Первадчук В.П., Масенко И.Б.
Пермский государственный технический университет

МАТЕМАТИЧЕСКАЯ МОДЕЛЬ ПРОГНОЗИРОВАНИЯ ФИНАНСОВОГО СОСТОЯНИЯ ПРЕДПРИЯТИЯ

В данной статье описано построение линейной множественной регрессионной модели прогнозирования финансового состояния предприятия путем расчета будущей величины его чистых активов на основе реальных данных финансовой отчетности организаций Пермского края.

Введение

Основными задачами экономико-математического моделирования являются построение модели, определение ее параметров и применение для решения современных проблем. При этом точность и обоснованность анализа, прогнозирования и, соответственно, планирования и управления зависят от того, насколько в разработанных моделях отражены реальные процессы и связи между показателями развития экономических объектов, ограничения, накладываемые на развитие системы объектов, достоверна информация, используемая при моделировании.

Поэтому в настоящее время при разработке методов моделирования экономических объектов все больше внимания уделяется адекватности структуры моделей реальным процессам [6].

Одна из наиболее сложных проблем современной экономики – предсказание финансового состояния предприятий, а особенно их неплатежеспособности. Среди различных способов решения этой задачи наиболее эффективным методом, на наш взгляд, является математическое моделирование величины чистых активов предприятия. Причем в основе множественной линейной регрессионной модели должны присутствовать реальные статистические данные российских предприятий. В предлагаемой модели использованы сведения о балансах 61 организации по состоянию на 1.10.2004 г. и на 1.10.2005 г.

В работе рассматривались линейные по параметрам и по переменным множественные регрессионные модели:

$$y_i = b_0 + b_1 x_{i1} + b_2 x_{i2} + \ldots b_k x_{ik} + \varepsilon_i, \qquad i = 1 \ldots n. \quad (1)$$

где $b_0, b_1, \ldots, b_k$ – неизвестные параметры модели, а ε_i – случайные ошибки модели.

Класс линейных множественных регрессионных моделей выбран не случайно. Коэффициенты, полученные в результате использования именно данного типа моделей, легко поддаются экономической интерпретации, чего нельзя сказать о других классах моделей.

Основные результаты

По данным балансов были рассчитаны 29 показателей, характеризующих финансовое состояние предприятий. Используя показатели деятельности организаций за 2004 год, была построена модель для прогнозирования величины чистых активов на последующий год, в которой зависимой переменной Y выступает величина чистых активов на 1.10.2005 г. (NET_ASSETS) как основной показатель платежеспособности организации. В качестве независимых переменных $X_1, \ldots, X_k$ выступают следующие финансовые показатели балансов организаций на 1.10.2004 г.:

К1 – отношение оборотного капитала к общим активам.
К2 – отношение нераспределенной прибыли к общим активам.
К3 – отношение прибыли до выплаты процентов и налогов к общим активам.
К4 – отношение рыночной стоимости акций к стоимости балансовых обязательств.
К5 – отношение выручки от реализации к общим активам.
К6 – отношение нераспределенной прибыли прошлых лет к общим активам.
К7 – рентабельность вложений в предприятие.
К8 – отношение денежного потока к обязательствам.
К9 – отношение оборотного капитала к обязательствам.
К10 – коэффициент абсолютной ликвидности.
К11 – промежуточный коэффициент покрытия.
К12 – коэффициент текущей ликвидности.

Таблица 1.

	K2	K8	K13	K14	K17	K24	K26	NET_ASSETS
Выборочное среднее	0,103	0,068	0,475	0,053	0,244	0,973	0,035	65497,670
Медиана	0,063	0,021	0,161	0,030	0,171	0,979	0,016	21944,000
Максимум	0,644	0,845	3,214	0,429	0,794	2,111	0,287	597558,000
Минимум	0,000	-0,118	-0,047	-0,098	0,002	0,166	-0,042	-436,000
Стандартное отклонение	0,135	0,141	0,662	0,079	0,236	0,357	0,055	119408,000
Выборочная дисперсия	0,018	0,020	0,438	0,006	0,056	0,128	0,003	14258270464,000
Коэффициент асимметрии	2,349	3,450	1,872	2,298	0,768	0,228	2,129	2,794
Коэффициент эксцесса	9,280	17,485	6,609	10,671	2,245	4,509	9,030	10,900

K13 – коэффициент соотношения собственных и заемных средств.

K14 – рентабельность продаж.

K15 – коэффициент рентабельности оборотных активов.

K16 – коэффициент рентабельности внеоборотных активов.

K17 – коэффициент наличия собственных средств.

K18 – рентабельность деятельности предприятия.

K19 – доля денежных средств в выручке.

K20 – отношение прибыли до налогообложения к собственному капиталу.

K21 – отношение (прибыль до налогообложения + проценты к уплате) к общим активам.

K22 – отношение прибыли до налогообложения к краткосрочным обязательствам.

K23 – отношение прибыли от реализации к краткосрочным обязательствам.

K24 – отношение оборотных активов к сумме обязательств.

K25 – отношение чистого собственного капитала к сумме задолженности.

K26 – отношение чистой прибыли к совокупным активам.

K27 – величина чистых активов.

K28 – отношение общих активов к краткосрочным обязательствам.

K29 – отношение совокупных активов к совокупному долгу.

Отбор переменных, вошедших в итоговую модель, производился на основе метода пошаговой регрессии. Критериями выбора наилучшей регрессионной модели являлись значение скорректированного коэффициента детерминации R^2_{adj}, значимость всех коэффициентов модели по критерию Стьюдента, значимость модели на основе критерия Фишера, а также значение коэффициента корреляции между фактическим значением зависимой переменной и расчетным, найденным по модели.

В итоговую же модель вошли следующие независимые переменные:

K2 – отношение нераспределенной прибыли к общим активам;

K8 – отношение денежного потока к обязательствам;

K13 – коэффициент соотношения собственных и заемных средств;

K14 – рентабельность продаж;

K17 – коэффициент наличия собственных средств;

K24 – отношение оборотных активов к сумме обязательств;

K26 – отношение чистой прибыли к совокупным активам.

Определив входящие в исходную базу данных переменные, вычислили описательные статистики по каждой из них (табл. 1).

Положительные значения коэффициента асимметрии показывают, что распределения имеют правостороннюю асимметрию по сравнению с нормальным распределением, то есть значения показателей, находящиеся справа от среднего значения ряда, имеют большую частоту. Положительные значения коэффициента эксцесса показывают, что распределения имеют более крутую вершину по сравнению с кривой нормального распределения.

Таблица 2.

	K2	K8	K13	K14	K17	K24	K26	NET_ASSETS
K2	1,000	0,741	0,502	0,338	0,517	0,429	0,769	-0,082
K8	0,741	1,000	0,618	0,395	0,501	0,346	0,926	-0,087
K13	0,502	0,618	1,000	0,232	0,886	0,538	0,458	0,145
K14	0,338	0,395	0,232	1,000	0,221	-0,073	0,452	0,179
K17	0,517	0,501	0,886	0,221	1,000	0,404	0,414	0,401
K24	0,429	0,346	0,538	-0,073	0,404	1,000	0,329	-0,051
K26	0,769	0,926	0,458	0,452	0,414	0,329	1,000	-0,131
NET_ASSETS	-0,082	-0,087	0,145	0,179	0,401	-0,051	-0,131	1,000

Полученные значения этих двух показателей обусловлены особенностями анализируемых выборочных данных.

Необходимость в нормировке данных, по мнению авторов, отсутствует, поскольку выборочные данные по независимым переменным достаточно однородны. Это связано с тем, что большинство показателей рассчитываются как отношения, в результате чего большинство из них принимают значения от 0 до 1. Кроме того, нормировка данных не позволит дать четкую интерпретацию коэффициентов линейной регрессии, а именно на сколько изменится величина чистых активов предприятия при изменении одного из показателей на единицу.

Далее была построена корреляционная матрица для рассматриваемых переменных (табл. 2).

Анализ значимых коэффициентов корреляции факторов с зависимой переменной позволил сделать следующие выводы:

• Существует выраженная прямая зависимость между величиной чистых активов и коэффициентом наличия собственных средств, между величиной чистых активов и коэффициентом соотношения собственных и заемных средств, между величиной чистых активов и рентабельностью продукции.

• Наблюдается обратная зависимость между величиной чистых активов и отношением чистой прибыли к совокупным активам.

Анализ корреляционной матрицы показал, что между переменными К8 и К26 наблюдается сильная зависимость, поскольку коэффициент корреляции между ними равен 0,926.

Обратная зависимость между отношением совокупных активов к совокупному долгу и рентабельностью деятельности предприятия может быть свидетельством наличия мультиколлинеарности. Ее наличие означает, что некоторые факторы всегда будут действовать однонаправленно. В результате вариация в исходных данных перестает быть полностью независимой и нельзя оценить воздействие каждого фактора в отдельности.

Далее была проведена проверка наличия мультиколлинеарности двумя различными способами.

Первый способ основан на применении метода вспомогательных регрессий, отражающих зависимость каждой из независимых переменных от остальных. В этих моделях нас интересовали только коэффициенты детерминации, выражающие силу этой зависимости.

$R^2(К2) = 0,67$; $R^2(К8) = 0,9$; $R^2(К13) = 0,88$; $R^2(К14) = 0,32$; $R^2(К17) = 0,85$; $R^2(К24) = 0,52$; $R^2(К26) = 0,89$, (в скобках указаны зависимые переменные).

Поскольку ни один из коэффициентов детерминации не превосходит критическое значение 0,9, то мультиколлинеарности нет. Соответственно можно рассчитать значения показателя толерантности (TOL) и коэффициента «вздутия» дисперсии (VIF):

$$TOL_1 = 1-R_1^2 = 0,33 \quad VIF_1 = 1/(1-R_1^2) = 3,03$$
$$TOL_2 = 1-R_2^2 = 0,1 \quad VIF_2 = 1/(1-R_2^2) = 10$$
$$TOL_3 = 1-R_3^2 = 0,12 \quad VIF_3 = 1/(1-R_3^2) = 8,33$$
$$TOL_4 = 1-R_4^2 = 0,68 \quad VIF_4 = 1/(1-R_4^2) = 1,47$$
$$TOL_5 = 1-R_5^2 = 0,15 \quad VIF_5 = 1/(1-R_5^2) = 6,67$$
$$TOL_6 = 1-R_6^2 = 0,48 \quad VIF_6 = 1/(1-R_6^2) = 2,08$$
$$TOL_7 = 1-R_7^2 = 0,11 \quad VIF_7 = 1/(1-R_7^2) = 9,09$$

Второй способ основан на проверке определителя матрицы парных коэффициентов корреляции между факторами (табл. 3).

Определитель этой матрицы равен $\Delta(R) = 0,0019$. Чем меньше корреляция между факторами, тем ближе значение определителя данной матрицы к единице. Чем выше корреляция между факторами, тем ближе

Таблица 3

	K2	K8	K13	K14	K17	K24	K26
K2	1,000	0,741	0,502	0,338	0,517	0,429	0,769
K8	0,741	1,000	0,618	0,395	0,501	0,346	0,926
K13	0,502	0,618	1,000	0,232	0,886	0,538	0,458
K14	0,338	0,395	0,232	1,000	0,221	- 0,073	0,452
K17	0,517	0,501	0,886	0,221	1,000	0,404	0,414
K24	0,429	0,346	0,538	- 0,073	0,404	1,000	0,329
K26	0,769	0,926	0,458	0,452	0,414	0,329	1,000

значение определителя к нулю. И, соответственно, тем сильнее мультиколлинеарность факторов и ненадежнее результаты множественной регрессии.

Оценка значимости мультиколлинеарности факторов может быть проведена путем проверки гипотезы о независимости переменных. Гипотеза H_o представляет собой не что иное, как предположение об отсутствии мультиколлинеарности в модели. Конкурирующая гипотеза H_1 предполагает обратное. Установим уровень значимости α = 0,05. Уровень значимости – это максимально допустимое значение вероятности отклонения верной гипотезы H_o. Нулевая гипотеза принимается, если вычисленное значение статистического критерия не превышает его табличного значения при заданном уровне значимости.

$$H_0: \Delta(R)=1;\ H_1: \Delta(R) \neq 1$$

$$k_{выч} = \left[n - 1 - \frac{1}{6}(2m+5)\lg \Delta(R) \right] = 69.52 .$$

$k_{таб} = x_{1-\alpha}\left[\chi^2 (\frac{1}{2} n(n-1)) \right] = 136.15$ – квантиль распределения хи-квадрат, найденная по статистическим таблицам [7], где α = 0,05, n – количество наблюдений (61 наблюдение), m – количество параметров модели, включая константу (равно 8).

Поскольку $k_{выч} < k_{таб}$, следовательно, гипотезу H_0 принимаем: мультиколлинеарность отсутствует на уровне значимости 0,05.

Отбор наилучшей множественной линейной регрессионной модели проводился по нескольким критериям. Прежде всего модель должна быть значима по критерию Фишера на уровне значимости 0,05. Кроме того, должны быть значимы все коэффициенты искомой модели на том же уровне значимости. Сравнение качества построенных моделей производилось по двум характеристикам: величине скорректированного коэффициента детерминации (Adjusted R-squared), а также значению коэффициента корреляции между фактическим значением зависимой переменной (величины чистых активов) и прогнозным значением, полученным по модели.

Коэффициент детерминации (R-squared) показывает качество подгонки регрессионной модели к наблюдаемым значениям зависимой переменной.

Если его значение равно нулю, то полученная регрессия не улучшает качество предсказания зависимой переменной Y_t по сравнению с тривиальным предсказанием $\hat{Y}_t = \bar{Y}$, где $\bar{Y}$ – выборочное среднее.

При коэффициенте детерминации, равном 1, можно говорить о точной подгонке: все точки наблюдений удовлетворяют уравнению регрессии.

Однако использование простого коэффициента детерминации в качестве критерия сравнения качества двух моделей связано с одной существенной проблемой: значение коэффициента детерминации возрастает при добавлении в модель еще одного регрессора.

Поэтому был использован скорректированный коэффициент детерминации. Коррекция производилась с учетом числа регрессоров в модели. Использование Adjusted R-squared позволило сравнить регрессии с разным количеством регрессоров.

Формула для расчета обычного коэффициента детерминации выглядит следующим образом:

$$R^2 = 1 - \frac{RSS}{TSS},$$

где RSS – сумма квадратов остатков по модели, TSS – сумма квадратов отклонений значения зависимой переменной от среднего значения.

При расчете скорректированного коэффициента Adjusted R-squared формула корректируется следующим образом:

$$R_{adj}^{\ 2} = 1 - \frac{RSS(n-1)}{TSS(n-k-1)},$$

где n – количество наблюдений, k – количество факторов в модели.

Его использование позволяет выбрать наилучшую линейную регрессионную модель, так как дает возможность сравнить регрессии с разным количеством переменных и корректно оценить качество модели.

В результате проведенного анализа была построена модель прогнозирования величины чистых активов предприятия:

$$\begin{gathered} NET\_ASSETS = -83453{,}59302 - \\ -359873{,}7576*K2 + 929388{,}9235*K8 - \\ -297979{,}5072*K13 + 588912{,}984*K14 + \\ +895066{,}1809*K17 + 101184{,}9771*K24 - \\ -2376563{,}111*K26. \qquad (1) \end{gathered}$$

Данная модель является наилучшей среди всех возможных моделей, поскольку отвечает всем нижеописанным критериям:

1. Проверка значимости коэффициентов регрессии проводилась на основе критерия Стьюдента на уровне значимости 0,05[4]. Поскольку все Prob.(t-statistic) < 0,05, то гипотезу о незначимости коэффициентов регрессии отвергаем. Все коэффициенты регрессии значимы на уровне значимости 0,05.

2. Модель в целом значима по критерию Фишера, поскольку Prob (F-statistic) также меньше 0,05.

3. Скорректированный коэффициент детерминации в модели наибольший по сравнению с другими моделями.

4. Коэффициент корреляции прогнозного значения по модели и фактического значения вектора значений зависимой переменной также наибольший.

Наиболее содержательные и точные выводы относительно модели (1) по результатам наблюдений могут быть получены при следующих предположениях.

1. Значения $x_1,\ldots,x_n$ являются неслучайными величинами.

2. Математическое ожидание случайной ошибки в каждом наблюдении равно нулю, т. е. $M(\varepsilon_i) = 0, \quad i = \overline{1,n}.$

3. Дисперсия случайной ошибки постоянна для всех наблюдений, т. е. $D(\varepsilon_i) = M(\varepsilon_i^2) = \sigma^2, \ i = \overline{1,n}.$

4. Случайные ошибки различных наблюдений статистически не связаны (некоррелированы) между собой, т. е. $M(\varepsilon_i \varepsilon_j) = 0$ при $i \neq j$.

5. Случайные ошибки имеют совместное нормальное распределение, в частности.

При выполнении этих условий модель (1) называют классической линейной регрессионной моделью наблюдений.

В регрессионном анализе часто вместо условия о неслучайности значений объясняющей переменной используется более слабое условие о независимости (некоррелированности) объясняющей переменной и случайной ошибки. Получаемые при этом условии МНК-оценки коэффициентов регрессии обладают теми же основными свойствами, что и оценки, полученные при использовании условия 1.

Условие постоянства дисперсии случайной ошибки называют также гомоскедастичностью случайных ошибок. Зависимость дисперсии случайной ошибки от номера наблюдения называется гетероскедастичностью.

Если условие гомоскедастичности не выполняется, то оценки коэффициентов регрессии по методу наименьших квадратов будут неэффективными, хотя и несмещенными. Для диагностики и устранения гетероскедастичности существуют специальные методы.

Четвертое условие указывает на некоррелированность случайных ошибок для разных наблюдений. Это условие часто нарушается, когда исходные данные являются временными рядами. В этом случае говорят об автокорреляции случайных ошибок.

Главным основанием широкого использования МНК-оценок является следующее утверждение [4].

Теорема Гаусса – Маркова. Если для модели (1) выполняются условия 1-4, то оценки $\hat{a}, \hat{b}$, найденные с помощью МНК, являются наилучшими линейными несмещенными оценками (Best Linear Unbiased Estimator, или BLUE) параметров a и b. При этом для любых действительных значений c_1, c_2 оценка $\hat{a}c_1 + \hat{b}c_2$ является наилучшей линейной несмещенной оценкой параметрической функции $ac_1 + bc_2$.

При работе с реальными статистическими данными важно проверить, действительно ли желаемые ограничения (условия, определяющие классическую парную линейную регрессионную модель) имеют место. После проведения этой проверки может также возникнуть вопрос, как поступить, если выяснится, что нарушение условий действительно имеет место.

Для проверки выполнения стандартных предположений о линейной модели наблюдений помимо графических существует довольно много процедур, использующих статистические критерии проверки гипотез.

Основу тестов, ориентированных на проверку классических предположений, составляют остатки $e_i = y_i - \hat{y}_i, i = \overline{1,n}$. Если все они выполнены, число наблюдений достаточно велико, то, как правило, остатки ведут себя как слабо коррелированные случайные величины, имеющие нормальный закон распределения с нулевым математическим ожиданием и дисперсией σ^2; при этом $\mathbf{M}|e_i| = const \cdot \sigma$.

Проверка стационарности последовательности остатков с помощью сериального критерия Вальда – Вольфовитца

Критерий Вальда – Вольфовитца ориентирован на проверку предположения о том, можно ли считать, что остатки получены в результате независимых наблюдений одной и той же случайной величины.

Основу статистики критерия составляет количество серий KS, вычисляемое по остаткам следующим образом. Последовательности остатков поставим в соответствие знаковую последовательность, элементы которой определяются по правилу: знак «+», если $e_i > 0$, знак «-», если $e_i < 0$. Нулевые остатки игнорируются. Серией называют последовательность из расположенных рядом одинаковых знаков. Обозначим через n_1 – общее количество плюсов в знаковой последовательности, а через n_2 – общее количество минусов. При $n_1 > 20$ или $n_2 > 20$ проверка стационарности осуществляется с помощью статистики,

$$Z = \frac{\left| KS - \left(\frac{2n_1 n_2}{n_1 + n_2} + 1 \right) \right| - 0,5}{\sqrt{\frac{2n_1 n_2 (2n_1 n_2 - n_1 - n_2)}{(n_1 + n_2)^2 (n_1 + n_2 - 1)}}},$$

имеющей при отсутствии в записи знака модуля приблизительно стандартное нормальное распределение. Гипотеза о стационарности остатков принимается, если выполняется неравенство $Z_{выч} < t_{1-\alpha}\left[N(0;1)\right]$, где $t_{1-\alpha}N(0,1)$ – двухсторонняя квантиль стандартного нормального распределения.

В нашем случае KS = 22, $n_1 = 24$, $n_2 = 37$, уровень значимости – 0,05. Отсюда

$$k_{выч} = -1,89 \quad k_{табл} = 1,96$$

Вычисленное значение критерия по модулю меньше табличного значения, тогда была принята нулевая гипотеза: вектор остатков является случайной величиной.

Проверка нормальности ошибок модели

Наиболее распространенным тестом на нормальность является тест Жарка – Бера (Jarque-Bera), который состоит в проверке одновременного равенства нулю коэффициентов асимметрии и эксцесса остатков. Этот критерий реализован в ряде пакетов статистического анализа данных (например, в EVIEWS).

Статистика Жарка – Бера выражается формулой:

$$K_{выч} = JB = n\left[\frac{\mathrm{Sk}^2}{6} + \frac{\mathrm{Ku}^2}{24}\right].$$

Здесь Sk – выборочный коэффициент асимметрии (sample skewness), Ku – выборочный коэффициент эксцесса (sample kurtosis), вычисляемые по значениям остатков.

Если распределение ошибок модели действительно является нормальным, то значения выборочного коэффициента асимметрии и выборочного коэффициента эксцесса близки к 0.

При нарушении условия нормальности распределения ошибок значения статистики JB имеют тенденцию к возрастанию. Поэтому гипотеза нормальности ошибок отвергается, если значения этой статистики «слишком велики», а именно, если

$$JB_{выч} > x_{1-\alpha}\left[\chi^2(2)\right].$$

Критерий применим лишь при достаточно большом объеме выборки n.

В данном случае n = 61, sk = 2,13, ku = 0,62, уровень значимости задан на уровне 0,05. Тогда $k_{выч} = 3{,}89$, $k_{табл} = 5{,}99$.

$k_{выч} < k_{табл}$, следовательно, нулевую гипотезу следует принять, т. е. на уровне значимости 0,05 можно утверждать, что данные имеют нормальное распределение.

Проверка остатков на автокорреляцию

Одно из предположений регрессионного анализа состоит в независимости результатов различных наблюдений, а именно $\mathbf{M}\left[\varepsilon_i\varepsilon_j\right] = 0$ при $i \neq j$.

Если данное условие не выполняется, то говорят, что случайный член подвержен автокорреляции (autocorrelation, serial correlation). В этом случае коэффициенты регрессии, получаемые по МНК, оказываются неэффективными, хотя и несмещенными, а их стандартные ошибки рассчитываются некорректно (занижаются). Наиболее распространенным критерием проверки автокорреляции является критерий Дарбина – Уотсона.

Выдвигается нулевая гипотеза об отсутствии автокорреляции первого порядка, т. е. $H_0 : \rho = 0$. Для проверки нулевой гипотезы используют статистику Дарбина-Уотсона:

$$DW = \frac{\sum_{i=2}^{n}\left(e_i - e_{i-1}\right)^2}{\sum_{i=1}^{n} e_i^2}.$$

Можно показать, что с ростом объема выборки $DW \xrightarrow{P} 2 - 2\rho$, где ρ – коэффициент автокорреляции первого порядка последовательности случайных ошибок $\varepsilon_i, i = \overline{1,n}$.

Если автокорреляция остатков отсутствует, то $DW \approx 0$. При положительной автокорреляции имеем $0 < DW \leq 2$, а при отрицательной – соответственно $2 < DW \leq 4$.

Проверка гипотезы осуществляется с помощью табл. 4, содержащей критические значения критерия Дарбина – Уотсона $d_U = d_U(\alpha,k,n)$ и $d_L = d_L(\alpha,k,n)$, определяемые по значениям входных параметров α, k, n.

При проверке гипотезы $H_0 : \rho = 0$ против гипотезы $H_1^+ : \rho > 0$ гипотеза об отсутствии автокорреляции первого порядка принимается, если выполняется неравенство $DW_{выч} > d_U(\alpha,k,n)$; гипотеза отклоняется в пользу альтернативной, если $DW_{выч} < d_L(\alpha,k,n)$.

При проверке гипотезы $H_0 : \rho = 0$ против гипотезы $H_1^- : \rho < 0$ гипотеза об отсутствии автокорреляции первого порядка принимается, если выполняется неравенство $DW_{выч} < 4 - d_U(\alpha,k,n)$; гипотеза отклоняется в пользу альтернативной, если $DW_{выч} > 4 - d_L(\alpha,k,n)$.

При проверке гипотезы $H_0 : \rho = 0$ против гипотезы $H_1 : \rho \neq 0$ гипотеза об отсутствии автокорреляции первого порядка принимается, если выполняется неравенство $4 - d_U\left(\frac{\alpha}{2},k,n\right) > DW_{выч} > d_U\left(\frac{\alpha}{2},k,n\right)$; гипотеза отклоняется в пользу альтернативной, если $DW_{выч} > 4 - d_L\left(\frac{\alpha}{2},k,n\right)$ или $DW_{выч} < d_L\left(\frac{\alpha}{2},k,n\right)$.

Недостатком критерия Дарбина – Уотсона является наличие области неопределенности критерия, не позволяющей при определенных значениях статистики Дарбина – Уотсона решить задачу проверки гипотез. Наличие области неопределенности критерия обусловлено тем, что распределение статистики зависит не только от числа наблюдений и числа объясняющих переменных, но и от значений объясняющих переменных. Кроме того, следует иметь в виду, что критические значения статистики DW определены для объемов выборки не менее 15.

В нашем случае: $DW_{выч} = 2{,}09$, $d_U = 1{,}77$. Поскольку $1{,}77 < DW_{выч} < 2{,}23$, то на уровне значимости 0,05 гипотезу об отсутствии автокорреляции принимаем.

Проверка гетероскедастичности случайных ошибок

Одной из предпосылок регрессионного анализа является предположение о постоянстве дисперсии случайной ошибки для всех наблюдений (свойство гомоскедастичности). Это значит, что для любых значений объясняющих переменных случайные ошибки имеют одинаковые дисперсии. Если это условие не соблюдается, то имеет место гетероскедастичность (heteroscedasticity). Вместо термина «гетероскедастичная модель» часто используют термин «модель с неравноточными наблюдениями». В случае гетероскедастичной модели МНК-оценки коэффициентов регрессии оказываются хотя и несмещенными, но неэффективными.

Влияние гетероскедастичности может существенно снизить точностные показатели модели и сделать ненадежными выводы статистической проверки гипотез о значимости параметров модели и связи между факторами и показателем.

Проверка модели на гетероскедастичность была осуществлена при помощи теста Уайта[4]:

$$H_0:\ \sigma_i^2 \equiv \sigma^2$$

$$H_1:\ \sigma_i^2 \equiv \sigma^2 \cdot (a_0 + a_1 \cdot k2 + a_2 \cdot k2^2 + a_3 \cdot k8 + a_4 \cdot k8^2 + a_5 \cdot k13 + a_6 \cdot k13^2 + a_7 \cdot k14 + a_8 \cdot k14^2 + a_9 \cdot k17 + a_{10} \cdot k17^2 + a_{11} \cdot k24 + a_{12} \cdot k24^2 + a_{13} \cdot k26 + a_{14} \cdot k26^2)$$

Таблица 4

White Heteroskedasticity Test:

F-statistic	1.649890	Probability	0.101303
Obs*R-squared	20.39128	Probability	0.118280

Test Equation:
Dependent Variable: RESID^2
Method: Least Squares
Sample: 1 61
Included observations: 61

Variable	Coefficient	Std. Error	t-Statistic	Prob.
C	-2.21E+10	1.71E+10	-1.289263	0.2038
K2	4.76E+09	7.71E+10	0.061772	0.9510
K2^2	-8.94E+10	1.58E+11	-0.565672	0.5744
K8	1.81E+11	1.74E+11	1.041593	0.3030
K8^2	-2.40E+11	2.29E+11	-1.049656	0.2994
K13	-2.25E+10	2.39E+10	-0.943295	0.3505
K13^2	3.02E+09	7.91E+09	0.381865	0.7043
K14	1.88E+11	7.04E+10	2.664239	0.0106
K14^2	-2.05E+11	2.14E+11	-0.958149	0.3430
K17	8.44E+10	6.45E+10	1.308699	0.1971
K17^2	-7.27E+10	7.89E+10	-0.921889	0.3614
K24	3.10E+10	3.72E+10	0.834188	0.4085
K24^2	-4.27E+09	2.05E+10	-0.208830	0.8355
K26	-7.80E+11	3.52E+11	-2.218179	0.0315
K26^2	3.20E+12	1.62E+12	1.976642	0.0541

R-squared	0.334283	Mean dependent var	5.58E+09
Adjusted R-squared	0.131674	S.D. dependent var	1.83E+10
S.E. of regression	1.71E+10	Akaike info criterion	50.16937
Sum squared resid	1.34E+22	Schwarz criterion	50.68844
Log likelihood	-1515.166	F-statistic	1.649890
Durbin-Watson stat	2.154154	Prob(F-statistic)	0.101303

Как видно из таблицы 4, вероятность ошибки первого рода равна 0,1. Следовательно, принимается нулевая гипотеза (об отсутствии гетероскедастичности). (Если Probability = 0,1 > 0,05, то принимается гипотеза H_0: (дисперсии одинаковые)).

Выводы

Анализ корреляционной матрицы позволил сделать предположение о мультиколлинеарности независимых переменных. Однако использование различных методов обнаружения мультиколлинеарности дало ответ – мультиколлинеарность в данных отсутствует.

В результате сравнения различных построенных регрессионных моделей была выбрана наиболее удачная модель.

Анализ модели показал, что для нее выполняются все классические предположения, кроме предпосылки о нормальности распределения вектора остатков. Тем не менее, это не сильно сказывается на эффективности оценок, полученных методом наименьших квадратов. Таким образом, была построена модель прогнозирования величины чистых активов, построенная на основе реальных данных финансовой отчетности предприятий Пермского края.

На рисунке представлены фактические значения величины чистых активов и их прогнозные значения по рассматриваемой модели. По оси абсцисс отображен порядковый номер предприятия, по оси ординат – значение величины чистых активов.

В Российской Федерации пока еще отсутствует статистика неплатежеспособности предприятий по причине молодости института банкротства в нашей стране, что затрудняет собственные разработки, основанные на реалиях нашей экономики и направленные на достоверное прогнозирование возможного банкротства предприятий. Существует также проблема достоверности информации о состоянии дел на конкретных предприятиях и трудности ее получения. Невозможность использования таких моделей, как модель Альтмана [1], Фулмера [2], Спрингейта [3] и других известных авторов, обусловливается прежде всего тем, что они были разработаны на основе данных финансовой отчетности иностранных государств, функционирующих в совершенно иных рыночных условиях. В данной ситуации перед нами было только два пути решения. Первый – это адаптация моделей к реалиям российской экономики и проведение дискриминантного анализа данных конкретных предприятий, второй – по-нашему мнению, наиболее правильный, но более трудоемкий – создание новой модели [5].

Кроме того, представляется возможным использование построенной модели совместно с пакетами анализа инвестиционных проектов, в результате чего становится возможен расчет и прогноз не только таких показателей, как NPV проекта, свободных денежных средств, нераспределенной прибыли, дисконтированного срока окупаемости, но и величины чистых активов (фактически финансовое состояние предприятия), что важно и во время и после реализации инвестиционного проекта, и, кроме того, влиять на конечный результат, изменяя график выборки денежных средств и/или график погашения основного долга.

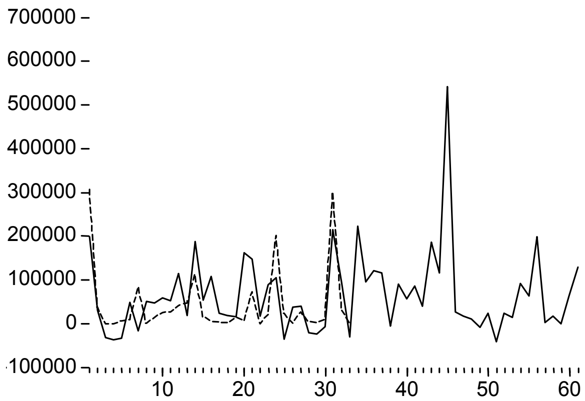

Рисунок 1. Величины чистых активов предприятий, рассчитанные из фактических данных и при помощи построенной модели

Рисунок 2. Величины чистых активов предприятий, рассчитанные из фактических данных и при помощи построенной модели (out-of-sample)

Анализ модели показал, что для нее выполняются все классические предположения, кроме предпосылки о нормальности распределения вектора остатков. Тем не менее, это не сильно сказывается на эффективности оценок, полученных методом наименьших квадратов. Вычислив при помощи построенной модели прогнозную величину чистых активов на 1.10.2005 г. 61 предприятия Пермского края и сравнив данные величины с фактическими, можно сделать вывод о том, что на качественном уровне построенная модель дала хорошее совпадение (см. рисунок).

Для проверки качества построенной множественной линейной регрессионной модели необходимо произвести ее тестирование на так называемом out-of-sample, то есть выборочных данных, не использованных при построении модели. Наиболее целесообразным следует признать возможность проверки модели на более поздних данных. Для этого были использованы данные финансовой отчетности 42 предприятий на 1.10.2005, а в качестве зависимой переменной было выбрано значение чистых активов этих предприятий на 1.10.2006.

В итоге были получены следующие результаты. Коэффициент корреляции между фактическим значением чистых активов и прогнозным по модели составил 0,721 (сильная зависимость). Это свидетельствует о высоком качестве построенной модели и ее потенциале прогнозирования. На рисунке 2 на графике представлены фактические значения величины чистых активов и их прогнозные значения по рассматриваемой модели.

В целом можно сделать вывод о том, что построенная модель оказалась весьма удачной, а ее проверка на out-of-sample показала, что ее применение не ограничивается только той выборкой данных, на которой она была построена. Модель обладает потенциалом прогнозирования величины чистых активов предприятия в будущем периоде на основе анализа его текущей финансовой отчетности.

Список использованной литературы:
1. Altman, Edward I., Financial Ratios, Discriminant Analysis and the Prediction of Corporate Bankruptcy // Journal of Finance (September 1968), pp. 589-609.
2. Fulmer, John G. Jr., Moon, James E., Gavin, Thomas A., Erwin, Michael J., A Bankruptcy Classification Model For Small Firms // Journal of Commercial Bank Lending (July 1984), pp. 25-37.
3. Springate, G. L.V. (1978) Predicting the Possibility of Failure in a Canadian Firm // Unpublished M.B.A. Research Project, Simon Fraser University, January. In: INSOLVENCY PREDICTION, E. Sands & Associates Inc.
4. Магнус Я.Р., Катышев П.К., Пересецкий А.А. Эконометрика. Начальный курс. М.: Дело, 1999.
5. Масенко И.Б. Построение модели анализа финансового состояния предприятия с учетом реализации инвестиционного проекта // Молодые исследователи – регионам: Материалы Всероссийской научной конференции студентов и аспирантов. В 2-х т. – Вологда: ВоГТУ, 2005. – Т. 2, стр. 49-50.
6. Сиразетдинов Т.К. // Динамическое моделирование экономических объектов. – Казань, «Фэн», 1996.
7. Большев Л.Н., Смирнов Н.В. Таблицы математической статистики. – М.: Наука, 1977.