

ИДЕНТИФИКАЦИЯ И ДИСКРИМИНАЦИЯ S-ОБРАЗНЫХ КРИВЫХ РОСТА

к.т.н. В.Ю. Дубницкий

(представил д.ф.-м.н., проф. С.В. Смеляков)

Изложена процедура оценки параметров кривых, описывающих процессы с насыщением и выбора наилучшего приближения экспериментальных данных одной из них.

Постановка задачи. При описании процессов, для которых характерно явление насыщения, используют S-образные кривые роста [1]:

– кривую Перла-Рида

$$y_1 = \frac{a_0}{1 + a_1 \exp(-a_2 t)}, \quad a_0 > 0; a_1 > 0; a_2 > 0; \quad (1)$$

– логистическую кривую

$$y_2 = (a_3 + a_4 b_1^t)^{-1}, \quad 0 < b_1 < 1; a_3 > 0; a_4 > 0; \quad (2)$$

– кривые Гомперца вида

$$y_3 = a_5 \exp(-a_6 e^{-a_7 t}), \quad a_5 > 0; a_6 > 0 \quad (3)$$

$$\text{или } y_4 = \exp(a_8 b_2^t + a_9), \quad 0 < a_8 < 1; a_9 > 0; \quad (4)$$

– кривую Джонсона

$$y_5 = \exp\left(a_{10} - \frac{a_{11}}{a_{12} + t}\right); \quad a_{10} > 0; a_{11} > 0. \quad (5)$$

Несмотря на внешнюю схожесть, кривые (1) – (5) являются решением различных дифференциальных уравнений [2], что свидетельствует о различной природе процессов, вызвавших их появление.

Целью данной работы является создание вычислительной процедуры, позволяющей оценить по экспериментальным данным численные значения параметров трехпараметрических кривых (1) – (5), (идентифицировать модели); проверить соответствие расчетных значений фактическим данным (проверить адекватность моделей) и, в случае, если адекватными окажутся несколько моделей, то выбрать лучшую из них (дискриминация моделей).

1. Решение задачи идентификации в общем виде. Идентификацию моделей проводят, используя методы нелинейного регрессионного анали-

за [3, 4]. Предположим, что все условия классического регрессионного анализа выполнены. Тогда примем, что вектор оцениваемых параметров равен $A = (\alpha_1, \alpha_2, \alpha_3)$. Оценкой $\hat{A}$ вектора A будет такой вектор A , для которого

$$Q(A) = \sum_{k=1}^{\ell} (y_k - f(A; t_k))^2 \quad (6)$$

принимает минимальное значение, т.е.

$$\hat{A} = \operatorname{arginf} Q(A), \quad A \in R^m \quad (7)$$

при условии, что $m = 3$, $k = 1, \dots, \ell$ – известные экспериментальные данные.

Решение задачи (6) – (7) выполняли методом Ньютона в варианте, описанном в [4]:

$$A^{(1)} = A^{(0)} + H^{-1}q; \quad \dots \quad A^{(q)} = A^{(q-1)} + H^{-1}q. \quad (8)$$

Правила прекращения процесса вычислений приведены в [4]. В условии (8) принято, что $A^{(0)}$ – вектор начального приближения, q – анти-градиент функции (6), равный $(-q)$:

$$q_i = -2 \left(Q(A)^{1/2} \right) \frac{\partial f_k}{\partial \alpha_i}, \quad i = 1, 2, \dots, m; k = 1, 2, \dots, \ell. \quad (9)$$

Условимся, что далее, если это не будет вызывать недоразумений, перечисления для i и k будут опущены.

Пусть H – гессиан функции (6), вычисленный в каждой k -й точке:

$$H_{ij} = 2 \left(\sum_{k=1}^{\ell} \frac{\partial f_k}{\partial \alpha_i} \cdot \frac{\partial f_k}{\partial \alpha_j} - \left(Q(A)^{1/2} \right) \frac{\partial^2 f_k}{\partial \alpha_i \cdot \partial \alpha_j} \right); \quad i = \overline{1, m}; j = \overline{1, m}. \quad (10)$$

2. Решение задачи идентификации моделей (1) – (5).

Примем, что
$$y_{\infty} = \lim_{t \rightarrow \infty} f(A; t) \quad (11)$$

и
$$y_0 = f(A; t)_{t=0}. \quad (12)$$

Тогда, для кривой (1) с учетом (11) и (12) получим систему:

$$\begin{cases} y_{\infty} = a_0; \\ y_0 = a_0 / (1 + a_1); \\ \ln \left((y^{-1} - 1) / a_1 \right) = -a_2 t_k, \end{cases} \quad (13)$$

откуда
$$a_1 = (y_{\infty} - y_0) / y_0 \quad (14)$$

и
$$a_2 = (-1) \frac{\sum_{k=1}^{\ell} \ln \frac{y_k^{-1} - 1}{a_1} \cdot t_k}{\sum_{k=1}^{\ell} t_k^2}. \quad (15)$$

Для кривых вида (2) – (5) решения аналогичных систем приведены в табл. 1. Оптимизация выбранных начальных условий выполнена по процедуре, описанной в [5].

Таблица 1

Параметры кривых (2) – (5)

Тип кривой (номер уравнения)	Параметры кривой	Начальные приближения
Тип 2(2)	a_3	y_{∞}^{-1}
	a_4	$\frac{y_{\infty} - y_0}{y_{\infty} \cdot y_0}$
	b_1	$\sum_{k=1}^1 \ln \frac{\ln(y_k / y_{\infty})}{\ln(y_0 / y_{\infty})} \cdot t_k \cdot \left(\sum_{k=1}^1 t_k^2 \right)^{-1}$
Тип 3(3)	a_5	y_{∞}
	a_6	$-\ln \frac{y_0}{y_{\infty}}$
	a_7	$-\sum_{k=1}^1 \ln \frac{\ln \frac{y_k}{y_{\infty}}}{\ln \frac{y_0}{y_{\infty}}} \cdot t_k \cdot \left(\sum_{k=1}^1 t_k^2 \right)^{-1}$
Тип 4(4)	a_8	$\ln \frac{y_0}{y_{\infty}}$
	a_9	$\ln y_{\infty}$
	b_2	$\exp \left[\sum_{k=1}^1 \ln \frac{\ln \frac{y_k}{y_{\infty}}}{\ln \frac{y_0}{y_{\infty}}} \cdot t_k \cdot \left(\sum_{k=1}^1 t_k^2 \right)^{-1} \right]$
Тип 5(5)	a_{10}	$\ln y_{\infty}$
	a_{11}	$(\ln y_{\infty} - \ln y_0) a_{12}$
	a_{12}	$\frac{\ln y_{\infty} - \ln \tau}{\ln y_{\tau} - \ln y_0} \ln \tau^*$

*Примечание: τ – одно, произвольно выбранное, значение t_k из множества значений $T = t_1, t_2, \dots, t_l$.

Составляющие градиента q_i (9) получены из следующих условий.

Для кривой (1):

$$\frac{\partial y_1}{\partial a_0} = \frac{y_1}{a_0}; \quad k = 1, 2, \dots, l; \quad (16)$$

$$\frac{\partial y}{\partial a_1} = (-y_1) \cdot \frac{\exp(-a_2 t_k)}{1 + a_1 (\exp(-a_2 t_k))}; \quad k = 1, 2, \dots, l; \quad (17)$$

$$\frac{\partial y}{\partial a_2} = y_1 \cdot \frac{t a_1 \exp(-a_2 t_k)}{1 + a_1 \exp(-a_2 t_k)}; \quad k = 1, 2, \dots, l. \quad (18)$$

Для кривой (2):

$$\frac{\partial y_2}{\partial a_3} = (-y_2) \cdot \frac{1}{(a_3 + a_4 b_1^{t_k})^2}; \quad k = 1, 2, \dots, l; \quad (19)$$

$$\frac{\partial y_2}{\partial a_4} = (-y_2) \cdot \frac{b_1^{t_k}}{a_3 + a_4 b_1^{t_k}}; \quad k = 1, 2, \dots, l; \quad (20)$$

$$\frac{\partial y_2}{\partial b_1} = (-y_2) a_4 b_1^{t_k-1} t_k; \quad k = 1, 2, \dots, l. \quad (21)$$

Для кривой (3):

$$\frac{\partial y_3}{\partial a_5} = \frac{y_3}{a_5}; \quad (22)$$

$$\frac{\partial y_3}{\partial a_6} = (-y_3) \exp(-a_7 t_k); \quad (23)$$

$$\frac{\partial y_3}{\partial a_7} = t_k y_3 a_6 \exp(-a_7 t_k). \quad (24)$$

Для кривой (4):

$$\frac{\partial y_4}{\partial a_8} = b_2^{t_k} y_4; \quad (25)$$

$$\frac{\partial y_4}{\partial y_9} = y_4; \quad (26)$$

$$\frac{\partial y_4}{\partial b_2} = t_k a_8 b_2^{t_k-1} y_4. \quad (27)$$

Для кривой (5):

$$\frac{\partial y_5}{\partial a_{10}} = y_5; \quad (28)$$

$$\frac{\partial y_5}{\partial a_{11}} = -\frac{y_5}{a_{12} + t_k}; \quad (29)$$

$$\frac{\partial y_5}{\partial a_{12}} = \left(\frac{a_{11} y_5}{a_{12} + t_k^2} \right). \quad (30)$$

В условиях (16) – (30) по умолчанию принято, что функции $y_p(p = \overline{1,5})$ – функции аргумента t_k , $k = 1, 2, \dots, l$.

Попарные произведения первых частных производных, входящие в (10), можно легко получить из выражений (16) – (30), поэтому в целях экономии места они опущены. Выражения для элементов, составляющих гессианы, соответствующие уравнениям кривых (1) – (5), приведены ниже.

Для кривой (1):

$$\frac{\partial^2 y_1}{\partial a_0 \partial a_1} = -y^2 \exp(-a_2 t) \cdot a_0^{-2}; \quad (31)$$

$$\frac{\partial^2 y_1}{\partial a_0 \partial a_2} = y^2 a_1 t \exp(-a_2 t) a_0^{-2}; \quad (32)$$

$$\frac{\partial^2 y_1}{\partial a_1 \partial a_2} = \frac{y^2}{a_0} t \left(\exp(-a_2 t) - \frac{2}{a_0} a_1 \exp(-a_2 t) \right); \quad (33)$$

$$\frac{\partial^2 y_1}{\partial^2 a_0} = 0; \quad (34)$$

$$\frac{\partial^2 y_1}{\partial a_1^2} = \frac{2 \exp(-a_2 t)^2 \cdot y^3}{a_0^2}; \quad (35)$$

$$\frac{\partial^2 y_1}{\partial a_2^2} = a_1 t^2 \left(\frac{2y}{a} a_1 \exp(-a_2 t)^2 - \exp(-a_2 t) \right). \quad (36)$$

Для кривой (2):

$$\frac{\partial^2 y_2}{\partial a_3 \partial a_4} = 2 y_2^3 b_1^t; \quad (37)$$

$$\frac{\partial^2 y_2}{\partial a_3 \partial b_1} = 2 a_4 b_1^{t-1} \cdot t y_2^3; \quad (38)$$

$$\frac{\partial^2 y_2}{\partial a_4 \partial b_1} = t y^2 b_1^{t-1} (2 a_4 y - 1); \quad (39)$$

$$\frac{\partial^2 y_2}{\partial a_3^2} = 2 y_2^3; \quad (40)$$

$$\frac{\partial^2 y_2}{\partial a_4^2} = 2 (b_1^t)^2 y_2^3; \quad (41)$$

$$\frac{\partial^2 y_2}{\partial b_1^2} = a_4 b_1^{t-2} (2y_2 a_4 b_1^t \cdot t - t + 1). \quad (42)$$

Для кривой (3):

$$\frac{\partial^2 y_3}{\partial a_5 \partial a_6} = \frac{y_3}{a_5} \exp(-a_7 t); \quad (43)$$

$$\frac{\partial^2 y_3}{\partial a_5 \partial a_7} = \frac{y_3}{a_5} a_6 t \exp(-a_7 t); \quad (44)$$

$$\frac{\partial^2 y_3}{\partial a_6 \partial a_7} = y_3 t (\exp(-a_7 t) - a_6 \exp(-a_7 t)^2); \quad (45)$$

$$\frac{\partial^2 y_3}{\partial a_5^2} = 0; \quad (46)$$

$$\frac{\partial^2 y_3}{\partial a_6^2} = y_3 \exp(-a_7 t)^2; \quad (47)$$

$$\frac{\partial^2 y_3}{\partial a_7^2} = y a_6 t^2 (a_6 \exp(-a_7 t)^2 - \exp(-a_7 t)). \quad (48)$$

Для кривой (4):

$$\frac{\partial^2 y_4}{\partial a_8 \partial a_9} = b_2^t y_4; \quad (49)$$

$$\frac{\partial^2 y_4}{\partial a_8 \partial b_2} = y t b_2^{-1} (b_2^{t-2} + a_8 (b_2^t)^2); \quad (50)$$

$$\frac{\partial^2 y_4}{\partial a_9 \partial b_2} = y_4 a_8 t b_2^{t-1}; \quad (51)$$

$$\frac{\partial^2 y_4}{\partial a_8} = (b_2)^t y_4; \quad (52)$$

$$\frac{\partial^2 y_4}{\partial a_9} = y_4; \quad (53)$$

$$\frac{\partial^2 y_4}{\partial b_2} = y_4 a_8 t b_2^{t-2} (t - 1 + a_8 b_2^t \cdot t). \quad (54)$$

Для кривой (5):

$$\frac{\partial^2 y_5}{\partial a_{10} \partial a_{11}} = -\frac{y_5}{a_{12} + t}; \quad (55)$$

$$\frac{\partial^2 y_5}{\partial a_{10} \partial a_{12}} = \frac{a_{11} y_5}{(a_{12} + t)^2}; \quad (56)$$

$$\frac{\partial^2 y_5}{\partial a_{11} \partial a_{12}} = \frac{y_5}{(a_{12} + t)^2} \left(1 - \frac{a_{11}}{a_{12} + t} \right); \quad (57)$$

$$\frac{\partial^2 y_5}{\partial a_{10}^2} = y_5; \quad (58)$$

$$\frac{\partial^2 y_5}{\partial a_{11}^2} = \frac{y_5}{(a_{12} + t)^2}; \quad (59)$$

$$\frac{\partial^2 y_5}{\partial a_{12}^2} = \frac{y_5 a_{11}}{(a_{12} + t)^3} \left(\frac{a_{11}}{a_{12} + t} - 2 \right). \quad (60)$$

Дальнейшая процедура получения оценок $\hat{A}$ вектора параметров A совпадает с описанной ранее в работе [5].

Оценка адекватности полученных уравнений. В работе [3] отмечено, что для нелинейной регрессии не существует критериев адекватности, аналогичных критерию Фишера для линейной модели. Поэтому в случае нелинейной регрессии под проверкой адекватности модели понимают проверку гипотезы о том, что остатки модели имеют нулевое среднее.

Пусть

$$\bar{e}_p = \frac{1}{\ell} \sum_{k=1}^{\ell} (y_k - f_p(\hat{A}_p; t_k)), \quad p = \overline{1, 5}. \quad (61)$$

Тогда модель считают адекватной, если

$$e_p - t_{\ell-1, \alpha} \frac{S_p}{\sqrt{\ell}} \leq E \leq e_p + t_{\ell-1, \alpha} \frac{S_p}{\sqrt{\ell}}, \quad (62)$$

причем число нуль принадлежит указанному в (62) доверительному интервалу.

В условии (62) принято, что S_p – среднеквадратическое отклонение остатков для p -й ($p = \overline{1, 5}$) модели; $t_{\ell-1, \alpha}$ – значение обратной функции распределения Стьюдента с $(\ell - 1)$ степенью свободы на уровне значимости α .

4. Дискриминация моделей. Пусть количество моделей, удовлетворяющих условию (62), равно ω ; $\omega \geq 2$. Выбор лучшей модели проводят, используя следующее эвристическое допущение. Наилучшей среди нескольких адекватных моделей будет та, которая несет наибольшее количество информации о выборочной совокупности, для которой составляется соответствующее уравнение нелинейной регрессии. В работе [6]

показано, что в данном случае количество информации равно

$$I_p = \log_2 \sqrt{F_p}, \quad p = \overline{1,5}, \quad (63)$$

где

$$F_p = \frac{\sum_{k=1}^{\ell} (y_k - \bar{y})^2 / (\ell - 1)}{\sum_{k=1}^{\ell} \left(y_k - f(\hat{A}_p; t_k) \right)^2 / (k - u)}; \quad (64)$$

u – число оцениваемых параметров; в нашем случае $p = \overline{1,5}$, $u = 3$.

Таким образом, из всех конкурирующих моделей следует выбрать модель с таким номером $p^* (p^* \in \overline{1,5})$, для которого

$$p^* = \max_p I_p. \quad (65)$$

Выводы. 1. Сформулирована задача выбора наилучшей S-образной кривой, описывающей процессы, протекающие с насыщением.

2. Предложен критерий выбора наилучшей модели среди множества допустимых моделей, основанный на теоретико-информационной оценке результатов регрессионного анализа.

3. Предложен метод решения поставленной задачи.

ЛИТЕРАТУРА

1. Лук'яненко І.Г., Краснікова Л.І. Економетрика. – К.: Знання, 1998. – 494 с.
2. Урмаев А. С. Основы моделирования на аналоговых вычислительных машинах. – М.: Наука, 1974. – 320 с.
3. Бард Й. Нелинейное оценивание параметров. – М.: Статистика, 1979. – 349 с.
4. Демиденко Е.З. Линейная и нелинейная регрессии. – М.: Финансы и статистика, 1981. – 301 с.
5. Дубницький В.Ю., Мусиенко М.В. Оптимизация выбора начальных условий для оценивания параметров нелинейной регрессии и идентификация кривой Лорффера // Системи обробки інформації. – Х.: ХВУ. – 2004. – Вип. 2. – С. 180 – 185.
6. Мосягин А.Б., Таранцев А.А. Информационно-энтропийный подход в регрессионном анализе // Автоматика и телемеханика. – 1999. – № 12. – С. 201 – 205.

Поступила 8.12.2004

ДУБНИЦКИЙ Валерий Юрьевич, канд. техн. наук, доцент Харьковского банковского института. В 1975 году окончил ХИРЭ. Область научных исследований – исследование операций.