

КОНЦЕПТУАЛЬНАЯ ОСНОВА ПОЛУЧЕНИЯ И ИССЛЕДОВАНИЯ МАКСИМАЛЬНОГО ВРЕМЕНИ ПРОГНОЗА С ЗАРАНЕЕ ЗАДАННОЙ ТОЧНОСТЬЮ¹

Кумратова Альфира Менлигуловна, канд. экон. наук, доц.

Кубанский государственный аграрный университет имени И.Т. Трубилина, Калинин, 13, Краснодар, Россия, 350044; e-mail: alfa_eo@rambler.ru

Цель: системно обобщающая комбинация моделей кластерного анализа и моделей конкордации Кенделла для агрегации значений частных прогнозов, доставленных разными методами и подходами. *Обсуждение:* в работе введено понятие «прогнозируемости» временных рядов различной природы, в котором впервые обоснован императив, связывающий длину горизонта прогноза, заданное качество прогноза и статистические и фрактальные характеристики прогнозируемого процесса. Математическая постановка задачи, а также модели, методологии и способы оценки степени «прогнозируемости», инструментальные и программные средства должны иметь реализующие их алгоритмы и инструменты, которые входят в состав профессиональных систем компьютерной математики. *Результаты:* проблемы количественной оценки «прогнозируемости» автор предлагает напрямую связать с длиной «горизонта будущего». Они имеют целью получение а priori оценки долготы и качества прогноза в зависимости от статистических свойств прогнозируемого сигнала при наперед заданной точности.

Ключевые слова: горизонт прогноза, метод конкордации Кенделла, частный прогноз, кластерный анализ.

DOI: 10.17308/meps.2017.6/1678

Введение

Развитие методов исследования, мониторинга, моделирования, системного анализа, представления, прогнозирования в настоящее время связано с непрерывным подъемом значения их формализации. Основой для данного процесса послужили требования существенным образом изменившейся в сторону возрастания стохастичности, турбулентности, волатильности природных и экономических процессов. Актуальность в анализе поведения природных и экономических временных рядов теперь приобретает систем-

¹ Работа выполнена при финансовой поддержке РФФИ (проект № 16-06-00158 А).

ная разработка комплекса разноплановых, взаимозависимых и взаимно дополняющих экономико-математических моделей. Модели связаны между собой, оперируют одним и тем же исходным материалом, а их подбор улучшит репрезентативность алгоритмов современным природным процессам зернового производства, что важно для трансформационных или переходных рыночных экономик. «Циклическое» или «блочное» прогнозирование ставит перед исследователем совершенно новые, значительные и нестандартные проблемы из-за необходимости графического и численного выделения и аналитического продолжения циклических образований в структуре волатильных временных рядов разной природы с тем, чтобы погрузить их в обобщённые или агрегированные индикаторы. Такой подход, естественно, приносит более ощутимые экономические результаты [5, 3].

В работе [11] сказано, что «в зависимости от статистических свойств прогнозируемого сигнала, при наперёд заданной точности эти проблемы имеют своей целью получение а priori оценки долготы и качества прогноза». В классических подходах прогнозирования на базе временных рядов данных как природных, так и экономических и финансовых процессов длина «горизонта прогноза» зависит от статистических свойств сигнала», во-первых, предполагается существование прогнозирующей системы с многими алгоритмами прогнозирования и вычисленными ими значениями частных прогнозов, во-вторых, потребностью генерировать значение обобщённого прогноза. Далее системное исследование указывает на то, что до настоящего времени при прогнозировании природных, биологических, социальных и финансово-экономических процессов недостаточно внимания уделялось на решение «обратной» задачи [11], то есть на получение и исследование максимального времени прогноза с заранее заданной точностью. Это время назовём «длиной прогноза», длиной «горизонта прогноза», «периодом упреждения», «периодом времени упреждения», «глубиной прогнозирования» или более строго длиной «горизонта будущего». При отсутствии понимания причин плохой или хорошей «прогнозируемости» разных временных рядов этот шаг оказывается только первым, но необходимым шагом в получении точного многокритериального прогноза. Скажем пока, что удовлетворительная глубина по качеству «горизонта будущего» достигается на таком промежутке, в котором риск принимаемого решения не превосходит указанной величины [11].

Сначала рассмотрим аппарат кластерного анализа как инструмент многокритериального сравнения в многомерном пространстве точек частных прогнозов и нахождения глобального прогноза в виде точки, находящейся посередине кластера локальных точек с минимумом суммарного «расстояния» от неё до всех локальных. Понятие близости, похожести, схожести между объектами или результатами является базовым в теории и методах аппроксимации, анализа и прогнозирования временных рядов. «Близость» лежит в основе постановки и решения таких классов задач, как анализ взаимосвязей переменных и анализ схожести объектов, которыми в нашем случае являются прогнозы.

Методология исследования

Важной характеристикой кластерного анализа является метрика, т.е. определение термина «функция расстояния». «Расстояния» обобщаются до расстояния между точкой и кластером (наш случай), до расстояниями между кластерами. В кластерном анализе определено, что неотрицательная вещественно-значная функция $d(X_i, X_j)$ называется функцией расстояния («метрикой»), если:

- $d(X_i, X_j) \geq 0$ для $\forall X_i, \forall X_j$ из пространства E_p ;
 - $d(X_i, X_j) = 0$ тогда и только тогда, когда $X_i = X_j$;
 - $d(X_i, X_j) = d(X_j, X_i)$;
 - $d(X_i, X_j) < d(X_i, X_k) + d(X_k, X_j)$ («неравенство треугольника»),
- где X_i, X_j, X_k – любые три вектора из E_p .

Значение $d(X_i, X_j)$ для заданных X_i, X_j называется «расстоянием между X_i и X_j » и оно эквивалентно «расстоянию» между значениями прогнозов в соответственно выбранном пространстве характеристик $\{C_1, C_2, \dots, C_p\}$. В связи со своеобразием «расстояния» в нашей задаче рассмотрим основные известные метрики и их характеристики:

- евклидова метрика:

$$d_E(X_i, X_j) = \left[\sum_{k=1}^p (X_{ki} - X_{kj})^2 \right]^{1/2}.$$

Данная метрика является популярной и особенно употребительной. Имеется также и недостаток: результат d_E зависит от масштаба выбранных единиц измерения характеристик X_{ki}, X_{kj} , этот недостаток распространяется и на многие другие метрики;

- l_1 -норма:

$$d_l(X_i, X_j) = \sum_{k=1}^p |X_{ki} - X_{kj}|,$$

где l_1 – метрика абсолютных значений, является наиболее простой с вычислительной точки зрения, иногда её называют манхеттенской метрикой;

- сюпремум-норма:

$$d_s(X_i, X_j) = \sup \{ |X_{ki} - X_{kj}| \}, k = 1..p.$$

Сюпремум-норма легко исчисляется и содержит внутри себя процедуру ранжирования;

- l_p -норма:

$$d_p(X_i, X_j) = \left[\sum_{k=1}^p (X_{ki} - X_{kj})^p \right]^{1/p}, p > 2.$$

l_p -норма охватывает несколько функций расстояния: евклидову; l_1 -метрику; сюпремум-норму, при этом $p = 2, 1$ и ∞ ;

- мера Махаланобиса (P.C. Mahalanobis):

$$D^2(X_i, X_j) = (X_i - X_j)^T W^{-1} (X_i - X_j).$$

В научной литературе расстояние Махаланобиса часто называют как «обобщённое евклидово расстояние». Здесь W^{-1} означает матрицу, которая

является обратной к матрице рассеяния. Расстояние Махаланобиса инвариантно относительно невырожденных линейных преобразований.

- мера Джеффриса-Матуситы (H. Jeffreys, K. Matasita):

$$DM = \left[\sum_{k=1}^p (\sqrt{X_{ki}} - \sqrt{X_{kj}})^2 \right]^{1/2}.$$

Данная мера изначально введена для измерения расстояния между двумя функциями плотностей вероятности. Она также может быть использована и в качестве измерения расстояния между двумя векторами;

- мера «коэффициент дивергенции»:

$$CM = \left[\sum_{k=1}^p \left(\frac{X_{ki} - X_{kj}}{X_{ki} + X_{kj}} \right)^2 \right]^{1/2}.$$

Обсуждение результатов

В первоначальном применении «коэффициента дивергенции» величины X выступали как действительно «средние» \bar{X} и рассматривались как расстояния между выборочными средними двух выборок.

Конкордация Кендалла или коэффициент множественной ранговой корреляции относится к анализу связи переменных или объектов, например, результатов значений частных прогнозов, полученных разными способами и моделями, посредством коэффициента конкордации. Эти методы следует отнести к так называемым «грубым» методам, необходимость применения которых возникает на начальных этапах решения разного рода задач анализа и прогнозирования. Например, эти методы неплохо смотрятся для порождения (формулирования) гипотез о существовании факторной структуры связей в изучаемой системе переменных или гипотез о существовании однотипных групп объектов.

В общем случае для ранговых переменных или переменных, чье распределение существенно отличается от нормального, используется коэффициент корреляции Спирмена или Кендалла. Отсюда же и общий случай конкордации («согласия», «согласованности») Кендалла:

$$W = 12 \cdot \Sigma D_i^2 / (m^2 \cdot (n^3 - n)),$$

где n – число испытываемых объектов; m – число ранговых последовательностей; R_{ij} – ранги всех объектов; $D_i = d_i - d^-$; $d_i = \Sigma R_{ij}$ (по j от 1 до m); $d^- = (1/n) \cdot \Sigma d_i$ (по i от 1 до n).

В наиболее частом, частном случае конкордация Кендалла применяется для определения согласованности мнений экспертов по нескольким факторам. Тогда формула для коэффициента конкордации будет:

$$W = 12 \cdot S / m^2 \cdot (n^3 - n),$$

где n – число факторов; m – число экспертов в группе; S – сумма квадратов разностей рангов (отклонений от среднего).

Нормировка с коэффициентом 12 в числителе выбрана так, чтобы W изменялась в пределах $0 \leq W \leq 1$. Если $W < 0,2-0,4$, то согласованность экс-

пертов (у нас – прогнозных значений) считается слабой. Если $W > 0,6-0,8$, то согласованность экспертов (прогнозов) сильная.

Под экспертизой понимают проведение измерений группой экспертных специалистов для суждения о явлении и подготовки принятия решения. Мы полагаем, что раз в качестве измерительных приборов выступают люди, каждый со своей объясняющей доминантой, то глобальный или обобщённый аддитивный прогноз будет частично состоять из этих доминант, количественным выражением которых будут коэффициенты в сумме. Особенно трудны те случаи, когда объекты и их характеристики ещё предстоит создать, как это имеет место при составлении прогнозов. Экспертиза, т.е. измерение или сравнение объектов, по определению обязана быть связанной с каким-либо оцениванием объектов. Количественные оценки, соответствующие, как правило, объективным измерениям объективных же показателей, представляются с помощью абсолютной шкалы.

Заметим, что в последнее время в прогностике начинают проявлять интерес не только к таким методам и алгоритмам прогнозирования, у которых на определенных отрезках одного и того же временного ряда демонстрируются отличные результаты, а на других – абсолютно слабые. Вызывает практический интерес изучение степени «прогнозируемости» временного ряда или степени «трендоустойчивости» природного, биологического, социального или экономического процесса, ему соответствующего временного ряда или его частей [11]. Так определяется а priori будущая «прогнозируемость» сигнала, которая будет рассчитывать длительность прогноза при заданной границе его погрешности и состава сигнала. Это некое инвариантное свойство, зависящее от структурных синергетических (фрактальных – мера хаотичности, «долговременная память», цвет «шума», персистентность и антиперсистентность, трендоустойчивость) и классических статистических (стандартное отклонение, дисперсия, коэффициент вариации, асимметрия, коэффициент эксцесса) его характеристик [4]. Предполагаемое объективно существующим свойство «прогнозируемости» тех или иных процессов означает, что в зависимости от их фрактальных и статистических свойств на протяжении одного и того же отчётного периода один процесс может прогнозироваться одним и тем же подходом лучше, а другой – хуже.

Длина «горизонта будущего» становится одним из важных показателей в области прогнозирования, тем самым представляя тот отрезок «будущего», в пределах которого прогноз не будет отличаться от истинного значения прогнозируемой переменной на некоторую наперёд заданной величины погрешности $\pm \varepsilon$. Конечно, более точно со статистической, вероятностной точки зрения следовало бы пользоваться интервальными понятиями, однако на начальном этапе достаточно грубая конструкция $c \pm \varepsilon$ неплохо выполнила свою роль. Как правило, в исследовании при всех вычислениях по умолчанию ε равняется 10%, что для прогнозируемых процессов оказывается совсем небольшой погрешностью.

Заключение

В литературе по футурологии не слишком много работ по измерению или проектированию длины «горизонта будущего», по определению «прогнозируемости» реальных процессов, по сравнению прогнозных оценок для одного процесса, исчисленных различными способами. Новый критерий соответствия – длина «горизонта будущего» – можно теперь использовать для сравнения прогнозных моделей. В условиях, когда поведение социальных, экономических, биологических, природно-климатических систем определяется одновременно разными структурными частями их временных рядов (тренды, сезонные составляющие, циклические конструкции, случайный шум, событийные изменения), особую актуальность приобретает поиск среди альтернативных экстраполирующих моделей тех, для которых максимальная длина «горизонта будущего» соответствует лучшей релевантности модели и экономического или природного процесса.

С другой стороны, окончательные решения по определению «прогнозируемости», по максимизации длины «горизонта будущего» и т.п. должны быть конструктивными, сопровождаясь точным математическим расчётом. Математическая постановка задачи, а также модели, методологии и способы оценки степени «прогнозируемости», инструментальные и программные средства должны иметь реализующие их алгоритмы и инструменты, которые входят в состав профессиональных систем компьютерной математики.

Таким образом, проблемы количественной оценки «прогнозируемости» можно напрямую связать с новым вычисляемым критерием качества – длиной «горизонта будущего». Они имеют целью получение а priori оценки долготы и качества прогноза в зависимости от статистических свойств прогнозируемого сигнала при наперёд заданной точности. Отдельно отметим, что в работе введено понятие «прогнозируемости» временных рядов различной природы, в котором впервые обоснован императив, связывающий длину горизонта прогноза, заданное качество прогноза и статистические и фрактальные характеристики прогнозируемого процесса.

Список источников

1. Иванюк В.А., Пашенко Ф.Ф. Обзор моделей и методов, используемых при прогнозировании долгосрочного социально-экономического развития экономики страны // *Управленческие науки в современном мире*, 2016, т. 2, no. 1, с. 21-24.
2. Крашенинников В.Р., Кувайскова Ю.Е., Клячкин В.Н., Шунина Ю.С. Обновление моделей прогнозирования состояния объектов в виде систем временных рядов и многомерных классификаторов // *Вестник компьютерных и информационных технологий*, 2017, no. 6 (156), с. 11-16.
3. Кумратова А.М., Попова Е.В. Снижение экономического риска на базе предпрогнозного анализа // *Современная экономика: проблемы и решения*, 2015, no. 3, с. 18-27.
4. Кумратова А.М., Попова Е.В., Мусатов И.С., Фролов П.Л., Петраков В.А. Предпрогнозный фазовый анализ эволюционного развития элементов финансового рынка // *Политематический сетевой электронный научный журнал КубГАУ*, 2017, no. 128, с. 772-785.
5. Кумратова А.М., Савинская Д.Н. Модифицированная система моделей и методов прогнозирования временных рядов с памятью // *Современная экономика:*

проблемы и решения, 2015, no. 1, с. 8-18.

6. Матвеева А.А. Прогнозирование курса доллара США на основании эконометрических моделей регрессии // *Экономика и социум*, 2016, no. 12-2 (31), с. 226-233.

7. Мелентьев Б.В. Перспективы прикладного использования межотраслевых оптимизационных моделей экономического прогнозирования // *Интерэкспо Гео-Сибирь*, 2017, т. 3, no. 2, с. 8-11.

8. Мелентьев Б.В. Прогнозирование финансовых потоков на основе межрегиональных межотраслевых моделей // *Экономика и математические методы*, 2016, т. 52, no. 3, с. 50-64.

9. Сазонов А.А., Сазонова М.В. Выявление свойств экзогенности в моделях прогнозирования экономического роста // *Фундаментальные и прикладные исследования: проблемы и результаты*, 2016, no. 24, с. 176-180.

10. Филиппов А.И. Проблемы экономического прогнозирования с помощью моделей авторегрессии // *Конкурентоспособность в глобальном мире: экономика, наука, технологии*, 2016, no. 7-2 (19), с. 157-162.

11. Чадранцев А. В. *Определение прогнозируемости экономических процессов*. Ставропольский гос. университет, Ставрополь, 2005.

THE CONCEPTUAL BASIS FOR MAXIMUM TIME OF THE FORECAST OBTAINMENT AND STUDY WITH A PREDETERMINED ACCURACY

Kumratova Alfira Menligulovna, Cand. Sc. (Econ.)

Kuban State Agrarian University, Kalinina st., 13, Krasnodar, Russia, 350044; e-mail: alfa_eo@rambler.ru

Purpose: the author constructs systematically summarizing combination of cluster analysis models and Kendall concordance models for aggregation the values of different private forecasts. *Discussion:* the author introduces the term of «predictability» for the different nature time series. Note that the term of «predictability» first justifies the imperative. This imperative links the length of the forecast horizon, the specified forecast quality and statistical and fractal characteristics of a predictable process. Mathematical formulation of the problem, as well as models, methodologies and assessment ways the degree of «predictability», instrumentation and software must have algorithms and tools for its implement. These algorithms and tools are the part of professional systems of computer mathematics. *Results:* the author related the problems of quantitative evaluation of «predictability» directly with the «horizon of the future» length. The writer obtained a priori estimates of longitude and forecast quality depending on the statistical properties of the predicted signal at preassigned precision.

Keywords: the forecast horizon, Kendall concordance method, private forecast, cluster analysis.

References

1. Ivanjuk V.A., Pashhenko F.F. Obzor modelej i metodov, ispol'zuemyh pri prognozirovanii dolgosrochnogo social'no-jekonomicheskogo razvitija jekonomiki strany [Overview of models and methods used in prediction long-term socio-economic development of the country's economy]. *Upravlencheskie nauki v sovremenom mire*, 2016, vol. 2, no. 1, pp. 21-24. (In Russ.)
2. Krashennikov V.R., Kuvajskova Ju.E., Kljachkin V.N., Shunina Ju.S. Obnovlenie modelej prognozirovanija sostojanija ob'ektov v vide sistem vremennyh rjadov i mnogomernyh klassifikatorov [Updating of forecasting models the state of objects in the form of systems of time series and multivariate classifiers]. *Vestnik komp'juternyh i informacionnyh tehnologij*, 2017, no. 6 (156), pp. 11-16. (In Russ.)
3. Kumratova A.M., Popova E.V. Snizhenie jekonomicheskogo riska na baze predprognoznogo analiza [Reduction of economic risk on the basis of preprocessor analysis]. *Sovremennaja jekonomika: problemy i reshenija*, 2015, no. 3, pp. 18-27. (In Russ.)
4. Kumratova A.M., Popova E.V., Musatov I.S., Frolov P.L., Petrakov V.A. Predprognoznij fazovyj analiz jevoljucionnogo razvitija jelementov finansovogo rynka [Pre-forecasting phase analysis of the evolutionary development of the elements of the financial market]. *Polythematic*

network electronic scientific journal of the Kuban state agrarian University, 2017, no. 128, pp. 772-785. (In Russ.)

5. Kumratova A.M., Savinskaja D.N. Modificirovannaja sistema modelej i metodov prognozirovanija vremennyh rjadov s pamjat'ju [The modified system of models and methods of forecasting time series with memory]. *Sovremennaja jekonomika: problemy i reshenija*, 2015, no. 1, pp. 8-18. (In Russ.)

6. Matveeva A.A. Prognozirovanie kursa dollara SShA na osnovanii jekonomicheskikh modelej regressii [Forecasting the US dollar rate on the basis of econometric regression models]. *Jekonomika i socium*, 2016, no. 12-2 (31), pp. 226-233. (In Russ.)

7. Melent'ev B.V. Perspektivy prikladnogo ispol'zovanija mezhotraslevykh optimizacionnyh modelej jekonomicheskogo prognozirovanija [Prospects for the application of interindustry optimization models for economic forecasting]. *Interjekspos Geo-Sibir'*, 2017, vol. 3, no. 2, pp. 8-11. (In Russ.)

8. Melent'ev B.V. Prognozirovanie finan-

sovyh potokov na osnove mezhregional'nyh mezhotraslevykh modelej [Forecasting of cash flows based on interregional inter-industry models]. *Jekonomika i matematicheskie metody*, 2016, vol. 52, no. 3, pp. 50-64. (In Russ.)

9. Sazonov A.A., Sazonova M.V. Vyjavlenie svojstv jezkogennosti v modeljah prognozirovanija jekonomicheskogo rosta [Identification the properties of echtgenote in forecasting models of economic growth]. *Fundamental'nye i prikladnye issledovanija: problemy i rezul'taty*, 2016, no. 24, pp. 176-180. (In Russ.)

10. Filippov A.I. Problemy jekonomicheskogo prognozirovanija s pomoshh'ju modelej avtoregressii [Problems of economic forecasting with autoregressive models]. *Konkurentosposobnost' v global'nom mire: jekonomika, nauka, tehnologii*, 2016, no. 7-2 (19), pp. 157-162. (In Russ.)

11. Chadrancev A.V. *Opredelenie prognoziruемости jekonomicheskikh processov* [Determination of the predictability for economic processes]. Stavropol'skij gos. universitet, Stavropol', 2005. (In Russ.)