

Тема 1. Основные категории анализа временных рядов



Содержание

- 1.1. Временные ряды, их характеристики и задачи анализа
- 1.2. Требования к исходной информации
- 1.3. Аномальные наблюдения, методы их выявления и анализа
- 1.4. Методы выявления закономерностей динамики исследуемых явлений

1.1. Временные ряды, их характеристики и задачи анализа

Динамическим рядом называется совокупность наблюдений какого-либо явления (показателя), упорядоченная в зависимости от возрастающих или убывающих значений другого явления (показателя, признака).

Временной ряд - это набор чисел, привязанных к последовательным, обычно равноотстоящим моментам времени. Временной ряд - это совокупность значений какого-либо показателя за несколько последовательных моментов или периодов.

Числа, составляющие временной ряд и полученные в результате наблюдения за ходом какого-либо процесса, называются **уровнями (или элементами) ряда**.

Длиной временного ряда называют количество входящих в него уровней.

1.1. Временные ряды, их характеристики и задачи анализа

* Следовательно, *динамические ряды* - понятие, относящееся к тем рядам уровней, в которых содержится тенденция изменения, а *временные ряды* - более общее понятие, включающее как динамические, так и статические последовательности уровней какого-либо показателя.

Итак, временной ряд – это последовательность упорядоченных во времени числовых показателей, характеризующих уровень состояния и изменения изучаемого явления.

Элементы временного ряда (члены, уровни ряда) нумеруются в соответствии с номером момента времени, к которому они относятся (например, Y_1 , $Y_2, \dots$), и порядок уровней ряда играет важную роль в дальнейшем исследовании.

1.1. Временные ряды, их характеристики и задачи анализа

Анализ временных рядов дает возможность последить развитие явления, показать его основные пути, тенденции и темпы.

Основные цели анализа временного ряда

- * описание характерных особенностей ряда;
- * выяснение механизма, порождающего временной ряд;
- * подбор статистической модели, описывающей временной ряд;
- * прогноз будущих значений ряда на основе прошлых наблюдений;
- * управление процессом, порождающим временной ряд.

1.1. Временные ряды, их характеристики и задачи анализа

Практическая реализация перечисленных целей выполняется в результате следующих этапов:

- * графическое представление временного ряда;
- * выделение и удаление: тренда, сезонных и циклических составляющих (выравнивание, сглаживание);
- * выделение и удаление низко- или высокочастотных составляющих процесса (фильтрация);
- * исследование случайной составляющей временного ряда, оставшейся после сглаживания и фильтрации;
- * построение математической модели для описания случайной составляющей и проверка ее адекватности;
- * прогнозирование будущего развития процесса, описанного временным рядом;
- * исследование взаимосвязей между различными временными рядами.

1.1. Временные ряды, их характеристики и задачи анализа

Для решения перечисленных выше задач используются следующие *методы и модели*.

- * методы корреляционного анализа, позволяющие выявить периодические зависимости и их лаги внутри одного процесса (автокорреляция) или между несколькими процессами (кросскорреляция);
- * методы спектрального анализа, нацеленные на выявление периодических и квазипериодических составляющих временного ряда;
- * методы преобразования временных рядов (сглаживание и фильтрация) с целью удаления из них высокочастотных или сезонных составляющих;
- * модели авторегрессии и скользящего среднего (для описания и прогнозирования случайной составляющей временного ряда).

1.1. Временные ряды, их характеристики и задачи анализа

Классификация временных рядов

Признак классификации	Виды временного ряда
1. В зависимости от качественной особенности изучаемого явления.	1. Абсолютных величин 2. Относительны величин 3. Средних величин
2. В зависимости от того, как уровни выражают состояния явлений во времени.	1. Интервальные ряды 2. Моментные ряды
3. В зависимости от расстояния между уровнями.	1. С равностоящими уровнями по времени 2. С не равностоящими уровнями по времени
4. В зависимости от наличия основной тенденции изучаемого процесса.	1. Стационарные ряды 2. Нестационарные ряды

1.1. Временные ряды, их характеристики и задачи анализа

Интервальные временные ряды представляются последовательностью значений показателей за определенный интервал времени (год, квартал, месяц и т.д.).

Моментные ряды представляются в виде последовательности показателей, относящихся к конкретным моментам времени (на 1 января, на 1 июля и т. д.).

Особенности уровней интервальных и моментных рядов

1. отдельные уровни моментного временного ряда абсолютных величин содержат **элементы повторного счета**, то есть в каждом последующем уровне содержится полностью или частично значение предыдущего уровня. Все это делает бессмысленным суммирование моментных рядов.

2. Значения же уровней абсолютных интервальных временных рядов, в отличие от уровней моментного ряда, не содержатся в предыдущих и последующих показателях, их можно просуммировать, что позволяет получать ряды более укрупненных периодов, или ряды с нарастающими итогами

1.1. Временные ряды, их характеристики и задачи анализа

Временные ряды подразделяются на **стационарные** и **нестационарные**.

Случайные процессы, протекающие во времени приблизительно однородно и имеющие вид непрерывных случайных колебаний вокруг некоторого среднего значения, причем ни средняя амплитуда, ни характеристика этих колебаний не обнаруживают существенных изменений с течением времени в математической статистике называются **стационарными**.

Всякий стационарный процесс можно рассматривать как процесс, неопределенно долго продолжающийся во времени.

1.1. Временные ряды, их характеристики и задачи анализа

Стационарным (строго стационарным или стационарным в узком смысле слова) называется такой ряд динамики, у которого совместное распределение вероятностей для n уровней $y(t_1), y(t_2), \dots, y(t_n)$, такое же, что и для n уровней $y(t_1+\tau), y(t_2+\tau), \dots, y(t_n+\tau)$ для любого $t_1, t_2, \dots, t_n$ и τ .

Стационарным в широком смысле называется процесс, если постоянны первый и второй моменты, а функция ковариации зависит только от лагового сдвига m и не зависит от времени:

$$1). \mu_t = \mu = \text{const};$$

$$2). D_t = \sigma^2 = \text{const};$$

$$3). \text{Cov}(t+m) = \text{cov}(m)$$

$$\text{где } \text{cov}_X(t+m, t) = \mu(X_{t+m} - \mu)(X_t - \mu) \text{ .}$$

1.1. Временные ряды, их характеристики и задачи анализа

Стационарность означает, что среднее значение уровней и их дисперсия - постоянные величины.

Можно сказать и так: ряд динамики стационарный, если значения его уровней не зависят от начала отсчета.

В экономической практике в большинстве случаев приходится иметь дело со случайными процессами, имеющими вполне определенную тенденцию развития во времени. Такие процессы называются нестационарными, и временные ряды также называются нестационарными.

Характеристики нестационарного случайного процесса меняются во времени, то есть зависят от начала отсчета.

нестационарный ряд: $x(k) \in (\mu, \sigma^2)$, $\mu = \text{var}$, $\sigma^2 = \text{const}$
где k - это номер наблюдения

1.1. Временные ряды, их характеристики и задачи анализа

Выбор вида временного ряда определяется целями анализа.

Временные ряды могут быть изображены *графически*. Графическое изображение позволяет наглядно представить развитие явлений во времени и способствует проведению анализа уровней.

Наиболее распространенным видом графического изображения для аналитических целей является **линейная диаграмма**, которая строится в прямоугольной системе координат.

Наряду с линейной диаграммой для графического изображения временных рядов в целях популяризации широко используются столбиковые, секторные и другие виды диаграмм.

1.2. Преобразования исходной информации

При обнаружении несопоставимости уровней ряда может применяться процедура смыкания рядов.

Смыкание может быть произведено двумя способами:

1. с помощью коэффициента перевода;
2. с помощью формирования относительных величин

Первый способ заключается в том, что данные за предыдущие периоды умножаются на коэффициент перевода, равный отношению показателей на этот момент времени, когда произошло изменение условий формирования уровней ряда. Например, в современных условиях переоценка основных производственных фондов происходит ежегодно, и, следовательно, во временном ряду каждый год становится переходным, что постоянно требует расчета коэффициента перевода.

1.2. Преобразования исходной информации

Смыкание может быть применено, когда для одной даты имеются значения уровней обоих рядов:

t_i			...			...	
y_i			...		—	...	—
	—	—	...			...	

Коэффициент перевода рассчитывается по формуле:

$$k = \frac{y_1^*}{y_m}$$

на который умножаем значение всех уровней. Получаем сомкнутый ряд динамики:

t_i			...			...	
	ky_1	ky_2	...			...	

1.2. Преобразования исходной информации

Второй способ — уровень переходного периода принимается для 2-ой части ряда за 100%, и от этого уровня определяются соответствующие показатели (вперед или назад). При этом получается сопоставимый ряд относительных величин.

То есть, последний уровень y_m первого ряда и первый уровень y_1^* второго ряда принимают за 100%, а остальные пересчитывают в процентах по отношению к ним:

t_i			...			...	
			...	100%		...	

1.2. Преобразования исходной информации

Способы получения пропущенных значений временного ряда

Интерполяция, интерполирование — в вычислительной математике это способ нахождения промежуточных значений величины по имеющемуся дискретному набору известных значений.

Рассмотрим систему несовпадающих точек x_i ($i \in 0, 1, \dots, N$) из некоторой области D . Пусть значения функции известны только в этих точках:

$$y_i = f(x_i), \quad i = 1, \dots, N.$$

Задача интерполяции состоит в поиске такой функции F из заданного класса функций, что

$$F(x_i) = y_i, \quad i = 1, \dots, N.$$

1.2. Преобразования исходной информации

Способы получения пропущенных значений временного ряда

Точки x_i называют узлами интерполяции, а их совокупность — интерполяционной сеткой.

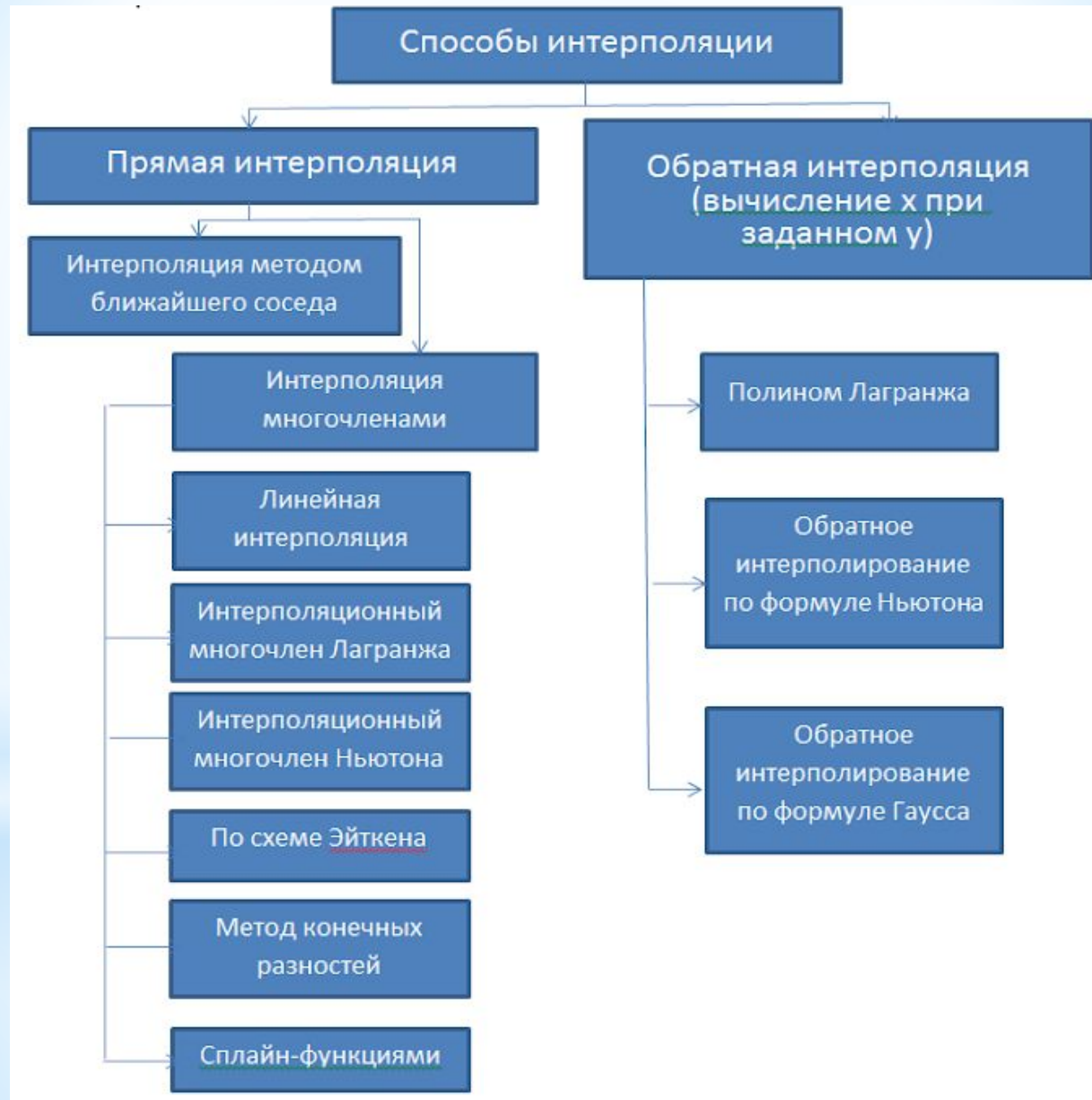
Пары (x_i, y_i) называют точками данных или базовыми точками.

Разность между «соседними» значениями

$\Delta x_i = x_i - x_{i-1}$ - шагом интерполяционной сетки. Он может быть как переменным, так и постоянным.

Функцию $F(x)$ — интерполирующей функцией или интерполянтом.

1.2. Преобразования исходной информации



1.2. Преобразования исходной информации

Способы интерполяции

1) Интерполяция методом ближайшего соседа - простейший способ нахождения недостающего значения уровня ряда

Интерполяция методом ближайшего соседа (ступенчатая интерполяция) — метод интерполяции, при котором в качестве промежуточного значения выбирается *ближайшее известное значение функции*. Интерполяция методом ближайшего соседа является самым простым методом интерполяции.

2) Интерполяция многочленами

На практике чаще всего применяют интерполяцию многочленами. Это связано прежде всего с тем, что многочлены легко вычислять, легко аналитически находить их производные и множество многочленов плотно в пространстве непрерывных функций (теорема Вейерштрасса).

1.2. Преобразования исходной информации

2.1. Линейная интерполяция — интерполяция алгебраическим двучленом вида $P_1(x) = ax + b$ функции f , заданной в двух точках x_0 и x_1 отрезка $[a, b]$.

В случае, если заданы значения в нескольких точках, функция заменяется кусочно-линейной функцией.

Геометрически это означает замену графика функции f прямой, проходящей через точки $(x_0, f(x_0))$ и $(x_1, f(x_1))$.

Уравнение такой прямой имеет вид:

отсюда для

$$\frac{y - f(x_0)}{f(x_1) - f(x_0)} = \frac{x - x_0}{x_1 - x_0}$$

$$x \in [x_0, x_1]$$

Это и есть формула линейной интерполяции, при этом

где $R_1(x)$ — погрешность формулы:

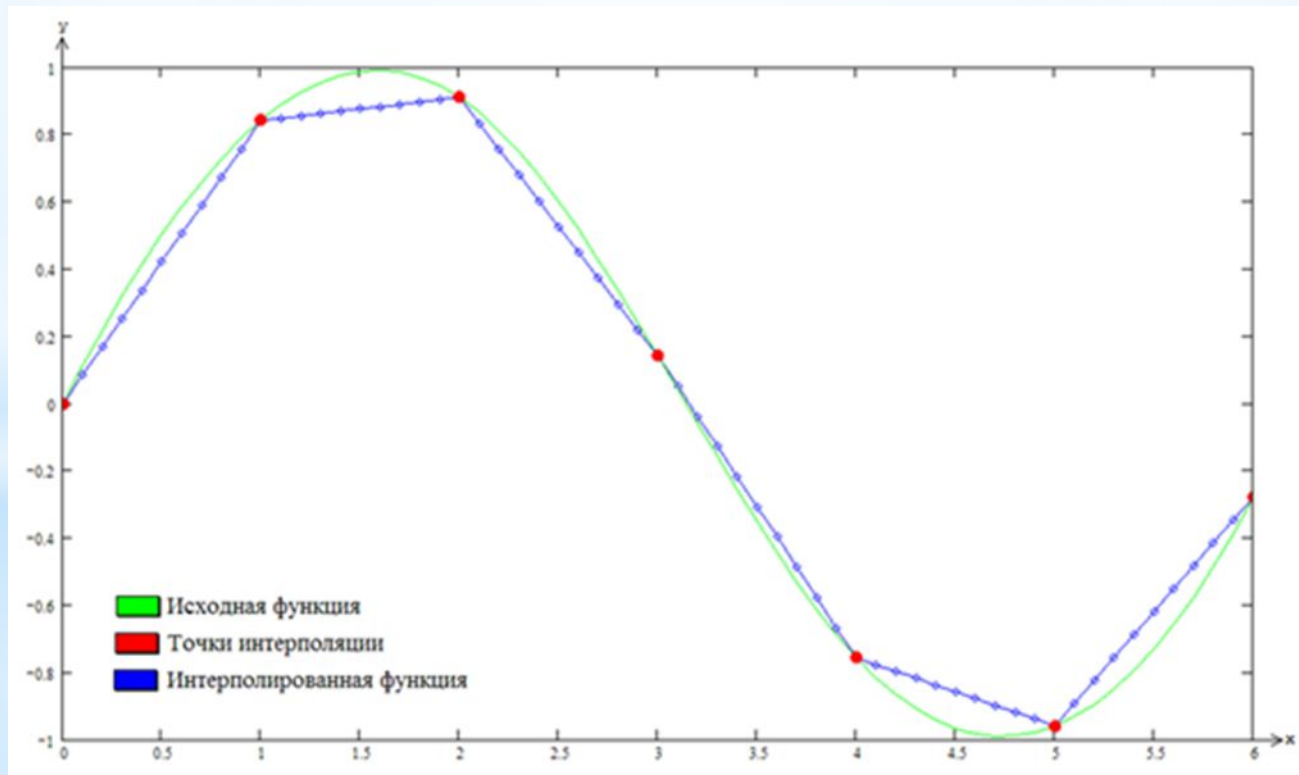
$$R_1(x) = \frac{f''(\psi)}{2} (x - x_0)(x - x_1), \quad \psi \in [x_0, x_1]$$

1.2. Преобразования исходной информации

Справедлива оценка

$$|R_1(x)| \leq \frac{M_2}{2} \max |(x-x_0)(x-x_1)| = \frac{M_2 h^2}{8}, \quad M_2 = \max_{[a,b]} |f''(x)|, \quad h = x_1 - x_0.$$

Пример линейной интерполяции



1.2. Преобразования исходной информации

2.2. Полином Лагранжа (интерполяционный многочлен) - многочлен минимальной степени, принимающий данные значения в данном наборе точек. Для $n + 1$ пар чисел $(x_0, y_0), (x_1, y_1) \dots, (x_n, y_n)$

где все x_j различны, существует единственный многочлен $L(x)$ степени не более n , для которого $L(x_j) = y_j$.

В простейшем случае ($n = 1$) — это линейный многочлен, график которого — прямая, проходящая через две заданные точки.

Лагранж предложил способ вычисления таких многочленов:

$$L(x) = \sum_{i=0}^n y_i l_i(x)$$

где базисные полиномы определяются по формуле:

$$l_i(x) = \prod_{j=0, j \neq i}^n \frac{x - x_j}{x_i - x_j} = \frac{x - x_0}{x_i - x_0} \dots \frac{x - x_{i-1}}{x_i - x_{i-1}} \frac{x - x_{i+1}}{x_i - x_{i+1}} \dots \frac{x - x_n}{x_i - x_n}$$

1.2. Преобразования исходной информации

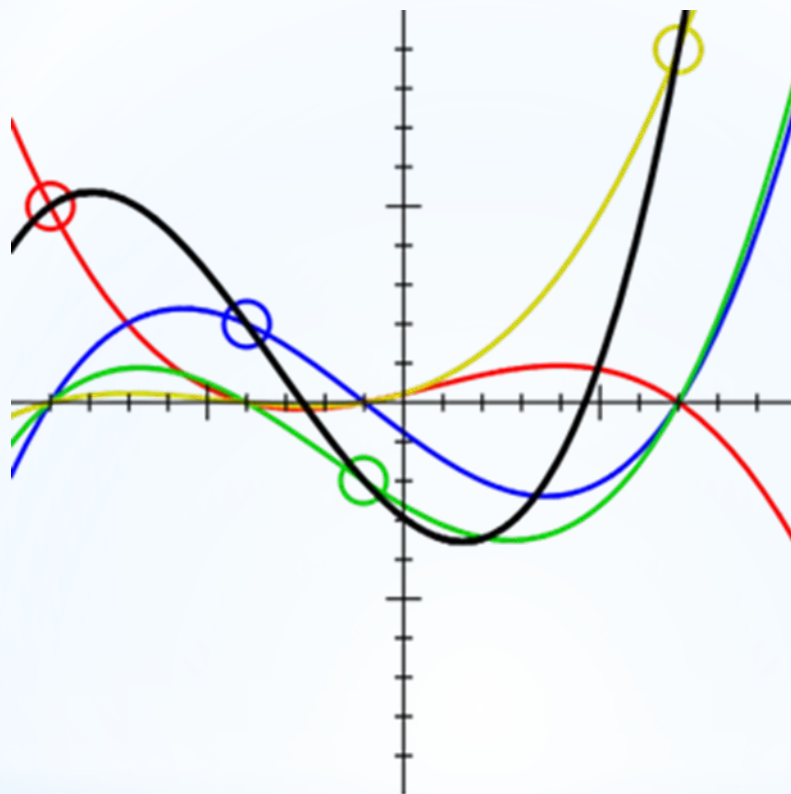
$l_i(x)$ обладают следующими свойствами:

- 1) являются многочленами степени n ;
- 2) $l_i(x_i) = 1$;
- 3) $l_i(x_j) = 0$ при

Отсюда следует, что $L(x)$, как линейная комбинация $l_i(x)$, может иметь степень не больше n , и $L(x_i) = y_i$,

Полиномы Лагранжа используются для интерполяции, а также для численного интегрирования.

1.2. Преобразования исходной информации



Этот пример показывает интерполяционный многочлен Лагранжа для четырёх точек $(-9, 5)$, $(-4, 2)$, $(-1, -2)$ и $(7, 9)$, а также полиномы $y_i, l_i(x)$, каждый из которых проходит через одну из выделенных точек, и принимает нулевое значение в остальных x_j .

1.2. Преобразования исходной информации

Требования к информации



1.2. Преобразования исходной информации

Сопоставимость достигается в результате одинакового подхода к наблюдениям на разных этапах формирования динамического ряда.

Однородность данных означает отсутствие сильных изломов тенденций, а также **аномальных** (т.е. резко выделяющихся, нетипичных для данного ряда) наблюдений. Аномальные наблюдения проявляются в виде сильного изменения уровня - скачка или спада - с последующим приблизительным восстановлением предыдущего уровня.

Устойчивость характеризуется преобладанием закономерности над случайностью в изменении уровней ряда.

Требование полноты данных обуславливается тем, что закономерность может обнаружиться лишь при наличии минимально допустимого объема наблюдений.

1.2. Преобразования исходной информации

Процедуры предварительного анализа временных рядов

