

Модель ARIMA

Модель авторегрессии порядка p описывается как

$$AR(p): y_t = c + \phi_1 y_{(t-1)} + \phi_2 y_{(t-2)} + \dots + \phi_p y_{(t-p)} + \varepsilon_t$$

и показывает зависимость значения нынешнего периода от прошлых значений p -периодов.

Модель скользящего среднего порядка q описывается как

$$MA(q): y_t = c + \varepsilon_t + \theta_1 \varepsilon_{(t-1)} + \theta_2 \varepsilon_{(t-2)} + \dots + \theta_q \varepsilon_{(t-q)}$$

и показывает зависимость значения нынешнего периода от ошибок предсказания предыдущих q периодов.

Модель авторегрессии с интегрированием и скользящим средним порядков (p, d, q) является суммой $AR(p)$ и $MA(q)$ моделей и может быть представлена в виде

$$ARIMA(p, d, q): (1 - \phi_1 L - \dots - \phi_p L^p)((1-L)^d y_t - \mu) = (1 + \theta_1 L + \dots + \theta_q L^q) \varepsilon_t,$$

где d – количество дифференцирований исходного временного ряда до достижения его стационарности, а L – величина лага.

Для выделения сезонности можно использовать модификацию базовой модели $ARIMA(p, d, q) - SARIMA(P, D, Q)_m$, где m – количество измерений в году, P, D, Q – порядки сезонной авторегрессии, сезонного интегрирования и сезонного скользящего среднего соответственно.

Функция `auto.arima()` принимает временной ряд и самостоятельно подбирает необходимые значения параметров. Для этого функция находит оптимальное количество дифференцирований, дающих стационарный ряд. Впоследствии происходит выбор оптимальной модели среди моделей $ARIMA$ с параметрами $(2,2), (0,0), (1,0), (0,1)$. После этого параметры выбранной модели меняются на ± 1 , добавляются/убираются константы. Подобные итерации происходят до тех пор, пока изменение параметров не приводит к уменьшению скорректированного критерия Акаике. Подбор сезонных параметров происходит по такой же схеме.