2.2. Методы адаптации моделей среднесрочного прогнозирования дисконтированием данных

Успех адаптивных методов краткосрочного прогнозирования в экономике породил желание прогнозистов трансформировать принципы этих методов и на задачи прогнозирования на среднесрочную перспективу.

Однако непосредственно метод Брауна или его модификации применить к решению это задачи не получается – метод Брауна нацелен на адаптацию модели к текущим краткосрочным изменениям, а при среднесрочном прогнозировании необходимо адаптировать модель к намечающимся отклонениям от тенденции, которые будут действовать более длительный срок.

Действительно, при нахождении оптимального значения постоянной сглаживания выполняется процедура ретропрогноза, когда с помощью модели Брауна:

$$\hat{Y}_{T+1} = \alpha Y_T + (1-\alpha)\hat{Y}_T$$

На основе текущих значений ряда вычисляются следующие значения, которые сравниваются с реальными данными, имеющимися в распоряжении прогнозиста, то есть, вычисляется ошибка ретропрогноза. Выбирается то значение постоянной сглаживания, для которой дисперсия ошибки ретропрогноза минимальна (либо иной критерий точности ретропрогноза). Модель, как видно, будет наилучшим способом прогнозировать следующее прогнозное значение на шаге *T+1*. А если прогнозисту необходимо выполнить прогноз не на один шаг вперёд, а, например, на десять, то есть – спрогнозировать показатель в момент времени *T+10*? Действительно, для такой цели метод Брауна представляется непригодным. Но на самом деле есть простой способ использования метода Брауна для поставленной задачи.

Поскольку метод Брауна применим для процессов не имеющих тенденцию к росту или снижению, то есть – он использует способ вычисления скользящей средней, то можно с помощью этого метода оценить скользящую среднюю и через два, и через три, и через десять наблюдений. Но для этого необходимо определить оптимальное значение постоянной сглаживания.

Пусть перед прогнозистом стоит задача выполнить прогноз не на один шаг вперёд, а на два шага. Модель Брауна в этом случае будет выглядеть так:

$$\hat{Y}_{T+2} = \alpha Y_T + (1-\alpha)\hat{Y}_T . \qquad (2.2.1)$$

Выбор оптимального значения постоянной сглаживания осуществляется следующим образом. Задаётся некоторое значение постоянной сглаживания, равное, например, *0,1*. Подставляя это значение в (2.2.1), рассчитываются прогнозные значения ряда на всей имеющейся у прогнозиста базе. От фактических значений отнимаются расчётные, тем самым вычисляются ошибки ретропрогноза. После того, как расчёты выполнены по всей базе, вычисляется некоторая обобщённая оценка точности прогноза методом Брауна при этой постоянной сглаживания. Затем задаётся другое значение постоянной сглаживания, например, *0,2*. Вновь запускается процедура ретропрогноза при этой постоянной сглаживания и вычисляется обобщённая характеристика точности прогноза модели на два шага вперёд. Проведя расчёты в допустимой области изменения постоянной сглаживания от нуля до двух, выбирается то значение, при котором достигается наивысшая точность ретропрогноза. Обозначим для определённости это наилучшее значение постоянной сглаживания как α_2. Тогда, подставляя это значение постоянной сглаживания, с помощью метода Брауна на имеющемся множестве делается прогноз на два наблюдения вперёд.

Если перед прогнозистом стоит задача выполнить прогноз на три наблюдения, то модель Брауна будет для такого случая иметь вид:

$$\hat{Y}_{T+3} = \alpha Y_T + (1-\alpha)\hat{Y}_T . \qquad (2.2.2)$$

Здесь оптимальное значение постоянной сглаживания находится точно также, как это было описано выше. Найдя это оптимальное значение постоянной сглаживания для прогноза на три шага вперёд α_3, прогнозист легко может по имеющимся у него значениям выполнить прогноз показателя на три наблюдения вперёд.

И вообще, если перед прогнозистом стоит задача выполнить прогноз на τ наблюдений вперёд, то модель Брауна будет для этого случая записана так:

$$\hat{Y}_{T+\tau} = \alpha Y_T + (1-\alpha)\hat{Y}_T . \qquad (2.2.3)$$

Для использования этой модели необходимо описанным выше способом найти оптимальное значение постоянной сглаживания α_τ, после чего и выполняется прогноз на заданный период упреждения.

Следует отметить, что оптимальные значения постоянных сглаживания для прогнозирования методом Брауна на различные периоды упреждения отличаются друг от друга, то есть:

$$\alpha_1 \neq \alpha_2 \neq \alpha_3 \neq \alpha_\tau . \qquad (2.2.4)$$

Этот подход распространим и на различные модификации метода Брауна, рассмотренные в предыдущей главе, если их применить для задач среднесрочного прогнозирования.

Как видно, при таком подходе принцип адаптивного прогнозирования: «текущая информация более ценна для прогнозирования, чем прошлая», полностью выполняется. Для среднесрочного прогнозирования при таком использовании метода Брауна выполняется дисконтирование прошлых данных.

Желание совместить аппарат математической статистики с новым подходом к прогнозированию, открытым Брауном, привёл к тому, что к началу 70-х годов XX века появился метод дисконтирования оценок МНК, который позволяет при оценивании параметров моделей учесть текущую информацию в большей степени, чем прошлую и приспособить модель к более поздним данным и использовать при дисконтировании веса, заданные по методу Брауна. Делается так.

Критерий МНК, как известно, имеет вид:

$$Q = \sum_{t} \left(Y_t - \hat{Y}_t \right)^2 \to \min . \qquad (2.2.5)$$

В соответствии с ним, находятся такие оценки прогнозной модели, при которых минимизируется сумма квадратов отклонений фактических значений от расчётных. Но, как уже отмечалось, для прогнозиста в случае прогноза эволюционно протекающих процессов важнее более точно описать последние наблюдения, нежели те, которые убывают в прошлое. Поэтому и ошибка аппроксимации последних наблюдений должна минимизироваться в большей степени, чем ошибки аппроксимации в начале ряда. Логично, поэтому задать этим ошибкам аппроксимации некоторые веса v_t так, чтобы их значения уменьшались с убыванием наблюдений в прошлое:

$$v_T > v_{T-1} > \ldots > v_t > \ldots > v_1 . \qquad (2.2.6)$$

Веса могут задаваться в числовой форме или в виде функциональной зависимости, но так, чтобы по мере продвижения в прошлое веса убывали.

Для удобства часто вводят дополнительное условие:

$$\sum_{t=1}^{T} v_t = 1 , \qquad (2.2.7)$$

но его выполнение, в отличие от ситуации метода Брауна, не является обязательным.

Возможно два варианта дисконтирования оценок МНК.

Первый вариант - когда взвешивается каждая ошибка аппроксимации, и эта взвешенная величина подставляется в сумму квадратов МНК. Тогда критерий взвешенного МНК будет иметь вид:

$$Q = \sum_{t} (v_t \left(Y_t - \hat{Y}_t \right))^2 \to \min . \qquad (2.2.8)$$

Применение этого критерия, например, для простой однофакторной линейной модели приведёт к необходимости решать такую систему уравнений:

$$\begin{cases}\sum_t v_t^2 Y_t = a_0 \sum_t v_t^2 + a_1 \sum_t v_t^2 x_t, \\ \sum_t v_t^2 Y_t x_t = a_0 \sum_t v_t^2 x_t + a_1 \sum_t v_t^2 {x_t}^2. \end{cases} \tag{2.2.9}$$

Такой способ взвешивания данных о наблюдении при построении адаптированных моделей используют не очень часто, поскольку в полученной системе уравнений не ясен смысл взвешивания различных сумм, поскольку веса везде представлены квадратами.

Значительно чаще используется другой метод взвешивания, а именно, взвешиваются не сами ошибки аппроксимации, а их квадраты. Тогда критерий МНК, взвешенного таким способом, имеет вид:

$$Q = \sum_t v_t \left(Y_t - \hat{Y}_t\right)^2 \to \min . \tag{2.2.10}$$

Использование этого критерия, например, для линейной однофакторной модели приведёт к необходимости решения системы двух таких уравнений:

$$\begin{cases}\sum_t v_t Y_t = a_0 \sum_t v_t + a_1 \sum_t v_t x_t, \\ \sum_t v_t Y_t x_t = a_0 \sum_t v_t x_t + a_1 \sum_t v_t {x_t}^2. \end{cases} \tag{2.2.11}$$

В этом случае уже появляется возможность смыслового толкования уравнений системы. Левая часть первого уравнения, как легко заметить, при выполнении равенства (2.2.7) означает вычисление взвешенной средней переменной Y_t, а второе слагаемое этого же уравнения представляет собой произведение коэффициента a_1 на взвешенную среднюю переменной x_t. В качестве дополнительного «плюса» такого метода взвешенного МНК является и то обстоятельство, что в правой части первого уравнения системы коэффициент a_0 умножается на сумму весов, а она в силу (2.2.7) равна единице. Поэтому система (2.2.11) ещё и упрощается.

Как задавать веса v_t? Ответ на этот вопрос напрашивается сам собой – так, как это сделано для краткосрочного прогнозирования Брауном, то есть:

$$v_T = \alpha, v_{T-1} = \alpha(1-\alpha), \ldots v_t = \alpha(1-\alpha)^{T-t}, \ldots \tag{2.2.12}$$

Вспомним, что этот способ задания весов позволяет получить взвешенную среднюю:

$$\hat{Y}_{T+1} = \alpha Y_T + (1-\alpha)\hat{Y}_T = \overline{Y}_{T+1}, \tag{2.2.13}$$

которая в краткосрочном прогнозировании используется как лучшая прогнозная оценка данного показателя Y на шаг вперёд, то есть:

Этот же способ взвешивания, как легко убедиться из первого уравнения системы (2.2.11) применяется и для факторной переменной x_t:

$$\hat{x}_{T+1} = \alpha x_T + (1-\alpha)\hat{x}_T = \overline{x}_{T+1}, \qquad (2.2.14)$$

Поскольку для метода Брауна сумма весов равна единице, подставляя такой способ взвешивания наблюдений в (2.2.11), и используя обозначения средней взвешенной, получим:

$$\begin{cases} \overline{Y}_{T+1} = a_0 + a_1\overline{x}_{T+1}, \\ \overline{Y_{T+1}x_{T+1}} = a_0\overline{x}_{T+1} + a_1\overline{x}^2_{T+1}, \end{cases} \qquad (2.2.15)$$

где: $\overline{Y_{T+1}x_{T+1}} = \alpha Y_T x_T + (1-\alpha)\overline{Y_T x_T}$, (2.2.16)

$$\overline{x}^2_{T+1} = \alpha x^2_T + (1-\alpha)\overline{x}^2_T. \qquad (2.2.17)$$

Такой способ более интересен и удобен, как видно не только тем, что позволяет построить адаптированную к последним наблюдениям модель, но и при появлении новых наблюдений $t=T+1$ легко пересчитывать коэффициенты модели.

Для успешного применения с помощью весов метода Брауна взвешенного МНК необходимо оптимизировать постоянную сглаживания α. Эта задача решается также, как и в случае краткосрочного прогнозирования методом Брауна.

Поскольку критерий (2.2.10) является универсальным для использования оценок взвешенного МНК, его можно использовать для оценивания коэффициентов нелинейных и многофакторных прогнозных моделей.

Теперь воспользуемся выводами первого параграфа этой главы для того, чтобы понять суть взвешенного МНК. Систему уравнений (2.2.11) можно получить с помощью общей схемы оценивания методом z-множителей (2.1.11) если задать z-множители так:

$$\begin{cases} z_{ot} = v_t, \\ z_{1t} = v_t x_t. \end{cases} \qquad (2.2.18)$$

Это означает, что решая систему (2.2.11) мы получаем такие оценки коэффициентов прогнозной модели, для которой выполняются условия (2.1.10), которые применительно к рассматриваемому случаю будут иметь вид:

$$\begin{cases} \sum_t v_t \varepsilon_t = 0, \\ \sum_t v_t \varepsilon_t x_t = 0. \end{cases} \qquad (2.2.19)$$

Смысл первого уравнения системы (2.2.19) очевиден, поскольку веса заданы по методу Брауна и они убывают в прошлое – прогнозная модель будет описывать исходный ряд данных так, что ошибки аппроксимации, убывающие в прошлое, будут больше, чем ошибки аппроксимации последних наблюдений. При этом модель обязательно будет иметь как положительные, так и отрицательные ошибки аппроксимации, иначе сумма взвешенных ошибок аппроксимации не будет равна нулю. Модель, как следует из сказанного, хорошо описывает текущие наблюдения и плохо – прошлые.

Смысл второго уравнения системы (2.2.19) менее ясен. Будет равна нулю сумма взвешенных произведений фактических значений фактора на ошибки аппроксимации.

Но метод z-множителей не только позволяет получить дополнительное толкование оценкам взешенного МНК, но и, используя общий принцип учёта текущих наблюдений в большей степени, чем более ранние, получить новые оценки. Например, можно задать такие z-множители:

$$\begin{cases} z_{ot} = v_t, \\ z_{1t} = v_t Y_t. \end{cases} \qquad (2.2.20)$$

Тогда будет получена такая система уравнений:

$$\begin{cases} \sum\limits_t v_t Y_t = a_0 \sum\limits_t v_t + a_1 \sum\limits_t v_t x_t, \\ \sum\limits_t v_t Y_t^2 = a_0 \sum\limits_t v_t Y_t + a_1 \sum\limits_t v_t x_t Y_t. \end{cases} \qquad (2.2.21)$$

Решая эту систему, прогнозист получит оценки адаптировной модели – ведь текущая информация используется в большей степени, чем прошлая, но эти оценки будут отличаться от оценок взвешенного МНК и, возможно, в некоторых случаях будут давать более точные прогнозы. Ряд различных способов дисконтирования данных, который открывает метод z-множителей, довольно широк. Это вооружает прогнозиста новым дополнительным инструментом построения адаптивных моделей среднесрочного прогнозирования.

Недостатком этого способа получения оценок прогнозной модели с помощью взвешенного МНК является показательный способ задания весов ошибок аппроксимации. Зачастую в экономике встречаются ситуации, когда наилучший способ задания весов не соответствует правилу (2.2.12), например, когда после некоторого периода использования новых технологий предприятие возвращается по различным причинам к старым технологиям. Тогда вес наблюдений в прошлом, когда использовались старые технологии, не будут менее важны для прогнозирования, чем текущие. В этом случае возможен и такой способ задания весов:

$$v_T > v_{T-1} > \ldots = v_t = v_{t-1} > \ldots > v_1 . \qquad (2.2.14)$$

Это тем более важно для случаев, когда прогнозные ряды имеют лаги – задержки во времени. Например, инвестиции в новые технологии дадут отдачу только через период, связанный с внедрением инноваций и началом их использования. В этой ситуации вес наблюдения в год инвестиций для прогноза может быть больше, чем вес наблюдения в последний год. Метод *z*-множителей позволяет устранить этот недостаток, подбирая к каждому ряду свой наилучший способ задания множителей, но эта задача не формализована, поэтому широко её использовать на практике нельзя – её успех зависит от личных качеств прогнозиста.