

## Глава вторая. Среднесрочное прогнозирование эволюционных социально-экономических процессов

### ***2.1. Общая схема оценивания коэффициентов прогнозных моделей***

Если задачей краткосрочного прогнозирования эволюционных процессов является приспособление модели к краткосрочно действующим отклонениям от общей тенденции, то задача прогнозирования среднесрочного прогнозирования эволюционной динамики заключается в том, чтобы, уловив наметившиеся отклонения от общей тенденции, вызванные адаптацией социально-экономического объекта к изменившимся внешним и внутренним условиям, оценить будущую динамику с учётом этих отклонений. Задача эта не очень проста.

Действительно, динамика движения, например, Земли и других планет вокруг Солнца стационарна – траектория этого движения, отмеренная во времени, не меняется. Более того, если перед прогнозистом стоит задача определить положение в солнечной системе, например, Земли через четыре месяца, то для этого он постарается собрать как можно больше данных о том, как менялось расстояние от Земли до Солнца  $Y_t$  для того, чтобы по этим данным построить точную модель движения Земли во времени и в пространстве  $Y_t=f(t)$ .

При этом очевидно правило – чем длиннее ряд наблюдений  $\{Y_t\}$ , тем точнее будет построена модель движения Земли вокруг Солнца и тем точнее будет оценка коэффициентов модели. Сто наблюдений для этой задачи будет предпочтительнее, чем пятьдесят; тысяча наблюдений – ещё предпочтительнее, чем сто; десять тысяч наблюдений позволят построить вполне адекватную прогнозную модель.

Валовой внутренний продукт России – тоже меняется во времени. Его чисто формально также можно представить как некоторый динамический ряд  $\{Y_t\}$ . Но для того, чтобы построить модель тренда этого показателя и оценить значения его коэффициентов, вряд ли кто из прогнозистов бросится собирать данные о ВВП России за весь XX, XIX, XVIII и XVII век в стремлении построить наиболее точную прогнозную модель. Даже данные за 1997 год мало пригодны для прогноза ВВП России в настоящее время, поскольку дефолт 1998 года существенно изменил экономическую систему России – высокая цена иностранной валюты после 1998 года сделала конкурентоспособными по цене отечественные товары, спрос на которые со стороны россиян послужил стартом роста производства.

К 2000 году проявилась во всей силе и другая тенденция – рост цен на топливно-энергетические ресурсы, что для ресурсоориентированной экономики России послужило ещё одним фактором роста экономики страны и ВВП – как её основного показателя.

Поэтому тенденции роста ВВП с 1999 года отражают другое состояние экономики, нежели тенденции в предыдущие годы. Не случайно с 2002 года как заклинания звучали с высоких трибун российской государственности

слова – «удвоить к 2010 году ВВП России по сравнению с 2000 годом». Если проанализировать сложившуюся к тому моменту тенденцию последних трёх-четырёх лет, и описать её различными трендами, с их помощью получался оптимистический прогноз – к 2010 году ВВП России становится в два раза больше ВВП в 2000 году. Другое дело, что аналитики, выполнявшие подобные прогнозы забыли, что имеют дело с эволюционными необратимыми процессами, что тенденции, сложившиеся и длящиеся определённый промежуток времени, в скором времени под влиянием внешних и внутренних факторов будут меняться. И действительно, вместо обещанного удвоения ВВП России к 2010 году страна находилась в состоянии экономического кризиса, поразившего экономику не только России, но и других стран мира.

Таким образом, главный принцип выборочного метода – чем больше собрано информации о прошлом состоянии объекта, тем лучше, поскольку можно точнее оценить коэффициенты модели, для необратимых процессов абсолютно не пригоден. В этом случае необходимо собрать информацию о динамике объекта только за тот период, когда сам объект не успел изменить свои основные свойства, когда наблюдаемые количественные наблюдения не привели объект к новому качеству. То есть, необходимо предварительно определить период инерционности объекта социально-экономического прогнозирования, а затем, ориентируясь на продолжительность этого периода, сформировать базу данных. К сожалению, формальных процедур, решающих эту задачу, пока не разработано. Приходится предварять сбор статистики серьёзным экономическим анализом (фундаментальным анализом), опираясь на опыт и интуицию экспертов.

Поскольку построенную в такой ситуации модель ни в коем случае нельзя воспринимать как одно из приближений к математическому ожиданию процесса (его не существует), а следует понимать как некоторое описание тенденции, действовавшей в определённый промежуток времени, то требования к оценкам таких моделей существенно смягчаются.

Если при прогнозировании обратимых процессов, в том числе и динамических, наилучшими оценками коэффициентов модели с позиций её приближения к математическому ожиданию процесса являлось требование минимизации суммы квадратов отклонений расчётных значений от фактических, то для построения трендов эволюционных процессов это требование не имеет смысла. Вновь возникает задача построения наилучшей модели, но критерий, по которому можно выбрать лучшую модель, не может быть перенесён из задач прогнозирования стационарной экономики. Критерий выбора наилучшей модели перестаёт быть строго статистическим.

Поскольку наблюдаемый признак (показатель)  $y_t$  представляет собой результат сложного взаимодействия детерминированной составляющей  $y'_t$ , случайной составляющей  $\varepsilon_t$  и неопределённой составляющей  $\mu_t$ :

$$y_t = y'_t + \varepsilon_t + \mu_t, \quad (2.1.1)$$

успех прогноза определяется не тем, насколько хорошо описаны детерминированная и случайная компоненты, а тем, насколько верно оценено направление и сила изменения неопределённой составляющей. Но ведь неопределённая составляющая прогнозисту не известна, её влияние он может оценить только по результирующему наблюдаемому признаку  $y_t$ . Как видно, задача представляется весьма сложной.

Тривиальной теперь представляется задача прогнозирования обратимых процессов, поскольку для них обоснованно предполагается, что чем лучше выбранная модель описывает имеющуюся выборку с позиций минимума дисперсии ошибки аппроксимации, тем в большей степени она пригодна для прогнозирования. В случае прогнозирования эволюционирующих процессов, хорошее описание прошлого вовсе не является гарантией такого же хорошего прогнозирования будущего. Более того, как уже говорилось ранее – для целей предсказания будущего ценнее текущая информация, чем прошлая. Поэтому, выбрав вид модели, необходимо так оценить её коэффициенты, чтобы она лучше прогнозировала будущее, чем аппроксимировала прошлое. Но если МНК не является априорно наилучшим методом для оценки коэффициентов тренда, то что же может использовать прогнозист из арсенала выборочного метода - элементарные методы конечных разностей и средних? Такой выбор представляется маловатым.

Вопрос о том, каким образом найти коэффициенты выбранной модели, в теории прогнозирования применительно к необратимым процессам в полной мере не ставился. Принято использовать МНК, после чего различными процедурами корректируют полученные оценки. И использование МНК для оценивания коэффициентов модели ничем не обосновывается, кроме привычки, выработанной десятилетиями применения этого метода для статистической обработки различных экономических зависимостей. К тому же в распоряжении прогнозиста есть множество прикладных программ, в которых МНК «зашит» по определению и применяется по умолчанию.

На самом деле МНК является лишь одним из возможных и далеко не лучших методов оценки коэффициентов моделей необратимых процессов социально-экономической динамики. Даже в случае применения выборочного подхода обратимых процессов МНК является лучшим способом оценки коэффициентов модели в ситуации нормального распределения случайных величин. Для других распределений случайных величин лучшими будут являться другие методы – метод моментов, метод Ньютона, метод спейсингов или разнообразные методы непараметрической регрессии.

Конечно, МНК используется чаще всего при обработке статистических выборок, но это вовсе не говорит о том, что он – самый лучший метод. Просто чаще всего в случайных совокупностях обратимых процессов проявляются условия, для которых наилучшим будет именно МНК – множество случайных факторов действуют незначительно на статистический показатель, причём действуют разнонаправлено, поэтому результат этого действия вполне описывается нормальным распределением.

Какие же методы могут выступить в виде альтернативы МНК для решения задачи оценивания коэффициентов моделей, описывающих необратимые процессы? Ведь эти процессы многообразны и наперёд узнать, к какому типу они относятся, чтобы применить лучший метод (или набор методов), чаще всего нельзя. Нужен некоторый набор различных методов, в том числе и МНК для того, чтобы, испробовав каждый из них, с помощью процедуры ретропрогноза выбрать лучший для этого процесса метод. То есть, вместо априорного подхода, который характерен для задач математической статистики, следует использовать апостериорный подход, когда тщательно изучив особенности прогнозируемого процесса, испробовав несколько подходов, выбрать наилучший из них.

Здесь следует сказать несколько слов о самом апостериорном подходе и его сравнении с априорным. Настоящий прорыв в науке произошёл как раз после того, как на смену Аристотелевому дедуктивному выводу пришёл индуктивный вывод. В соответствии с ним учёный, обнаружив некоторую закономерность в исследуемом процессе, высказывает гипотезу о том, что эта закономерность присуща и для всех остальных подобных процессов, протекающих в аналогичных условиях. Такой вывод освобождал учёного от необходимости подкреплять свои выводы многочисленными и трудоёмкими натурными экспериментами и опытами, описывающими генеральную совокупность. Решающая роль в науке стала отводиться гипотезе – предварительному высказыванию о сути изучаемого процесса высказыванию, базирующемуся на научных принципах, подходах и знаниях, но носящему, всё же, характер возможности, но не истины. Наука стала строиться, опираясь на гипотетико-эмпирический вывод, более дешёвый и потому эффективный. В соответствии с его логикой, в начале исследования высказывается научная гипотеза, которая проверяется на имеющихся данных. Если гипотеза не отвергается, то по вероятности делается вывод о верности высказанной гипотезы.

Математическая статистика, позволяющая моделировать обратимые процессы, как раз и является математическим инструментом индуктивного вывода, инструментом априорного подхода. Она позволяет по нескольким наблюдениям судить о генеральной выборке в целом.

Апостериорный подход можно отнести к области Аристотелевого вывода. Он требует предварительного тщательного анализа процесса, вместо гипотетико-эмпирического вывода априорного подхода в данном случае равнодопустимыми являются все возможные гипотезы, которые вытекают из свойств эмпирических данных. Именно тщательное изучение свойств наблюдаемого объекта позволяет генерировать гипотезы о том, почему объект ведёт себя так, а не иначе. И только после того, как высказаны все возможные гипотезы, каждая из них проходит тщательную проверку. Это – эмпирико-дедуктивный вывод. Очевидно, что он значительно более трудоёмок, нежели гипотетико-эмпирический вывод. Но применительно к необратимым процессам априорное предположение о характере изучаемого

процесса сделать невозможно – свойства необратимых процессов меняются со временем.

Но поскольку современная наука вооружена вычислительной техникой, позволяющей легко решать ранее считавшиеся трудоёмкими задачи, то и апостериорный подход может использоваться наряду с априорным, в том числе и в задачах социально-экономического прогнозирования.

При построении прогнозных моделей необратимых процессов прогнозист не может утверждать о том, какая модель лучше всего будет прогнозировать следующие будущие наблюдения – та ли, которая имеет в прошлом минимальную дисперсию, та ли, которая имеет в прошлом минимальную среднюю абсолютную ошибку аппроксимации или ещё какая-нибудь. Нет никаких оснований для такого априорного вывода. Необходимо сгенерировать различными методами множество различных оценок выбранной модели на некоторой части имеющейся базы; проверить с помощью процедуры ретропрогноза точность прогноза каждого из методов оценки коэффициентов модели на проверочном множестве и отдать предпочтение тому методу, который показал наилучшие прогнозные оценки. Здесь, конечно, мы возвращаемся к индуктивному подходу, предполагая, что если некоторый выбранный метод в прошлом давал лучшие прогнозные оценки, то и в будущем он будет обладать подобными же свойствами, что вовсе не очевидно. Но другого варианта у прогнозиста нет.

Покажем, как можно получить множество способов оценивания коэффициентов прогнозных моделей с помощью метода  $z$ -множителей<sup>1</sup>. Вспомним, что для нахождения значений коэффициентов прогнозной модели мы должны каким-то образом получить такое число уравнений  $n$ , которое бы соответствовало числу  $n$  неизвестных коэффициентов этой модели. Решая эту систему из  $n$  уравнений с  $n$  неизвестными, можно найти численные значения коэффициентов.

Действительно, если, например, перед прогнозистом стоит задача найти коэффициенты линейного тренда (с двумя коэффициентами), то тот же МНК предлагает ему решить систему двух нормальных уравнений, в результате чего вычисляются значения двух неизвестных коэффициентов тренда. Если же прогнозисту необходимо оценить значения коэффициентов квадратичной функции с тремя коэффициентами, тот же МНК приводит его к необходимости решения системы трёх нормальных уравнений и т.д.

Рассмотрим, в соответствии с общенаучным принципом «от простого – к сложному» самую простую модель линейной однофакторной зависимости, на примере которой будет ясен смысл метода  $z$ -множителей. После этого легко можно будет использовать метод и для оценки коэффициентов более сложных моделей.

Любая модель, очевидно, описывает реальный процесс с некоторой ошибкой аппроксимации  $\varepsilon_t$ , поэтому для любого значения  $t$  выполняется такое равенство:

---

<sup>1</sup> Светуных С.Г. Эконометрические методы прогнозирования спроса (на примере промышленной энергетики). - М.: Изд-во МГУ, 1993. - С.86.

$$y_t = \hat{y}_t + \varepsilon_t = a_0 + a_1 x_t + \varepsilon_t, \quad (2.1.2)$$

Для использования этой модели при прогнозировании необходимо на имеющемся множестве значений  $y_t$  найти значения двух коэффициентов –  $a_0$  и  $a_1$ . Значит, надо каким-то образом построить два уравнения с этими двумя коэффициентами и, решая эту систему из двух уравнений, оценить значения коэффициентов модели.

Очевидно также, что равенство (2.1.2) не нарушится, если его левую и правую части умножить на некоторый заранее известный заданный прогнозом множитель  $z_{0t} \neq 0$ :

$$y_t z_{0t} = a_0 z_{0t} + a_1 x_t z_{0t} + \varepsilon_t z_{0t}. \quad (2.1.3)$$

Если теперь просуммировать левую и правую части полученного равенства по всем наблюдениям  $t$ , получим уравнение:

$$\sum_t y_t z_{0t} = \sum_t a_0 z_{0t} + \sum_t a_1 x_t z_{0t} + \sum_t \varepsilon_t z_{0t}. \quad (2.1.4)$$

Теперь умножим левую и правую части равенства (2.1.2) на другой, также заранее известный заданный прогнозом множитель  $z_{1t} \neq 0$ , не являющийся линейным преобразованием множителя  $z_{0t}$ :

$$Y_t z_{1t} = a_0 z_{1t} + a_1 x_t z_{1t} + \varepsilon_t z_{1t}. \quad (2.1.5)$$

Просуммировав теперь и это уравнение по всем наблюдениям  $t$ , получим второе уравнение:

$$\sum_t Y_t z_{1t} = a_0 \sum_t z_{1t} + a_1 \sum_t x_t z_{1t} + \sum_t \varepsilon_t z_{1t}. \quad (2.1.6)$$

Сведем уравнения (2.1.3) и (2.1.6) в одну систему:

$$\begin{cases} \sum_t Y_t z_{0t} = a_0 \sum_t z_{0t} + a_1 \sum_t x_t z_{0t} + \sum_t \varepsilon_t z_{0t}, \\ \sum_t Y_t z_{1t} = a_0 \sum_t z_{1t} + a_1 \sum_t x_t z_{1t} + \sum_t \varepsilon_t z_{1t}. \end{cases} \quad (2.1.7)$$

Данная система - система двух линейных уравнений с  $(T+2)$  неизвестными -  $a_0$ ,  $a_1$  и  $\varepsilon_t$  (численные значения множителей  $z_{0t}$  и  $z_{1t}$  задаются прогнозом). Очевидно, что эта система имеет множество возможных решений и поэтому для задачи нахождения оценок коэффициентов линейной однофакторной модели она не пригодна. Но на её основе можно добиться

решения поставленной задачи, для чего необходимо задать некоторые дополнительные условия к этой системе.

Для задания этих условий будем исходить из того очевидного положения, что точность описания некоторого процесса с помощью любой модели определяется характером ошибок аппроксимации  $\varepsilon_t$ . Поэтому, если и следует задавать некоторые условия к задаче (2.1.7), то их следует связывать именно с этими ошибками аппроксимации. Можно, например, задать такие дополнительные условия именно относительно этих ошибок аппроксимации:

$$\left\{ \begin{array}{l} \sum_t \varepsilon_t z_{0t} = k_0 = const, \\ \sum_t \varepsilon_t z_{1t} = k_1 = const. \end{array} \right. \quad (2.1.8)$$

где  $k_0$  и  $k_1$  – наперёд заданные числа.

Тогда при выполнении условий (2.1.8) система уравнений (2.1.7) будет записана так:

$$\left\{ \begin{array}{l} \sum_t Y_t z_{0t} = a_0 \sum_t z_{0t} + a_1 \sum_t x_t z_{0t} + k_0, \\ \sum_t Y_t z_{1t} = a_0 \sum_t z_{1t} + a_1 \sum_t x_t z_{1t} + k_1. \end{array} \right. \quad (2.1.9)$$

Поскольку  $k_0$  и  $k_1$  заданы, то получена система двух уравнений с двумя неизвестными, которое имеет одно решение. Очевидно, что значения коэффициентов модели будут определяться как характером задания  $z$ -множителей, так и значениями констант  $k_0$  и  $k_1$ .

Самый простой случай рассматриваемой задачи соответствует ситуации, когда  $k_0 = k_1 = 0$ , то есть:

$$\left\{ \begin{array}{l} \sum_t \varepsilon_t z_{0t} = 0, \\ \sum_t \varepsilon_t z_{1t} = 0. \end{array} \right. \quad (2.1.10)$$

для него будет получена система:

$$\left\{ \begin{array}{l} \sum_t Y_t z_{0t} = a_0 \sum_t z_{0t} + a_1 \sum_t x_t z_{0t}, \\ \sum_t Y_t z_{1t} = a_0 \sum_t z_{1t} + a_1 \sum_t x_t z_{1t}. \end{array} \right. \quad (2.1.11)$$

С её помощью коэффициенты  $a_0$  и  $a_1$  которые могут быть легко найдены.

Этот случай предпочтительнее случая, задаваемого условием (2.1.8), поскольку при этом появляется возможность интерпретации свойств получаемых оценок.

Действительно, пусть, например, используются такие  $z$ -множители:

$$\begin{cases} z_{0t} = C_o = const, \\ z_{1t} = (-1)^t, \end{cases} \quad (2.1.12)$$

а относительно ошибок аппроксимации выполняется (2.1.10). Тогда в случае, если число наблюдений  $T$  чётное, получим:

$$\begin{cases} \sum_t Y_t = a_0 T + a_1 \sum_t x_t \\ \sum_t Y_t (-1)^t = a_1 \sum_t x_t (-1)^t \end{cases} \quad (2.1.13)$$

Коэффициенты прогнозной модели для этого способа находятся очень просто – из второго уравнения сразу же вычисляются значения коэффициента  $a_1$ , а после этого, подставляя полученные значения в первое уравнение системы, легко найти  $a_0$ . Следует сразу же отметить, что для этой модели выполняются условия

$$\begin{cases} \sum_t \varepsilon_t = 0, \\ \sum_t \varepsilon_t (-1)^t = 0, \end{cases} \quad (2.1.14)$$

которые со всей очевидностью следуют из (2.1.10).

Из этой системы равенств следует понимание того, что будет собой представлять модель линейной модели, если использовать этот способ задания  $z$ -множителей.

Во-первых, сумма отклонений расчётных значений модели от фактических всегда будет равна нулю. Это, означает, что модель с коэффициентами, полученными таким способом задания  $z$ -множителей, будет всегда проходить через среднюю арифметическую точку. Это со всей очевидностью следует из первого равенства системы (2.1.13).

Второе равенство (2.1.14), свидетельствует о том, что сумма ряда ошибок аппроксимации  $\varepsilon_t$ , умноженных на знакопеременный ряд единиц, будет равна нулю (второе равенство системы (2.1.14)).

Итак, с помощью системы (2.1.11), задавая различные значения  $z$ -множителей, можно получить и различные значения коэффициентов прогнозной модели, это - во-первых. Во-вторых, с учётом одновременного выполнения системы равенств (2.1.10) и (2.1.11) прогнозист имеет дополнительную информацию о том, какими свойствами обладает ошибка аппроксимации.

Сразу же возникает вопрос: а как соотносится этот метод и метод наименьших квадратов? Ответ прост – если множители задать так:



$$\begin{cases} z_{ot} = C_o = const \neq 0, \\ z_{1t} = x_t, \end{cases} \quad (2.1.15)$$

то, подставляя их в (2.1.11), получим систему двух уравнений, которая в точности будет соответствовать системе нормальных уравнений МНК. Действительно, в этом случае получим:

$$\begin{cases} \sum_t Y_t = a_0 T + a_1 \sum_t x_t, \\ \sum_t Y_t x_t = a_0 \sum_t x_t + a_1 \sum_t x_t^2. \end{cases} \quad (2.1.16)$$

Но теперь метод  $z$ - множителей позволяет не только получить оценки МНК применительно к линейной модели, но и получить дополнительную интерпретацию оценок МНК, поскольку будет выполняться и условие (2.1.10). Мы можем с полным основанием утверждать, что МНК, применённый к линейной функции, даёт такие значения коэффициентов модели, при которых всегда выполняются условия:

$$\begin{cases} \sum_t \varepsilon_t = 0, \\ \sum_t \varepsilon_t x_t = 0, \end{cases} \quad (2.1.17)$$

Какой смысл имеют эти условия?

Первое равенство уравнения (2.1.17) свидетельствует о том, что для оценок МНК сумма ошибок аппроксимации всегда будет равна нулю и всегда оценки МНК линейной однофакторной зависимости будут такими, что модель будет проходить через среднюю арифметическую наблюдений. А если описываемый с помощью линейной модели процесс будет нелинейным, как будет вести себя модель, коэффициенты которой найдены с помощью МНК? Как следует из первого равенства (2.1.17), модель пройдёт через среднюю точку, а сумма отклонений фактических значений от расчётных будет равна нулю. Очень редко, но иногда всё же приходится сталкиваться с экономистами, которые считают, что если для построенной модели сумма ошибок аппроксимации равна нулю, то модель лучше всего описывает исследуемый процесс. Ошибочность этой точки зрения теперь очевидна – МНК, применённый к любому процессу, к нелинейному в том числе, будет для линейной однофакторной модели всегда давать такие оценки коэффициентов, при которых сумма ошибок аппроксимации равна нулю. О пригодности или непригодности модели сумма ошибок аппроксимации ничего не говорит.

Задавая различные  $z$ -множители, прогнозист, решая систему (2.1.11), получает различные значения коэффициентов линейной однофакторной модели, и выбирает ту пару значений  $z$ -множителей, при которой ошибка

ретропрогноза минимальна. Это могут быть и оценки МНК, но для необратимых процессов, чаще всего это будут другие оценки.

Рассмотрим теперь, как можно использовать метод  $z$ -множителей для более сложных моделей, например, квадратичной модели:

$$\hat{Y}_t = a_0 + a_1 x_t + a_2 x_t^2. \quad (2.1.18)$$

Поскольку модель содержит три неизвестных коэффициента  $a_0$ ,  $a_1$  и  $a_2$ , то необходимо использовать три множителя:  $z_{0t} \neq z_{1t} \neq z_{2t}$  для получения системы из трёх уравнений с тремя неизвестными. Опуская вывод системы уравнений метода  $z$ -множителя, аналогичный выводу системы уравнений для линейной однофакторной модели, получим следующую систему для нахождения значений коэффициентов квадратичной модели (2.1.18):

$$\begin{cases} \sum_t Y_t z_{0t} = a_0 \sum_t z_{0t} + a_1 \sum_t x_t z_{0t} + a_2 \sum_t x_t^2 z_{0t}, \\ \sum_t Y_t z_{1t} = a_0 \sum_t z_{1t} + a_1 \sum_t x_t z_{1t} + a_2 \sum_t x_t^2 z_{1t}, \\ \sum_t Y_t z_{2t} = a_0 \sum_t z_{2t} + a_1 \sum_t x_t z_{2t} + a_2 \sum_t x_t^2 z_{2t}, \end{cases} \quad (2.1.19)$$

которой соответствует система равенств, задающая условия для ошибки аппроксимации:

$$\begin{cases} \sum_t \varepsilon_t z_{0t} = 0, \\ \sum_t \varepsilon_t z_{1t} = 0, \\ \sum_t \varepsilon_t z_{2t} = 0. \end{cases} \quad (2.1.20)$$

Опять покажем, как получить из этой общей системы оценивания коэффициентов квадратичной модели такие коэффициенты, которые будут соответствовать оценкам МНК. Для этого зададим такие множители:

$$z_{0t} = 1, \quad z_{1t} = x_t, \quad z_{2t} = x_t^2.$$

Тогда система (2.1.19) с такими множителями будет соответствовать системе нормальных уравнений МНК. Действительно, подставляя эти множители в (2.1.19), получим:

$$\begin{cases} \sum_t Y_t = a_0 T + a_1 \sum_t x_t + a_2 \sum_t x_t^2, \\ \sum_t Y_t x_t = a_0 \sum_t x_t + a_1 \sum_t x_t^2 + a_2 \sum_t x_t^3, \\ \sum_t Y_t x_t^2 = a_0 \sum_t x_t^2 + a_1 \sum_t x_t^3 + a_2 \sum_t x_t^4. \end{cases} \quad (2.1.21)$$

что, как легко убедиться, полностью соответствует системе уравнений МНК. Если теперь задать, например, такую совокупность множителей:

$$z_{0t}=t, z_{1t}=x_t, z_{2t}=x_t^2,$$

то полученные оценки будут близки к оценкам МНК, но всё же отличаться от них.

Изложенный метод  $z$ -множителей позволяет предложить бесконечное множество способов оценки коэффициентов моделей прогнозирования, причём МНК – только один из этого множества. Поэтому, используя разумное число возможных комбинаций и способов задания множителей, прогнозист может из этого множества выбрать тот из них, который демонстрирует свои лучшие свойства в процедуре ретропрогноза.

Методическое свойство метода  $z$ -множителей заключается в том, что он позволяет легко сформулировать систему нормальных уравнений МНК для любых аддитивных моделей, что позволяет сформировать систему нормальных уравнений для различных моделей, не прибегая к утомительному выводу этой системы традиционным путём через вычисление производных по коэффициентам функции минимизации суммы квадратов отклонений фактических значений от расчётных.

Действительно, сравнивая (2.1.11) и (2.1.16) можно убедиться в том, что для получения системы нормальных уравнений МНК следует левую и правую части линейной модели  $y_t = a_0 + a_1 x_t$  умножить сначала на множитель при свободном члене  $a_0$ , и просуммировать по всем  $t$  (этим множителем является единица,  $z_{0t}=1$ ); затем левую и правую части этой же модели умножить на множитель перед вторым коэффициентом  $a_1 (z_{1t}=x_t)$ , и вновь просуммировать по всем наблюдениям – будет получено второе равенство, а вместе - система нормальных уравнений.

Пусть, например, прогнозист хочет с помощью МНК оценить коэффициенты такой модели:

$$y_t = a_1 \sin x_t + a_2 \ln x_t. \quad (2.1.22)$$

Для построения системы нормальных уравнений МНК оценивания коэффициентов этой модели следует вначале левую и правые части равенства умножить на  $z_{0t}=\sin x_t$ , просуммировать по всем  $t$ . После этого левую и правую части равенства умножить на  $z_{1t}=\ln x_t$ , после чего вновь

полученные произведения просуммировать по всем  $t$ . Сведём два этих уравнения в систему:

$$\begin{cases} \sum_t y_t \sin x_t = a_1 \sum_t \sin^2 x_t + a_2 \sum_t \ln x_t \sin x_t, \\ \sum_t y_t \ln x_t = a_1 \sum_t \sin x_t \ln x_t + a_2 \sum_t \ln^2 x_t. \end{cases} \quad (2.1.23)$$

Не менее просто получить систему нормальных уравнений МНК для многофакторной аддитивной модели, например такой:

$$y_t = a_0 + a_1 x_{1t} x_{2t} + a_2 \ln x_{2t} + a_3 e^{x_{3t}}. \quad (2.1.24)$$

$Z$ -множители такой модели для получения системы нормальных уравнений очевидны:

$$z_{0t} = 1, \quad z_{1t} = x_{1t} x_{2t}, \quad z_{2t} = \ln x_{2t}, \quad z_{3t} = a_3 e^{x_{3t}}. \quad (2.1.25)$$

С их помощью легко получить искомую систему уравнений:

$$\begin{cases} \sum_t y_t = a_0 T + a_1 \sum_t x_{1t} x_{2t} + a_2 \sum_t \ln x_{2t} + a_3 \sum_t e^{x_{3t}}, \\ \sum_t y_t x_{1t} x_{2t} = a_0 \sum_t x_{1t} x_{2t} + a_1 \sum_t (x_{1t} x_{2t})^2 + a_2 \sum_t x_{1t} x_{2t} \ln x_{2t} + a_3 \sum_t x_{1t} x_{2t} e^{x_{3t}}, \\ \sum_t y_t \ln x_{2t} = a_0 \sum_t \ln x_{2t} + a_1 \sum_t x_{1t} x_{2t} \ln x_{2t} + a_2 \sum_t \ln^2 x_{2t} + a_3 \sum_t \ln x_{2t} e^{x_{3t}}, \\ \sum_t y_t e^{x_{3t}} = a_0 \sum_t e^{x_{3t}} + a_1 \sum_t x_{1t} x_{2t} e^{x_{3t}} + a_2 \sum_t e^{x_{3t}} \ln x_{2t} + a_3 \sum_t e^{2x_{3t}}. \end{cases} \quad (2.1.26)$$

Естественно, что в силу (2.1.10) будет выполняться для полученных оценок МНК следующие условия:

$$\begin{cases} \sum_t \varepsilon_t = 0, \\ \sum_t \varepsilon_t x_{1t} x_{2t} = 0, \\ \sum_t \varepsilon_t \ln x_{2t} = 0, \\ \sum_t \varepsilon_t e^{x_{3t}} = 0. \end{cases} \quad (2.1.27)$$

Но поскольку для эволюционных процессов оценки МНК не являются самыми лучшими, используя различные  $z$ -множители, можно получать самые различные системы уравнений, решая которые, формируется семейство оценок коэффициентов, из которого выбирается лучший набор  $z$ -множителей, используя процедуру ретропрогноза. Практика применения метода  $z$ -множителей показывает, что для задач прогнозирования эволюционных процессов МНК дают лучшие прогнозы примерно в 10%

случаях – в остальных 90% есть такие системы, оценки коэффициентов которых дают более точные прогнозы.

Рассмотрим простой пример<sup>2</sup>. В табл. 2.1 приведены статистические данные, которые будем использовать для демонстрации метода  $z$ -множителей.

На статистических данных с первого по восемнадцатое наблюдение будем с помощью разных способов задания  $z$ -множителей оценивать коэффициенте однофакторной линейной модели. Данные с девятнадцатого по двадцать третье наблюдение будем использовать как проверочные для определения точности ретропрогноза. Будем использовать три способа задания множителей:

- 1)  $z_{0t}=1, z_{1t}=x_t$ ,
- 2)  $z_{0t}=1, z_{1t}=(-1)^t$ ,
- 3)  $z_{0t}=x_t, z_{1t}=(-1)^t/x_{t-1}$ ,

Табл. 2.1  
Исходные данные для использования метода  $z$ -множителей

Год наблюдения, $t$	Электропотребление промышленностью, млрд. кВт час, $Y_t$	Численность занятых в промышленности, млн. чел., $x_t$	Год наблюдения, $t$	Электропотребление промышленностью, млрд. кВт час, $Y_t$	Численность занятых в промышленности, млн. чел., $x_t$
1	5,57	1,953	13	17,10	
2	6,36	2,055	14	17,49	
3	7,50	2,291	15	17,90	
4	8,28	2,350	16	18,48	
5	9,06	2,443	17	19,22	
6	9,74	2,535	18	19,91	
7	10,36	2,634	19	21,10	
8	11,60	2,773	20	22,10	
9	12,79	2,878	21	23,40	
10	13,92	2,965	22	24,30	
11	14,95	3,056	23	25,05	
12	16,12	3,153			

Первый способ задания  $z$ -множителей, очевидно, соответствует оценкам МНК. Все остальные способы – дают иные оценки. Использование этих  $z$ -множителей позволило построить такие модели:

- 1)  $\hat{Y}_t = 8,965x_t - 12,64$ , средняя абсолютная ошибка аппроксимации этой модели равна 0,32;

<sup>2</sup> Светуных С.Г. Эконометрические методы прогнозирования спроса (на примере промышленной энергетики). - М.: Изд-во МГУ, 1993. – С. 88 – 95.

- 2)  $\hat{Y}_t = 9,304x_t - 13,60$ , средняя абсолютная ошибка аппроксимации модели равна 0,34;
- 3)  $\hat{Y}_t = 9,257x_t - 13,50$ , средняя абсолютная ошибка аппроксимации этой модели равна 0,32.

Модели не только по-разному аппроксимируют прошлое, но и обладают различной прогнозной точностью, что видно из результатов прогноза данных с двадцатого по двадцать третье наблюдение. Для наглядности результаты приведены в табл. 2.2.

Как видно из таблицы, с позиций минимума средней абсолютной ошибки аппроксимации оценки МНК не хуже, чем оценки, полученные с помощью третьей модели и лучше, чем оценки второй модели.

Но оценки МНК оказались хуже для целей прогнозирования - лучшими с позиций минимума ошибки средней абсолютной ошибки ретропрогноза являются второй и третий способы задания  $z$ -множителей. Они же дают самые точные ретропрогнозы с позиций минимума дисперсии ошибки ретропрогноза.

Табл. 2.2

Результаты аппроксимации и ретропрогноза при разном задании  $z$ -множителей

Показатель	<i>Модель 1 (оценки МНК)</i>	<i>Модель 2</i>	<i>Модель 3</i>
Средняя абсолютная ошибка аппроксимации	0,32	0,34	0,32
Средняя абсолютная ошибка ретропрогноза	0,42	0,18	0,20
Дисперсия ретропрогноза	0,19	0,04	0,05

Подобные же результаты получаются и для множества других эволюционно развивающихся процессов, для каждого из них наилучшим способом оценки коэффициентов прогнозных моделей является свой оригинальный способ задания  $z$ -множителей. Поэтому можно сделать вывод о том, что для необратимых процессов, для которых не выполняются предпосылки выборочного метода, оценки МНК не являются лучшими для целей прогнозирования динамики.