кронекеровское произведение матриц $\Sigma$ и $\mathrm{I}_n$.

**3. Итерационная процедура оценивания коэффициентов в системе регрессионных уравнений**

В выражение (19) для матрицы $\Sigma_{\xi}$ входит неизвестная матрица $\Sigma$. Используем это для построения итерационной процедуры вычисления неизвестных коэффициентов.

Пусть $\hat{\mathrm{d}}(r)$ – оценка вектора коэффициентов $\overset{\mathrm{o}}{\theta}$, а $\hat{\mathrm{d}}(k;r)$, $k = 1, 2, \ldots, h$ – оценка коэффициентов $\overset{\mathrm{o}}{\theta}(k)$ в виде (9), полученные на итерации с номером $r$. Тогда для системы регрессионных уравнений выполняется

$$\mathrm{y}(k) = \mathrm{X}(k)\hat{\mathrm{d}}(k;r) + \mathrm{u}(k;r) = \hat{\mathrm{y}}(k;r) + \mathrm{u}(k;r), \quad k = 1, 2, \ldots, h, \tag{20}$$

где $\hat{\mathrm{y}}(k;r)$ – $(n \times 1)$-вектор выходов модели, а $\mathrm{u}(k;r)$ – $(n \times 1)$-вектор остатков модели для $k$-й переменной на итерации с номером $r$.

Введем в рассмотрение матрицу наблюдений $\mathrm{Y}$, матрицу выходов $\hat{\mathrm{Y}}$ и матрицу остатков $\mathrm{U}$ системы моделей (20):

$$\mathrm{Y} = [\mathrm{y}(1), \mathrm{y}(2), \ldots, \mathrm{y}(h)], \quad \hat{\mathrm{Y}}(r) = [\hat{\mathrm{y}}(1,r), \hat{\mathrm{y}}(2,r), \ldots, \hat{\mathrm{y}}(h,r)], \tag{21}$$

$$\mathrm{U}(r) = [\mathrm{u}(1,r), \mathrm{u}(2,r), \ldots, \mathrm{u}(h,r)], \tag{22}$$

для которых выполняется

$$\mathrm{U}(r) = \mathrm{Y} - \hat{\mathrm{Y}}(r) = [\mathrm{y}(1) - \hat{\mathrm{y}}(1,r), \mathrm{y}(2) - \hat{\mathrm{y}}(2,r), \ldots, \mathrm{y}(h) - \hat{\mathrm{y}}(h,r)]. \tag{23}$$

Итерационная процедура вычисления коэффициентов предусматривает ряд итераций.

1. На итерации $r = 0$, полагая $\Sigma = \mathrm{I}_N$ – единичная $(N \times N)$-матрица, в качестве начального приближения получаем оценку обычного метода наименьших квадратов (МНК):

$$\hat{\mathrm{d}}(0) = \left( \underline{\underline{\mathrm{X}}}^{\mathrm{T}} \underline{\underline{\mathrm{X}}} \right)^{-1} \underline{\underline{\mathrm{X}}}^{\mathrm{T}} \mathrm{y}\,. \qquad (24)$$

2. На итерации $r = 1, 2, \ldots, r^*$ производим вычисления в такой последовательности.

a) Вычисляем выходы моделей:

$$\hat{\mathrm{y}}(k; r-1) = \mathrm{X}(k)\,\hat{\mathrm{d}}(k; r-1), \;\; k = 1, 2, \ldots, h\,. \qquad (25)$$

b) Вычисляем остатки моделей:

$$\mathrm{u}(k; r-1) = \mathrm{y}(k) - \hat{\mathrm{y}}(k; r-1) \;\; k = 1, 2, \ldots, h\,. \qquad (26)$$

c) Объединяем остатки в матрицу:

$$\mathrm{U}(r-1) = [\mathrm{u}(1, r-1), \mathrm{u}(2, r-1), \ldots, \mathrm{u}(h, r-1)]\,. \qquad (27)$$

d) Вычисляем оценку ковариационной матрицы $\Sigma_{\xi}$:

$$\hat{\Sigma}(r-1) = (n-1)^{-1}\,[\mathrm{U}^{\mathrm{T}}(r-1)\,\mathrm{U}(r-1)]\,, \qquad (28)$$

$$\hat{\Sigma}_{\xi}(r-1) = \hat{\Sigma}(r-1) \otimes \mathrm{I}_{n}\,. \qquad (29)$$

e) Вычисляем оценку $\hat{\mathrm{d}}(r)$:

$$\hat{\mathrm{d}}(r) = \left( \underline{\underline{\mathrm{X}}}^{\mathrm{T}} [\hat{\Sigma}_{\xi}(r-1)]^{-1} \underline{\underline{\mathrm{X}}} \right)^{-1} \underline{\underline{\mathrm{X}}}^{\mathrm{T}} [\hat{\Sigma}_{\xi}(r-1)]^{-1} \mathrm{y}\,, \qquad (30)$$

f) Вычисляем целевой функционал:

$$\Phi(r) = (\,(n-1)^{-1} \det\,[\mathrm{U}^{\mathrm{T}}(r)\mathrm{U}(r)]\,)^{1/h}\,. \qquad (31)$$

3. Итерационный процесс заканчиваем на итерации $r^*$ при выполнении условия

$$\delta^{2} = \Phi(r^* - 1) - \Phi(r^*) \le \delta_{0}^{2}, \qquad (32)$$

где $\delta_{0}^{2}$ – заданная величина.

Эффективность итерационной процедуры (24)–(32) подтверждена методом статистических испытаний.

**4. Заключение**

Рассмотрена задача моделирования в классе систем регрессионных уравнений в условиях неопределенности по степени статистической зависимости между случайными составляющими разных выходных переменных объекта. Разработана итерационная процедура оценивания параметров системы регрессионных уравнений, эффективность которой подтверждена методом статистических испытаний. Результаты, аналогичные приведенным, использованы для моделирования в классе систем регрессионных уравнений в условиях неопределенности по количеству та составу входных переменных – разработан критерий структурной идентификации метода группового учета аргументов [2].

**ЛИТЕРАТУРА**

1. Greene W. H. Econometric analysis : – 5th edition. – New Jersey : Prentice Hall, 2003. – 1056 p.
2. Сарычев А. П. Идентификация состояний структурно-неопределенных систем / А. П. Сарычев – Днепропетровск:НАН Украины и НКА Украины, Ин-т технической механики,2008.–268 с.

 ISSN 1562-9945