

ПРАКТИЧЕСКАЯ РАБОТА № 5

Методы анализа временных рядов

Общие сведения

Цель работы

Приобрести навыки прогнозирования значений временного ряда, в частности, выделения устойчивых тенденции и учета сезонной составляющей, а также навыки использования средств Пакет Анализа и Поиск решения, входящих в MS Excel.

План выполнения

1. Изучить теоретическую часть;
2. Получить задание преподавателя;
3. Выполнить задания 1 и 2:
 - 3.1. Построить график значений временного ряда;
 - 3.2. Построить графики прогнозируемых значений (для каждого из полученных прогнозов);
 - 3.3. Вычислить среднее абсолютных отклонений или среднее относительных ошибок;
 - 3.4. Ответить на вопросы задачи;
4. Составить отчёт по лабораторной работе. Отчёт должен иметь следующую структуру:
 - 4.1. Титульный лист, который должен содержать следующую информацию:
 - 4.1.1. Название университета и кафедры, ответственной за дисциплину;
 - 4.1.2. ФИО и должности студента и преподавателя;
 - 4.2. Отчёт о решении заданиях 1 и 2, содержащий следующее информационное наполнение:
 - 4.2.1. Формулировка индивидуального задания;
 - 4.2.2. Графики значений временного ряда и прогноза;
 - 4.2.3. При необходимости, снимки экрана монитора, содержащие основные моменты решения задачи;
 - 4.2.4. Ответы на вопросы задания;
 - 4.2.5. Результаты решения и выводы.

Теоретическая часть

Общие сведения

Временным рядом называется последовательность значений некоторого показателя во времени (например, объемов продаж, см. Рисунок 54).

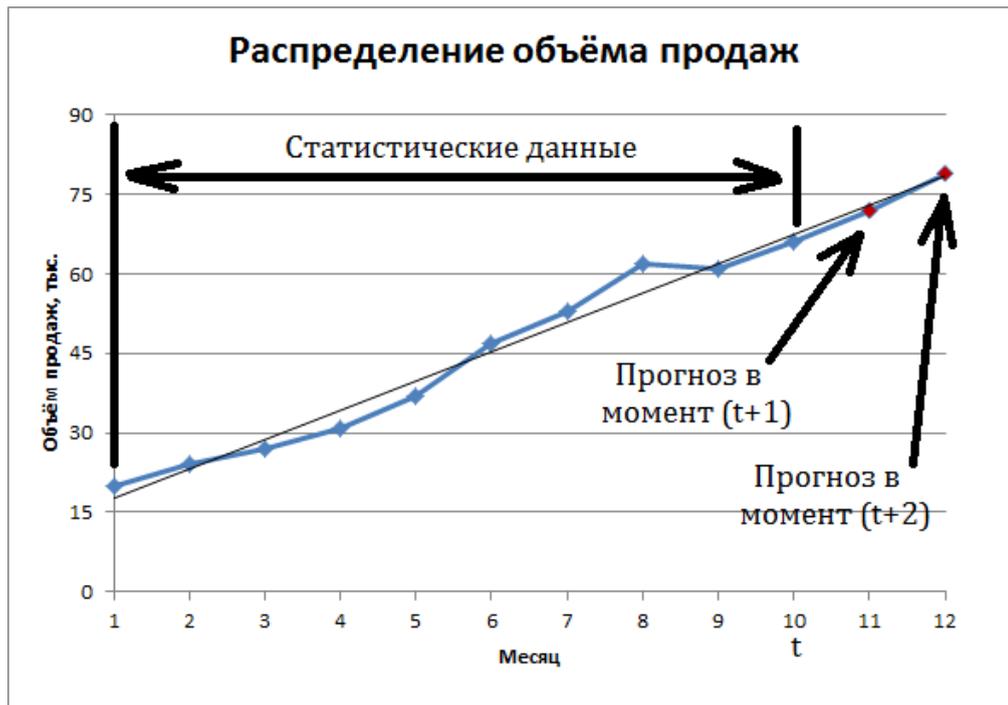


Рисунок 54. Временный ряд объёма продаж

Анализ временных рядов является способом выявления тенденций прошлого и продления их в будущее. Методы анализа временных рядов осуществляют прогноз путем экстраполяции значений отдельной переменной на основе статистических данных за прошлый временной период. *Основное допущение*, которое при этом делается, заключается в том, что происшедшее в прошлом дает хорошее приближение в оценке будущего.

Развитие процессов, реально наблюдаемых в жизни, складывается из некоторой *устойчивой тенденции* и некоторой *случайной составляющей*, выражающейся в колебании значений показателя вокруг тренда. Рисунок 55 показывает, как могут зависеть объемы продаж одного и того же товара на двух стадиях его жизненного цикла (в начале и в конце продаж).

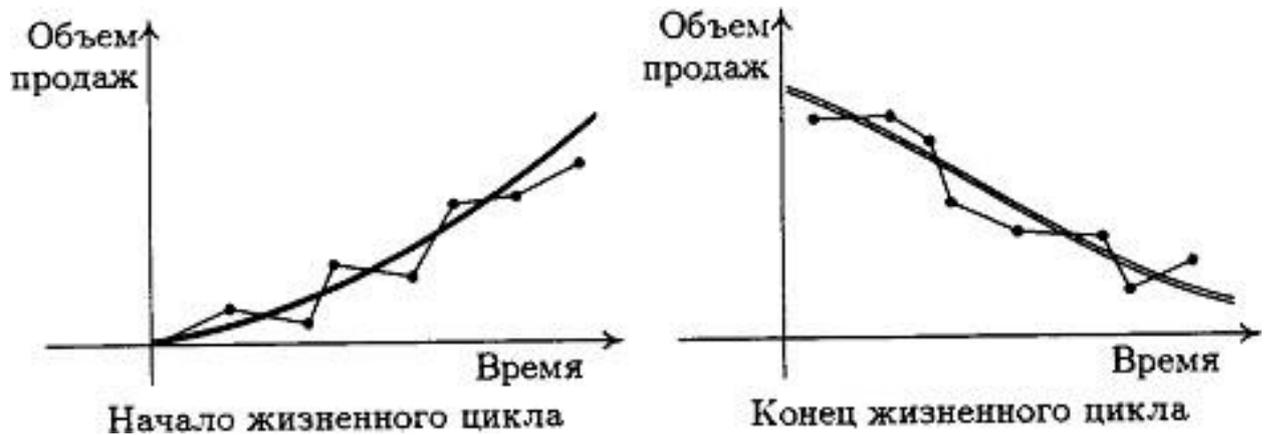


Рисунок 55. Зависимость объёма продаж от стадии жизненного цикла продукта

Кривые тренда сглаживают временной ряд значений показателя, выделяя общую тенденцию. Именно выбор кривой тенденции во многом определяет результаты прогнозирования.

В большинстве случаев временной ряд, кроме тенденции и случайных отклонений от него, характеризуется еще *сезонной составляющей*. Сезонная составляющая — это периодические изменения показателя. Обычная продолжительность сезонной составляющей измеряется днями, неделями или месяцами.

Методы без сезонной составляющей

Вначале рассмотрим несколько простейших методов прогнозирования, не учитывающих наличия сезонности во временном ряде. Предположим, что в журнале РБК приведена сводка за последние 12 дней (включая сегодняшний) цен на апельсины, сложившихся на момент закрытия биржи. Используя эти данные, нужно предсказать завтрашнюю цену на какао (также на момент закрытия биржи). Рассмотрим несколько способов сделать это.

1. Если последнее (сегодняшнее) значение наиболее значимо по сравнению с остальными, то оно является наилучшим прогнозом на завтра.
2. Возможно, из-за быстрого изменения цен на бирже первые шесть значений уже устарели и не актуальны, в то время как последние шесть значимы и имеют равную ценность для прогноза. Тогда в качестве прогноза на завтра можно взять среднее последних шести значений.

3. Если все значения существенны, но сегодняшнее 12-е значение наиболее значимо, а предыдущие 11-е, 10-е, 9-е и т.д. имеют все меньшую и меньшую значимость, следует найти взвешенное среднее всех 12 значений. Причем весовые коэффициенты для последних значений должны быть больше, чем для предыдущих, и сумма всех весовых коэффициентов должна равняться 1.

Первый способ называется «наивным» прогнозом и достаточно очевиден. Рассмотрим подробнее остальные способы.

Метод скользящего среднего

Одним из предположений, лежащих в основе данного метода, является то, что более точный прогноз на будущее можно получить, если использовались недавние наблюдения, причем, чем «новее» данные, тем их вес для прогноза должен быть больше. Удивительно, но такой «наивный» подход оказывается чрезвычайно полезным для практики. Например, многие авиакомпании используют частный тип скользящего среднего для создания прогнозов спроса на авиаперелеты, которые, в свою очередь, используются в сложных механизмах управления и оптимизации доходов. Более того, практически все программные пакеты управления запасами содержат модули, выполняющие прогнозы на основе того или иного типа скользящего среднего.

Рассмотрим следующий пример. Маркетологу нужно спрогнозировать спрос на производимые его компанией станки. Данные по объемам продаж за последний год работы компании находятся в файле «ЛР6.Пример 1.Станки.xls».

Простое скользящее среднее. В этом методе среднее фиксированного числа N последних наблюдений используется для оценки следующего значения временно ряда. Например, используя данные о продажах станков за первые три месяца года, менеджер получает для апреля значение, используя формулу, приведённую ниже:

$$\hat{x}_4 = \frac{x_1 + x_2 + x_3}{3} = \frac{20 + 24 + 27}{3} \approx 23,67 \quad (1).$$

В случае произвольного числа N узлов расчетная формула обобщается следующим образом:

$$\hat{x}_k = \frac{1}{N} \sum_{i=1}^N x_{k-i} \quad (2).$$

Менеджер вычислил объем продаж на основе простого скользящего среднего за 3 и 4 месяца. Однако требуется определить, какое количество узлов даёт более точный прогноз. Для оценки точности прогнозов используются *среднее абсолютных отклонений* (САО) и *среднее относительных ошибок*, в процентах (СООП), вычисляемые по формулам (3) и (4).

$$\text{САО} = \frac{\sum_N |x_i - x'_i|}{N} \quad (3),$$

$$\text{СОО} = \frac{\sum_N \frac{|x_i - x'_i|}{x_i} * 100\%}{N} \quad (4).$$

где x_i – i -ое реальное значение переменной в i -й момент времени, а x'_i – i -ое спрогнозированное значение переменной в i -й момент времени, N — количество прогнозов.

Согласно результатам, полученным на листе «Простое ск. среднее» рабочей книги «ЛР6.Пример 1.Станки.xls» (см. Рисунок 56), скользящее среднее за три месяца имеет значение САО равное 12,67 (ячейка D16), тогда как для скользящего среднего за 4 месяца значение САО равно 15,59 (ячейка F16). Тогда можно выдвинуть гипотезу, что использование большего количества статистических данных скорее ухудшает, чем улучшает точность прогноза методом скользящего среднего.

	A	B	C	D	E	F
1	Месяц	Объемы продаж, тыс.	Прогноз на основе ск. среднего за 3 месяца	Абсолютная ошибка	Прогноз на основе ск. среднего за 4 месяца	Абсолютная ошибка
2	Январь	20	-	-	-	-
3	Февраль	24	-	-	-	-
4	Март	27	-	-	-	-
5	Апрель	31	23,67	7,33	-	-
6	Май	37	27,33	9,67	25,50	11,50
7	Июнь	47	31,67	15,33	29,75	17,25
8	Июль	53	38,33	14,67	35,50	17,50
9	Август	62	45,67	16,33	42,00	20,00
10	Сентябрь	54	54,00	0,00	49,75	4,25
11	Октябрь	36	56,33	20,33	54,00	18,00
12	Ноябрь	32	50,67	18,67	51,25	19,25
13	Декабрь	29	40,67	11,67	46,00	17,00
14						
15			Сумма:	114,00	Сумма:	124,75
16			САО:	12,67	САО:	15,59

Рисунок 56. Пример 1 – результаты прогнозирования методом простого скользящего среднего

На графике (см. Рисунок 57), построенном по результатам наблюдений и прогнозов с интервалом 3 месяца, можно заметить ряд особенностей, общих для всех применений метода скользящего среднего.

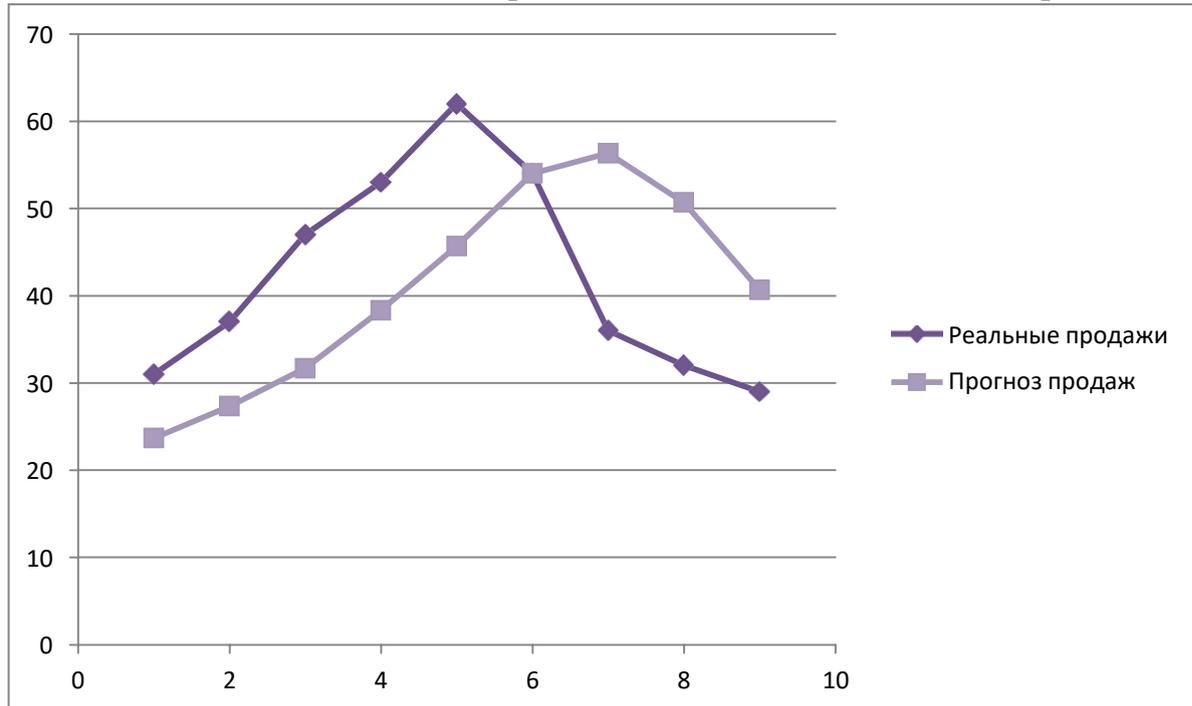


Рисунок 57. Пример 1 – график кривой прогноза методом простого скользящего среднего и график реального объёма продаж

Значение прогноза, полученное методом простого скользящего среднего, всегда меньше фактического значения, если исходные данные монотонно возрастают, и больше фактического значения, если исходные данные монотонно убывают. Поэтому, если данные монотонно возрастают или убывают, то с помощью простого скользящего среднего нельзя получить точных прогнозов. Этот метод лучше всего подходит для данных с небольшими случайными отклонениями от некоторого постоянного или медленно меняющегося значения.

Основной недостаток метода простого скользящего среднего возникает в результате того, что при вычислении прогнозируемого значения самое последнее наблюдение имеет такой же вес (т. е. значимость), как и предыдущие. Это происходит потому, что вес всех N последних наблюдений, участвующих в вычислении скользящего среднего, равен $1/N$. Присвоение равного веса противоречит интуитивному представлению о том, что во многих случаях последние данные могут больше сказать о том, что произойдет в ближайшем будущем, чем предыдущие.

Взвешенное скользящее среднее. Вклад различных моментов времени можно учесть, вводя вес для каждого значения показателя в скользящем интервале. В результате получается метод взвешенного скользящего среднего, который математически можно записать так:

$$\hat{x}_k = \frac{\sum_{i=1}^N (w_{k-i} * x_{k-i})}{\sum_{i=1}^N w_{k-i}} \quad (4).$$

где w_{k-i} — вес, с которым используется показатель x_{k-i} при расчете.

Вес — это всегда положительное число. В случае, когда все веса одинаковы, вырождается метод простого скользящего среднего.

Теперь маркетолог может использовать метод взвешенного скользящего среднего за 3 месяца. Но прежде требуется понять, как выбрать веса. Используя средство Поиск решения, можно определить оптимальный вес узлов. Чтобы определить вес узлов с помощью средства Поиск решения, при котором значение среднего абсолютных отклонений было бы минимально, выполните следующие действия:

1. Выберите команду Сервис -> Поиск решения.
2. В диалоговом окне Поиска решения установите ячейку G16 целевой (см. лист «Веса»), минимизируя её.
3. Изменяемыми ячейками укажите диапазон B1:B3.
4. Установите ограничения $B4 = 1,0$; $B1:B3 \geq 0$; $B1:B3 \leq 1$; $B1 \leq B2$ и $B2 \leq B3$.
5. Запустите поиск решения (результат отображает).

	A	B	C	D	E	F	G
1	альфа 2 =	0,000		Месяц	Объемы продаж, тыс.	Прогноз на основе ск. среднего за 3 месяца	Абсолютная ошибка
2	альфа 1 =	0,000		Январь	20	-	-
3	альфа 0 =	1,000		Февраль	24	-	-
4	Сумма весов =	1,00		Март	27	-	-
5				Апрель	31	27,00	4,00
6				Май	37	31,00	6,00
7				Июнь	47	37,00	10,00
8				Июль	53	47,00	6,00
9				Август	62	53,00	9,00
10				Сентябрь	54	62,00	8,00
11				Октябрь	36	54,00	18,00
12				Ноябрь	32	36,00	4,00
13				Декабрь	29	32,00	3,00
14							
15						Сумма =	68,00
16						CAO =	7,56

Рисунок 58. Пример 1 – результат поиска весов значений показателей при использовании метода взвешенного скользящего среднего

Полученные результаты показывают, что оптимальное распределение весов таково, что весь вес сосредоточен на самом последнем наблюдении, при этом значение среднего абсолютных отклонений равно 7,56 (см. также Рисунок 59). Этот результат подтверждает предположение о том, что более поздние наблюдения должны иметь больший вес.

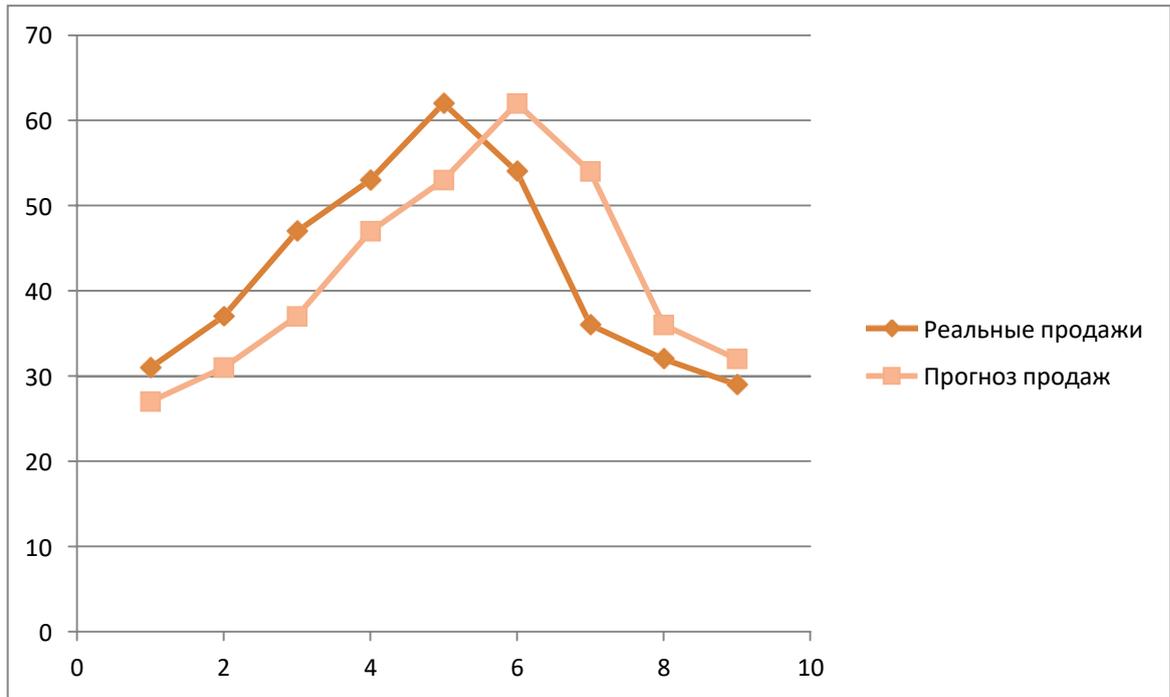


Рисунок 59. Пример 1 – график кривой прогноза методом взвешенного скользящего среднего и график реального объема продаж

Метод экспоненциального сглаживания

Прогнозы в методах скользящего среднего зависят от предыдущих значений показателя временного ряда, но не от качества предыдущих прогнозов. Рассмотрим один из методов прогнозирования, который учитывает отклонение предыдущего прогноза от реального значения показателя ряда.

Очевидно, что в методе взвешенного скользящего среднего существует множество способов задавать значения весов так, чтобы их сумма была равной 1, и экспоненциальное сглаживание – один из таких способов. В этой схеме метода взвешенного среднего для любого $t > 1$ прогнозируемое значение \hat{x}_{t+1} в момент времени $(t+1)$ представляет собой взвешенную сумму фактического объема продаж x_t , за период времени t и прогнозируемого объема продаж \hat{x}_t , за период времени t (см. формулу (5)).

$$\hat{x}_{t+1} = \alpha x_t + (1-\alpha)\hat{x}_t \quad (5).$$

Экспоненциальное сглаживание имеет вычислительные преимущества перед скользящим средним. Здесь, чтобы вычислить \hat{x}_{t+1} , необходимо знать только значения x_t и \hat{x}_t , (а также значение α), вместо значений показателя ряда во всех узлах, по которым происходит сглаживание. Сохраняя значение α и последний прогноз, мы также неявно сохраняем и все предыдущие прогнозы.

Рассмотрим некоторые свойства модели экспоненциального сглаживания. Для начала заметим, что если $t > 2$, то в формуле (5) t можно заменить на $t-1$, т.е.:

$$\hat{x}_t = \alpha x_{t-1} + (1-\alpha)\hat{x}_{t-1} \quad (6).$$

Подставив формулу (6) в первоначальную формулу (5), получим:

$$\hat{x}_{t+1} = \alpha x_t + \alpha (1-\alpha)x_{t-1} + (1-\alpha)^2\hat{x}_{t-1} \quad (7).$$

Выполняя последовательно аналогичные подстановки, получим следующее выражение для \hat{x}_{t+1} :

$$\hat{x}_{t-1} = \alpha x_t + \alpha (1-\alpha)x_{t-1} + \alpha (1-\alpha)^2x_{t-2} + \dots + \alpha (1-\alpha)^{t-1}x_1 + (1-\alpha)^t\hat{x}_1 \quad (8).$$

Поскольку из неравенства $0 < \alpha < 1$ следует, что $0 < 1 - \alpha < 1$, то $\alpha > \alpha(1-\alpha) > \alpha(1-\alpha)^2 \dots$. Другими словами, наблюдение x_t , имеет больший вес, чем наблюдение x_{t-1} , которое, в свою очередь, имеет больший вес, чем x_{t-2} . Это иллюстрирует основное свойство модели экспоненциального сглаживания — коэффициенты при x_k убывают при уменьшении номера k . Также можно показать, что сумма всех коэффициентов (включая коэффициент при x_1), равна 1.

Из приведенной формулы видно также, что значение \hat{x}_{t+1} является взвешенной суммой всех предыдущих наблюдений (включая последнее наблюдение x_t). Последнее слагаемое этой суммы является не статистическим наблюдением, а «предположением» \hat{x}_1 (можно предположить, например, что $\hat{x}_1 = x_1$). Очевидно, что с ростом t влияние \hat{x}_1 , на прогноз уменьшается, и в определенный момент им можно будет пренебречь. Даже если значение α достаточно малое (такое, что $(1 - \alpha)$ приблизительно равно 1), значение $(1-\alpha)^t$ будет быстро убывать.

Значение параметра α сильно влияет на функционирование модели прогнозирования, поскольку α представляет собой вес самого последнего наблюдения x_t . Это значит, что следует назначать большее значение α в том случае, когда в модели наиболее прогностическим является именно последнее наблюдение. Если же α близко к 0, это означает практически полное доверие к прошлому прогнозу и игнорирование последнего наблюдения.

Перед маркетологом возникает проблема: как наилучшим образом подобрать значение α ? В этом поможет средство Поиск решения. Чтобы найти оптимальное значение α (т.е. такое, при котором прогнозная кривая будет менее всего отклоняться от кривой значений временного ряда), выполните следующие действия.

1. Выберите команду Сервис -> Поиск решения.
2. В диалоговом окне Поиск решения установите целевую ячейку G16 (см. лист «Экспо»), минимизируя значение.
3. Укажите изменяемой ячейкой ячейку B1.
4. Установите ограничения $B1 > 0$ и $B1 < 1$
5. Запустите расчёт (см. результат, Рисунок 60).

	A	B	C	D	E	F	G
1	альфа =	1,000		Месяц	Объемы продаж, тыс.	Прогноз	Абсолютная ошибка
2				Январь	20	20,00	-
3				Февраль	24	20,00	4,00
4				Март	27	24,00	3,00
5				Апрель	31	27,00	4,00
6				Май	37	31,00	6,00
7				Июнь	47	37,00	10,00
8				Июль	53	47,00	6,00
9				Август	62	53,00	9,00
10				Сентябрь	54	62,00	8,00
11				Октябрь	36	54,00	18,00
12				Ноябрь	32	36,00	4,00
13				Декабрь	29	32,00	3,00
14							
15						Сумма =	75,00
16						САО =	6,82

Рисунок 60. Пример 1 – результат использования метода экспоненциального сглаживания

Опять, как и в методе взвешенного скользящего среднего, наилучший прогноз будет получен, если назначить весь вес последнему наблюдению (см. Рисунок 61). Следовательно, оптимальное значение α равно 1, при этом среднее абсолютных отклонений равно 6,82 (ячейка G16).

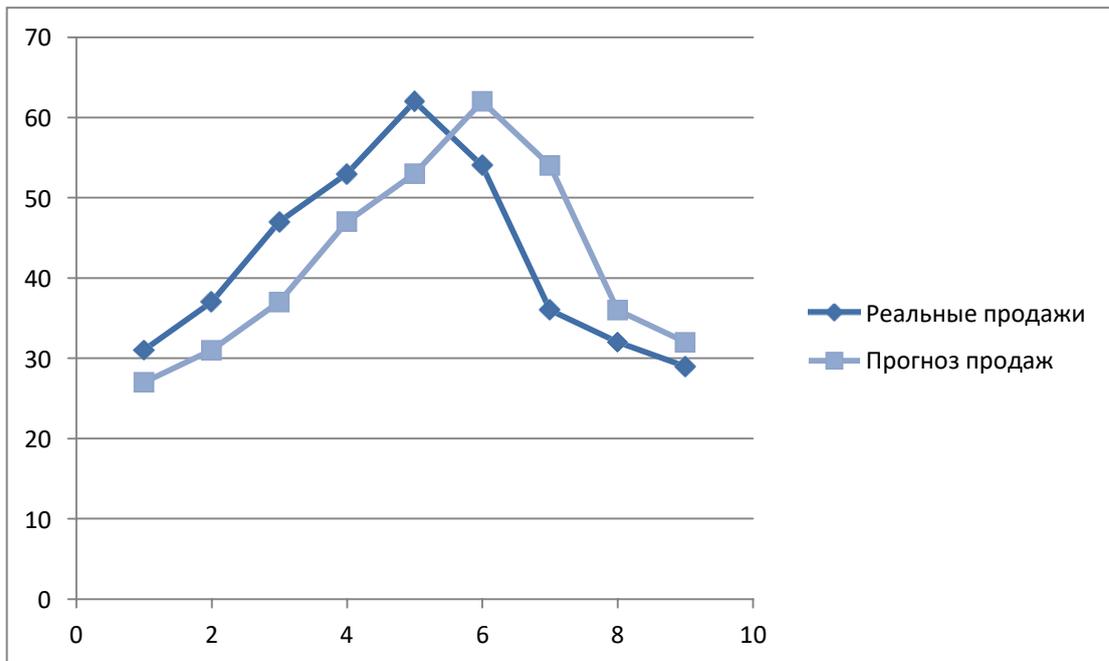


Рисунок 61. Пример 1 – график кривой прогноза методом взвешенного скользящего среднего и график реального объема продаж

Метод экспоненциального сглаживания хорошо работает в ситуациях, когда переменная имеет стационарный характер, а ее отклонения от постоянного значения вызваны случайными факторами и не носят регулярного характера. Но этим методом, как и методами скользящего среднего не удастся спрогнозировать монотонно возрастающие или монотонно убывающие данные. Прогнозируемые значения будут всегда меньше или больше наблюдаемых, соответственно, а точность данных будет сравнима с точностью «наивного прогноза». Эти методы также не учитывают сезонных изменений показателя ряда.

Если статистические данные монотонно изменяются или подвержены сезонным изменениям, необходимы специальные методы прогнозирования, которые будут рассмотрены ниже.

Подбор кривой тренда

В качестве примера, воспользуемся данными объемов продаж из файла «ЛР6.Пример 2.Продажи.xls». Вначале построим точечную диаграмму, отображающие реальные объемы продаж. Чтобы построить по этим данным линию тренда, отражающую тенденцию в изменении объемов продаж, надо выполнить такие действия.

1. Выберите любую точку выбранного ряда данных. В результате будут выделены все точки ряда.
2. Вызовите через контекстное меню функцию «Добавить линию тренда».
3. В диалоговом окне Линия тренда по умолчанию будет выбран линейный тип функции.
4. Нажмите «Ок».

После этого на графике появится прямая линия тренда (см. Рисунок 62).

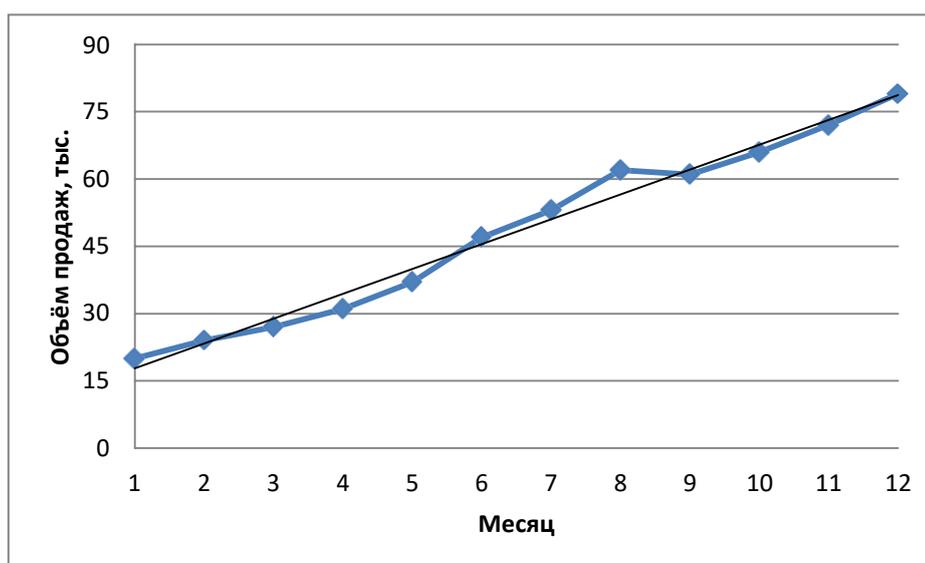


Рисунок 62. Пример 2 – линия тренда на графике Excel

Для того чтобы осуществить прогноз, нужно в диалоговом окне Линия тренда отметить интересующий интервал времени в пункте «вперед на» (или «назад на»).

В меню Линия тренда можно также задать параметры подбираемой кривой. Например, он может быть экспоненциальной или полиномом заданной степени.

В рассматриваемом случае кривой тренда является прямая линия с уравнением $y = ax + b$. Коэффициенты a и b для этой кривой можно также найти с помощью надстройки Пакет Анализа, выбрав средство Регрессия.

Метод Хольта

Метод Хольта представляет собой развитие метода экспоненциального сглаживания, с учетом наличия тренда. Формулировка метода имеет вид

$$\hat{y}_{+k} = L_t + kT_t \quad (8),$$

где

$$L_t = a x_t + (1-a)(L_{t-1} + T_{t-1}) \quad (9),$$

$$T_t = \beta(L_t - L_{t-1}) + (1 - \beta)T_{t-1} \quad (10).$$

Метод Хольта позволяет прогнозировать на k периодов времени вперед. Метод, как видно, использует два параметра a и β , значения которых находятся в пределах от 0 до 1. Переменная L , указывает на долгосрочный уровень значений или базовое значение данных временного ряда. Переменная T указывает на возможное возрастание или убывание значений за один период, т.е. на присутствие тренда.

Рассмотрим работу этого метода на следующем примере. Светлана работает аналитиком в большой брокерской фирме. На основе имеющихся у нее квартальных отчетов компании RusWindAirlines она хочет спрогнозировать доход этой компании в следующем квартале. Имеющиеся данные и диаграмма, построенная на их основе, находятся в файле «ЛР6.Пример 3.RusWindAirlines.xls» (см. Рисунок 63).

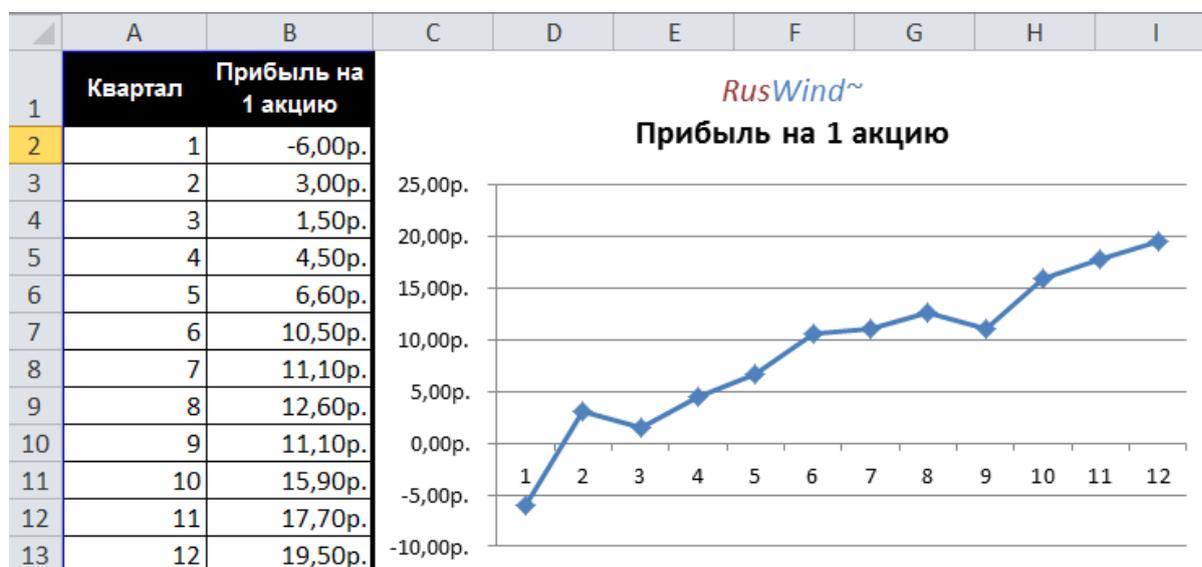


Рисунок 63. Пример 3 - данные о доходе авиакомпании RusWind

Видно, что данные имеют явный тренд (почти монотонно возрастают). Светлана хочет применить метод Хольта, чтобы спрогнозировать значение прибыли на одну акцию на тринадцатый квартал. Для этого необходимо задать начальные значения для L и T . Есть несколько вариантов выбора:

1. L равно значению прибыли на одну акцию за первый квартал и $T = 0$;

2. L равно среднему значению прибыли на одну акцию за 12 кварталов и T равно среднему изменению за все 12 кварталов.

Существуют и другие варианты начальных значений для L и T , но Светлана выбрала первый вариант.

Она решила воспользоваться средством Поиск решения, чтобы найти оптимальное значение параметров α и β , при которых значение среднего абсолютных ошибок в процентах было бы минимально. Для этого нужно выполнить следующие действия.

1. Выбрать команду Сервис -> Поиск решения.
2. В диалоговом окне Поиск решения задать ячейку F18 целевой, минимизируя её.
3. Выбрать диапазон ячеек B1:B2 изменяемыми.
4. Добавить ограничения B1:B2 > 0 и B1:B2 < 1.
5. Выполнить поиск.

Полученный прогноз показан отображают Рисунок 64, Рисунок 65, Рисунок 66. Как видно, оптимальными оказались значения $\alpha = 0,59$ и $\beta = 0,42$, при этом среднее абсолютных ошибок в процентах равно около 38%.

	A	B	C	D	E	F
1	$\alpha =$	0,59				
2	$\beta =$	0,42				
3						
4	Квартал	П/А	L_i	T_i	Прогноз П/А	Абс. ошибка, %
5	1	-6,00р.	-6,00р.	0	-	-
6	2	3,00р.	-0,69р.	2,230	-6,00р.	300,0%
7	3	1,50р.	1,52р.	2,220	1,54р.	2,7%
8	4	4,50р.	4,19р.	2,409	3,74р.	17,0%
9	5	6,60р.	6,60р.	2,410	6,60р.	0,1%
10	6	10,50р.	9,89р.	2,780	9,01р.	14,2%
11	7	11,10р.	11,74р.	2,391	12,67р.	14,1%
12	8	12,60р.	13,23р.	2,011	14,13р.	12,2%
13	9	11,10р.	12,80р.	0,985	15,24р.	37,3%
14	10	15,90р.	15,03р.	1,510	13,78р.	13,3%
15	11	17,70р.	17,23р.	1,797	16,54р.	6,5%
16	12	19,50р.	19,30р.	1,915	19,02р.	2,5%
17					21,22р.	
18					СООП =	38,2%

Рисунок 64. Пример 3 - результат прогноза с использованием метода Хольта

	A	B	C	D	E	F
1	$\alpha =$	0,59				
2	$\beta =$	0,42				
3						
4	Квартал	П/А	L_i	T_i	Прогноз П/А	Абс. ошибка, %
5	1	=Данные!B2	=B5	0		-
6	2	=Данные!B3	= $\$B\$1*B6+(1-\$B\$1)*(C5+D5)$	= $\$B\$2*(C6-C5)+(1-\$B\$2)*D5$	=СУММ(C5:D5)	=ABS(B6-E6)/B6
7	3	=Данные!B4	= $\$B\$1*B7+(1-\$B\$1)*(C6+D6)$	= $\$B\$2*(C7-C6)+(1-\$B\$2)*D6$	=СУММ(C6:D6)	=ABS(B7-E7)/B7
8	4	=Данные!B5	= $\$B\$1*B8+(1-\$B\$1)*(C7+D7)$	= $\$B\$2*(C8-C7)+(1-\$B\$2)*D7$	=СУММ(C7:D7)	=ABS(B8-E8)/B8
9	5	=Данные!B6	= $\$B\$1*B9+(1-\$B\$1)*(C8+D8)$	= $\$B\$2*(C9-C8)+(1-\$B\$2)*D8$	=СУММ(C8:D8)	=ABS(B9-E9)/B9
10	6	=Данные!B7	= $\$B\$1*B10+(1-\$B\$1)*(C9+D9)$	= $\$B\$2*(C10-C9)+(1-\$B\$2)*D9$	=СУММ(C9:D9)	=ABS(B10-E10)/B10
11	7	=Данные!B8	= $\$B\$1*B11+(1-\$B\$1)*(C10+D10)$	= $\$B\$2*(C11-C10)+(1-\$B\$2)*D10$	=СУММ(C10:D10)	=ABS(B11-E11)/B11
12	8	=Данные!B9	= $\$B\$1*B12+(1-\$B\$1)*(C11+D11)$	= $\$B\$2*(C12-C11)+(1-\$B\$2)*D11$	=СУММ(C11:D11)	=ABS(B12-E12)/B12
13	9	=Данные!B10	= $\$B\$1*B13+(1-\$B\$1)*(C12+D12)$	= $\$B\$2*(C13-C12)+(1-\$B\$2)*D12$	=СУММ(C12:D12)	=ABS(B13-E13)/B13
14	10	=Данные!B11	= $\$B\$1*B14+(1-\$B\$1)*(C13+D13)$	= $\$B\$2*(C14-C13)+(1-\$B\$2)*D13$	=СУММ(C13:D13)	=ABS(B14-E14)/B14
15	11	=Данные!B12	= $\$B\$1*B15+(1-\$B\$1)*(C14+D14)$	= $\$B\$2*(C15-C14)+(1-\$B\$2)*D14$	=СУММ(C14:D14)	=ABS(B15-E15)/B15
16	12	=Данные!B13	= $\$B\$1*B16+(1-\$B\$1)*(C15+D15)$	= $\$B\$2*(C16-C15)+(1-\$B\$2)*D15$	=СУММ(C15:D15)	=ABS(B16-E16)/B16
17					=СУММ(C16:D16)	
18					СООП =	=СРЗНАЧ(F6:F16)

Рисунок 65. Пример 3 - результат прогноза с использованием метода Хольта.
Формулы

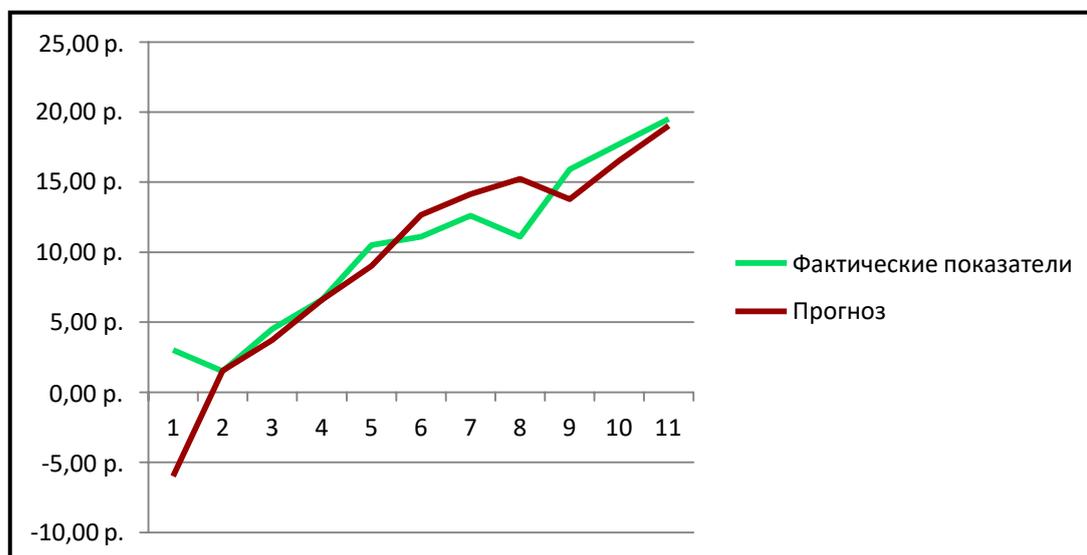


Рисунок 66. Пример 3 - результат прогноза с использованием метода Хольта.
График

Однако метод Хольта, как и рассмотренные ранее простейшие методы прогнозирования, не учитывает наличие во временном ряде сезонных изменений.

Учет сезонных изменений

Спрос на значительное число товаров меняется в течение года. Например, если посмотреть на объемы продаж мороженого по месяцам, то можно увидеть в теплые месяцы (с июня по август в северном полушарии) более высокий уровень продаж, чем зимой, и так каждый год. Здесь сезонные колебания имеют период в 12 месяцев. Другой пример:

анализируются еженедельные отчеты о количестве постояльцев, которые оставались на ночь в отеле, расположенном в бизнес-центре города. Предположительно можно сказать, что большое число клиентов ожидается в ночи на вторник, среду и четверг, меньше всего клиентов будет в ночи на субботу и воскресенье, и среднее число постояльцев ожидается в ночи на пятницу и понедельник. Такая структура данных, отображающая количество клиентов в разные дни недели, будет повторяться через каждые семь дней.

Подобные циклические изменения показателя временного ряда носят название сезонных колебаний (хотя сезон, как мы видели, может продлиться и неделю и год). Процедура, которая позволяет сделать прогноз с учетом сезонных изменений, состоит из следующих этапов.

1. На основе исходных данных определяется структура сезонных колебаний и период этих колебаний.
2. Используя численный метод, описанный далее, из исходных данных исключают сезонную составляющую.
3. На основе данных, из которых исключена сезонная составляющая, делается наилучший возможный прогноз.
4. К полученному прогнозу добавляется сезонная составляющая.

Проиллюстрируем этот подход на данных об объемах сбыта угля (измеряемого в тысячах тонн) в США на протяжении девяти лет. Пусть некто Фрэнк работает менеджером в компании Gillette Coal Mine, и ему необходимо спрогнозировать спрос на уголь на ближайшие два квартала. Он ввел данные по всей угольной отрасли в рабочую книгу «ЛР6.Пример 4.Уголь.xls» и построил по этим данным график (см. Рисунок 67).

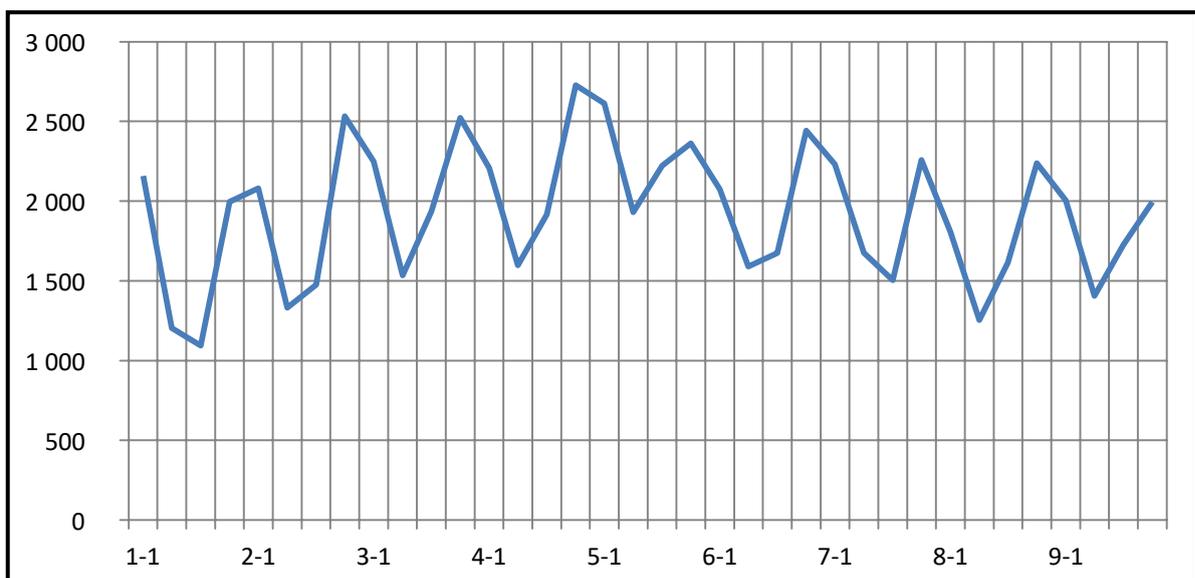


Рисунок 67. Пример 4 - данные о спросе на уголь в компании GilletteCoalMine

Определение структуры и периода сезонных колебаний. Из графика, приведённого выше видно, что объёмы продаж выше среднего уровня в первом и четвертом кварталах (зимнее время года) и ниже среднего во втором и третьем кварталах (весенне-летние месяцы).

Исключение сезонной составляющей. Прежде всего необходимо вычислить среднее значение всех отклонений за один период сезонных изменений. Чтобы исключить сезонную составляющую в пределах одного года, используются данные за четыре периода (квартала). А чтобы исключить сезонную составляющую из всего временного ряда, вычисляется последовательность скользящих средних по T узлам, где T — продолжительность сезонных колебаний. Для выполнения необходимых вычислений Фрэнк использовал столбцы C и D , как показано на рисунке ниже. Столбец C содержит значения скользящего среднего по 4 узлам на основе данных, которые находятся в столбце B .

Далее требуется назначить полученные значения скользящего среднего средним точкам последовательности данных, на основе которых эти значения были вычислены. Эта операция называется центрированием значений. Если T нечетное, то первое значение скользящего среднего (среднее значений от первой до T -й точки) надо присвоить $(T + 1)/2$ точке (например, если $T = 7$, то первое скользящее среднее будет назначено четвертой точке). Аналогично среднее значений от второй до $(T + 1)$ -й точки центрируется в $(T + 3)/2$ точке и т. д. Центр n -го интервала находится в точке $(T + (2n - 1))/2$.

Если T четное, как в рассматриваемом случае, то задача несколько усложняется, поскольку центральные (средние) точки расположены между точками, по которым вычислялось значение скользящего среднего. Поэтому центрированное значение для третьей точки вычисляется как среднее первого и второго значений скользящего среднего. Например, первое число в столбце D отцентрированных средних (см. Рисунок 68), слева равняется $(1613 + 1594) / 2 = 1603$.

7	Период год-квартал	Исходные данные	Скользящие средние	Центрированные средние
8	1-1	2 159	-	-
9	1-2	1 203	-	-
10	1-3	1 094	1 613	1 603
11	1-4	1 996	1 594	1 610
12	2-1	2 081	1 626	1 674
13	2-2	1 332	1 721	1 788
14	2-3	1 476	1 856	1 877
15	2-4	2 533	1 898	1 923
16	3-1	2 249	1 948	2 005
17	3-2	1 533	2 063	2 061
18	3-3	1 935	2 060	2 055
19	3-4	2 523	2 090	2 058
20	4-1	2 208	2 066	2 064
21	4-2	1 597	2 061	2 087
22	4-3	1 917	2 112	2 163
23	4-4	2 726	2 213	2 255
24	5-1	2 612	2 297	2 335
25	5-2	1 931	2 373	2 328
26	5-3	2 223	2 282	2 215
27	5-4	2 363	2 148	2 105
28	6-1	2 074	2 062	1 994
29	6-2	1 589	1 925	1 935
30	6-3	1 673	1 945	1 964
31	6-4	2 443	1 984	1 995
32	7-1	2 231	2 006	1 984
33	7-2	1 675	1 963	1 940
34	7-3	1 503	1 917	1 864
35	7-4	2 259	1 812	1 759
36	8-1	1 809	1 706	1 720
37	8-2	1 254	1 734	1 731
38	8-3	1 613	1 729	1 753
39	8-4	2 238	1 777	1 796
40	9-1	2 004	1 815	1 829
41	9-2	1 406	1 843	1 813
42	9-3	1 725	1 782	-
43	9-4	1 994	-	-
44	10-1			
45	10-2			

7	Период год-квартал	Исходные данные	Скользящие средние	Центрированные средние
8	1-1	2159	-	-
9	1-2	1203	-	-
10	1-3	1094	=СРЗНАЧ(В8:В11)	=СРЗНАЧ(С10:С11)
11	1-4	1996	=СРЗНАЧ(В9:В12)	=СРЗНАЧ(С11:С12)
12	2-1	2081	=СРЗНАЧ(В10:В13)	=СРЗНАЧ(С12:С13)
13	2-2	1332	=СРЗНАЧ(В11:В14)	=СРЗНАЧ(С13:С14)
14	2-3	1476	=СРЗНАЧ(В12:В15)	=СРЗНАЧ(С14:С15)
15	2-4	2533	=СРЗНАЧ(В13:В16)	=СРЗНАЧ(С15:С16)
16	3-1	2249	=СРЗНАЧ(В14:В17)	=СРЗНАЧ(С16:С17)
17	3-2	1533	=СРЗНАЧ(В15:В18)	=СРЗНАЧ(С17:С18)
18	3-3	1935	=СРЗНАЧ(В16:В19)	=СРЗНАЧ(С18:С19)
19	3-4	2523	=СРЗНАЧ(В17:В20)	=СРЗНАЧ(С19:С20)
20	4-1	2208	=СРЗНАЧ(В18:В21)	=СРЗНАЧ(С20:С21)
21	4-2	1597	=СРЗНАЧ(В19:В22)	=СРЗНАЧ(С21:С22)
22	4-3	1917	=СРЗНАЧ(В20:В23)	=СРЗНАЧ(С22:С23)
23	4-4	2726	=СРЗНАЧ(В21:В24)	=СРЗНАЧ(С23:С24)
24	5-1	2612	=СРЗНАЧ(В22:В25)	=СРЗНАЧ(С24:С25)
25	5-2	1931	=СРЗНАЧ(В23:В26)	=СРЗНАЧ(С25:С26)
26	5-3	2223	=СРЗНАЧ(В24:В27)	=СРЗНАЧ(С26:С27)
27	5-4	2363	=СРЗНАЧ(В25:В28)	=СРЗНАЧ(С27:С28)
28	6-1	2074	=СРЗНАЧ(В26:В29)	=СРЗНАЧ(С28:С29)
29	6-2	1589	=СРЗНАЧ(В27:В30)	=СРЗНАЧ(С29:С30)
30	6-3	1673	=СРЗНАЧ(В28:В31)	=СРЗНАЧ(С30:С31)
31	6-4	2443	=СРЗНАЧ(В29:В32)	=СРЗНАЧ(С31:С32)
32	7-1	2231	=СРЗНАЧ(В30:В33)	=СРЗНАЧ(С32:С33)
33	7-2	1675	=СРЗНАЧ(В31:В34)	=СРЗНАЧ(С33:С34)
34	7-3	1503	=СРЗНАЧ(В32:В35)	=СРЗНАЧ(С34:С35)
35	7-4	2259	=СРЗНАЧ(В33:В36)	=СРЗНАЧ(С35:С36)
36	8-1	1809	=СРЗНАЧ(В34:В37)	=СРЗНАЧ(С36:С37)
37	8-2	1254	=СРЗНАЧ(В35:В38)	=СРЗНАЧ(С37:С38)
38	8-3	1613	=СРЗНАЧ(В36:В39)	=СРЗНАЧ(С38:С39)
39	8-4	2238	=СРЗНАЧ(В37:В40)	=СРЗНАЧ(С39:С40)
40	9-1	2004	=СРЗНАЧ(В38:В41)	=СРЗНАЧ(С40:С41)
41	9-2	1406	=СРЗНАЧ(В39:В42)	=СРЗНАЧ(С41:С42)
42	9-3	1725	=СРЗНАЧ(В40:В43)	-
43	9-4	1994	-	-
44	10-1			
45	10-2			

Рисунок 68. Пример 4 – нахождение центрированного среднего по исходным данным

Графики исходных данных и отцентрированных средних показаны ниже (см. Рисунок 69).

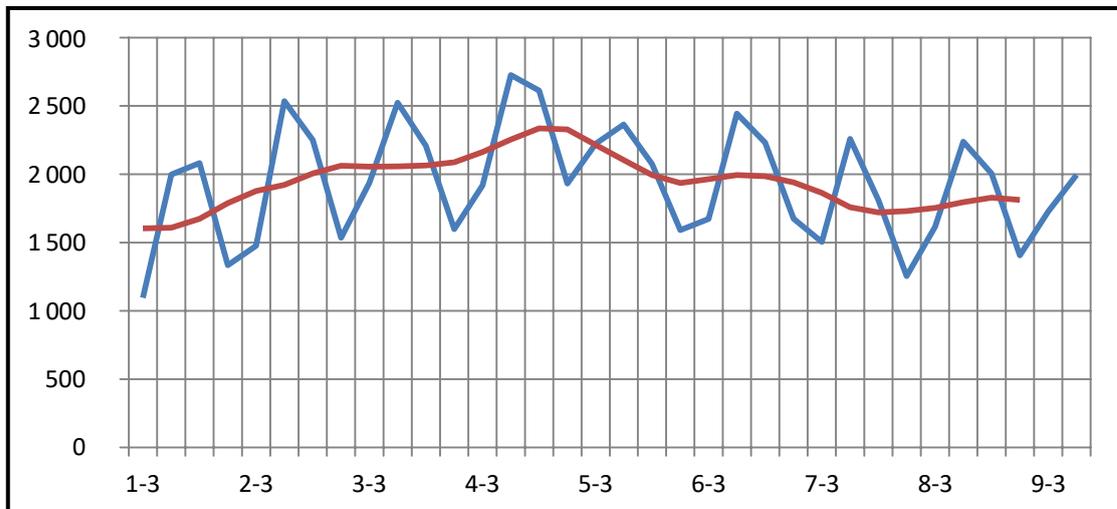


Рисунок 69. Пример 4 - график центрированного скользящего среднего

Далее найдём отношения значений точек данных к соответствующим значениям отцентрированных средних (см. Рисунок 70). Поскольку точкам в начале и конце последовательности данных нет соответствующих отцентрированных средних (см. первые и последние значения в столбце D), такое действие на эти точки не распространяется. Эти отношения показывают степень отклонения значений данных относительно типового уровня, определяемого отцентрированными средними. Заметим, что значения отношений для третьих кварталов меньше 1, а для четвертых — больше 1.

	A	E
7	Период год-квартал	Отношения данных к центрированным средним
8	1-1	-
9	1-2	-
10	1-3	0,682
11	1-4	1,240
12	2-1	1,244
13	2-2	0,745
14	2-3	0,787
15	2-4	1,317
16	3-1	1,122
17	3-2	0,744
18	3-3	0,942
19	3-4	1,226
20	4-1	1,070
21	4-2	0,765
22	4-3	0,886
23	4-4	1,209
24	5-1	1,119
25	5-2	0,830
26	5-3	1,004
27	5-4	1,123
28	6-1	1,040
29	6-2	0,821
30	6-3	0,852
31	6-4	1,225
32	7-1	1,124
33	7-2	0,863
34	7-3	0,806
35	7-4	1,284
36	8-1	1,052
37	8-2	0,724
38	8-3	0,920
39	8-4	1,246
40	9-1	1,096
41	9-2	0,776
42	9-3	-
43	9-4	-
44	10-1	
45	10-2	

	A	B	D	E
7	Период год-квартал	Исходные данные	Центрированные средние	Отношения данных к центрированным
8	1-1	2159	-	-
9	1-2	1203	-	-
10	1-3	1094	=СРЗНАЧ(C10:C11)	=B10/D10
11	1-4	1996	=СРЗНАЧ(C11:C12)	=B11/D11
12	2-1	2081	=СРЗНАЧ(C12:C13)	=B12/D12
13	2-2	1332	=СРЗНАЧ(C13:C14)	=B13/D13
14	2-3	1476	=СРЗНАЧ(C14:C15)	=B14/D14
15	2-4	2533	=СРЗНАЧ(C15:C16)	=B15/D15
16	3-1	2249	=СРЗНАЧ(C16:C17)	=B16/D16
17	3-2	1533	=СРЗНАЧ(C17:C18)	=B17/D17
18	3-3	1935	=СРЗНАЧ(C18:C19)	=B18/D18
19	3-4	2523	=СРЗНАЧ(C19:C20)	=B19/D19
20	4-1	2208	=СРЗНАЧ(C20:C21)	=B20/D20
21	4-2	1597	=СРЗНАЧ(C21:C22)	=B21/D21
22	4-3	1917	=СРЗНАЧ(C22:C23)	=B22/D22
23	4-4	2726	=СРЗНАЧ(C23:C24)	=B23/D23
24	5-1	2612	=СРЗНАЧ(C24:C25)	=B24/D24
25	5-2	1931	=СРЗНАЧ(C25:C26)	=B25/D25
26	5-3	2223	=СРЗНАЧ(C26:C27)	=B26/D26
27	5-4	2363	=СРЗНАЧ(C27:C28)	=B27/D27
28	6-1	2074	=СРЗНАЧ(C28:C29)	=B28/D28
29	6-2	1589	=СРЗНАЧ(C29:C30)	=B29/D29
30	6-3	1673	=СРЗНАЧ(C30:C31)	=B30/D30
31	6-4	2443	=СРЗНАЧ(C31:C32)	=B31/D31
32	7-1	2231	=СРЗНАЧ(C32:C33)	=B32/D32
33	7-2	1675	=СРЗНАЧ(C33:C34)	=B33/D33
34	7-3	1503	=СРЗНАЧ(C34:C35)	=B34/D34
35	7-4	2259	=СРЗНАЧ(C35:C36)	=B35/D35
36	8-1	1809	=СРЗНАЧ(C36:C37)	=B36/D36
37	8-2	1254	=СРЗНАЧ(C37:C38)	=B37/D37
38	8-3	1613	=СРЗНАЧ(C38:C39)	=B38/D38
39	8-4	2238	=СРЗНАЧ(C39:C40)	=B39/D39
40	9-1	2004	=СРЗНАЧ(C40:C41)	=B40/D40
41	9-2	1406	=СРЗНАЧ(C41:C42)	=B41/D41
42	9-3	1725	-	-
43	9-4	1994	-	-
44	10-1			2260,25
45	10-2			2170,75

Рисунок 70. Пример 4 – отношение данных к центрированным средним скользящим

Эти отношения являются основой для создания сезонных индексов. Для их вычисления группируются вычисленные отношения по кварталам, как показано в столбцах С—К (см. Рисунок 71).

	A	B	C	D	E	F	G	H	I	J	K
1	1,108	1-й кв.	-	1,244	1,122	1,070	1,119	1,040	1,124	1,052	1,096
2	0,784	2-й кв.	-	0,745	0,744	0,765	0,830	0,821	0,863	0,724	0,776
3	0,860	3-й кв.	0,682	0,787	0,942	0,886	1,004	0,852	0,806	0,920	-
4	1,234	4-й кв.	1,240	1,317	1,226	1,209	1,123	1,225	1,284	1,246	-

Рисунок 71. Пример 4 - отношения исходных данных к центрированным скользящим средним, сгруппированные по сезонам

Затем требуется найти средние значения отношений по каждому кварталу (см. столбец А, Рисунок 71). Например, среднее всех отношений для первого квартала равно 1,108. Это значение является сезонным индексом первого квартала, на основе которого можно сделать вывод, что объем сбыта угля за первый квартал составляет в среднем около 110,8% относительного среднего годового объема сбыта.

Сезонный индекс — это среднее отношение данных, относящихся к одному сезону (в данном случае сезоном является квартал), ко всем данным. Если сезонный индекс больше 1, значит, показатели этого сезона выше средних показателей за год, аналогично, если сезонный индекс ниже 1, то показатели сезона ниже средних показателей за год.

Наконец, чтобы исключить из исходных данных сезонную составляющую, следует поделить значения исходных данных на соответствующий сезонный индекс. Результаты этой операции приведены в столбцах F и G (см. Рисунок 72). График данных, которые уже не содержат сезонной составляющей, отражает Рисунок 73.

Прогнозирование без сезонной составляющей. На основе данных, из которых исключена сезонная составляющая, строится прогноз. Для этого используется соответствующий метод, который учитывает характер поведения данных (например, данные имеют тренд или относительно постоянны). В этом примере прогноз строится с помощью простого экспоненциального сглаживания. Оптимальное значение параметра α находится с помощью средства Поиск решения. Графики прогноза и реальных данных с исключенной сезонной составляющей приведены ниже (см. Рисунок 74).

	В	Ф	Г	Н	И		В	Ф	Г
7	Исходные данные	Сезонные индексы	Данные без сезонной составляющей	Прогноз		7	Исходные данные	Сезонные индексы	Данные без сезонной составляющей
8	2 159	1,108	1 948,1	1 948,1		8	2 159	=SA\$1	=B8/F8
9	1 203	0,784	1 535,4	1 948,1		9	1 203	=SA\$2	=B9/F9
10	1 094	0,860	1 272,3	1 678,5		10	1 094	=SA\$3	=B10/F10
11	1 996	1,234	1 617,8	1 413,1		11	1 996	=SA\$4	=B11/F11
12	2 081	1,108	1 877,8	1 546,8		12	2 081	=SA\$1	=B12/F12
13	1 332	0,784	1 700,0	1 763,0		13	1 332	=SA\$2	=B13/F13
14	1 476	0,860	1 716,6	1 721,9		14	1 476	=SA\$3	=B14/F14
15	2 533	1,234	2 053,1	1 718,4		15	2 533	=SA\$4	=B15/F15
16	2 249	1,108	2 029,3	1 937,1		16	2 249	=SA\$1	=B16/F16
17	1 533	0,784	1 956,5	1 997,4		17	1 533	=SA\$2	=B17/F17
18	1 935	0,860	2 250,4	1 970,7		18	1 935	=SA\$3	=B18/F18
19	2 523	1,234	2 045,0	2 153,4		19	2 523	=SA\$4	=B19/F19
20	2 208	1,108	1 992,3	2 082,6		20	2 208	=SA\$1	=B20/F20
21	1 597	0,784	2 038,2	2 023,6		21	1 597	=SA\$2	=B21/F21
22	1 917	0,860	2 229,5	2 033,2		22	1 917	=SA\$3	=B22/F22
23	2 726	1,234	2 209,5	2 161,4		23	2 726	=SA\$4	=B23/F23
24	2 612	1,108	2 356,9	2 192,8		24	2 612	=SA\$1	=B24/F24
25	1 931	0,784	2 464,5	2 300,0		25	1 931	=SA\$2	=B25/F25
26	2 223	0,860	2 585,3	2 407,5		26	2 223	=SA\$3	=B26/F26
27	2 363	1,234	1 915,3	2 523,7		27	2 363	=SA\$4	=B27/F27
28	2 074	1,108	1 871,4	2 126,2		28	2 074	=SA\$1	=B28/F28
29	1 589	0,784	2 028,0	1 959,7		29	1 589	=SA\$2	=B29/F29
30	1 673	0,860	1 945,7	2 004,4		30	1 673	=SA\$3	=B30/F30
31	2 443	1,234	1 980,1	1 966,0		31	2 443	=SA\$4	=B31/F31
32	2 231	1,108	2 013,1	1 975,2		32	2 231	=SA\$1	=B32/F32
33	1 675	0,784	2 137,8	2 000,0		33	1 675	=SA\$2	=B33/F33
34	1 503	0,860	1 748,0	2 090,0		34	1 503	=SA\$3	=B34/F34
35	2 259	1,234	1 831,0	1 866,5		35	2 259	=SA\$4	=B35/F35
36	1 809	1,108	1 632,3	1 843,3		36	1 809	=SA\$1	=B36/F36
37	1 254	0,784	1 600,5	1 705,5		37	1 254	=SA\$2	=B37/F37
38	1 613	0,860	1 875,9	1 636,9		38	1 613	=SA\$3	=B38/F38
39	2 238	1,234	1 814,0	1 793,0		39	2 238	=SA\$4	=B39/F39
40	2 004	1,108	1 808,3	1 806,7		40	2 004	=SA\$1	=B40/F40
41	1 406	0,784	1 794,5	1 807,7		41	1 406	=SA\$2	=B41/F41
42	1 725	0,860	2 006,2	1 799,1		42	1 725	=SA\$3	=B42/F42
43	1 994	1,234	1 616,2	1 934,4		43	1 994	=SA\$4	=B43/F43
44				1 726,5		44			
45						45			

Рисунок 72. Пример 4 – сезонные индексы и данные без сезонной составляющей

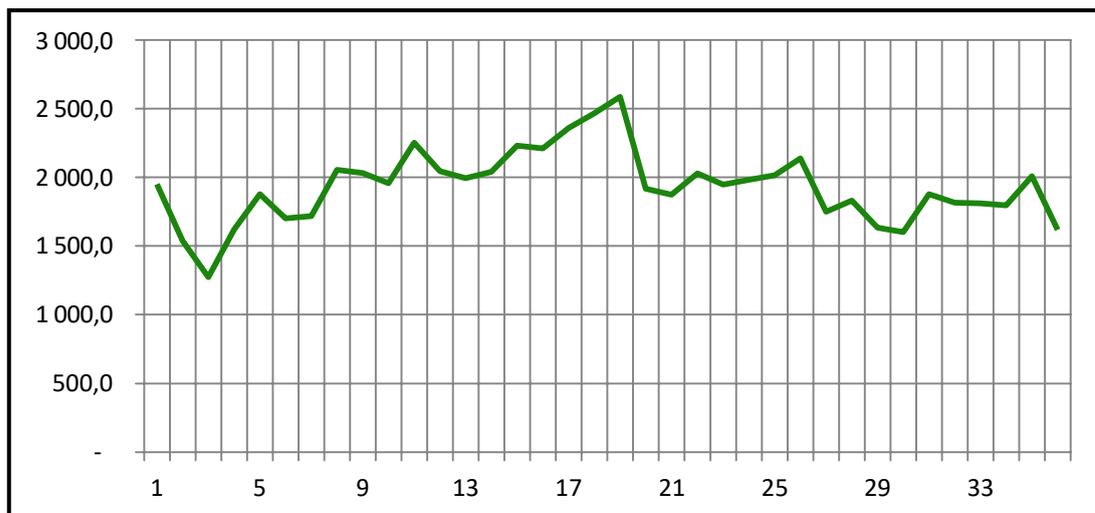


Рисунок 73. Пример 4 - данные без сезонной составляющей

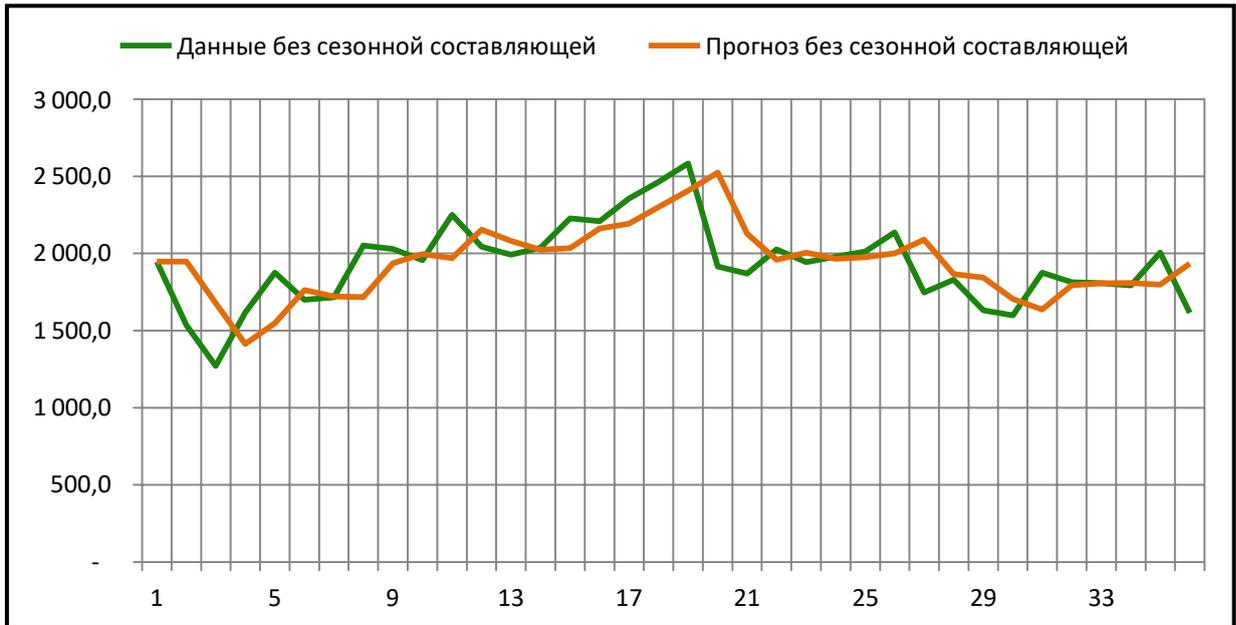


Рисунок 74. Пример 4 – графики прогноза и реальных данных с исключенной сезонной составляющей

Учет сезонной структуры. После нахождения прогноза без сезонной составляющей требуется учесть в полученном прогнозе (1726,5) сезонную составляющую. Для этого следует умножить 1726 на сезонный индекс первого квартала 1,108, в результате чего получим значение 1912. Аналогичная операция (умножение 1726 на сезонный индекс 0,784) даст прогноз на второй квартал, равный 1353. Результат добавления сезонной структуры к полученному прогнозу показан ниже (см. Рисунок 75).



Рисунок 75. Пример 4 – итоговый результат прогнозирования с сезонной составляющей

Литература

1. Шикин Е.В., Чхартишвили А.Г. Математические методы и модели в управлении. — М.: Дело, 2000. — 431 с.
2. Мур Дж., Уэдерфорд Л. и др. Экономическое моделирование в MS Excel. — М.: Вильямс, 2004. — 1024 с.
3. Эддоус М., Стэнсфилд Р. Методы принятия решений. — М.: Юнити, 1997. — 587 с.

