МЕТОДИЧЕСКАЯ РАЗРАБОТКА
для проведения лекции

Занятие № 6: «Моделирование процессов на основе прогнозирования временных рядов»

Учебные вопросы:

1. Общие подходы к прогнозированию временных рядов. Классификация методов прогнозирования

2. Обоснование выбора модели временного ряда

Заключительная часть

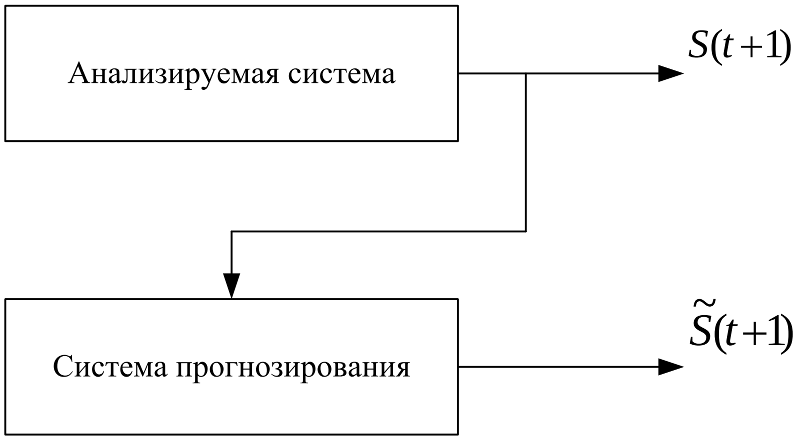
1. Общие подходы к прогнозированию временных рядов. Классификация методов прогнозирования

В ходе решения разнородных научно-практических задач полезным инструментом исследований могут выступать математические модели прогнозирования. При этом для увеличения точности прогнозов в изменяющихся условиях, в условиях неопределенности или неполной информации необходима работа по совершенствованию моделей. Важную роль в этом играют адаптивные методы прогнозирования. Отличие адаптивных моделей от других прогностических моделей состоит в том, что они отражают текущие свойства ряда и способны непрерывно учитывать эволюцию динамических характеристик изучаемых процессов. Цель адаптивных методов заключается в построении самокорректирующихся (самонастраивающихся) математических моделей, которые способны отражать изменяющиеся во времени условия, учитывать информационную ценность различных членов временной последовательности и давать достаточно точные оценки будущих членов данного ряда. Именно поэтому такие модели предназначаются, прежде всего, для краткосрочного прогнозирования.

На временной ряд воздействуют в разное время различные факторы. Одни из них по тем или иным причинам ослабляют свое влияние, другие воздействуют активнее. Таким образом, реальный процесс протекает в изменяющихся условиях, составляющих его внешнюю среду, с которой он взаимодействует, что и отражает временной ряд. При этом модель адаптируется к ряду, характеризующему реальный процесс. В силу того, что анализируемые ряды в большинстве своем являются нестационарными, модель всегда будет находиться в движении.

Таким образом, говоря о классификации методов прогнозирования временных рядов, следует выделять *неадаптивные (полуадаптивные)* методы прогнозирования и *адаптивные*.

Строгой границы между адаптивными и неадаптивными методами прогнозирования не существует. Обобщенная схема системы **полуадаптивного** прогнозирования имеет следующий вид.



Обобщенная схема системы полуадаптивного прогнозирования

При этом даже неадаптивные методы прогнозирования содержат некоторый элемент адаптации. Например, рассмотрим метод экстраполяции регрессионных кривых, когда с каждым новым получением фактических данных параметры регрессионных кривых пересчитываются, уточняются. Через достаточно большой промежуток времени может быть заменен даже тип кривой. Однако здесь степень адаптации весьма незначительна, к тому же с течением времени она падает вместе с увеличением общего количества наблюденных точек и соответственно с уменьшением в выборке удельного веса каждой новой точки.

Модификацией этого метода является метод кусочно-линейной аппроксимации, использование которого ведет к уменьшению «памяти» модели, к «забыванию» старых данных и построению линии регрессии на искусственно ограниченном количестве информации. Этот метод лучше учитывает новые тенденции, быстрее приспосабливается к изменившимся характеристикам процесса, но зато сильнее реагирует на помехи, случайные отклонения и искажения в связи с уменьшением доли «наследственности». Соотношение между «изменчивостью» и «наследственностью» в кусочно-линейном варианте регрессионного анализа определяется субъективным выбором интервалов аппроксимации. Недостатком является также то, что ценность информации в пределах интервала аппроксимации считается одинаковой независимо от возраста, а вне его пределов скачком падает до нуля.

Ценность информации в зависимости от возраста можно учесть с помощью геометрически убывающих весовых коэффициентов (взвешенная регрессия).

В частности можно ввести веса для квадратов ошибок. В этом случае совокупность весовых коэффициентов представляет собой функцию ценности информации от времени. Тогда параметры a_i регрессионных кривых будут отыскиваться из условия

$$\min \sum_i \beta_i [s_{T-i} - \tilde{s}_{T-i}(a_1, a_2, \ldots, a_n)]^2, \ 0 < \beta < 1,$$

где T - текущий момент времени;

s_{T-i} - фактическое значение процесса в момент $T-i$;

$\tilde{s}_{T-i}(a_1, a_2, \ldots, a_n)$ - значение подбираемой функции в момент $T-i$.

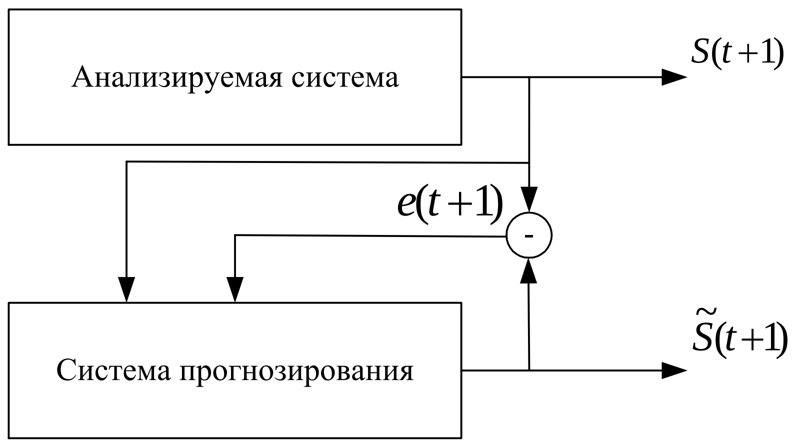
Пусть веса ошибок для более ранних моментов времени уменьшаются по закону убывающей геометрической прогрессии. Тогда функция ценности информации будет иметь экспоненциальную форму в отличие от прямоугольной в случае простого или кусочно-линейного метода построения регрессионных кривых. Такое взвешивание ошибок обеспечивает лучший подгон регрессионной кривой к более свежим данным.

Выбор величин β_i зависит, прежде всего, от характера моделируемого процесса, от его динамических свойств и статистических характеристик. Во многом выбор β_i зависит от опыта исследователя. Наилучшие их значения можно установить экспериментально методом проб.

Положительным свойством моделей такого вида является их способность лучше приспосабливаться к динамике процесса. Следует подчеркнуть, что если при выборе типа кривой регрессии была допущена ошибка, то в данной модели она обычно частично компенсируется.

Однако такой подход обладает рядом недостатков, главными из которых являются субъективность выбора структуры модели, ее чрезмерная жесткость и в силу этого автокоррелированность остатков. Громоздкий пересчет параметров регрессионных кривых с получением каждой новой точки требует повторного вовлечения в вычислительную процедуру всего объема информации.

Адаптивное моделирование позволяет в известной мере избавиться от представленных недостатков. Обобщенная схема системы адаптивного прогнозирования имеет следующий вид.



Обобщенная схема системы адаптивного прогнозирования

Основу адаптивного направления составляет простейшая модель экспоненциального сглаживания. Модификации и обобщения этой модели привели к появлению целого семейства адаптивных моделей с различными свойствами.

Основные особенности моделей указанного типа состоят в следующем. Прежде всего, все рассматриваемые модели имеют жесткий алгоритм поведения. Вместе с тем, адаптация в данных моделях слагается из небольших дискретных сдвигов. В основе процедуры адаптации лежит метод проб и ошибок, который считается универсальным путем выработки нового поведения.

Последовательность процесса адаптации в обобщенном виде выглядит следующим образом. Пусть модель находится в некотором исходном состоянии (определены текущие значения ее коэффициентов) и по ней делается прогноз. По истечении шага моделирования проводится анализ, насколько далек результат, полученный с использованием модели, от фактического значения ряда. Ошибка прогнозирования через обратную связь поступает на вход системы и используется моделью в соответствии с ее логикой для перехода из одного состояния в другое с целью большего согласования своего поведения с динамикой ряда. На изменения ряда модель должна отвечать «компенсирующими» изменениями. Затем делается прогноз на следующий момент времени, и весь процесс повторяется.

Таким образом, адаптация осуществляется итеративно с получением каждой новой фактической точки ряда. При этом в рамках различных моделей используется отличающийся механизм адаптации. Вопрос выбора этого механизма решается каждым исследователем интуитивно. «Принцип работы» механизма адаптации задается априорно, а затем проверяется эмпирически.

Быстроту реакции модели на изменения в динамике процесса характеризует параметр адаптации. Процесс «обучения» модели состоит в выборе наилучшего параметра адаптации на основе проб на ретроспективном материале. Например, воздействие белого шума должно вызывать у модели адекватную оборонительную реакцию, проявляющуюся в том, что модель не адаптируется к белому шуму, а должна отфильтровывать его. При наличии некоторых закономерностей с случайном процессе наилучшей реакцией модели является определенный компромисс между двумя крайними ситуациями, обеспечивающий отражение присутствующих закономерностей и одновременно фильтрацию случайных отклонений. После выбора параметра адаптации самообучение модели происходит в процессе переработки новых статистических данных.

В силу простоты каждой отдельно взятой модели и ограниченности исходной информации, зачастую представленной единственным рядом, нельзя ожидать, что какая-либо одна адаптивная модель годится для прогнозирования любого ряда, любых вариаций поведения, что говорит об отсутствии некоторой универсальной модели. Поэтому при обосновании выбора конкретной модели необходимо учитывать наиболее вероятные закономерности развития реального процесса, динамические свойства ряда соотносить с возможностями выбранной модели. Вместе с тем, нельзя надеяться на успешную самоадаптацию модели, более общей по отношению к той, которая необходима для отражения данного процесса, поскольку увеличение числа параметров придает системе излишнюю чувствительность, приводит к ее «раскачке» и ухудшению получаемых по ней прогнозов.

Таким образом, при построении адаптивной модели приходится выбирать между общей и частной моделью и, взвешивая их достоинства и недостатки, отдавать предпочтение той, от которой можно ожидать наименьшей ошибки прогнозирования. Только при этом условии можно надеяться, что последовательность проб и ошибок постепенно приведет к наиболее эффективному прогнозированию.

Для сравнения возможных альтернатив необходим критерий полезности модели. В самом общем случае такой критерий сам является предметом исследований, однако в случае краткосрочного прогнозирования признанным критерием обычно является средний квадрат ошибки прогнозирования. О качестве модели судят также по наличию автокорреляции в ошибках.

Таким образом, обобщая сказанное выше, следует отметить, что к основным задачам моделирования временных рядов и прогнозирования их параметров относятся: выбор подходящей параметрической модели временного ряда - временной последовательности значений некоторой случайной величины, оценивание ее параметров, а также получение формул для прогноза параметров ряда (экстраполяции значений случайной величины на упреждающий момент времени). Важным моментом в моделировании временного ряда является то, что он, в отличие от классической модели случайного процесса, несет информацию обо всей генеральной совокупности его значений, является лишь одной из ее выборок. Поэтому для обеспечения состоятельности, достаточности и эффективности результатов обработки этих статистических данных необходимо выполнение для временного ряда условий стационарности и эргодичности, что достигается при выборе времени обработки много меньшем чем интервал корреляции значений ряда.

Под этапом выбора модели понимается обоснование некоторого класса стохастической модели и идентификации ее параметров на основе знания автокорреляционных и частных корреляционных функций элементов временного ряда, а также метода диагностической проверки модели. Далее проводится этап оптимального прогнозирования параметров потока требований на основе рекуррентных вычислений. Качество прогноза определяется временем упреждения и точностью. Время упреждения, в свою очередь, определяется временем запаздывания в принятии решений по изменению плана распределения ресурсов, а точность прогноза определяется вероятностью достижения заданной погрешности прогнозирования.

2. Обоснование выбора модели временного ряда

Вопрос построения, оценки и использования моделей прогнозирования является одним из ключевых в рамках анализа временных рядов.

Если имеется достаточно длинный отрезок ряда (50 и более точек) и принято решение искать адекватную модель, то обоснование и выбор модели к имеющимся данным лучше всего достигаются с помощью трехстадийной итеративной процедуры, включающей идентификацию, оценку и диагностическую проверку модели. После этого модель можно использовать для прогнозирования.

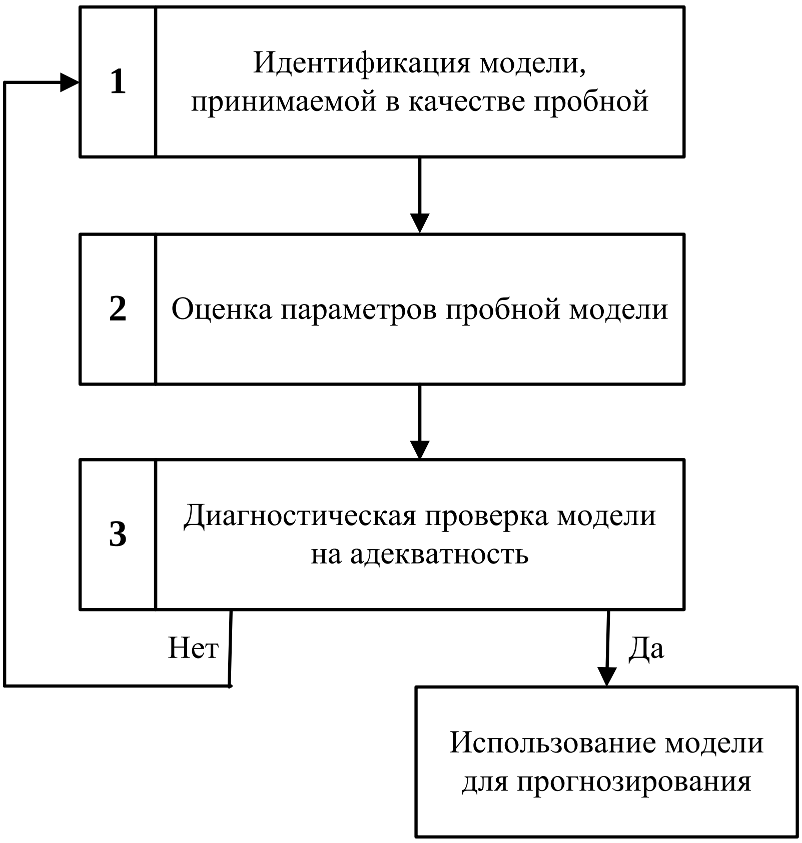
Под идентификацией имеется в виду использование накопленных данных и любой другой информации для определения подкласса моделей, оптимальных с точки зрения настраиваемых параметров, среди которых следует искать адекватную.

Под оценкой понимается эффективное использование данных для получения численных значений параметров модели при предположении ее адекватности процессу.

Диагностическая проверка имеет целью анализ адекватности подобранной модели и ее улучшение.

Для оценивания степени адекватности выбранной модели временного ряда анализируемому процессу на основе полученной выборки длины n наибольшее распространение получил критерий минимума среднеквадратической ошибки прогнозирования

$$M[e^2(t+1)] = \frac{1}{n}\sum_{t=0}^{n-1}[s(t+1) - \tilde{s}(t+1)]^2 . \qquad (24)$$



Методика выбора модели временного ряда

Следящий контрольный сигнал

Для скорейшего обнаружения неадекватности модели реальному процессу, что необходимо для внесения соответствующих изменений в модель прогнозирования, разработан способ анализа прогнозирующей системы, состоящей в подсчете величины следящего контрольного сигнала. Следящий контрольный сигнал K_t определяется как сумма ошибок прогнозирования $e(t)$, деленная на величину их сглаженного абсолютного значения

$$\tilde{e}(t) = (1-\gamma)\tilde{e}(t-1) + \gamma|e(t)|$$

т.е.

$$K_t = \frac{\sum_{i=0}^{t} e(i)}{(1-\gamma)\tilde{e}(t-1) + \gamma|e(t)|}, \tag{25}$$

где $0 < \gamma < 1$ - постоянная сглаживания.

Такой подход имеет два недостатка. Во-первых, в случае когда контрольный сигнал вышел за установленные пределы, он не обязательно вернется в эти же пределы, даже если рассматриваемый процесс вновь будет развиваться по прежним законам и прогнозирующая модель окажется адекватной реальному процессу. Следовательно, потребуется вмешательство, чтобы сделать сумму ошибок снова близкой к нулю и таким образом избежать ложных сигналов тревоги.

Во-вторых, возможна и другая ситуация, когда контрольный сигнал выходит из указанных пределов, а система начинает давать более точные прогнозы. Например, если с некоторого момента имеют место совершенные прогнозы, то среднее абсолютное отклонение будет стремиться к нулю, в то время как сумма ошибок останется неизменной. Таким образом, контрольный сигнал будет стремиться к бесконечности.

С учетом этих обстоятельств была предложена модификация рассмотренного правила, в рамках которого вместо суммы ошибок используется сглаженная ошибка

$$\hat{e}(t) = (1-\gamma)\hat{e}(t-1) + \gamma e(t). \tag{26}$$

Следящий контрольный сигнал K_t определяется отношением

$$K_t = \frac{\hat{e}(t)}{\tilde{e}(t)}.$$

Если прогнозирующая система окажется настолько неадекватной изучаемому процессу, что все ошибки будут одного знака, то контрольный сигнал будет стремиться к +1 или -1. Если известно, что прогнозирующая система адекватна реальному процессу и полученные ошибки образуют неавтокоррелированную, нормально распределенную случайную последовательность с нулевым средним и стандартным отклонением σ, то для контрольного сигнала могут быть определены доверительные интервалы.

Уравнение для сглаженной ошибки можно переписать в виде

$$\hat{e}(t) = \sum_{i=0}^{t} \gamma(1-\gamma)^{t-i} e(t-i), \tag{27}$$

ее дисперсия будет равна сумме дисперсий отдельных составляющих

$$\tilde{\sigma}^2[\hat{e}] = \sum_{i=0}^{t} (\gamma(1-\gamma)^{t-i}\sigma[e(i)])^2. \tag{28}$$

Так как $0 < (1-\gamma) < 1$, то при $t \to \infty$ этот ряд сходится и существует предельный переход

$$\tilde{\sigma}^2[\hat{e}] = \frac{\gamma^2\sigma^2[e(i)]}{1-(1-\gamma)^2} = \frac{\gamma\sigma^2[e(i)]}{2-\gamma}. \tag{29}$$

Пределы для сглаженной ошибки, определяемые величиной $2\tilde{\sigma}[\hat{e}]$, равны, поэтому

$$\pm 2\sigma[e(i)]\sqrt{\frac{\gamma}{2-\gamma}}\,.$$

Известно, что СКО примерно равно 1,2 среднего абсолютного отклонения. Если γ достаточно мало, то можно принять, что локальная оценка среднего абсолютного отклонения относительно постоянна и приблизительно равна истинному среднему абсолютному отклонению, т.е. равна $\frac{\sigma}{1{,}2}$.

Таким образом, при малых γ, при принятых допущениях относительно ошибки прогнозирования величина $\hat{e}(t)$ является случайной, нормально распределенной, а $\tilde{e}(t)$ приблизительно постоянной величиной. Это дает возможность, несмотря на то, что $-1 \le K_t \le +1$, аппроксимировать распределение K_t нормальным законом.

Следовательно, пределы для контрольного сигнала, определяемые величиной $2\sigma_K$, приблизительно равны

$$\pm 2{,}4\sqrt{\frac{\gamma}{2-\gamma}}\,.$$

Для $\gamma = 0{,}1$ получим $\pm 0{,}55$, т.е. с вероятностью 0,95

$$-0{,}55 \le K_t \le +0{,}55\,.$$

Пределы в $3\sigma_K$ (соответствующие вероятности 0,99) при этом же значении γ составят

$$-0{,}83 \le K_t \le +0{,}83\,.$$

Для значений γ, которые не очень малы, эти рассуждения теряют справедливость, и доверительные интервалы целесообразно получать другими методами (например, Монте-Карло).

Таким образом, следящий контрольный сигнал является мерой неадекватности модели реальному процессу. При превышении контрольным сигналом заданного доверительного уровня (обычно $2\sigma_K$) целесообразно пересмотреть модель и заменить ее другой.

Заключительная часть

– напомнить изученные вопросы и цели занятия;

– подвести итоги занятия, определить полноту достижения целей занятия.

Рекомендованная литература:

1. Лукашин Ю.П. Адаптивные методы краткосрочного прогнозирования временных рядов. – М.: Финансы и статистика, 2003.

2. Статистические методы обработки результатов наблюдений / под ред. Р.М. Юсупова. – Л.: МО СССР, 1984.