

УДК 512.55 (076)

Лаврентик О.А.<sup>1</sup>, Лаврентик А.И.<sup>2</sup>, Симченко О.И.<sup>3</sup>, Федоренко В.В.<sup>4</sup>

### ПОСТРОЕНИЕ МАТЕМАТИЧЕСКИХ МОДЕЛЕЙ ЭКОНОМЕТРИЧЕСКИХ СИСТЕМ НА ОСНОВЕ ЯВЛЕНИЯ МУЛЬТИКОЛЛИНЕАРНОСТИ

*Построены математические модели анализа эконометрических систем в условиях мультиколлинеарности, на основе которых разработано программное обеспечение, позволяющее проводить анализ результатов в зависимости от различных факторов и прогнозировать развитие системы в будущем.*

Экономические системы характеризуются определенными экономическими параметрами. Поведение и значение параметров зависит от большого количества факторов, которые влияют на эти параметры, причем полностью учесть все факторы невозможно. Одной из важных проблем, которые возникают при построении моделей множественной регрессии на основе статистических данных, является наличие мультиколлинеарности – линейной зависимости между некоторыми аргументами [1]. Это приводит к смещению оценок параметров модели, которые рассчитываются по методу наименьших квадратов. На основе этих оценок невозможно сделать конкретные выводы о результатах взаимосвязи между показателем и факторами.

Предлагаемая работа направлена на устранение вышеперечисленных недостатков и разработку математических моделей и программного обеспечения, позволяющего проанализировать построенную эконометрическую модель, выявить мультиколлинеарность, если таковая имеется, и попытаться устранить ее наиболее подходящим методом, а также строить зависимости при различных исходных данных, способствовать более полному учету влияния факторов и прогнозировать развитие системы в будущем.

Цель работы – построение математических моделей и программного обеспечения анализа эконометрических систем в условиях мультиколлинеарности, что позволит существенно ускорить процесс обработки информации, выявить из возможных вариантов развития системы оптимальные, получить адекватные модели.

Каждой области экономических исследований, связанной с анализом эмпирических данных соответствуют свои эконометрические модели. При этом можно выделить следующие основные этапы [1, 2]:

1) Постановочный.

Формируется цель исследования, набор участвующих в модели экономических переменных. Обычно рассматривают анализ исследуемого экономического объекта (процесса); прогноз его экономических показателей, имитацию развития объекта при различных значениях экзогенных переменных (отражая их случайный характер, изменение во времени), выработку управленческих решений. При выборе экономических переменных необходимо теоретическое обоснование каждой переменной. Объясняющие переменные не должны быть связаны функциональной или тесной корреляционной зависимостью, так как это может привести к невозможности оценки параметров модели или к получению неустойчивых, не имеющих реального смысла оценок, т. е. к явлению мультиколлинеарности.

2) Априорный.

Проводится анализ сущности изучаемого объекта, формирование и формализация априорной (известной до начала моделирования) информации.

---

<sup>1</sup>ПГТУ, канд. техн. наук, доц.

<sup>2</sup>ПГТУ, ст. препод.

<sup>3</sup>ПГТУ, соискатель

<sup>4</sup>ПГТУ, спец. лабор. сетевых технологий

### 3) Этап параметризации.

Осуществляется непосредственно моделирование, т.е. выбор общего вида модели, выявление входящих в нее связей. Основная задача, решаемая на этом этапе, – выбор вида функции в эконометрической модели, в частности, возможность использования линейной модели как наиболее простой и надежной. Весьма важной проблемой на этом (и предыдущих) этапе эконометрического моделирования является проблема спецификации модели, в частности: выражение в математической форме обнаруженных связей и соотношений; установление состава экзогенных и эндогенных переменных, в том числе лаговых; формулировка исходных предпосылок и ограничений модели. От того, насколько удачно решена проблема спецификации модели, в значительной степени зависит успех всего эконометрического моделирования.

### 4) Информационный.

Осуществляется сбор необходимой статистической информации – наблюдаемых значений экономических переменных. Здесь могут быть наблюдения, полученные как с участием исследователя, так и без его участия (в условиях активного или пассивного эксперимента).

### 5) Этап идентификации.

Осуществляется статистический анализ модели и оценка ее параметров.

### 6) Этап верификации модели.

Проводится проверка истинности, адекватности модели. Выясняется, насколько удачно решены проблемы спецификации, идентификации и идентифицируемости модели, какова точность расчетов по данной модели, в конечном счете, насколько соответствует построенная модель моделируемому реальному экономическому объекту или процессу.

Некоторые организации используют широкомасштабные эконометрические модели, чтобы на основании прогнозов таких факторов сделать прогнозы относительно развития системы в будущем. Разработчики таких широкомасштабных моделей обычно предусматривают несколько «стандартных» прогнозов, основанных на определенном наборе экзогенных переменных. Некоторые модели содержат вероятность, с которой может осуществляться тот или иной прогноз. В других случаях пользователи могут включать сделанные ими самими предположения и анализировать полученные в результате этих предположений прогнозы [2 – 4]. Эконометрические модели такого типа насчитывают большое число уравнений, которые описывают большое число важных взаимосвязей. Несмотря на то, что оценки таких взаимосвязей основаны на данных за прошедший период, эти оценки могут позволить (или не позволить) модели эффективно работать в будущем. Когда прогнозы оказываются неудачными, это значит, что лежащая в основе модели экономическая взаимосвязь претерпела структурные изменения. То есть, возникает ситуация, когда имеющихся данных может быть недостаточно для того, чтобы определить функциональную связь между переменными, либо они недостаточно варьируются, чтобы можно было отличить влияние одного фактора от влияния другого. В связи с этим возникает необходимость анализа системы на наличие мультиколлинеарности между различными факторами. На практике можно столкнуться с очень высокой (или близкой к ней) мультиколлинеарностью – сильной корреляционной зависимостью между объясняющими переменными. Оценки становятся очень чувствительными к незначительному изменению результатов наблюдений и объема выборки. Уравнения регрессии в этом случае, как правило, не имеют реального смысла, так как некоторые из его коэффициентов могут иметь неправильные с точки зрения экономической теории знаки и неоправданно большие значения.

Как известно, при выполнении определенных предпосылок метод наименьших квадратов [1] дает наилучшие линейные несмещенные оценки. Причем свойство несмещенности и эффективности оценок остается в силе даже, если несколько коэффициентов регрессии оказываются статистически незначимыми. Однако несмещенность фактически означает лишь то, что при многократном повторении наблюдений (при постоянных объемах выборок) за исследуемыми величинами средние значения оценок стремятся к их истинным значениям. Наименьшая возможная дисперсия вовсе не означает, что дисперсия оценок будет мала по сравнению с самими оценками. В ряде случаев такая дисперсия достаточно велика, чтобы оценки коэффициентов стали статистически незначимыми.

Обычно выделяются следующие последствия мультиколлинеарности:

1) Большие дисперсии (стандартные ошибки) оценок. Это затрудняет нахождение истинных значений определяемых величин и расширяет интервальные оценки, ухудшая их точность.

2) Уменьшаются значения коэффициентов, что может привести к неоправданному выводу о существенности влияния соответствующей объясняющей переменной на зависимую переменную.

3) Оценки коэффициентов по методу наименьших квадратов и их стандартные ошибки становятся очень чувствительными к малейшим изменениям данных, т. е. они становятся неустойчивыми.

4) Затрудняется определение вклада каждой из переменных в объясняемую уравнением регрессии дисперсию зависимой переменной.

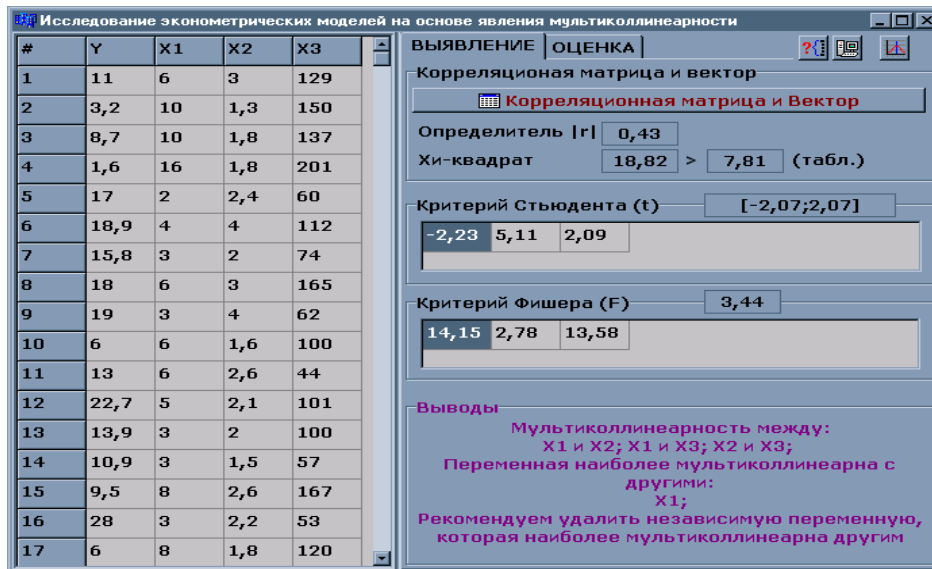
5) Возможно получение неверного знака у коэффициента регрессии.

Точных количественных критериев для определения наличия или отсутствия мультиколлинеарности не существует. Один из подходов заключается в анализе корреляционной матрицы между объясняющими переменными  $X_1, X_2, \dots, X_p$  и выявлении пар переменных, имеющих высокие коэффициенты корреляции (обычно больше 0,8). Если такие переменные существуют, то говорят о мультиколлинеарности между ними. Полезно также находить множественные коэффициенты детерминации между одной из объясняющих переменных и некоторой группой из них. Наличие высокого множественного коэффициента детерминации (обычно больше 0,6) свидетельствует о мультиколлинеарности.

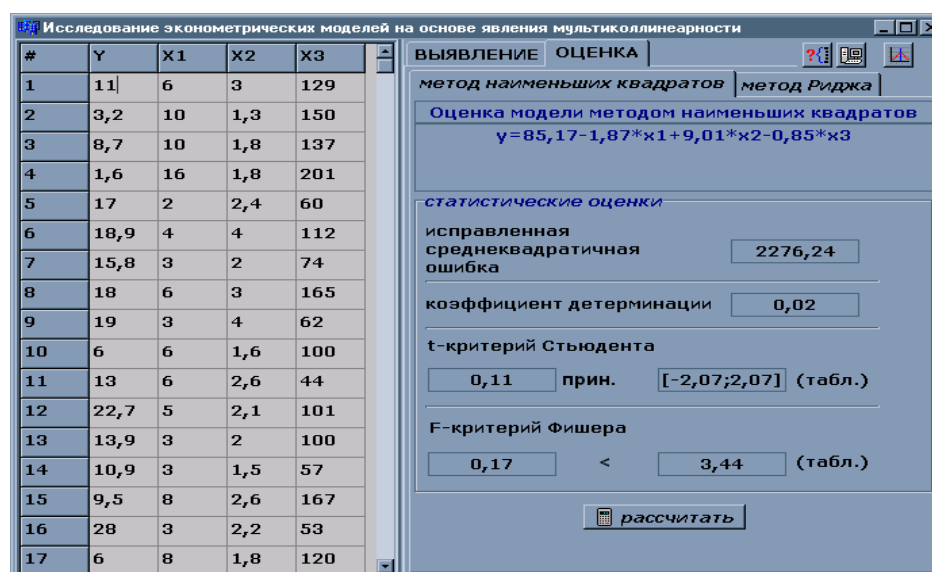
Исходя из вышесказанного возникает необходимость создания специализированного программного обеспечения для проведения анализа и оценки параметров эконометрической модели, которое позволит автоматизировать расчеты необходимых величин, повысит их точность и поможет провести многосторонний анализ создаваемой модели на предмет ее адекватности. В качестве математической модели выбран комплексный метод исследования модели на наличие мультиколлинеарности по алгоритму Фаррара-Глобера [1].

На основе математических моделей разработано программное обеспечение, позволяющее проводить анализ эконометрических систем на основе явления мультиколлинеарности, оценивать степень корреляции данных, устранять мультиколлинеарность, а также проводить оценку модели методом наименьших квадратов и методом гребневой регрессии. Для реализации был выбран язык программирования C++ в среде быстрой разработки (RAD) C++ Builder 6.0.

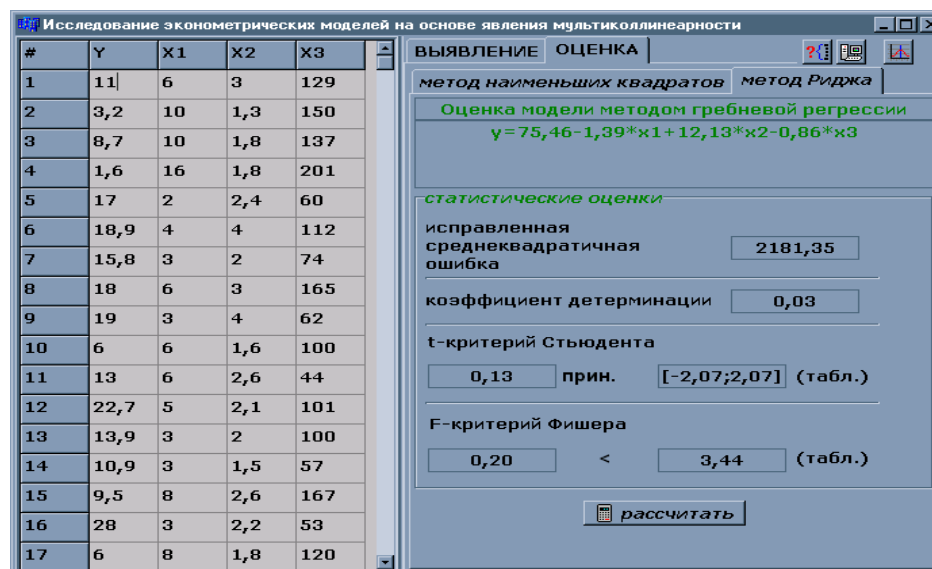
Для начала работы необходимо загрузить в программу данные, представленные выборками зависимой и независимых переменных. Существует возможность ручного поэлементного ввода и редактирования элементов выборок либо загрузка их из подготовленных заранее текстовых файлов данных. Если исходные данные были изменены, то имеется возможность сохранить их в отдельные файлы. Далее происходит выявление мультиколлинеарности (рис. 1а). В зависимости от результатов работы программы на стадии выявления мультиколлинеарности, исследователь может оценить модель методом наименьших квадратов (при отсутствии мультиколлинеарности) либо оценить модель методом гребневой регрессии (при наличии мультиколлинеарности). При оценке модели методом наименьших квадратов на вкладке отображаются полученное уравнение регрессии и статистические оценки (рис. 1б). Если статистические оценки метода наименьших квадратов не удовлетворяют исследователя, то в соответствии с результатами, представленными программой при анализе данных на мультиколлинеарность, можно вывести из модели наиболее коррелирующую переменную либо увеличить число элементов выборки. Если принято решение оставить модель без изменений, то программа предоставляет возможность перейти к смещенным оценкам, то есть оценить модель методом гребневой регрессии («ридж» регрессии). При оценке модели методом гребневой регрессии на вкладке будут представлены результаты оценивания и полученной уравнение регрессии (рис. 1в). Наглядные результаты работы программы представлены на графиках (рис. 2), где отображены текущие данные выборки, а также уравнения регрессий, полученных при оценке модели методами наименьших квадратов и гребневой регрессии. При этом существует возможность просмотра графиков уравнения регрессий, как для каждого метода в отдельности, так и сводного графика для методов несмещенных и смещенных оценок.



а)



б)



в)

Рис. 1 – Результаты моделирования: а) выявление мультиколлинеарности, б) оценка методом наименьших квадратов, в) оценка методом гребневой регрессии.

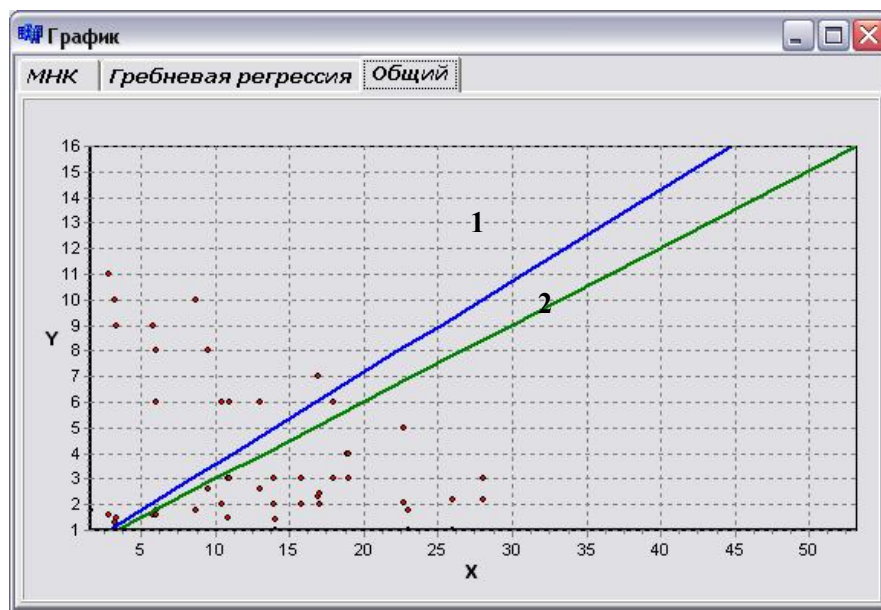


Рис. 2 – Сводный график уравнений линейной регрессии, полученный методом наименьших квадратов (прямая 1) и гребневой регрессии (прямая 2).

В результате компьютерного моделирования было выявлено, что результаты исследований, полученные различными методами, практически совпадают. Ошибка находится в пределах нормы (до 5 – 10 %), тогда как при расчете без учета мультиколлинеарности ошибки были достаточно существенны. Анализ полученных результатов позволил сделать вывод о правильности работы системы и адекватности полученной математической модели.

В дальнейшем возможно более детальное исследование эконометрических систем, автоматический подбор и корректировка параметров модели при выявлении между факторами мультиколлинеарности.

#### Выводы

1. Построены математические модели и разработано программное обеспечение, позволяющее выявлять и устранять мультиколлинеарность, что дает возможность всесторонне проводить анализ результатов в зависимости от различных факторов и правильно прогнозировать характер изменения показателей системы.
2. Использование данных моделей дает возможность более корректно обрабатывать исходную информацию, обоснованно принимать оптимальные решения.

#### Перечень ссылок

1. *Дугерти К.* Введение в эконометрику: Учеб. для вузов / *К. Дугерти.* – М.: ИНФРА-М, 1999. – 401 с.
2. *Бабешко Л.О.* Основы эконометрического моделирования / *О.Л. Бабешко.* – К.: КомКнига, 2006. – 432 с.
3. *Айвазян С.А.* Прикладная статистика и основы эконометрики: Учеб. для вузов / *С.А. Айвазян, В.С. Мхитарян.* – М.: ЮНИТИ, 1998. – 1022 с.
4. *Лаврентик О.А.* Построение математических моделей эконометрических систем методами прогнозирования / *О.А. Лаврентик, О.П. Павлова, О.В. Усольцев* // *Вісник Приазов. держ. техн. ун-ту: Зб. наук.пр.* – Маріуполь, 2005. – Вип. 15. – С. 254 – 257.

Рецензент: Ю.Е. Коляда  
д-р физ-мат. наук, проф., ПГТУ

Статья поступила 20.03.2008