

Автокорреляция (Autocorrelation)

Разделы: [Алгоритмы](#)

LogInom: [Автокорреляция \(обработчик\)](#).

В математической статистике автокорреляцией называют меру статистической связи (корреляции) между функцией и ее копией, сдвинутой на некоторый интервал, называемый **лагом**.

В регрессионном анализе автокорреляция остатков регрессии может иметь место, если модель плохо согласуется с данными. Это происходит, если на зависимую переменную воздействуют факторы, неучтенные в модели.

В анализе данных автокорреляция широко используется для анализа и моделирования временных рядов и позволяет описывать динамические свойства временного ряда. Чем выше автокорреляция временного ряда, тем сильнее связаны его наблюдения. Для обнаружения автокорреляции во временных рядах используется критерий Дарбина-Уотсона.

Количественно автокорреляцию можно измерить с помощью коэффициента автокорреляции. Для временного ряда X_1, X_2, \dots, X_{n-L} и его копии $X_{1+L}, X_{2+L}, \dots, X_n$, сдвинутой на лаг L , коэффициент автокорреляции вычисляется следующим образом:

$$r_{t,t-L} = \frac{\overline{X_t X_{t-L}} - \overline{X_t} \overline{X_{t-L}}}{s_t s_{t-L}},$$

где:

$$\overline{X_t X_{t-L}} = \left(\sum_{t+L}^n X_t X_{t-L} \right) / (n - L);$$

$$\overline{X_t} = \left(\sum_{t+L}^n X_t \right) / (n - L) \text{ — среднее значение исходного ряда;}$$

$$\overline{X_{t-L}} = \left(\sum_{t+L}^n X_{t-L} \right) / (n - L) \text{ — среднее значение сдвинутого ряда;}$$

s_t и s_{t-L} — среднеквадратические отклонения исходного и сдвинутого ряда соответственно.

Временной лаг определяет порядок коэффициента автокорреляции. Например, если $L = 1$, то получим коэффициент автокорреляции 1-го порядка r_{t-L} . Следует учитывать, что с увеличением лага на единицу число пар значений, по которым рассчитывается коэффициент автокорреляции, уменьшается на 1. Поэтому обычно рекомендуют максимальный порядок коэффициента автокорреляции, равный $n/4$.

Если рассчитать несколько коэффициентов автокорреляции, можно определить лаг L' при котором автокорреляция наиболее высокая, определив тем самым структуру временного ряда. Если наиболее высоким оказывается значение $r_{t,t-L'}$, то исследуемый ряд содержит только тенденцию, а циклическая составляющая отсутствует.

Если наиболее высоким оказался $r_{t,t-L}$, то ряд, помимо тенденции, содержит циклические колебания с периодом L .

Последовательность коэффициентов автокорреляции называют автокорреляционной функцией $f(L)$ временного ряда. Она принимает максимальное значение, равное 1 при $L = 0$ (ряд полностью коррелирован сам собой) и равное 0, когда исходный ряд и его сдвинутая копия не коррелированы.

Чем быстрее убывает автокорреляционная функция с ростом L , тем слабее автокорреляция, и наоборот.