

УДК 519.86

Е.В. Царькова

ФГБОУ ВО «Российский государственный университет правосудия», Москва,
email: e.v.tsarkova@mail.ru

ПРОГНОЗИРОВАНИЕ В УПРАВЛЕНИИ НА ОСНОВЕ EXCEL И MATHCAD

Ключевые слова: прогнозирование, регрессионный анализ, аппроксимация, линия тренда, метод наименьших квадратов.

В статье рассмотрена актуальная проблема прогнозирования и регрессионного анализа динамических рядов на основе метода экстраполяции по множеству наблюдений, полученных последовательно во времени. Построены примеры использования для прогнозирования ситуаций в управлении предприятием полученных в Excel и MathCad регрессионных моделей.

E.V. Tsarkova

Russian state University of justice, Moscow, email: e.v.tsarkova@mail.ru

FORECASTING IN MANAGEMENT BASED ON EXCEL AND MATHCAD

Keywords: forecasting, regression analysis, approximation, trend line, least squares method.

The article considers the problem of forecasting and regression analysis of dynamic series based on the method of extrapolation from a set of observations obtained sequentially over time. Examples of the use of regression models obtained in Excel and MathCad for predicting situations in enterprise management are constructed.

На разных горизонтах планирования формы организации и задачи существенно различаются. В оперативном планировании разработка мероприятий, бюджетов, заданий по разным направлениям ведется на небольшой временной интервал (неделя, месяц, квартал, год), текущее планирование на среднесрочную перспективу – на 1-3 года, перспективное стратегическое планирование – на 3-5 лет и более. Для расчета, планирования и прогнозирования основных производственных показателей деятельности организации используются различные методы. Программно-целевой метод формирует задачи на основе имеющихся ресурсов организации, при нормативном методе количественные показатели рассчитываются на основе системы стандартных показателей и норм, методы экономико-математического моделирования и экономической кибернетики (имитационного моделирования) применяются для оптимизации организации производства и рационального использования материальных ресурсов и времени выполнения проектов для минимизации расходов и максимизации прибыли, методами математической статистики и эконометрики

(стохастического факторного анализа) производят обработку пространственно-временных показателей и прогнозирования результатов принятых решений.

Бывает так, что руководству организации не хватает данных для принятия решений. При отсутствии количественных данных используют качественные методы прогнозирования, например, метод экспертных оценок [1, 2].

Устойчивое и ритмичное функционирование организации, стабильно высокая прибыль могут быть обеспечены только грамотными управленческими решениями, основанными на социально-экономическом прогнозировании, что невозможно без знания соответствующих математических методов и моделей. При неиспользовании экономико-статистических методов или неправильном их применении возможны простои, затоваривание, нехватка кадров, убытки и другие негативные последствия.

Цель исследования

Работа направлена на совершенствование научно-методической базы теории принятия обоснованных управленческих решений.

Обзор литературы

Литература по проблемам прогнозирования и планирования экономического развития в условиях цифровой экономики на всех уровнях управления от фирмы, муниципального образования, региона и до государства [3-8] показывает место и задачи прогнозирования и планирования в системе управления экономикой и методы решения этих задач на уровне фирм малого, среднего бизнеса, а также крупных фирм (кластеров). Особое место занимает исследование проблем прогнозирования и развития отраслевых рынков и системных объектов экономики и на его основе формирование прогноза бюджета социально-экономического развития территорий соответствующего уровня. Формализация и решение задачи ситуационного управления функциональной устойчивостью крупномасштабных эргатических систем рассмотрены в [9], [10], инструментальные методы моделирования с использованием пакета «Statistika» и пакета «Анализ данных» в Excel – в [11], организационно-технический регламент обработки наборов данных – в [12], проблемы управления проектами в условиях информационной неопределенности – в [13].

Материал и методы исследования

Метод управления на основе статистик состоит не только в сборе и анализе данных, это точная научно-обоснованная технология, применение которой дает однозначный рост.

В экономико-статистических методах исследуются социально-экономические явления и осуществляется прогнозирование с использованием правил сбора информации, специальных способах ее обработки, вычисления числовых характеристик обобщающих статистических показателей, проведения анализа на основе экономических законов. Статистическая отчетность организаций как важнейший и наиболее распространенный тип статистического наблюдения отражает структуру и динамику рынка, его колебания, что позволяет проследить влияние различных рыночных факторов и построить прогнозы развития.

В регрессионном анализе процессы и явления разной природы изучаются

в связи с показателями, характеризующими различные стороны этих процессов. Из множества причинно-следственных связей выбирают наиболее существенные, отбрасывая второстепенные и случайные. Вначале исследователь решает, какие виды данных характеризуют причину (факторные или независимые), а какие рассматриваются как результативные (зависимые, являющиеся следствием). Большое число переменных может быть сведено к меньшему количеству факторных величин за счет объединения в один фактор переменных (признаков), сильно коррелирующих между собой. Факторы могут быть как позитивными, так и негативными. Разумеется, для выяснения того, насколько эффективно работает компания, необходима система количественного измерения произведённой продукции или предоставленных услуг. Результативными показателями обычно считаются финансовые показатели: выручка от реализации продукции, рентабельность продаж, рентабельность товаров, производственных ресурсов, капитальных вложений. Могут использоваться многие статистические показатели: одни отражают эффективность компании в целом, другие – уровень работы отделов и подразделений, третьи характеризуют работу отдельных сотрудников. В слабо структурированных проблемах теснота связи между фактором и результатом невысока, да и в целом связи не остаются неизменными.

На факторный показатель с разной степенью влияния могут воздействовать множество факторов. Существующие закономерности необходимо учитывать при анализе рассматриваемого процесса, планировании и прогнозировании. Формы связи между показателем и факторами на основе статистических данных изучает регрессионный анализ: устанавливается форма зависимости между переменными, оценивается функция регрессии и прогнозируются значения зависимой переменной.

Для прогнозирования состояний экономических систем, исходя из полученных экспериментальных данных о значениях переменных x и y , предлагается гипотетическая математическая модель $\hat{y} = f(x)$, причем числовые значения

параметров моделей находятся с учетом требования сделать минимальными расхождения между исходными (y) и вычисленными по формуле ($\hat{y}$) значениями показателей. Это расхождение может оцениваться по разным критериям. Метод наименьших квадратов является наиболее распространенным для оценивания параметров моделей. Предварительно выбирается класс функций, аппроксимирующих изучаемую зависимость, например,

линейные функции $\hat{y}(x) = a_0 + a_1x$, и в заданном классе функций находится та, для которой сумма всех квадратов отклонений эмпирических данных от значений, вычисленных по этой формуле, была бы минимальной. Предположим, что собраны экспериментальные данные за 10 периодов, т.е. для значений факторного признака от 0 до 10 значения результативного признака приведены в таблице 1.

Таблица 1

Значения результативного признака

x	0	1	2	3	4	5	6	7	8	9	10
y	2,1	2,7	3,3	4,9	5,2	5,1	4,8	4,6	4,3	3,6	3,1

Для множества точек наблюдения (x_i, y_i) , $i = 0, 1, 2, \dots, 10$ выберем тип кривой, аппроксимирующей исходные данные. На полигоне распределения (рис.1) видно, что линейная зависимость не соответствует динамике изменения признака, т.к. значения результативного признака сначала возрастают, затем убывают, что может соответствовать параболе с отрицательным старшим коэффициентом.

Поэтому функцию $\hat{y}(x)$ будем искать в виде $\hat{y}(x) = a_0 + a_1x + a_2x^2$, а неизвестные параметры a_0, a_1, a_2 найдем из условия минимума функционала

$$F(a) = \sum_{i=1}^n (y_i - \hat{y}(x_i))^2,$$

приравнивая к нулю его частные производные по всем параметрам:

$$\begin{cases} \sum y = na_0 + a_1 \sum x + a_2 \sum x^2, \\ \sum xy = a_0 \sum x + a_1 \sum x^2 + a_2 \sum x^3, \\ \sum x^2 y = a_0 \sum x^2 + a_1 \sum x^3 + a_2 \sum x^4. \end{cases}$$

Формулы для общего решения системы имеют громоздкий вид, неудобный и для записи, и для вычислений. Обычно составляют вспомогательную таблицу со столбцами $x_i, y_i, x_i^2, x_i^3, x_i^4, x_i y_i, x_i^2 y_i$, суммируют по столбцам. Найденные суммы являются коэффициентами си-

стемы, которые можно решить, например, методом Гаусса.

Функции вида $\hat{y}(x) = a_0 + a_1x + a_2x^2$ при $a_0 > 0, a_1 > 0, a_2 < 0$ могут характеризовать зависимость урожайности сельскохозяйственных культур от количества внесенных удобрений. В экономике степенная производственная функция на имеет младших членов и описывает-

ся формулой $y = a_0x^{\alpha_1}$ или $y = y_0K^{\alpha}L^{1-\alpha}$ (функция Кобба-Дугласа, где y – экономический результат, L – затраты труда, K – объем производственных фондов), а в случае для многофакторных производственных функций $y = \alpha \prod_{i \in M} x^{\alpha_i}$.

Для аппроксимации точечных данных в MathCad в случае полиномиальной аппроксимации используются встроенные функции *regress* и *interp*. Набор значений величины X должен быть записан в порядке возрастания. Функция *regress*{ X, Y, n } выдает набор z коэффициентов искомого полинома порядка n , наилучшим образом приближающего данные X и Y . Функция *interp*(z, X, Y, x) возвращает результат полиномиальной регрессии, не записанный в явном виде, но выдающей его значение при подстановке конкретного значения аргумента x . Определяя функцию $f(x) = \text{interp}(z, X, Y, x)$, при $x = 11$ в рассмотренном примере MathCad получаем результат $f(11) = 1.83$.

```

X := (0
      1
      2
      3
      4
      5
      6
      7
      8
      9
      10)

Y := (2.1
      2.7
      3.3
      4.9
      5.2
      5.1
      4.8
      4.6
      4.3
      3.6
      3.1)

n := 2    z := regress(X,Y,n)    z = ( 3
                                         3
                                         2
                                         1.921
                                         1.143
                                         -0.105)

f(x) := interp(z,X,Y,x)
f(11) = 1.83

Coeffs := submatrix(z,4,length(z),1,1)    CoeffsT = (1.921 1.143 -0.105)
    
```

Рис. 1.

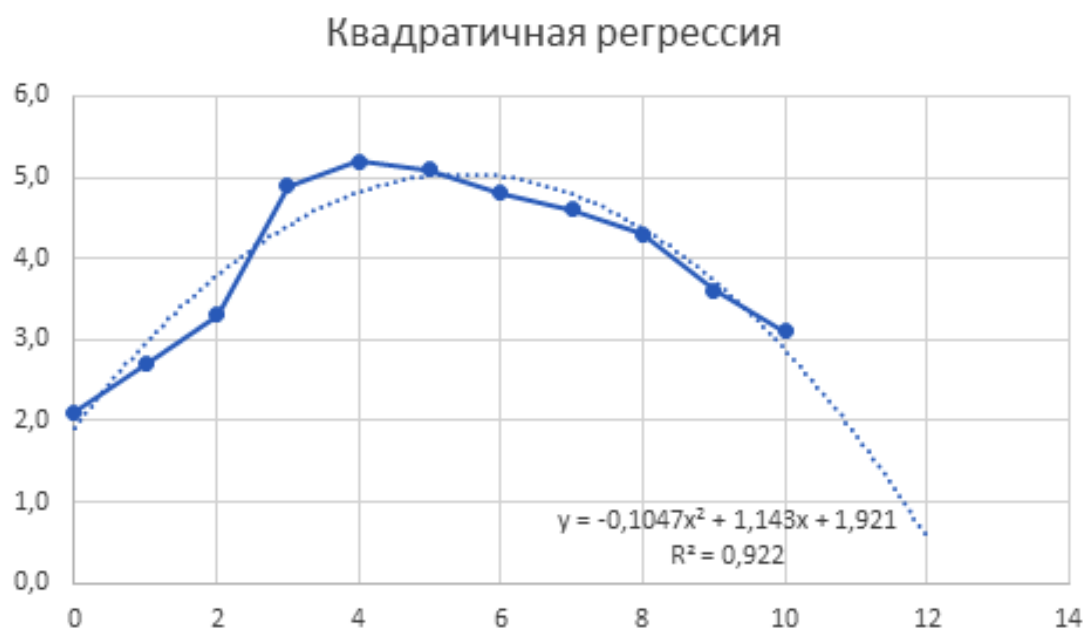


Рис. 2.

Качество аппроксимации оценивается по величине суммы квадратов отклонений.

Для исходного набора данных (табл. 1) в MathCad в качестве $\hat{y}(x)$ получаем (рис.1):

$$f_2(x) = -0.105x^2 + 1.143x + 1,921$$

Такой же результат получается в Excel (рис.2) при выборе следующих параметров линии тренда: полиномиальная, степень 2, вперед на 2 периода.

С помощью средней ошибки аппроксимации (среднего процента расхождения теоретических значений и фактических) можно установить адекватность построенной модели изучаемому процессу:

$$\bar{k} = \frac{1}{n} \sum \frac{|y - \hat{y}|}{\hat{y}} \cdot 100\%$$

Для моделирования экономических показателей допустимой считается 5%-ная погрешность (иногда 7%, редко

10%), т.е. при $\bar{k} \leq 5\%$ модель считается адекватной и пригодной для дальнейшего экономического анализа.

Для оценки силы взаимосвязи показателей эффективным является коэффициент детерминации R^2 , вычисляемый как квадрат эмпирического коэффициента корреляции для двух наборов наблюдений: фактическими значениями y_i и теоретическими, вычисленными по регрессионной формуле значениями зависимой переменной:

$$R^2 = \frac{\left(\sum_{i=1}^n (y_i - \bar{y}) \left(\hat{y}(x_i) - \bar{y} \right) \right)^2}{\sum_{i=1}^n (y_i - \bar{y})^2 \sum_{i=1}^n \left(\hat{y}(x_i) - \bar{y} \right)^2} = 1 - \frac{\sum_{i=1}^n (y_i - \hat{y}(x_i))^2}{\sum_{i=1}^n (y_i - \bar{y})^2}$$

Здесь $\bar{y} = \frac{1}{n} \sum_{i=1}^n y_i$ – среднее значение.

Близость к единице значение коэффициента детерминации означает, что теоретические значения $\hat{y}$ достаточно

точно аппроксимируют эмпирические значения y . Чем он ближе к единице коэффициент детерминации, тем лучше оценивается регрессионное уравнение.

Величина $R^2 \cdot 100\%$ выражает, на сколько процентов изменения Y обусловлено изменением X . В рассмотрен-

ном примере $R^2 = 0,922$, т.е. R^2 близок к единице, значит, регрессионное уравнение оценено хорошо. Иначе говоря, фактор X на 92,2% предопределяет изменение Y . В Excel надстройка «Пакет анализа» представляет широкие возможности для проведения статистического анализа.

Для линейной зависимости между случайными величинами X и Y тесноту линейной связи характеризует коэффициент корреляции. Чем он ближе к ± 1 , тем интенсивнее исследуемая связь между X и Y , чем он ближе к 0, тем связь слабее. Для нелинейных регрессий для оценки интенсивности связи между двумя величинами Y и X рассчитывают индекс корреляции как квадратный корень

из коэффициента детерминации R^2 .

Приведем примеры других функций и процедур в MathCad:

процедура $line(x, y)$ определяет коэффициенты линейного уравнения регрессии

$$f(x) = a_0 + a_1 x;$$

процедура $Infit(x, y)$ определяет коэффициенты логарифмической функции

$$f(x) = a \ln x + b;$$

процедура $Expfit(x, y)$ определяет коэффициенты экспоненциальной функции

$$f(x) = a \cdot e^{bx} + c;$$

процедура $Pwrfit(x, y, g)$ определяет коэффициенты степенной функции

$$f(x) = a \cdot x^b + c.$$

Уравнение регрессии позволяет получить прогноз о будущем состоянии объекта не в категорической, а в вероятностной форме. Экономические явления в развитии отражаются временными

рядами, т.е. наборами некоторых показателей изучаемого объекта или процесса, упорядоченными во времени и берущимися через равные календарные отрезки: год, квартал, месяц и т. д. Отрезок времени от момента времени, для которого имеются последние статистические данные об изучаемом процессе, до момента, к которому относится прогноз, называется периодом упреждения (или периодом прогноза). Прогноз – это вероятностное суждение о возможном состоянии объекта в будущем. При прогнозировании либо дают точечный прогноз, т.е. оценивают в точке прогноза математическое

ожидание величины Y , либо составляют интервальный прогноз – величину интервала, в который с заданной вероятностью попадет прогнозируемое значение изучаемой характеристики.

При прогнозировании широко применяется легко реализуемый метод наименьших квадратов, достоинством которого является несложная реализация, а недостаток состоит в том, что адекватный прогноз можно составить лишь на короткий период упреждения, поэтому он относится к методам краткосрочного прогнозирования. В рассмотренном примере для прогноза на 11-й период в фор-

мулу $f_2(x) = -0,1047x^2 + 1,143x + 1,921$

подставим $x = 11$ и получим $f_2(11) = 1,825$.

Важен правильный выбор вида модели, что является непростой задачей. Представляют трудность обоснование и выбор весов во взвешенном методе наименьших квадратов. На практике для построения прогноза более весомое значение имеет информация о процессе в моменты времени, приближенные к текущему моменту времени. Отстоящая далеко в прошлом информация менее ценна для прогноза. Суть взвешенного метода как раз и состоит в том, что элементам имеющегося динамического ряда присваиваются некоторые веса в зависимости от близости к началу периода прогноза.

Метод наименьших квадратов прост в реализации для линейных зависимостей. Если для прогноза используется не линеаризуемая функциональ-

ная зависимость, то метод значительно усложняется.

Модель тренда используется для прогнозирования при условии, что период, для которого построен тренд, достаточен для выявления тенденций (содержит не менее 10 данных), процесс является динамически стабильным и не ожидается существенных внешних воздействий и случайных колебаний, которые могут изменить тенденции развития.

При учете многих факторов составляется множественная регрессионная модель, где значения выручки Y исследуются в зависимости от нескольких по-

казателей: X_1 – оклады, X_2 – надбавки

за стаж, X_3 – премии, X_4 – социальные выплаты, в том числе материальная помощь, добровольное медицинское страхование и санаторно-курортное обеспечение. Пусть, к примеру, получена линейная модель вида

$$Y = 20000 + 0,86 X_1 + 3,60 X_2 + 8,32 X_3 + 13,34 X_4$$

Социальное стимулирование X_4 оказывает наибольшее влияние на экономический результат, каждый вложенный в социальную сферу рубль увеличивает выручку на 13,34 рублей, менее эффективны премии и надбавки за стаж. Оклады воздействуют слабее, один выплаченный по окладам рубль приводит к росту выручки лишь на 0,86 рубля. Руководство имеет основания предположить, что для роста мотивации работников при найме нового персонала целесообразно снизить размеры окладов при увеличении социального пакета, процента премий и надбавок за стаж, что должно привести к существенному росту.

Выводы

Прогнозирование воздействия различных факторов стимулирования на показатели результативности и финансовые результаты предприятия необходимо в современном менеджменте. Обеспечение управляемого воздействия эффективно лишь при использовании компьютерных технологий и инструментальных средств проведения исследований.

Библиографический список

1. Ващекин А.Н. Моделирование и выбор рациональных стратегий коммерческой деятельности предприятий оптовой торговли: монография. М.: изд-во ВЗФЭИ, 2004, 194 с.
2. Ващекин А.Н. Применение математических методов теории нечетких множеств при моделировании принятия решений в экономической и правовой сфере // Экономика. Статистика. Информатика. Вестник УМО. 2013. № 6. С. 18-21.
3. Бережная Е.В., Бережной В.И. Математические методы моделирования экономических систем. М.: Финансы и статистика, 2006. 432 с.
4. Кирилук И.Л. Модели производственных функций для российской экономики // Компьютерные исследования и моделирование. 2013. Т. 5. № 2. С. 293-312.
5. Кирилук И.Л., Малков С.Ю., Малков А.С. Особенности долгосрочной экономической динамики мировой системы: анализ статистических данных // Прикладная эконометрика. 2009. № 4 (16). С. 34-45.
6. Машунин Ю.К. Прогнозирование и планирование социально-экономических систем: учебник для вузов. М.: Издательство Юрайт, 2023. 330 с.
7. Орлов А.И. Организационно-экономическое моделирование: теория принятия решений: учебное пособие. М.: КноРус, 2013. 576 с.
8. Светуных И.С., Светуных С.Г. Методы социально-экономического прогнозирования в 2 т. Т. 1 теория и методология: учебник и практикум для вузов. М.: Издательство Юрайт, 2023. 351 с.
9. Ловцов Д.А., Сергеев Н.А. Имитационно-игровое моделирование функционирования и прогнозирования развития крупномасштабных эргасистем // Трансформация национальной социально-экономической системы России: Материалы V Международной научно-практической конференции. Москва, 2023. С. 495-502.
10. Ловцов Д.А., Сергеев Н.А. Ситуационное управление устойчивостью функционирования и прогнозирования развития крупномасштабных эргасистем // Трансформация национальной социально-экономической системы России: Тр. IV Междунар. науч.-прак. конф. (22 ноября 2021 г.). М.: РГУП, 2022. С. 367-373.
11. Стельмашонок Е.В., Стельмашонок В.Л., Еникеева Л.А., Соколовская С.А. Моделирование процессов и систем: учебник и практикум для вузов / под редакцией Е. В. Стельмашонок. М.: Издательство Юрайт, 2023. 289 с.
12. Борисов Р.С., Ефименко А.А. Протокол обработки наборов данных для их публикации в открытых источниках // Правовая информатика. 2021. № 2. С. 59-70.
13. Царькова Е.В. Методы управления проектами в условиях информационной неопределенности // Правовая информатика. 2019. № 4. С. 29-39.