Получение нелинейных уравнений регрессии с помощью параболической оптимизации

к.э.н. Ивашнев Л.И.
Университет машиностроения
8(985) 284-26-98

Аннотация. В статье предложен новый приближенный метод получения нелинейных уравнений регрессии – параболическая оптимизация. Дан пример его применения для получения нелинейного уравнения, которое известными методами получить невозможно. В качестве статистического критерия оценки качества полученного уравнения используется среднеквадратичное отклонение.

Ключевые слова: *метод получения, нелинейное уравнение регрессии, параболическая оптимизация, статистический критерий, оценка качества уравнения регрессии, среднеквадратичное отклонение.*

1. Нелинеаризуемые уравнения регрессии

Если задана некоторая выборка:

$$y_1\ x_{11}\ x_{21} \ldots x_{m1},$$
$$y_2\ x_{121}\ x_{22} \ldots x_{m21},$$
$$\ldots\ldots\ldots\ldots\ldots\ldots\ldots$$
$$y_n\ x_{1n}\ x_{2n} \ldots x_{mn}$$

и требуется получить приближенную зависимость между показателем Y и факторами X_1, X_2, … X_m

$$Y = f(X_1, X_2, \ldots X_m), \text{ или } F(X_1, X_2, \ldots X_m) = 1,$$

то для получения такой зависимости можно воспользоваться методом наименьших квадратов [1, 2], или методом получения вырожденных уравнений регрессии [4], или методом получения взвешенных уравнений регрессии [5], или методом получения уравнений регрессии общего вида [6], или «методом трех точек» [3].

Однако, ни одним из этих методов нельзя получить сравнительно простое уравнение регрессии вида:

$$Y = (a_0 + a_1 X)^{a_2}. \qquad (1)$$

Более того, можно указать бесчисленное множество вариантов нелинейных уравнений регрессии, которые получить указанными методами нельзя.

Действительно, для линеаризации уравнения (1) из обеих его частей можно извлечь корень степени $1/a_2$. Тогда это уравнение примет вид:

$$Y^{1/a_2} = a_0 + a_1 X. \qquad (2)$$

Если теперь выполнить замену $V = Y^{1/a_2}$, то уравнение (2) примет вид:

$$V = a_0 + a_1 X. \qquad (3)$$

Отмечаем, что уравнение (3) является линейным уравнением, однако дополнительный показатель $V = Y^{1/a_2}$ не является простым слагаемым, поскольку значения элементов столбца V расширенной выборки зависят от неизвестного коэффициента a_2. Следовательно, применить только метод наименьших квадратов для расчета коэффициентов уравнения (1) нельзя.

В [4] отмечено, что некоторые разновидности нелинейных уравнений регрессии не могут быть линеаризованы и для получения таких уравнений **необходимо** использование оптимизации. Применение оптимизации для получения нелинейных уравнений регрессии достаточно подробно изложено в [4], причем предлагается использовать «метод покоординатного спуска» и «метод луча».

Оптимизация, т.е. поиск точки минимума или максимума, может выполняться с помощью «метода случайного поиска», «метода покоординатного спуска» или «метода луча» по схеме «с оракулом», которые подробно изложены в [4]. Однако эти методы характеризуются слишком большим временем использования ЭВМ, причем дают приближенное решение.

Значительно более эффективным является метод параболической оптимизации, для реализации которого может быть использована система EXCEL и который будет представлен в последующем изложении.

2. Уравнение параболы, проведенной через 3 точки

Через 3 точки можно провести параболу, уравнение которой определяется следующими зависимостями:

$$\begin{cases} y_1 = a_1 x_1^{\ 2} + a_2 x_1 + a_3, \\ y_2 = a_1 x_2^{\ 2} + a_2 x_2 + a_3, \\ y_3 = a_1 x_3^{\ 2} + a_2 x_3 + a_3, \end{cases} \qquad (4)$$

где: $(x_1; y_1), (x_2; y_2), (x_3; y_3)$ – координаты точек параболы.

Если из 2-го уравнения вычесть 1-е, а из 3-го – второе, то получим систему уравнений:

$$\begin{cases} y_2 - y_1 = a_1(x_2^{\ 2} - x_1^{\ 2}) + a_2(x_2 - x_1), \\ y_3 - y_2 = a_1(x_3^{\ 2} - x_2^{\ 2}) + a_2(x_3 - x_2). \end{cases} \qquad (5)$$

Далее обе части каждого уравнения разделим на коэффициент при a_2. Получим систему выражений:

$$\begin{cases} \dfrac{y_2 - y_1}{x_2 - x_1} = a_1 \dfrac{x_2^{\ 2} - x_1^{\ 2}}{x_2 - x_1} + a_2, \\ \dfrac{y_3 - y_2}{x_3 - x_2} = a_1 \dfrac{x_3^{\ 2} - x_2^{\ 2}}{x_3 - x_2} + a_2, \end{cases}$$

откуда, пользуясь формулой $x_2^{\ 2} - x_1^{\ 2} = (x_2 - x_1)(x_2 + x_1)$, получаем систему:

$$\begin{cases} \dfrac{y_2 - y_1}{x_2 - x_1} = a_1(x_2 + x_1) + a_2, \\ \dfrac{y_3 - y_2}{x_3 - x_2} = a_1(x_3 + x_2) + a_2. \end{cases} \qquad (6)$$

Вычитая из 2-го уравнения 1-е получаем уравнение:

$$\frac{y_3 - y_2}{x_3 - x_2} - \frac{y_2 - y_1}{x_2 - x_1} = a_1(x_3 - x_1),$$

откуда:

$$a_1 = \frac{\dfrac{y_3 - y_2}{x_3 - x_2} - \dfrac{y_2 - y_1}{x_2 - x_1}}{x_3 - x_1}. \qquad (7)$$

Из первого уравнения системы (6) получаем:

$$a_2 = \frac{y_2 - y_1}{x_2 - x_1} - a_1(x_2 + x_1),$$

а из 1-го уравнения системы (4):

$$a_3 = y_1 - a_1 x_1^{\ 2} - a_2 x_1.$$

Итак, уравнение параболы $y = a_1 x^2 + a_2 x + a_3$, проходящей через три заданные точки, имеет следующие коэффициенты:

$$\begin{cases} a_1 = \dfrac{\dfrac{y_3 - y_2}{x_3 - x_2} - \dfrac{y_2 - y_1}{x_2 - x_1}}{x_3 - x_1}, \\ a_2 = \dfrac{y_2 - y_1}{x_2 - x_1} - a_1(x_2 + x_1), \\ a_3 = y_1 - a_1 x_1^2 - a_2 x_1. \end{cases} \quad (8)$$

Отмечаем, что если $a_1 < 0$ (рисунок 1*а*), то ветви параболы направлены вниз, если же $a_1 > 0$ (рисунок 1*б*), то ветви параболы направлены вверх.

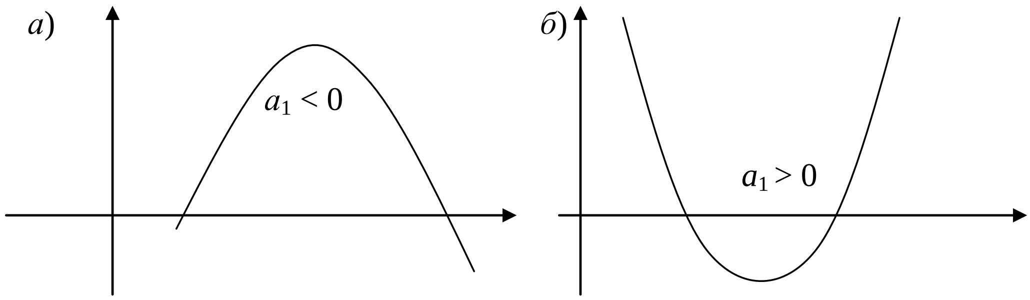

Рисунок 1. Вид параболы в зависимости от знака старшего члена

Поэтому, если $a_1 > 0$, то точкой минимума параболы является ее вершина, координата x которой вычисляется по формуле:

$$x = -\frac{a_2}{2a_1}. \quad (9)$$

Если же $a_1 < 0$, то точкой минимума является один из концов ветвей параболы (на рисунке 1*а* это конец правой ветви).

3. Алгоритм получения уравнений регрессии с помощью параболической оптимизации

Для получения нелинейного уравнения регрессии $y = f(x_1, x_2, \ldots, x_k)$ или $F(x_1, x_2, \ldots, x_k) = 1$ любым из рассмотренных методов определяется точка минимума для суммы квадратов отклонений $Q = \sum_{i=1}^{n} \left[y - f(x_1, x_2, \ldots, x_k) \right]^2$ или $Q = \sum_{i=1}^{n} \left[F(x_1, x_2, \ldots, x_k) - 1 \right]^2$ точек выбранного уравнения регрессии от точек исходной выборки. При этом может использоваться либо метод наименьших квадратов с заменой факторов, либо метод оптимизации, либо совместное использование обоих этих методов для получения оптимального результата.

Если уравнение регрессии получено, например, методом наименьших квадратов и ему соответствует минимальная сумма квадратов отклонений, то это уравнение может считаться наилучшим решением по заданной выборке и его дальнейшего уточнения не требуется. Если выполняется пошаговый процесс оптимизации совместно с заменой факторов и методом наименьших квадратов, то процесс поиска уравнения регрессии ориентировочно может представляться следующим алгоритмом:

1) выбрать вид искомого уравнения регрессии;
2) линеаризовать это уравнение;
3) вычислить значения дополнительных факторов и подставить их в исходную выборку, т.е. сформировать расширенную выборку;
4) определить слагаемые, не являющиеся простыми слагаемыми, т.е. содержащие оптими-

зируемые факторы;

5) выбрать исходные (опорные) значения для оптимизируемых факторов и задать шаг перемещения для каждого из них. Выбрать текущий фактор, т.е. фактор с которого будет начата оптимизация;
6) вычислить коэффициенты уравнения регрессии по текущим значениям дополнительных и оптимизируемых факторов и вычислить соответствующее значение СКО;
7) задать приращение оптимизируемого фактора и рассчитать координаты точек, лежащих по обе стороны от текущей точки. Для этих точек вычислить значения СКО;
8) получить координату точки минимума, т. е. точки, которой соответствует минимальное значение СКО. Зафиксировать точку минимума в качестве текущей точки;
9) вычислить коэффициенты уравнения регрессии по приращенным значениям дополнительных и оптимизируемых факторов и вычислить соответствующее значение СКО;
10) если перебор оптимизируемых факторов закончен, то проверить было ли улучшение СКО на цикле перебора. Если улучшения не было, то перейти к п. 13;
11) если СКО не уменьшилось, то поменять знак приращения и перейти к п. 6;
12) выбрать следующий оптимизируемый фактор и перейти к п. 7;
13) вычислить коэффициенты уравнения регрессии по приращенным значениям дополнительных и оптимизируемых факторов и вычислить соответствующее значение СКО;
14) печать наилучшего уравнения регрессии.

Выполняя предложенный алгоритм можно получить уравнение регрессии даже в том случае, если его линеаризация окажется невозможной, т.е. если при линеаризации получена сумма слагаемых, причем некоторые из них содержат неизвестные коэффициенты, т.е. не являются простыми слагаемыми.

4. Процесс получения уравнения регрессии с помощью параболической оптимизации

Теперь можно выполнить в EXCEL процесс получения уравнения регрессии с помощью параболической оптимизации по данным следующей выборки:

Таблица 1

Выборка для параболической оптимизации

x	*Y*	*x*	*Y*	*x*	*Y*
0,1	0,3125	1,3	7,595	2,5	27,565
0,2	0,7021	1,4	7,959	2,6	27,573
0,3	0,9654	1,5	10,375	2,7	39,438
0,4	1,062	1,6	12,376	2,8	37,863
0,5	1,7241	1,7	10,794	2,9	35,599
0,6	2,4014	1,8	12,772	3,0	48,534
0,7	2,3865	1,9	19,197	3,1	43,46
0,8	3,4591	2,0	17,707	3,2	52,656
0,9	3,7008	2,1	16,666	3,3	59,051
1,0	4,4696	2,2	17,991	3,4	58,165
1,1	5,975	2,3	24,05	3,5	51,191
1,2	5,4618	2,4	27,58	3,6	58,333

Эта выборка сгенерирована в системе EXCEL и предназначена для демонстрации процесса расчета коэффициентов уравнения регрессии, которое невозможно получить любым из известных к настоящему времени методов. Значения всех параметров перед началом процесса оптимизации показаны на рисунке 2, где в столбце A показаны значения фактора X, а в столбце D – значения показателя Y, причем в правой части рисунка 2 внизу показан график рассматриваемой кривой, построенный по опорному значению $a_2 = 2$ (кривая напоминает параболу), и точки исходной выборки, соединенные отрезками прямых. Из общего вида этого графика можно сделать вывод, что эта зависимость напоминает квадратичную параболу и описывается уравнением вида (1), для линеаризации которого из обеих частей уравнения (1) можно извлечь корень степени $1/a_2$. Тогда это уравнение примет вид: $Y^{1/a_2} = a_0 + a_1 X$.

Заменяя $V = Y^{1/a_2}$ получаем линейное уравнение $V = a_0 + a_1 X$, где V – дополнительный показатель, значения которого при изменении значения a_2 меняются. В столбце F получены опорные значения дополнительного показателя V, полученные при опорном значении $a_2 = 2$. Наличие столбцов X и V позволяет получить коэффициенты линейного уравнение регрессии $V = a_0 + a_1 X$ решая систему нормальных уравнений:

$$\begin{cases} na_0 + a_1 \sum_i x_i = \sum_i v_i \\ a_0 \sum_i x_i + a_1 \sum_i x_i x_i = \sum_i x_i v_i \end{cases}. \qquad (10)$$

На рисунке 2 эта система представлена в явном виде в диапазоне $J_1{:}N_2$, ниже показано ее решение.

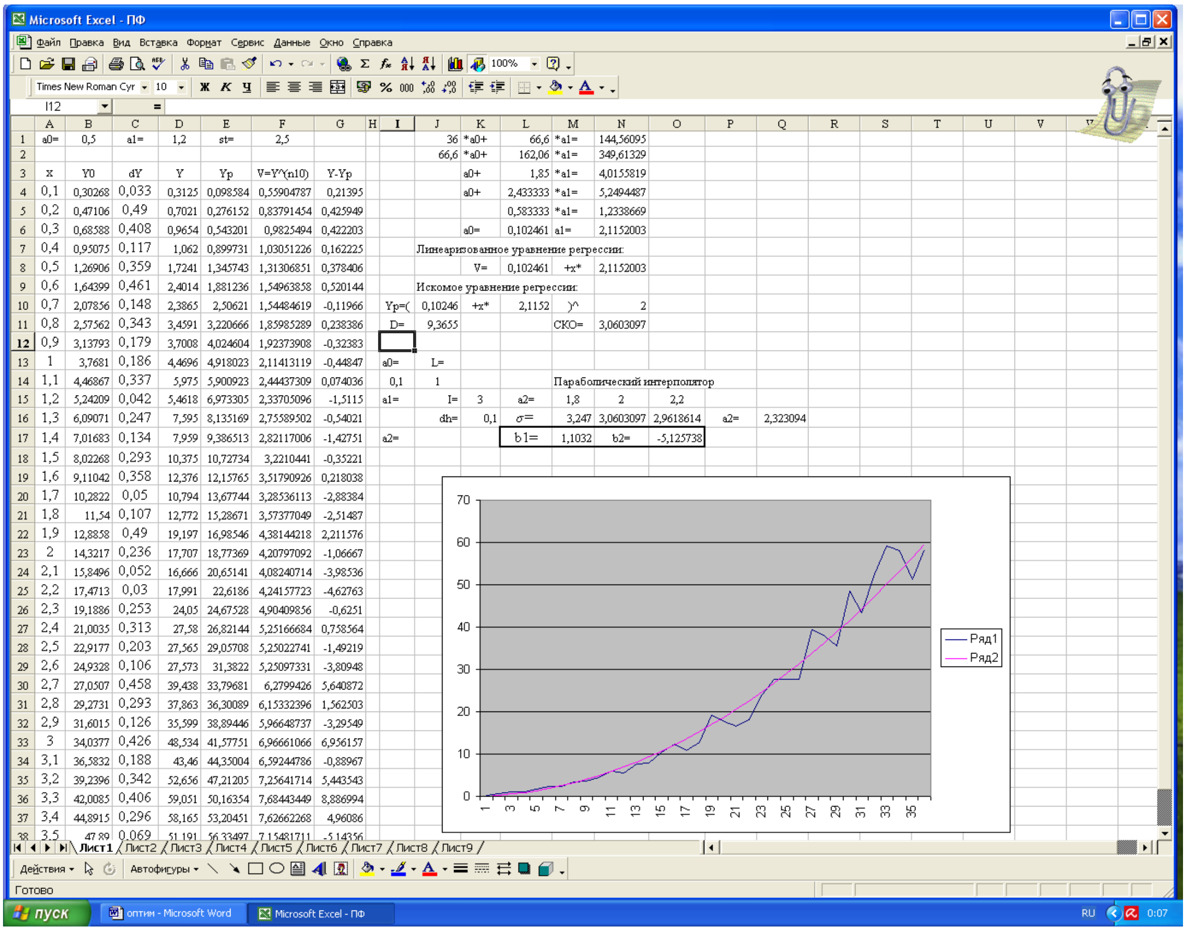

Рисунок 2. Окно монитора перед оптимизацией

Решая систему (10) получены следующие опорные значения: $a_0 = 0{,}102461$, $a_1 = 2{,}1152$ и использовано $a_2 = 2$, т.е. получено опорное уравнение регрессии:

$$Y = (0{,}102461 + 2{,}1152 \cdot X)^2 .$$

По формуле (1) получены расчетные значения $Y_р$, которые показаны в столбце E. Отклонения $Y - Y_р$ показаны в столбце G. Затем по формуле:

$$\sigma = \sqrt{\frac{\sum_i (y_i - y_{pi})^2}{n}} \qquad (11)$$

вычислено среднеквадратичное отклонение точек исходной выборки от поученного уравнения регрессии $\sigma = 2{,}9618614$, которое показано в ячейке $N11$.

Отмечаем, что, судя по рисунку 2, график полученной кривой очень неплохо описывает исходную выборку. Тем не менее, выполним оптимизацию полученного уравнения регрессии. Для этого в ячейку $J14$ помещаем длину интервала интерполяции $L = 1$ и, пользуясь позицией «значение» в «специальной вставке», в ячейку $N15$ записываем значение $a_2 = 2$, тогда в ячейках $M15$ и $O15$ появляются абсциссы x двух дополнительных точек, отстоящих на $L/2$ от середины интервала интерполяции и необходимых для построения параболы и получения координаты $x_{мин}$ вершины или нижней точки ветвей параболы. Эта координата появляется в ячейке $Q16$.

Расчет координаты $x_{мин}$ выполняется параболическим интерполятором, помещенным в ячейки $M17$, $O17$ и $Q16$, причем в ячейке $M17$ помещен оператор b1: =((O16-N16)/(O15-N15)-(N16-M16)/(N15-M15))/(O15-M15), в ячейке $O17$ – b2: =(N16-M16)/(N15-M15)-M17*(N15+M15) и в ячейке $Q16$ – оператор: =ЕСЛИ(M17>0;-O17/(2*M17);ЕСЛИ(O16>M16;M16;O16)).

При этом, в ячейке M17 рассчитывается коэффициент a_1 для системы (8) (он обозначен как b_1), в ячейке O17 рассчитывается a_2 (обозначен как b_2) и в ячейке Q16 вычисляется $x_{мин}$ по формуле (9) $x_{мин} = -\dfrac{a_2}{2a_1}$.

Отмечаем, что на первом цикле оптимизации получено значение $x_{мин} = 2{,}323094$. Это значение подставляем в $N10$ и получаем новое значение СКО $\sigma = 2{,}9389408$. Полученное значение СКО на 0,23 меньше ранее полученного значения.

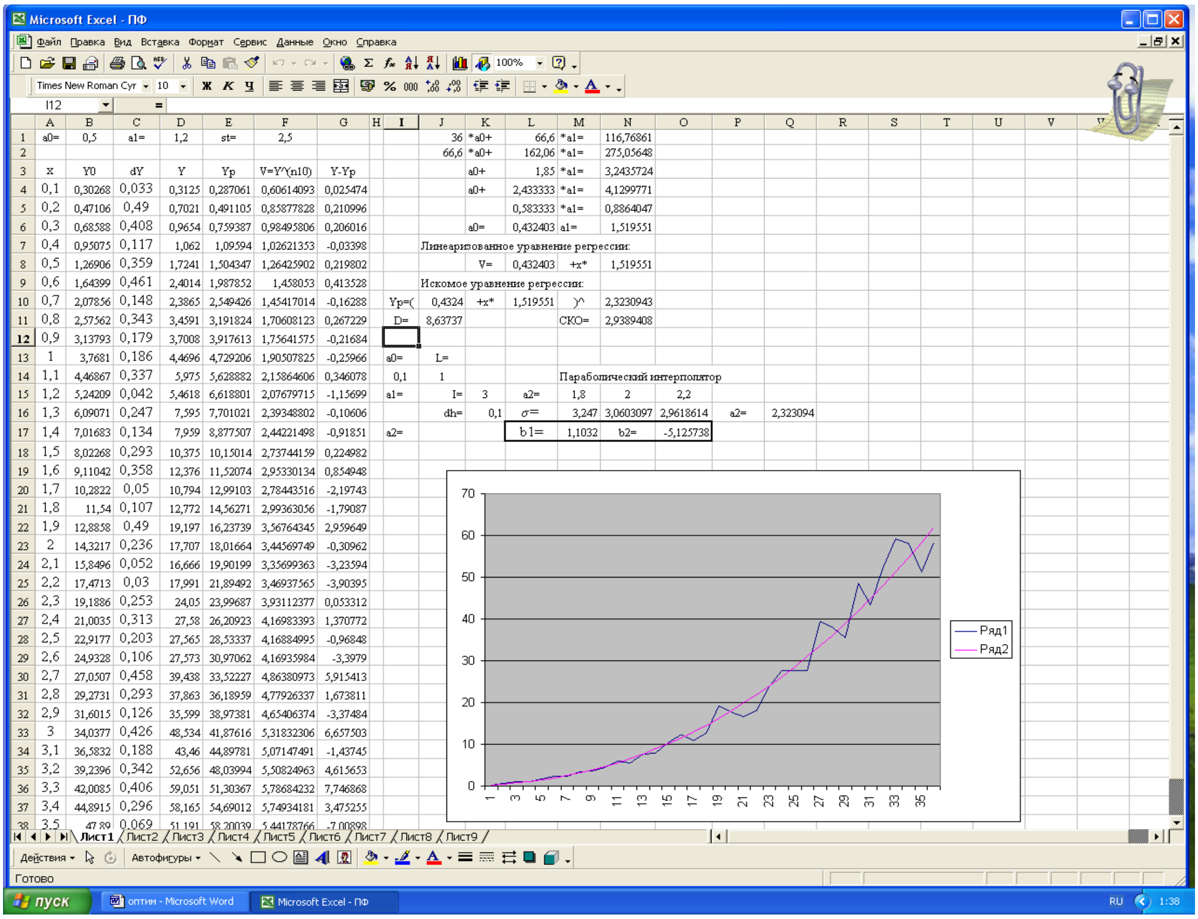

Рисунок 3. Результат оптимизации уравнения регрессии

Далее можно выполнить оптимизацию и по другим коэффициентам a_0 и a_1. Можно также последовательно уменьшать длину интервала интерполяции в 2 или в 3 раза. При этом величина СКО еще несколько уменьшится. Однако принципиального значения это уже не имеет, поскольку, судя по рисунку 3, полученный результат оптимизации вполне соответствует погрешности метода.

Итак, в результате выполненной оптимизации получено нелинейное уравнение регрессии:

$$Y = \left(0{,}4324 + 1{,}5196 \cdot x\right)^{2{,}3231}.$$

Этому уравнению соответствует СКО σ = 2,939. Сформулированная задача решена, поскольку мы получили уравнение регрессии, которое известными методами получить невозможно.

В заключение отмечаем, что как метод наименьших квадратов, так и методы вырожденной и взвешенной регрессии дают уравнение регрессии, минимизирующее среднеквадратичное отклонение значений показателя *Y* от точек полученного уравнения регрессии. Это означает, что использование дополнительного показателя *V*, который используется вместо показателя *Y*, может являться причиной возникновения дополнительных погрешностей. В этом случае оптимизация может служить для проверки и уточнения зависимости, полученной в результате применения методов регрессии.

Литература

1. Айвазян С.А., Мхитарян В.С. Прикладная статистика и основы эконометрики. Учебник для ВУЗов. -М.:ЮНИТИ, 1998. – 1022 с.
2. Кремер Н.Ш., Путко Б.А. Эконометрика: Учебник для вузов / Под ред. проф. Н.Ш. Кремера. – М.: ЮНИТИ-ДАНА, 2002. – 311 с.
3. Кремер Н.Ш. Теория вероятностей и математическая статистика: Учебник для вузов. – М.: ЮНИТИ-ДАНА, 2003. – 543 с.
4. Эконометрика: Учебник / И.И. Елисеева, С.В. Курышева, Т.В. Костеева и др.: Под ред. И.И. Елисеевой. – 2-е изд., перераб. и доп. – М.: Финансы и статистика, 2006. – 576 с.: ил.
5. Ивашнев Л.И. Методы и модели в экономике: Учеб. пособие. – М.: Издательский дом «Лидер-М», 2011. – 328 с.
6. Ивашнев Л.И. Методы регрессии в экономической математике: Монография. – М.: Изд-во МГОУ, 2005.