УДК 519.86

С. И. Спивак[1], д-р физ.-мат. наук, проф., зав. кафедрой, semen.spivak.@mail.ru,
О. Г. Кантор[2], канд. физ.-мат. наук, доц., ст. науч. сотрудник, e-mail: o_kantor@mail.ru
[1] ФГБОУ ВПО Башкирский государственный университет
[2] ФГБУН Институт социально-экономических исследований Уфимского научного центра РАН

Чебышевское приближение в задачах эконометрического моделирования[1]

Исследована возможность применения идеи задачи чебышеского приближения к проблеме идентификации эконометрических моделей. Показана связь оптимального решения задачи чебышевского приближения со средней ошибкой аппроксимации — критерием точности модели. Предложенные подходы к оценке параметров эконометрических моделей нивелируют ряд объективно существующих проблем и позволяют учитывать важные с точки зрения исследователя требования непосредственно на этапе идентификации.

Ключевые слова: *идентификация модели, верификация модели, чебышевское приближение*

Введение

В качестве основного инструментария для эконометрического моделирования используются математико-статистические методы [1]. Их применение осуществляется в соответствии с шестью основными этапами, в рамках которых проводятся, соответственно, постановка проблемы, предмодельный анализ, параметризация, сбор необходимой информации, идентификация модели и ее верификация. К числу основных проблем при этом относятся спецификация модели (т. е. определение непосредственного вида модели), ее идентификация — проблема, связанная с выбором и непосредственной реализацией методов оценивания параметров модели, и верификация — проверка соответствия полученной модели ряду характеристик, обусловливающих ее качественные признаки, такие как, например, точность и адекватность [1].

Важное значение имеют субъективные характеристики исследователя. Так, спецификация модели во многом зависит от того, насколько хорошо исследователь разбирается в сути проблемы и, безусловно, от его профессионализма и компетенции. Однако помимо субъективных существуют и некоторые объективные обстоятельства, оказывающие влияние на процесс и результаты эконометрического моделирования. В их числе можно отметить, как минимум, следующие два. Во-первых, непосредственное определение параметров модели (ее идентификация) осуществляется с учетом имеющейся экспериментальной информации, состав и качество которой во многом определяют возможности применения тех или методов. Так, например, известно, что для успешного применения математико-статистического инструментария число наблюдений (экспериментальных данных) должно в 6—7 раз превышать число параметров при экзогенных факторах [2]. Однако далеко не во всех практических исследованиях такое количество данных может быть обеспечено. Во-вторых, в процессе верификации необходимо проверять ограничивающие условия на значения параметров и (или) переменных модели, которые, очевидно, должны быть формализованы в виде каких-либо математических выражений. Классические математико-статистические методы, применяемые в эконометрическом моделировании, не позволяют использовать такие условия на

[1] Исследование выполнено при финансовой поддержке РФФИ (проект № 13-01-00749 "Качество моделей математической обработки наблюдений в социальных и экономических системах").

стадии идентификации модели, что может стать причиной многократных корректировок модели и, как следствие, затягивания процесса получения результата.

В связи с этим представляется целесообразной разработка специфических методов моделирования, нивелирующих указанные сложности. По мнению авторов, с позиций повышения эффективности проводимых исследований, наиболее перспективными являются следующие два направления: снижение "нагрузки" на этап верификации модели и достижение высокой точности аппроксимации как основного качественного показателя получаемой модели.

Сказанное означает осуществление, во-первых, уже на стадии определения параметров модели формализации в качестве ограничений тех требований, соблюдение которых предполагается контролировать на этапе верификации и, во-вторых, выбора в качестве критерия оптимальности функционала, характеризующего соответствие расчетных и экспериментальных данных.

Основу большинства математико-статистических подходов к решению эконометрических задач составляет метод наименьших квадратов [1, 2], суть которого заключается в определении параметров зависимостей, минимизирующих сумму квадратов невязок — отклонений расчетных значений эндогенной величины от экспериментальных. При этом предполагается, что невязки являются случайными величинами, подчиняются нормальному закону распределения с нулевым математическим ожиданием, гетероскедастичны и не автокоррелированы [2]. Проверка этих предположений представляет собой самостоятельную задачу, решение которой в случае ограниченного количества экспериментальной информации будет затруднено.

Критерий, используемый в методе наименьших квадратов, характеризует совокупную погрешность аппроксимации имеющихся экспериментальных данных. При этом точность аппроксимации оценивается усредненно, несмотря на то что значения некоторых невязок могут быть очень большими.

Альтернативой такому подходу к пониманию точности аппроксимации и соответствующему формированию критерия оптимальности параметров искомой зависимости является анализ значений невязок в каждом отдельном наблюдении. Примером такого подхода является метод выравнивания по Чебышеву, основная идея которого заключается в приближении экспериментальных данных таким способом, чтобы обеспечивалась равномерная точность описания во всей исследуемой области, т. е. каждая невязка по модулю не должна превышать некоторого оптимального уровня.

В настоящей работе исследована возможность применения задачи чебышевского выравнивания для идентификации моделей и показана связь оптимальных решений подобных задач с оценками для значений таких показателей совокупной погрешности аппроксимации, как сумма модулей невязок и средняя ошибка аппроксимации.

1. Чебышевское приближение

Задача чебышевского приближения системы m несовместных линейных уравнений с n неизвестными

$$\eta_i(\mathbf{x}) \equiv a_{i1}x_1 + \ldots + a_{in}x_n + a_i = 0,\ i = \overline{1, m} \qquad (1)$$

состоит в том [3], чтобы определить чебышевскую точку системы (1) $\mathbf{x}^* = \{x_1^*, \ldots, x_n^*\}$ такую, что

$$\max_{i = \overline{1, m}} |\eta_i(\mathbf{x}^*)| = \inf_{\mathbf{x}} \max_{i = \overline{1, m}} |\eta_i(\mathbf{x})| = \delta^*. \qquad (2)$$

Соотношения (1) определяют в n-мерном пространстве систему из m плоскостей. В соответствии с (2) точка $\mathbf{x}^*$ является наименее отклоняющейся по модулю от этой системы. Любая другая точка будет отклоняться по модулю, хотя бы от одной из плоскостей, на значение, большее δ^*.

Задача чебышевского приближения системы (1) сводится к задаче линейного программирования посредством введения дополнительной переменной λ, гарантирующей выполнение системы линейных неравенств

$$|\eta_i(\mathbf{x}) \leqslant \lambda,\ i = \overline{1, m}.$$

Это позволяет, используя очевидные свойства модулей, записать следующую задачу линейного программирования:

$$\min_{\mathbf{x},\, \lambda} \lambda, \qquad (3)$$

$$a_{i1}x_1 + \ldots + a_{in}x_n + a_i - \lambda \leqslant 0,\ i = \overline{1, m}, \qquad (4)$$

$$a_{i1}x_1 + \ldots + a_{in}x_n + a_i + \lambda \geqslant 0,\ i = \overline{1, m}. \qquad (5)$$

В том случае, когда имеются ограничивающие условия на значения переменных задачи (1)—(2), заданных в виде линейных ограничений, определяющих некоторый выпуклый многогранник X в n-мерном пространстве, задача определения параметров, подобная задаче чебышевского приближения, примет вид

$$\max_{i = \overline{1, m}} |\eta_i(\mathbf{x})| = \min_{\mathbf{x} \in X} \max_{i = \overline{1, m}} |\eta_i(\mathbf{x})| = \delta^*,$$

а эквивалентная ей задача линейного программирования запишется следующим образом:

$$\min_{\mathbf{x},\,\lambda} \lambda, \tag{6}$$

$$a_{i1}x_1 + ... + a_{in}x_n + a_i - \lambda \leqslant 0,\ i = \overline{1, m}, \tag{7}$$

$$a_{i1}x_1 + ... + a_{in}x_n + a_i + \lambda \geqslant 0,\ i = \overline{1, m}, \tag{8}$$

$$\mathbf{x} \in X. \tag{9}$$

Задача (3)—(5) в силу теоремы Вейерштрасса о достижении непрерывной на компакте функции своей точной нижней границы всегда имеет решение. По этой же причине, в случае если условия (9) не нарушают совместности множества допустимых решений (7)—(9), задача (6)—(9) также всегда имеет решение. Приведенным задачам можно дать следующие геометрические интерпретации. При некотором λ и при каждом фиксированном i ограничения (4) и (5) определяют в соответствующем пространстве полосу — часть пространства, ограниченную двумя плоскостями. Пересечение конечного числа таких полос может либо образовывать многогранник (область S на рис. 1, *а*), либо быть пустым (рис. 1, *б*). Поэтому геометрически задача (4)—(6) сводится к тому, чтобы найти такое λ^*, при котором множество S будет не пустым и имеющим минимальный объем (минимальную площадь в случае двух переменных).

Дополнительные условия на значения переменных, заданные в виде ограничений (9), и не противоречащие (7)—(8), "урезают" множество S или оставляют его неизменным, что, в свою очередь, может приводить к увеличению оптимального значения λ^*.

Задача чебышевского приближения системы линейных уравнений является частным случаем задачи минимизации суммы модулей линейных функций:

$$z(\mathbf{x}) = \sum_{i=1}^{m} |\eta_i(\mathbf{x})|. \tag{10}$$

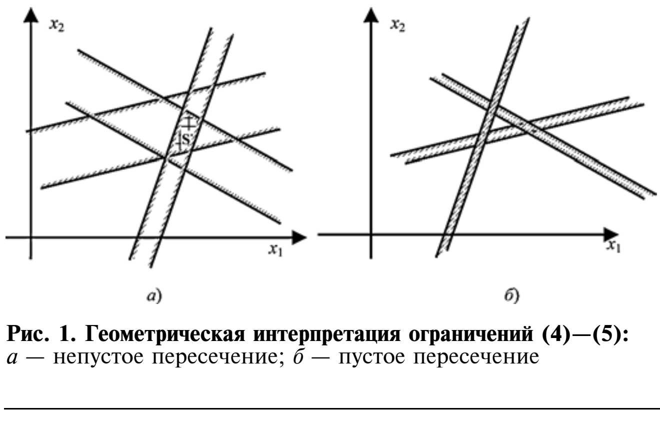

Рис. 1. Геометрическая интерпретация ограничений (4)—(5):
а — непустое пересечение; *б* — пустое пересечение

Введем дополнительные переменные $\{\lambda_i|,\ \lambda_i \geqslant 0,\ i = \overline{1, m}\}$ такие, чтобы выполнялись соотношения

$$|\eta_i(\mathbf{x})| \leqslant \lambda_i,\ i = \overline{1, m}. \tag{11}$$

Тогда задаче (10)—(11) может соответствовать эквивалентная задача линейного программирования [3]:

$$\min_{\mathbf{x},\,\{\lambda_i\}} \sum_{i=1}^{m} \lambda_i, \tag{12}$$

$$\lambda_i + a_{i1}x_1 + ... + a_{in}x_n + a_i \geqslant 0,\ i = \overline{1, m}, \tag{13}$$

$$\lambda_i - a_{i1}x_1 - ... - a_{in}x_n - a_i \geqslant 0,\ i = \overline{1, m}. \tag{14}$$

Из теоремы Вейерштрасса следует, что задача (12)—(14) всегда имеет решение, в том числе и при существовании дополнительных ограничений на переменные $\mathbf{x}$ в виде линейных неравенств и (или) равенств, не противоречащих условиям (13), (14).

Если в задачу (12)—(14) ввести условие равенства всех переменных $\{\lambda_i,\ i = \overline{1, m}\}$, то задача линейного программирования, эквивалентная задаче определения минимума суммы модулей линейных функций, будет иметь следующий вид:

$$\min_{\mathbf{x},\,\lambda} m\lambda,$$

$$\lambda + a_{i1}x_1 + ... + a_{in}x_n + a_i \geqslant 0,\ i = \overline{1, m},$$

$$\lambda - a_{i1}x_1 - ... - a_{in}x_n - a_i \geqslant 0,\ i = \overline{1, m},$$

и будет с точностью до константы m в целевой функции совпадать с задачей (3)—(5). Из этого следует очевидный вывод: оптимальное решение задачи чебышевского приближения системы линейных уравнений является допустимым решением задачи минимизации суммы модулей линейных функций, при этом оптимальное значение f^* целевой функции (12) не превышает оптимального значения λ^* целевой функции (3), увеличенного в m раз, т. е. справедливо соотношение $f^* \leqslant m\lambda^*$.

2. Связь идентификации эконометрической модели и задачи чебышевского приближения

Идентификация линейной модели как одного из этапов эконометрического моделирования заключается в том, чтобы на основании имеющейся экспериментальной информации $\{(\mathbf{x}_t, y_t),\ t = \overline{1, m}\}$ (где $\mathbf{x}_t = \{x_1|_t, ..., x_n|_t\}$ — вектор значений экзогенных факторов в момент времени t; y_t — значение эндогенного фактора в момент времени t; m — число наблюдений) определить параметры $\mathbf{a} = \{a_0, a_1, ..., a_n\}$ линейной зависимости

$$\hat{y} = a_0 + a_1x_1 + ... + a_nx_n. \tag{15}$$

Очевидно, что $\forall t = \overline{1, m}$ — расчетные значения эндогенной переменной в моменты времени t ($\hat{y}|_t$) должны быть близки к фактическим значениям y_t. Выражения для невязок имеют следующий вид:

$$\varepsilon_t = \hat{y}|_t - y_t = (a_0 + a_1x_1 + ... + a_nx_n)|_t - y_t,\ i = \overline{1, m}. \quad (16)$$

Абсолютное соответствие расчетных и экспериментальных данных эндогенной переменной означает равенство всех невязок нулю:

$$\varepsilon_t = (a_0 + a_1x_1 + ... + a_nx_n)|_t - y_t = 0,\ i = \overline{1, m}.$$

Последнее соотношение с точностью до обозначений совпадает с задачей чебышевского приближения (1). Таким образом, определение параметров линейной зависимости (15) может проводиться на основе решения задачи линейного программирования вида (3)—(5), а в том случае, если исследователь считает необходимым учесть некоторые дополнительные условия, на основе решения задачи линейного программирования вида (6)—(9). Важным преимуществом такого подхода, помимо априорного учета дополнительных условий, является отсутствие требования на число наблюдений, что следует из существования решений задач линейного программирования (3)—(5) и (6)—(9) вне зависимости от параметра m.

Одним из общеупотребительных числовых критериев, позволяющих делать вывод о качестве эконометрической модели, является средняя ошибка аппроксимации [1, 2]:

$$\bar{A} = \frac{1}{m}\sum_{t=1}^{m}\left|\frac{\hat{y}|_t - y_t}{y_t}\right|. \quad (17)$$

Величина (17) является безразмерной и допускает удобную интерпретацию в процентах. Считается, что эконометрическая модель имеет хорошую точность аппроксимации, если $\bar{A}$ не превышает 8...10 % [2].

Справедливо следующее.

Утверждение. В случае идентификации модели линейной структуры (15), задача определения ее параметров, минимизирующих выражение (17), эквивалентна задаче минимизации суммы модулей линейных функций (1), (10).

Доказательство. Задача определения точного вида модели (15) в соответствии с критерием (17) заключается в определении на основании экспериментальной информации $\{(\mathbf{x}_t, y_t),\ t = \overline{1, m}\}$ значений параметров $\{a_0, a_1, ..., a_n\}$, минимизирующих функционал (17); $\hat{y}|_t = \hat{y}(x_1|_t, ..., x_n|_t)$, где $x_j|_t$ — экспериментальное значение j-й экзогенной переменной в момент времени t.

Для характеристики близости расчетных и экспериментальных значений фактора y будем использовать значения невязок (16).

Умножим невязки ε_t на соответствующие константы $\frac{1}{my_t}$ и преобразуем полученные выражения следующим образом:

$$\varepsilon_t\frac{1}{my_t} = a_0\frac{1}{my_t} + a_1\frac{x_1|_t}{my_t} + ... + a_n\frac{x_n|_t}{my_t} - \frac{y_t}{my_t} =$$
$$= a_0\frac{1}{my_t} + a_1\frac{x_1|_t}{my_t} + ... + a_n\frac{x_n|_t}{my_t} - \frac{1}{m}.$$

Введем следующие обозначения:

$$b_{t0} = \frac{1}{my_t};\ b_{t1} = \frac{x_1|_t}{my_t},\ ...,\ b_{tn} = \frac{x_n|_t}{my_t};\ b_t = -\frac{1}{m},\ t = \overline{1, m}.$$

Тогда задача идентификации модели (15) может быть сформулирована следующим образом:

$$\min_{a_0, a_1, ..., a_n}\bar{A} = \min_{a_0, a_1, ..., a_n}\sum_{t=1}^{m}|\hat{\varepsilon}_t(a_0, a_1, ..., a_n)|, \quad (18)$$

$$\hat{\varepsilon}_t(a_0, a_1, ..., a_n) \equiv b_{t0}a_0 + b_{t1}a_1 + ... + b_{tn}a_n + b_t,$$
$$t = \overline{1, m}. \quad (19)$$

Задача (18)—(19) с точностью до обозначения переменных совпадает с задачей минимизации суммы модулей линейных функций (1), (10). Утверждение доказано.

Заметим, что значение средней ошибки аппроксимации очевидным образом ограничивается сверху:

$$\bar{A} = \frac{1}{m}\sum_{t=1}^{m}\left|\frac{\hat{y}|_t - y_t}{y_t}\right| \leqslant \frac{1}{m\min\limits_{t=\overline{1,m}}\{|y_t|\}}\sum_{t=1}^{m}|\hat{y}|_t - y_t|.$$

Таким образом, с учетом ранее доказанного неравенства $f^* \leqslant m\lambda^*$ для идентифицируемой модели линейного вида справедливо соотношение

$$\bar{A}^* \leqslant \frac{f^*}{m\min\limits_{t=\overline{1,m}}\{|y_t|\}} \leqslant \frac{\lambda}{\min\limits_{t=\overline{1,m}}\{|y_t|\}}, \quad (20)$$

где $\bar{A}^*$ — оптимальное значение критерия (18).

На практике соотношение (20) может давать достаточно грубую оценку $\bar{A}^*$ (см. ниже п. 3 и пп. 4.2). Однако если по результатам решений задач определения параметров линейных моделей с помощью сведения к задачам чебышевского приближения и минимизации суммы модулей линейных функций (в том числе и при наличии дополнительных условий) рассчитывать средние ошибки аппроксимации ($\bar{A}^{\text{чеб}}$ и $\bar{A}^{\text{мод}}$ соответственно), то будет выполняться соотношение

$$\bar{A}^* \leqslant \bar{A}^{\text{мод}} \leqslant \bar{A}^{\text{чеб}}, \quad (21)$$

Таблица 1

Результаты идентификации линейной модели урожайности зерновых культур

Модель	Критерий	Характеристики		
		R^2	F	$\bar{A}$, %
5 экзогенных факторов				
Задача чебышевского приближения с дополнительными условиями: $\hat{y} = 4{,}130 + 19{,}332x_2 + 3{,}945x_4 - 5{,}010x_5$	$\lambda^* = 2{,}258$	0,442	6,73	14,99
Задача минимизации суммы модулей невязок: $\hat{y} = 6{,}665 + 0{,}445x_1 + 4{,}560x_4 - 1{,}330x_5$	$f^* = 17{,}273$	0,460	7,24	8,32
Задача минимизации средней ошибки аппроксимации: $\hat{y} = 6{,}504 + 0{,}435x_1 + 1{,}562x_2 + 2{,}943x_4 - 0{,}645x_5$	$\bar{A}^* = 0{,}081$	0,464	7,38	8,09
2 экзогенных фактора (x_3 и x_4)				
Задача чебышевского приближения с дополнительными условиями: $\hat{y} = 8{,}858 + 0{,}185x_3 + 2{,}353x_4$	$\lambda^* = 2{,}519$	0,430	6,41	17,00
Задача минимизации суммы модулей невязок: $\hat{y} = 6{,}754 + 0{,}336x_3 + 3{,}777x_4$	$f^* = 18{,}735$	0,474	7,66	9,01
Задача минимизации средней ошибки аппроксимации: $\hat{y} = 6{,}807 + 0{,}372x_3 + 2{,}425x_4$	$\bar{A}^* = 0{,}086$	0,439	6,65	8,59

справедливость которого следует из смысла $\bar{A}^*$ и из того факта, что задача чебышевского приближения является частным случаем задачи минимизации суммы модулей линейных функций, и поэтому $\bar{A}^{\text{мод}} \leqslant \bar{A}^{\text{чеб}}$.

Таким образом, результаты решения задачи чебышевского приближения позволяют получить оценку лучшего значения средней ошибки аппроксимации — критерия близости расчетных и экспериментальных данных (табл. 1).

3. Моделирование линейных зависимостей

Возможности описанного подхода продемонстрируем на примере, рассмотренном в работе [1], посвященном исследованию зависимости урожайности зерновых культур (y, ц/га) от следующих пяти переменных:

- число тракторов приведенной мощности на 100 га (x_1, ед./100 га);
- число зерноуборочных комбайнов на 100 га (x_2, ед./100 га);
- число орудий поверхностной обработки почвы на 100 га (x_3, ед./100 га);
- количество удобрений, расходуемых на гектар (x_4, т/га);
- количество химических средств защиты растений, расходуемых на гектар (x_5, ц/га).

Исходя из имеющихся 20 наблюдений по методу наименьших квадратов была получена следующая линейная зависимость:

$$\hat{y} = 3{,}515 - 0{,}006x_1 + 15{,}542x_2 + 0{,}110x_3 + 4{,}475x_4 - 2{,}932x_5. \quad (22)$$

Линейная зависимость (22) имеет недостаточно хорошие характеристики: коэффициент детерминации $R^2 = 0{,}517$, F-критерий Фишера $F = 3{,}001$, средняя ошибка аппроксимации $\bar{A} = 10{,}42$ %, t-статистика значима только у коэффициента при переменной x_4. В дальнейшем, посредством исключения из рассмотрения "несущественных" экзогенных переменных была получена зависимость

$$\hat{y} = 7{,}29 + 3{,}48x_3 + 3{,}48x_4$$

со статистически значимыми коэффициентами и следующими характеристиками: $R^2 = 0{,}482$, $F = 7{,}921$, $\bar{A} = 10{,}47$ %.

Относительно приведенного примера обращает внимание тот факт, что параметры при первых трех переменных в идентифицируемой линейной модели (15), очевидно, не могут быть отрицательными ввиду того, что для обеспечения высокой урожайности важен требуемый объем работ, а не количество технических средств труда, которыми этот объем будет выполнен.

Исходя из этого при формировании системы ограничений задач линейного программирования на значения искомых параметров были наложены дополнительные условия: $a_1 \geqslant 0$, $a_2 \geqslant 0$, $a_3 \geqslant 0$.

Заметим, что приведенные результаты подтверждают справедливость соотношений (20)[1] и (21). Соотношение (20) для случая зависимости с пятью и с двумя переменными имеет вид $0{,}081 \leqslant 0{,}125 \leqslant 0{,}327$ и $0{,}086 \leqslant 0{,}136 \leqslant 0{,}365$ соответственно.

4. Идентификация нелинейных моделей

4.1. Формализация задач

Если спецификация модели $\hat{y} = \hat{y}(\mathbf{x}, \mathbf{a})$ не является линейной, то определение параметров **a** может происходить на основе решения задачи математического программирования:

$$\min_{\mathbf{a},\, \lambda} \lambda, \quad (23)$$

$$|\varepsilon_t| \leqslant \lambda,\ t = \overline{1, m}, \quad (24)$$

$$\mathbf{a} \in A, \quad (25)$$

[1] В рассматриваемом примере $\min\limits_{i = \overline{1, m}} \{y_t\} = 6{,}9$.

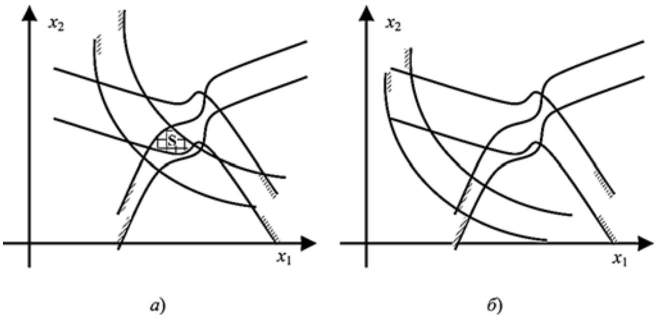

Рис. 2. Геометрическая интерпретация ограничений (24):
а — непустое пересечение; *б* — пустое пересечение

где по аналогии с приведенными выше задачами (3)—(5) и (6)—(9) ограничения (24) характеризуют условия на значения невязок в каждом наблюдении, а ограничения (25), суть — формализованные требования, подлежащие учету.

Для случая, когда невязки (16) представляют собой монотонные функции, каждое ограничение (24) геометрически представляет собой часть пространства соответствующей размерности между поверхностями, задаваемыми уравнениями одинаковой спецификации и отличающимися лишь свободными константами (рис. 2). Поэтому если ограничения задачи (23)—(25) являются совместными, то, очевидно, что ее решение существует.

Если же в качестве критерия оптимальности будет выбрана сумма модулей невязок, то задача определения параметров идентифицируемой зависимости примет вид

$$\min_{\mathbf{a},\ \{\lambda_t\}} \sum_{t=1}^{m} \lambda_t, \tag{26}$$

$$|\varepsilon_t| \leqslant \lambda_t,\ t = \overline{1, m},$$

$$\mathbf{a} \in A.$$

При этом, учитывая отсутствие привязки формулы для вычисления средней ошибки аппроксимации (17) к спецификации модели, соотношения, аналогичные (20), также будут выполняться, с той лишь разницей, что f^* — оптимальное значение критерия (26), λ^* — оптимальное значение критерия оптимизации в задаче (23)—(25), а $\bar{A}^*$ — оптимальное значение целевой функции соответствующей задачи определения параметров идентифицируемой нелинейной функции, в которой средняя ошибка аппроксимации является критерием минимизации. В этом случае величина λ^*, строго говоря, не является решением задачи чебышевского приближения, но ее смысл остается прежним: она показывает минимальное значение всех возможных максимальных по модулю невязок.

4.2. Моделирование нелинейных зависимостей

Существенные трудности в применении математико-статистических методов для определения модельных параметров возникают как при недостаточном количестве экспериментальных данных в расчете на каждый искомый параметр, так и в ситуации, когда невозможно осуществить линеаризацию модели с нелинейной спецификацией. Описанные в настоящей работе подходы позволяют определять параметры модели и при наличии такого рода проблем. Для подтверждения сказанного рассмотрим пример моделирования гипотетического временного ряда (табл. 2, рис. 3).

По результатам визуального анализа уровней временного ряда (рис. 3) была установлена возможность использования следующей спецификации модели:

$$\hat{y} = a \cdot t^b + c. \tag{27}$$

Модель (27) не допускает осуществления линеаризации, а количество имеющихся данных (10 на три оцениваемых параметра) является недостаточным для получения значимых характеристик, рассчитываемых в рамках математико-статистических методов.

В качестве дополнительного ограничения на значения параметров модели (27) использовалось формализованное представление требования на уровень временного ряда в момент времени $t = 12$: :$9{,}2 \leqslant \hat{y}|_{12} \leqslant 9{,}5$. (Такого рода условия могут формироваться на основе объективных данных или мнений экспертов, и их априорный учет может повысить степень доверия к результирующим моделям.)

Результаты проведенных расчетов (табл. 3, рис. 3) свидетельствуют о практически одинаковой значимости каждой из идентифицированных моделей, что предоставляет исследователю определенную свободу выбора, и может быть особенно существенным в случае получения дополнительной инфор-

Таблица 2

Исходные данные для моделирования временного ряда

t	1	2	3	4	5	6	7	8	9	10
y_t	1,6	3,6	3,7	5,4	6,1	6,2	6,6	8,0	8,3	8,7

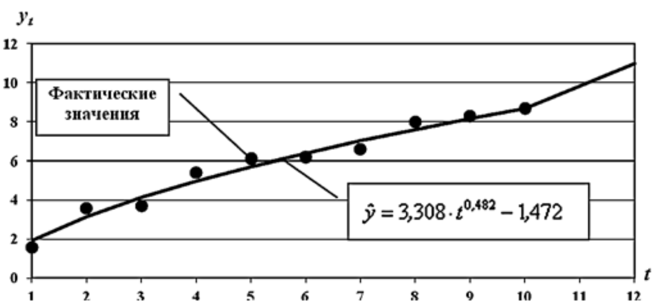

Рис. 3. Геометрическая интерпретация модели временного ряда

Таблица 3

Результаты идентификации модели временного ряда $\hat{y} = a \cdot t^b + c$

Модель	Критерий	Характеристики			
		R^2	F	$\bar{A}$, %	$\hat{y}\|_{12}$
Задача чебышевского выравнивания с дополнительными условиями: $\hat{y} = 3{,}308 \cdot t^{0{,}482} - 1{,}472$	$\lambda^* = 0{,}449$	0,975	79,51	7,18	9,5
Задача минимизации суммы модулей невязок: $\hat{y} = 7{,}302 \cdot t^{0{,}295} - 5{,}704$	$f^* = 2{,}617$	0,974	75,35	5,25	9,5
Задача минимизации средней ошибки аппроксимации: $\hat{y} = 8{,}299 \cdot t^{0{,}269} - 6{,}699$	$\bar{A}^* = 0{,}052$	0,973	72,83	5,17	9,5

мации о значениях параметров модели или ее переменных.

Оценки (20) для полученных моделей выполняются: $0{,}052 \leqslant 0{,}164 \leqslant 0{,}281$.

5. Направление развития исследований

Результаты проведенных расчетов по идентификации моделей (п. 3, пп. 4.2) свидетельствуют о существовании множества наборов значений параметров, обеспечивающих приемлемые качественные характеристики моделей. В связи с этим является целесообразным проведение на основе использования методов математического моделирования дополнительных исследований, направленных на определение лучшего из числа таких наборов. При этом очевидно, что в зависимости от предъявляемых к модели требований, вполне возможна ситуация, при которой оптимальными будут являться не единичные значения параметров, а некоторые интервалы их значений. В этом смысле совокупность оптимальных значений параметров идентифицируемой модели может трактоваться как "область неопределенности" параметров модели, а окончательный вид модели может быть выявлен на основе индивидуальных предпочтений исследователя и (или) с учетом каких-либо дополнительных требований при непосредственном использовании необходимых численных методов и программных продуктов.

В приведенном контексте понимание "областей неопределенности" параметров эконометрических моделей тесно переплетается с идеями, высказанными в работе [4]. Применение подхода Л. В. Канторовича в сочетании со специально разработанными моделями (аналогичными приведенным в настоящей работе), алгоритмами и программными продуктами позволило авторам [5—8] решить проблему идентификации модели системной динамики численности населения Российской Федерации.

Заключение

Большое разнообразие практических задач далеко не всегда позволяет использовать математико-статистические методы при построении эконометрических моделей, что обусловлено в том числе следующими причинами:

- применение данных методов детально разработано для линейных и линеаризуемых моделей;
- ограниченное число наблюдений не позволяет получать значимые характеристики для идентифицируемой модели;
- наличие большого количества требований, на соответствие которым модель должна проверяться на этапе ее верификации, части из которых она может не удовлетворять.

При этом необходимо принимать во внимание и объективно существующие трудности, связанные с возможной неточностью исходных данных, что можно объяснить например, многостадийностью процесса сбора информации, существенной долей субъективизма, а в некоторых случаях и умышленным искажением информации с целью "приукрасить" реальное положение дел. Оценить точность полученных наблюдений чаще всего не представляется возможным в силу отсутствия эталонных значений и возможности многократного проведения наблюдений в одних и тех же условиях, а это, в свою очередь, может стать причиной, по которой будет построена "точная модель неточных данных".

Сказанное обусловливает необходимость разработки специальных методов идентификации эконометрических моделей. И одним из возможных подходов к реализации этого является использование идей задач чебышевского приближения. Приведенные в работе постановки задач позволяют определять параметры эконометрических зависимостей с учетом требований, предъявляемых как к самим моделям, так и к составляющим их элементам. При этом существует возможность оценки степени соответствия экспериментальных и расчетных данных, определяемой показателем средней ошибки аппроксимации. Следует особо отметить, что приведенные в работе способы определения параметров идентифицируемых моделей позволяют учитывать условия на будущие значения эндогенных переменных. Это является существенным при анализе факторов, имеющих высокую инерционность, в силу чего для них может быть оценен диапазон возможных значений в краткосрочной перспективе.

Приведенные в работе расчеты в силу достаточно хороших качественных характеристик позволяют не только констатировать целесообразность использования описанного подхода для практического применения, но и могут составить основу для проведения более детальных исследований, направленных на определение точного вида модели.

Список литературы

1. **Айвазян С. А., Мхитарян В. С.** Прикладная статистика и основы эконометрики. М.: ЮНИТИ, 1998. 1022 с.
2. **Эконометрика:** учебник / Под ред. И. И. Елисеевой. М.: Финансы и статистика, 2008. 575 с.
3. **Зуховицкий С. И., Авдеева Л. И.** Линейное и выпуклое программирование. М.: Наука, 1967. 460 с.
4. **Канторович Л. В.** О некоторых новых подходах к вычислительным методам и обработке наблюдений // Сибирский математический журнал. 1962. Т. 3, № 5. С. 701—709.
5. **Спивак С. И., Кантор О. Г.** Построение модели системной динамики в условиях ограниченной экспертной информации // Информатика и ее применения. 2014. Т. 8, № 2. С. 111—121.
6. **Спивак С. И., Кантор О. Г.** Риск ошибочных прогнозных оценок параметров социально-экономических систем // Управление риском. 2013. № 1. С. 2—6.
7. **Спивак С. И., Кантор О. Г., Юсупова Г. Н.** Алгоритм распараллеливания в задачах перебора значений многомерных величин на основе декомпозиции по данным // Системы управления и информационные технологии. 2014. № 2 (56). С. 52—57.
8. **Kantor O. G., Spivak S. I.** Interval Estimation of System Dynamics Model Parameters // International Journal of Applied Engineering Research. 2014. V. 9, no. 19. P. 5689—5695.

S. I. Spivak, Head of Department of Mathematical Modeling; semen.spivak.@mail.ru,
Bashkir State University, Ufa,
O. G. Kantor, Senior scientist; o_kantor@mail.ru,
Institute of Social and Economic Research Ufa Research Centre Russian Academy of Sciences, Ufa

Chebyshev Approximation in Problems of Econometric Modeling

A wide variety of practical problems is not always permit to use mathematical and statistical methods in the construction of econometric models. This may be due, for example, by the following reasons: the use of these methods in detail designed for linear and linearizable models; a limited number of observations do not allow to get meaningful statistical characteristics. In addition, the resulting model should meet the requirements ensuring its adequacy. In the event of any breach of the adequacy researcher have to change the form of the function and carry out a complete cycle of remodeling.

The possibility of using the Chebyshev approximation for identification of econometric models is studied at the paper. The connection between the optimal solution of Chebyshev approximation with an average error of approximation — the criterion of model accuracy is shown. The proposed approaches to the estimation of econometric models parameters allow for researcher to take into account a number of objectively existing problems and consider the important requirements at the stage of identification directly.

Keywords: *model identification, model verification, Chebyshev approximation, average approximation error, uncertainty sets of the model parameters*

References

1. **Ajvazjan S. A., Mhitarjan V. S.** *Prikladnaja statistika i osnovy jekonometriki.* M.: JuNITI, 1998. 1022 p.
2. **Jekonometrika:** Uchebnik / Pod red. I. I. Eliseevoj. M.: Finansy i statistika, 2008. 575 p.
3. **Zuhovickij S. I., Avdeeva L. I.** *Linejnoe i vypukloe programmirovanie.* M.: Nauka, 1967. 460 p.
4. **Kantorovich L. V.** O nekotoryh novyh podhodah k vychislitel'nym metodam i obrabotke nabljudenij. *Sibirskij matematicheskij zhurnal.* 1962, vol. 3, no. 5, pp. 701—709.
5. **Spivak S. I., Kantor O. G.** Postroenie modeli sistemnoj dinamiki v uslovijah ogranichennoj jekspertnoj informacii, *Informatika i ee primenenija*, 2014, vol. 8, no. 2, pp. 111—121.
6. **Spivak S. I., Kantor O. G.** Risk oshibochnyh prognoznyh ocenok parametrov social'no-jekonomicheskih sistem, Upravlenie riskom, 2013, no. 1, pp. 2—6.
7. **Spivak S. I., Kantor O. G., Jusupova G. N.** Algontm rasparallelivanija v zadachah perebora znachenij mnogomernyh velichin na osnove dekompozicii po dannym, *Sistemy upravlenija i informacionnye tehnologii*, 2014, no. 2 (56), pp. 52—57.
8. **Kantor O. G., Spivak S. I.** Interval Estimation of System Dynamics Model Parameters, *International Journal of Applied Engineering Research*, 2014, vol. 9, no. 19, pp. 5689—5695.