Временные ряды

Временно́й ряд – это совокупность значений, какого – либо показателя за несколько последовательных моментов или периодов времени. Каждое значение (уровень) временного ряда формируется под воздействием большого числа факторов, которые можно условно разделить на три группы:

- факторы, формирующие тенденцию ряда;
- факторы, формирующие циклические колебания ряда;
- случайные факторы.

Тенденция характеризует долговременное воздействие факторов на динамику показателя. Тенденция может быть возрастающей или убывающей.

Циклические колебания могут носить сезонный характер или отражать динамику конъюнктуры рынка, а также фазу бизнес – цикла, в которой находится экономика страны.

Реальные данные часто содержат все три компоненты. В большинстве случаев временной ряд можно представить как сумму или произведение трендовой , циклическойи случайнойкомпонент. В случае суммы имеет место**аддитивная** модель временного ряда:

(1)

в случае произведения – **мультипликативная** модель:

. (2)

Основная задача эконометрического исследования отдельного временного ряда – выявление количественного выражения каждой из компонент и использование полученной информации для прогноза будущих значений ряда или построение модели взаимосвязи двух или более временных рядов.

Сначала рассмотрим основные подходы к анализу отдельного временного ряда. Такой ряд может содержать, помимо случайной составляющей, либо только тенденцию, либо только сезонную (циклическую) компоненту, либо все компоненты вместе. Для того, чтобы выявить наличие той или иной неслучайной компоненты, исследуется корреляционная зависимость между последовательными уровнями временного ряда, или **автокорреляция уровней ряда.** Основная идея такого анализа заключается в том, что при наличии во временном ряде тенденции и циклических колебаний значения каждого последующего уровня ряда зависят от предыдущих.

Количественно автокорреляцию можно измерить с помощью линейного коэффициента корреляции между уровнями исходного временного ряда и уровнями этого ряда, сдвинутыми на несколько шагов во времени.

Коэффициент автокорреляции уровней ряда первого порядка измеряет зависимость между соседними уровнями ряда ит.е. при лаге 1.

Он вычисляется по следующей формуле:

(3)

где в качестве средних величин берутся значения:

(4)

В первом случае усредняются значения ряда, начиная со второго до последнего, во втором случае – значения ряда с первого до предпоследнего.

Формулу (3) можно представить как формулу выборочного коэффициента корреляции:

(5)

где в качестве переменной берется ряда в качестве переменнойряд

Если значение коэффициента (3) близко к единице, это указывает на очень тесную зависимость между соседними уровнями временного ряда и о наличии во временном ряде сильной линейной тенденции.

Аналогично определяются коэффициенты автокорреляции более высоких порядков. Так, коэффициент автокорреляции второго порядка характеризует тесноту связи между уровнями ии определяется по формуле:

(6)

где в качестве одной средней величины берут среднюю уровней ряда с третьего до последнего, а в качестве другой – среднюю с первого уровня до

(7)

Число периодов, по которым рассчитывается коэффициент автокорреляции, называют **лагом.** С увеличением лага число пар значений, по которым рассчитывается коэффициент автокорреляции, уменьшается. Для обеспечения статистической достоверности максимальный лаг, как считают некоторые известные эконометристы, не должен превышать четверти общего объема выборки.

Коэффициент автокорреляции строится по аналогии с линейным коэффициентом корреляции, и поэтому он характеризует тесноту только линейной связи текущего и предыдущего уровней ряда. По нему можно судить о наличии линейной или близкой к линейной тенденции. Однако для некоторых временных рядов с сильной нелинейной тенденцией (например, параболической или экспоненциальной), коэффициент автокорреляции уровней ряда может приближаться к нулю.

Кроме того, по знаку коэффициента автокорреляции нельзя делать вывод о возрастающей или убывающей тенденции в уровнях ряда. Большинство временных рядов экономических данных имеют положительную автокорреляцию уровней, однако при этом не исключается убывающая тенденция.

Последовательность коэффициентов автокорреляции уровней различных порядков, начиная с первого, называется **автокорреляционной функцией временного ряда.** График зависимости ее значений от величины лага называется **коррелограммой.** Анализ автокорреляционной функции и коррелограммы помогает выявить структуру ряда. Здесь уместно привести следующие качественные рассуждения.

Если наиболее высоким является коэффициент автокорреляции первого порядка, очевидно, исследуемый ряд содержит только тенденцию. Если наиболее высоким оказался коэффициент автокорреляции порядка τ,ряд содержит циклические колебания с периодичностью вτ моментов времени. Если ни один из коэффициентов автокорреляции не является значимым, то либо ряд не содержит тенденции и циклических колебаний и имеет только случайную составляющую, либо ряд содержит сильную нелинейную тенденцию, для исследования которой нужно провести дополнительный анализ.

Пример 1. Пусть имеются данные об объёмах потребления электроэнергии жителями района за 16 кварталов, млн. квт.-ч:

t	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16
y_t	6,0	4,4	5,0	9,0	7,2	4,8	6,0	10,0	8,0	5,6	6,4	11,0	9,0	6,6	7,0	10,8

Нанесем эти значения на график:

Определим автокорреляционную функцию данного временного ряда. Рассчитаем коэффициент автокорреляции первого порядка. Для этого определим средние значения:

С учетом этих значений можно построить вспомогательную таблицу:

t	y_t					
1	6,0		-1,0667			1,137778
2	4,4	-2,9867	-2,6667	3,185778	8,920178	7,111111
3	5,0	-2,3867	-2,0667	6,364444	5,696178	4,271111
4	9,0	1,6133	1,9333	-3,33422	2,602844	3,737778
5	7,2	-0,1867	0,1333	-0,36089	0,034844	0,017778
6	4,8	-2,5867	-2,2667	-0,34489	6,690844	5,137778
7	6,0	-1,3867	-1,0667	3,143111	1,922844	1,137778
8	10,0	2,6133	2,9333	-2,78756	6,829511	8,604444
9	8,0	0,6133	0,9333	1,799111	0,376178	0,871111
10	5,6	-1,7867	-1,4667	-1,66756	3,192178	2,151111
11	6,4	-0,9867	-0,6667	1,447111	0,973511	0,444444
12	11,0	3,6133	3,9333	-2,40889	13,05618	15,47111

t	y_t					
13	9,0	1,6133	1,9333	6,345778	2,602844	3,737778
14	6,6	-0,7867	-0,4667	-1,52089	0,618844	0,217778
15	7,0	-0,3867	-0,0667	0,180444	0,149511	0,004444
16	10,8	3,4133		-0,22756	11,65084	
Итого				9,813333	65,3173	54,0533

С помощью итоговых сумм подсчитаем величину коэффициента автокорреляции первого порядка:

.

Это значение свидетельствует о слабой зависимости текущих уровней ряда от непосредственно им предшествующих. Однако из графика очевидно наличие возрастающей тенденции уровней ряда, на которую накладываются циклические колебания.

Продолжая аналогичные расчеты для второго, третьего и т.д. порядков, получим автокорреляционную функцию, значения которой сведем в таблицу и построим по ней коррелограмму:

Лаг	1	2	3	4	5	6	7	8
	0,16515	0,56687	0,11355	0,98302	0,11871	0,72204	0,00336	0,97384

Из коррелограммы видно, что наиболее высокий коэффициент корреляции наблюдается при значении лага, равном четырем, следовательно, ряд имеет циклические колебания периодичностью в четыре квартала. Это подтверждается и графическим анализом структуры ряда.

Вслучае если при анализе структуры временного ряда обнаружена только тенденция и отсутствуют циклические колебания (случайная составляющая присутствует всегда), следует приступать к моделированию тенденции. Если же во временном ряде имеют место и циклические колебания, прежде всего, следует исключить именно циклическую составляющую, и лишь затем приступать к моделированию тенденции. Выявление тенденции состоит в построении аналитической функции, характеризующей зависимость уровней ряда от времени, или**тренда**. Этот способ называют **аналитическим выравниванием временного ряда**.

Зависимость от времени может принимать разные формы, поэтому для её формализации используют различные виды функций:

- линейный тренд: ;
- гипербола: ;
- экспоненциальный тренд: (или);
- степенной тренд: ;

- параболический тренд второго и более высоких порядков:

.

Параметры каждого из трендов можно определить обычным МНК, используя в качестве независимой переменной время , а в качестве зависимой переменной – фактические уровни временного ряда y_t (или уровни за вычетом циклической составляющей, если таковая была обнаружена). Для нелинейных трендов предварительно проводят стандартную процедуру их линеаризации.

Существует несколько способов определения типа тенденции. Чаще всего используют качественный анализ изучаемого процесса, построение и визуальный анализ графика зависимости уровней ряда от времени, расчет некоторых основных показателей динамики. В этих же целях можно использовать и коэффициенты автокорреляции уровней ряда. Тип тенденции можно определить путем сравнения коэффициентов автокорреляции первого порядка, рассчитанных по исходным и преобразованным уровням ряда. Если временной ряд имеет линейную тенденцию, то его соседние уровни y_t и y_{t-1} тесно коррелируют. В этом случае коэффициент автокорреляции первого порядка уровней исходного ряда должен быть высоким. Если временной ряд содержит нелинейную тенденцию, например, в форме экспоненты, то коэффициент автокорреляции первого порядка по логарифмам уровней исходного ряда будет выше, чем соответствующий коэффициент, рассчитанный по уровням ряда. Чем сильнее выражена нелинейная тенденция в изучаемом временном ряде, тем в большей степени будут различаться значения указанных коэффициентов.

Выбор наилучшего уравнения в случае, если ряд содержит нелинейную тенденцию, можно осуществить путем перебора основных форм тренда, расчета по каждому уравнению скорректированного коэффициента детерминации и выбора уравнения тренда с максимальным значением этого коэффициента. Реализация этого метода относительно проста при компьютерной обработке данных.

При анализе временных рядов, содержащих сезонные или циклические колебания, наиболее простым подходом является расчет значений сезонной компоненты методом скользящей средней и построение аддитивной или мультипликативной модели временнóго ряда в форме (1) или (2).

Если амплитуда колебаний приблизительно постоянна, строят аддитивную модель (1), в которой значения сезонной компоненты предполагаются постоянными для различных циклов. Если амплитуда сезонных колебаний возрастает или уменьшается, строят мультипликативную модель (2), которая ставит уровни ряда в зависимость от значений сезонной компоненты.

Построение модели (1) или (2) сводится к расчету значений T, S или E для каждого уровня ряда. Процесс построения модели включает в себя следующие шаги:

1. Выравнивание исходного ряда методом скользящей средней.

2. Расчет значений сезонной компоненты S.
3. Устранение сезонной компоненты из исходных уровней ряда и получение выровненных данных ($T+E$) в аддитивной или ($T \cdot E$) в мультипликативной модели.
4. Аналитическое выравнивание уровней ($T+E$) или ($T \cdot E$) и расчет значений T с использованием полученного уравнения тренда.
5. Расчет полученных по модели значений ($T+S$) или ($T \cdot S$)
6. Расчет абсолютных и относительных ошибок.

Пример 2. ***Построение аддитивной модели временного ряда***. Рассмотрим данные об объёме потребления электроэнергии жителями района из ранее приведенного примера. Из анализа автокорреляционной функции было показано, что данный временно́й ряд содержит сезонные колебания периодичностью в 4 квартала. Объёмы потребления электроэнергии в осенне – зимний период (I и IV кварталы) выше, чем весной и летом (II и III кварталы). По графику этого ряда можно установить наличие приблизительно равной амплитуды колебаний. Это говорит о возможном наличии аддитивной модели. Рассчитаем её компоненты.

Шаг 1. Проведем выравнивание исходных уровней ряда методом скользящей средней.

Поскольку циклические колебания имеют периодичность в 4 квартала, просуммируем уровни ряда последовательно за каждые 4 квартала со сдвигом на один момент времени и определим условные годовые объёмы потребления электроэнергии (колонка 3 в таблице 1).

Разделив полученные суммы на 4, найдем скользящие средние (колонка 4 таблицы 1). Полученные таким образом выровненные значения уже не содержат сезонной компоненты.

Поскольку скользящие средние получены осреднением четырех соседних уровней ряда, т.е. четного числа значений, они соответствуют серединам подынтервалов, состоящих из четверок чисел, т.е. должны располагаться между третьим и четвертым значениями четверок исходного ряда. Для того, чтобы скользящие средние располагались на одних временных отметках с исходным рядом, пары соседних скользящих средних ещё раз усредняются и получаются центрированные скользящие средние (колонка 5 таблицы 1). При этом теряются первые две и последние две отметки временного ряда, что связано с осреднением по четырем точкам.

Таблица 1

№ квартала	Потребление электроэнергии y_t	Итого за четыре квартала	Скользящая средняя за четыре квартала	Центрированная скользящая средняя	Оценка сезонной компоненты
1	**2**	**3**	**4**	**5**	**6**
1	6,0				
2	4,4				
3	5,0	24,4	6,10	6,25	-1,250
4	9,0	25,6	6,40	6,45	2,550
5	7,2	26,0	6,50	6,625	0,575
6	4,8	27,0	6,75	6,875	-2,075
7	6,0	28,0	7,00	7,1	-1,100
8	10,0	28,8	7,20	7,3	2,700
9	8,0	29,6	7,40	7,45	0,550
10	5,6	30,0	7,50	7,625	-2,025
11	6,4	31,0	7,75	7,875	-1,475
12	11,0	32,0	8,00	8,125	2,875
13	9,0	33,0	8,25	8,325	0,675
14	6,6	33,6	8,40	8,375	-1,775
15	7,0	33,4	8,35		
16	10,8				

Шаг 2. Найдем оценки сезонной компоненты как разность между фактическими уровнями ряда (колонка 2 таблицы 1) и центрированными скользящими средними (колонка 5). Эти значения помещаем в колонку 6 таблицы 1 и используем для расчета значений сезонной компоненты (таблица 2), которые представляют собой средние за каждый квартал (по всем годам) оценки сезонной компоненты S_i. В моделях с сезонной компонентой обычно предполагается, что сезонные воздействия за период (в данном случае – за год) взаимопогашаются. В аддитивной модели это выражается в том, что сумма значений сезонной компоненты по всем точкам (здесь – по четырем кварталам) должна быть равна нулю.

Таблица 2

Показатели	Год	№ квартала, i			
		I	II	III	IV

	1	-	-	-1,250	2,550
	2	0,575	-2,075	-1,100	2,700
	3	0,550	-2,025	-1,475	2,875
	4	0,675	-1,775	-	-
Итого за *I*-й квартал (за все годы)		1,800	-5,875	-3,825	8,125
Средняя оценка сезонной компоненты для *I*-го квартала,		0,600	-1,958	-1,275	2,708
Скорректированная сезонная компонента,		0,581	-1,977	-1,294	2,690

Для данной модели сумма средних оценок сезонной компоненты равна:

$$0,6-1,958-1,275+2,708=0,075.$$

Эта сумма оказалась не равной нулю, поэтому каждую оценку уменьшим на величину поправки, равной одной четверти полученного значения:

$$\Delta=0,075/4=0,01875.$$

Рассчитаем скорректированные значения сезонной компоненты (они записаны в последней строке таблицы 2):

(8)

Эти значения при суммировании уже равны нулю:

$$0,581-1,977-1,294+2,69=0.$$

Шаг 3. Исключаем влияние сезонной компоненты, вычитая её значения из каждого уровня исходного временного ряда. Получаем величины:

$T+E=Y-S$ (9)

Эти значения рассчитываются в каждый момент времени и содержат только тенденцию и случайную компоненту (колонка 4 следующей таблицы):

Таблица 3

t				T	$T+S$		E^2
1	**2**	**3**	**4**	**5**	**6**	**7**	**8**
1	6,0	0,581	5,419	5,902	6,483	-0,483	0,2332
2	4,4	-1,977	6,377	6,088	4,111	0,289	0,0833
3	5,0	-1,294	6,294	6,275	4,981	0,019	0,0004
4	9,0	2,69	6,310	6,461	9,151	-0,151	0,0228
5	7,2	0,581	6,619	6,648	7,229	-0,029	0,0008

t				*T*	*T+S*		*E*²
1	**2**	**3**	**4**	**5**	**6**	**7**	**8**
6	4,8	-1,977	6,777	6,834	4,857	-0,057	0,0032
7	6,0	-1,294	7,294	7,020	5,726	0,274	0,0749
8	10,0	2,69	7,310	7,207	9,897	0,103	0,0107
9	8,0	0,581	7,419	7,393	7,974	0,026	0,0007
10	5,6	-1,977	7,577	7,580	5,603	-0,003	0,0000
11	6,4	-1,294	7,694	7,766	6,472	-0,072	0,0052
12	11,0	2,69	8,310	7,952	10,642	0,358	0,1278
13	9,0	0,581	8,419	8,139	8,720	0,280	0,0785
14	6,6	-1,977	8,577	8,325	6,348	0,252	0,0634
15	7,0	-1,294	8,294	8,512	7,218	-0,218	0,0474
16	10,8	2,69	8,110	8,698	11,388	-0,588	0,3458

Шаг 4. Определим трендовую компоненту данной модели. Для этого проведем выравнивание ряда ($T+E$) с помощью линейного тренда:

Подставляя в это уравнение значения , найдем уровниT для каждого момента времени (колонка 5 таблицы 3).

Шаг 5. Найдем значения уровней ряда, полученные по аддитивной модели. Для этого прибавим к уровням T значения сезонной компоненты для соответствующих кварталов, т.е. к значениям в колонке 5 таблицы 3 прибавим значения в колонке 3. Результаты операции представлены в колонке 6 таблицы 3.

Шаг 6. В соответствии с методикой построения аддитивной модели расчет ошибки производим по формуле:

(10)

Это абсолютная ошибка. Численные значения абсолютных ошибок приведены в колонке 7 таблицы 3.

По аналогии с моделью регрессии для оценки качества построения модели или для выбора наилучшей модели можно применять сумму квадратов полученных абсолютных ошибок. Для данной аддитивной модели сумма квадратов абсолютных ошибок равна 1,10. По отношению к общей сумме квадратов отклонений уровней ряда от его среднего уровня, равной 71,59, эта величина составляет чуть более 1,5%. Следовательно, можно сказать, что аддитивная модель объясняет 98,5% общей вариации уровней временного ряда потребления электроэнергии за последние 16 кварталов.

Примеры

Перекрестные данные

Фермерские хозяйства	1	2	3	4	5	6	7	8	9	10
Урожайность (2020 г.)	25,5	31	27,3	21,4	28	27,8	29,1	26,6	26,7	28,5

Временной ряд

Годы	2011	2012	2013	2014	2015	2016	2017	2018	2019	2020
Урожайность (фермерское хозяйство 1)	24,5	22,6	24,3	21,4	28,4	27,1	23,1	26,6	25,2	25,5