

МИНИСТЕРСТВО ОБРАЗОВАНИЯ РЕСПУБЛИКИ БЕЛАРУСЬ
БЕЛОРУССКИЙ ГОСУДАРСТВЕННЫЙ УНИВЕРСИТЕТ

ФАКУЛЬТЕТ ПРИКЛАДНОЙ МАТЕМАТИКИ И
ИНФОРМАТИКИ

КАФЕДРА МАТЕМАТИЧЕСКОГО МОДЕЛИРОВАНИЯ И АНАЛИЗА ДАННЫХ

Курсовой проект

**«Краткосрочное прогнозирование и наукастинг
макроэкономических временных рядов на основе векторных
моделей авторегрессии по смешанным данным»**

Бовта Тимофея Анатольевича
студента 4 курса 7 группы
специальности «прикладная
математика»

Научный руководитель:

В. И. Малюгин
зав. кафедрой ММАД,
доктор экономических наук,
доцент

Минск, 2024 г.

МИНИСТЕРСТВО ОБРАЗОВАНИЯ РЕСПУБЛИКИ БЕЛАРУСЬ
БЕЛОРУССКИЙ ГОСУДАРСТВЕННЫЙ УНИВЕРСИТЕТ
ФАКУЛЬТЕТ ПРИКЛАДНОЙ МАТЕМАТИКИ И ИНФОРМАТИКИ
Кафедра математического моделирования и анализа данных

**ЗАДАНИЕ
НА КУРСОВОЙ ПРОЕКТ**

Студент Бовт Тимофей Анатольевич

1. Тема курсового проекта

«КРАТКОСРОЧНОЕ ПРОГНОЗИРОВАНИЕ И НАУКАСТИНГ МАКРОЭКОНОМИЧЕСКИХ ВРЕМЕННЫХ РЯДОВ НА ОСНОВЕ ВЕКТОРНЫХ МОДЕЛЕЙ АВТОРЕГРЕССИИ ПО СМЕШАННЫМ ДАННЫМ».

2. Исходные данные к курсовому проекту

2.1. Малюгин В. И. Краткосрочное прогнозирование и наукастинг темпов роста инфляции на основе моделей по смешанным данным — Журнал «Банковский вестник» №1/726 — С. 23-36.

2.2. Макеева, Н.М. Наукастинг элементов использования ВВП России / Н.М. Макеева, И.П. Станкевич – Статья 2022/10, Экономический журнал ВШЭ.

2.3. Макеева, Н.М. Наукастинг макроэкономических показателей экономики России на основе анализа новостного фона и регулярных данных Росстата / Н. М. Макеева, И. П. Станкевич

3. Перечень подлежащих разработке вопросов:

3.1. Провести обзор по теме «Проблема краткосрочного прогнозирования и наукастинга для данных смешанной частоты».

3.2. Подготовить математическое описание модели MFVAR.

3.3. Провести экспериментальное исследование модели MFVAR в задаче «Краткосрочное прогнозирование и наукастинг темпов роста ВВП Республики Беларусь по его компонентам (элементам использования, источникам использования доходов)».

3.4. Подготовить отчёт по курсовому проекту.

4. Дата выдачи задания: 18.09.2024.

5. Срок сдачи законченного курсового проекта: 20.12.2024.

Руководитель

В. И. Малюгин

Подпись обучающегося

Т. А. Бовт

Дата

20.12.2024

ОГЛАВЛЕНИЕ

ВВЕДЕНИЕ	5
1 МАТЕМАТИЧЕСКОЕ ОПИСАНИЕ РАССМАТРИВАЕМЫХ МОДЕЛЕЙ И АЛГОРИТМОВ	7
1.1 Временные ряды и лаговый оператор	7
1.2 Стационарные временные ряды	8
1.3 Стационарные относительно тренда временные ряды	9
1.4 Интегрированные временные ряды	11
1.5 Тест Дики-Фуллера	12
1.6 Расширенный тест Дики-Фуллера	13
1.6.1 ADF-тест без константы	13
1.6.2 ADF-тест с константой	15
1.6.3 ADF тест с трендом	16
1.6.4 Альтернативный подход к тестированию	17
1.6.5 Выбор типа ADF теста и порядка лагов	17
1.6.6 Тесты на единичный корень, допускающие структур- ный сдвиг в известное время	18
1.7 Сезонная корректировка временных рядов	20
1.7.1 Темпы роста как метод сезонной корректировки	20
1.7.2 Разложение временного ряда на сезонный фактор, тренд и случайную составляющую	21
1.7.3 Стандартные методы сезонной корректировки	21
1.8 Авторегрессия, векторная авторегрессия	23
1.9 Модель векторной авторегрессии по смешанным данным MF- VAR	24
1.9.1 Недостающие наблюдения и оценки	26
1.10 Байесовский подход к модели векторной авторегрессии по смешанным данным MF-BVAR	26
1.11 Оценка точности моделей	27
2 СТАТИСТИЧЕСКИЙ АНАЛИЗ МОДЕЛЕЙ ВРЕМЕННЫХ РЯДОВ ПО РЕАЛЬНЫМ ДАННЫМ	29
2.1 Сезонная корректировка	30
2.2 Исключение тренда	30
2.3 Тестирование стационарности	33
3 МОДЕЛИРОВАНИЕ И ПРОГНОЗИРОВАНИЕ ТЕМПОВ РОСТА ВВП БЕЛАРУСИ ПО ИСТОЧНИКАМ ДОХОДОВ	35
3.1 MFVAR на основе промышленности, товарооборота и сель- ского хозяйства	36

3.2	MFVAR на основе промышленности, товарооборота сельского хозяйства, строительства	37
3.3	MFVAR на основе промышленности, товарооборота, сельского хозяйства, строительства, объема инвестиций	39
3.4	MFVAR на основе промышленности, товарооборота, сельского хозяйства, строительства, объема инвестиций, доходов населения	40
3.5	Оценки моделей и сравнительный анализ	41
4	МОДЕЛИРОВАНИЕ И ПРОГНОЗИРОВАНИЕ ТЕМПОВ РОСТА И ПРОМЫШЛЕННОСТИ ВВП БЕЛАРУСИ ПО ИСТОЧНИКАМ ДОХОДОВ	43
	ЗАКЛЮЧЕНИЕ	44
	СПИСОК ИСПОЛЬЗОВАННОЙ ЛИТЕРАТУРЫ	46

ВВЕДЕНИЕ

Все применяемые регрессионные модели классического машинного обучения работают с данными, заданными в одной частоте. Нередко на практике при анализе собранных данных можно столкнуться с такой проблемой, как различная частота этих данных. К примеру, некоторые данные из сферы экономики, как правило, формируются в квартальных представлениях (например, ВВП). Параллельно с этим какие-либо объясняющие факторы могут быть собраны с более высокой частотой, будь то ежемесячные, еженедельные или ежедневные представления. Стандартные регрессионные модели не заточены под такое представление данных. Соответственно в ходе предварительного анализа приходится преобразовывать данные к одной частоте. Как правило, выбирается один из следующих способов.

1. Агрегация данных высокой частоты.

Приведем пример: если исследуемая зависимая переменная находится в квартальном представлении, а независимые данные — в ежемесячном, то мы можем составить новый набор независимых переменных, взяв в качестве квартального значения последний месяц квартала. Таким образом, мы получим все данные в одной частоте, что позволяет нам использовать большое количество моделей машинного обучения для предсказания необходимого нам показателя.

Однако такой подход имеет свой главный недостаток: возникает потеря некоторой информации о динамике объясняющих данных, которая может быть крайне полезна при построении модели.

2. Интерполяция низкочастотных переменных.

Для этого используются специальные подходы для заполнения пропущенных значений, рассматривать которые мы не будем. Этот вариант используется редко, и зачастую предпочтение отдается первому варианту. Данный подход также может способствовать появлению различного рода проблем при построении модели.

В связи с этим возникает вопрос: как можно без преобразования данных и потери какой-либо информации строить регрессионную модель для предсказания исследуемых показателей.

В данной работе для решения задачи прогнозирования по смешанным данным рассматриваются модели векторной авторегрессии по данным смешанной частоты (MFVAR). В последнее время эти модели используются для прогнозирования макроэкономических временных рядов, где обычно квартальный рост ВВП прогнозируется по ежемесячным макроэкономическим и финансовым показателям.

Векторные авторегрессионные модели (VAR) для прогнозирования используют не только прошлые значения объясняющих факторов, но и прошлые значения предсказываемой переменной. Также модели векторной авторегрессии позволяют прогнозировать сразу несколько эндогенных переменных, что выделяет данные модели на фоне известных MIDAS моделей, которые также работают с данными смешанной частоты.

Разработка моделей, способных работать с переменными, отбираемыми с разной частотой вызывает значительный интерес в сфере эконометрии. В работе описывается и приводится процесс предварительной обработки временных рядов. Поскольку прежде, чем решать задачу прогнозирования, необходимо убедиться в том, что модели смогут корректно работать на исследуемых данных. Следовательно, сперва мы приведем временные ряды к нужной нам форме, а затем займемся исследованием рассматриваемых моделей.

Опишем подробно ход работы:

1. исследовать вероятностно-статистические свойства предоставленных данных;
2. рассмотреть математические аппараты, используемые для моделирования поставленной задачи;
3. обработать входные данные с целью улучшения качества прогноза;
4. построить эконометрическую модель, описывающую поведение эндогенных переменных в зависимости от поведения экзогенных переменных;
5. исследовать статистическую адекватность построенных моделей и результатов прогнозирования.

ГЛАВА 1

МАТЕМАТИЧЕСКОЕ ОПИСАНИЕ РАССМАТРИВАЕМЫХ МОДЕЛЕЙ И АЛГОРИТМОВ

Рассмотрим математический аппарат, с помощью которого далее будем проводить моделирование. Для описания данных, исследованием которых мы будем заниматься, используются модели временных рядов.

1.1 Временные ряды и лаговый оператор

Пусть $\{y_t\}_{t=-\infty}^{t=+\infty}$ — это временной ряд, то есть последовательность наблюдений, упорядоченная по параметру $t \in T \subset \mathbb{R}$. Будем считать, что под параметром t мы подразумеваем время.

Одно случайное значение y_t временного ряда в момент времени $t \in T$ будем называть *отсчетом* временного ряда. А расстояние между отсчетами по времени будем называть *лагом*. Совокупность всех отсчетов временного ряда будем называть *реализацией* [4].

В эконометрике наиболее распространенными моделями временных рядов (одномерными и многомерными) являются:

1. стационарные временные ряды;
2. стационарные относительно тренда временные ряды (Trend Stationarity);
3. интегрированные временные ряды (Difference Stationarity, DS).

В теории временных рядов важным является понятие стационарности. Использование в регрессионной модели нестационарных временных рядов может привести к фиктивным результатам или к построению так называемой «мнимой» или ложной регрессии. Это обуславливает необходимость учитывать при моделировании, являются ли временные ряды стационарными или нет. Далее мы подробнее рассмотрим все описанные модели временных рядов.

Введем определение, на основании которого и будут строиться все рассматриваемые нами модели.

Лаговый оператор L — это оператор сдвига, позволяющий получить значения элементов временного ряда на основании ряда предыдущих значений.

То есть для реализации временного ряда $\{y_1, y_2, \dots, y_t\}$ лаговый оператор будет действовать следующим образом

$$L^k y_t = y_{t-k}. \quad (1.1.1)$$

Этот оператор обладает следующими свойствами:

1. $L^0 = 1$;
2. $Lc = c, c \in \mathbb{R}$;
3. $L^{-1}y_t = y_{t+1}$.

Наряду с лаговым оператором определяются лаговые многочлены. Например, лаговый многочлен p -ой степени может быть записан как

$$\psi(L) = \psi_0 + \psi_1 L + \dots + \psi_p L^p = \sum_{j=0}^p L^j. \quad (1.1.2)$$

Тогда

$$\psi(L)y_t = \psi_0 y_t + \psi_1 y_{t-1} + \dots + \psi_p y_{t-p} = \sum_{j=0}^p L^j y_t. \quad (1.1.3)$$

1.2 Стационарные временные ряды

Временной ряд $y_t, t \in T$ называется *стационарным в узком смысле (строго стационарным)*, если для любых $n \geq 1$ его отсчетов совместная функция распределения вероятностей этих отсчетов не зависит от сдвига во времени.

$$F_n(y_1, \dots, y_n; t_1, \dots, t_n) = F_n(y_1, \dots, y_n; t_1 + \tau, \dots, t_n + \tau). \quad (1.2.1)$$

Временной ряд $y_t, t \in T$ называется *стационарным в широком смысле (слабо стационарным)*, если выполняются следующие условия:

1. отсчеты временного ряда как случайные величины имеют конечные первый и второй моменты, то есть

$$|\mathbb{E}\{y_t\}| < +\infty, \quad \mathbb{E}\{y_t^2\} < +\infty;$$

2. математическое ожидание временного ряда не зависит от времени, то есть

$$\mathbb{E}\{y_t\} = \mu \in \mathbb{R}, \quad \forall t \in T;$$

3. для ковариационной функции выполняется

$$\text{cov}(y_{t_1}, y_{t_2}) = \text{cov}(y_{t_1+\tau}, y_{t_2+\tau}), \quad \forall t_1, t_2 \in T, \tau \in \mathbb{Z}.$$

Если существуют первый и второй моменты отсчетов стационарного в узком смысле временного ряда, то этот временной ряд является стационарным в широком смысле [4].

Два разных типа нестационарных по отношению к среднему временных рядов могут быть приведены к стационарному виду с помощью взятия последовательных разностей (конечных разностей, лагов).

Стационарные временные ряды вполне эффективны, так как имеют ряд полезных свойств, присущих также и модели случайной выборки. В частности, для стационарных временных рядов мы можем рассчитать постоянные математическое ожидание, дисперсию, ковариацию и, следовательно, корреляцию. Это свойство позволяет нам строить правдоподобные модели для прогнозирования будущих значений. Также все позже определенные нами модели будут корректно работать именно со стационарными временными рядами.

В общем случае, если у нас есть какие-либо данные с временными рядами, то эти временные ряды будут являться нестационарными, поскольку один и тот же интервал по времени нельзя прожить более чем один раз. Поэтому для наших исследований нам важно проверять временные ряды на стационарность и в случае нестационарности приводить временные ряды к стационарной форме.

1.3 Стационарные относительно тренда временные ряды

В экономическом анализе встречаются временные ряды имеющие устойчивую тенденцию с течением времени. Поведение таких временных рядов можно описывать регрессионной моделью тренда, где в качестве объясняющей переменной выступает фактор времени. Рассмотрим класс рядов, имеющих выраженный тренд.

Временной ряд y_t , $t \in T$ называется *стационарным относительно тренда (TS)*, если для него имеет место представление

$$y_t = f(t) + u_t, \quad (1.3.1)$$

где $f(t)$ — детерминированная функция, называемая *трендом*, или долгосрочной тенденцией, u_t — стационарный временной ряд с нулевым средним, то есть $\mathbb{E}\{u_t\} = 0$. В частности, к TS-рядам относятся все стационарные временные ряды. Однако, многие TS-ряды являются нестационарными.

Из данного определения очевидно следует, что

$$\mathbb{E}\{y_t\} = f(t)$$

и

$$\text{Var}\{y_t\} = \text{Var}\{u_t\} = \text{const}.$$

В прикладных нередко можно столкнуться с ситуацией, когда в модели (1.3.1) временной ряд u_t автокоррелирован, то есть существует статистическая взаимосвязь между последовательностями величин этого ряда, взятыми со сдвигом. Рассмотрим основные модели тренда.

Модель линейного тренда представляется в виде

$$y_t = \beta_0 + \beta_1 t + u_t, \quad t = 0, \pm 1, \dots \quad (1.3.2)$$

где коэффициент β_1 отвечает за среднее изменение за один период времени.

В предположении, что ошибки u_t удовлетворяют условиям теоремы Гаусса-Маркова, к модели линейного тренда применимы выводы модели классической линейной регрессии. Легко видеть, что среднее значение $\mathbb{E}\{y_t\}$ временного ряда будет линейно зависеть от времени t :

$$\mathbb{E}\{y_t\} = \beta_0 + \beta_1 t. \quad (1.3.3)$$

Коэффициент β_1 считают средним приращением временного ряда за один период времени, он может быть выражен следующим образом

$$\Delta \mathbb{E}\{y_t\} = \mathbb{E}\{y_t\} - \mathbb{E}\{y_{t-1}\} = \beta_1. \quad (1.3.4)$$

Таким образом, с увеличением времени, при $\beta_1 > 0$ ($\beta_1 < 0$) во временном ряду есть «тенденция к возрастанию» (убыванию), причем средняя скорость изменения временного ряда за один период времени постоянна.

Модель экспоненциального тренда представляется в виде

$$\ln y_t = \beta_0 + \beta_1 t + u_t, \quad t = 0, \pm 1, \dots \quad (1.3.5)$$

где коэффициент $\beta_1 \cdot 100\%$ отвечает за среднее процентное изменение за один период;

По аналогии с моделью линейного тренда, будем считать, что ошибки u_t удовлетворяют условиям теоремы Гаусса-Маркова. Также среднее значение модели экспоненциального тренда будет линейно зависеть от времени t :

$$\mathbb{E}\{\ln(y_t)\} = \beta_0 + \beta_1 t. \quad (1.3.6)$$

Для коэффициента β_1 можно получить следующую интерпретацию:

$$\Delta \mathbb{E}\{\ln(y_t)\} = \mathbb{E}\{\ln(y_t)\} - \mathbb{E}\{\ln(y_{t-1})\} = \mathbb{E}\left\{\ln\left(\frac{y_t}{y_{t-1}}\right)\right\} = \beta_1, \quad (1.3.7)$$

то есть за один период времени (в среднем) значение y_t изменяется в $\exp(\beta_1)$ раз. Если β_1 мало, то $\exp(\beta_1) \approx 1 + \beta_1$ и за один период времени в среднем значение y_t изменяется (в первом приближении) на $\beta_1 \cdot 100\%$.

По аналогии с линейными и экспоненциальными моделями тренда можно строить и другие модели. Например, *модель квадратичного тренда*, которая представляется в виде

$$y_t = \beta_0 + \beta_1 t + \beta_2 t^2 + u_t$$

и так далее. Выбор функциональной модели тренда является особой задачей, для которой применяются отдельные методы и тесты модели регрессии на функциональную форму.

1.4 Интегрированные временные ряды

Временной ряд y_t называется *интегрированным порядка k* , $k = 1, 2, \dots$, если

1. ряд y_t не является TS-рядом, то есть нестационарный;
2. k – это такое минимальная кратность взятия конечных разностей, что ряд $\Delta^k y_t$ является стационарным рядом или TS-рядом.
3. все временные ряды $\Delta^{k-1} y$ до порядка $(k - 1)$ включительно не являются TS-рядами.

Обозначают $y_t \sim I(k)$.

В принятой системе обозначений соотношение $y_t \sim I(0)$ соответствует ряду, который является стационарным и при этом не является результатом дифференцирования TS-ряда.

Совокупность интегрированных рядов различных порядков $k = 1, 2, \dots$ образует класс *разностно стационарных (DS)*, или *стационарных относительно взятия разностей временных рядов*.

При построении эконометрических моделей на основе реальных статистических данных, как правило, в случае DS-рядов исследователь работает с рядами, интегрированными первого порядка.

Понятие «единичного корня» (unit root) используется в анализе временных рядов как характеризующее свойство некоторых нестационарных временных рядов [2].

Временной ряд y_t *имеет единичный корень*, или порядок интеграции один, если его конечные разности образуют стационарный ряд, то есть $y_t \sim I(1)$, если

$$\Delta y_t = y_t - y_{t-1} \sim I(0).$$

Рассмотрим *авторегрессионное уравнение первого порядка (AR(1))*

$$y_t = a \cdot y_{t-1} + u_t, \quad a = \text{const} \quad (1.4.1)$$

где y_t – это временной ряд, а u_t – это случайная ошибка. Если в уравнении (1.4.1) коэффициент $a = 1$, то временной ряд имеет единичный корень, то есть в этом случае ряд y_t нестационарен, при этом является интегрированным временным рядом первого порядка $I(1)$. Если же $|a| < 1$, то временной ряд y_t стационарный, то есть $y_t \sim I(0)$.

Для финансово-экономических процессов значение $|a| > 1$ не свойственно, так как в этом случае процесс является «взрывным». Возникновение таких процессов маловероятно, так как финансово-экономическая среда достаточно инерционная, что не позволяет принимать бесконечно большие значения за малые промежутки времени [1].

Методика проверки временных рядов на стационарность включает в себя целый ряд тестов, направленных на выявление «единичного корня» (unit root tests), базовым из которых является тест Дики-Фуллера (DF-test).

1.5 Тест Дики-Фуллера

Как мы ранее выяснили, одной из ключевых задач в статистическом анализе временных рядов является классификация ряда как стационарного, TS или DS-ряда. Для решения данной поставленной задачи применяются различные методы тестирования, среди которых особое место занимают тесты на наличие единичного корня (unit root tests).

Тест Дики — Фуллера (DF-тест) – это методика, которая используется в эконометрике для статистического анализа временных рядов для проверки на стационарность. Является одним из тестов на единичные корни. Был предложен в 1979 году Дэвидом Дики и Уэйном Фуллером.

Авторегрессионное уравнение (1.4.1) можно записать как

$$\Delta y_t = b \cdot y_{t-1} + u_t, \quad (1.5.1)$$

где $b = a - 1$, а Δ — оператор разности первого порядка:

$$\Delta y_t = y_t - y_{t-1}.$$

Проверка гипотезы о наличии единичного корня в данном представлении сводится к проверке следующей нулевой гипотезы:

$$H_0 : b = 0, \quad (1.5.2)$$

против альтернативной гипотезы:

$$H_1 : b < 0. \quad (1.5.3)$$

Поскольку случай «взрывных» процессов ($b > 0$) исключается, тест является односторонним.

Статистика теста, называемая DF-статистикой, является обычной t -статистикой для проверки значимости коэффициента b в уравнении линейной регрессии. Однако распределение этой статистики отличается от классического t -распределения (распределения Стьюдента). Распределение DF-статистики выражается через винеровский процесс и называется распределением Дики-Фуллера.

Существует три версии теста:

1. Без константы и тренда:

$$\Delta y_t = b \cdot y_{t-1} + u_t. \quad (1.5.4)$$

2. С константой, но без тренда:

$$\Delta y_t = b_0 + b \cdot y_{t-1} + u_t. \quad (1.5.5)$$

3. С константой и линейным трендом:

$$\Delta y_t = b_0 + b_1 \cdot t + b \cdot y_{t-1} + u_t. \quad (1.5.6)$$

Для каждой из трёх тестовых регрессий существуют свои критические значения DF-статистики, которые берутся из специальной таблицы Дики — Фуллера (МакКиннона). Если значение статистики лежит левее критического значения (критические значения — отрицательные) при данном уровне значимости, то нулевая гипотеза о единичном корне отклоняется и процесс признается стационарным (в смысле данного теста). В противном случае гипотеза не отвергается и процесс может содержать единичные корни, то есть быть нестационарным (интегрированным) временным рядом [1].

1.6 Расширенный тест Дики-Фуллера

Широко используемым и стандартным тестом на единичный корень является *расширенный тест Дики-Фуллера (ADF-test)*. Он является модификацией теста Дики-Фуллера в тех случаях, когда предполагается авторегрессионность отклонений модели. В каждое уравнение теста вводятся авторегрессионные переменные (лаги) переменной разности Δy_t для коррекции возможной коррелированности случайных отклонений тестируемой модели.

1.6.1 ADF-тест без константы

Пусть y_t — временной ряд с нулевым средним $\mathbb{E}\{y_t\} = 0$.

Для некоторого заданного p для модели

$$y_t = \beta_1 y_{t-1} + \dots + \beta_p y_{t-p} + u_t, \quad u_t \sim \text{WN}(0, \sigma_u^2) \quad (1.6.1)$$

тестируем гипотезу, что авторегрессионный многочлен

$$\beta(z) = 1 - \beta_1 z - \dots - \beta_p z^p \quad (1.6.2)$$

имеет единичный корень кратности 1 (остальные корни многочлена по модулю больше 1).

Очевидно, $z = 1$ будет корнем многочлена $\beta(z)$ тогда и только тогда, когда $\sum_{j=1}^p \beta_j = 1$. Преобразуем уравнение. Перепишем (1.6.1) в виде

$$\Delta y_t = \varphi y_{t-1} + \theta_1 \Delta y_{t-1} + \dots + \theta_{p-1} \Delta y_{t-p+1} + u_t, \quad (1.6.3)$$

где $\varphi = \beta_1 + \dots + \beta_p - 1$, $\theta_i = -\sum_{j=i+1}^p \beta_j$. Тогда наличие единичного корня означает, что

$$\varphi = 0,$$

а стационарность означает

$$\varphi \leq 0.$$

Идея теста состоит в том, чтобы для преобразованного уравнения (1.6.3) тестируется гипотеза

$$H_0 : \varphi = 0,$$

$$H_1 : \varphi \leq 0.$$

При справедливости H_0 t -статистика

$$t = \frac{\hat{\varphi}_{\text{МНК}}}{\text{s.e.}(\varphi)} \quad (1.6.4)$$

не имеет распределения Стьюдента.

Приведем алгоритм теста:

1. оцениваем метод наименьших квадратов преобразованное уравнение (1.6.3);
2. вычисляем тестовую статистику

$$\text{ADF}_t = \text{DF}_\tau = \frac{\hat{\varphi}_{\text{МНК}}}{\text{s.e.}(\varphi)}; \quad (1.6.5)$$

3. сравниваем значение статистики с критическим значением τ_{cr} Дики-Фуллера без константы.

4. отвергаем H_0 (гипотезу единичного корня) при

$$\text{ADF}_t \leq -\tau_{cr} \leq 0; \quad (1.6.6)$$

иначе принимаем гипотезу H_0 .

1.6.2 ADF-тест с константой

В этом случае имеем временной ряд y_t с постоянным средним $\mathbb{E}\{y_t\} = a$ и допускается $a \neq 0$. Тогда

$$y_t = a + z_t, \quad \mathbb{E}\{z_t\} \equiv 0. \quad (1.6.1)$$

Для заданного p для модели

$$y_t = \mu + \beta_1 y_{t-1} + \dots + \beta_p y_{t-p} + u_t, \quad u_t \sim \text{WN}(0, \sigma_u^2) \quad (1.6.2)$$

тестируем, что авторегрессионный многочлен $\beta(z)$ имеет единичный корень кратности один и $\mu = 0$.

Аналогично предыдущему преобразуем уравнение (1.6.2)

$$\Delta y_t = \mu + \varphi y_{t-1} + \theta_1 \Delta y_{t-1} + \dots + \theta_{p-1} \Delta y_{t-p+1} + u_t \quad (1.6.3)$$

Тогда гипотеза единичного корня означает, что

$$\mu = 0, \quad \varphi = 0,$$

а стационарность будет при

$$\varphi \leq 0.$$

Можем записать следующий алгоритм:

- (а) Оцениваем с помощью метода наименьших квадратов преобразованное уравнение (1.6.3);
- (b) вычисляем тестовую статистику

$$\text{ADF}_t = \text{DF}_\tau = \frac{\hat{\varphi}_{\text{МНК}}}{\text{s.e.}(\varphi)}; \quad (1.6.4)$$

- (с) сравниваем значение статистики с критическим значением τ_{cr} теста Дики-Фуллера с константой.
- (d) отвергаем H_0 (гипотезу единичного корня) при

$$\text{ADF}_t \leq -\tau_{cr} \leq 0, \quad (1.6.5)$$

иначе принимаем гипотезу.

1.6.3 ADF тест с трендом

Пусть y_t – временной ряд с линейным трендом. Тогда

$$y_t = a_0 + a_1 t + z_t, \quad \mathbb{E}\{z_t\} \equiv 0. \quad (1.6.1)$$

Для заданного p для модели

$$y_t = \mu + at + \beta_1 y_{t-1} + \dots + \beta_p y_{t-p} + u_t \quad (1.6.2)$$

тестируем, что авторегрессионный многочлен $\beta(z)$ имеет единичный корень кратности один и $a = 0$.

Аналогично предыдущему преобразуем уравнение (1.6.2)

$$\Delta y_t = \mu + at + \varphi y_{t-1} + \theta_1 \Delta y_{t-1} + \dots + \theta_{p-1} \Delta y_{t-p+1} + u_t \quad (1.6.3)$$

Тогда гипотеза единичного корня означает, что

$$a = 0, \quad \varphi = 0,$$

а стационарность будет при

$$\varphi \leq 0.$$

Можем записать следующий алгоритм:

- (а) оцениваем с помощью метода наименьших квадратов преобразованное уравнение (1.6.3);
- (б) вычисляем тестовую статистику

$$ADF_t = DF_\tau = \frac{\widehat{\varphi}_{\text{МНК}}}{\text{s.e.}(\varphi)}; \quad (1.6.4)$$

- (с) сравниваем значение статистики с критическим значением τ_{cr} теста Дики-Фуллера с линейным трендом.
- (д) отвергаем H_0 (гипотезу единичного корня) при

$$ADF_t \leq -\tau_{cr} \leq 0, \quad (1.6.5)$$

иначе принимаем.

1.6.4 Альтернативный подход к тестированию

Для преобразованной регрессии

$$y_t = \varphi y_{t-1} + \begin{pmatrix} 0 \\ \text{const} \\ \text{trend} \end{pmatrix} + \sum_{j=1}^{p-1} \beta_j \Delta x_{t-j} + u_t \quad (1.6.1)$$

используем тестовые статистики

$$\text{ADF}_t = \frac{\hat{\varphi} - 1}{\text{s.e.}(\varphi)}, \quad (1.6.2)$$

$$\text{ADF}_n = \frac{n(\hat{\varphi} - 1)}{1 - \hat{\beta}_1 - \dots - \hat{\beta}_p}. \quad (1.6.3)$$

Отметим, что $\text{ADF}_n = n(\hat{\varphi} - 1)$ в случае $p = 0$.

Критические значения для ADF_t такие же, как и раньше. А для критических значений ADF_n теста имеются отдельные таблицы.

Статистическое правило, как и раньше, заключается в том, что мы отвергаем гипотезу единичного корня, когда тестовая статистика меньше критического значения (меньше 0).

1.6.5 Выбор типа ADF теста и порядка лагов

Выбор типа теста зависит от априорной информации о временном ряде:

- если считаем, что есть детерминированный долгосрочный тренд, то используем тест с трендом;
- если считаем, что детерминированного тренда нет, то используем тест только с константой (и без тренда, это повышает мощность теста);
- если думаем, что ряд имеет нулевое среднее, то используем первый тест.

Неформальный подход для выбора порядка лагов состоит в использовании экономического смысла или экономической модели. Формальный подход, как правило, сводится к применению правила Schwert'a (1989):

1. определяем максимальный лаг $p_{max} = [12\sqrt[4]{\frac{n}{100}}]$;
2. выбираем порядок лага $p \leq p_{max}$ по информационным критериям Akaike или Schwarz (Bayesian), либо по t -статистикам (чтобы старший коэффициент был значим).

1.6.6 Тесты на единичный корень, допускающие структурный сдвиг в известное время

В исследованиях П. Перрона допускалось наличие «иногда» возникающих существенных изменений в функции тренда рассматриваемый временных рядов [14, 15]. Эти изменения могут быть связаны с «глобальными» шоками, такими как мировой финансовый кризис в 2008 г., пандемия Covid-19 в 2020 г., и внутренними факторами страны, данные которой анализируются. В связи с тем, что большие изменения имеют нечастый характер, это мешает оцениванию и определению вероятностных распределений для них. Следовательно, такие изменения нельзя рассматривать как стохастические, выбранные из распределения, отличного от того, из которого выбираются шоки, происходящие в каждый период. Поэтому необходимо такие большие изменения в тренде моделировать как структурные изменения. При проведении ADF-теста допускается наличие нескольких больших перманентных изменений в функции тренда.

Перейдем к описанию современных методов тестирования гипотезы о наличии единичного корня при возможном наличии структурных сдвигов.

Изначально в работах П. Перона предполагалось однократное изменение в функции тренда и две версии трёх различных структур:

1. изменение в уровнях (модель «краха»);
2. изменение в наклоне тренда (модель «изменения роста»);
3. изменение и в уровнях, и в наклоне тренда [16].

Для каждого из этих случаев существует две различные версии, учитывающие различные эффекты перехода:

- инновационный выброс (ИО), предполагающий, что изменение в тренде происходит постепенно;
- аддитивный выброс (АО), предполагающий, что изменение в функции тренда происходит мгновенно.

Различие между ИО и АО моделями значимо не только из-за различных траекторий перехода, но и по причине различий в статистических методах, используемых для тестирования на наличие единичного корня в этих случаях. Структурный сдвиг в модели Перрона предполагается заданных экзогенно. Данный факт является недостатком такого типа моделей, поскольку структурный сдвиг не обязательно совпадает со временем шока, что означает, что сдвиг может иметь запаздывающий характер (имеет место некий лаг).

Основная концепция исследований Перрона заключается в том, что в случае наличия структурного сдвига в истинном процессе мощность расширенного теста Дики-Фуллера может значительно снижаться и зачастую оказывается равной нулю. Перрон предлагает статистику для моделей инновационного выброса, которая строится на основе регрессии

$$y_t = \mu_0 + \mu_1 DU_t + a_0 t + a_1 DT_t + \delta D_t(T_1) + \varphi y_{t-1} + \sum_{j=1}^{p-1} \beta_j \Delta y_{t-j} + u_t. \quad (1.6.1)$$

В этой модели Перрон добавляет дополнительные дамми-переменные,

$$DU_t = \begin{cases} 1, & \text{если } t \geq T_1 \\ 0, & \text{если } t \leq T_1 \end{cases}$$

для учета сдвига в уровнях и

$$DT_t = \begin{cases} t - T_1, & \text{если } t \geq T_1 \\ 0, & \text{если } t \leq T_1 \end{cases}$$

для учета сдвига в тренде.

Тестовой статистикой является обычная t -статистика для тестирования нулевой гипотезы $\varphi = 0$ против альтернативы $\varphi \leq 0$ и обозначается как $t_\varphi(\lambda_1)$, где $\lambda_1 = \frac{T_1}{T}$ — доля даты структурного сдвига. Если дата сдвига является истинной, то распределение статистики остаётся инвариантным относительно параметров функции тренда, что означает, что оно не зависит от параметров, определяющих изменения в константе и наклоне тренда.

Для моделей АО процедура немного отличается от ранее описанной и включает два этапа. На первом этапе оценивается трендовая функция временного ряда, после чего эта функция удаляется из исходного временного ряда с помощью следующей регрессии:

$$y_t = \mu_0 + a_0 t + \mu_1 DU_t + a_1 DT_t + \tilde{u}_t, \quad (1.6.2)$$

Соответствующие детрендрованные ряды обозначаются как остатки \tilde{u}_t . Второй шаг различается в зависимости от того, содержится ли на первом шаге компонента DU_t , представляющая собой дамми, определяющая сдвиг в константе. Если DU_t добавляется на первом шаге, используется t -статистика для проверки нулевой гипотезы $\varphi = 0$ в следующей регрессии:

$$\tilde{u}_t = \varphi \tilde{u}_{t-1} + \sum_{j=0}^{p-1} d_j D_{t-j}(T_1) + \sum_{j=1}^{p-1} \beta_j \Delta y_{t-j} + u_t, \quad (1.6.3)$$

где

$$D_t(T_1) = \begin{cases} 1, & \text{если } t = T_1 + 1 \\ 0, & \text{если } t \neq T_1 + 1 \end{cases}$$

является импульсной дамми.

1.7 Сезонная корректировка временных рядов

На динамику экономических показателей существенно влияют сезонные колебания — повторяющиеся изменения в структуре ряда, связанные с временем года. Эти колебания могут маскировать истинные тренды и искажать результаты прогнозирования, что делает необходимым применение сезонной корректировки. *Сезонная корректировка* — это статистический метод, направленный на устранение влияния сезонных факторов на исходные данные временного ряда. Она позволяет получить «очищенные» данные, которые более точно отражают базовую тенденцию (тренд) и циклические колебания, исключая влияние предсказуемых сезонных эффектов. Без этой процедуры анализ может привести к ложным выводам и неточным прогнозам.

Сезонно скорректированный временной ряд сохраняет не только тренд, но также случайные изменения, выбросы и структурные изменения, вызванные различными внутренними и внешними шоками. Полученные результаты позволяют проводить точные расчеты темпов роста в текущем периоде, анализировать динамику соответствующих индикаторов в контексте других экономических показателей и осуществлять краткосрочные прогнозы экономической ситуации.

1.7.1 Темпы роста как метод сезонной корректировки

Один из распространенных, но достаточно приблизительных методов исключения сезонного фактора — это расчет темпов роста показателя к аналогичному периоду предыдущего года, то есть теоретически мы можем получить сезонно скорректированный временной ряд путем преобразования

$$y_t^* = \frac{y_t}{y_{t-1}}. \quad (1.7.1)$$

Однако такой подход имеет ряд недостатков:

- хотя этот подход может уменьшить влияние постоянного годового сезонного эффекта, полное устранение сезонности не всегда возможно, так как сезонный фактор может постепенно изменяться в пределах года;
- он не учитывает другие эффекты (например, эффект количества рабочих дней, скользящих праздничных дней), а также не отражает изменение сезонного эффекта по мере изменения во времени уровня показателя (предполагается, что сезонный фактор остается неизменным);

- данный метод чувствителен к различным случайным эффектам; например, если экономический показатель в предшествующем периоде существенно снизился из-за неожиданного шока, то в соответствующем квартале следующего года будет наблюдаться аномально высокий рост (эффект базы);
- использование темпов роста в текущем периоде к аналогичному периоду прошлого года может неадекватно отражать поворотные точки в динамике, так как не учитывает поквартальное изменение показателя в пределах года [11].

1.7.2 Разложение временного ряда на сезонный фактор, тренд и случайную составляющую

Метод разложения временного ряда позволяет выделить три ключевые компоненты временного ряда: тренд, сезонный фактор и случайную составляющую. Разделяют два вида компонентного разложения:

- аддитивное;
- мультипликативное.

В аддитивной модели временной ряд y_t в момент t представляется как сумма трех составляющих

$$y_t = T_t + S_t + E_t, \quad (1.7.1)$$

где T_t — тренд, отражающий общее направление изменений во времени; S_t — сезонный фактор, фиксирующий регулярные колебания, происходящие в определенные временные интервалы; E_t — случайная составляющая, представляющая собой случайные отклонения от ожидаемого значения.

Аддитивное разложение применяется, когда величина сезонных колебаний остается относительно постоянной и не зависит от общего уровня временного ряда.

В мультипликативной модели временной ряд представляется как произведение трех составляющих

$$y_t = T_t \times S_t \times E_t. \quad (1.7.2)$$

1.7.3 Стандартные методы сезонной корректировки

Анализ сезонности и корректировка временных рядов, как правило, проводится при помощи программного обеспечения. Наиболее популярным

является метод TRAMO/SEATS, предложенный банком Испании [13]. Метод состоит из двух этапов: TRAMO, который отвечает за предварительную обработку и моделирование; SEATS, который выполняет разложение временного ряда на компоненты.

TRAMO

1. *Предварительная обработка данных.* На первом этапе TRAMO проводит очистку данных от выбросов и пропусков. Для этого применяются регрессионные модели, которые могут учитывать внешние факторы, влияющие на временной ряд. Формально, модель может быть записана как

$$y_t = \beta_0 + \beta_1 X_{1t} + \beta_2 X_{2t} + \dots + \beta_k X_{kt} + u_t \quad (1.7.1)$$

где y_t — значения временного ряда, которые наблюдаются в момент времени t ; X_{it} — внешние регрессионные переменных (например, экономические показатели); β_i — коэффициенты модели; u_t — белый шум, предполагается, что он подчиняется нормальному распределению.

2. *Моделирование ARIMA.* После обработки данных TRAMO определяет модель ARIMA, которая лучше всего описывает шумовые компоненты. Модель ARIMA обозначается как $ARIMA(p, d, q)$, где p — порядок авторегрессии, d — степень интегрирования, q — порядок скользящей средней.

Общая форма модели ARIMA может быть записана с помощью лагового многочлена следующим образом:

$$\Phi(L)\Delta^d y_t = \Theta(L)u_t \quad (1.7.2)$$

где L — лаговый оператор,

$$\Phi(L) = 1 - \varphi_1 L - \varphi_2 L^2 - \dots - \varphi_p L^p$$

— это лаговый полином авторегрессии,

$$\Theta(L) = 1 + \theta_1 L + \theta_2 L^2 + \dots + \theta_q L^q$$

— это лаговый полином скользящей средней, Δ — оператор взятия конечных разностей.

SEATS

1. *Сезонная декомпозиция.* На этапе SEATS разложение временного ряда осуществляется на три компоненты: тренд T_t , сезонный фактор S_t и случайную составляющую E_t . Данный этап был описан в подразделе 1.7.2.

2. *Извлечение тренда и сезонности.* Для извлечения тренда и сезонного компонента используются скользящие средние и фильтры. В аддитивной модели тренд может быть оценён с помощью сглаживания:

$$T_t = \frac{1}{k} \sum_{j=0}^{k-1} y_{t-j} \quad (1.7.3)$$

где k — размер окна сглаживания.

Для сезонного фактора S_t можно использовать

$$S_t = y_t - T_t - E_t \quad (1.7.4)$$

или в случае мультипликативной модели

$$S_t = \frac{y_t}{T_t \times E_t} \quad (1.7.5)$$

3. *Оценка случайной составляющей.* В конце производится оценка свойств и поведения случайной компоненты.

Таким образом, мы рассмотрели такую модель описания данных как временные ряды и математически сформулировали основные свойства этой модели, которые будут играть важную роль при построении математических моделей для прогнозирования. Далее мы подробно остановимся на конкретных моделях, которые мы будем использовать для прогнозирования и получения результатов.

1.8 Авторегрессия, векторная авторегрессия

Пусть $\{\varepsilon_t\}_{t=-\infty}^{t=+\infty}$ — это последовательность случайных величин таких, что

$$\mathbb{E}\{\varepsilon_t\} = 0, \quad \mathbb{D}\{\varepsilon_t\} = \sigma^2 < +\infty, \quad t \in \mathbb{Z}.$$

Причем пусть выполняется одно из условий:

- случайные величины ε_t некоррелированы:

$$\mathbb{E}\{\varepsilon_t \varepsilon_l\} = \begin{cases} \sigma^2, & l = k, \\ 0, & l \neq k \end{cases}, \quad l, k \in \mathbb{Z},$$

- случайные величины ε_t независимы в совокупности и одинаково распределены;
- случайные величины ε_t некоррелированы и имеют нормальное распределение $\mathcal{N}(0, \sigma^2)$.

Временной ряд $\{y_t\}_{t=-\infty}^{t=+\infty}$ называется *временным рядом авторегрессии* порядка p , т.е. $AR(p)$, если

$$\sum_{i=0}^p \beta_i y_{t-i} = \varepsilon_t, \quad (1.8.1)$$

где $\beta_0 = 1$, а β_1, \dots, β_p — коэффициенты авторегрессии, причем $\beta_p \neq 0$.

Коэффициенты авторегрессии можно оценивать с помощью известных из курса математической статистики метода максимального правдоподобия и метода наименьших квадратов.

Также возможна запись $AR(p)$ в виде

$$y_t = \beta_0 + \sum_{i=1}^p \beta_i y_{t-i} + \varepsilon_t, \quad (1.8.2)$$

здесь уже β_0 — это свободный член.

Перенесем эту формулу из \mathbb{R} в \mathbb{R}^n . Пусть у нас теперь вектор временных рядов

$$y_t = (y_t^1, y_t^2, \dots, y_t^k), \quad (1.8.3)$$

$\varepsilon_t \in \mathbb{R}^k$ и $B_i = (\beta_{ij})$ — матрица. Тогда формулу (1.6.2) можно переписать как

$$y_t = \beta_0 + \sum_{i=1}^p B_i y_{t-i} + \varepsilon_t. \quad (1.8.4)$$

Формула (1.6.4) задает векторную авторегрессию порядка p , то есть $VAR(p)$.

Приведенная модель является замкнутой, в том смысле, что в качестве объясняющих переменных выступают только лаги объясняемых переменных.

1.9 Модель векторной авторегрессии по смешанным данным MF-VAR

Модели векторной авторегрессии (VAR модели) предполагают использование совершенно другого подхода для работы с данными смешанной частоты. Теперь же определим модель векторной регрессии по данным смешанной частоты (mixed frequency VAR, или MF-VAR). Предположим, что полученные нами данные на самом деле измерены на одной частоте, причем на той частоте, на которой измерена независимая переменная. И пусть среди значений зависимой переменной имеются пропуски. Причем пропуски в этих данных не случайные, а периодические.

Пусть у нас получены данные, где предсказываемая переменная y_t (например, ВВП) наблюдается в квартальной частоте, а объясняющие переменные x_t получены в ежемесячной частоте. Модель MF-VAR может опре-

делить совместную динамику месячной зависимой переменной, которая получается из квартальных значений зависимой переменной с разбивкой по времени, и месячных значений независимых переменных.

В соответствии с обозначениями Mariano и Mugasawa (2010), разбивка квартального роста ВВП y_{t_m} на ненаблюдаемые значения месячного роста ВВП $y_{t_m}^*$ основана на следующем отношении агрегации

$$\begin{aligned} y_{t_m} &= \frac{1}{3}(y_{t_m}^* + y_{t_m-1}^* + y_{t_m-2}^*) + \frac{1}{3}(y_{t_m-1}^* + y_{t_m-2}^* + y_{t_m-3}^*) + \\ &+ \frac{1}{3}(y_{t_m-2}^* + y_{t_m-3}^* + y_{t_m-4}^*) = \frac{1}{3}y_{t_m}^* + \frac{2}{3}y_{t_m-1}^* + y_{t_m-2}^* + \frac{2}{3}y_{t_m-3}^* + \frac{1}{3}y_{t_m-4}^*, \end{aligned} \quad (1.9.1)$$

которое рассматривается на каждом $t_m = 3, 6, 9, \dots, T_m$, поскольку у нас есть данные о ВВП лишь каждый третий месяц каждого квартала.

Пусть для всех t_m рост месячного ВВП $y_{t_m}^*$ и соответствующее этому $y_{t_m}^*$ значение месячной объясняющей переменной x_{t_m} соответствуют двумерному VAR(p) процессу

$$\Phi(L_m) \begin{pmatrix} y_{t_m}^* - \mu_y^* \\ x_{t_m} - \mu_x \end{pmatrix} = \varepsilon_{t_m}, \quad (1.9.2)$$

где $\Phi(L_m) = \sum_{i=1}^p \Phi_i L_m^i$ и $u_{t_m} \sim \mathcal{N}(0, \Sigma)$, $\mu_y^* = \mathbb{E}\{y_{t_m}^*\}$, $\mu_x = \mathbb{E}\{x_{t_m}\}$. Таким образом, VAR(p) процесс из уравнения (1.8.2) вместе с уравнением (1.8.1) позволяют определить представление в пространстве состояний. Определим вектор состояний

$$s_{t_m} = \begin{pmatrix} z_{t_m} \\ \vdots \\ z_{t_m-4} \end{pmatrix}, \quad z_{t_m} = \begin{pmatrix} y_{t_m}^* - \mu_y^* \\ x_{t_m} - \mu_x \end{pmatrix}. \quad (1.9.3)$$

Тогда представление MF-VAR модели в пространстве состояний будет записано как

$$s_{t_m} = F s_{t_m-1} + G v_{t_m} \quad (1.9.4)$$

$$\begin{pmatrix} y_{t_m} - \mu_y \\ x_{t_m} - \mu_x \end{pmatrix} = H s_{t_m}, \quad (1.9.5)$$

где $\mu_y = 3\mu_y^*$ и $v_{t_m} \sim \mathcal{N}(0, I_2)$. Матрицы, введенные в формуле (7.5) определим как

$$F = \begin{bmatrix} F_1 \\ F_2 \end{bmatrix}, \quad F_1 = [\Phi_1 \ \dots \ \Phi_p \ 0_{2 \times 2(5-p)}], \quad F_2 = [I_8 \ 0_{8 \times 2}] \quad (1.9.6)$$

$$G = \begin{bmatrix} \Sigma^{1/2} \\ 0_{8 \times 2} \end{bmatrix}, \quad H = [H_0 \ \dots \ H_4], \quad (1.9.7)$$

где матрица H содержит коэффициенты лагового многочлена

$$H(L_m) = \sum_{i=0}^4 H_i L_m^i, \quad (1.9.8)$$

который определен как

$$H(L_m) = \begin{bmatrix} 1/3 & 0 \\ 0 & 1 \end{bmatrix} + \begin{bmatrix} 2/3 & 0 \\ 0 & 0 \end{bmatrix} L_m + \begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix} L_m^2 + \begin{bmatrix} 2/3 & 0 \\ 0 & 0 \end{bmatrix} L_m^3 + \begin{bmatrix} 1/3 & 0 \\ 0 & 0 \end{bmatrix} L_m^4, \quad (1.9.9)$$

исходя из формулы (1.8.1).

Для удобства записи мы рассматриваем только $p < 4$ для матрицы F и G , однако представление для $p > 4$ может быть получено простым способом путем соответствующего изменения вектора состояния и матриц систем.

1.9.1 Недостающие наблюдения и оценки

Модель пространства состояний, состоящая из формул (1.8.4) и (1.8.5), может быть оценена с помощью метода максимального правдоподобия или алгоритма максимизации математического ожидания (ЕМ), где мы должны учитывать недостающие наблюдения из-за низкочастотного характера ВВП. Для того, чтобы заполнить недостающие значения, мы сначала заменяем все пропущенные значения нулями, предполагая, что пропущенные значения являются реализациями некоторой стандартной нормальной случайной величины, независимой и одинаково распределенной. Во-вторых, уравнение сигнала (1.8.5) также модифицируется соответствующим образом: для первых двух месяцев каждого квартала верхняя строка матрицы H устанавливается равной нулю и добавляется стандартный элемент нормальной ошибки. Затем для оценки параметров используется алгоритм ЕМ.

За счет заполнения недостающих наблюдений MF-VAR модели считаются более продвинутыми с технической точки зрения среди моделей, используемых для наукастинга.

1.10 Байесовский подход к модели векторной авторегрессии по смешанным данным MF-BVAR

Байесовские векторные авторегрессии смешанной частоты (Mixed-Frequency Bayesian VAR, MF-BVAR) представляют собой версию стандартных векторных авторегрессий, модифицированную для использования с данными разной частоты. Предполагается, что при исходном высокочастотном (в нашем случае месячном) процессе, описываемом стандартной VAR-моделью,

мы наблюдаем часть переменных только на более низкой частоте. При этом, к примеру, наблюдаемые квартальные значения являются средними из ненаблюдаемых месячных значений.

В качестве примера байесовской оценки MF-VAR мы представляем алгоритм, разработанный Schorfheide и Song (2011). Авторы представляют MF-VAR как модель пространства состояний и используют методы марковской цепи Монте-Карло (MCMC) для проведения байесовского вывода для параметров модели и ненаблюдаемых ежемесячных переменных.

Уравнение состояния модели можно представить моделью VAR(p) с использованием формул (1.8.3)-(1.8.7), записанной в форме-компаньоне:

$$z_{t_m} = F_1(\Phi)z_{t_m-1} + F_c(\Phi) + v_{t_m}, \quad v_{t_m} \sim iid\mathcal{N}(0, \Omega(\Sigma)). \quad (1.10.1)$$

Для того, чтобы написать уравнение измерения, авторы записывают уравнение агрегации, которое в данном случае отличается от того, которое рассматривалось в формуле (1.8.1). В этом случае квартальная переменная рассматривается как среднее значение месячного процесса за три месяца, которое в предыдущем обозначении равно:

$$y_{t_m} = \frac{1}{3}(y_{t_m}^* + y_{t_m-1}^* + y_{t_m-2}^*) = \Lambda_{mz}z_{t_m}. \quad (1.10.2)$$

Однако поскольку y_{t_m} наблюдается только каждый третий месяц, то существует необходимость в матрице выбора M_{t_m} , которая равна единичной матрице, если t_m соответствует последнему месяцу квартала, и нулевая в противном случае. Следовательно, уравнение измерения можно записать как

$$\begin{pmatrix} y_{t_m} \\ x_{t_m} \end{pmatrix} = M_{t_m}\Lambda_z z_{t_m}. \quad (1.10.3)$$

Для решения проблемы размерности введен метод Миннесоты, который сокращает коэффициенты VAR до одномерных представлений случайного блуждания.

1.11 Оценка точности моделей

Для оценки качества прогнозов моделей наиболее популярными являются три следующих критерия:

- средняя абсолютная ошибка (MAE)

$$MAE = \frac{1}{T} \sum_{t=1}^T |y_t - \hat{y}_t|; \quad (1.11.1)$$

- средняя абсолютная ошибка в процентах (MAPE)

$$MAPE = \frac{1}{T} \sum_{t=1}^T \frac{|y_t - \hat{y}_t|}{y_t}; \quad (1.11.2)$$

- корень из среднеквадратической ошибки (RMSE)

$$RMSE = \sqrt{\frac{1}{T} \sum_{t=1}^T (y_t - \hat{y}_t)^2}, \quad (1.11.3)$$

- y_t — фактическое значение эндогенной переменной в период t ;
- \hat{y}_t — спрогнозированное с помощью модели значение эндогенной переменной в период t ;
- T — количество периодов, на которых тестируется модель.

При построении наукастов не учитывается информация о последнем доступном квартале: перед оцениванием модели из выборки удаляются значения зависимых переменных и соответствующие данному кварталу месячные значения объясняющих переменных. Далее в выборку возвращаются удаленные значения регрессоров и для них рассчитывается прогнозное значение зависимой переменной (наукаст). Рассматриваемые модели сравниваются по последним 12 точкам, например это может быть период с третьего квартала 2019 года по первый квартал 2022 года в случае ВВП. Таким образом, модели тестируются на достаточно разнородных данных: в тестирование попадают как относительно спокойный период 2019 г., так и кризисные периоды.

Для проверки устойчивости к добавлению новых данных все модели тестируются трижды: с использованием данных по объясняющим переменным за все три месяца квартала, для которого рассчитывается наукаст; без данных за последний месяц и без данных за два последних месяца. В случае удаления части данных в объясняющих переменных, «пустым» месяцам в объясняющих переменных проставляется последнее доступное значение показателя (за второй или за первый месяц квартала в зависимости от метода тестирования). Такая проверка позволяет определить, насколько методы устойчивы к объему используемых данных, и смоделировать встречающиеся в реальной жизни условия, когда наукаст показателя за текущий квартал рассчитывается еще до окончания квартала.

ГЛАВА 2

СТАТИСТИЧЕСКИЙ АНАЛИЗ МОДЕЛЕЙ ВРЕМЕННЫХ РЯДОВ ПО РЕАЛЬНЫМ ДАННЫМ

Для проведения исследований нам даны следующие экономические показатели

- GDP — ВВП Беларуси по источникам использования доходов в среднегодовых ценах 2018 г. на квартальной частоте;
- IPV — объём промышленного производства в среднегодовых ценах 2018 г. на месячной частоте;
- RTV — объём розничного товарооборота в среднегодовых ценах 1995 г. на месячной частоте;
- BIV — базисный индекс объёма строительно монтажных работ на месячной частоте;
- ICV — объём инвестиций в основной капитал в среднегодовых ценах 2018 г. на месячной частоте;
- BIVMP — базисный индекс объёма денежных доходов населения на месячной частоте;
- APV — объём продукции сельского хозяйства в среднегодовых ценах 2018 г. на месячной частоте [5].

Период наблюдения данных следующий

- для квартальных: 1 квартал 2009 года — 2 квартал 2024 года;
- для месячных: 1 месяц 2009 года — 6 месяц 2024 года.

Для дальнейшего математического моделирования каждый экономический показатель мы будем описывать как отдельный временной ряд.

Считая все входные данные временными рядами, мы будем предполагать, что они являются нестационарными временными рядами, так как эти данные пришли из реального мира. Для того, чтобы строить модели для прогнозирования, нам необходимо привести эти временные ряды к стационарному виду. В соответствии со всей рассмотренной ранее теорией можно заключить, что мы сможем достичь стационарности во временных рядах путем исключения трендовой и сезонной компонент. Для проверки стационарности будем использовать ADF-тест.

2.1 Сезонная корректировка

В соответствии с 1.7.3 мы можем исключить сезонную компоненту при помощи метода TRAMO/SEAT. Применяя последовательно этот метод к каждому из рассматриваемых временных рядов GDP_t , IPV_t , RTV_t , BIV_t , ICV_t , $BIVMP_t$, APV_t , мы получим новые временные ряды GDP_SA_t , IPV_SA_t , RTV_SA_t , BIV_SA_t , ICV_SA_t , $BIVMP_SA_t$, APV_SA_t , которые будут являться сезонно скорректированными (Рис.2.1, Рис.2.2)

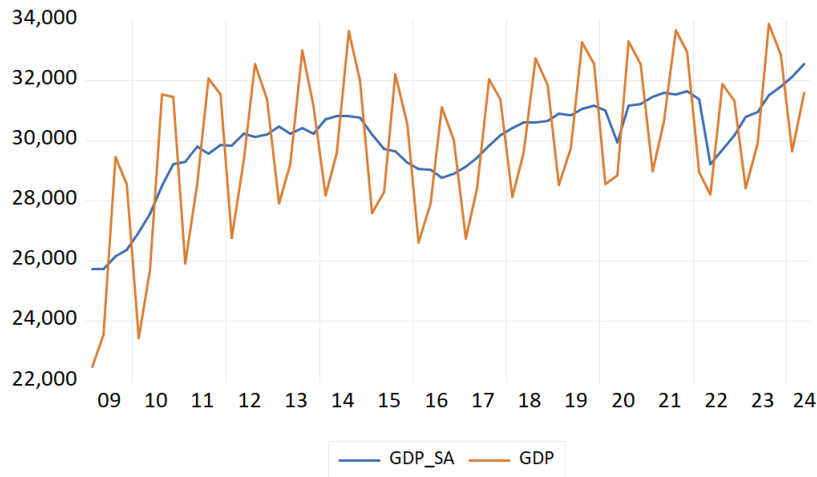


Рис. 2.1: ВВП Беларуси (GDP) и сезонно скорректированный ВВП Беларуси (GDP_SA)

Однако заметно, что в этих временных рядах все еще присутствует тренд. Поэтому следующим важным этапом будет исключение тренда.

2.2 Исключение тренда

В соответствии с 1.7.1 для исключения трендовой компоненты перейдем к темпам роста для каждого временного ряда. Переходя к темпам роста для каждого временного ряда, получим новые временные ряды

$$GGDP_SA_t = \frac{GDP_SA_t}{GDP_SA_{t-1}} \quad (2.2.1)$$

$$GIPV_SA_t = \frac{IPV_SA_t}{IPV_SA_{t-1}} \quad (2.2.2)$$

$$GRTV_SA_t = \frac{RTV_SA_t}{RTV_SA_{t-1}} \quad (2.2.3)$$

$$GICV_SA_t = \frac{ICV_SA_t}{ICV_SA_{t-1}} \quad (2.2.4)$$

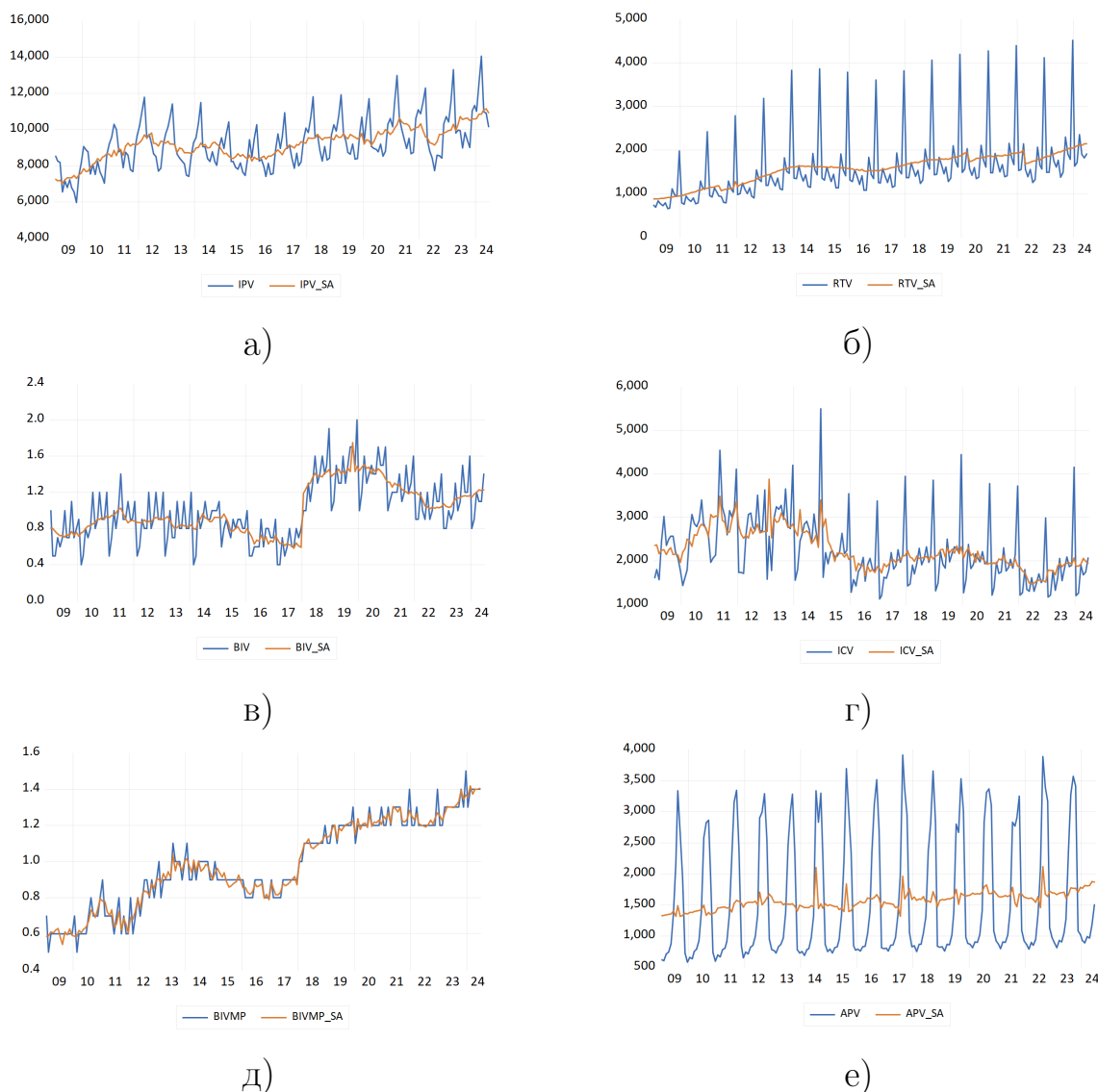


Рис. 2.2: — Исходные ряды и сезонно скорректированные:
а) IPV и IPV_SA ; б) RTV и RTV_SA ; в) BIV и BIV_SA ;
г) ICV и ICV_SA ; д) $BIVMP$ и $BIVMP_SA$; е) APV и APV_SA

$$GBIV_SA_t = \frac{BIV_SA_t}{BIV_SA_{t-1}} \quad (2.2.5)$$

$$GBIVMP_SA_t = \frac{BIVMP_SA_t}{BIVMP_SA_{t-1}} \quad (2.2.6)$$

$$GAPV_SA_t = \frac{APV_SA_t}{APV_SA_{t-1}} \quad (2.2.7)$$

Временной ряд $GGDP_SA$ отражает квартальные темпы роста ВВП Беларуси (Рис. 2.3). Аналогично все месячные переменные также отображают темпы роста продукции или цен месяц к месяцу (Рис. 2.4).

Таким образом, мы получили новые временные ряды, в которых отсутствуют трендовая и сезонная компоненты. Однако этого мало для того, чтобы сделать заключение о стационарности временных рядов. Необходимо

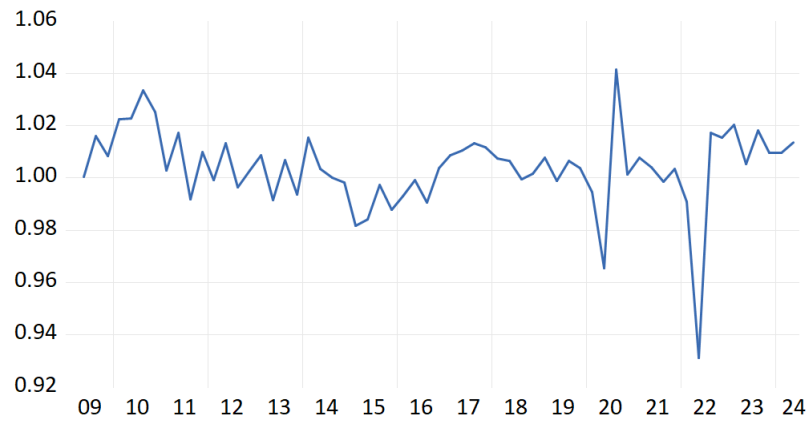
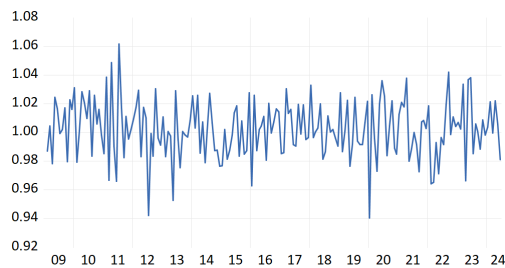
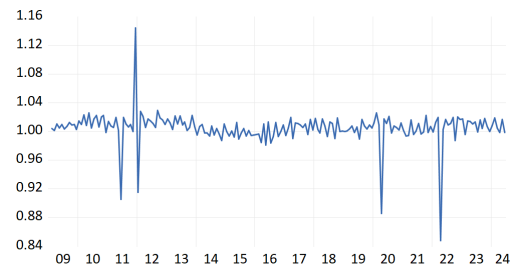


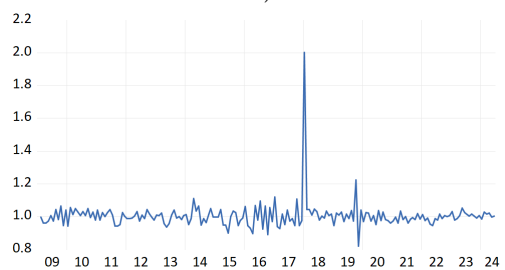
Рис. 2.3: Темпы роста ВВП ($GGDP_SA$) квартал к кварталу



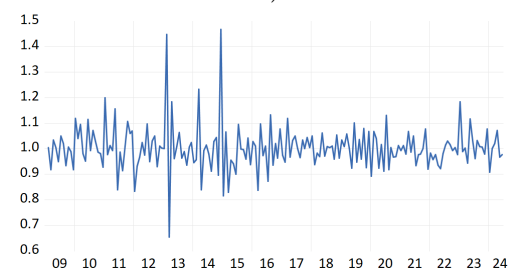
а)



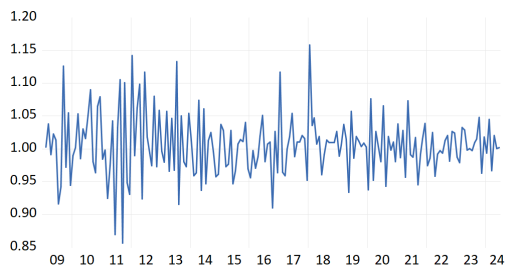
б)



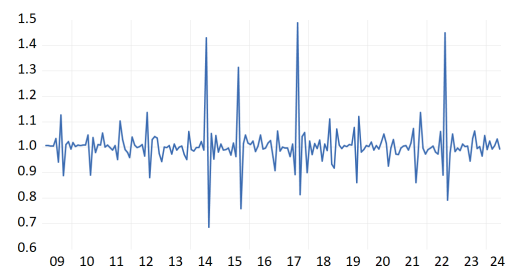
в)



г)



д)



е)

Рис. 2.4: — Темпы роста месяц к месяцу:
а) $GIPV_SA$; б) $GRTV_SA$; в) $GBIV_SA$;
г) $GICV_SA$; д) $GBIVMP_SA$; е) $GAPV_SA$

также статистически подкрепить предположения о стационарности полученных рядов, для чего мы воспользуемся тестом ADF.

2.3 Тестирование стационарности

Для тестирования стационарности и определения типа нестационарности мы проведем расширенный тест Дики-Фуллера (ADF). Этот тест является стандартным инструментом анализа временных рядов, однако следует отметить, что он может быть склонен к ошибочным выводам в случаях наличия структурных изменений, что часто приводит к предпочтению моделей с единичным корнем (DS-моделей).

В связи с этой проблемой мы также применим модифицированный ADF-тест, известный как Brake Point Unit Root (BPUR-тест). Этот тест позволяет выявить одно наиболее выраженное структурное изменение во временном ряде, описываемом DS-моделью, что делает его полезным инструментом для более точного анализа.

Алгоритмы работы указанных тестов и их применение подробно представлены в разделах 1.6 и 1.6.6 соответственно.

Таблицы 2.1 и 2.2 содержат результаты тестирования временных рядов, проведенного с использованием тестов ADF и BPUR. Нулевая гипотеза для этих тестов формулируется следующим образом: «временной ряд является нестационарным и описывается DS-моделью». Это означает, что если нулевая гипотеза оказывается верной, то временной ряд не демонстрирует стабильности во времени и может находиться под воздействием различных структурных изменений.

Таблица 2.1 – Результаты тестирования временных рядов с помощью ADF-теста

Временной ряд	Значение ADF-статистики	Пороговые значения, $\varepsilon = 0.05$	Тип модели
GGDP_SA _t	-6,796	-2,911	TS
IPV_SAG _t	-16,478	-2,877	TS
RTV_SAG _t	-16,678	-2,877	TS
ICV_SAG _t	-22,472	-2,877	TS
BIV_SAG _t	-14,380	-2,877	TS
BIVMP_SAG _t	-20,402	-2,877	TS
APV_SAG _t	-15,410	-2,877	TS

Если статистика критерия в результате тестирования превышает установленное пороговое значение, то в таком случае нулевая гипотеза не отклоняется на уровне значимости 0,05. Это указывает на то, что времен-

Таблица 2.2 – Результаты тестирования временных рядов с помощью BPUR-теста

Временной ряд	Значение ADF-статистики	Пороговые значения, $\varepsilon = 0.05$	Моменты структурных изменений	Тип модели
GGDP_SA _t	-9,138	-4,444	2022q2	TS
IPV_SAG _t	-17,236	-4,444	2012m7	TS
RTV_SAG _t	-18,945	-4,444	2022m4	TS
ICV_SAG _t	-24,902	-4,444	2013m2	TS
BIV_SAG _t	-15,405	-4,444	2017m1	TS
BIVMP_SAG _t	-20,985	-4,444	2012m7	TS
APV_SAG _t	-24,331	-4,444	2022m10	TS

ной ряд является нестационарным и, следовательно, требует применения соответствующих методов анализа. В противном случае, если статистика критерия оказывается ниже порогового значения, временной ряд считается стационарным или может быть отнесен к TS-моделям.

Результаты тестирования указывают на то, что все временные ряды являются стационарными. Это означает, что их статистические свойства, такие как среднее и дисперсия, остаются постоянными во времени. Ниже приведены ключевые выводы и значения стационарности:

1. *Корректность модели.* Стационарные ряды позволяют использовать модели временных рядов, включая MFVAR, без необходимости их трансформации.
2. *Прогнозирование.* Стационарные ряды обеспечивают более надёжные прогнозы благодаря предсказуемости их поведения.
3. *Стабильность взаимосвязей.* Стационарные ряды предполагают стабильность взаимосвязей между переменными, что важно для интерпретации результатов.

ГЛАВА 3

МОДЕЛИРОВАНИЕ И ПРОГНОЗИРОВАНИЕ ТЕМПОВ РОСТА ВВП БЕЛАРУСИ ПО ИСТОЧНИКАМ ДОХОДОВ

В данной главе будут представлены результаты прогнозирования с помощью моделей MFVAR, описанных в 1.9. В качестве низкочастотной эндогенной переменной в данной главе будет выступать сезонно скорректированные темпы роста ВВП Беларуси, полученные в предыдущей главе. Остальные же переменные будут выступать в качестве экзогенных. При чем добавлять в модель экзогенные переменные высокой частоты мы будем последовательно. Выбор порядка добавления параметров обусловлен их экономической значимостью, предсказательной силой и взаимосвязи с динамикой ВВП.

1. *Темпы роста объёма промышленного производства.* Промышленный сектор является ключевым элементом экономики, который оказывает непосредственное влияние на формирование ВВП. Темпы роста объёма промышленного производства отражают динамику производства товаров и услуг, что, в свою очередь, напрямую связано с объемом добавленной стоимости. Увеличение $GIPV_SA$ указывает на активизацию производственной деятельности, что приводит к увеличению ВВП в краткосрочной перспективе. Поэтому этот показатель становится первым в порядке добавления.
2. *Темпы роста объёма розничного товарооборота.* Розничный товарооборот является индикатором потребительского спроса и активности в сфере услуг. Он отражает уровень потребления домохозяйств и может служить предвестником экономической активности. Увеличение розничных продаж часто ведет к росту производства и, как следствие, к росту ВВП.
3. *Темпы роста объёма сельскохозяйственных работ.* Сельское хозяйство играет значительную роль в экономике Беларуси, однако его влияние на ВВП может быть менее предсказуемым в краткосрочной перспективе. Темпы роста объёма сельскохозяйственных работ могут колебаться из-за сезонных факторов и погодных условий. Тем не менее рост в этом секторе может позитивно сказаться на ВВП.
4. *Темпы роста строительно-монтажных работ.* Строительство является важным драйвером экономического роста, так как этот сектор создает рабочие места и способствует инвестициям. Темпы роста объёма строительно-монтажных работ могут служить индикатором

будущих экономических изменений, так как рост в строительстве часто предшествует увеличению экономической активности в других секторах.

5. *Темпы роста доходов населения.* Темпы роста доходов населения влияют на потребление и инвестиции, однако они могут проявлять свою предсказательную силу с некоторым запаздыванием. Повышение доходов может способствовать увеличению потребительского спроса, что положительно сказывается на росте ВВП.
6. *Темпы роста инвестиций в основной капитал.* Инвестиции в основной капитал являются важным индикатором долгосрочного экономического роста, но их краткосрочная предсказательная сила может быть ограниченной. Хотя инвестиции могут способствовать росту производительных мощностей и, следовательно, ВВП в долгосрочной перспективе, их влияние на краткосрочные изменения в ВВП может быть менее заметным.

Помимо вышеперечисленных переменных мы также в качестве экзогенной переменной добавим фиктивную переменную с квартальной частотой DUM2022Q2. Так как из Таблицы 2.2 можно увидеть, что момент 2022q2 является моментом структурных изменений, что может сильно сказаться на прогнозах. Поэтому мы добавляем переменную, которая принимает на каждом квартале значение 0, а на 2-ом квартале 2022-ого года – значение 1.

В соответствии с методологией построения наукастов, прогнозы мы будем строить по описанному в 1.11 алгоритму: исключать последний квартал из обучающей выборки, обучать модель, а затем оценивать модель оп последним 12 точкам и отдельно по 13-ой вневыборочной точке.

MFVAR модели также требуют от исследователя задать количество лагов. Ниже будут представлены модели с заранее определенным количеством лагом, основанном на сравнении метрик. То есть будет выбрано такое число лагов, при котором значение метрик наилучшее.

3.1 MFVAR на основе промышленности, товарооборота и сельского хозяйства

Для построения MFVAR модели в качестве экзогенных выступали переменные GIPV_SA (промышленность), GRTV_SA (товарооборот), GAPV_SA (сельское хозяйство), а также дамми-переменная DUM2022q2, а в качестве эндогенной переменной - GGDP_SA.

Число лагов выберем равное 4.

Результаты построения модели иллюстрируется на Рис.3.1.

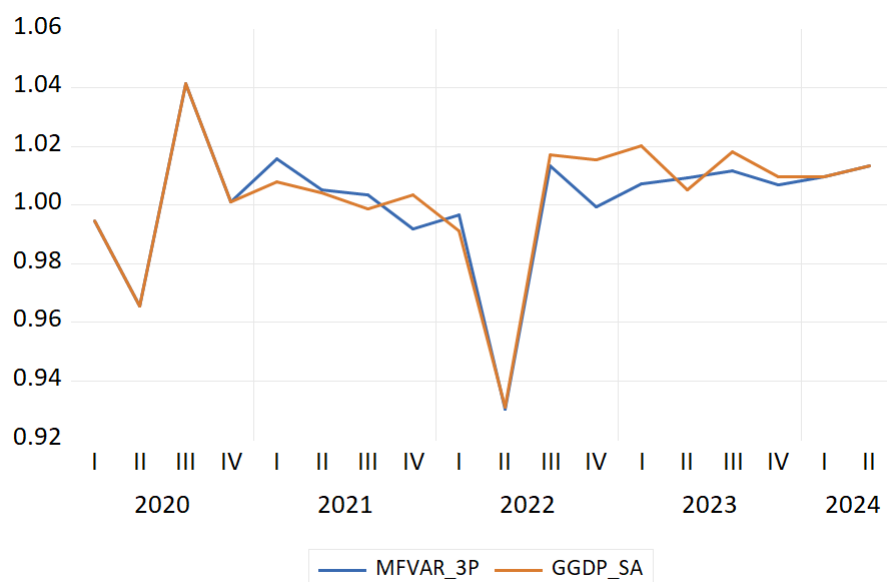


Рис. 3.1: Ретроспективный прогноз модели MFVAR по 3 параметрам

Рассмотрим поведение остатков модели на наличие закономерностей на Рис. 3.2.

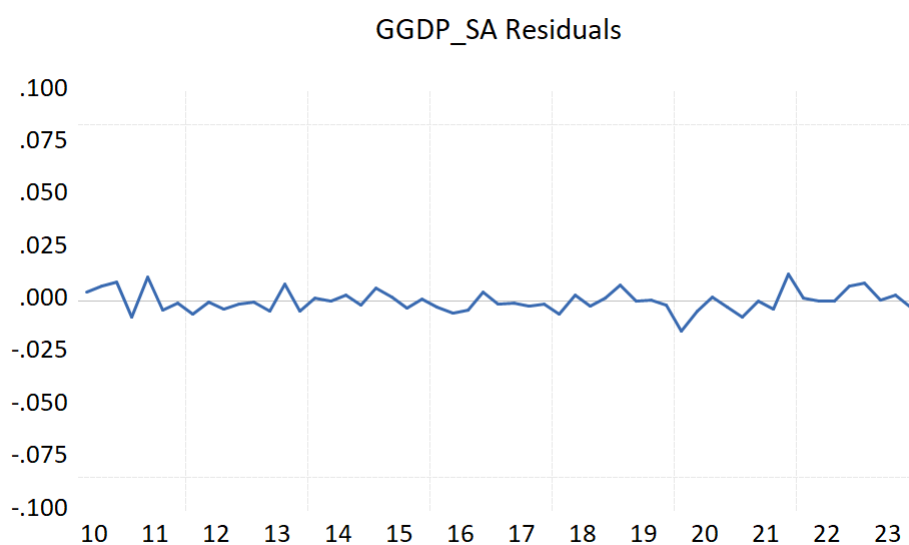


Рис. 3.2: Остатки модели MFVAR по 3 параметрам

Как можно видеть, никаких явных закономерностей нет, поэтому мы можем считать построенную MFVAR модель адекватной для прогнозирования.

3.2 MFVAR на основе промышленности, товарооборота сельского хозяйства, строительства

Для построения MFVAR модели в качестве экзогенных выступали переменные GIPV_SA (промышленность), GRTV_SA (товарооборот), GAPV_SA

(сельское хозяйство), GBIV_SA (строительство) а также дамми-переменная DUM2022q2, а в качестве эндогенной переменной – GGDP_SA.

Число лагов выберем равное 4.

Результаты построения модели иллюстрируется на Рис.3.3.

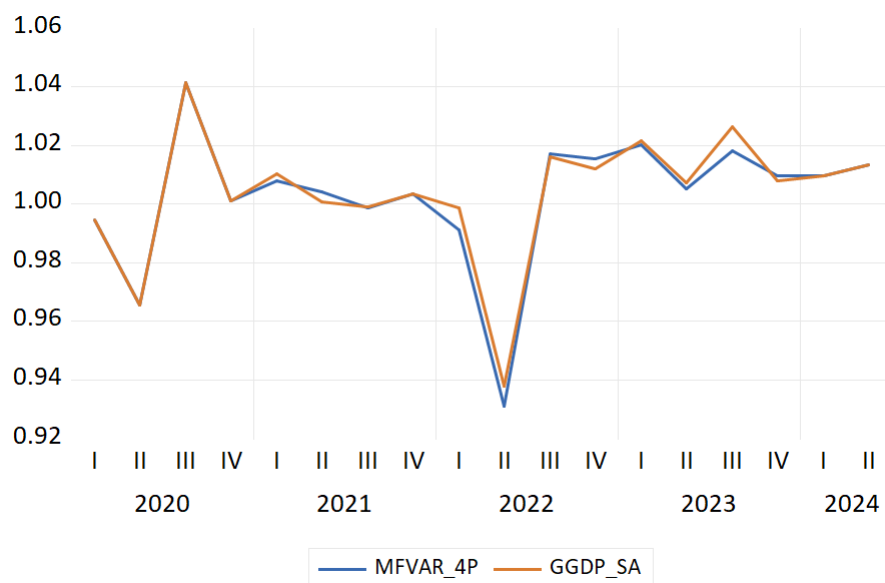


Рис. 3.3: Ретроспективный прогноз модели MFVAR по 4 параметрам

Рассмотрим поведение остатков модели на наличие закономерностей на Рис. 3.4.

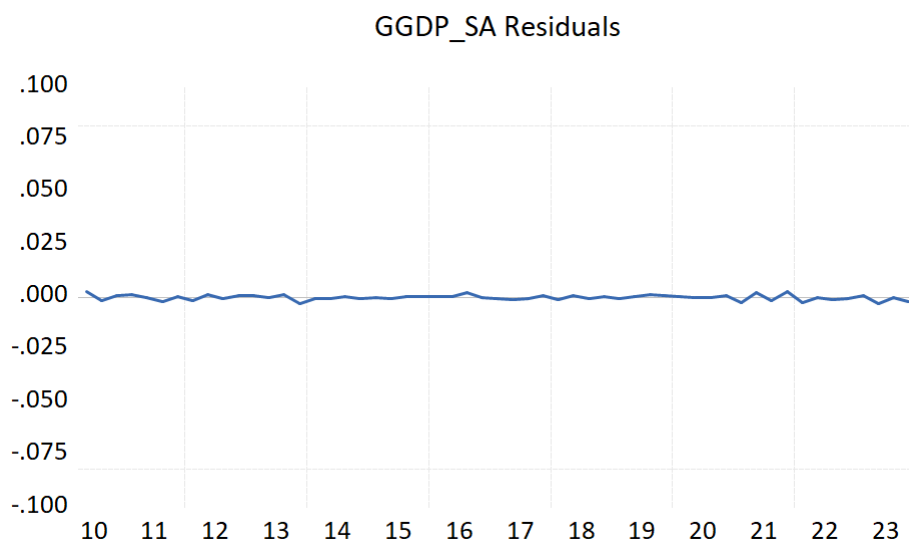


Рис. 3.4: Остатки модели MFVAR по 4 параметрам

Как можно видеть, никаких явных закономерностей нет, поэтому мы можем считать построенную MFVAR модель адекватной для прогнозирования.

3.3 MFVAR на основе промышленности, товарооборота, сельского хозяйства, строительства, объема инвестиций

Для построения MFVAR модели в качестве экзогенных выступали переменные GIPV_SA (промышленность), GRTV_SA (товарооборот), GAPV_SA (сельское хозяйство), GBIV_SA (строительство), GICV_SA (объем инвестиций), а также дамми-переменная DUM2022q2, а в качестве эндогенной переменной – GGDP_SA.

Число лагов выберем равное 3.

Результаты построения модели иллюстрируется на Рис. 3.5.

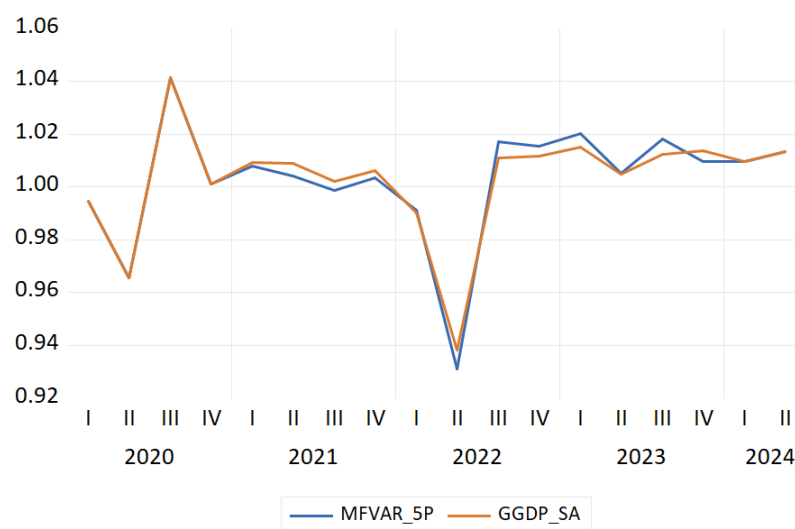


Рис. 3.5: Ретроспективный прогноз модели MFVAR по 5 параметрам

Рассмотрим поведение остатков модели на наличие закономерностей на Рис. 3.6.

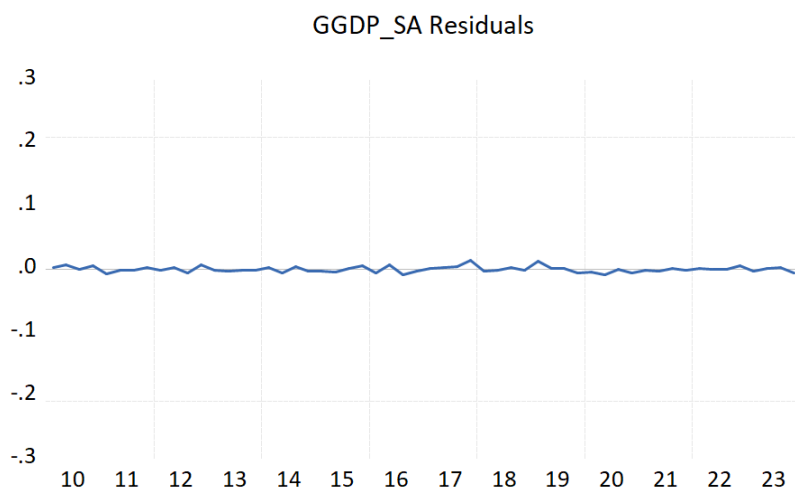


Рис. 3.6: Остатки модели MFVAR по 5 параметрам

Как можно видеть, никаких явных закономерностей нет, поэтому мы можем считать построенную MFVAR модель адекватной для прогнозирования.

3.4 MFVAR на основе промышленности, товарооборота, сельского хозяйства, строительства, объема инвестиций, доходов населения

Для построения MFVAR модели в качестве экзогенных выступали переменные GIPV_SA (промышленность), GRTV_SA (товарооборот), GAPV_SA (сельское хозяйство), GBIV_SA (строительство), GICV_SA (объем инвестиций), GBIVM_SA (доходы населения), а также дамми-переменная DUM2022q2, а в качестве эндогенной переменной – GGDP_SA.

Число лагов выберем равное 2.

Результаты построения модели иллюстрируется на Рис.3.7.

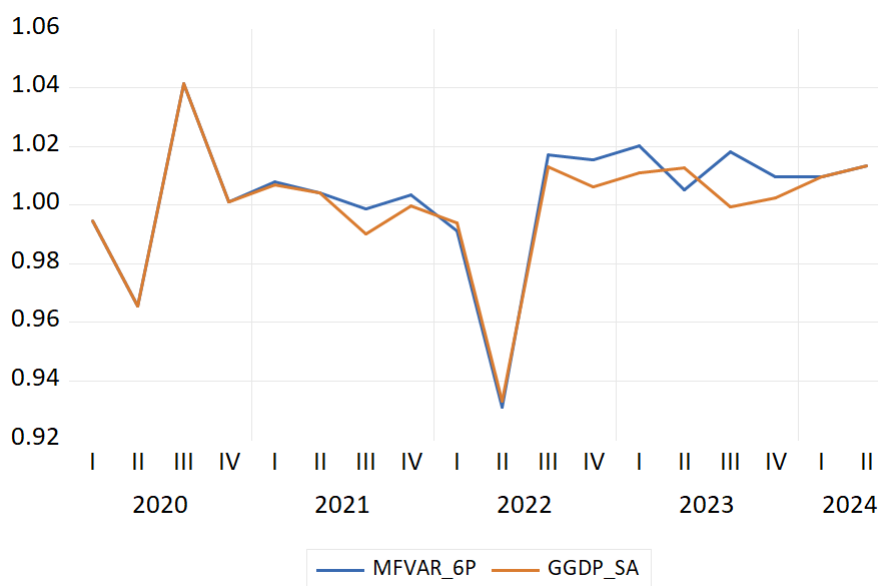


Рис. 3.7: Ретроспективный прогноз модели MFVAR по 6 параметрам

Рассмотрим поведение остатков модели на наличие закономерностей на Рис. 3.8.

Как можно видеть, никаких явных закономерностей нет, поэтому мы можем считать построенную MFVAR модель адекватной для прогнозирования.

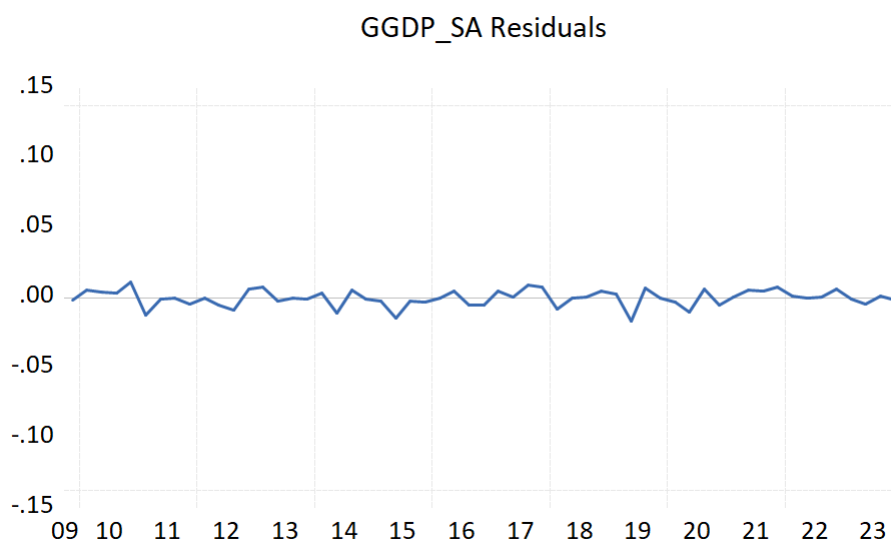


Рис. 3.8: Остатки модели MFVAR по 6 параметрам

3.5 Оценки моделей и сравнительный анализ

Таким образом, мы построили MFVAR модели и проделали с помощью них ретроспективные прогнозы. Теперь проделаем сам наукаст: спрогнозируем значение на следующий квартал и оценим его точность, а затем сравним результаты.

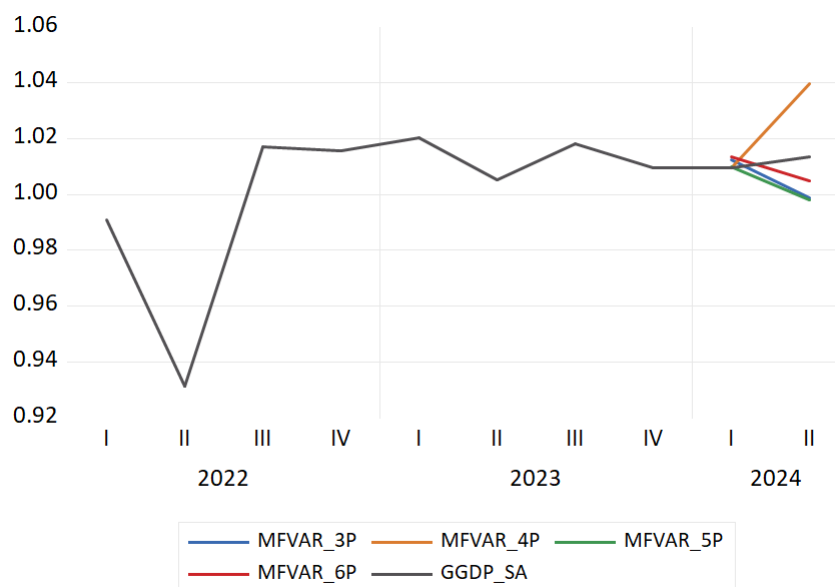


Рис. 3.9: Краткосрочные прогнозы всех MFVAR моделей на вневыборочном квартале

Заметно, что большинство моделей спрогнозировало спад ВВП, но четырехпараметрическая модель наоборот спрогнозировала сильный рост.

Рассмотрим таблицу с оценками всех указанных моделей и сделаем на основании этого выводы.

Таблица 3.1 — Показатели точности прогнозов годовых темпов роста ВВП РБ на основе альтернативных моделей, период оценивания моделей: 2021q1-2023q4

Прогнозный период: 2021q1 – 2023q4 (ретроспективные прогнозы)					
Модель	Число наблюдений	RMSE	MAE	MAPE, %	Lag
MFVAR_3P	12	0.007958	0.006491	0.647089	4
MFVAR_4P	12	0.004152	0.003195	0.320232	4
MFVAR_5P	12	0.004336	0.003847	0.385297	3
MFVAR_6P	12	0.007863	0.006145	0.613890	2
Прогнозный период: 2024q1 – 2024q2 (краткосрочный прогноз)					
MFVAR_3P	2	0.010541	0.008671	0.866359	4
MFVAR_4P	2	0.018522	0.013152	1.265229	4
MFVAR_5P	2	0.010820	0.007860	0.787171	3
MFVAR_6P	2	0.006649	0.006215	0.616806	2

Исходя из данных таблицы можем сделать следующие выводы:

- по оценкам RMSE и MAPE на ретроспективном прогнозе лучше всех оказалась модель с 4 переменными;
- по оценкам RMSE и MAPE на краткосрочном прогнозе лучше всех оказалась модель с 6 переменными, которая спрогнозировала самый слабый спад относительно всех остальных моделей.

ЗАКЛЮЧЕНИЕ

В данной работе была рассмотрена задача прогнозирования компонент ВВП на основе регрессионных моделей MFVAR. В ходе исследования

1. было дано теоретическое описание всех используемых для решения данной задачи моделей;
2. были рассмотрены свойства и особенности, которые возникают в ходе работы с исследуемыми моделями;
3. был проведен полный цикл исследования и преобразования моделей временных рядов для приведения к стационарному виду;
4. были построены модели MFVAR для прогнозирования и оценки темпов роста ВВП Беларуси по его компонентам, таким как объем промышленности, объем товарооборота, сельского хозяйства, объем строительства, объем инвестиций, объем доходов населения;
5. был проведен сравнительный анализ точности прогнозных значений и ретроспективных прогнозов моделей MFVAR в зависимости от переменных.

Таким образом, данная работа вносит вклад в развитие методов прогнозирования на основе моделей временных рядов по смешанным данным и может быть использована в дальнейших исследованиях и практических применениях в области экономики и финансов.

СПИСОК ИСТОЧНИКОВ

1. Тест Дики — Фуллера. [Электронный ресурс] – https://ru.wikipedia.org/wiki/Тест_Дики_-_Фуллера – Дата доступа: 02.11.2024.
2. Понятие стационарности временного ряда. Процессы «единичного корня». [Электронный ресурс] – <https://bsu.by/upload/page/544153.pdf> – Дата доступа: 02.11.2024
3. Процессы «единичного корня». Тесты «единичного корня»: ADF, PP, KPSS. [Электронный ресурс] – <https://bsu.by/upload/page/546923.pdf> – Дата доступа: 02.11.2024
4. Харин, Ю. С. Теория вероятностей, математическая и прикладная статистика / Ю. С. Харин, Н. М. Зуев, Е. Е. Жук // Минск : БГУ, 2011.
5. Интерактивная информационно-аналитическая система распространения официальной статистической информации. [Электронный ресурс] – <https://dataportal.belstat.gov.by/> – Дата доступа: 02.12.2024.
6. Макеева, Н.М., Наукастинг элементов использования ВВП России / Н.М. Макеева, И.П. Станкевич // Статья 2022/10, Экономический журнал ВШЭ.
7. Станкевич И.П. Сравнение методов наукастинга макроэкономических индикаторов на примере российского ВВП // Прикладная эконометрика 2020. С. 113–127.
8. Малюгин, В. Краткосрочное прогнозирование и наукастинг темпов роста инфляции на основе моделей по смешанным данным / В.И. Малюгин // Банкаўскі веснік. – 2024. – С. 1–13.
9. Ghysels, E. Regression models with mixed sampling frequencies / E. Andreou, A. Kourtellos // Journal of Econometrics 2010.
10. Soybilgen, B. Nowcasting the New Turkish GDP / B. Soybilgen, E. Yazgan // Economics Bulletin, Volume 38, Issue 2, С. 1083-1089
11. Kuzin, V. MIDAS vs. Mixed-Frequency VAR: Nowcasting GDP in the Euro Area / V. Kuzin, M. Marcellino, C. Shumacher // EUI Working Paper.
12. X-13ARIMA-SEATS. [Электронный ресурс] – <https://en.wikipedia.org/wiki/X-13ARIMA-SEATS> – Дата доступа: 26.11.2024.

13. «Сгладить нельзя не сгладить!» или пара слов о том, как строить свои отношения с сезонностью в данных. [Электронный ресурс] – https://www.econ.msu.ru/ext/lib/Category/x81/xf9/33273/file/Встреча%207_%20Сезонность.pdf – Дата доступа: 26.11.2024.
14. Perron, P. Testing for a Unit Root in a Time Series with a Changing Mean. – Journal of Business and Economic Statistics, Volume 8, Issue 2. – pp. 153–162.
15. Perron, P. A Note on the Additive Outlier Model with Breaks. / Perron, P., Vogelsang, T.J. – Revista de Econometria Volume 13. – pp. 181–201.
16. Perron, P. The Great Crash, the Oil Price Shock and the Unit Root Hypothesis: Erratum / Perron, P., Vogelsang, T.J. – Econometrica, Volume 61. – pp. 248–249.