

Анализ временных рядов и прогнозирование в Microsoft Excel

Информационные технологии
управления

Временные ряды (ВР)

При практическом изучении ВР ставятся следующие цели:

- кратко описать характерные особенности ряда;
- подобрать статистическую модель, описывающую ВР;
- предсказать будущие значения на основе прошлых наблюдений

Достижение перечисленных целей позволит управлять процессом, породившим исследуемый ВР

Стадии анализа временного ряда:

- графическое представление и описание поведения ВР;
- выделение и удаление закономерных составляющих, зависящих от времени;
- выделение и удаление низко- и высокочастотных составляющих процесса;
- построение математической модели для описания оставшейся случайной составляющей и проверка ее адекватности;
- прогнозирование будущего развития процесса, представленного временным рядом.

Составляющие временного ряда

- Под детерминированной составляющей ВР x_1, \dots, x_n мы будем понимать числовую последовательность d_1, \dots, d_n , элементы которой d_t вычисляются по определенному правилу как функция времени t .
- Оставшаяся после выделения закономерной составляющей случайная компонента ВР обозначается обычно $\varepsilon_1, \dots, \varepsilon_t, \dots, \varepsilon_n$. Для ее описания используют понятия и методы теории вероятностей.

Элементы ВР представляют

либо в виде суммы, либо в виде произведения вышеназванных

компонент:

$$x_t = d_t + \varepsilon_t$$

- $x_t = d_t + \varepsilon_t$ при $t=1, \dots, n$ или $X=D+E$
- $x_t = d_t \times \varepsilon_t$ при $t=1, \dots, n$ или $X=D \times E$

В экономических приложениях детерминированную компоненту ВР обычно разделяют на три части: тренд tr_t , сезонную s_t и циклическую c_t составляющие. Так что можно записать следующее:

$$d_t = tr_t + s_t + c_t \quad \text{при } t = 1, \dots, n$$

Трендом ВР называют плавно изменяющуюся, нециклическую компоненту, описывающую чистое влияние долговременных факторов, эффект которых сказывается постепенно.

Сезонная компонента ВР описывает поведение, изменяющееся регулярно в течение заданного периода (года, месяца и т.д.). Она состоит из последовательности почти повторяющихся циклов.

Циклическая компонента ВР описывает длительные периоды относительного подъема и спада. Она состоит из циклов, которые меняются по амплитуде и протяженности.

Модели тренда

$$tr_t = b_0 + b_1 t$$

линейная модель

$$tr_t = b_0 + b_1 t + b_2 t^2 + \dots + b_n t^n$$

ПОЛИНОМ

$$tr_t = \exp(b_0 + b_1 t)$$

экспоненциальная

Выделение тренда

Для оценки и удаления тренда из ВР чаще всего используется метод *наименьших квадратов* - основу регрессионного анализа временных рядов. При этом значения ВР x_t рассматриваются как отклик, а время t - как фактор, влияющий на него:

$$x(t_i) = f(t_i, \theta) + \varepsilon_i \quad i=1, \dots, n$$

Метод наименьших квадратов состоит в том, что мы выбираем функцию тренда так, чтобы

$$\sum_{i=1}^n [x(t_i) - f(t_i, \theta)]^2 \rightarrow \min_{\theta}$$

Выделение сезонной компоненты

Предположим, что рассматриваемый ВР может быть описан аддитивной моделью. Пусть p - период последовательности s_t так что $s_t = s_{t+p}$ для всякого t . Наша задача - оценить значения s_t по наблюдениям x_t при том, что величина p известна.

Получив оценку тренда по одной из описанных выше методик, рассмотрим для каждого сезона i , $1 \leq i \leq p$, все относящиеся к нему разности (предполагая, что $n = (m+1)p$)

Выделение сезонной компоненты

Усреднение этих разностей дает нам оценку сезонной компоненты, например простое среднее:

$$\hat{s}_i = \frac{1}{m+1} \sum_{l=0}^m (x_{i+l p} - \hat{t}r_{i+l p})$$

В практических задачах распространена ситуация, когда сезонные колебания пропорциональны среднему значению процесса в рассматриваемый момент времени. Для описания подобных данных следует использовать либо смешанную мультипликативно-аддитивную модель $x_t = tr_t \times s_t + \varepsilon_t$, либо простую мультипликативную модель ВР. Если выбрать первую из них, то при оценке сезонных эффектов удобно рассмотреть частные

Выделение сезонной компоненты

$$\frac{x_{i+lp}}{\hat{tr}_{i+lp}} \times 100\% \quad \text{при } l = 0, 1, 2, \dots, m$$

В этом случае оценкой сезонной компоненты или сезонным индексом называют величину

$$\hat{s}_i = \frac{1}{m+1} \sum_{l=0}^m \left(\frac{x_{i+lp}}{\hat{tr}_{i+lp}} \times 100\% \right) \quad \text{где } 1 \leq i \leq p$$

Получив оценки сезонной компоненты, следует провести ее удаление из рассматриваемого ВР: в случае аддитивной модели, вычитая ее из начальных значений ряда; в случае мультипликативно-аддитивной модели - деля значения исходного ряда на соответствующие сезонные индексы и умножая полученный результат на 100%.

Метод скользящих средних (СС)

Метод СС основан на переходе от исходных значений ряда к их средним значениям на интервале времени, длина которого выбрана заранее. При этом сам выбранный интервал времени скользит вдоль ряда.

Величина интервала сглаживания должна быть кратна периоду сезонности или равна ему. При этом каждый интервал вычисления СС будет содержать данные, отвечающие всему периоду сезонности.

Метод экспоненциального сглаживания

Основная идея данного метода состоит в том, что каждый новый прогноз получается посредством перемещения предыдущего прогноза в направлении, которое дало бы лучшие результаты по сравнению со старым прогнозом. Базовое уравнение этой модели имеет следующий вид:

$$\hat{x}_{t+1} = \hat{x}_t + a \times (x_t - \hat{x}_t)$$

где a - константа сглаживания. Таким образом, каждая новый прогноз представляет собой сумму предыдущего прогноза и поправочного коэффициента, который и передвигает новый прогноз в направлении, делающем оценку более точной.

Графические средства анализа

В Excel существует механизм вставки *Линии тренда* в диаграмму, построенную на основании базовой линии ВР. Этот способ наиболее нагляден и прост в обращении. Добавление линии тренда к рядам данных производится следующим образом:

- Выберите ряд данных, к которому нужно добавить линию тренда.
- Выберите команду *Добавить линию тренда* в меню *Диаграмма*.
- На вкладке *Тип* выберите нужный тип линии тренда.
- При выборе типа *Полиномиальный* введите в поле *Степень* оптимальное на Ваш взгляд значение для независимой переменной.

Графические средства анализа

- При выборе типа *Скользящее среднее* введите в поле *Период* число периодов, используемых для расчета скользящего среднего.
- На вкладке *Параметры* можно указать программе, чтобы линия тренда снабжалась уравнением, по которому она строилась и среднеквадратичным отклонением от реального ВР.
- Здесь же можно указать число временных интервалов экстраполирования тренда для построения прогноза.